BONN-TH-2024-05

# Cutting-Edge Tools for Cutting Edges

Ruth Britto [1], Claude Duhr [2], Holmfridur S. Hannesdottir [3] and Sebastian Mizera [3]

[1] *School of Mathematics and Hamilton Mathematics Institute, Trinity College, Dublin 2, Ireland*
[2] *Bethe Center for Theoretical Physics, Universität Bonn, D-53115, Germany*
[3] *Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA*

We review different notions of cuts appearing throughout the literature on scattering amplitudes. Despite similar names, such as *unitarity cuts* or *generalized cuts*, they often represent distinct computations and distinct physics. We consolidate this knowledge, summarize how cuts are used in various computational strategies, and explain their relations to other quantities including imaginary parts, discontinuities, and monodromies. Differences and nuances are illustrated on explicit examples.

arXiv:2402.19415v1 [hep-th] 29 Feb 2024

## CONTENTS

I. Introduction 1

II. Preliminaries 2

III. Unitarity cuts and total discontinuities 3
A. Nonperturbative unitarity 3
1. Largest time equation 4
B. Perturbative unitarity 4
1. Proofs 5
2. The largest time equation in perturbation theory 6
3. Unstable particles 7
4. Reverse unitarity 7
C. Branch cuts and dispersion relations 7
1. Mathematical preliminaries 7
2. Dispersion relations for amplitudes 8
D. Double discontinuity 9
1. S-matrix bootstrap 9
2. Steinmann relations 9
E. Other applications 10
1. Gravitational radiation 10
2. Thermal physics 10
3. Cosmology and (anti-)de Sitter space 10
4. String theory 10

IV. Cuts and monodromies 10
A. Encircling singularities 10
B. Connection to cuts 13
C. Discontinuities and monodromies 13
D. Cuts for anomalous thresholds 14
1. Non-Mandelstam variables 15
2. Loop-tree theorems 16
E. Symbols and monodromies 17
1. Definitions 17
2. Symbols of Feynman integrals 17
F. Picard–Lefschetz theory for Feynman integrals 18

V. Generalized cuts 18
A. Cuts of arbitrary sets of edges 18
B. Unitarity methods for amplitudes 19
C. Solution space of generalized cuts 19

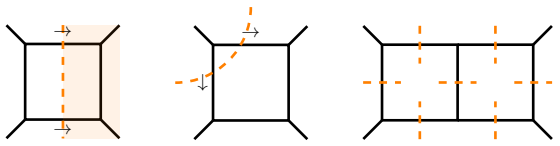
VI. Summary 20

Acknowledgments 20

References 20

## I. INTRODUCTION

At the heart of modern research into scattering amplitudes are *cuts*. Diagrammatically, cutting a diagram means tagging a subset of its internal edges, here denoted with dashed lines, e.g.



(1)

However, this operation can have distinct meanings depending on the physical or mathematical context in which it is employed. The purpose of this article is to review different flavors of cuts appearing in the literature, as well as explain how they are used in physical applications and as a computational scheme.

While details differ depending on the application, cuts always come with a notion of placing particles with momentum $p$ and mass $m$ on their mass shell, i.e., $p^2 = m^2$, which can be implemented as a Dirac delta function or a residue prescription. As a result, in perturbation theory the loop integrand simplifies and the number of loop integrations drops, which is ultimately why cuts become a useful tool. Here, we will categorize cuts by their purpose. Broadly, they are employed to compute:

- **Imaginary parts and total discontinuities.** Historically the oldest version, which is codified by the *Cutkosky rules* and directly descends from unitarity. Unitarity cuts appear as a consequence of integrating over the phase-space of on-shell states and become useful for computing cross sections through the optical

theorem, as well as sum rules via dispersion relations, see Sec. III. For unitarity cuts, we use the convention (1, left) where the energy flows from the unshaded region to the shaded one.

- **Individual discontinuities and monodromies.** A finer version, which relates cuts to discontinuities around individual branch points of scattering amplitudes. They give constraints on the analytic structure, which can be used in perturbative bootstrap approaches to scattering amplitudes, see Sec. IV. For this type of cut, we use the convention (1, middle) with arrows indicating the energy flow.
- **Integrands and bases of functions.** Generalized cuts work at the level of loop integrands and are useful for recursively constructing them by gluing lower-point amplitudes. They can be successfully employed in a variety of applications, including generalized unitarity, differential equations, and double-copy relations, see Sec. V. For generalized cuts, we use the convention (1, right) where energies of cut particles are not constrained at all.

## II. PRELIMINARIES

Throughout this review we will use the following standard notation. We stick to all-incoming conventions and mostly-minus signature, i.e., momentum conservation reads $\sum_{i=1}^{n} p_i = 0$ with $p_i^0 > 0$ for incoming particles and $p_i^0 < 0$ for outgoing. We use the short-hand notation $p_{ij\cdots k} = p_i + p_j + \ldots + p_k$. For $n = 4$, we use the Mandelstam invariants

$$s = (p_1 + p_2)^2, \quad t = (p_2 + p_3)^2, \quad u = (p_1 + p_3)^2\,, \quad (2)$$

which satisfy $s + t + u = \sum_{i=1}^{4} p_i^2$ due to momentum conservation.

**Definition 1** A Feynman integral $\mathcal{I}$ of a connected diagram with L loops and E internal edges in $\mathrm{D} = \mathrm{D}_0 - 2\epsilon$ space-time dimensions, where $\mathrm{D}_0$ is a positive integer, is given by

$$\mathcal{I} := \int_{\ell_1,\ldots,\ell_{\mathrm{L}}} \frac{\mathcal{N}}{\prod_{e=1}^{\mathrm{E}} (q_e^2 - m_e^2 + i\varepsilon)}\,, \quad (3)$$

where $q_e$ is the momentum flowing through the $e^{\text{th}}$ edge, $\ell_1, \ldots, \ell_{\mathrm{L}}$ form a basis of loop momenta, and $\mathcal{N}$ is any numerator that is polynomial in the loop momenta. Note that $q_e$ and $\mathcal{N}$ can depend on both the loop momenta $\ell_i$, as well as the external momenta labelled with $p_i$. We use the short-hand notation

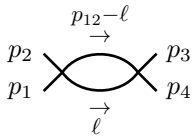
$$\int_{\ell_1,\ldots,\ell_{\mathrm{L}}} := \lim_{\varepsilon\to 0^+} \int \prod_{a=1}^{\mathrm{L}} \frac{\mathrm{e}^{\gamma_E \epsilon} \mu^{2\epsilon}}{i\pi^{\mathrm{D}/2}} \mathrm{d}^{\mathrm{D}}\ell_a\,, \quad (4)$$

where $\gamma_E$ is the Euler–Mascheroni constant and $\mu$ is a mass scale. We henceforth work in units in which $\mu = 1$.

We refer to [93] for an introduction to perturbative scattering amplitudes. For reviews of nonperturbative aspects, see [23, 29, 65].

In all examples, we take the square root, logarithms and dilogarithms to be on their principal branches.

**Example 1** We label the bubble diagram with both internal masses equal to $m$ as



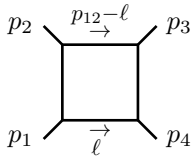

(5)

In two dimensions, the corresponding Feynman integral is given by

$$\mathcal{I}^{\mathrm{D}=2}_{\mathrm{bub}} := \int_\ell \frac{1}{(\ell^2 - m^2)((\ell - p_{12})^2 - m^2)} = \frac{2i}{\sqrt{s(4m^2 - s)}} \log\left(\frac{\sqrt{4m^2 - s} - i\sqrt{s}}{\sqrt{4m^2 - s} + i\sqrt{s}}\right). \quad (6)$$

This expression is valid for any complex $s$, except on the branch cut $s > 4m^2$.

**Example 2** The massless box diagram is labelled with



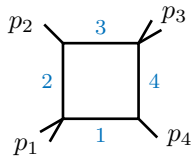

(7)

with all $p_i^2 = 0$. In "Euclidean kinematics" $s, t < 0$, it is given by

$$\mathcal{I}^{0m}(s,t) := \int_\ell \frac{1}{\ell^2(\ell - p_1)^2(\ell - p_{12})^2(\ell - p_{123})^2} = \frac{1}{st}\Big[\frac{2}{\epsilon^2}\big((-s)^{-\epsilon} + (-t)^{-\epsilon}\big) - \big(\log(-s) - \log(-t)\big)^2 - 4\pi^2/3 + \mathcal{O}(\epsilon)\Big] \quad (8)$$

where we have expanded up to $\mathcal{O}(\epsilon^0)$ around $\mathrm{D} = 4 - 2\epsilon$ spacetime dimensions. The answer is written in such a way that the analytic continuation to other kinematic channels is given by replacing $s \to s + i\varepsilon$ and $t \to t + i\varepsilon$. Note that analytic continuation in general does not amount to such a simple replacement.

**Example 3** The "two-mass easy box" diagram corresponds to (7) with $p_1^2 \neq 0$ and $p_3^2 \neq 0$. It can be interpreted as either a contribution to a 4-point amplitude with two massive legs, or to a 6-point massless amplitude:




(9)

Either way, in an expansion around $\mathrm{D} = 4 - 2\epsilon$ space-time dimensions, it evaluates to

$$\mathcal{I}^{2m}(s,t,p_1^2,p_3^2) = \frac{1}{st - p_1^2 p_3^2} \tag{10}$$
$$\times\Bigg\{\left(\frac{2}{\epsilon^2} - \frac{\pi^2}{6}\right)\left[(-s)^{-\epsilon}+(-t)^{-\epsilon}-(-p_1^2)^{-\epsilon}-(-p_3^2)^{-\epsilon}\right]$$
$$-2\mathrm{Li}_2\left(1-\frac{p_1^2}{s}\right) - 2\mathrm{Li}_2\left(1-\frac{p_1^2}{t}\right) - 2\mathrm{Li}_2\left(1-\frac{p_3^2}{s}\right)$$
$$-2\mathrm{Li}_2\left(1-\frac{p_3^2}{t}\right) + 2\mathrm{Li}_2\left(1-\frac{p_1^2 p_3^2}{st}\right) - \log^2\left(\frac{s}{t}\right)\Bigg\} + \mathcal{O}(\epsilon)\,.$$

when $s,t,p_1^2,p_3^2 < 0$. The analytic continuation of this expression to other kinematic regions is nontrivial, and must be done carefully so that it matches onto the $i\varepsilon$ prescription of the Feynman integral. We can, for example, rotate the invariants from negative to positive in the upper half-plane and analytically continue the expression in (10). Note that this does not correspond to numerically substituting $s \to s + i\varepsilon$ etc. for the kinematic invariants, since we might need to change the branch of the dilogarithms during the analytic continuation; recall that (10) assumes they are written on their principal branch.

We define the "one-mass box" as

$$\mathcal{I}^{1m}(s,t,p_1^2) = \lim_{p_3^2 \to 0} \mathcal{I}^{2m}(s,t,p_1^2,p_3^2). \tag{11}$$

This equation holds for a fixed dimension D. Note that the limit in $p_3^2 \to 0$ and the expansion in $\epsilon$ do not commute in general, but (10) is written in such a way that substituting $p_3^2 = 0$ on the right-hand side gives $\mathcal{I}^{1m}(s,t,p_1^2)$ up to terms of $\mathcal{O}(\epsilon)$.

Cuts of $\mathcal{I}$ will be defined by replacing propagators with delta functions of the form

$$\frac{-1}{2\pi i}\hat{\delta}^{\pm}(q_e^2 - m_e^2) = \delta^{\pm}(q_e^2 - m_e^2) \tag{12a}$$
$$= \delta(q_e^2 - m_e^2)\theta(\pm q_e^0)\,. \tag{12b}$$

The $\pm$ sign selects a definite flow of the energy component $q_e^0$ across the cut. The specific choice of which signs to use will depend on the application.

## III. UNITARITY CUTS AND TOTAL DISCONTINUITIES

### A. Nonperturbative unitarity

Historically, the first incarnation of cut integrals were unitarity cuts. They are a direct consequence of the unitarity of the $S$-matrix, $S^\dagger S = \mathbb{1}$, which embodies the physical principle of probability conservation. If we define the transfer matrix $T$ in the usual way by $S = \mathbb{1} + iT$, then the unitarity of the $S$-matrix translates into the statement that

$$T - T^\dagger = iT^\dagger T\,. \tag{13}$$

By definition, the scattering amplitude $\mathcal{M}(\text{in} \to \text{out})$ corresponds to the matrix elements of the transfer matrix with asymptotic external states $\langle\text{out}|$ and $|\text{in}\rangle$,

$$\hat{\delta}_{\text{in,out}}\,\mathcal{M}(\text{in} \to \text{out}) = \langle\text{out}|T|\text{in}\rangle\,, \tag{14a}$$
$$\hat{\delta}_{\text{in,out}}\mathcal{M}(\text{out} \to \text{in})^* = \langle\text{out}|T^\dagger|\text{in}\rangle\,, \tag{14b}$$

where $\hat{\delta}_{\text{in,out}} = (2\pi)^{\mathrm{D}}\delta^{\mathrm{D}}(p_{\text{out}} + p_{\text{in}})$ is the overall momentum-conserving delta function. Note that $\mathcal{M}$ contains all connected and disconnected terms apart from the fully disconnected term $\hat{\delta}_{\text{in,out}}\langle\text{out}|\mathbb{1}|\text{in}\rangle$. Hence, if we contract (13) with asymptotic sates, and we insert a complete set of asymptotic states into the right-hand side, we can cast the unitarity of the $S$-matrix a nonlinear equation for the scattering amplitudes:

$$\mathcal{M}(\text{in} \to \text{out}) - \mathcal{M}(\text{out} \to \text{in})^* = \tag{15}$$
$$i \sum\!\!\!\!\!\!\int_X \hat{\delta}_{\text{in,X}}\mathcal{M}(\text{in} \to X)\,\mathcal{M}(\text{out} \to X)^*\,.$$

The term on the right-hand side is called the *unitarity cut*.[1] It involves summing over all possible intermediate states $X$ in the theory and integrating over their on-shell phase-space. More concretely,

$$\sum\!\!\!\!\!\!\int_X := \sum_{X_j}\prod_{i\in X_j}\int\frac{\mathrm{d}^{\mathrm{D}}\ell_i}{(2\pi)^{\mathrm{D}-1}}\delta^+(\ell_i^2 - m_i^2)\,, \tag{18}$$

where $m_i$ is the mass of particle $i$, and each $X_j$ consists of a possible set of particles that can be exchanged.

[1] Recall that in even space-time dimensions D, CPT invariance gives

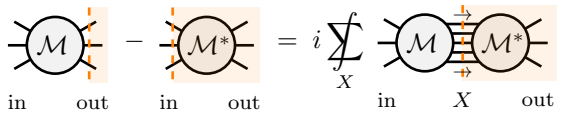
$$\mathcal{M}(\text{in} \to \text{out}) = \mathcal{M}(\overline{\text{out}} \to \overline{\text{in}})\,, \tag{16}$$

where the bar denotes applying charge and parity conjugation to the state. Likewise, CP invariance (or equivalently, T invariance) means

$$\mathcal{M}(\text{in} \to \text{out}) = \mathcal{M}(\overline{\text{in}} \to \overline{\text{out}}) = \mathcal{M}(\text{out} \to \text{in}) = \mathcal{M}(\overline{\text{out}} \to \overline{\text{in}})\,, \tag{17}$$

so the left-hand side of (15) can be replaced with $2i\,\mathrm{Im}\,\mathcal{M}(\text{in} \to \text{out})$ for CP-invariant theories.

Graphically, we depict (15) as




(19)

Only the states in $X$ can cross the dashed line, and the shading to the right of it indicates the complex conjugation. In particular, the two terms on the left-hand side can be interpreted as trivial cuts. In our pictures, energy always flows from left to right.

The unitarity equation (15) indicates that scattering amplitudes $\mathcal{M}$ have discontinuities whenever a new unitarity cut starts being kinematically allowed. These are called *normal thresholds* and are associated with singularities of $\mathcal{M}$, e.g., poles, or square-root or logarithmic branch points. Once iterated (by plugging in $T^\dagger$ from the left-hand side into the right-hand side), unitarity also implies existence of more complicated singularities, see, e.g., [64]. They are called *anomalous thresholds* or *Landau singularities* [14, 72, 79] and are associated with the cuts explored in Sec. IV and V. Indeed, anomalous thresholds are central objects of those sections.

We emphasize that, since (15) is a direct consequence of the unitarity of the $S$-matrix, (15) is valid fully non-perturbatively. In particular, if $|\text{in}\rangle = |\text{out}\rangle = |p\rangle$ is a single-particle state for a particle species with on-shell momentum $p$, then (15) implies that the imaginary part of the two-point function $\mathcal{M}(p \to p)$ is proportional to the total decay width of the particle. Similarly, if $|\text{in}\rangle = |\text{out}\rangle = |p_1 p_2\rangle$ is a two-particle state, then (15) reduces to the well-known *optical theorem*, which relates the imaginary part of the forward ($t = 0$) scattering amplitude $\mathcal{M}(p_1 p_2 \to p_1 p_2)$ to the total cross section,

$$\sigma^{\text{tot}}_{p_1p_2\to X}(s) = \frac{1}{2\sqrt{\lambda(s,p_1^2,p_2^2)}} \sum\!\!\!\!\!\!\int_X \hat{\delta}_{p_1p_2,\text{X}} |\mathcal{M}(p_1p_2\to X)|^2 = \frac{\text{Im}\,\mathcal{M}(p_1p_2\to p_1p_2)}{\sqrt{\lambda(s,p_1^2,p_2^2)}}\,, \tag{20}$$

where $s = (p_1 + p_2)^2$. The Källén function is

$$\lambda(a,b,c) := a^2 + b^2 + c^2 - 2ab - 2bc - 2ca\,. \tag{21}$$

We will come back to the relation between unitarity cuts and cross section computations in Sec. III B 4.

### 1. *Largest time equation*

Unitarity cuts are not independent of each other. One constraint they satisfy descends from the following algebraic identity:

$$\sum_{k=0}^{n} (-1)^k \sum_{\sigma\in P(k,n-k)} \bar{\mathcal{T}}[\mathcal{O}_{\sigma_1}\cdots\mathcal{O}_{\sigma_k}]\,\mathcal{T}[\mathcal{O}_{\sigma_{k+1}}\cdots\mathcal{O}_{\sigma_n}] = 0\,, \tag{22}$$

where $P(k, n-k)$ is the set of partitions of $n$ labels into two sets of size $k$ and $n-k$. Here, $\mathcal{T}$ and $\bar{\mathcal{T}}$ denote the time and anti-time ordering of the operators $\mathcal{O}_i = \mathcal{O}_i(x_i)$. After sandwiching (22) between vacuum states $\langle 0|$ and $|0\rangle$, inserting a complete basis of states $\mathbb{1} = \sum\!\!\!\!\!\!\int_X |X\rangle\langle X|$ between $\bar{\mathcal{T}}$ and $\mathcal{T}$, and applying the LSZ reduction on every term, we obtain the *largest time equation*. It states that the sum over all possible unitarity cuts (including trivial ones) with appropriate $\pm$ signs, vanishes. This statement was already illustrated in (19).

The largest time equation can be also derived in perturbation theory, see Sec. III B 2.

**Example 4** Consider $n = 4$. The identity (22) expands to

$$\mathcal{T}[\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3\mathcal{O}_4] + \bar{\mathcal{T}}[\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3\mathcal{O}_4] + \bar{\mathcal{T}}[\mathcal{O}_3\mathcal{O}_4]\mathcal{T}[\mathcal{O}_1\mathcal{O}_2] + (\text{all other partitions}) = 0\,. \tag{23}$$

Except for the first two, all the remaining terms have support only in one kinematic channel. For example, specializing to $s$-channel kinematics (with $p_1^0, p_2^0 > 0$ and $p_3^0, p_4^0 < 0$), only the first term in the second line of (23) gives a nonzero contribution. After applying the LSZ procedure and taking $\mathcal{O}_i$ to be operators that create the external states, it evaluates to

$$\prod_{i=1}^4 \lim_{p_i^2\to m_i^2} \sum\!\!\!\!\!\!\int_X \int \prod_{j=1}^4 \mathrm{d}^{\mathrm{D}}x_j\, \mathrm{e}^{-ip_j\cdot x_j}(\Box_{x_j}+m_j^2) \langle 0|\bar{\mathcal{T}}[\mathcal{O}_3\mathcal{O}_4]|X\rangle\langle X|\mathcal{T}[\mathcal{O}_1\mathcal{O}_2]|0\rangle = \sum\!\!\!\!\!\!\int_X \hat{\delta}_{p_1p_2,X}\,\mathcal{M}(p_1p_2\to X)\,\mathcal{M}(p_3p_4\to X)^*\,. \tag{24}$$

Likewise, the first two terms in (23) lead to the time-ordered and anti-time-ordered amplitudes $i\mathcal{M}$ and $-i\mathcal{M}^*$. The result is an instance of the unitarity equation (15).

## B. Perturbative unitarity

The computational power of the unitarity equation (15) is that it can be applied order-by-order in perturbation theory. In fact, it also applies to individual Feynman integrals. Perturbative unitarity cuts are natural extensions of those we encountered above.

**Definition 2** For any connected Feynman diagram, a unitarity cut is a curve (illustrated by a dashed line) through a subset $C$ of its edges that splits the diagram into two parts. Let us call the (possibly disconnected) set of edges to the left of the cut $C_L$ and likewise $C_R$ for those to the right. The unitarity cut of the

corresponding Feynman integral $\mathcal{I}$ is then defined as

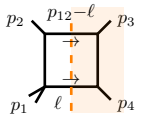
$$\text{Cut}^U_C \mathcal{I} := \int_{\ell_1,\dots\ell_L} \frac{\mathcal{N}_L}{\prod_{e\in C_L}(q_e^2 - m_e^2 + i\varepsilon)} \times \prod_{e\in C} \hat{\delta}^+(q_e^2 - m_e^2) \frac{\mathcal{N}_R^*}{\prod_{e\in C_R}(q_e^2 - m_e^2 - i\varepsilon)}\,, \tag{25}$$

where $\mathcal{N}_L$ and $\mathcal{N}_R$ are the sets of numerators to the left and right of the cut respectively.

Note that unitarity cuts are in general not analytic and depend on the kinematic channel they are evaluated in. We illustrate this property and the notation on the following examples.

**Example 5** Consider $\mathcal{I}^{1m}$ from Ex. 3 in the $s$-channel for which $s > 0$ and $t < 0$. In this example, we always have $\mathcal{N} = \mathcal{N}_L = \mathcal{N}_R = 1$. The unitarity cut through the edges $C = \{1, 3\}$, with $C_L = \{2\}$ and $C_R = \{4\}$, (with the edges labelled as in Ex. 3) gives:

$$\text{Cut}^U_{\{1,3\}} \mathcal{I}^{1m}\big|^{s>0}_{t<0} = $$

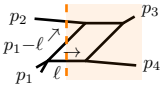



$$= \int_\ell \frac{\hat{\delta}^+(\ell^2)\hat{\delta}^+[(p_{12}-\ell)^2]}{[(p_1-\ell)^2+i\varepsilon][(p_{123}-\ell)^2-i\varepsilon]} = \frac{4i}{st}\Big[\sin(\pi\epsilon)\Big(\frac{1}{\epsilon^2} - \frac{\pi^2}{12}\Big)s^{-\epsilon} + \pi \log\Big(\frac{p_1^2+i\varepsilon - s}{t}\Big)\Big] + \mathcal{O}(\epsilon)\,. \tag{26}$$

Note that the same cut evaluated in the $t$-channel would be zero.

Similarly, the unitarity cut through $C = \{1, 2\}$ is nonzero only if $p_1^2 > 0$:

$$\text{Cut}^U_{\{1,2\}} \mathcal{I}^{1m}\big|^{s>0}_{t<0} = $$




$$= \frac{4i}{st}\Big[-\sin(\pi\epsilon)\Big(\frac{1}{\epsilon^2} - \frac{\pi^2}{12}\Big)(p_1^2)^{-\epsilon} + \pi \log\Big(\frac{st}{(p_1^2+i\varepsilon-s)(p_1^2-t)}\Big) + \mathcal{O}(\epsilon)\Big]\theta(p_1^2)\,. \tag{27}$$

If $p_1^2 \leqslant 0$, this cut is not kinematically allowed because the bottom-left vertex cannot be realized as an on-shell interaction. We therefore explicitly included the step function $\theta(p_1^2)$ above. Notice that the computation did not depend on the sign of energy of $p_2$, and therefore the same unitarity cut also contributes in the $p_1^2$-channel kinematics (the $1 \to 234$ decay process), where the energy of $p_2$ has the opposite sign. Note that when $p_1^2 > 0$, the $i\varepsilon$ in the argument of the logarithm is crucial to obtain the correct branches of the unitarity cut. It also illustrates that individual unitarity cuts can have nonvanishing imaginary parts.

In order to state the perturbative version of the unitarity equation (15) more concisely, let us introduce the notation for the sum of all kinematically-allowed unitarity cuts in a given physical region:

$$\text{Cut}^U_{s_J} \mathcal{I} := \sum_{C\in C_{s_J}} \pm\, \text{Cut}^U_C \mathcal{I}\,, \tag{28}$$

where $C_{s_J}$ denotes the set of all such admissible cuts and the sign depends on the number of disconnected components to the left and right of the cut, see, e.g., [82] for details. Unitarity in the form of (15) then gives[2]

$$2i\,\text{Im}\,\mathcal{I}\big|_{s_J\text{-channel}} = \text{Cut}^U_{s_J}\mathcal{I}\,. \tag{29}$$

Note once again that it is important to include all disconnected contributions. We illustrate this formula with the following examples.

**Example 6** Consider the running example of $\mathcal{I}^{1m}$ in the $s$-channel. Once again, the value of the imaginary part depends on the sign of $p_1^2$. Using the identity $(-x+i\varepsilon)^a - (-x-i\varepsilon)^a = 2i\sin(\pi a)x^a\theta(x)$, we find

$$2i\,\text{Im}\,\mathcal{I}^{1m}\big|^{s>0}_{p_1^2,t<0} = \frac{4i}{st}\Big[\sin(\pi\epsilon)\Big(\frac{1}{\epsilon^2} - \frac{\pi^2}{12}\Big)s^{-\epsilon} + \pi\log\Big(\frac{p_1^2 - s}{t}\Big)\Big] + \mathcal{O}(\epsilon) = \text{Cut}^U_{\{1,3\}}\mathcal{I}^{1m}\big|^{s>0}_{t,p_1^2<0}\,, \tag{30}$$

where the last step follows from comparing to (26) in Ex. 5. Similarly, we have

$$2i\,\text{Im}\,\mathcal{I}^{1m}\big|^{s,p_1^2>0}_{t<0} = \frac{4i}{st}\Big[\sin(\pi\epsilon)\Big(\frac{1}{\epsilon^2} - \frac{\pi^2}{12}\Big)\left(s^{-\epsilon} - (p_1^2)^{-\epsilon}\right) + \pi\log\Big(\frac{s}{p_1^2 - t}\Big)\Big] + \mathcal{O}(\epsilon) = \big(\text{Cut}^U_{\{1,3\}}\mathcal{I}^{1m} + \text{Cut}^U_{\{2,3\}}\mathcal{I}^{1m}\big)\big|^{s,p_1^2>0}_{t<0}\,. \tag{31}$$

Note that in the second case the real parts of the two unitarity cuts cancelled out in such a way that $2i\,\text{Im}\mathcal{I}^{1m}$ is purely imaginary.

### 1. *Proofs*

There are various proofs of (29) at different levels of rigor. The original derivation was sketched by Cutkosky

[2] Since most of the perturbative computations are done at the level of scalar "master integrals", we will now assume CP invariance.

based on unitarity arguments [40, Sec. II.C], which is why this method for computing imaginary parts is often referred to as the *Cutkosky (or cutting) rules*.

Another approach is that of the largest time equation [94], further reviewed in Sec. III A 1, which proves (29) as an algebraic identity using the combinatorics of retarded, advanced, and cut propagators. Yet another strategy is to use time-ordered (or old-fashioned) perturbation theory [92], where one splits every term in (29) into different time orderings of vertices and proves the equality term-by-term. See also [82] for a proof using contour manipulations.

Let us explain why (29) is intuitively true. One starts by engineering a Lagrangian that produces a single Feynman diagram at the leading order in the couplings. For example, the bubble diagram from Ex. 1 needs

$$\mathcal{L} = \sum_{i=1}^{6} \tfrac{1}{2}[(\partial\phi_i)^2 - m^2\phi_i^2] + (g_1\phi_1\phi_2 + g_2\phi_3\phi_4)\phi_5\phi_6\,, \tag{32}$$

such that $\mathcal{M}(\phi_1\phi_2 \to \phi_3\phi_4)$ at the leading order $\mathcal{O}(g_1 g_2)$ equals to the Feynman diagram from Ex. 1. Expanding the nonperturbative unitarity equation (15) order-by-order, the leading term reproduces (29).

### 2. *The largest time equation in perturbation theory*

The largest time equation can also be shown to hold diagram-by-diagram in perturbation theory [97]. We sketch the derivation for scalar Feynman diagrams for simplicity. We start with writing scalar Feynman propagators in position space,

$$\Delta_F(x) = \lim_{\varepsilon\to 0^+} \int \frac{\mathrm{d}^{\mathrm{D}}q}{(2\pi)^{\mathrm{D}}} \frac{i\mathrm{e}^{-iq\cdot x}}{q^2 - m^2 + i\varepsilon}\,, \tag{33}$$

as a sum of positive- and negative-frequency components,

$$\Delta_F(x) = \theta(x^0)\Delta^+(x) + \theta(-x^0)\Delta^-(x)\,, \tag{34}$$

where

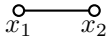
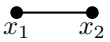
$$\Delta^{\pm}(x) = i\int \frac{\mathrm{d}^{\mathrm{D}}q}{(2\pi)^{\mathrm{D}}} \mathrm{e}^{-iq\cdot x}\,\hat{\delta}^{\pm}(q^2 - m^2)\,. \tag{35}$$

Similarly, the complex conjugated propagator satisfies $\Delta_F^*(x) = \theta(x^0)\Delta^- + \theta(-x^0)\Delta^+(x)$. The largest time equation for the simplest Feynman diagram consisting of just one propagator reads

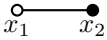
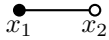
$$\Delta_F + \Delta_F^* = \Delta^+ + \Delta^-\,, \tag{36}$$

for any $x^0$. It easily follows from (34) and its conjugate, since if $x^0 > 0$, $\Delta_F = \Delta^+$ and $\Delta_F^* = \Delta^-$, and vice versa if $x^0 < 0$.

Anticipating a generalization to more complicated diagrams, we associate the different propagators to colorings of vertices,

$\Delta_F(x_2-x_1) =$  , $\Delta_F^*(x_2-x_1) =$ 



$$\Delta^+(x_2-x_1) = -$$  , $\Delta^-(x_2-x_1) = -$ 



(37)

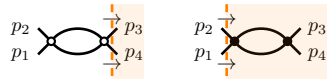
Following [97], we use the convention that black vertices come with a minus sign. Note that energy always flows from white to black vertices. With this notation, the generalization of (36) to a Feynman integral $\mathcal{I}$ with multiple propagators and V vertices becomes

$$\sum_{\text{colorings } c} \mathcal{I}_c = 0\,, \tag{38}$$

where each of the $2^{\mathrm{V}}$ colorings $c$ corresponds to assigning black or white color to a vertex, following which $\mathcal{I}_c$ is computed by using the rules in (37). To show that this equation holds, one first notes that one of the vertices, say $x_i$, must correspond to the one with the largest time, i.e., $x_i^0 > x_j^0$ for all $j \neq i$. The diagrams in (38) then cancel pairwise between those with $x_i$ colored black and those with $x_i$ white. This equation is simply (22) applied to an individual Feynman diagram. Note that here we have used sign conventions for $\mathcal{I}_c$ consistent with [97].

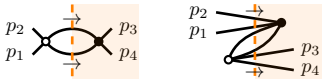
To tie (38) back to the unitarity equation from (15), we note that the terms with all white or all black vertices correspond to the amplitude and complex-conjugated amplitude, respectively. Note that the sets of white and black vertices might be each disconnected. Propagators connecting white and black vertices lie on the unitarity cut.

**Example 7** The bubble Feynman integral from Ex. 1 admits four distinct colorings of vertices,





(39a)





(39b)

Assuming that $p_1$, $p_2$ are incoming, and $p_3$, $p_4$ are outgoing, and recalling that energy flows from white to black vertices, the bottom-right configuration vanishes by energy conservation. The largest-time equation (38) says that

$$\mathcal{I}_{\text{bub}} - \mathcal{I}_{\text{bub}}^* - \int_\ell \hat{\delta}^+(\ell^2 - m^2)\hat{\delta}^+[(p_{12}-\ell)^2 - m^2] = 0\,. \tag{40}$$

Note that the extra minus sign between $\mathcal{I}_{\text{bub}}$ and $\mathcal{I}_{\text{bub}}^*$ traces its roots to the $i$ in the definition of the loop-momentum phase space in (4). We can recognize the integral as the unitarity cut $\mathrm{Cut}_C^U\,\mathcal{I}$ from Def. 2 for the bubble Feynman integral.

The unitarity cut of the scalar bubble is simply a phase-space integral for the two cut particles. The result of the integration in $\mathrm{D} = 2$ spacetime dimensions

is

$$\mathrm{Cut}_C^U \mathcal{I}_{\mathrm{bub}}^{\mathrm{D}=2} = \frac{4\pi i}{\sqrt{s(s-4m^2)}}\,, \tag{41}$$

which agrees with $2i\,\mathrm{Im}\,\mathcal{I}_{\mathrm{bub}}^{\mathrm{D}=2}$ for $s > 4m^2$ according to (6) and thus verifies (40).

#### 3. *Unstable particles*

If the field theory under consideration contains unstable particles (which is typically the case for phenomenologically relevant theories), the unitarity of the $S$-matrix, and consequently the notion of unitarity cuts, is more subtle. A consistent way of defining a unitary $S$-matrix in the presence of unstable particles is to only allow stable particles as asymptotic states, and to replace all propagators by their dressed version, where all 1PI contributions to the two-point function have been resummed [96]. This amounts to working with complex propagator poles, which can be implemented, e.g., when working in the so-called *complex mass scheme* [43–45]. Importantly, it can be shown that in this scheme all contributions from unitarity cuts which involve a cut propagator of an unstable particle cancel in the final contribution to the cut of the scattering amplitude [96]. For the nonperturbative perspective, see e.g., [65].

#### 4. *Reverse unitarity*

The optical theorem (20) allows one to relate total cross sections to the imaginary part of the forward scattering amplitude, which may be written as a sum of Feynman integrals. Since the imaginary parts of Feynman integrals are computable from unitarity cuts, the latter provide an effective tool to compute (total) cross sections in perturbation theory.

Unitarity cuts, and thus the imaginary parts of Feynman integrals, satisfy the same algebraic relations as the full Feynman integrals since they can be defined using the same integrand but different integration contours. We will come back to this point in Sec. V from a different perspective. In particular, the unitarity cuts satisfy the same linear *integration-by-parts* (IBP) relations [36, 95] (note that a diagram vanishes if an absent propagator is cut). IBP relations are the basis for computing Feynman integrals using the differential equations technique [68–70, 85]. Hence, this technique can be applied to compute unitarity cuts relevant to cross section computations. This result, which is known as *reverse unitarity* [5, 7], forms the basis for many analytic state-of-the-art computations for (total) cross sections for collider processes, e.g. [6, 35, 48–50, 52, 78].

### C. Branch cuts and dispersion relations

#### 1. *Mathematical preliminaries*

A multivalued analytic function $f(z)$ is typically introduced as a function that has *branch cuts* in the complex plane, i.e., a curve such that $f(z)$ behaves discontinuously when it is crossed. Multivalued functions and branch cuts in particular arise when integrating a function that has a pole in the complex plane. If the residue at the pole is nonvanishing, the result of the integration will depend on how the integration contour encircles the singularity. The notion of encircling a singularity will be made more precise in Sec. IV.

**Example 8** The prime example of a multivalued function is the (complex) logarithm,

$$\log(z) = \int_1^z \frac{\mathrm{d}t}{t}\,. \tag{42}$$

As long as $z$ is nonnegative, we can integrate along the straight line $[1, z]$. This defines the *principal sheet* of the (complex) logarithm function. For negative $z$, however, the integral along the straight path is ill-defined, because of the pole at $t = 0$. As a consequence, the function is discontinuous as we approach the negative real axis from above or below:

$$\lim_{\delta\to 0^+} \log(z \pm i\delta) = \log|z| \pm i\pi\,, \quad z < 0\,. \tag{43}$$

We therefore say that $\log z$ has a branch cut extending from $z = 0$ to $z = -\infty$ in the complex plane, and $z = 0, -\infty$ are called *branch points*.

We would like to stress the following important point: while it is conventional to choose the principal sheet of the logarithm so that the branch cut is along the negative real axis, this choice is completely arbitrary and conventional, and we could have chosen the branch cut to be any other semi-infinite non-self-intersecting curve extending from the origin. In the following, we will always assume that branch cuts for real branch points are located along the real axis.

**Definition 3** The discontinuity $\mathrm{Disc}_z f(z)$ of a function $f(z)$ across the real axis is defined as the difference

$$\mathrm{Disc}_z f(z) := \lim_{\delta\to 0^+} \left[f(z + i\delta) - f(z - i\delta)\right]\,. \tag{44}$$

The discontinuity is nonzero if $f(z)$ has a branch cut on the real line. For example, for the logarithm with the branch cut as in Ex. 8, we have $\mathrm{Disc}_z \log z = 2\pi i\,\theta(-z)$. Note that the discontinuity is in general not an analytic function, e.g., $\mathrm{Disc}_z \frac{1}{z} = -2\pi i\delta(z)$. For $\mathrm{Disc}_z \log z$ the non-analyticity is apparent from the step function, and $\mathrm{Disc}_z \log z$ is only piecewise analytic. More generally, a

discontinuity can be thought of as the difference between the value of a function after analytic continuation to the second sheet of (possibly multiple) branch cuts and the value on the original sheet. For this reason, $\mathrm{Disc}_z f(z)$ may have singularities that do not exist on the first sheet of $f(z)$. Finally, we mention that there is a connection between the discontinuity and the imaginary part. If $f(z)$ takes real values on some real interval, then Schwarz's reflection principle implies that for real $z$, $f(z - i\delta) = f(z + i\delta)^*$, and so in those cases we find

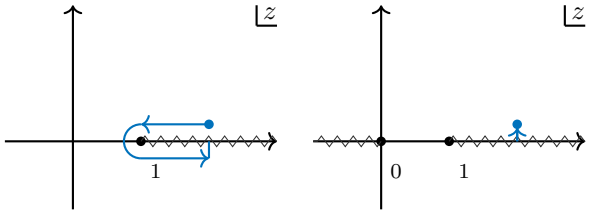
$$\mathrm{Disc}_z f(z) = \lim_{\delta \to 0^+} 2i \,\mathrm{Im}\, f(z + i\delta)\,. \tag{45}$$

For instance, $f(z) = \log(z)$ satisfies this identity because the logarithm is real for $z > 0$, but $f(z) = i\log(z)$ does not.

**Example 9** The dilogarithm

$$\mathrm{Li}_2(z) := -\int_0^z \frac{\log(1-t)}{t}\mathrm{d}t \tag{46}$$

on the principal sheet has a logarithmic branch point at $z = 1$ with a branch cut going to the right. On the second sheet of this branch cut, it has another logarithmic singularity at $z = 0$ which is not present on the first. The analytic structure on the two sheets can be illustrated as, respectively:

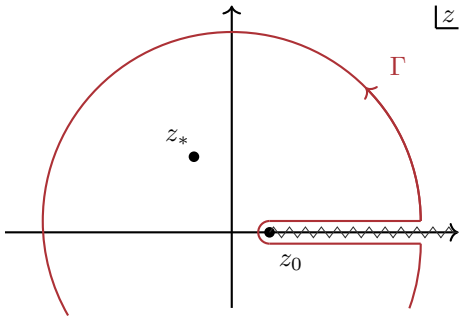



The discontinuity of the dilogarithm is computed as the difference between the blue dots on both sheets (in the limit as $\delta \to 0^+$):

$$\mathrm{Disc}_z \mathrm{Li}_2(z) = 2\pi i \log(z)\,\theta(z-1)\,. \tag{47}$$

The identity theorem guarantees that its extension from $z > 1$ to the $z$-plane is unique. The result (47) has a branch point at $z = 0$, which is a singularity of $\mathrm{Li}_2(z)$ on the second sheet.

It is well known that for multivalued complex functions it is possible to recover the full function from its discontinuity via a *dispersion relation*. To simplify the discussion, we assume that $f(z)$ vanishes at infinity and that it has a branch point at $z_0 \geqslant 0$ (and the branch cut extends along the half-line $[z_0, +\infty]$), but no other singularities. Consider the following contour $\Gamma$:




(48)

Applying the residue theorem, we find:

$$\frac{1}{2\pi i}\oint_\Gamma \frac{\mathrm{d}z}{z - z_*} f(z) = f(z_*)\,. \tag{49}$$

Since $f(z)$ vanishes at infinity, the contribution from the (infinitely large) circle vanishes, while the integration along the cuts gives rise to the discontinuity:

$$f(z_*) = \frac{1}{2\pi i}\int_{z_0}^{\infty} \frac{\mathrm{d}z}{z - z_*}\mathrm{Disc}_z f(z)\,. \tag{50}$$

It is possible to extend this result to the case where $f(z)$ has additional singularities by appropriately subtracting them before performing the contour integral.

### 2. *Dispersion relations for amplitudes*

We can apply the previous results to scattering amplitudes and individual Feynman integrals. A detailed discussion is given in [22]. Here we only give some highlights. In general, establishing the validity of dispersion relations depends on proving analyticity in the complexified kinematics, e.g., the center-of-mass energy squared $z = s$. Further, relating discontinuities to unitarity cuts hinges on the statement (45). The latter requires the existence of a *Euclidean region* in which the amplitude is real and free of branch cuts.

In the simplest case of off-shell two-point functions, dispersion relations are well established nonperturbatively. The following example illustrates that we can also use dispersion relations to compute individual Feynman integrals from their unitarity cuts, cf. [84, 97].

**Example 10** Consider the bubble diagram from Ex. 1 treated as a $1 \to 1$ process. In this case, the Euclidean region is $s < 0$, and so the relation between $\mathrm{Disc}_s$ and the imaginary part from (45) applies. We have seen in Ex. 7 that the unitarity cut $\mathrm{Cut}_C^U \mathcal{I}_{\mathrm{bub}}^{\mathrm{D}=2}$ computes the imaginary part of the bubble integral. Consequently, we also have

$$\mathrm{Disc}_s \mathcal{I}_{\mathrm{bub}}^{\mathrm{D}=2} = \mathrm{Cut}_C^U \mathcal{I}_{\mathrm{bub}}^{\mathrm{D}=2}\,. \tag{51}$$

Since we can recover a multivalued function from its discontinuity via dispersion relations, we see that we

can write the bubble integral as a dispersion integral over its unitarity cut:

$$\mathcal{I}^{\mathrm{D}=2}_{\mathrm{bub}}(s) = \frac{1}{2\pi i}\int_{4m^2}^{\infty}\frac{\mathrm{d}s'}{s'-s}\mathrm{Cut}^{U}_{C}\mathcal{I}^{\mathrm{D}=2}_{\mathrm{bub}}(s')\,. \tag{52}$$

Beyond two-point functions, the best-studied case is that of $2\to 2$ amplitudes $\mathcal{A}(s,t)$ of stable particles in gapped theories. Provided that the momentum transfer $-t$ is sufficiently low compared to the mass gap scale, dispersion relations can be written down in the complex $s$-plane minus the $s$-channel (where $\mathcal{A}(s,t) = \lim_{\delta\to 0^+}\mathcal{A}(s+i\delta,t)$) and $u$-channel (where $\mathcal{A}(s,t) = \lim_{\delta\to 0^+}\mathcal{A}(s-i\delta,t)$) branch cuts [74]. For general $t<0$, the domain of analyticity established with axiomatic quantum field theory [24] is not sufficiently large to write down dispersion relations in the same form nonperturbatively. For recent approaches with applications to bounds on Wilson coefficients, see, e.g., [4, 42]. We refer to [22, 23] for a more thorough discussion.

Complications arise for higher-point amplitudes or when external unstable particles are involved. For example, it might happen that $\mathcal{A}(s,t) \neq \lim_{\delta\to 0^+}\mathcal{A}(s+i\delta,t)$ in the $s$-channel or that $2i\,\mathrm{Im}\,\mathcal{A} \neq \mathrm{Disc}_s\mathcal{A}$, see [64] for examples. Dispersion relations for amplitudes with more than four external legs have not been exhaustively studied; see, e.g., [86, 89] for partial results.

## D. Double discontinuity

One can iterate the computation of discontinuities. Let us consider the simplest case of a double discontinuity for $2\to 2$ scattering. We can define the *double spectral density*:

$$\rho(s,t) := \lim_{\delta\to 0^+}\sum_{\eta_s,\eta_t\in\{\pm 1\}}\eta_s\,\eta_t\,\mathcal{A}(s+i\eta_s\delta,\,t+i\eta_t\delta)\,. \tag{53}$$

Often, the notation $\mathrm{Disc}_t\mathrm{Disc}_s\mathcal{A}(s,t)$ is used for the same quantity. This formula requires a level of care. It is predicated on the assumption of *maximal analyticity*, which states that $\mathcal{I}(s,t)$ is analytic everywhere in $(s,t)\in\mathbb{C}^2\setminus\{s,t,u\geqslant \mathfrak{m}^2\}$, where $\mathfrak{m}$ is the mass of the lightest (multi-particle) exchanged state. In this case, the four terms in (53) are to be understood as the boundary values of $\mathcal{I}(s,t)$ approached from different directions.

**Example 11** For the massless box $\mathcal{I}^{0m}$ from Ex. 2, only the term proportional to $\log(-s)\log(-t)$ contributes to the double spectral density, and it is given by

$$\begin{aligned}\rho^{0m}(s,t) &= \frac{2}{st}\lim_{\delta\to 0^+}\sum_{\eta_s,\eta_t\in\{\pm 1\}}\eta_s\eta_t\log(-s-i\eta_s\delta)\log(-t-i\eta_t\delta)\\ &= -\frac{8\pi^2}{st}\theta(s)\theta(t)\,.\end{aligned} \tag{54}$$

### 1. S-matrix bootstrap

The double spectral density $\rho(s,t)$ has support in unphysical kinematics, e.g., when $s,t>0$, which can be realized by embedding the external momenta in $(2,2)$ space-time signature. In this case, $\rho(s,t)$ can be related to a generalized version of a cut, but only in the sense of a specific analytic continuation of $\mathrm{Disc}_s\mathcal{I}$ to $t>0$ [74].

The *Mandelstam representation* is a natural generalization of dispersion relations to two complex variables, which relates $\mathcal{I}(s,t)$ to two-fold integrals of $\rho(s,t)$ and possibly subtraction terms [73]. It sets the starting point for the modern nonperturbative *S-matrix bootstrap* program, as it imposes a nonlinear constraint on the scattering amplitude embodying the assumption of maximal analyticity. See [39, 71] for reviews.

### 2. Steinmann relations

For the scattering of stable states in gapped theories, it is known that $\rho(s,t)=0$ within the physical region. This is the simplest instance of *Steinmann relations*, which state that double discontinuities in partially overlapping channels vanish when evaluated in physical kinematics:

$$\mathrm{Disc}_{s_J}\mathrm{Disc}_{s_I}\mathcal{A}\big|_{\mathrm{physical}} = 0 \tag{55}$$

if the sets $I,J$ are partially overlapping: neither disjoint nor contained in one another [90, 91]. This relation is trivial for $2\to 2$ scattering, since only one out of $s$, $t$, $u$ can be timelike and have a nonzero discontinuity. Steinmann relations may not hold in theories with massless particles, as illustrated in examples below.

**Example 12** We can interpret the diagram $\mathcal{I}^{2m}$ with $s>0$ and $t>0$ from Ex. 3 as a $2\to 4$ scattering process, if $p_1^2$ is large. Then, $s$ and $t$ are partially-overlapping physical kinematic channels, so we can check (55). The only terms that can potentially lead to a double discontinuity are

$$\mathcal{I}^{2m} \supset \frac{1}{st-p_1^2p_3^2}\left[2\mathrm{Li}_2\left(1-\frac{p_1^2p_3^2}{st}\right)-\log^2\left(\frac{s}{t}\right)\right]. \tag{56}$$

We have to perform an analytic continuation of the expression given in (10), making sure that the branch cuts of the resulting expression are at $s>0$ and $t>0$.

After doing so, we find that in physical-kinematic regions,

$$\mathrm{Disc}_s\mathrm{Disc}_t\mathcal{I}^{2m}\big|_{\mathrm{physical}} = 0\,, \tag{57}$$

in accordance with (55). Even though the Steinmann relations are not proven for this diagram since massless particles are involved, they happen to work in this case.

**Example 13** The Feynman integral $\mathcal{I}^{1m}$ from Ex. 3 can be made into a $2 \to 3$ scalar scattering process where $p_1$ is the sum of the momenta of the two incoming particles, and we take $p_1^2 > s > 0$, $p_1^2 > t > 0$ and $s + t > p_1^2$ for the kinematics to be physical. For such a configuration, the Steinmann relations from (55) do not hold, since

$$\mathrm{Disc}_s \mathrm{Disc}_t \, \mathcal{I}^{1m}\big|_{p_1^2 > s, t > 0} = -\frac{8\pi^2}{st} \,. \tag{58}$$

We computed this expression by noticing that the term proportional to $\log(s)\log(t)$ in (10) is the only one that contributes to the double discontinuity in $s$ and $t$, and after analytic continuation to $s > 0$ and $t > 0$, it becomes $\log(-s)\log(-t)$. Thus, the double discontinuity in the partially overlapping channels $s$ and $t$ does *not* vanish, even though $s > 0, t > 0$ correspond to a physical-kinematic channel. The breakdown of the Steinmann relations can be traced to the presence of massless particles.

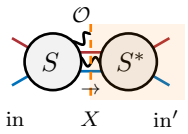
## E. Other applications

So far we have discussed unitarity cuts primarily in the context of quantum field theory and high-energy physics. Unitarity cuts also appear in many other applications throughout physics. Here, we collect some of them.

### 1. Gravitational radiation

The expectation value $\langle \mathrm{in}|\mathcal{O}|\mathrm{in}\rangle$ of an operator $\mathcal{O}$ in the in-state $|\mathrm{in}\rangle$ can be computed using quantum field theory techniques. A notable application is that of gravitational Bremsstrahlung emitted in a scattering event of two heavy bodies in the classical limit [67], where $\mathcal{O}$ (say, the Riemann tensor) is an outgoing mode sourced by a graviton field and $|\mathrm{in}\rangle = |p_1 p_2\rangle$. After inserting a complete basis of out-states $X$ before $\mathcal{O}$, the computation involves unitarity cuts:

$$\langle \mathrm{in}'|\mathcal{O}|\mathrm{in}\rangle = \sum\!\!\!\!\!\!\int_X$$




. (59)

Here, $S$ is the full $S$-matrix operator including the identity. The two heavy bodies are represented by solid, colored lines and the graviton by a wavy line. The sum over $X$ involves the two bodies themselves, but also gravitational radiation. Therefore, cut-based techniques are useful even in classical physics.

### 2. Thermal physics

At finite temperature or in out-of-equilibrium systems one is interested in computing thermal correlations functions, often evaluated in the Schwinger–Keldysh formalism [66, 88], where operators are distributed among two time-folds, one running forward and the other one backwards. One encounters correlations functions involving mixed time orderings of the type $\bar{\mathcal{T}}[\cdots]\mathcal{T}[\cdots]$ encountered in Sec. III A 1 and hence also generalized versions of unitarity cuts. We refer to [37] for details. Conversely, Schwinger–Keldysh techniques can be also used to study scattering amplitudes and other quantities at zero temperature [32, 59].

### 3. Cosmology and (anti-)de Sitter space

A number of generalizations of cuts have been introduced in curved backgrounds. For example, one can extend the optical theorem and the largest-time equation to (anti-)de Sitter space correlation functions and use them practically on Witten diagrams [77]. Different versions of cutting rules extending those from time-ordered perturbation theory [92] can be applied in cosmological backgrounds. See e.g. [10] for a review.

### 4. String theory

Unitarity cuts in superstring theory require separate treatment, because the discussion so far relied on the existence of local operators, while strings interact nonlocally. There have been two approaches to this problem. In the context of string field theory, where amplitudes are sums of infinitely many Feynman diagrams, Cutkosky rules can be derived in $\mathrm{D} > 4$ with a careful treatment of loop-momentum contours [82]. In the worldsheet formulation, unitarity cuts are seen as a modified contour prescription on the moduli space of Riemann surfaces and computed practically at one-loop level [53].

# IV. CUTS AND MONODROMIES

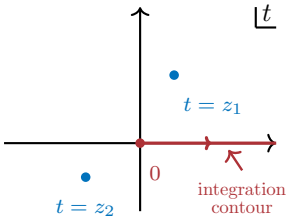
## A. Encircling singularities

In the previous section, we related unitarity cuts, imaginary parts, and total discontinuities. The total discontinuity corresponds to simultaneously crossing multiple branch cuts in the kinematic space. Using unitarity, we related this discontinuity to a sum over all allowed unitarity cuts in physical kinematics. In this section, we will instead ask a question from a more mathematical point of view: what happens if we analytically continue around an *individual* branch point in the complexified kinematic space?

**Example 14** Let us consider the following function of two complex variables $z_1$ and $z_2$:

$$g(z_1,z_2) := \int_0^\infty \frac{\mathrm{d}t}{(t-z_1)(t-z_2)} = \frac{\log(-z_2)-\log(-z_1)}{z_1-z_2}\,. \tag{60}$$

The integrand in the complex $t$ plane can be represented as

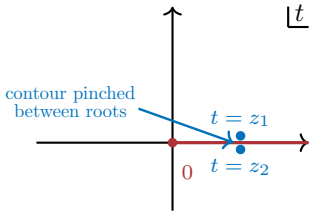



(61)

where the two simple poles of the integrand are indicated with blue dots. The integral representation of (60) is ambiguous when the denominator goes to zero, which happens when $t = z_1$ or $t = z_2$ for $z_1 \geqslant 0$ and $z_2 \geqslant 0$, i.e., when either of the two roots approaches the integration contour. In such situations, Cauchy's theorem can be used to deform the contour to avoid the singularities at $t = z_1$ and $t = z_2$. Thus, the function $g(z_1, z_2)$ has *branch cuts* for $z_1 \geqslant 0$ and $z_2 \geqslant 0$. Deforming the integration contour to avoid the singularities amounts to an analytic continuation of the integral in (60), which effectively moves away the branch cuts. For example, deforming the contour $t \to te^{-i\varepsilon t}$ makes the integral well-defined directly for positive $z_1$ and $z_2$ since the branch cuts moved to slightly below the positive $z_1$ and $z_2$ axes. More generally, the choice of the integration contour translates to the choice of how branch cuts of the resulting integral are placed.

However, when one of the singular points of the denominator, $t = z_1$ or $t = z_2$, coincides with the endpoints of the integral at $t = 0$ or $t \to \infty$, one cannot use a contour deformation to avoid the denominator becoming zero on the integration contour. Thus, these are places of potential *singularities* of the integral. In this example, the values $z_1 = 0, \infty$ and $z_2 = 0, \infty$ are precisely the singular points of the functions $\log(-z_1)$ and $\log(-z_2)$.

The denominator of the function in (60) has an additional singularity when $z_1 = z_2$. However, on the principal sheets of the logarithms, the numerator cancels the pole to render the function finite at this point. This is generically not the case on other sheets. One example is when one of the roots, say $z_1$, approaches a real and positive value from the upper half-plane, while the other root $z_2$ approaches the same value for the lower half-plane. The limit in that case is $\lim_{z_1\to z_2} g(z_1,z_2) = \frac{2\pi i}{z_1-z_2}$, so with this approach, $g(z_1, z_2)$ is singular at $z_1 = z_2$. What happens is that the integration contour is pinched between the two roots, so it cannot be deformed to simultaneously avoid both of them:




(62)

This is the simplest example of a *pinch singularity* in one complex dimension.

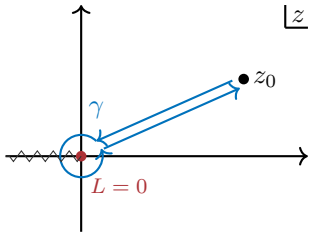
Let us study singularity of an analytic function at a given codimension-1 locus $L = 0$ in a multidimensional kinematic space. Singularity is any place where the function is ill-defined: it could be a pole or a branch point. We pick a one-dimensional linear subspace parametrized by a complex variable $z$, such that it intersects $L = 0$ transversely at $z = 0$. We write the function as $f(z; w)$ where $w$ represents all other arguments.

Analytically continuing around a singularity is made precise with a notion of monodromy.

**Definition 4** The *monodromy* (also called *monodromy variation*) $\Delta_{L=0} f(z_0; w)$ of a function $f(z; w)$ around $L = 0$, as described above, is defined as the difference

$$\Delta_{L=0} f(z_0; w) := f(z_0; w) - f(z_0 + \gamma; w)\,, \tag{63}$$

where $f(z_0+\gamma; w)$ is the result of analytically continuing $f(z_0; w)$ along a path $\gamma$ in the complex $z$-plane, starting at the base point $z = z_0$:



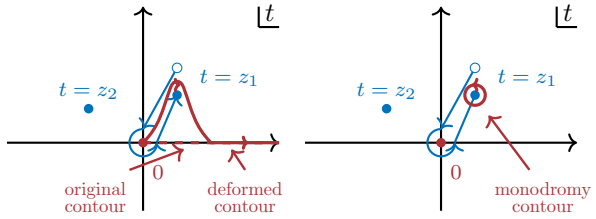

(64)

which goes from $z_0$ to a neighborhood of the origin, along an infinitesimal counterclockwise circle, and goes back to $z_0$, in such a way that no other singularities of $f(z; w)$ are enclosed.

If (63) is nonzero, the function $f(z; w)$ is said to have a *branch point* at $z = 0$. The value of the monodromy is independent of the choice of branch cuts.

The action of $\Delta_{L=0}$ is often called *variation* or simply *discontinuity* around $L = 0$, but we use the term *monodromy* to distinguish this action from Disc specified in Def. 3.

**Example 15** We can compute the monodromy around the branch point at $z_1 = 0$ of the function $g(z_1, z_2)$ from Ex. 14 by tracking the root at $t = z_1$ as $z_1$ encircles 0 in an counterclockwise direction. As the root in the $t$ complex plane encircles $t = 0$, the integration contour needs to be deformed accordingly:




(65)

As a result, the difference between the original integration contour and the one obtained after analytic continuation is a contour encircling the singularity at $t = z_1$, as depicted above. As a result, the monodromy is proportional to the residue around $t = z_1$,

$$\Delta_{z_1=0} g(z_1, z_2) = 2\pi i \,\mathrm{Res}_{t=z_1} \frac{1}{(t-z_1)(t-z_2)} = \frac{2\pi i}{z_1 - z_2}. \tag{66}$$

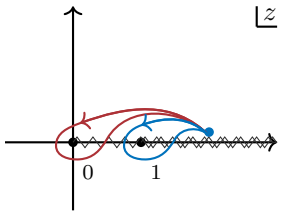
This result can also be obtained using the functional form of the right-hand side of (60), recognizing that a monodromy around the branch point of a logarithm is equal to $2\pi i$.

It is worth illustrating the distinction between the imaginary part, discontinuity, and monodromy on a simple example.

**Example 16** Consider the function

$$f(z) = i \log(-z)\sqrt{1-z}\,. \tag{67}$$

Its analytic structure in the $z$-plane is as follows:




(68)

We have

$$\mathrm{Im}\, f(z) = \begin{cases} \log(|z|)\sqrt{1-z} & \text{if} \quad z \leqslant 1\,, \\ -\pi\sqrt{z-1} & \text{if} \quad z > 1\,, \end{cases} \tag{69a}$$

$$\mathrm{Disc}_z f(z) = \begin{cases} 0 & \text{if} \quad z \leqslant 0\,, \\ 2\pi\sqrt{1-z} & \text{if} \quad 0 < z \leqslant 1\,, \\ 2\log(z)\sqrt{z-1} & \text{if} \quad z > 1\,, \end{cases} \tag{69b}$$

$$\Delta_{z=0} f(z) = 2\pi\sqrt{1-z}\,, \tag{69c}$$

$$\Delta_{z=1} f(z) = 2i \log(-z)\sqrt{1-z}\,. \tag{69d}$$

The two monodromies are computed using the red and blue contours in (68). Note that $\mathrm{Disc}_z f(z) = \Delta_{z=0} f(z)$ for $0 < z \leqslant 1$, because they are computed using the same contour. Moreover, $\Delta_{z=z_*} f(z)$ vanishes for any point $z_* \in \mathbb{C} \setminus \{0, 1\}$.

The monodromy is an analytic function, while the discontinuity is only piecewise analytic, i.e., its value jumps as the discontinuity contour encloses new singularities.

It is worth asking how to detect singularities of Feynman integrals in the first place, without direct computation. It is clear that they can arise only if something violent happens to the analytic properties of the *integrand*, e.g., when some of its poles collide (see Ex. 14) or when one of them degenerates from a hyperboloid to a lightcone. More generally, singularities can occur for special values of the kinematics for which the pole locus of the integrand changes topology.

Let us apply this logic to Feynman integrals $\mathcal{I}$ from (3). In fact, the combinatorics of pole degenerations can be streamlined by introducing Schwinger (or Feynman) parameters $\alpha_e$:

$$\mathcal{I} = \int_{\ell_1,\ldots,\ell_{\mathrm{L}}} \int_0^\infty \frac{\mathrm{d}^{\mathrm{E}}\alpha}{\mathrm{GL}(1)} \frac{\mathcal{N}\,\Gamma(\mathrm{E})}{\left[\sum_{e=1}^{\mathrm{E}} \alpha_e (q_e^2 - m_e^2 + i\varepsilon)\right]^{\mathrm{E}}}\,. \tag{70}$$

The invariance under GL(1) rescalings ($\alpha_e \to \lambda\alpha_e$ for any $\lambda > 0$) is commonly used to fix $\sum_{e=1}^{\mathrm{E}} \alpha_e = 1$.

Were it not for the insertion of $i\varepsilon$, the integral $\mathcal{I}$ would be ambiguous whenever its denominator goes to zero on the integration contour. Points in the external kinematic space for which this happens correspond to branch cuts (see Ex. 14). At such points, the $i\varepsilon$ prescription is needed to resolve which side of the branch cut the integral is evaluated on. While branch cuts of analytic functions can be placed arbitrarily without changing the function, the integral representation (70) selects how such cuts are placed in the kinematic space. This defines the *physical sheet* of $\mathcal{I}$.

The integral $\mathcal{I}$ may be genuinely singular when the denominator vanishes and, in addition, all the integration variables $\alpha_e$ and $\ell_a$ are either at the endpoint of the integration contour ($\alpha_e = 0$), or the contour is pinched (all derivatives $\frac{\partial}{\partial\alpha_e}$ and $\frac{\partial}{\partial\ell_a}$ of the denominator in (70) vanish). These conditions can be summarized in the textbook form of *Landau equations* [14, 72, 79]:

$$\alpha_e (q_e^2 - m_e^2) = 0 \qquad \text{for all } e = 1, 2, \ldots, \mathrm{E}\,, \tag{71a}$$

$$\sum_{e \in \text{loop } a} \pm\, \alpha_e q_e = 0 \qquad \text{for all } a = 1, 2, \ldots, \mathrm{L}\,. \tag{71b}$$

In the last equation, the sum goes over all edges $e$ forming a given loop $a$, and the sign in front of $\alpha_e q_e$ is $+$ if $q_e$ is oriented in the same direction as $\ell_a$ and $-$ otherwise.

The system (71) splits into different families depending on which subset of edges $\mathcal{E}$ satisfies $\alpha_e = 0$ for $e \in \mathcal{E}$. Each

such family corresponds to the *reduced diagram* obtained from the original Feynman diagram by contracting all the edges in $\mathcal{E}$ and putting the surviving propagators on their mass shell.

In general, the equations (71) impose conditions on the external kinematics. The most interesting case, though not the only one, is that in which they yield a codimension-1 variety $L = 0$ (UV and IR divergences would show up as codimension-0 solutions). Likewise, the solution set in the internal $(\alpha_e, \ell_a)$ variables might be isolated, but it could also form continuous manifolds.

The Landau equations capture a subset of possible singularities of massive Feynman integrals that occur at finite values of the loop momenta. Singularities at infinite loop momenta are sometimes referred to as *second and mixed-type* singularities [40, 54]. In the presence of massless particles, the Landau equations (71) must be treated more carefully since they mix kinematic singularities with infrared divergences. In a more careful treatment, one needs to consider the cases in which $\alpha_e$ and $q_e^2 - m_e^2$ in (71a) vanish, but at different relative rates [55]. In the mathematical language, one uses a compactification of the loop (and/or Schwinger parameter) integration domain. For related geometric constructions see [27, 80].

**Definition 5** The original integration contour of $\mathcal{I}$ passes through the solutions to the Landau equations (71) for which

$$\alpha_e \geqslant 0 \text{ and } \ell_a \in \mathbb{R}^{1,D-1}\,. \tag{72}$$

Singularities corresponding to such solutions are referred to as *physical-sheet singularities*. Further, those obtained for physical values of the external momenta $p_i$, are called *physical singularities*.

Physical singularities of a Feynman integral $\mathcal{I}$ can be interpreted as those for which the corresponding Feynman diagram can be realized as a *classical* spacetime process. To see why, one can define $\tau_e = m_e \alpha_e$ as the proper time of a particle of positive mass $m_e$. The equation (71b) then indicates that the distance between any two interaction points is given by the path taken by classical particles moving with momentum $q_e$. That is, physical singularities occur when all particles in the Feynman diagram are on shell and the interactions happen at fixed spacetime points $x_e$, with $x_{e+1} - x_e = \alpha_e q_e$. This perspective on the Landau equations is referred to as the *Coleman–Norton interpretation* [38].

Most of the singularities of Feynman integrals are not physical-sheet and are instead encountered when analytically continuing $\mathcal{I}$ on other Riemann sheets.

### B. Connection to cuts

We will also need a refined definition of cuts, also due to Cutkosky [40, Sec. II.A], which allow us to encircle any combination of poles in the integrand.

**Definition 6** For a Feynman integral $\mathcal{I}$ as in (3) and a subset of edges $\{e_1^\pm, e_2^\pm, \dots, e_k^\pm\}$ with associated signs $\pm$, the *cut* is defined as

$$\mathrm{Cut}_{e_1^\pm e_2^\pm \dots e_k^\pm} \mathcal{I} := \int_{\ell_1 \ell_2 \dots \ell_L} \frac{\mathcal{N} \prod_{i=1}^k \hat{\delta}^\pm(q_i^2 - m_i^2)}{\prod_{e=k+1}^E (q_e^2 - m_e^2 + i\varepsilon)}\,, \tag{73}$$

where the definitions of $q_e$, $\mathcal{N}$, $\ell$ and $\hat{\delta}$ were given in and around (3) and (12). The choice of $\pm$ sign of each edge selects which $\hat{\delta}^\pm$ is used.

The difference between this definition and the unitarity cut from Def. 2 is that we do not complex-conjugate any part of the diagram, and energies of cut particles may have different signs.

Let us now connect Def. 6 and the computation of monodromies from Def. 4. We consider the reduced diagram obtained imposing on-shell conditions $\delta^\pm(q_{e_i}^2 - m_{e_i}^2)$ on the set of edges $\{e_1^\pm, e_2^\pm, \dots, e_k^\pm\}$ and $\alpha_e = 0$ on the complementary set. Consider further the case in which the resulting singularity is physical and given by a codimension-1 variety $L = 0$ coming from a dimension-0 pinch in the loop-momentum space (i.e., where the Landau equations uniquely determine the Schwinger parameters and the loop momenta). Provided that the $L = 0$ singularity does not arise from any other pinch, we have

$$\Delta_{L=0} \mathcal{I} \overset{?}{=} \pm\, \mathrm{Cut}_{e_1^\pm e_2^\pm \dots e_k^\pm} \mathcal{I}\,, \tag{74}$$

where the $\pm$ sign has to do with the choice of orientation. The equality is marked with a question mark because it is only known to hold in favorable cases. A proof was given by Pham for finite Feynman integrals with generic masses where the above conditions are satisfied [81] (see also [15]). Despite this restriction, there are some examples of divergent integrals or those involving massless particles for which (74) still holds, as illustrated below in Ex. 21.

The mathematical theory used to study monodromies around individual branch points and which cuts they correspond to is called Picard–Lefschetz theory [81], which we discuss in Sec. IV F.

### C. Discontinuities and monodromies

The simplest way to connect individual discontinuities and cuts is to work in a kinematic region where only one cut is allowed, and using the unitarity equations from the previous section.

**Example 17** Going back to the example of the one-mass box from Ex. 3, whose cuts were computed in Ex. 5, we consider the kinematic region $s > 0$, $p_1^2 < 0$ and $t < 0$. The only nonvanishing cut is in the $s$ channel. We can also easily connect the discontinuity and monodromy, since there we are above the a branch cut starting at $s = 0$, but we are below the branch cuts

in both $t$ and $p_1^2$. In this case, we get

$$\Delta_{s=0}\mathcal{I}^{1m}(t,p_1^2;s) = \mathrm{Disc}_s\mathcal{I}^{1m}(s,t,p_1^2) \tag{75a}$$
$$= \mathrm{Cut}^{U}_{\{1,3\}}\mathcal{I}^{1m} \tag{75b}$$
$$= \mathrm{Cut}_{1^+3^+}\mathcal{I}^{1m}\,, \tag{75c}$$

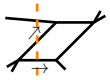
if we start at a kinematic point where $\mathrm{Re}(s) > 0$ and $s \to s + i\varepsilon$. Note that the unitarity cut $\mathrm{Cut}^{U}_{\{1,3\}}$ and $\mathrm{Cut}_{1^+3^+}$ are the same in this case, since they impose the same energy flow for the cut propagators and the $i\varepsilon$'s of remaining propagators do not matter as they can never go on shell. Hence, in this case where only one cut is allowed in the $s$-channel, we have shown that $\Delta_{s=0}\mathcal{I}^{1m}(t,p_1^2;s) = \mathrm{Cut}_{1^+3^+}\mathcal{I}^{1m}$.

With the definition of a monodromy, we can start in different kinematic regions, and hold a subset of the variables fixed. In some cases, they can be related to the cuts from Def. 6.

**Example 18** In the one-mass box example from Ex. 3, the monodromies in different channels can be computed from the full expression. We can, for example, start at a physical kinematic point where $s, p_1^2 > 0$, $t < 0$ and take $p_1^2 \to p_1^2 + i\varepsilon$, $s \to s + i\varepsilon$. If we analytically continue counterclockwise around $s = 0$ while holding $p_1^2$ fixed, we get

$$\Delta_{s=0}\mathcal{I}^{1m}\big|^{p_1^2,s>0}_{t<0} \tag{76}$$
$$= \frac{4i}{st}\left[\frac{\sin(\pi\epsilon)}{\epsilon^2}s^{-\epsilon} + \pi\log\left(\frac{p_1^2+i\varepsilon-s}{t}\right) + \mathcal{O}(\epsilon)\right],$$

where $p_1^2 > s$ by the assumption that $\Delta_{s=0}$ corresponds to a small circle around $s = 0$. In this case, we see from (26) that $\Delta_{s=0}\mathcal{I}^{1m}\big|^{p_1^2,s>0}_{t<0} = \mathrm{Cut}_{1^+3^+}\mathcal{I}^{1m}$. Similarly, we can compute the monodromy in $p_1^2$, $\Delta_{p_1^2=0}\mathcal{I}^{1m}\big|^{p_1^2,s>0}_{t<0}$, which agrees with the following cut (using the labeling from Ex. 5):

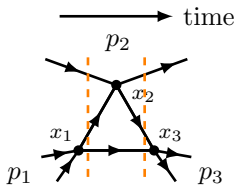
$$\mathrm{Cut}_{1^+2^+}\mathcal{I}^{1m}\big|^{p_1^2,s>0}_{t<0} = \tag{77}$$

$$= \int_\ell \frac{\delta^+(\ell^2)\delta^+[(p_1-\ell)^2]}{[(p_{12}-\ell)^2+i\varepsilon](p_{123}-\ell)^2}$$
$$= \frac{4i}{st}\left[\frac{-\sin(\pi\epsilon)}{\epsilon^2}(p_1^2)^{-\epsilon} + \pi\log\left(\frac{st}{(p_1^2-i\varepsilon-s)(p_1^2-t)}\right) + \mathcal{O}(\epsilon)\right].$$

In this simple one-loop example, there was a way to relate the individual monodromies around $s = 0$ and $p_1^2 = 0$ to the cuts in the $s$-channel and $p_1^2$, respectively, with the proper $i\varepsilon$ prescription for the remaining propagators. We should not be misled into generalizing from this example: it is not generally true that individual unitarity cuts agree with individual discontinuities. This one-loop example is however deceivingly simple, since the last propagator $(\ell+p_{123})^2$ can never go on shell, so we can fully separate out the dependence on the imaginary parts of $p_1^2$ and $s$. Beyond such examples, no such correspondence is known to hold.

### D. Cuts for anomalous thresholds

The correspondence between monodromies around individual branch points and cuts in different kinematic channels is much more intricate, as previously discussed around (74). As a first step, we must solve the Landau equations to find the points in the kinematic space that are potential singularities or branch points of the scattering amplitude. We sometimes find that such solutions are physical.

**Example 19** In a $3 \to 3$ scattering process of particles with mass $m$, the location of the anomalous threshold can be found by solving for the momenta at which the classical scattering process can occur, i.e., when it factorizes into a product of three $2 \to 2$ on-shell scattering processes as follows:



(78)

Such a classical configuration occurs when the momenta $q_i$ of all the cut particles are put on shell, $q_i^2 = m^2$ with $i = 1,2,3$, and interaction vertices happen at fixed space-time points $x_i$. Since $x_{i+1} - x_i = \alpha_i q_i$ (with $x_4 = x_1$) for Schwinger proper time $\alpha_i$, the latter constraint reads

$$\sum_{i=1}^{3}\alpha_i q_i = 0\,. \tag{79}$$

By dotting this constraint with $q_j$, we can rephrase it in terms of the $3\times 3$ matrix $y$ with entries:

$$y_{ij} = -\frac{q_i\cdot q_j}{m^2} = \frac{p_{ij}^2 - 2m^2}{2m^2} \quad i,j = 1,2,3\,. \tag{80}$$

Equation (79) implies that the determinant of $y$ vanishes:

$$\det y = \frac{p_1^2p_2^2p_3^2 + m^2\lambda}{4m^6} = 0\,, \tag{81}$$

where the $\lambda = \lambda(p_1^2,p_2^2,p_3^2)$ is the Källén function from (21). This is a second-order equation in $p_2^2$, and

when $p_1^2 > 4m^2$ and $p_3^2 > 4m^2$, one of the solutions in $p_2^2$ is found to be a physical singularity, i.e., $\alpha_i \geqslant 0$.

We follow up with an example of a concrete formula in which such singularities are displayed.

**Example 20** In D = 3 space-time dimensions in the region $y_{ij} < -1$, the triangle diagram from Ex. 19 is given by

$$\mathcal{I}^{\text{D=3}}_{\text{tri}}(y_{12}, y_{23}, y_{13})\Big|_{y_{ij}<-1} = \frac{\sqrt{\pi}}{4\sqrt{\det y}} \log\left(\frac{y_{12}+y_{23}y_{13}+\sqrt{\det y}}{y_{12}+y_{23}y_{13}-\sqrt{\det y}} \times \frac{y_{23}+y_{12}y_{13}+\sqrt{\det y}}{y_{23}+y_{12}y_{13}-\sqrt{\det y}} \times \frac{y_{13}+y_{23}y_{12}+\sqrt{\det y}}{y_{13}+y_{23}y_{12}-\sqrt{\det y}}\right) \tag{82}$$

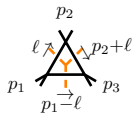
with $y_{ij}$ given in (80). In this kinematic region, there is no singularity at $\det y = 0$, since the naive square-root branch cuts cancel between the prefactor and the square roots inside the logarithm.

After analytic continuation to other physical-kinematic regions, the logarithms can pick up imaginary parts, and the singularity at $\det y = 0$ can become a physical singularity.

In cases where anomalous thresholds can be realized in physical kinematics, one can compute cuts corresponding to putting all particles in the loop simultaneously on shell.

**Example 21** For the triangle integral we can compute unitarity cuts in the $p_1^2$, $p_2^2$ and $p_3^2$ channels using the same methods as in Ex. 5 and 7. We also have a new type of a cut in addition: the one where all particles in the triangle can simultaneously go on shell. In the kinematic channel where $p_1^2 > 4m^2$, $p_2^2 < 0$ and $p_3^2 > 4m^2$, it evaluates to

$$\text{Cut}_{1+2+3+}\mathcal{I}^{\text{D=3}}_{\text{tri}} = \tag{83}$$




$$= \int_\ell \hat{\delta}^+(\ell^2-m^2)\hat{\delta}^+[(p_1-\ell)^2-m^2]\hat{\delta}^+[(\ell+p_2)^2-m^2] = \frac{2\pi^{3/2}}{\sqrt{-\det y}}\,.$$

It is natural to ask whether this cut corresponds to a monodromy in the kinematic space. In this simple example, it turns out that the cut is precisely the monodromy around the anomalous threshold at $\det y = 0$. Let us verify it.

The value of the monodromy around $\det y = 0$ depends heavily on the value of the kinematics. As an example, let us assume that we fix $y_{12}$ and $y_{23}$, and analytically continue $y_{13}$ in the counterclockwise direction such that it encircles $\det y = 0$. This amounts to fixing $p_1^2$ and $p_3^2$ and analytically continuing in $p_2^2$. Since we already established in Ex. 20 that $\det y = 0$ is not a singularity of $\mathcal{I}^{\text{D=3}}_{\text{tri}}$ when all the invariants $p_i^2$ are negative, we get

$$\Delta_{\det y=0}\,\mathcal{I}^{\text{D=3}}_{\text{tri}}\big|_{p_1^2,p_2^2,p_3^2<0} = 0\,. \tag{84}$$

In this kinematic region, the cut is also zero. The reason why there is no singularity at $\det y = 0$ when all the invariants are negative can be explained with the physical intuition mentioned before: we cannot factorize the process into on-shell processes unless at least two of the invariants are positive. However, if we instead fix $p_1^2, p_3^2 > 4m^2$, with $p_1^2 \to p_1^2 + i\varepsilon$ and $p_3^2 \to p_3^2 + i\varepsilon$, and take a monodromy around the branch of $\det y = 0$ which is physical (i.e., its solutions of (79) have $\alpha_i \geqslant 0$ in physical kinematics), we get

$$\Delta_{\det y=0}\,\mathcal{I}^{\text{D=3}}_{\text{tri}}\big|^{p_1^2,p_3^2>4m^2}_{p_2^2<0} = \frac{2\pi^{3/2}}{\sqrt{-\det y}}\,. \tag{85}$$

When $p_1^2 > 4m^2$ and $p_3^2 > 4m^2$, the anomalous-threshold process can be allowed in physical kinematics as in (78), and the monodromy has support. Furthermore, the monodromy is, in this case, equal to the cut, and thus

$$\Delta_{\det y=0}\,\mathcal{I}^{\text{D=3}}_{\text{tri}}\big|^{p_1^2,p_3^2>4m^2}_{p_2^2<0} \tag{86}$$

$$= \text{Cut}_{1+2+3+}\mathcal{I}^{\text{D=3}}_{\text{tri}}\big|^{p_1^2,p_3^2>4m^2}_{p_2^2<0}\,. \tag{87}$$

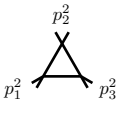
Note that this last expression does not have a hope of being obtained by only using the unitarity equation in (13). We are relying on keeping $p_1^2$ and $p_3^2$ fixed to be positive and with a small positive imaginary part, and only analytically continuing in $p_2^2$. This is a very different manipulation from the one using unitarity constraints, where we cross branch cuts of all invariants involved.

### 1. *Non-Mandelstam variables*

The discussion of analytic structure depends heavily on the choice of variables for the problem.

**Example 22** Consider the triangle Feynman integral for a $3 \to 3$ scattering process in a theory of scalar massless particles. The external kinematic invariants are allowed to take any complex values. This integral evaluates to

$$\mathcal{I}^{\text{D=4}}_{\text{tri}} = \tag{}$$




$$= \frac{2}{\sqrt{\lambda}}\left[\text{Li}_2\left(\frac{p_1^2+p_2^2-p_3^2+\sqrt{\lambda}}{2p_1^2}\right)\right.$$

$$- \mathrm{Li}_2\left(\frac{p_1^2+p_2^2-p_3^2-\sqrt{\lambda}}{2p_1^2}\right) \tag{88}$$
$$+\frac{1}{2}\log\left(\frac{p_2^2}{p_1^2}\right)\log\left(\frac{p_3^2+p_1^2-p_2^2-\sqrt{\lambda}}{p_3^2+p_1^2-p_2^2+\sqrt{\lambda}}\right)\Bigg],$$

where $\lambda = \lambda(p_1^2, p_2^2, p_3^2)$ is the Källén function given in (21). The expression is valid in the kinematic region where $p_1^2, p_2^2, p_3^2 < 0$ and $\lambda < 0$, when the square root, $\mathrm{Li}_2$ and log are taken to be on their principal branches. We see from the expression that it has branch points at $p_1^2 = 0$, $p_2^2 = 0$, $p_3^2 = 0$ and at $\lambda = 0$. The last one, $\lambda = 0$, is a second-type singularity and does not correspond to the triangle anomalous threshold.

**Example 23** The triangle integral from Ex. 22 can also be written as

$$\mathcal{I}^{\mathrm{D}=4}_{\mathrm{tri}} = \frac{2}{p_1^2(z-\bar{z})}\left[\mathrm{Li}_2(z)-\mathrm{Li}_2(\bar{z})+\frac{1}{2}\log(z\bar{z})\log\left(\frac{1-z}{1-\bar{z}}\right)\right], \tag{89}$$

where $z$ and $\bar{z}$ are two independent variables defined by $z\bar{z} = p_2^2/p_1^2$ and $(1-z)(1-\bar{z}) = p_3^2/p_1^2$. We take $\bar{z} < z$ without loss of generality, and then the expression above is valid in the region where $0 < \bar{z} < z < 1$. In terms of the Mandelstam invariants, this region is the one where either $\{p_1^2, p_2^2, p_3^2 < 0, \lambda > 0\}$ or $\{p_1^2, p_2^2, p_3^2 > 0, \lambda > 0\}$. In other kinematic regions, $z$ and $\bar{z}$ pick up $\pm i\varepsilon$'s to capture the correct analytic continuation of the logarithms and dilogarithms in (89).

The triangle anomalous threshold at $\lambda = z - \bar{z} = 0$, which is a *square root* singularity of the expression in (88) has become a *pole* at $z - \bar{z} = 0$ in the $z$, $\bar{z}$ variable representation in (89). Thus, the change of variables (which itself is branched) has washed out the square root branch cut, and changed the singularity into a pole instead. This phenomenon is not surprising; it is a common feature of analytic functions that changes of variables generally modifies the analytic structure.

Let us briefly comment on the analytic structure of the one-loop triangle in (89) seen as a function of $z$ and $\bar{z}$. First, we note that in the region where $\lambda < 0$, $z$ and $\bar{z}$ are complex conjugates of each other. In that region we may see (89) as a function of the single complex variable $z$. Although the individual logarithms and dilogarithms in (89) have branch cuts, (89) is a single-valued function of the complex variable $z$, cf., e.g., [26, 33]. There is no contradiction with the fact that $\mathcal{I}^{\mathrm{D}=4}_{\mathrm{tri}}$ must have branch cuts dictated by unitarity. In fact, the branch cut structure of $\mathcal{I}^{\mathrm{D}=4}_{\mathrm{tri}}$ seen as a function of the invariants $p_i^2$ translates directly into the fact that (89) is a single-valued function of the complex variable $z$ in the region where $\lambda < 0$.

### 2. *Loop-tree theorems*

There are multiple theorems relating loop-level scattering amplitudes to on-shell phase integrals over tree amplitudes. Let us explain the simplest of them: the Feynman tree theorem [56]. Consider a one-loop diagram computed with retarded (instead of Feynman) propagators $G_R(q,m) = 1/(q^2 - m^2 + i\varepsilon q^0)$:

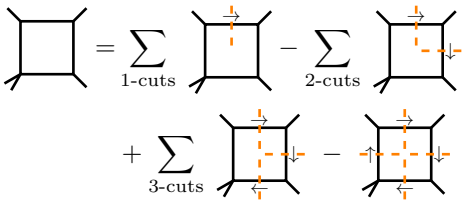
$$\int_\ell \mathcal{N} \prod_{i=0}^{n-1} G_R(\ell + p_{1\cdots i}, m_i) = 0\,. \tag{90}$$

Causality dictates that such an amplitude has to vanish, as otherwise we would create a closed time-like loop. One way to see it practically is by deforming the loop-energy $\ell^0$ contour into the upper half-plane, which vanishes since all the poles are located in the lower half-plane.

Plugging in the relation between retarded and Feynman propagators $G_F(q,m) = 1/(q^2 - m^2 + i\varepsilon)$,

$$G_R(q,m) = G_F(q,m) - \hat{\delta}^+(q^2 - m^2) \tag{91}$$

into (90), one obtains a relation between the time-ordered amplitude, all of its single cuts, double cuts, etc. Note here that we require real external kinematics. For example, we have



(92)

Many of the cuts might vanish due to kinematical reasons, e.g., the final term always vanishes since it involves positive energy circulating in a loop. The same result can of course be obtained by carefully deforming the $\ell^0$-contour in the original Feynman integral. This strategy can be extended to higher loops and used in numerical evaluations of amplitudes [30, 41, 87].

The distinction between the Feynman tree theorem and the optical theorem is that the former involves the full integral (not just its imaginary part) and the cuts do not have a one-directional flow of energy.

**Example 24** Consider the bubble diagram in Ex. 1 in the $s$-channel kinematics. In this case, the Feynman tree theorem reads

$\mathcal{I}^{\mathrm{D}=2}_{\mathrm{bub}} =$  (93)

As above, the maximal cut does not contribute. In $\mathrm{D} = 2$ dimensions, the first cut evaluates to

 $= \int_\ell \frac{\hat{\delta}^+(\ell^2 - m^2)}{(p_{12} - \ell)^2 - m^2 + i\varepsilon}$ (94)

$$= \frac{i}{\sqrt{s(4m^2-s)}}\left[\log\left(\frac{\sqrt{4m^2-s}-i\sqrt{s}}{\sqrt{4m^2-s}+i\sqrt{s}}\right) - i\pi\right],$$

where we take $s \to s+i\varepsilon$. The remaining cut doubles the contribution from the first term in the square brackets and cancels the $i\pi$ term, such that the identity (93) is satisfied.

## E. Symbols and monodromies

### 1. Definitions

A convenient way of probing the analytic structure of (multiple) polylogarithms (cf. Ex. 9) is through the *symbol* [28, 34, 51, 61, 62]. We refer to [47] for a review and notation. Let us say that a symbol of a function $f$ is given by

$$\mathcal{S}[f] = \sum_{i_1,i_2,\ldots,i_k} c_{i_1,i_2,\ldots,i_k} a_{i_1} \otimes a_{i_2} \otimes \cdots \otimes a_{i_k}\,, \tag{95}$$

where $a_i$ are the symbol letters and $c_{i_1,i_2,\ldots,i_k}$ are the coefficients, which are typically rational numbers. Then, the symbol of the monodromy around a given singularity at $a_j = 0$ is given by

$$\mathcal{S}[\Delta_{a_j=0} f] = 2\pi i \sum_{i_2,\ldots,i_k} c_{j,i_2,\ldots,i_k} a_{i_2} \otimes \cdots \otimes a_{i_k}\,. \tag{96}$$

In other words, only terms in the symbol which have $a_j$ in the first entry survive. A similar statement can be made at the level of the coaction [25, 46]. More generally, we can iterate this procedure, and we see that the sequence of $a_j$'s in a symbol captures the effect of computing iterated monodromies, e.g.,

$$\mathcal{S}[\Delta_{a_l=0}\Delta_{a_j=0} f] = (2\pi i)^2 \sum_{i_3,\ldots,i_k} c_{j,l,i_3,\ldots,i_k} a_{i_3} \otimes \cdots \otimes a_{i_k}\,. \tag{97}$$

Note that while the symbol captures the (iterated) monodromies of a polylogarithmic function, its value is independent of the Riemann sheet on which the function is evaluated.

**Example 25** Consider the dilogarithm $\mathrm{Li}_2(z)$ from Ex. 9. Its symbol is

$$\mathcal{S}[\mathrm{Li}_2(z)] = -(1-z)\otimes z\,. \tag{98}$$

On the other hand, computing the monodromy around $z = 1$ gives

$$\Delta_{z=1}\mathrm{Li}_2(z) = 2\pi i \log(z)\,, \tag{99}$$

which results in

$$\mathcal{S}[\Delta_{z=1}\mathrm{Li}_2(z)] = 2\pi i\, z\,. \tag{100}$$

This expression is an instance of (96).

Note that we can read off the analytic structure of the dilogarithm directly from the symbol in (98). The first entry of the symbol vanishes or becomes infinite when $z = 1$ or $z = \infty$, which reflects the fact that on its principal sheet (where the function is real), the dilogarithm only has a branch cut starting at $z = 1$ along the positive real axis, cf. Ex. 9. The second entry vanishes for $z = 0$, and encodes the fact that on the second sheet, the dilogarithm also has a branch point at $z = 0$.

### 2. Symbols of Feynman integrals

If a scattering amplitude or an individual Feynman integral can be expressed in terms of polylogarithms, we can compute its symbol. Since the symbol captures the (iterated) monodromies, the possible symbol letters that can arise are constrained by the position of the branch points.

In particular, the first entries in the symbol vanish only at physical-sheet singularities. In massless theories it is common to use the *first entry condition* [58], which says that the first entries in a symbol can only be Mandelstam invariants. Since the symbol also encodes iterated monodromies, the second entries are constrained by relations inspired by those of Steinmann (55) extrapolated to massless theories and with the monodromy $\Delta$ instead of the discontinuity Disc [31]:

$$\Delta_{s_J}\Delta_{s_I}\mathcal{A} = 0\,. \tag{101}$$

Generalizations to multiple iterated monodromies exist in some cases [11, 63].

The constraints on the symbol allow one to put constraints on the coefficients $c$ in (95). For example, this strategy becomes useful in the perturbative bootstrap of $\mathcal{N} = 4$ SYM amplitudes in the planar limit, see, e.g., [31].

**Example 26** We illustrate the structure of the symbol on the example of the "two-mass easy box" function from Ex. 3. If we define $\mathcal{J}^{2m} = \frac{1}{2}(st - p_1^2 p_3^2)\mathcal{I}^{2m}$, then the symbol is given by

$$\begin{aligned}\mathcal{S}\big[\mathcal{J}^{2m}\big] = &-\frac{1}{\epsilon}(s+t-p_1^2-p_3^2)\\ &+\big[s\otimes s - s\otimes(s-p_1^2) - s\otimes(s-p_3^2)\\ &\quad + s\otimes(st-p_1^2p_3^2) + (s\leftrightarrow t)\big] +\\ &-\big[p_1^2\otimes p_1^2 - p_1^2\otimes(s-p_1^2) - p_1^2\otimes(t-p_1^2)\\ &\quad + p_1^2\otimes(st-p_1^2p_3^2) + (p_1^2\leftrightarrow p_3^2)\big] + \mathcal{O}(\epsilon)\,.\end{aligned} \tag{102}$$

We see that the first entries are all drawn from the set $\{s, t, p_1^2, p_3^2\}$, in agreement with the first entry condition. We also know from Ex. 12 that the Steinmann relations

(in the from of (101)) apply. Equation (97) then implies

$$\mathcal{S}\big[\Delta_s\Delta_t\mathcal{J}^{2m}\big] = 0\,. \tag{103}$$

This equation is indeed true because the symbol in (102) does not contain any terms of the form $s\otimes t$ or $t\otimes s$.

**Example 27** For the one-mass box function $\mathcal{I}^{1m}$, we define $\mathcal{I}^{1m} = \frac{1}{2}st\mathcal{J}^{2m}$. The symbol is given by

$$\begin{aligned}\mathcal{S}\big[\mathcal{J}^{1m}\big] &= \frac{1}{\epsilon^2} + \frac{1}{\epsilon}\big(p_1^2 - s - t\big) + \Big[p_1^2\otimes(p_1^2-s)\\ &\quad - p_1^2\otimes s - s\otimes(p_1^2-s) - \frac{1}{2}p_1^2\otimes p_1^2 + t\otimes s\\ &\quad + s\otimes s + (s\leftrightarrow t)\Big] + \mathcal{O}(\epsilon)\,.\end{aligned} \tag{104}$$

The first entry condition is clearly satisfied. This time, however, the symbol contains terms of the form $s\otimes t$ and $t\otimes s$, which implies

$$\mathcal{S}\big[\Delta_s\Delta_t\mathcal{J}^{1m}\big] \neq 0\,. \tag{105}$$

This is a manifestation of the Steinmann relations not being satisfied for the one-mass box function, cf. Ex. 13.

### F. Picard–Lefschetz theory for Feynman integrals

There have been attempts to systematize computations of monodromies using tools from algebraic topology, with a goal of putting (74) on a firm mathematical footing. Consider the favorable cases discussed above (74), in which a pinch associated with $L=0$ is an isolated point. Notice that a Feynman integral $\mathcal{I}$, as written in (3), is an integral over a specific contour $\Gamma$ implementing the $i\varepsilon$ prescription. The Picard–Lefschetz theorem prescribes how the monodromy $\Delta_{L=0}\mathcal{I}$ around a branch point $L=0$ can be rewritten in terms of a new integration contour $c$:

$$\Delta_{L=0}\mathcal{I} = N_{\Gamma,c}\int_c \frac{\mathcal{N}}{\prod_{e=1}^{\mathrm{E}}(q_e^2-m_e^2)}\,, \tag{106}$$

where $N_{\Gamma,c}\in\mathbb{Z}$ is called the *intersection number* or *Kronecker index* of $\Gamma$ and $c$. Here, $c$ is the *vanishing cycle* homologous to a sphere around the pinch point [81].

In practice, one starts with identifying which denominators pinch the integration contour at a physical singularity corresponding to $L=0$. Labeling the denominators which take part in this pinch as $1,2,\ldots,k$, the result is

$$\Delta_{L=0}\mathcal{I} = N_{\Gamma,c}\int_{\ell_1,\ell_2\ldots,\ell_{\mathrm{L}}} \frac{\mathcal{N}\prod_{e=1}^{k}\hat{\delta}^{\pm}(q_e^2-m_e^2)}{\prod_{e'=k+1}^{\mathrm{E}}(q_{e'}^2-m_{e'}^2+i\varepsilon)}\,, \tag{107}$$

where the signs $\pm$ are determined by which of the two roots of the polynomial $q_e^2-m_e^2$ takes part in the pinch. The monodromy is therefore proportional to the individual cuts $\mathrm{Cut}_{e_1^{\pm}e_2^{\pm}\ldots e_k^{\pm}}\mathcal{I}$ defined in (73). For the cases under consideration, one can show that the intersection number equals either $N_{\Gamma,c}=\pm1$ or 0, depending on whether the branch point corresponding to the physical singularity is actually present or not, see, e.g., [18, 60] for examples.

Some extensions to cases with non-generic masses, where pinches could become non-isolated or degenerate, exist for finite integrals [11].

## V. GENERALIZED CUTS

### A. Cuts of arbitrary sets of edges

Cutting a diagram can be considered as an operation per se, liberated from any interpretation as a specific imaginary part or discontinuity. The operation is the restriction of some subset of propagators to their mass shells, using delta functions or residues, with or without energy-flow conditions. It is an operation on an integral (integrand and/or contour). The idea of performing such an operation is to apply a specific generalized cut on two sides of an equation to obtain a set of constraints.

**Definition 7** The generalized unitarity cut through the set of edges $C$ of the Feynman integral $\mathcal{I}$ is defined as

$$\mathrm{Cut}_C^G\mathcal{I} = \int_{\ell_1,\ldots,\ell_{\mathrm{L}}} \frac{\mathcal{N}}{\prod_{e\notin C}(q_e^2-m_e^2)}\prod_{c\in C}\hat{\delta}(q_c^2-m_c^2)\,, \tag{108}$$

where the integration contour is deformed so that it captures the solutions of the on-shell conditions implied by the delta functions.

**Example 28** All four edges of a one-loop box in $\mathrm{D}=4-2\epsilon$ dimensions can be cut simultaneously, by an operation known as a quadruple cut. Although the four delta functions have no common solution for real momenta in Euclidean or Minkowski signature, a pair of solutions can be identified in (2,2) signature or with complex-valued momenta [21]. In the case of the one-mass box,

$$\begin{aligned}\mathrm{Cut}_{\{1,2,3,4\}}^G\mathcal{I}^{1m}(s,t,p_1^2) &= \\ &= \int_\ell \hat{\delta}(\ell^2)\,\hat{\delta}[(\ell+p_1)^2]\,\hat{\delta}[(\ell+p_{12})^2]\,\hat{\delta}[(\ell-p_4)^2]\\ &= \frac{8\pi^2}{ist} + \mathcal{O}(\epsilon)\,.\end{aligned} \tag{109}$$



Although we have written delta functions in Def. 7, the on-shell conditions can be freely interpreted in any of the other ways discussed above, such as by residues. Parametric representations are also possible (see for example [8, 17, 20]) and remain to be more fully developed.

### B. Unitarity methods for amplitudes

Generalized unitarity is commonly applied to identify coefficients of a basis of master integrals for a given scattering amplitude, at a fixed order in perturbation theory. Since the coefficients are rational or algebraic by construction, the branch cuts and discontinuities of the type probed by the cutting operation operate only on the master integrals. Therefore, the cut amplitude can be expanded in a basis of cut master integrals, with the same coefficients;

$$\mathcal{A} = \sum_i c_i \mathcal{I}_i \,, \tag{110}$$

$$\mathrm{Cut}_C^G \mathcal{A} = \sum_i c_i \,\mathrm{Cut}_C^G \mathcal{I}_i \,. \tag{111}$$

In the unitarity method presented in [12, 13], this idea was used with $\mathrm{Cut}_C^G \to \mathrm{Cut}_C^U$, collecting information from various unitarity cuts. One advantage of applying cuts is that $\mathrm{Cut}_C^G \mathcal{I}_i = 0$ whenever $C$ contains an edge that is not present in $\mathcal{I}_i$. Another is that information can be collected from a variety of cuts. A third advantage is that the left-hand side of (111) can be given the interpretation of a pair of amplitudes, integrated over a momentum space restricted to the on-shell conditions of the cut propagators, as in Sec. III A.

In Ex. 28, the same result can be obtained by taking successive unitarity cuts in the $s$ and $t$ channels after suitable analytic continuation, or by taking the residues at the singular points, which can be given the interpretation of a monodromy. The latter formalism was developed in [2]. We stress that the idea of generalized unitarity is to choose any possible interpretation of integrating with the on-shell delta functions. However, one must be precise and consistent within a given calculation, in order to restrict an ansatz for a scattering amplitude based on its functional form.

**Example 29** One-loop reduction uses a well known set of master integrals. The leading-color partial amplitude for 5-particle scattering in $\mathcal{N} = 4$ supersymmetric Yang–Mills theory can be expressed in terms of scalar 1-mass boxes.

$$\mathcal{A} = c_1 \; \text{box}(5,2,3,4) + c_2 \; \text{box}(1,3,4,5) + c_3 \; \text{box}(2,4,5,1) + c_4 \; \text{box}(3,5,1,2) + c_5 \; \text{box}(4,1,2,3) \tag{112}$$

The quadruple cut of Ex. 28 annihilates all master integrals except the last, leaving a linear equation for the single coefficient $c_5$:

$$\mathrm{Cut}^G_{\{1,2,3,4\}} \mathcal{A} = c_5 \,\mathrm{Cut}^G_{\{1,2,3,4\}} \; \text{box}(4,1,2,3) = \frac{8\pi^2 c_5}{ist} \,, \tag{113}$$

where

$$c_5 = \frac{st}{2} \mathcal{A}_1^{\text{tree}} \mathcal{A}_2^{\text{tree}} \mathcal{A}_3^{\text{tree}} \mathcal{A}_4^{\text{tree}} \,. \tag{114}$$

Here $\mathcal{A}_i^{\text{tree}}$ are the 3- and 4-particle amplitudes glued together by the cuts. Likewise, each of the remaining coefficients $c_i$ can be isolated by a single quadruple cut.

In Ex. 29, we used a set of generalized cuts that isolated individual box integrals. In general, for example in a one-loop amplitude in a theory involving triangle and bubble master integrals, their coefficients could only be obtained from cuts of fewer propagators, which also act nontrivially on the boxes. Generalized unitarity methods are most complete in cases where the master integrals are fully polylogarithmic, giving good control of the basis of functions and their associated branch cuts. See [9, 19, 98] for some recent reviews and further discussion.

### C. Solution space of generalized cuts

The *maximal cut* of a Feynman integral is a cut of all propagators, such as in the Ex. 7, 21 and 28. Maximal cuts are closely related to the original integral but typically far simpler.

Beyond one loop, cuts of a given set of propagators are not always unique. We illustrate this fact on the maximal cut of a sunrise integral.

**Example 30** Consider the two-loop sunrise integral with a single massive propagator in $D = 2 - 2\epsilon$ dimensions. The first loop integration gives the maximal cut of a scalar bubble:

$$\mathrm{Cut}^G_{\{1,2,3\}} \mathcal{I}_{\text{sunrise}}(p^2, m^2) = \; \text{sunrise}(p_1, p_2, p_3, p_4) \tag{115}$$

$$= \int_{\ell_1,\ell_2} \hat{\delta}(\ell_1^2)\,\hat{\delta}(\ell_2^2)\,\hat{\delta}[(\ell_1+\ell_2+p)^2 - m^2]$$

$$= \alpha_1 \int_\ell \hat{\delta}(\ell^2)\left[(\ell+p)^2 - m^2\right]^{-1-2\epsilon}\left[(\ell+p)^2\right]^{\epsilon}$$

$$= \alpha_2 \int \mathrm{d}\ell^0 \, \frac{(\ell^0)^{-1-2\epsilon}\left(p^2 + 2\sqrt{p^2}\ell^0\right)^{2\epsilon}}{\left(p^2 - m^2 + 2\sqrt{p^2}\ell^0\right)^{1+2\epsilon}} \,.$$

See [3] for the normalization factors $\alpha_1$ and $\alpha_2$, and further details.

The contour of the last integral is unspecified, and we can identify two independent choices of the maximal cut. The natural choices of contour extend between two

endpoints that are branch points of the integrand. In this case, we might recognize this integral as taking the form of the Gauss hypergeometric function. There are several related ways to understand that we expect two solutions to the maximal cut of this integral. One is by observing the two independent integration contours to complete the maximal cut calculation as above.

Another is by considering the basis of master integrals related to the original one by IBP and dimension-shift identities, and observing that there is a linearly independent integral in this basis with the same set of propagators. For example, it could be taken to contain an additional factor of $(\ell_1 + \ell_2 + p)^2$ in the numerator of the integrand.

A third is to consider the system of first-order linear differential equations satisfied by the set of master integrals. Generalized cuts of the master integrals satisfy the same system of differential equations [16, 57, 83], with possibly different boundary conditions, and possibly spanning a smaller subspace than the master integrals. In this example of the one-mass sunrise, there are no master integrals with fewer propagators, and so the maximal cuts span the full solution space. It is possible to express both the uncut integral and the unitarity cut in $p^2$ in terms of any two independent maximal cuts.

Especially in cases such as Ex. 30, with multiple solutions for a given set of cut propagators, it is helpful to formulate a generalized cut as a choice of a new integration contour along with the original, uncut, integrand. The number of cut solutions is the dimension of the homology group associated to a space with singularities identified with the chosen propagators. It also becomes apparent that generalized cut integrals satisfy the same linear (IBP) relations and differential equations as the original integral.[3]

This observation also gives a new perspective on why computational techniques for loop computations can be applied to cross section computations, cf. the discussion of reverse-unitarity in Sec. III B 4. Thus, associated to a given Feynman integral, it is instructive to generate the so-called period matrix, whose entries are pairings of all homology and cohomology classes[4] obtained from the singular loci of the Feynman integral. The entries of the period matrix are then understood as the related master integrals, and all of their independent generalized cuts. See [1] for a review of these topics with additional references.

## VI. SUMMARY

This article summarized different notions of cuts used throughout the literature on scattering amplitudes and related topics. We grouped them into three categories, depending on which energies are allowed to flow through a cut and whether the right-hand of the cut is conjugated or not. Such choices can make an enormous difference for the result of the computations, as highlighted in the explicit examples above. We moreover emphasized that cuts, and relations derived from them, strongly depend on the kinematic domain in which they are evaluated. Applications, ranging from dispersion relations, through symbol bootstrap, to generalized unitarity, serve as evidence that cuts are comprehensive and versatile computational tools.

## ACKNOWLEDGMENTS

The work of C.D. is funded by the European Union (ERC Consolidator Grant LoCoMotive 101043686). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. H.S.H. gratefully acknowledges funding provided by the William D. Loughlin Membership, an endowed fund of the Institute for Advanced Study. S.M. gratefully acknowledges funding provided by the Sivian Fund and the Roger Dashen Member Fund at the Institute for Advanced Study. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC0009988.

[3] Cuts of propagators raised to nonpositive powers are zero. Therefore maximal cuts satisfy a highly simplified set of differential equations, explaining their relative simplicity as functions.

[4] In dimensional regularization, we require twisted homology and cohomology groups, due to the multivaluedness of the integrand in noninteger dimension [75, 76].

[1] S. Abreu, R. Britto and C. Duhr, *The SAGEX review on scattering amplitudes Chapter 3: Mathematical structures in Feynman integrals*, *J. Phys. A* **55** (2022) 443004 [2203.13014].

[2] S. Abreu, R. Britto, C. Duhr and E. Gardi, *Cuts from residues: the one-loop case*, *JHEP* **06** (2017) 114 [1702.03163].

[3] S. Abreu, R. Britto, C. Duhr, E. Gardi and J. Matthew, *The diagrammatic coaction beyond one loop*, *JHEP* **10** (2021) 131 [2106.01280].

[4] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis and R. Rattazzi, *Causality, analyticity and an IR obstruction to UV completion*, *JHEP* **10** (2006) 014 [hep-th/0602178].

[5] C. Anastasiou, L.J. Dixon, K. Melnikov and F. Petriello, *Dilepton rapidity distribution in the Drell-Yan process at NNLO in QCD*, *Phys. Rev. Lett.* **91** (2003) 182002

[hep-ph/0306192].

[6] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, *Phys. Rev. Lett.* **114** (2015) 212001 [1503.06056].

[7] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, *Nucl. Phys. B* **646** (2002) 220 [hep-ph/0207004].

[8] N. Arkani-Hamed and E.Y. Yuan, *One-Loop Integrals from Spherical Projections of Planes and Quadrics*, 1712.09991.

[9] S. Badger, J. Henn, J.C. Plefka and S. Zoia, *Scattering Amplitudes in Quantum Field Theory*, *Lect. Notes Phys.* **1021** (2024) pp. [2306.05976].

[10] D. Baumann, D. Green, A. Joyce, E. Pajer, G.L. Pimentel, C. Sleight et al., *Snowmass White Paper: The Cosmological Bootstrap*, in *Snowmass 2021*, 3, 2022 [2203.08121].

[11] M. Berghoff and E. Panzer, *Hierarchies in relative Picard-Lefschetz theory*, 2212.06661.

[12] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *One-Loop n-Point Gauge Theory Amplitudes, Unitarity and Collinear Limits*, *Nucl. Phys.* **B425** (1994) 217 [hep-ph/9403226].

[13] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *Fusing gauge theory tree amplitudes into loop amplitudes*, *Nucl. Phys.* **B435** (1995) 59 [hep-ph/9409265].

[14] J.D. Bjorken, *Experimental tests of Quantum electrodynamics and spectral representations of Green's functions in perturbation theory*, Ph.D. thesis, Stanford U., 1959.

[15] S. Bloch and D. Kreimer, *Cutkosky Rules and Outer Space*, 1512.01705.

[16] J. Bosma, K.J. Larsen and Y. Zhang, *Differential equations for loop integrals in Baikov representation*, *Phys. Rev. D* **97** (2018) 105014 [1712.03760].

[17] J.B. Boyling, *A homological approach to parametric feynman integrals*, *Il Nuovo Cimento A (1965-1970)* **53** (1968) 351.

[18] J.B. Boyling, *Construction of Vanishing Cycles for Integrals over Hyperspheres*, *Journal of Mathematical Physics* **7** (2004) 1749.

[19] A. Brandhuber, J. Plefka and G. Travaglini, *The SAGEX Review on Scattering Amplitudes Chapter 1: Modern Fundamentals of Amplitudes*, *J. Phys. A* **55** (2022) 443002 [2203.13012].

[20] R. Britto, *Generalized Cuts of Feynman Integrals in Parameter Space*, *Phys. Rev. Lett.* **131** (2023) 091601 [2305.15369].

[21] R. Britto, F. Cachazo and B. Feng, *Generalized unitarity and one-loop amplitudes in N=4 super-Yang-Mills*, *Nucl.Phys.* **B725** (2005) 275 [hep-th/0412103].

[22] J. Bros, *Dispersion relations*, in *Encyclopedia of Mathematical Physics*, J.-P. Françoise, G.L. Naber and T.S. Tsun, eds., (Oxford), pp. 87–101, Academic Press (2006), DOI.

[23] J. Bros, *Scattering in relativistic quantum field theory: The analytic program*, in *Encyclopedia of Mathematical Physics*, J.-P. Françoise, G.L. Naber and T.S. Tsun, eds., (Oxford), pp. 465–475, Academic Press (2006), DOI.

[24] J. Bros, H. Epstein and V. Glaser, *A proof of the crossing property for two-particle amplitudes in general quantum field theory*, *Commun. Math. Phys.* **1** (1965) 240.

[25] F. Brown, *Notes on Motivic Periods*, *Commun. Num. Theor Phys.* **11** (2017) 557 [1512.06410].

[26] F. Brown and O. Schnetz, *Single-valued multiple polylogarithms and a proof of the zig–zag conjecture*, *J. Number Theor.* **148** (2015) 478.

[27] F.C.S. Brown, *On the periods of some Feynman integrals*, 0910.0114.

[28] F.C. Brown, *Multiple zeta values and periods of moduli spaces $\mathfrak{M}_{0,n}$*, *Annales Sci.Ecole Norm.Sup.* **42** (2009) 371 [math/0606419].

[29] D. Buchholz and S. Summers, *Scattering in relativistic quantum field theory: Fundamental concepts and tools*, in *Encyclopedia of Mathematical Physics*, J.-P. Françoise, G.L. Naber and T.S. Tsun, eds., (Oxford), pp. 456–465, Academic Press (2006), DOI.

[30] Z. Capatti, V. Hirschi, D. Kermanschah and B. Ruijl, *Loop-Tree Duality for Multiloop Numerical Integration*, *Phys. Rev. Lett.* **123** (2019) 151602 [1906.06138].

[31] S. Caron-Huot, L.J. Dixon, F. Dulat, M. Von Hippel, A.J. McLeod and G. Papathanasiou, *The Cosmic Galois Group and Extended Steinmann Relations for Planar $\mathcal{N} = 4$ SYM Amplitudes*, *JHEP* **09** (2019) 061 [1906.07116].

[32] S. Caron-Huot, M. Giroux, H.S. Hannesdottir and S. Mizera, *What can be measured asymptotically?*, *JHEP* **01** (2024) 139 [2308.02125].

[33] F. Chavez and C. Duhr, *Three-mass triangle integrals and single-valued polylogarithms*, *JHEP* **11** (2012) 114 [1209.2722].

[34] K.-T. Chen, *Iterated path integrals*, *Bull.Am.Math.Soc.* **83** (1977) 831.

[35] X. Chen, T. Gehrmann, E.W.N. Glover, A. Huss, B. Mistlberger and A. Pelloni, *Fully Differential Higgs Boson Production to Third Order in QCD*, *Phys. Rev. Lett.* **127** (2021) 072002 [2102.07607].

[36] K. Chetyrkin and F. Tkachov, *Integration by Parts: The Algorithm to Calculate beta Functions in 4 Loops*, *Nucl. Phys. B* **192** (1981) 159.

[37] K.-c. Chou, Z.-b. Su, B.-l. Hao and L. Yu, *Equilibrium and Nonequilibrium Formalisms Made Unified*, *Phys. Rept.* **118** (1985) 1.

[38] S. Coleman and R. Norton, *Singularities in the physical region*, *Nuovo Cim.* **38** (1965) 438.

[39] M. Correia, A. Sever and A. Zhiboedov, *An analytical toolkit for the S-matrix bootstrap*, *JHEP* **03** (2021) 013 [2006.08221].

[40] R.E. Cutkosky, *Singularities and discontinuities of Feynman amplitudes*, *J. Math. Phys.* **1** (1960) 429.

[41] J. de Jesús Aguilera-Verdugo et al., *A Stroll through the Loop-Tree Duality*, *Symmetry* **13** (2021) 1029 [2104.14621].

[42] C. de Rham, S. Kundu, M. Reece, A.J. Tolley and S.-Y. Zhou, *Snowmass White Paper: UV Constraints on IR Physics*, in *Snowmass 2021*, 3, 2022 [2203.06805].

[43] A. Denner and S. Dittmaier, *The Complex-mass scheme for perturbative calculations with unstable particles*, *Nucl. Phys. B Proc. Suppl.* **160** (2006) 22 [hep-ph/0605312].

[44] A. Denner, S. Dittmaier, M. Roth and D. Wackeroth, *Predictions for all processes e+ e- —> 4 fermions + gamma*, *Nucl. Phys. B* **560** (1999) 33 [hep-ph/9904472].

[45] A. Denner, S. Dittmaier, M. Roth and L.H. Wieders, *Electroweak corrections to charged-current e+ e- —> 4 fermion processes: Technical details and further results*, *Nucl. Phys. B* **724** (2005) 247 [hep-ph/0505042].

[46] C. Duhr, *Hopf algebras, coproducts and symbols: an application to Higgs boson amplitudes*, *JHEP* **1208** (2012) 043 [1203.0454].

[47] C. Duhr and F. Dulat, *PolyLogTools — polylogs for the masses*, *JHEP* **08** (2019) 135 [1904.07279].

[48] C. Duhr, F. Dulat and B. Mistlberger, *Charged current Drell-Yan production at $N^3LO$*, *JHEP* **11** (2020) 143 [2007.13313].

[49] C. Duhr, F. Dulat and B. Mistlberger, *Drell-Yan Cross Section to Third Order in the Strong Coupling Constant*, *Phys. Rev. Lett.* **125** (2020) 172001 [2001.07717].

[50] C. Duhr, F. Dulat and B. Mistlberger, *Higgs Boson Production in Bottom-Quark Fusion to Third Order in the Strong Coupling*, *Phys. Rev. Lett.* **125** (2020) 051804 [1904.09990].

[51] C. Duhr, H. Gangl and J.R. Rhodes, *From polygons and symbols to polylogarithmic functions*, *JHEP* **1210** (2012) 075 [1110.0458].

[52] C. Duhr and B. Mistlberger, *Lepton-pair production at hadron colliders at $N^3LO$ in QCD*, *JHEP* **03** (2022) 116 [2111.10379].

[53] L. Eberhardt and S. Mizera, *Unitarity cuts of the worldsheet*, *SciPost Phys.* **14** (2023) 015 [2208.12233].

[54] D.B. Fairlie, P.V. Landshoff, J. Nuttall and J.C. Polkinghorne, *Singularities of the Second Type*, *Journal of Mathematical Physics* **3** (1962) 594.

[55] C. Fevola, S. Mizera and S. Telen, *Landau Singularities Revisited*, 2311.14669.

[56] R.P. Feynman, *Quantum theory of gravitation*, *Acta Phys. Polon.* **24** (1963) 697.

[57] H. Frellesvig and C.G. Papadopoulos, *Cuts of Feynman Integrals in Baikov representation*, *JHEP* **04** (2017) 083 [1701.07356].

[58] D. Gaiotto, J. Maldacena, A. Sever and P. Vieira, *Pulling the straps of polygons*, *JHEP* **12** (2011) 011 [1102.0062].

[59] F. Gelis and N. Tanji, *Schwinger mechanism revisited*, *Prog. Part. Nucl. Phys.* **87** (2016) 1 [1510.05451].

[60] P. Goddard, *Nonphysical Region Singularities of the S Matrix*, *Journal of Mathematical Physics* **11** (2003) 960.

[61] A. Goncharov, *A simple construction of Grassmannian polylogarithms*, 0908.2238.

[62] A.B. Goncharov, M. Spradlin, C. Vergu and A. Volovich, *Classical Polylogarithms for Amplitudes and Wilson Loops*, *Phys. Rev. Lett.* **105** (2010) 151605 [1006.5703].

[63] H.S. Hannesdottir, A.J. McLeod, M.D. Schwartz and C. Vergu, *Constraints on sequential discontinuities from the geometry of on-shell spaces*, *JHEP* **07** (2023) 236 [2211.07633].

[64] H.S. Hannesdottir and S. Mizera, *What is the $i\varepsilon$ for the S-matrix?*, SpringerBriefs in Physics, Springer (1, 2023), 10.1007/978-3-031-18258-7, [2204.02988].

[65] D. Iagolnitzer and J. Magnen, *Scattering, asymptotic completeness and bound states*, in *Encyclopedia of Mathematical Physics*, J.-P. Françoise, G.L. Naber and T.S. Tsun, eds., (Oxford), pp. 475–487, Academic Press (2006), DOI.

[66] L.V. Keldysh, *Diagram technique for nonequilibrium processes*, *Zh. Eksp. Teor. Fiz.* **47** (1964) 1515.

[67] D.A. Kosower, B. Maybee and D. O'Connell, *Amplitudes, Observables, and Classical Scattering*, *JHEP* **02** (2019) 137 [1811.10950].

[68] A.V. Kotikov, *Differential equation method: The Calculation of N point Feynman diagrams*, *Phys. Lett. B* **267** (1991) 123.

[69] A.V. Kotikov, *Differential equations method: New technique for massive Feynman diagrams calculation*, *Phys. Lett. B* **254** (1991) 158.

[70] A.V. Kotikov, *Differential equations method: The Calculation of vertex type Feynman diagrams*, *Phys. Lett. B* **259** (1991) 314.

[71] M. Kruczenski, J. Penedones and B.C. van Rees, *Snowmass White Paper: S-matrix Bootstrap*, 2203.02421.

[72] L.D. Landau, *On analytic properties of vertex parts in quantum field theory*, *Nucl. Phys.* **13** (1959) 181.

[73] S. Mandelstam, *Determination of the pion - nucleon scattering amplitude from dispersion relations and unitarity. General theory*, *Phys. Rev.* **112** (1958) 1344.

[74] A. Martin, *Scattering Theory: Unitarity, Analyticity and Crossing*, vol. 3 (1969), 10.1007/BFb0101043.

[75] P. Mastrolia and S. Mizera, *Feynman Integrals and Intersection Theory*, *JHEP* **02** (2019) 139 [1810.03818].

[76] S.-J. Matsubara-Heo, S. Mizera and S. Telen, *Four lectures on Euler integrals*, *SciPost Phys. Lect. Notes* **75** (2023) 1 [2306.13578].

[77] D. Meltzer and A. Sivaramakrishnan, *CFT unitarity and the AdS Cutkosky rules*, *JHEP* **11** (2020) 073 [2008.11730].

[78] B. Mistlberger, *Higgs boson production at hadron colliders at $N^3LO$ in QCD*, *JHEP* **05** (2018) 028 [1802.00833].

[79] N. Nakanishi, *Ordinary and Anomalous Thresholds in Perturbation Theory*, *Prog. Theor. Phys.* **22** (1959) 128.

[80] F. Pham, *Singularités des processus de diffusion multiple*, *Annales de l'institut Henri Poincaré. Section A, Physique Théorique* **6** (1967) 89.

[81] F. Pham, *Singularities of Integrals: Homology, hyperfunctions and microlocal Analysis*, Springer Science & Business Media (2011).

[82] R. Pius and A. Sen, *Cutkosky rules for superstring field theory*, *JHEP* **10** (2016) 024 [1604.01783].

[83] A. Primo and L. Tancredi, *On the maximal cut of Feynman integrals and the solution of their differential equations*, *Nucl. Phys. B* **916** (2017) 94 [1610.08397].

[84] E. Remiddi, *Dispersion Relations for Feynman Graphs*, *Helv. Phys. Acta* **54** (1982) 364.

[85] E. Remiddi, *Differential equations for Feynman graph amplitudes*, *Nuovo Cim. A* **110** (1997) 1435 [hep-th/9711188].

[86] M. Rubin, R. Sugar and G. Tiktopoulos, *Dispersion Relations for Three-Particle Scattering Amplitudes. I*, *Phys. Rev.* **146** (1966) 1130.

[87] R. Runkel, Z. Szor, J.P. Vesga and S. Weinzierl, *Causality and loop-tree duality at higher loops*, *Phys. Rev. Lett.* **122** (2019) 111603 [1902.02135].

[88] J.S. Schwinger, *Brownian motion of a quantum oscillator*, *J. Math. Phys.* **2** (1961) 407.

[89] H.P. Stapp and A.R. White, *An Asymptotic Dispersion Relation for the Six Particle Amplitude*, *Phys. Rev. D* **26** (1982) 2145.

[90] O. Steinmann, *Über den Zusammenhang zwischen den Wightmanfunktionen und den retardierten Kommutatoren*, *Helv. Phys. Acta* **33** (1960) 257.

[91] O. Steinmann, *Wightman-Funktionen und retardierte Kommutatoren. II*, *Helv. Phys. Acta* **33** (1960) 347.

[92] G.F. Sterman, *An Introduction to quantum field theory*, Cambridge University Press (8, 1993).

[93] R. Szabo, *Perturbation theory and its techniques*, in *Encyclopedia of Mathematical Physics*, J.-P. Françoise, G.L. Naber and T.S. Tsun, eds., (Oxford), pp. 28–41, Academic Press (2006), DOI.

[94] G. 't Hooft and M.J.G. Veltman, *Diagrammar*, *NATO Sci. Ser. B* **4** (1974) 177.

[95] F. Tkachov, *A Theorem on Analytical Calculability of Four Loop Renormalization Group Functions*, *Phys. Lett. B* **100** (1981) 65.

[96] M.J.G. Veltman, *Unitarity and causality in a renormalizable field theory with unstable particles*, *Physica* **29** (1963) 186.

[97] M.J.G. Veltman, *Diagrammatica: The Path to Feynman rules*, vol. 4, Cambridge University Press (5, 2012).

[98] S. Weinzierl, *Feynman Integrals* (1, 2022), 10.1007/978-3-030-99558-4, [2201.03593].

# Testing gravitational waveforms in full General Relativity

Fabio D'Ambrosio,[1,*] Francesco Gozzini,[2,†] Lavinia Heisenberg,[2,‡]
Henri Inchauspé,[2,§] David Maibach,[2,¶] and Jann Zosso[1,**]

[1] *Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland*
[2] *Institute for Theoretical Physics, University of Heidelberg, Philosophenweg 16 D-69120 Heidelberg Germany*

(Dated: March 1, 2024)

We perform a comprehensive analysis of state-of-the-art waveform models, focusing on their predictions concerning kick velocity and inferred gravitational wave memory. In our investigation we assess the accuracy of waveform models using energy-momentum balance laws, which were derived in the framework of full, non-linear General Relativity. The numerical accuracy assessment is performed for precessing as well as non-precessing scenarios for models belonging to the *EOB*, *Phenom*, and *Surrogate* families. We analyze the deviations of these models from each other and from Numerical Relativity waveforms. Our analysis reveals statistically significant deviations, which we trace back to inaccuracies in modelling subdominant modes and inherent systematic errors in the chosen models. We corroborate our findings through analytical considerations regarding the mixing of harmonic modes in the computed kick velocities and inferred memories.

arXiv:2402.19397v1 [gr-qc] 29 Feb 2024

## I. INTRODUCTION

The first direct detection of gravitational waves (GWs) from the merger of two black holes in 2015 [1] marked a milestone in the field of astrophysics, confirming a key prediction of Einstein's General Theory of Relativity (GR) and ushering in a new era of observational astronomy. The waveforms matched against the GW signal encode a wealth of information, including the masses and spins of the binaries, the distances to the sources, and the geometry of their motion. By analyzing these waveforms, astrophysicists can uncover the underlying physical processes, discern the properties of exotic objects, and validate theoretical models with unprecedented precision.

To date, the process of parameter estimation and the testing of GR necessitate numerical modeling of gravitational waveforms across a wide range of source parameters, including masses, spins, and other relevant factors pertaining to the merging objects. These modelled waveforms are used as fitting templates with respect to actual data. The effectiveness of this template-to-signal match hinges on the precision of the estimated waveforms. Therefore, in order to extract meaningful information from signals, it is crucial to construct comprehensive and accurate waveform templates that faithfully capture the physics of GR. This requirement is further substantiated by the expected increase in signal-to-noise-ratio of future GW observatories such as LISA [2], the Einstein Telescope [3], and the Cosmic Explorer [4]. Their increased resolution and sensitivity opens them up to more subtle effects like the gravitational memory [5–10] and they may even reveal physics beyond GR [11].

However, deviations between templates and the actual waveforms contained in the observational data introduce systematic biases, compromising the reliability of the information extracted from the signal. To counteract such biases, a diverse set of template waveforms is employed when analyzing data from GW instruments. Among these template waveforms, the ones obtained from Numerical Relativity (NR) simulations are the most reliable. Their disadvantage, however, is their consumption of vast amounts of computational resources for each simulation, i.e., for each choice of parameters describing an individual merging scenario. This time-consuming and resource-intense process poses a significant challenge, particularly as the volume of data to be processed is expected to increase drastically in the upcoming years with the advent of multiple ground- and space-based instruments like LISA [12] and LIGO/Virgo [13, 14]. Furthermore, as the measuring precision advances, deviations from GR [11] may reveal themselves in the observed data. Detecting such deviations necessitates an expanded parameter space to account for alternative descriptions of gravity, which consequently amplifies the number of waveforms against which the data must be tested.

To address these challenges, and in particular to tackle the efficiency issue of the waveform generation process, multiple alternatives to NR simulations were established. Prominent representatives of alternative waveform models are the *Surrogate* models [15–17], phenomenological models [18–21] and effective-one-body simulations [22–27]. To obtain reliable results within reasonable timescales, the models adopt distinct strategies to compute the gravitational strain. Each model focuses on different physical aspects of compact binary coalescence and is capable of producing a waveform within certain parameter ranges in an efficient manner. To understand what distinguishes the different approximant families, it is helpful to divide the compact binary coalescence into three phases: the inspiral, the merger, and the ringdown phase. The objective of each approximant family is to generate waveforms that replicate NR without re-

* fabio.dambrosio@gmx.ch
† gozzini@thphys.uni-heidelberg.de
‡ l.heisenberg@thphys.uni-heidelberg.de
§ h.inchauspe@thphys.uni-heidelberg.de
¶ d.maibach@thphys.uni-heidelberg.de
** zosso.jann@bluewin.ch

sorting to the computationally expensive procedure of numerically solving GR in the strong-field regime. While most models follow a Post-Newtonian (PN) ansatz for the inspiral (see also section II), merger and ringdown are simulated using various approaches, such as effective-one-body or phenomenological models, as mentioned above. After modeling each of the three phases separately, it is necessary to patch these segments together to obtain a coherent waveform for a specific event. Given that two of the three phases are modelled in fundamentally different ways, it is evident that the process of patching waveform segments together introduces errors and further discrepancies between different approximants.

In this work, we provide an extensive and detailed comparison of state-of-the-art waveform models based on their prediction of physical quantities such as the remnant's kick velocity and the inferred gravitational wave memory. We compare the models to events extracted from NR simulations, as well as between themselves for events beyond the scope of the available NR catalogue. Our analysis prioritizes non-precessing events for the evaluation of kick velocities, and we extend the analysis of the memory to include precessing as well as non-precessing events. Our primary tool for effecting these comparisons are the so-called *energy-momentum balance laws* [28]. These exact mathematical results were derived within full, non-linear GR and provide an infinite tower of constraints for waveform models. After performing a decomposition of the balance laws into spherical harmonics, we obtain a simple constraint equation as well as equations for directly computing kick velocity and inferred memory for any given waveform model. A proof of concept that the balance laws can be used as diagnostic tools for assessing waveform accuracy was provided in [29, 30]. Here, our objective is to fully leverage the strength of these tools for a comparative analysis between different waveform models. Our study builds upon prior research, which concentrated exclusively on analytical models [30], the kick velocity [31, 32], or gravitational memory [29, 33]. We extend these earlier investigations and reveal systematic trends in waveform templates.

Our paper is structured as follows. In section II we review the families of the waveform generators which are the subject of our analysis in later sections. We provide a detailed account of their conventions, and we describe their generation procedures and implementations. We further compare the use of reference frames and strategies employed for handling waveform alignment and precession within each waveform family. Section III delves into a thorough discussion of our own alignment procedures, where we explore various strategies and discuss the outcomes for each cases. In section IV we list the binary black hole (BBH) merger events that enter our comparative analysis. We categorize them into two groups: those generated from NR (referred to as "catalogued") and those falling outside the NR parameter space (referred to as "non-catalogued"). Section V revisits the framework and formulation of the energy-momentum balance laws, with a focus on a qualitative understanding and an emphasis on their power for assessing the accuracy of waveform models. We also perform a decomposition of the balance laws into spherical harmonics in order to distill equations for kick velocity and GW memory.

The core of our investigations is found in section VI, where we analyze kick velocity, memory, and the relative differences between the selected models, based on the equations derived in section V. Special attention is given to the impact of mode content and selection. Finally, in section VII, we discuss the outcomes of the model comparison, considering aspects such as model performance, alignment strategies, and mode selection.

## II. OVERVIEW OF WAVEFORM MODELS

We start by presenting the most recent waveform model families that have been prominent in GW research and that we will employ in this study. We establish a unified parameter space for all models, provide an overview of their modeling procedures, and discuss their strengths and weaknesses qualitatively. For a more comprehensive overview encompassing the models considered in this work, we direct readers to [34].

In this investigation, we focus on an ensemble of four different waveform families available through the *LALSuite* software collaboration [35]. The list of models contains four main families, which we label *NR*, *Surrogate*, *Phenom*, *EOB* according to their generation mechanisms. Among these four families, we focus more extensively on individual representatives. Specifically, we analyze the following implementations: *SXS* and its *LALSuite* interface [36], the *Surrogate* model *NRSur7dq4* [17], *EOB* being represented by *SEOBNRv4PHM* [26] and *IMRPhenomTPHM* [21] as the phenomenological model. For simplicity, in the remainder of this work we shall often identify the specific model (e.g. *NRSur7dq4*) with its family (e.g. *Surrogate*). The numerical relativity data from the *SXS* collaboration is tested through the available LALsuite interpolation [36] which we refer to as *NR*.

The chosen families represent distinct approaches to waveform simulations. The *NR* family represents full numerical relativity simulations, offering the most comprehensive and reliable waveforms against which the other models are competing. The alternative "approximant" models are based on the phenomenology of binary mergers (*Phenom*), effective-one-body simulations (*EOB*), and interpolation of NR simulations (*Surrogate*). In Table I, we present an overview of the selected models, specifying for each the included (spin-weighted) harmonic modes and the applicable domains concerning mass-ratio $q$ and spin magnitude $|\chi_i|$.

| **Family** | **Implementation** | **Branch** | **Mode Content** | $q$**-range** | $\lvert\chi_i\rvert$**-range** |
|---|---|---|---|---|---|
| Numerical Relativity | *SXS* (NRhdf5) | precessing | $\{(\ell,m)\lvert\ell\leq 8\}$ | $\lesssim 4$ | $\lesssim 0.8$ |
| Surrogate | NRSur7dq4 | precessing | $\{(\ell,m)\lvert\ell\leq 4\}$ | $\leq 4$ | $\leq 0.8$ |
| Effective One Body | SEOBNRv4PHM | precessing | $\{(\ell,m)\lvert\ell\leq 5\}$ | - | - |
| Phenomenological | IMRPhenomTPHM | precessing | $\{(\ell,m)\lvert\ell\leq 5\}$ | - | - |

TABLE I. Summary of the families, specific implementations, mode content and parameter space of the waveform models considered in this work.

### A. Waveform families

Generally speaking, the precessing BBH problem for quasicircular orbits is parametrized by seven parameters: the mass ratio $q = m_1/m_2 \geq 1$ and two spin vectors $\chi_{1,2}$, each containing three independent components. Here, the index 1 (2) refers to the heavier (lighter) BH. In the following, we abbreviate all parameters relevant to a BBH merger using the vector $\boldsymbol{\lambda}$. For the non-precessing case, the parameter space reduces to a total of only three parameters as the spins $\chi_{1,2}$ are aligned (or anti-aligned) and only their magnitude is relevant for the calculation of the strain.

Using this common parameter vector, in what follows, we briefly outline the general working methods of the waveform families and highlight certain steps that are essential to identify strengths and shortcomings of each model on a phenomenological level.

*Numerical Relativity:* The bedrock of waveform modelling is formed by numerical simulations of the nonlinear Einstein field equations, which until today are the only means of obtaining accurate waveforms for the strong gravity regime of a BBH coalescence. The publicly available simulations are distributed across various catalogues, with one of the most comprehensive being the catalogue produced by the SXS collaboration [37]. The NR code employed in these simulation is the Spectral Einstein Code (SpEC) [38]. The distinguishing feature of SpEC is to utilize so-called *spectral methods* for the numerical approximation of the derivative operators in the relevant differential equations at hand. In essence, while the conventional approach for approximating the derivative of a function on a given grid involves variations of the finite-difference method, spectral methods instead consider the expansion of the function to be differentiated into a finite basis of orthogonal functions on the grid.

Every numerical simulation begins by generating initial data that adheres to the constraint equations of the system, with the initial parameters $\boldsymbol{\lambda}$ serving as input. The evolution typically commences with an initial burst of spurious or "junk" radiation before the systems settle into a stable inspiral phase. This initial junk radiation constitutes a primary source of numerical errors in the *SXS* waveforms, as varying resolutions in the simulations may lead to distinct initial bursts and consequently result in slightly different estimated parameters during subsequent evolution. The simulation progresses until a shared apparent horizon forms, indicating the onset of the coalescence phase. At this juncture, the simulation is momentarily halted, and the data are interpolated onto a new grid tailored to a single extraction boundary. Subsequently, the evolution recommences through the coalescence and ringdown phases until all ringdown gravitational waves are emitted. For the resulting waveforms, the numerical error estimation involves a comparison of waveforms generated at different resolutions, specifically the highest and the next-to-highest.

The NR waveforms obtained via the described procedure and collected in the main catalogue considered in this study do not exhibit GW memory effects. According to [39], this is attributed to the extrapolation procedure of finite-grid data to null infinity, which utilizes polynomial interpolation. The more accurate extrapolation method based on Cauchy-characteristic extraction (CCE) [39, 40] has, so far, only been implemented in a few events within the public catalogue but contains gravitational memory in the catalogued strain modes.

#### 1. *Effective One Body*

The Effective One Body (*EOB*) [41, 42] models attempt to approximate the complete waveform of binary mergers by replacing the two-body dynamics by an effective one-body description, i.e., by solving the problem of a single particle of reduced mass immersed in an "effective" background metric, $g^{\rm eff}_{\mu\nu}(x^\lambda)$ [22]. Solving the equations resulting from the effective background metric, one obtains the full waveform for the complete BBH merger process.

There have been several versions of *EOB* models over the past two decades. Arguably, the most significant advancements manifested with the introduction of the NR-enhanced models known as the Spinning Effective One Body Numerical Relativity models (*SEOBNR*). These models leverage NR simulations to calibrate the free parameters intrinsic to the *EOB* waveform. For instance, *SEOBNR* improves and calibrates its waveforms by requiring the amplitudes, time derivatives of the amplitudes, GW frequencies, and time derivatives of the GW frequencies to coincide with that of the NR simulations at the matching times. Given that NR simulations only cover discrete values of the parameter space, interpolation is employed to bridge the gaps.

Naturally, the accuracy of interpolation is limited and hence the obtained *SEOBNR* waveforms span only across a constrained region of parameter space. To extend

*SEOBNR*'s coverage across a broader region, BH perturbation theory is utilized to generate waveforms for mass ratios significantly smaller than one. Combining all techniques and latest advancements, *SEOBNR* constitutes a waveform model that, in principle, covers any mass ratio and nearly any spin configuration.

In general, several other adjustable parameters may be incorporated into *EOB*'s generation pipeline, depending on whether the model accounts for the spins of the BBH constituents and whether it describes the precession of the BBH system. For instance, *SEOBNRv1-v4* models describe spinning non-precessing BBHs, *SEOBNRv4HM* includes higher modes in the waveform, and *SEOBNRv4PHM* incorporates both higher modes and precession effects [22–26]. Notably, the latest iteration of this family, *SEOBNRv5PHM* [27], is not included in this study.

### 2. *Phenom*

Phenomenological models (*Phenom*) attempt to approximate waveforms derived through alternative methods and interpolate within the parameter space of BBH mergers. A key objective is to rapidly generate these waveforms, enabling their utilization in matched filtering searches and for parameter estimation. Here we focus on the *IMRPhenom*-family of phenomenological models. In particular, we consider *IMRPhenomD* [18, 19] and the most recent versions *IMRPhenomXPHM* [20] and *IMRPhemonTPHM* [21] models. The latter two mainly differ by the domain in which the waveform is generated, where $X$ refers to the frequency domain and $T$ to the time domain.

The starting point of the *IMRPhenomD* model is the decomposition of the $\ell = |m| = 2$ component of a NR waveform in the frequency domain as $\tilde{h}_{22}(f) = A(f)\, e^{-i\phi(f)}$. The functions $A(f)$ and $\phi(f)$ can be approximated by specifically designed parameterized functions which yield a good fit across a diverse range of binary parameters. Within the *IMRPhenomD* model, these functions are piecewise smooth and effectively capture the inspiral, intermediate, and merger-ringdown regions. For the inspiral region, an augmented *EOB* waveform is employed. The augmentation involves incorporating a sum of five parametrized powers of frequency for the phase, and three parametrized powers of frequency for the amplitude. These supplementary terms facilitate a more precise fitting of waveforms. The fitting functions for the intermediate and merger-ringdown regions are entirely phenomenological. They consist of parameterized polynomials as well as more general functions to capture specific features of the waveform.

In the calibration step, each phenomenological waveform parameter $\mathbf{\Lambda}$ (17 in total) is allowed to depend on the binary parameters $\boldsymbol{\lambda}$ (11 in total) that went into the generation of the NR waveforms. This greatly improves the flexibility and accuracy of the model, which can now depend on as many as $17 \times 11 = 187$ free parameters for the fit (listed in Appendix C of [19]). Other optimizations that have been implemented in subsequent versions include additional parameters in the fundamental $h_{2,2}$ mode, the introduction of parameters for higher modes, and the inclusion of NR waveforms of higher mass ratios.

Two additional aspects worth noting are time-domain phenomenological models and precession. Conceptually, the only distinction for time-domain models is that instead of starting with the frequency domain waveform $\tilde{h}_{2,2}(f)$, the process initiates with the time-domain waveform $h_{2,2}(t)$. One then proceeds as for the *IMRPhenomD* model by decomposing it into an amplitude and phase function. Time-domain models are part of the *IMRPhenomT** series of models, such as *IMRPhenomTHM* [43]. Compatibility with precessing events has been included in more recent models by "twisting up" non-precessing waveforms [44]. This procedure distributes the power across different $m$-modes for a fixed $\ell$ of the strain's mode decomposition based on the precession of the individual spin vectors of the BH's involved in the merger. In this work, we solely consider the time-domain precessing model *IMRPhemonTPHM* [21].

### 3. *Surrogate*

A surrogate waveform model utilizes a collection of precomputed waveforms across a defined set of parameters and subsequently performs an interpolation in the parameter space between these waveforms. The objective is to rapidly generate a waveform for any given parameter values. A surrogate waveform can be computed much more rapidly than the underlying (typically NR) calculation, and its accuracy can match that of the underlying model provided there is a sufficiently extensive set of precomputed waveforms covering a broad range of parameter space.

One can broadly distinguish between "pure" surrogate models and "hybrid" surrogate models. Pure surrogate models [15–17] take a set of NR waveforms (e.g. 866 different simulated events for [17]) close to the desired point in parameter space as input and employ parametric fits to extrapolate the waveform across a continuous set of parameters. The extrapolation involves several assumptions, with the alignment of the NR waveforms considered for a given simulation being a critical factor. In general, each NR waveform must be associated with a single value of the parameter vector $\boldsymbol{\lambda}$, even though some parameters (especially spin directions) are time-dependent. This procedure involves various subtleties, one of which is the alignment procedure further elaborated upon in section III.

In order to address some of the major shortcomings of numerical simulations — namely, a low number of inspiral orbits and a narrow range in parameter space, especially in the mass ratio — so-called *hybrid* models have been developed. Hybrid surrogate models [45] construct

a surrogate waveform by treating the inspiral independently from the NR waveforms with respect to which the model interpolates, using the same masses and spin components. Each full waveform consists of a semi-analytical waveform at early times (up to roughly 20 orbits before the ringdown) that is smoothly attached to a NR waveform at late times which is constructed as for the pure surrogate models. For the early inspiral, a mixture of PN and EOB calculations is used. The PN waveform corresponding to the given parameters $\boldsymbol{\lambda}$ is improved using the *SEOBNRv4* model [25]. This results in a EOB-corrected waveform that remains faithful to the NR waveform until much later times, whilst the pure PN waveform becomes inaccurate. Note that since one restricts simulations to the three-dimensional space of non-precessing BBHs for hybrid surrogate models, fewer simulations are necessary compared to the seven-dimensional precessing case. The resulting hybrid model is an NR-based surrogate model to span the entire LIGO frequency band for stellar mass binaries; assuming a low-frequency detector cutoff of 20 Hz, this model is valid for total masses as low as $2.25\,M_\odot$. It is based on 104 NR waveforms in the range given in Table I. For the practical purpose of including precessing events in our analysis, however, we focus on the pure *NRSur7d4* [17] model[1].

### B. Common difficulties of waveform models

#### 1. Reference frames

In the process of calculating waveforms, each model uses multiple reference frames which are best suited for their individual computations. While the waveform families are generally different, there is an overlap in some commonly used frames that simplify numerical calculations. There are essentially four commonly used frames reoccurring in the algorithms that process the waveforms. Following the notation in [36], these are:

**$\hat{\mathbf{L}}_0$- or source-frame:** This frame uses the orbital momentum vector $\vec{L}$ at a certain reference frequency $f_{\rm ref}$ (or time $t_{\rm ref}$) to define the $z$-axis of a right-handed Cartesian coordinate system. This is a time-independent frame.

**$\hat{\mathbf{J}}$- or inertial-frame:** Here the total angular momentum vector $\vec{J}$ is used to define the $z$-axis of a right-handed Cartesian coordinate system. This is another time-independent frame.

**$\hat{\mathbf{N}}$- or observer-frame:** This frame's $z$-axis is defined by the line-of-sight vector $\hat{N}$ from the observer to the centre-of-mass of the binary. This frame is also time-independent.

**$\hat{\mathbf{L}}$- or coprecessing-frame:** This frame coincides with the $\hat{\mathbf{L}}_0$ frame when the orbital frequency $f$ (or time $t$) is equal to $f_{\rm ref}$ (or $t_{\rm ref}$). This is a time-dependent frame. In fact, it is obtained by evolving the $\hat{\mathbf{L}}_0$ in time and it is used to keep track of the precession of the orbital plane. With a small modification of the coprecessing frame, one obtains the so-called *coorbital-frame*, in which the BHs are always on the $x$-axis, with the heavier BH on the positive half.

While these descriptions established a point of origin and one axis of a Cartesian coordinate system, a full definition of the different frames is yet to be provided, i.e., we are left with rotational freedom around the defined axis. One can either constrain this freedom through a more complete definition of frames or leave the remaining freedom to be minimized through an alignment of waveforms. In this work, we adopt the latter approach, as detailed in section III.

#### 2. Alignment:

All the approximants under consideration encounter the challenge of comparing waveforms originating from different sources, each potentially using a distinct coordinate system. To perform an alignment of different waveforms, one aims to identify an initial state in the inspiral from which on each model evolves the BBH forward in time. If this initial state is found, there is a unique rotation translating one frame into each other.
A common strategy is to first identify the time of coalescence for each waveform by considering the peak time of the strain amplitude, and set it to $t = 0$. Then, when comparing to NR simulations, it is possible to obtain the full state of the binary system at some fixed earlier time $t_{\rm ref} < 0$, sufficiently deep in the inspiral phase (but not too early, to avoid incurring in so-called "junk radiation"). Subsequently, the unique rotation of the system that puts the two black holes in the desired configuration in the chosen reference frame, for example with the more massive black hole along the positive $x$-axis, can be determined. An instance of this procedure for precessing *Surrogate* models is described in [16]. There, at reference time $t_{\rm ref} = -4500M$ the data from the numerical simulations are used to find the direction of the orbital angular momentum of the binary, which define the instantaneous orbital plane at $t_{\rm ref}$. A unique rotation and a small error correction are then applied to move the more massive black hole on the $x$-axis and fix any residual ambiguity of the frame definition mentioned above. Note that this implies that the "initial" parameters $\boldsymbol{\lambda}$ that define the waveform are set at $t_{\rm ref}$ only *after* the alignment. However, for the purpose of model comparison, it is necessary to obtain waveforms corresponding to the *same* initial conditions at the same initial time. One possible approach is to perform numerical optimization over the binary parameters at a specific fixed reference time.

[1] Note that recently, the first *Surrogate* model with memory effects, obtained by training the model with CCE-extrapolated numerical simulations [39], has been published [46].

This strategy has been employed for benchmarking *IMRPhenomXPHM* [47], and we also adopt this approach in our study. We refer to section III for a full discussion of the subtleties involved in this choice of the alignment procedure as well as the practical implementation.

#### 3. Precession

Generating waveforms for precessing events, each waveform model family has evolved different strategies. Recent *Phenom* models describe precession effects via the so-called twisting up of the waveform [44]. This procedure is applied to a waveform obtained for the corresponding non-precessing BBH merger. Assuming that the precession frequency is much lower than the orbital frequency, one can employ a stationary-phase approximation and obtain a map between precessing and non-precessing BBH mergers. In this regime, the predominant impact of precession manifests as amplitude and phase modulation, expressible through a time-dependent rotation of the non-precessing system. Once the specific form of this rotation and its dependence on the binary parameters are established, it can be directly applied to the waveforms. Such a time-dependent rotation naturally mixes different modes. In particular, if the non-precessing waveform contains any mode of a certain $\ell_0$, then the precessing waveform receives contributions from *all* modes up to that $\ell_0$. As such, all precessing *Phenom* models contain strain modes of the form $\{(\ell, m) \mid \ell \leq \ell_0\}$ for some $\ell_0$, even if the underlying non-precessing model does not include modes for all $m$ within that range. Crucially, the absent modes in the underlying waveform do not remain absent in the precessing waveform, but the lacking information is now distributed across all modes of that specific $\ell_0$. In such cases, none of the modes of that $\ell_0$ can be expected to be accurately modelled. For instance, the *IMRPhenomXHM* model includes the $h_{4,4}$ mode in $\ell = 4$, hence the twisted up *IMRPhenomXPHM* model [47] receives contributions from all $(4, m)$ modes with $|m| \leq 4$ for a precessing binary, but none of them is "complete" due to lack of, e.g., the $(4, 3)$ mode in the *IMRPhenomXHM* model.

Similar considerations apply to the *EOB* model *SEOBNRv4PHM*. The latter is an augmented version of the non-precessing model *SEOBNRv4HM* (hence sharing effectively the same mode content), where precession is introduced by projecting the spin of the effective background Kerr spacetime onto the orbital angular momentum [26].

On the contrary, precessing *Surrogate* models adopt a different approach that refrains from making specific assumptions about the modelling of the underlying waveform [16]. In principle, they contain the full mode content up to the claimed value of $\ell$ for precessing as well as non-precessing BBH mergers, where each mode is subject to the uncertainties introduced in the fitting procedure with NR waveforms. Consequently, we anticipate superior performance from these models compared to other approximants when precession is a factor. Our analysis, as detailed in section VI, confirms this expectation.

### C. Strengths and weaknesses

Fully numerical simulations can be regarded as the most accurate means to obtain BBH waveforms. Except for the early inspiral, where PN approximations are highly accurate, the intricate equations governing coalescence can only be comprehensively solved through numerical simulations. These sophisticated computations are resource-intensive, demanding substantial computational power. Consequently, only a finite number of discrete points in parameter space have been simulated and compiled in catalogs thus far. The sparse parameter space coverage of the *SXS* catalogue [37], which is still the most comprehensive one, motivated alternative GW models.

As previously explained, the *Surrogate* models utilize a specific set of NR waveforms to parametrically fit a new waveform located in their vicinity in parameter space. However, this implies that *Surrogate* models are once again confined to a specific region of parameter space, albeit to a significantly lesser extent than when only considering the available catalogs.

For *EOB* models, the restrictions in parameter space are weaker than for *Surrogate* models. Due to the analyticity and adjustability of the *EOB* models, they can cover a large portion of the BBH parameter space, once the calibration is performed against NR simulations. For instance, *SEOBNRv1* covers arbitrary mass ratios and black hole spins in the range $-1 \leq \chi_c \leq 0.7$ for the dominant $h_{2,2}$ mode, with further improvements in later versions.

*Phenom* models, on the other hand, again rely predominantly on the calibration to NR simulations. However, as they produce only an approximate waveform these models are still very fast and can be easily extended by introducing more parameters and different fitting functions. The precision can be increased by fitting to more NR simulations, which will naturally yield a drop in performance.

As one might anticipate, the expansion in parameter space comes with certain trade-offs. For instance, in the case of *Surrogate*, the non-precessing waveforms encompass only a limited number of orbits before the merger. As mentioned, this limitation is addressed in hybrid *Surrogate* models to attain an accuracy comparable to *EOB* for the inspiral segment. In extensive data analysis scenarios, however, the evaluation of *EOB* models can be computationally expensive, potentially undermining the purpose of being an approximant.

In summary, each model has its "region of proficiency" and delivers a reliable, accurate waveform for a given point in parameter space. Although the region of applicability may be smaller for some models, they might excel in precision in other aspects. The choice of the waveform

model should be based on the specific requirements at hand, considering factors such as relevant mode content, the segment of the waveform of interest, the importance of quality versus quantity, and so on.

## III. WAVEFORM ALIGNMENT

In what follows, we delineate the alignment procedure employed in this work, which, in some aspects, shares similarities with the previously presented approaches of the waveform models. It constitutes the first step towards a comprehensive comparison of different waveform models.

### A. Alignment and residual ambiguities

To ensure a meaningful comparison, the output waveforms computed by each model must be translated into a common frame. This task has been primarily handled by the *LALSuite* framework, which encompasses the relevant approximants and many more. The waveforms generated through *LALSuite* are already aligned in a common frame, up to a rotation in the orbital plane, or equivalently, up to a phase factor denoted hereafter as $\phi_{\rm ref}$. Provided the waveforms of the approximants are of equal length (i.e., start at a common reference frequency), the orbital phase remains the only obstacle in the alignment of non-precessing waveforms.

An other issue arises when waveform models are compared to the *SXS* catalogue: For a given event associated with a specific set of spins and initial masses, it is possible that the maximal waveform domain of each approximant varies due to different restrictions on the reference frequency $f_{\rm ref}$[2]. In this case, we re-generate the corresponding events for each approximant using a new minimal common $f_{\rm ref}$. While the *LALSuite* approximants can produce waveforms with different initial or reference frequencies $f_{\rm ref}$ for a given event (within intrinsic constraints, see Table I), the *SXS* waveforms come with a fixed reference frequency that cannot be altered. As a result, the necessity of adjusting NR waveforms "by hand" arises, which in turn requires an additional alignment step with respect to the phase factor $\phi_{\rm ref}$.

For precessing events, analogous arguments apply. Here, the ambiguity to be resolved by the alignment extends to the initial directions of the individual spin vectors of the black holes, introducing four more independent parameters (the angles subtended by the spin vectors) to the alignment process, in addition to $\phi_{\rm ref}$.

Both the adjustments of $\phi_{\rm ref}$ and the initial spin vectors are achieved through minimizing a mismatch function $\mathcal{M}(\phi_{\rm ref}, \Omega_1, \Omega_2)$, where $\Omega_i = (\phi_{\chi_i}, \theta_{\chi_i})$ are the pairs of angles for each spin vector. Once minimization is accomplished, we posit that any discrepancies identified when comparing physical quantities on the aligned waveforms can be attributed to inherently distinct evaluations of the binary configuration for each individual model[3].

The mismatch function requires an input that is related to the gravitational strain. Our primary strategy is to align with respect to the dominant $h_{2,\pm2}$ mode, which corresponds to choosing as mismatch function

$$\mathcal{M}(\phi_{\rm ref}, \Omega_1, \Omega_2) \coloneqq 1 - \frac{\langle \tilde{h}^{\rm ref}_{2,2}, \tilde{h}^{\rm align}_{2,2} \rangle}{\|\tilde{h}^{\rm ref}_{2,2}\| \, \|\tilde{h}^{\rm align}_{2,2}\|} \,. \tag{3.1}$$

Here, $\tilde{h}_{2,2}$ corresponds to the Fourier transformed ($\ell = m = 2$)-mode of the strain in the orthogonal decomposition

$$h = \sum_{\ell,m} h_{\ell,m}(u) \, {}_{-2}Y_{\ell,m}(\theta, \phi) \,, \tag{3.2}$$

where ${}_{-2}Y_{\ell,m}$ are the spin-weighted spherical harmonics of weight $-2$. The inner product $\langle \cdot, \cdot \rangle$ is defined in the frequency domain as

$$\langle \tilde{h}_1, \tilde{h}_2 \rangle \coloneqq \int \mathrm{d}f \, \tilde{h}_1(f) \, \tilde{h}_2^*(f) \,, \tag{3.3}$$

where $\tilde{h}^*$ is the complex conjugate of $\tilde{h}$. Naturally, this inner product induces a norm, which we denote as $\|\cdot\|$. The modes $\tilde{h}^{\rm ref}_{2,2}$ and $\tilde{h}^{\rm align}_{2,2}$ are extracted from the reference model and the model to be aligned, respectively. Throughout the analysis, we adopt *SXS* waveforms as reference. For regions in parameter space where no *SXS* waveforms are available, we take *Surrogate* as reference instead.

Generally, the mismatch can be computed with respect to any strain mode or the full gravitational strain as in equation (3.2). It is established common practice to align with respect to the dominant $h_{2,2}$ mode in frequency domain. However, there is a caveat to this approach. When minimizing $\mathcal{M}$ for the $h_{2,2}$ mode with respect to the complete Fourier-space waveform (i.e., its maximal available extent in Fourier space), the resulting alignment can obscure systematic differences between the models. This is a direct result of the procedure of minimizing over all distinctive features of the individual waveforms, including the ones relevant to qualitative comparisons of the approximants, i.e., different dynamics during the merger and ringdown stage. Since our primary interest is to

[2] Here, we refer to the reference frequency as the initial time step from which on the generated waveform is non-trivial. This frequency can be translated to the number of orbits included before the merger, which generally varies throughout waveform models.

[3] Naturally, this is a crude assumption. However, an in-depth analysis of potential shortcomings in the alignment procedure beyond what is discussed in this section lays out of the scope of this work.

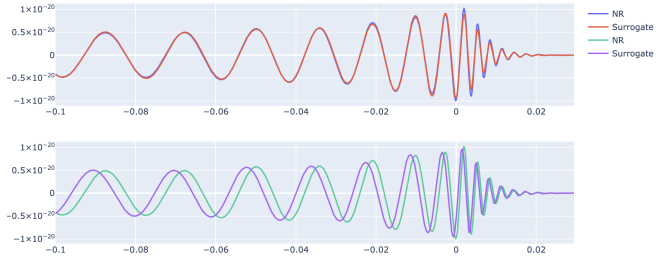



FIG. 1. Aligned (top) and non-aligned (bottom) $h_{2,2}$ waveform mode for a precessing event (SXS:BBH:1011).

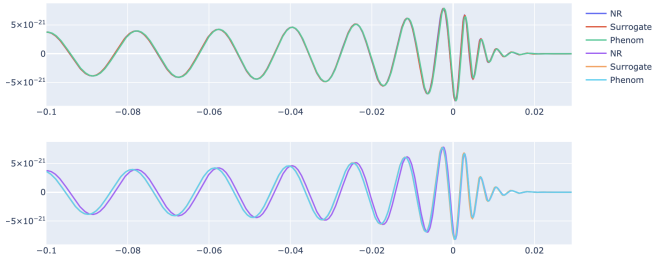



FIG. 2. Aligned (top) and non-aligned (bottom) $h_{2,2}$ waveform mode for a non-precessing event (SXS:BBH:0374).

align the initial sections of the waveforms, i.e., initial frequency, phase, and spins, it is more appropriate to restrict the integration of the inner product (3.3) to the inspiral phase only. Additional support for this alignment strategy can be given by noticing that the inspiral is almost exclusively modeled via a PN approximation throughout all approximants. Thus, intrinsic differences of waveform models in this frequency range can be expected to be minimal.

In the following subsection we scrutinize this alignment strategy and determine the most suitable alignment procedure, which will be employed throughout the rest of the article.

### B. Testing alignment methods

For the following comparison we use the mismatch function (3.1) but allow the use of shorter integration intervals in frequency space. An application of the standard mismatch formula (3.1) (i.e., where the integral ranges over the full frequency spectrum of $h_{2,2}$) is displayed in Figures 1 and 2, for a non-precessing and a precessing waveform. In this case, the two waveforms from *SXS* and *Surrogate* demonstrate excellent agreement after the alignment procedure. The non-precessing waveform is perfectly aligned after optimization over the orbital phase parameter. The precessing event shows very good agreement in phase but some noticeable difference in the amplitude of the peaks close to the merger.

The limitations of the standard alignment procedure become evident when subdominant modes are compared pre- and post-alignment. Figures 3 and 4 show the $h_{2,1}$ modes for precessing and non-precessing events prior to and after the standard alignment procedure. The aligned $h_{2,1}$ modes in the precessing case display pronounced differences, even in the inspiral phase, while in the non-precessing case there are no noticeable differences. This simple comparison illustrates multiple issues arising in the alignment procedure. First of all, aligning with respect to the $h_{2,2}$ mode does not guarantee a perfect alignment of the remaining, subdominant modes, in particular for the precessing case. As the $h_{2,2}$ mode is the dominant one, however, the mismatch in the other modes is not our primary concern. Furthermore, while the overall mode content varies throughout waveform models, the inclusion of the $h_{2,\pm2}$ mode is guaranteed in any case, making it the preferred choice for an alignment strategy. Nonetheless, to address the poor alignment of subdominant modes, aligning with respect to the full strain could present a viable solution. Due to the dominance of $h_{2,2}$ within the full strain $h$, aligning with respect to the $h$, rather than with respect to solely $h_{2,2}$, results only in minor differences. In a limited test including a small sample of non-precessing events, we found that aligning with respect to the full strain results in slightly larger mismatch and differences with respect to NR regarding the quantities studied in section VI, i.e., the remnant velocity magnitude and direction, compared to standard alignment with $h_{2,2}$. This might be partially due to the models' primary focus on generating a maximally precise dominant strain mode component, rendering the subdominant modes intrinsically less exact.

Moreover, Figures 1 and 3 illustrate that larger differences in the waveforms predominantly appear in the merger and ringdown phases. As mentioned, this raises the question whether aligning with respect to only the inspiral phase suffices and whether it is more transparent regarding intrinsic model differences in the merger and ringdown phases. In practice, when working with time-domain waveforms, this can be tested as follows: One identifies the merger time as the peak of the strain norm and then truncates the template waveforms entering equation (3.1) (before taking the Fourier transform) at a time well before the merger, deep in the inspiral phase. Here, we choose to cut at two-thirds of the waveform's full-time interval. This specific cutoff time corresponds to a cut-off frequency which enters the frequency domain integral in (3.3). We compare the alignment with respect to $h_{2,2}$ with and without temporal (frequency) cut-off, over a sample of randomly chosen non-precessing and precessing events. We find that the presence of the cut-off in the alignment procedure marginally worsens the mismatch of full waveform compared to the full mismatch computation. Anticipating the analysis in section VI, we also find that not only the mismatch, but also the relative error in the kick velocity and direction increase as well, albeit only marginally. This result in-

dicates that when aligning with respect to the full waveform, (intrinsic) differences of the model families are obscured compared to aligning with respect to only the inspiral (i.e., with a frequency cut-off). Although this serves as clear evidence that the alignment strategy can overshadow intrinsic model differences, there is currently no direct means to distinguish between alignment residuals and intrinsic waveform dissimilarities. Hence, at this point, we simply emphasize the potential shortcomings of the alignment procedure used throughout this work, which computes (3.3) for $h_{2,2}$ over the full available time (frequency) domain in concordance with the majority of the literature.

Lastly, Figures 1-4 clearly illustrate that the alignment procedure for precessing events exhibits significantly poorer performance compared to non-precessing events. Indeed, the alignment of precessing BBH mergers remains a subject of ongoing investigations within the GW community. The primary challenge arises from the need to identify a specific instance in time when the spins of both black holes point in the same direction for both the reference model and the one to be aligned to it. However, due to the model-dependent length of the generated waveform, this point in time with congruent spin orientations might not even exist for any given event. Disregarding these instances in parameter space, precessing events can, in principle, be aligned using the mismatch function as outlined in equation (3.1), involving adjustments for $\phi_{\text{ref}}$ as well as four angles (two for each spin). Admittedly, it is not guaranteed that this procedure results in the correct congruent spin configuration. Instead, a physically similar orbital configuration sufficiently close with respect to the distance measure induced by (3.1) might be chosen accidentally.

Despite the mentioned caveats, in this work, we employ the above procedure, computing the mismatch (3.1) with respect to $\phi_{\text{ref}}$ as well as four spin angles for all precessing events. Unsurprisingly, the mismatch, on average, turns out to be much larger than for the non-precessing case. Overall, we found the outcome of the alignment procedure for precessing events to be poorly reliable. As a consequence, in section VI, we will refrain from utilizing the alignment for precessing events. That is, we will not compare alignment dependent physical quantities in different models for precessing events but instead focus on other aspects of these events that are not (as) sensitive to the alignment process.

## IV. EVENTS UNDER CONSIDERATION

Given the substantial number of events considered in this study, it is instructive to discuss the covered regions in parameter space. This allows us to establish meaningful connections between our findings and phenomenological insights, especially in cases where one or more approximants exhibit large deviations with respect to the reference model. Additionally, by carefully spreading the considered events over a large parameter space, we avoid introducing selection biases.

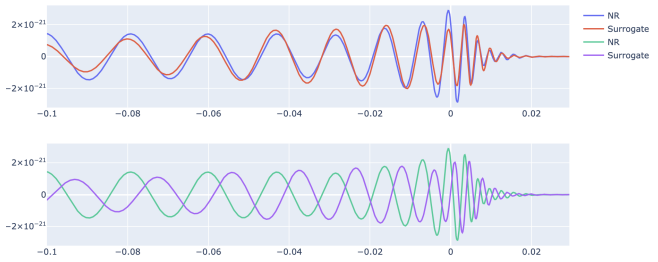



FIG. 3. Aligned (top) and non-aligned (bottom) $h_{2,1}$ waveform mode for a precessing event (SXS:BBH:1011).

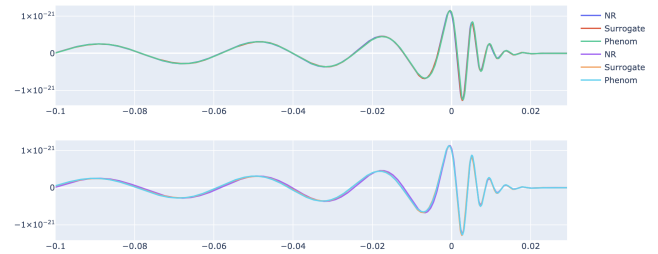



FIG. 4. Aligned (top) and non-aligned (bottom) $h_{2,1}$ waveform mode for a non-precessing event (SXS:BBH:0374).

For each case, that is precessing and non-precessing BBH mergers, we discuss the considered events for which NR counterparts are available and simulate additional ones to diversify the parameter space under investigation.

### A. Non-precessing events

The majority of our analysis resides in examining non-precessing events within the *SXS* catalogue. In Figure 5 we illustrate the distribution across parameter space, distinguishing between aligned and anti-aligned spin configurations and employing color-coding based on the mass ratio $\eta := \frac{q}{(1+q)^2}$ with $q := M_{i^+}/M_{i^\circ}$, corresponding to the ratio of final and initial mass. While a total of 175 events with non-negligible kick, i.e., $v > 20$km/s, is selected for the non-precessing case, the parameter space displayed in Figure 5 seems to be only sparsely covered. This is due to many events depicted in Figure 5 overlapping in the sense that either only their mass ratio or the alignment of spins changes.

From Figure 5 it is evident that regions characterized by lower spins $\chi_1$ and $\chi_2$ are insufficiently represented within this excerpt of cataloged data. To address sparsely populated areas within the parameter space, we introduce additional simulated events without a *SXS* counterpart. These events feature aligned as well as anti-aligned spins and span across an extensive range of mass ratios $\eta$, as

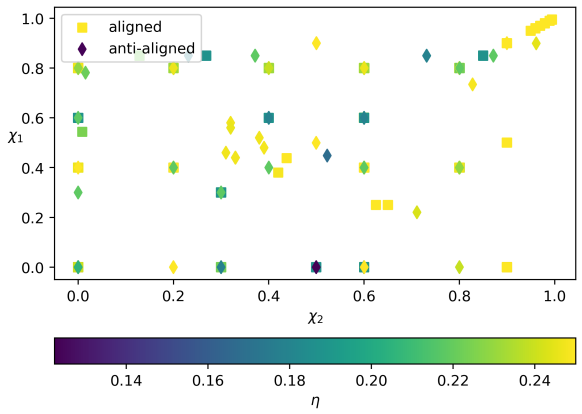



FIG. 5. Parameter space for the non-precessing *SXS* data.

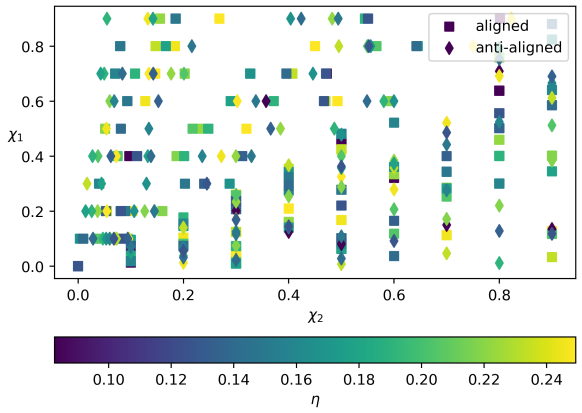



FIG. 6. Parameter space for the non-precessing events without NR counterpart.

depicted in Figure 6. This augmented set adds an additional 220 instances to our set of non-precessing events.

Together, catalogued and non-catalogued non-precessing events form a well-distributed set, which is suitable for an unbiased systematic investigation. Despite all efforts of accurately modeling GW waveforms, for some instances of non-precessing events, certain models may deviate strongly from the reference model and the other approximants in terms of remnant velocity or gravitational memory (again anticipating results of section VI). A closer investigation of such instances reveals that in these cases one of the modes of the deviating model (e.g. the $h_{2,1}$ mode) significantly differs from the corresponding mode of the remaining approximants and the reference model. This behavior predominantly occurs when one of the normalized spin components $\chi_{1,2}$ is extremal, i.e., either close to 0 or 1. Examples of these types of events include for instance *SXS:BBH:0222*, *SXS:BBH:0223*, and *SXS:BBH:0251*, for which *EOB* results in a kick velocity much larger or smaller than for the remaining waveform models. In order to not affect our statistical considerations, such events are excluded from our investigation.

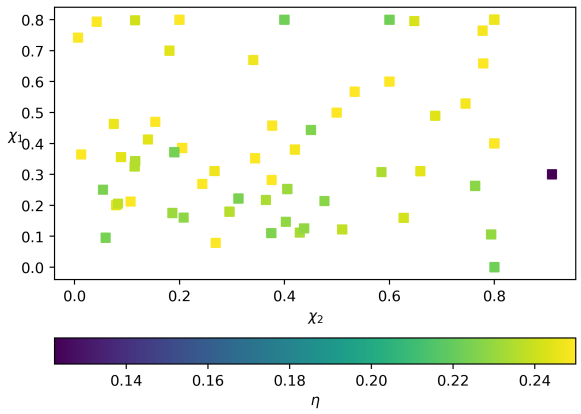



FIG. 7. Parameter space for the precessing *SXS* data.

### B. Precessing events

For the analysis of precessing BBH, we make use of 130 events with *SXS* counterparts, as displayed in Figure 7. In terms of the two spin vectors, the events are homogeneously distributed over the available parameter space. Notably, the catalogued events do not extend to lower mass ratios except for the event *SXS:BBH:0165*, which involves two black holes of masses $51.4\,M_\odot$ and $8.6\,M_\odot$. As this lack of low mass ratio binaries ($\eta < 0.2$) potentially demonstrates a parameter bias, we supplement these precessing *SXS* events with a set of 75 additional instances. The added events are displayed in Figure 8. The difference in number and randomness of the added events in the precessing and non-precessing case is a direct result of the computational complexity of aligning precessing events. For these events, we optimize the mismatch with respect to 5 variables instead of only $\phi_{\rm ref}$, resulting in a significantly larger data generation time. Due to time constraints, the added events in the precessing case are less numerous and more grid-like distributed.

## V. ENERGY-MOMENTUM BALANCE LAWS

To assess a waveform model's performance against a chosen reference model, an adequate measure of comparison has to be selected. We choose the remnant's kick velocity and the gravitational wave memory, which are both physical observables that can be calculated solely from the strain of a GW. In what follows we derive explicit expressions for the kick velocity and the memory from constraint equations which are known as *energy-momentum balance laws* [28]. These laws are exact mathematical results derived from full, non-linear GR when applied to the coalescence of compact binaries. For an exhaustive review and derivation of the balance laws, we refer to [28, 48]. For a proof of concept that the balance laws can be used as diagnostic tools for assessing waveforms

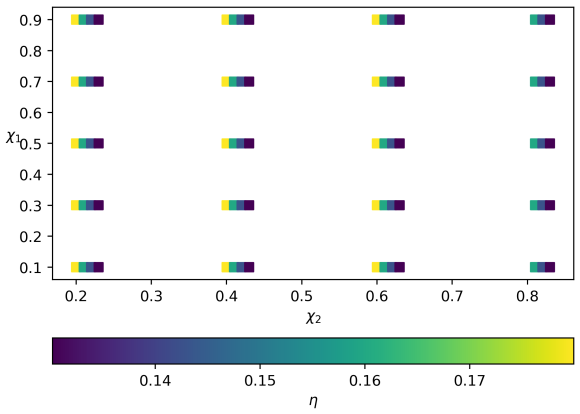



FIG. 8. Parameter space for the precessing events without NR counterpart. For every bundle of data points, the spin configuration is identical. For illustrative reasons, the plot shows a slight misplacement.

we refer to [30], which uses a simple analytical waveform model, and to [29]. Here, we will simply state the main result, explain its physical content, and perform a decomposition into spherical harmonics in order to derive explicit expressions for kick velocity and memory. These expressions then serve as the basis of our comparative analysis.

### A. The balance laws

Waveform models describe GWs emitted from isolated systems composed of two compact objects which orbit each other and coalesce. The coalescence is caused by a loss of orbital energy, due to the emission of GWs. When the two compact objects finally merge, the remnant can be subjected to a "kick". This is a consequence of the fact that GWs are emitted anisotropically and because they do not only carry energy, but also momentum. Based on this simple qualitative picture of the physics of compact binary coalescence, one would expect that there is a mathematical law which describes the energy loss of the system and balances it against the energy carried away by GWs. One may furthermore expect that this law establishes a relation between *(i)* the initial and final mass of the system, *(ii)* the kick velocity, and *(iii)* the strains $h_+$, $h_\times$ of the GW. Such a mathematical law does indeed exist, and it can be derived from full, non-linear GR when applied to compact binary coalescence [28]. As a matter of fact, there is not only one law, but infinitely many. Namely, one per point on the 2-sphere, i.e., one so-called energy-momentum balance law per choice of $(\theta, \phi)$. These laws can be brought into the form

$$c^2\left(\frac{M_{\rm remnant}}{\gamma^3\left(1-\frac{\vec{v}}{c}\cdot\hat{x}\right)^3}-M_{\rm binary}\right)=-\frac{1}{4}\frac{D_L^2\,c^3}{G}\int_{-\infty}^{\infty}\left(\dot{h}_+^2+\dot{h}_\times^2\right)\mathrm{d}t+\frac{1}{2}\frac{D_L\,c^4}{G}\mathrm{Re}\left[\eth^2\left(h_+-i\,h_\times\right)\right]\Big|_{t=-\infty}^{t=+\infty}, \tag{5.4}$$

where $\vec{v}$ stands for the kick velocity, $\gamma \equiv \gamma(v)$ is the usual Lorentz factor from Special Relativity, $\hat{x} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the unit radial vector in spherical coordinates, $D_L$ is the luminosity distance between source and observer, and $\eth$ is a differential operator called "eth" whose precise definition can be found in [28, 48].

By recalling our qualitative description of the physics of compact binary coalescence, we can give a tentative interpretation to the balance laws. Since $E = M\,c^2$, according to Einstein, we can interpret the left-hand side of (5.4) as the difference of two energies measured at two different instants of time. Namely, the energy of the remnant, which corresponds to the energy of the system after the merger, minus the energy of the binary system, which is the energy measured well before the merger. The denominator containing the Lorentz factor and the kick velocity can be understood as a correction due to the fact that the energy of the binary system is measured in its instantaneous rest frame, while the remnant is, in general, moving with respect to that frame with a velocity $\vec{v}$. As we will see below, when we decompose (5.4), for the $\ell = 0$ mode this factor reduces precisely to the special relativistic formula for the kinetic energy of a body moving relative to an inertial frame, $E = \gamma(v)Mc^2$. For a derivation of this factor and technical details we refer the reader to [28]. Since the binary system is losing energy due to the emission of GWs, it is natural to expect that the difference of energies is negative. Indeed, on the right-hand side of (5.4) we find an integral over a manifestly positive quantity, $\left(\dot{h}_+^2+\dot{h}_\times^2\right)$, with a minus sign in front. This integral computes the energy carried by GWs, and it is a well-known result of linearized GR. However, we stress again that these balance laws hold well beyond the linear regime and, in fact, no linearization is ever used. Finally, the second term on the right-hand side represents the GW memory. Because the left side of (5.4) depends on $\hat{x}$, which is a function of $(\theta, \phi)$, and because the strains $h_+$, $h_\times$ are not only functions of time but also functions of the angular variables $(\theta, \phi)$, we find that there is indeed one balance law per choice of $(\theta, \phi)$.

The strength and diagnostic capabilities of the balance laws derive from the fact that they are a precise math-

ematical result which is valid in full GR. The quantities which appear in their formulation, i.e., initial and final mass of the binary system, kick velocity, and the GW strains, are all quantities which either enter as parameters or are provided in an approximate fashion by waveform models. Thus, the exact balance laws can be used to test the accuracy of waveform models.

The balance laws are based on the pioneering work of Bondi, Metzner, Sachs, and van der Burg, who laid the foundations for describing GWs beyond the linear approximation [49, 50]. Their work was subsequently extended and formalized by Newman, Penrose, Geroch, Ashtekar, and others, who introduced precise mathematical definitions for "asymptotically flat" and "asymptotically Minkowski" spacetimes [51–53], which, qualitatively speaking, capture the idea of having an isolated system contained within a finite spacetime region. An other key element in the derivation of the balance laws are the symmetries of asymptotic Minkowski spacetimes. These are described by the so-called Bondi-Metzner-Sachs (BMS) group [49]. By definition, this group leaves the asymptotic structure on null infinity $\mathscr{I}^+$ (read: scri plus) invariant. A variant of Noether's theorem can then be applied, which leads to a description of the (non-)conservation of Noether charges: The energy-momentum balance laws [54, 55]. The balance laws (5.4) shown here are an adaptation to the special case of compact binary coalescence. We refer to [48] for a didactic introduction and recent review of the fascinating mathematical subjects (asymptotic Minkowski spacetimes, the BMS group, balance laws), which here we only brushed over.

In what follows, we will need the mode-decomposition of equation (5.4) with respect to spherical harmonics $Y_{\ell m}(\theta,\phi)$. In order to compactify the notation, we set[4]

$$M_{\text{binary}} \coloneqq M_{i^\circ} \qquad \text{and} \qquad M_{\text{remnant}} \coloneqq M_{i^+}\,, \tag{5.5}$$

and we also introduce the asymptotic shear

$$h(t,\theta,\phi) \coloneqq \frac{1}{2}\left(h_+ + i\,h_\times\right)(t,\theta,\phi)\,, \tag{5.6}$$

where $i$ is the imaginary unit. This last definition in particular allows us to re-write the memory term compactly as $\eth^2\Delta\bar{h}$, with $\bar{h}$ denoting the complex conjugate of $h$ and where

$$\Delta\bar{h} \coloneqq \int_{-\infty}^{+\infty} \dot{\bar{h}}\,\mathrm{d}t\,. \tag{5.7}$$

[4] This is standard notation in the balance law literature and it hints at its origin: The balance laws formally hold on null infinity, $\mathscr{I}^+$. If one takes the limit into the distant past, i.e., well before the binaries merge, one reaches spacelike infinity, $i^\circ$. On the other hand, if one moves along $\mathscr{I}^+$ into the distant future, well after the binaries have merged, one reaches timelike infinity, denoted by $i^+$.

Re-writing the memory term in this fashion is possible thanks to the fact that $\mathrm{Im}(\eth^2\bar{h})$ vanishes at $t=\pm\infty$ (cf. [28]).

Using these definitions and expanding both sides of equation (5.4) into spherical harmonics, we obtain a tower of constraints, namely one for each pair of $(\ell,m)$ with $|m|\leq\ell$. Concretely, this decomposition reads

$$\left(M_{i^\circ} - \frac{M_{i^+}}{\gamma^3\left(1-\frac{\vec{v}}{c}\cdot\hat{x}\right)^3}\right)_{\ell,m} = \frac{D_L^2 c}{G}\int_{-\infty}^{+\infty}\left(|\dot{h}|^2\right)_{\ell,m}\mathrm{d}t - \frac{D_L c^2}{G}C_\ell\Delta\bar{h}_{\ell,m}\,, \tag{5.8}$$

where the $C_\ell$ coefficients are defined as

$$C_\ell \coloneqq \sqrt{(\ell-1)\ell(\ell+1)(\ell+2)}\,. \tag{5.9}$$

To arrive at this expression, we used well-known properties of spherical harmonics and the explicit expression of $\eth^2\bar{h}$, which can for instance be found in [48]. Note that $C_\ell$ tells us that not each term in the balance laws (5.4) contributes to each mode. In fact, $C_\ell$ vanishes for $\ell=0$ and $\ell=1$. Thus, the memory term only contributes for $\ell\geq 2$. Treating the $\ell<2$ and $\ell\geq 2$ modes of the balance laws separately allows us to devise a method for computing the remnant mass, in case it is not provided by the waveform model, the kick velocity, and the memory solely based on the GW strain.

For carrying out the mode-decomposition of the left hand side of equation (5.8), it is convenient to first align the $z$-axis of the rest frame in which $M_{i^\circ}$ is measured with $\vec{v}$ using a rotation in the plane spanned by $\vec{v}$ and $\hat{z}$. The first few $(\ell,m)$ coefficients then read

$$\begin{aligned}
(\ell=0,m=0): &\quad 2\sqrt{\pi}(\gamma M_{i^+} - M_{i^\circ}),\\
(\ell=1,m=0): &\quad 2\sqrt{3\pi}\gamma M_{i^+} v,\\
(\ell=2,m=0): &\quad \frac{\sqrt{5\pi}\gamma M_{i^+}\left(5v^3 + 3\gamma^{-4}\tanh^{-1}(v) - 3v\right)}{v^3}\,.
\end{aligned}$$

Observe that the first term is, up to a factor of $c^2$, simply the difference between the rest energy $M_{i^\circ}$ of the binary and the kinetic energy $\gamma\,M_{i^+}$ of the remnant, while the second term is the ($z$-component) of the momentum of the remnant. The third term does not admit a simple physical interpretation, but we emphasize that the inverse of the hyperbolic tangent of $v$ caused numerical instabilities in the code we used for our analysis. We therefore Taylor expanded $\tanh^{-1}(v)$ up to sixth order in $v/c$. To obtain the corresponding modes in the original frame, we transform the above modes using Wigner $D$-matrices $D^\ell_{m,m'}$ for the inverse rotation used to align $\vec{v}$ and $\hat{z}$.

The decomposition of the GW energy integral in (5.8) can easily be carried out. In fact, the integrand can be written as

$$|\dot{h}|^2 = \sum_{\ell,m}\alpha_{\ell m}Y_{\ell m}(\theta,\phi)\,. \tag{5.10}$$

A straightforward computation reveals that the $\alpha_{\ell m}$ coefficients are given by (see also [29])

$$\alpha_{\ell m} = \sum_{\ell_1=2}^{\infty}\sum_{\ell_2=2}^{\infty}\sum_{|m_1|\leq\ell_1}\sum_{|m_2|\leq\ell_2}(-1)^{m_2+m}\dot{h}_{\ell_1 m_1}\dot{\bar{h}}_{\ell_2 m_2}\sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell+1)}{4\pi}}\begin{pmatrix}\ell_1 & \ell_2 & \ell\\ m_1 & -m_2 & -m\end{pmatrix}\begin{pmatrix}\ell_1 & \ell_2 & \ell\\ 2 & -2 & 0\end{pmatrix}. \tag{5.11}$$

The $\left(\begin{smallmatrix}\ell_1 & \ell_2 & \ell\\ m_1 & m_2 & m\end{smallmatrix}\right)$ denotes the Wigner-3$j$ symbol, which effectively determines which strain modes $\dot{h}_{\ell_1,m_1}\dot{\bar{h}}_{\ell_2,m_2}$ couple to each other in (5.11).

### B. Remnant mass and velocity

For astrophysically realistic kick velocities one can assume $\gamma \approx 1$. This approximation reduces the $\ell = 0$ mode of the balance laws to an energy conservation equation of the form

$$c^2(M_{i^\circ} - M_{i^+}) = \frac{D_L^2 c^3}{16\pi G}\int_{-\infty}^{\infty} \mathrm{d}t \oint \mathrm{d}\Omega\, |\dot{h}|^2\,, \tag{5.12}$$

where $\oint \mathrm{d}\Omega$ is an integral over the unit 2-sphere. In agreement with our intuition, the mass loss of the system is accounted for by the energy radiated away by GWs. For waveform models which provide the mass of the remnant, this formula can be used as a consistency check. On the other hand, when a waveform model does not provide $M_{i^+}$, the above formula can be used to determine the remnant's mass using only the total mass of the binary system, the luminosity distance, and the GW strain as input. Upon using the expansion coefficient (5.11) for $\ell = m = 0$, the energy conservation equation can also be expressed as

$$c^2(M_{i^\circ} - M_{i^+}) = \frac{D_L^2 c^3}{8\sqrt{\pi} G}\int_{-\infty}^{\infty} \mathrm{d}t\, \alpha_{0,0}. \tag{5.13}$$

The $\ell = 1$ mode of the balance laws encode conservation of linear momentum, as one might have expected. This allows us to solve the momentum equations for the components of the kick velocity, which can be expressed as

$$v_i = \frac{D_L^2 c^2}{16\pi G M_{i^+}}\int_{-\infty}^{\infty} \mathrm{d}t \oint \mathrm{d}\Omega\, \hat{x}_i\, |\dot{h}|^2, \tag{5.14}$$

in the original reference frame (i.e., the frame where $\vec{v}$ and $\hat{z}$ are not aligned). Here, $\hat{x}_i$ stand for the $i$-th component of the radial unit vector $\hat{x} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. The kick velocity components can of course also be expressed in terms of the $\alpha_{\ell m}$ coefficients for $\ell = 1$ and $m = -1, 0, +1$. One obtains

$$v_1 = \frac{D_L^2 c^2}{16\pi G M_{i^+}}\sqrt{\frac{2\pi}{3}}\int_{-\infty}^{\infty} \mathrm{d}t\, (\alpha_{1,-1} - \alpha_{1,1}) \tag{5.15}$$

$$v_2 = \frac{-iD_L^2 c^2}{16\pi G M_{i^+}}\sqrt{\frac{2\pi}{3}}\int_{-\infty}^{\infty} \mathrm{d}t\, (\alpha_{1,-1} + \alpha_{1,1}) \tag{5.16}$$

$$v_3 = \frac{D_L^2 c^2}{8\pi G M_{i^+}}\sqrt{\frac{\pi}{3}}\int_{-\infty}^{\infty} \mathrm{d}t\, \alpha_{1,0}\,. \tag{5.17}$$

Equations (5.15)-(5.17) determine the kick velocity using luminosity distance, remnant mass, and strain as input. If $M_{i^+}$ is not provided by a waveform model, it can be determined via the energy conservation equation (5.12).

### C. Gravitational memory

As we have seen above, modes with $\ell \geq 2$ contain contributions from the GW memory. The memory term can be decomposed into two contributions,

$$\Delta\bar{h}_{\ell,m} = \Delta\bar{h}^{\text{lin}}_{\ell,m} + \Delta\bar{h}^{\text{non-lin}}_{\ell,m}\,. \tag{5.18}$$

Following standard conventions in the literature, we define the first contribution as

$$\Delta\bar{h}^{\text{lin}}_{\ell,m} \coloneqq \frac{G}{C_\ell D_L c^2}\left(\frac{M_{i^+}}{\gamma^3(1-\frac{\vec{v}}{c}\cdot\hat{x})^3} - M_{i^\circ}\right)_{\ell,m} \tag{5.19}$$

and call it the *linear memory* [5, 56]. The second term, defined as

$$\Delta\bar{h}^{\text{non-lin}}_{\ell,m} \coloneqq \frac{D_L}{4C_\ell c}\int_{-\infty}^{\infty} \mathrm{d}t\, \alpha_{\ell m} \tag{5.20}$$

is known as the *non-linear memory* [6, 7]. Other names for these two terms are *ordinary memory* for the former and *null memory* for the latter [33]. More importantly, we interpret the $\ell \geq 2$ modes of the balance laws (5.8) as constraints on the memory terms. The same constraints have been used in previous studies to either add memory to waveform models or correct waveform models which did not accurately incorporate the memory effect [29, 33]. In this work, we instead use the linear and non-linear memory inferred from (5.19) and (5.20), respectively, as a means of comparison between different waveform models.

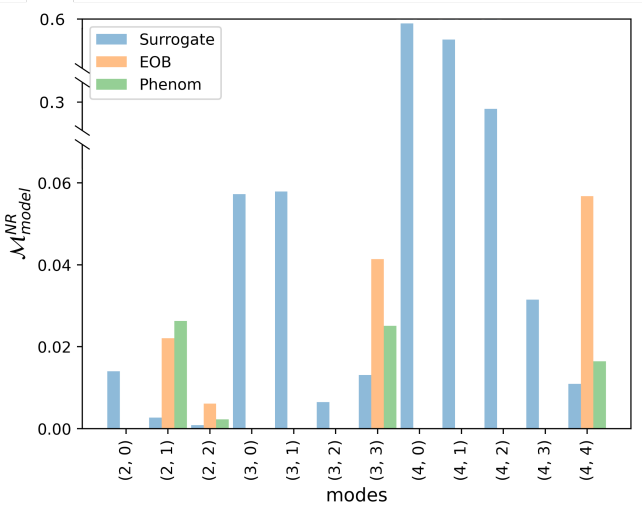



FIG. 9. Mode mismatch with respect to NR, averaged over all non-precessing events.

To summarize this section, we decomposed the balance laws (5.4) into spherical harmonic modes to obtain analytical equations for the remnant mass, the kick velocity, and the GW memory. The remnant mass $M_{i^+}$ is inferred from the $\ell = 0$ mode of the balance laws, while the kick velocity is deduced from the conservation of linear momentum implied by the $\ell = 1$ mode. The $\ell \geq 2$ modes are used to infer linear and non-linear GW memory. Equations (5.12), (5.14), (5.19), and (5.20) are implemented numerically and applied to all the models and events considered in this work. The resulting waveform assessment is the subject of the next section.

## VI. WAVEFORM ASSESSMENT

Computing the remnant velocity following equations (5.15)-(5.17) and evaluating the gravitational memory, encoded in equations (5.19) and (5.20), we assess the performance of the chosen waveform approximants across the parameter space specified in III. Specifically, we examine the impact of the subdominant strain modes on the kick velocity and gravitational wave memory. We restrict our analysis of alignment-sensitive quantities to non-precessing events only.

### A. Mode Mismatch

We start by considering the non-precessing events detailed in section IV, for which NR waveforms are accessible. A natural measure of comparability between waveforms (or individual modes) is the mismatch function (3.1). In the alignment procedure, the mismatch function $\mathcal{M}$ is minimized for the dominant $h_{2,2}$ mode. The remaining modes are not exposed to any direct mismatch minimization. Thus, a first intuition about potential shortcomings of waveform models can be gained based on the residual mismatch of the subdominant modes, and how the latter compares to the mismatch in the dominant $h_{2,2}$. In Figure 9, we present the mismatch calculated via (3.1) for the mode content common to all waveform models[5], averaged over all events under consideration. As expected, we observe a low mismatch for the dominant $h_{2,2}$ mode throughout approximants, with *Surrogate* performing best in this metric. For subdominant $\ell$ modes, the mismatch becomes larger. Particularly for $\ell = 4$ modes but also for $h_{2,0}$, $h_{3,2}$, $h_{3,1}$, $h_{3,0}$ and their complex conjugates, the mismatch is multiple orders of magnitude larger than for $h_{2,2}$. The higher mismatch between *Surrogate* and NR in subdominant modes is attributed to *Surrogate* prioritizing the modelling accuracy of the dominant $h_{2,2}$ in its interpolation procedure.

For multiple subdominant modes, *EOB* and *Phenom* obtain trivial contributions. As this would result in a mismatch $\mathcal{M} = 1$, we choose not to include those instances for *EOB* and *Phenom* in Figure 9. The absence of these modes extends to precessing events as well, despite internal procedures modifying the mode content via, for instance, the twisting-up procedure in *Phenom*, as outlined in section II B 3. However, the methods by which precessing waveforms are obtained from non-precessing ones cannot compensate the lack of information stored in the subdominant modes of a model. Thus, for precessing events a similar pattern regarding the mode-by-mode mismatches as in Figure 9 can be found.

We conducted a similar analysis for non-catalogued events, i.e., events for which no NR waveform is available. In this case, we chose *Surrogate* as the reference model and observe an analogous relationship between the mismatch of *EOB* and *Phenom* versus *Surrogate*. Specifically, for $\ell = 2$, the mismatch with respect to *Surrogate* for both *EOB* and *Phenom* is low. However, for $h_{3,3}$ and $h_{4,4}$, *EOB* waveforms exhibit visibly stronger dissimilarities with respect to *Surrogate* than *Phenom* does.

In addition to intrinsic differences in the waveform modes, the mismatch might be influenced by what we term "numerical artifacts", as discussed in appendix A. These artifacts, particularly prominent in higher modes, require separate consideration. In the literature, some artifacts similar in nature to what is displayed in Figure 15 have been addressed and mitigated, for instance, through the alignment of BMS charges [57]. Here, we will not delve into a detailed discussion about their origin or impact but simply point out that even with a well-defined alignment procedure, the numerical nature of the analysis can potentially affect the mismatch results.

Despite the higher mismatch of certain modes, the full waveforms provide an adequate approximation to NR due to the dominance of the well-modeled $h_{2,2}$ mode for the

[5] Note that we display only up to $\ell = 4$ due to the restricted mode content of *Surrogate*.

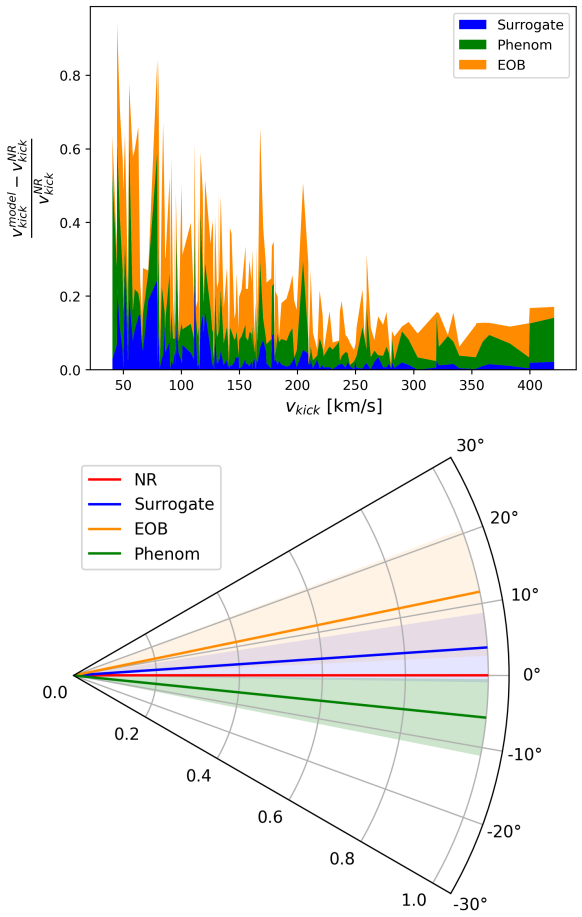



FIG. 10. Relative error for the kick magnitude (top) and absolute deviation angle (bottom) of the selected approximants with respect to NR for non-precessing events. The top stack plot shows the relative errors for each individual event labeled by its kick magnitude. The bottom plot shows the mean deviation over all tested events with respect to NR. A $2\sigma$-interval around the mean is colorized correspondingly.

majority of events. At the current level of precision, addressing the individual subdominant mode mismatch is thus only of secondary interest. However, in the perspective of future GW probes, the precision benchmark will increase significantly, potentially rendering higher modes more relevant [58]. Hence, it is imperative for future GW waveform models to confront the evident challenges depicted in Figure 9, such as incomplete mode contents and decreasing precision of higher-order (subdominant) modes. In the subsequent discussion, we establish the groundwork to address these issues for individual waveform models by examining the feasibility of subdominant mode testing using the analytical tools presented in section V.

### B. Testing Subdominant Modes

In order to investigate the impact of subdominant modes, we analyze physical quantities of BBH mergers with regard to their sensitivity towards mode content. Concretely, we compute the relevant quantities using two sets of strain modes. The first set includes all available modes for each model, the second set does only include high-mismatch subdominant modes, e.g. we exclude modes belonging to the set[6]

$$H_{\rm sub} = \{h_{2,0}, h_{3,\pm2}, h_{3,\pm1}, h_{3,0}, h_{4,\pm3}, h_{4,\pm2}, h_{4,\pm1}, h_{4,0}\}.$$

#### 1. *Remnant Velocity*

We first consider the recoil velocity of the remnant BH. Our attention thereby is limited to events with kick velocities varying between 20 and 400 km/s. Events with a negligible kick (i.e., $v < 20$ km/s) are excluded from the analysis. We compare the magnitude and direction of the remnant velocities (5.14) by computing each models' relative error with respect to NR. Computed using the full mode content, the relative error for the kick magnitude and the directional deviation of the remnant velocity vectors are displayed in Figure 10. In the top plot, we stack the relative errors and plot them against the corresponding kick magnitude for all events. The bottom plot displays the average of the absolute deviation in terms of kick direction with respect to NR. Negative angles for *Phenom* are assigned for better visibility in the plot.
From a quantitative point of view, Figure 10 shows a good agreement between *Surrogate* and NR, as expected. In particular, for larger velocities, the relative errors are small and the directional deviation is less than 5 degrees on average. *EOB* and *Phenom* waveforms cannot reproduce the reference velocity vector with comparable accuracy. They also display a similar trend of larger relative errors for lower remnant velocities. The strong fluctuations indicate a higher dependency on the individual event and its corresponding extrinsic parameters. For both the kick magnitude and direction, *EOB* exhibits on average the largest errors throughout the tested parameter space. In total, these results agree with the findings of previous assessments based on kick velocities [31].
Computing similar statistics without the subdominant high-mismatch modes $H_{\rm sub}$, we find that neither the kick magnitude nor its direction exhibit significant differences compared to the values computed using the full mode content. For NR and *Surrogate*, the change in directional deviation and the relative magnitude error, on average, corresponds to only a few percent of the corresponding full-mode content values. For *EOB* and *Phenom* we do

[6] Note that modes that exceed the mode content of *Surrogate* model, i.e., $\ell > 4$ are not excluded for the remaining models (including NR).

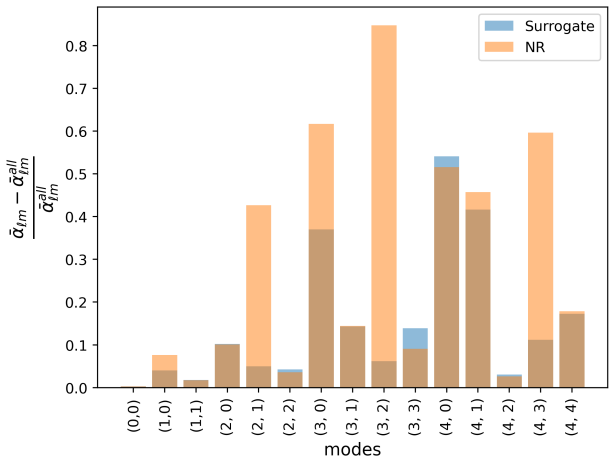



FIG. 11. Time-integrated, relative error for $\alpha_{\ell m}$ containing all strain modes compared to neglecting high-mismatch modes for NR and *Surrogate*.

not find statistically significant differences between excluding or including $H_{\rm sub}$. The same applies to non-catalogued non-precessing events.

Our results indicate that quantities related to the remnant velocity of the final BH are not sensitive to the subdominant mode content of the strain produced by a GW waveform model. This observation extends to the computation of the final mass of the remnant black hole (cf. equation (5.12)). The insensitivity of the remnant's mass and kick velocity with respect to the set of modes $H_{\rm sub}$ can be readily explained by the analytical expressions (5.11) for the $\alpha$-coefficients and equations (5.15)-(5.17): The Wigner-3$j$ symbols appearing in equation (5.11) single out individual terms $h_{\ell_1,m_1}h_{\ell_2,m_2}$[7] which contribute to each $\alpha_{\ell m}$. These coefficients, in turn, determine physical quantities such as the kick vector. According to equations (5.15)-(5.17), the coefficients relevant to the kick are $\alpha_{1,\pm1}$ and $\alpha_{1,0}$. Among all the modes combining to pairs within these coefficients, the most relevant mode combinations involve at least one dominant strain mode, $h_{2,\pm2}$. For the velocity components, such a mixing appears exclusively in $\alpha_{1,0}$, where it is observed that the term proportional to $h_{2,2}h_{2,2}$ cancels the term containing $h_{2,-2}h_{2,-2}$. Therefore, the most significant contributions to the kick components arise from the mixing of $h_{2,\pm1}$ modes with $h_{2,\pm2}$ in $\alpha_{1,\pm1}$[8]. Contrarily, for $\alpha_{1,0}$ there is no non-vanishing contribution combining two $\ell = 2$ strain modes, which demonstrates analytically the comparatively small $z$-component of the kick in non-precessing events. Naturally, the latter statement is not true in the precessing case as more power is distributed across subdominant modes, while $h_{2,2}$ still remains dominant.

Taking into account $\ell > 2$ modes, for $\alpha_{1,\pm1}$, the strain mode $h_{2,2}$ also appears in combinations such as $h_{2,\pm2}h_{3,\pm1}$ and $h_{2,\pm2}h_{3,\pm3}$, which cancel in the sum (5.11). Furthermore, when combining $\alpha_{1,1}$ and $\alpha_{1,-1}$ in the $v_x$ and $v_y$ velocity components, one of the components receives a larger contribution due to the cancellation of $h_{2,\pm2}h_{2,\pm1}$ terms in the sum $\alpha_{1,1} \pm \alpha_{1,-1}$.

Using the above arguments, it follows that the dominant contribution in the $\alpha$-components relevant to the kick and proportional to $h_{2,\pm1} \notin H_{\rm sub}$ are only marginally affected by the exclusion of subdominant modes $H_{\rm sub}$; therefore, neither magnitude nor direction changes significantly when disabling $H_{\rm sub}$ in the computation. A similar conclusion applies also to the final mass, which is determined by $\alpha_{0,0}$ and the strain mode mixing contained therein.

The qualitative analysis of the mode contributions to the kick is quantified in Figure 11. It displays the relative error of the $\alpha$-coefficients for NR and *Surrogate* for $H_{\rm sub}$ included versus excluded[9]. The plot clearly illustrates that the alteration in the $\alpha$-coefficients relevant for the kick velocity is minimal. The more pronounced difference in NR is explained by its expanded mode content ($\ell \leq 8$) compared to *Surrogate* ($\ell \leq 4$), influencing a greater number of mode mixing terms in the summation of $\alpha_{1\pm1}$ when $H_{\rm sub}$ is disregarded.

In conclusion, the fundamental limit of the kick velocity in assessing subdominant modes results from the restricted mixing appearing in $\alpha_{1m}$. Based on the Wigner-3$j$ symbols in (5.11), the allowed coupling of modes is restricted to $|\ell_1 - \ell_2| \leq \ell$ and, correspondingly, the dominant $h_{2,2}$ mode does not combine with any higher modes. This, however, changes when computing the memory.

### 2. *Memory Components*

The linear and non-linear GW memory induced by the waveforms is computed and decomposed in equation (5.18). Figure 12 illustrates the relative errors for both linear and non-linear contributions to the memory in non-precessing events, considering the full mode content. In the upper plot, the total linear memory is normalized to the NR contribution for each waveform model. The colorized regions indicate the $2\sigma$ interval of the average over all events. The solid lines represent the mean values. Similar quantities are presented in the bottom plot for the non-linear memory, depicted as a time series to observe the error accumulation as the non-linear memory builds up. The time $t = 0$ corresponds to the merger

[7] We suppress time derivative and conjugation symbols for ease of notation.

[8] For $\alpha_{1\pm1}$, one finds also terms such as $h_{2,\pm2}h_{3,\pm1}$ and $h_{2,\pm2}h_{3,\pm1}$. However, the $\ell = 2$ modes dominate over $\ell = 3$.

[9] *EOB* and *Phenom* are excluded from the presentation as their $\alpha$-coefficients do not show a noticeable change when $H_{\rm sub}$ is omitted. In these models, every mode in $H_{\rm sub}$ is already vanishing.

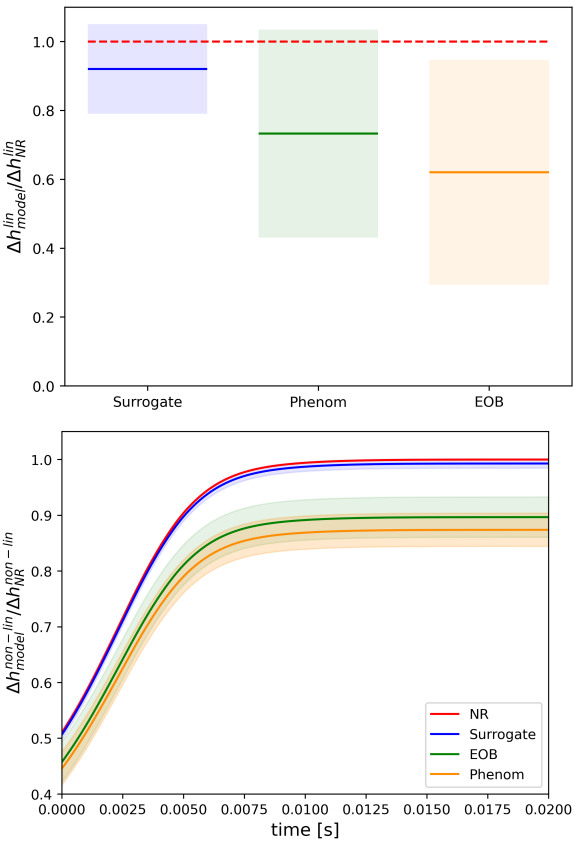



FIG. 12. Deviation of waveform models' linear (top) and non-linear (bottom) memory with respect to NR. Corresponding $2\sigma$-intervals are colorized. The non-linear memory is plotted as a time series. Both memory contributions are normalized to the NR value and averaged over all events.

time, and the normalization is performed with respect to the time-integrated non-linear NR memory.

In terms of both linear and non-linear memory, *EOB* and *Phenom* exhibit notably poorer performance compared to the *Surrogate* approximant. These two models consistently produce smaller memory values. In contrast, *Surrogate* performs well, particularly in the case of the dominant non-linear GW memory. For *EOB* and *Phenom*, the NR values fall outside the $2\sigma$ range. *EOB* also shows a similar trend for the linear part, suggesting systematic deficiencies regarding the memory in these models.

In general, the linear memory is highly non-trivial to model. Equation (5.19) was derived under the assumption that the waveform (and modes) is well-defined throughout time, which is a crude approximation. A more exact formulation of the linear memory is given in [33], which would require the extraction of certain Newman-Penrose scalars not provided for the majority of events. In addition to the considered approximation, the linear part is many orders of magnitude smaller than the non-linear part, making it more prone to numerical errors in the models. Consequently, the linear memory exhibits large statistical fluctuations.

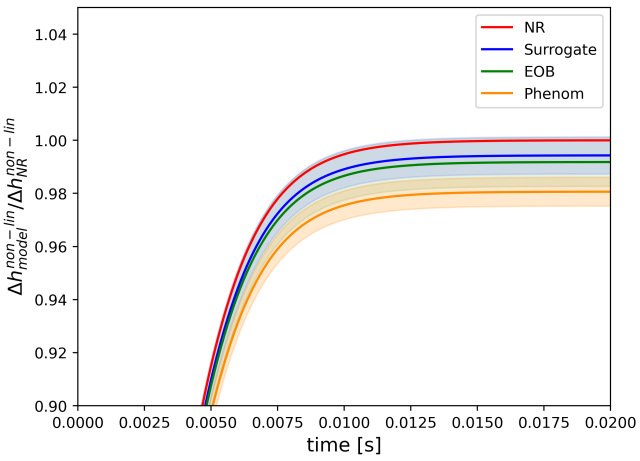



FIG. 13. Non-linear memory as depicted in Figure 12, here computed without high-mismatch modes, i.e., excluding $H_{\rm sub}$.

As before, we now compute the memory without the use of the subdominant modes $H_{\rm sub}$. We find that, for the linear memory component, the change of mode content, similar to the kick velocity, results only in a marginal change of the models individual contributions. The reason is found in (5.19), where the defined linear memory is solely determined by the remnant velocity. This dependency on the kick is a result of the approximation used in the derivation of the balance laws as we state them, i.e., (5.4), where the Newman-Penrose scalars originally appearing in the balance balance laws can be replaced by functions of the remnant velocity in certain limits[10]. Consequently, as the linear memory is determined by the kick, the discussion regarding the strain modes entering its computation follows similar arguments as for the kick in the previous subsection.

A completely different picture is drawn by the non-linear memory when disabling the set of modes $H_{\rm sub}$. As depicted in Figure 13, the relative errors throughout models are approaching each other when the mode content is reduced. While the absolute values for NR and *Surrogate* have decreased by $\mathcal{O}(10)$ percent, *EOB* and *Phenom* have virtually not changed. Consequently, after excluding $H_{\rm sub}$, the non-linear memory contribution of all considered approximants is roughly equal, and, by the definition of the balance laws, so is their energy flux. This statement is independent of the alignment procedure as the non-linear memory is not sensitive to the chosen reference phase[11]. Unlike the non-linear memory, the kick velocity, and thus the linear memory, is sensitive to the alignment procedure. Therefore, some of the discrepan-

[10] For a detailed discussion on the relevant transformations and physical assumptions we refer to [28].

[11] This is, in principle, also true for both non-precessing and precessing events, where for the latter spins have to be taken into consideration as well.

cies shown in the top of Figure 12 may be attributed to suboptimal alignment.

As anticipated by previous literature (e.g. [32] and compare to findings in [29]), *EOB* and *Phenom* models do not reproduce the full information content with their generated strain modes. Comparing Figures 12 and 13, it becomes apparent that the modes contained in $H_{\rm sub}$ roughly constitute the discrepancy in (memory) information content transmitted through the waveforms of *EOB* and *Phenom* with respect to NR (and *Surrogate*). To explicitly demonstrate this, we once again turn to the analysis of strain mode mixing in the memory components: According to equation (5.20), the total non-linear memory contains all $\alpha$-coefficients with $\ell > 1$, where $\alpha_{2,0}$ dominates the sum of coefficients. This is a consequence of the mode mixing present in the $\alpha$-coefficients[12]. As we already elaborated, two dominant strain modes $h_{2,2}$, $h_{2,-2}$ can only mix for $\alpha$-coefficients such as $\alpha_{\ell,0}$. However, the $h_{2,2}h_{2,2}$ and $h_{2,-2}h_{2,-2}$ terms cancel for all $\ell$, except for $\alpha_{2,0}$ and $\alpha_{4,0}$. For $\alpha_{\ell,m}$ with $\ell > 4$ such couplings are forbidden by virtue of the Wigner-3$j$ symbols in (5.11). For $\alpha_{4,0}$ the prefactor in (5.11) is of order $\mathcal{O}(10^{-2})$, suppressing its contribution to the full non-linear part. Consequently, on average, $\alpha_{2,0}$ accounts for roughly 96% of the non-linear memory.

Naturally, the majority of $\alpha_{2,0}$ is due to the term including two dominant modes. However, there is a significant part of the $\alpha_{2,0}$ component as well as the contributions from all other $\alpha$-coefficients that are due to subdominant modes coupling to $h_{2,\pm2}$. As many more $\alpha$'s enter the non-linear memory and for $\ell > 1$ more mixing of the strain modes is permitted by the Wigner-3$j$ symbols, the impact of subdominant modes is non-negligible (in contrast to what was observed for the kick velocity) and grows with the number of strain modes included in the computation of the non-linear memory. The difference in the non-linear memory between Figures 12 and 13 is, hence, largely due to the coupling of $h_{2,2}$ with subdominant modes in $\alpha$'s with $\ell > 1$. Switching of the (for *EOB* and *Phenom*) missing subdominant modes in NR and *Surrogate*, reduces the memory mainly to the $h_{2,\pm2}h_{2,\pm2}$ couplings that, due to the well modeled dominant strain mode throughout approximants, are roughly equal to *EOB* and *Phenom* as indicated by Figure 13.

Similar results are obtained in the precessing case, as shown in Figure 14. In the top plot, the average relative errors for the full mode content are displayed as time series. It is observed that *EOB* and *Phenom* exhibit similar deviations from NR as in the non-precessing case. Additionally, *Surrogate* obtains considerably larger deviations compared to non-precessing events and displays a large standard deviation. The latter persists when $H_{\rm sub}$ is excluded, as shown in the bottom plot of Figure 14. Nonetheless, for the non-linear memory, it still performs better than *EOB* and *Phenom* including the full mode content. When the modes in $H_{\rm sub}$ are disabled, *Surrogate*, *EOB* and *Phenom* exhibit identical memory contributions. Notably here, *Surrogate* has the largest $2\sigma$ interval.

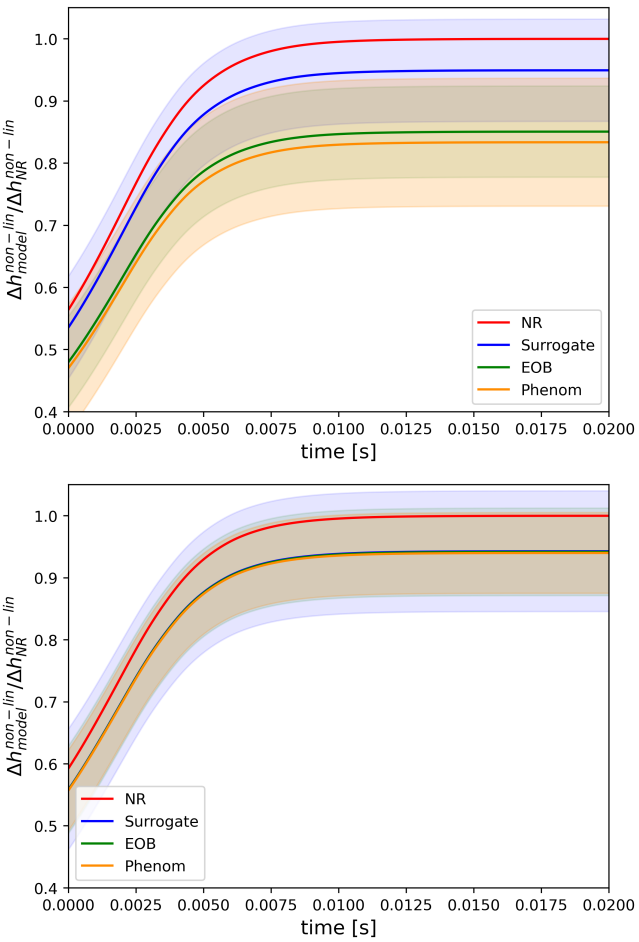



FIG. 14. Non-linear memory as depicted in Figure 12, here computed with (top) and without (bottom) high-mismatch modes, i.e., excluding $H_{\rm sub}$, for precessing events.

Although the non-linear memory contributions of the approximants closely resemble each other, we do not observe a similar convergence to NR as in Figure 13. Instead, all models settle on average on 90% of NR's total non-linear memory. This indicates that for the precessing case, *Surrogate*, *EOB* and *Phenom* differ mainly by modes included in $H_{\rm sub}$, however, to reproduce NR all approximants are evidently missing information beyond what is contained in $H_{\rm sub}$ as indicated by the gap between approximants and NR in the bottom plot of Figure 14. Additional discrepancies might manifest in the remnant velocity for precessing events. However, since the alignment procedure for precessing events is not effective as outlined in subsection III A, we cannot assess the kick ap-

[12] Physically, this results from the fact that the majority of energy flux away from the binary is transported by isotropic gravitational radiation. The azimuthally symmetric flux is sourced by $h_{2,\pm2}$. Its manifestation in terms of memory is carried by $h_{2,0}$.

proximation of the waveform models for precessing events in a meaningful way.

Leveraging the complete parameter space detailed in III, we examine whether the conclusions concerning the memory are influenced by the choice of NR events. By considering the supplementary parameter space beyond the *SXS* data, we verify that the results outlined earlier remain consistent for both precessing and non-precessing events. The relations between the memory contributions of *Surrogate*, *EOB*, and *Phenom* are validated for these instances, affirming that our observations are not affected by potential biases in the selection of extrinsic parameters.

Concluding our findings, we observe that neither the kick nor the memory in this form are able to assess a models subdominant waveform contribution. They are, however, excellent probes for the dominant $h_{2,\pm2}$ mode and for the mode content in general. The non-linear memory serves as a valuable tool to distinctly highlight the deficiencies in the mode content offered by *EOB* and *Phenom* models considered in this study. However, memory and kick as computed in this work indicate that the missing modes are not the only source of discrepancies of *EOB* and *Phenom* with respect to NR (or *Surrogate*). Our results imply that disparities emerge even at the level of subdominant modes common to all models. Specifically, we highlight the mixing of $h_{2,\pm1}$ with $h_{2,2}$ in the $\alpha$ coefficients relevant to the kick, which significantly contributes to the observed dissimilarities in kick magnitude and its direction as illustrated in Figure 10.

## VII. DISCUSSION

We conducted an extensive comparison of gravitational waveform models based on the calculation of the remnant's kick velocity and the gravitational memory for BBH mergers. Four state-of-the-art waveform models were considered (cf. Table I), each representing a distinct approach in efficiently generating a gravitational waveform for a given BBH configuration. We reviewed their main features, dissimilarities, and conventions. Moreover, we discussed issues in the commonly used waveform alignment procedures, pointing out difficulties in the meaningful distinction between alignment residuals and intrinsic waveform dissimilarities. After applying a well-tested alignment strategy, we utilized so-called balance laws to compute the kick velocity and the memory. In order to minimize a potential selection bias, we extended our analysis beyond the catalogued *SXS* BBH events and additionally considered randomly generated events in parameter space. We compared the computed kick velocities and memory terms with respect to reference models. In the case of catalogued events we chose NR waveforms as reference, while *Surrogate* was used for events not contained in the *SXS* database.

Our results demonstrate inconsistencies between the selected approximants and the reference model regarding both the computed kick velocity and inferred memory. This corroborates earlier findings in the literature [32]. We found that the *Surrogate* models considered here clearly outperform *EOB* and *Phenom* waveforms. It was further demonstrated that *EOB* and *Phenom* models do not accurately determine the gravitational wave memory due to the reduced subdominant mode content. This is true for both precessing and non-precessing events, and in good approximation independent of the alignment. For precessing events, the non-linear memory of all approximants, including *Surrogate*, display inconsistencies with respect to the *SXS* waveform memory as shown in Figure 14. On the other hand, compared to the gravitational wave memory, our results show that kick velocity is more strongly influenced by the dominant strain modes $h_{2,\pm2}$ and less sensitive to subdominant modes, both in amplitude and direction. Removing certain subdominant modes did not yield a significant deviation in the computed kick velocity vector.

We support our findings by analytical considerations based on strain mode mixing in the spherical harmonic decomposition of $(|\dot{h}|^2)_{\ell,m}$, as described by the $\alpha$-coefficients of equation (5.11). Notice that the way the modes mix is ultimately determined by the balance laws (5.8) and that with the help of the $\alpha$-coefficients one can express physical quantities of the BBH merger, such as radiated energy or remnant mass, kick velocity, and gravitational memory[13]. Another use of the $\alpha$-coefficients, which we discussed in VI, is to explain the relationship between the azimuthal symmetry of the energy flux away from the binary and the dominant memory contribution residing in the (azimuthally independent) $h_{2,0}$ mode.

Generally, our findings indicate that while the kick velocity as defined in section V is an excellent estimator of the quality of $h_{2,\pm2}$ mode and highly sensitive to the alignment procedure, the GW memory serves as a good indicator for differences in the mode content of the waveforms, both for precessing and non-precessing events. However, towards individual subdominant modes, or sets of those, we find that quantities like kick and memory in their original form exhibit very low sensitivity.

In fact, one key conclusion which can be drawn from our analysis is that kick velocity and gravitational memory are not suitable as high-precision metrics for modes with $\ell > 2$. In light of this realization, we propose a new three-step strategy for future waveform comparisons: Firstly, for each approximant, waveforms are generated and set on equal footing regarding the spatial and temporal frame (for instance via the *LALSuite*). After the

[13] When calibrating our numerical pipeline, we test all computations of physical quantities against the metadata contained in the *SXS* database. For instance, the kick velocity magnitudes computed from (5.14) were compared against the values in the *SXS* data. The substantial agreement found in these comparisons validate the applicability of the balance laws.

application of a suitable alignment scheme, such as for instance the one applied in this article, residual ambiguities in the waveforms are resolved. For each approximant, the waveforms and their (provided) mode decomposition are now ready for the computation of physical quantities. Thus, we compute the $\alpha$-coefficients using equation (5.11) in combination with the aligned waveforms.

The first step is repeated for different strain mode contents, i.e., in each iteration, the set of strain modes included in the computation of the $\alpha$-coefficients changes depending on which section of the mode content shall be tested. For instance, being interested in the overall impact of $\ell = 3, 4$ modes on the precision of the waveform, the first step yields a set $\{\alpha^{\text{approx.}}_{\ell,m}(h_{\ell,m}), \alpha^{\text{approx.}}_{\ell,m}(h_{\ell,m}|\ell \neq 3), \alpha^{\text{approx.}}_{\ell,m}(h_{\ell,m}|\ell \neq 4)\}$ for each approximant[14].

In the second step, we use $\alpha^{\text{approx.}}_{\ell,m}(h_{\ell,m})$ to compute the remnant velocity vector. Using its overall amplitude and direction, we test and compare the waveforms' dominant $h_{2,\pm2}$ and subdominant $h_{2,\pm1}$ mode under the assumption that an preceding alignment procedure minimized the mismatch and only intrinsic dissimilarities between waveforms stemming form different approximates remain[15]. The kick tests both of the latter modes simultaneously as it is dominated by the strain mode $h_{2,\pm1}h_{2,\pm2}$. For a proper distinction of the origin of dissimilarities, the $h_{2,2}$-mismatch and kick can be used in synergy.

Provided that the dominant $h_{2,\pm2}$ mode is modeled sufficiently accurate, the third step of the proposed strategy comprises a comparison of the computed $\alpha$'s between approximants. Here, we one can proceed twofold, i) the $\alpha$-coefficients are compared in accumulated form as the non-linear memory, or ii) $\alpha$-coefficients are compared individually. In the gravitational wave memory, the effect of individual subdominant strain modes accumulates and, thus, can be quantified despite the mode in question being highly subdominant. Concretely, the memory contributions computed using the different sets of $\alpha$'s within an approximant are compared with each other to gauge the impact of individual modes for a given approximant. The corresponding errors are then compared to the same quantities computed for *NR* waveforms. Shortcomings in certain modes are exposed if excluding these yields a larger impact for the *NR* than for the approximant's memory contribution. The shortcomings can be addressed by phenomenological corrections in the waveform or individual strain modes.

In case of a successful implementation of i), we can further improve precision of the analysis and proceed with ii): Each $\alpha_{\ell,m}$ contains different subdominant modes coupling to $h_{2,\pm2}$ (or the next-to-dominant mode, e.g, $h_{2,\pm1}$) which can be tested by simply comparing the corresponding $\alpha$-coefficients of *NR* and the approximants against each other. Here, the previous removal of subdominant modes isolates the impact of a given mode in question within a specific $\alpha_{\ell,m}$. For instance, removing all strain modes with $m = 0$ in the computation of the $\alpha$-coefficients will isolate the $h_{2,\pm2}h_{4,\pm4}$ mixing in $\alpha_{2,2}$ as the only dominant mode mixing. Hence, with the previous steps executed, comparing the so computed $\alpha_{2,2}$ between *NR* and the approximants will indicate potential discrepancies in the $h_{4,\pm4}$ mode. Note, however, that only strain modes coupling directly to $h_{2,\pm2}$ in the $\alpha$-coefficients can be tested in this way. For modes that only couple to subdominant modes in the presence of another dominant-mode-coupling, the metrics of comparison computed according to the above strategy are statistically not significant when considering an ensemble of events. On the other hand, the advantage of ii) over i) resides in a more precise mode testing. That is, while the memory influenced by many subdominant modes simultaneously, its individual components depend only a small number of couplings to $h_{2,\pm2}$. Testing an individual mode using the memory, the accumulation of mode impact in the sum over $\alpha$'s might pick up discrepancies in other subdominant modes coupling to the one in question.

The third step of the prescribed strategy targets predominantly the mode content with $\ell > 2$. As pointed out in section VI, these modes exhibit particularly large mismatch or are trivial for *EOB* and *Phenom*. Thus, the latter step is especially relevant for updated models of *EOB* and *Phenom* (e.g. [27, 59] and future *Phenom* models) that contain more modes than the versions tested in this work.

Physically speaking, testing individual $\alpha$-coefficients corresponds to analysing the anisotropic non-linear memory, or equivalently, the anisotropic energy flux radiated to infinity from the binary system. Naturally, the anisotropic energy flux is highly configuration-dependent and subdominant. It nevertheless constitutes a powerful tool in high-precision tests of subdominant waveform contributions, especially for precessing events.

The advantage of assessing a waveform's quality by systematically deleting subdominant modes from the computation of physical quantities, as opposed to a mode-by-mode mismatch analysis as the one depicted in Figure 9, is that no fit to the corresponding NR counterpart is needed. Such a fitting procedure would defeat the purpose of the waveform approximants, since they are supposed to produce accurate waveforms in a more time-efficient manner than NR.

We leave an implementation of the outlined strategy for future work. Here, we have established the groundwork for these forthcoming systematic studies by developing codes, methods, and guidelines that are applicable to the important task of comparing and assessing the accuracy and reliability of waveform models.

---

[14] We can even be more precise here and additionally specify the $m$ of an individual mode that we wish to analyze. Note, however, that it is crucial to remove both $h_{\ell,\pm m}$ in this case.

[15] This step reproduces previous investigations such as for instance [32]

## ACKNOWLEDGEMENTS

LH would like to acknowledge financial support from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme grant agreement No 801781. LH further acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2181/1 - 390900948 (the Heidelberg STRUCTURES Excellence Cluster). The authors thank the Heidelberg STRUCTURES Excellence Cluster for financial support. The authors acknowledge support by the state of Baden-Württemberg, Germany, through bwHPC. The authors thank Cecilio Garcia-Quiros, Keefe Mitman and Nils Deppe for fruitful discussions.


## Appendix A: Numerical artifacts

Numerical artifacts can significantly influence the outcome of the mismatch function discussed in section III. The morphology with which artefacts can appear varies. One of the most prominent examples is displayed in Figure 15. Here, the waveform mode carries a tail, which shall not be confused with the gravitational memory. The latter is not contained in the waveforms displayed. Instead, the non-trivial extent of the waveform after the ring down results from numerical artifacts and is to be disregarded as non-physical. Since the effect is rather small and effects only subdominant modes, we ignored it in the main analysis. In future works tackling the precision of higher strain modes, this issue has to be addressed.

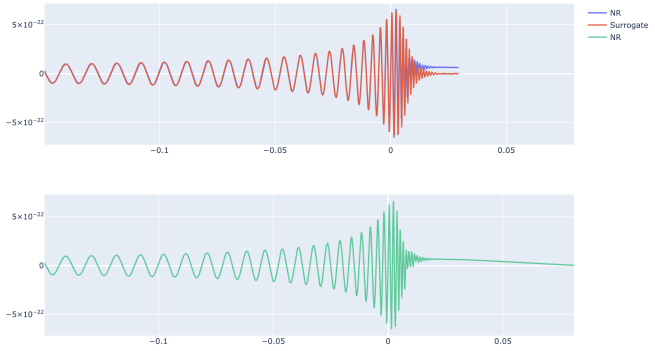



FIG. 15. $(4,4)$ waveform mode for a non-precessing event SXS:BBH:0155. Displayed are the truncated and aligned waveforms (above) for NR and *Surrogate* as well as the "raw" full waveform for NR.


[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).

[2] M. Colpi, K. Danzmann, M. Hewitson, K. Holley-Bockelmann, P. Jetzer, G. Nelemans, A. Petiteau, D. Shoemaker, C. Sopuerta, R. Stebbins, N. Tanvir, H. Ward, W. J. Weber, I. Thorpe, A. Daurskikh, A. Deep, I. F. Núñez, C. G. Marirrodriga, M. Gehler, J.-P. Halain, O. Jennrich, U. Lammers, J. Larrañaga, M. Lieser, N. Lützgendorf, W. Martens, L. Mondin, A. P. Niño, P. Amaro-Seoane, M. A. Sedda, P. Auclair, S. Babak, Q. Baghi, V. Baibhav, T. Baker, J.-B. Bayle, C. Berry, E. Berti, G. Boileau, M. Bonetti, R. Brito, R. Buscicchio, G. Calcagni, P. R. Capelo, C. Caprini, A. Caputo, E. Castelli, H.-Y. Chen, X. Chen, A. Chua, G. Davies, A. Derdzinski, V. F. Domcke, D. Doneva, I. Dvorkin, J. M. Ezquiaga, J. Gair, Z. Haiman, I. Harry, O. Hartwig, A. Hees, A. Heffernan, S. Husa, D. Izquierdo, N. Karnesis, A. Klein, V. Korol, N. Korsakova, T. Kupfer, D. Laghi, A. Lamberts, S. Larson, M. L. Jeune, M. Lewicki, T. Littenberg, E. Madge, A. Mangiagli, S. Marsat, I. M. Vilchez, A. Maselli, J. Mathews, M. van de Meent, M. Muratore, G. Nardini, P. Pani, M. Peloso, M. Pieroni, A. Pound, H. Quelquejay-Leclere, A. Ricciardone, E. M. Rossi, A. Sartirana, E. Savalle, L. Sberna, A. Sesana, D. Shoemaker, J. Slutsky, T. Sotiriou, L. Speri, M. Staab, D. Steer, N. Tamanini, G. Tasinato, J. Torrado, A. Torres-Orjuela, A. Toubiana, M. Vallisneri, A. Vecchio, M. Volonteri, K. Yagi, and L. Zwick, LISA Definition Study Report (2024), arXiv:2402.07571 [astro-ph, physics:gr-qc].

[3] M. Maggiore, C. V. D. Broeck, N. Bartolo, E. Belgacem, D. Bertacca, M. A. Bizouard, M. Branchesi, S. Clesse, S. Foffa, J. García-Bellido, S. Grimm, J. Harms, T. Hinderer, S. Matarrese, C. Palomba, M. Peloso, A. Ricciardone, and M. Sakellariadou, Journal of Cosmology and Astroparticle Physics **2020** (03), 050.

[4] M. Evans, A. Corsi, C. Afle, A. Ananyeva, K. G. Arun, S. Ballmer, A. Bandopadhyay, L. Barsotti, M. Baryakhtar, E. Berger, E. Berti, S. Biscoveanu, S. Borhanian, F. Broekgaarden, D. A. Brown, C. Cahillane, L. Campbell, H.-Y. Chen, K. J. Daniel, A. Dhani, J. C. Driggers, A. Effler, R. Eisenstein, S. Fairhurst, J. Feicht, P. Fritschel, P. Fulda, I. Gupta, E. D. Hall, G. Hammond, O. A. Hannuksela, H. Hansen, C.-J. Haster, K. Kacanja, B. Kamai, R. Kashyap, J. S. Key, S. Khadkikar, A. Kontos, K. Kuns, M. Landry, P. Landry, B. Lantz, T. G. F. Li, G. Lovelace, V. Mandic, G. L. Mansell, D. Martynov, L. McCuller, A. L. Miller, A. H. Nitz, B. J. Owen, C. Palomba, J. Read, H. Phurailatpam, S. Reddy, J. Richardson, J. Rollins, J. D. Romano, B. S. Sathyaprakash, R. Schofield, D. H. Shoemaker, D. Sigg, D. Singh, B. Slagmolen, P. Sledge, J. Smith, M. Soares-Santos, A. Strunk, L. Sun, D. Tanner, L. A. C. van Son, S. Vitale, B. Willke, H. Yamamoto, and M. Zucker, Cosmic Explorer: A Submission to the NSF MPSAC ngGW Subcommittee (2023), arXiv:2306.13745 [astro-ph].

[5] Ya. B. Zel'dovich and A. G. Polnarev, Soviet Astronomy **18**, 17 (1974).
[6] D. Christodoulou, Physical Review Letters **67**, 1486 (1991).
[7] K. S. Thorne, Physical Review D **45**, 520 (1992).
[8] A. G. Wiseman and C. M. Will, Phys. Rev. D **44**, R2945 (1991).
[9] M. Favata, Phys. Rev. D **80**, 024002 (2009).
[10] M. Favata, The Astrophysical Journal **696**, L159 (2009).
[11] L. Heisenberg, Phys. Rept. **796**, 1 (2019), arXiv:1807.01725 [gr-qc].
[12] P. Amaro-Seoane *et al.*, Living Reviews in Relativity **26**, 10.1007/s41114-022-00041-y (2023).
[13] J. Aasi *et al.*, Classical and Quantum Gravity **32**, 074001 (2015).
[14] F. Acernese *et al.*, Classical and Quantum Gravity **32**, 024001 (2014).
[15] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi, Phys. Rev. D **96**, 024058 (2017), arXiv:1705.07089 [gr-qc].
[16] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, D. A. Hemberger, P. Schmidt, and R. Smith, Phys. Rev. D **95**, 104023 (2017), arXiv:1701.00550 [gr-qc].
[17] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, Phys. Rev. Res. **1**, 033015 (2019).
[18] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, Phys. Rev. D **93**, 044006 (2016).
[19] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Phys. Rev. D **93**, 044007 (2016).
[20] G. Pratten, C. García-Quirós, M. Colleoni, A. Ramos-Buades, H. Estellés, M. Mateu-Lucena, R. Jaume, M. Haney, D. Keitel, J. E. Thompson, and S. Husa, Phys. Rev. D **103**, 104056 (2021).
[21] H. Estellés, M. Colleoni, C. García-Quirós, S. Husa, D. Keitel, M. Mateu-Lucena, M. d. L. Planas, and A. Ramos-Buades, Phys. Rev. D **105**, 084040 (2022).
[22] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel, Physical Review D **86**, 10.1103/physrevd.86.024011 (2012).
[23] B. D. Lackey, K. Kyutoku, M. Shibata, P. R. Brady, and J. L. Friedman, Physical Review D **89**, 10.1103/physrevd.89.043009 (2014).
[24] Y. Pan, A. Buonanno, A. Taracchini, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, Phys. Rev. D **89**, 084006 (2014).
[25] A. Bohé , L. Shao, A. Taracchini, A. Buonanno, S. Babak, I. W. Harry, I. Hinder, S. Ossokine, M. Pürrer, V. Raymond, T. Chu, H. Fong, P. Kumar, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, M. A. Scheel, and B. Szilágyi, Physical Review D **95**, 10.1103/physrevd.95.044028 (2017).
[26] S. Ossokine, A. Buonanno, S. Marsat, R. Cotesta, S. Babak, T. Dietrich, R. Haas, I. Hinder, H. P. Pfeiffer, M. Pürrer, C. J. Woodford, M. Boyle, L. E. Kidder, M. A. Scheel, and B. Szilágyi, Phys. Rev. D **102**, 044055 (2020).
[27] A. Ramos-Buades, A. Buonanno, H. Estellés, M. Khalil, D. P. Mihaylov, S. Ossokine, L. Pompili, and M. Shiferaw, SEOBNRv5PHM: Next generation of accurate and efficient multipolar precessing-spin effective-one-body waveforms for binary black holes (2023), arXiv:2303.18046 [gr-qc].
[28] A. Ashtekar, T. De Lorenzo, and N. Khera, Gen. Rel. Grav. **52**, 107 (2020), arXiv:1906.00913 [gr-qc].
[29] N. Khera, B. Krishnan, A. Ashtekar, and T. De Lorenzo, Phys. Rev. D **103**, 044012 (2021), arXiv:2009.06351 [gr-qc].
[30] L. Heisenberg, Phil. Trans. Roy. Soc. Lond. A **382**, 20230086 (2023), arXiv:2309.12505 [gr-qc].
[31] A. Borchers and F. Ohme, in *55th Rencontres de Moriond on Gravitation* (2021) arXiv:2106.02414 [gr-qc].
[32] A. Borchers and F. Ohme, Classical and Quantum Gravity **40**, 095008 (2023).
[33] K. Mitman *et al.*, Phys. Rev. D **103**, 024031 (2021), arXiv:2011.01309 [gr-qc].
[34] L. C. W. W. Group *et al.*, Waveform Modelling for the Laser Interferometer Space Antenna (2023), arxiv:2311.01300 [astro-ph, physics:gr-qc].
[35] LIGO Scientific Collaboration, Astrophysics Source Code Library , ascl:2012.021 (2020).
[36] P. Schmidt, I. W. Harry, and H. P. Pfeiffer, (2017), arXiv:1703.01076 [gr-qc].
[37] M. Boyle *et al.*, Classical and Quantum Gravity **36**, 195006 (2019), arXiv:1904.04831 [astro-ph, physics:gr-qc].
[38] L. E. Kidder, M. A. Scheel, S. A. Teukolsky, E. D. Carlson, and G. B. Cook, Physical Review D **62**, 084032 (2000), publisher: American Physical Society.
[39] K. Mitman, J. Moxon, M. A. Scheel, S. A. Teukolsky, M. Boyle, N. Deppe, L. E. Kidder, and W. Throwe, Physical Review D **102**, 104007 (2020), arxiv:2007.11562 [gr-qc].
[40] N. W. Taylor, M. Boyle, C. Reisswig, M. A. Scheel, T. Chu, L. E. Kidder, and B. Szilágyi, Physical Review D **88**, 124010 (2013).
[41] T. Damour, in *General Relativity, Cosmology and Astrophysics* (Springer International Publishing, 2014) pp. 111–145.
[42] A. Buonanno, Y. Chen, and T. Damour, Physical Review D **74**, 10.1103/physrevd.74.104005 (2006).
[43] H. Estellés, S. Husa, M. Colleoni, D. Keitel, M. Mateu-Lucena, C. García-Quirós, A. Ramos-Buades, and A. Borchers, Phys. Rev. D **105**, 084039 (2022).
[44] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Phys. Rev. Lett. **113**, 151101 (2014).
[45] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, Physical Review D **99**, 10.1103/physrevd.99.064045 (2019).
[46] J. Yoo, K. Mitman, V. Varma, M. Boyle, S. E. Field, N. Deppe, F. Hébert, L. E. Kidder, J. Moxon, H. P. Pfeiffer, M. A. Scheel, L. C. Stein, S. A. Teukolsky, W. Throwe, and N. L. Vu, Numerical relativity surrogate model with memory effects and post-Newtonian hybridization (2023), arxiv:2306.03148 [gr-qc].
[47] G. Pratten *et al.*, Phys. Rev. D **103**, 104056 (2021), arXiv:2004.06503 [gr-qc].
[48] D'Ambrosio, Fabio and Fell, Shaun D. B. and Heisenberg, Lavinia and Maibach, David and Zentarra, Stefan and Zosso, Jann, (2022), arXiv:2201.11634 [gr-qc].
[49] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **269**, 21 (1962), 2414436.

[50] R. K. . Sachs and H. Bondi, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **270**, 103 (1962).
[51] R. Penrose, Physical Review Letters **10**, 66 (1963).
[52] R. Geroch, in *Asymptotic Structure of Space-Time*, edited by F. P. Esposito and L. Witten (Springer US, Boston, MA, 1977) pp. 1–105.
[53] A. Ashtekar, arXiv:1409.1800 [gr-qc, physics:hep-th, physics:math-ph] (2015), arxiv:1409.1800 [gr-qc, physics:hep-th, physics:math-ph].
[54] A. Ashtekar and M. Streubel, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **376**, 585 (1981).
[55] R. M. Wald and A. Zoupas, Physical Review D **61**, 084027 (2000), arxiv:gr-qc/9911095.
[56] V. B. Braginsky and K. S. Thorne, Nature **327**, 123 (1987).
[57] K. Mitman *et al.*, Phys. Rev. D **106**, 084029 (2022), arXiv:2208.04356 [gr-qc].
[58] C. Pitte, Q. Baghi, S. Marsat, M. Besançon, and A. Petiteau, Physical Review D **108**, 044053 (2023), publisher: American Physical Society.
[59] L. Pompili *et al.*, Laying the foundation of the effective-one-body waveform models seobnrv5: improved accuracy and efficiency for spinning non-precessing binary black holes (2023), arXiv:2303.18039 [gr-qc].

Prepared for submission to JHEP

# Trans-series from condensates

**Marcos Mariño**[a] **and Ramon Miravitllas**[b]

[a] *Département de Physique Théorique et Section de Mathématiques*
*Université de Genève, Genève, CH-1211 Switzerland*

[b] *HUN-REN Wigner Research Centre for Physics*
*Konkoly-Thege Miklós u. 29-33, 1121 Budapest, Hungary*

*E-mail:* Marcos.Marino@unige.ch, ramon.miravitllas.mas@wigner.hu

Abstract: The Shifman–Vainshtein–Zakharov (SVZ) sum rules provide a method to obtain trans-series expansions in many quantum field theories, in which exponentially small corrections are calculated by combining the operator product expansion with the assumption of vacuum condensates. In some solvable models, exact expressions for trans-series can be obtained from non-perturbative results, and this makes it possible to test the SVZ method by comparing its predictions to these exact trans-series. In this paper we perform such a precision test in the example of the fermion self-energy in the Gross–Neveu model. Its exact trans-series expansion can be extracted from the large $N$ solution, at the first non-trivial order in $1/N$. It is given by an infinite series of exponentially small corrections involving factorially divergent power series in the 't Hooft parameter. We show that the first two corrections are associated to two-quark and four-quark condensates, and we reproduce the corresponding power series exactly, and at all loops, by using the SVZ method. In addition, the numerical values of the condensates can be extracted from the exact result, up to order $1/N$.

arXiv:2402.19356v1 [hep-th] 29 Feb 2024

## Contents


**1 Introduction** **1**

**2 The Gross–Neveu model** **3**

**3 Trans-series from the $1/N$ expansion** **8**
3.1 An exact result for the self-energy 8
3.2 Trans-series expression 10

**4 Trans-series from condensates** **15**
4.1 Perturbative series 15
4.2 General aspects of perturbation theory with condensates 17
4.3 Trans-series for the two-quark condensate 20
4.4 Trans-series for the four-quark condensate 26

**5 Conclusions and open problems** **34**

**A Bubbles** **36**
A.1 The fermion polarization loop 36
A.2 Summing over bubbles 37


## 1 Introduction

In quantum field theory (QFT), perturbative series give the asymptotic expansion of observables at small coupling. There are many indications that this expansion can be upgraded to a *trans-series*, i.e. a generalization of the perturbative series which includes exponentially small corrections, in such a way that the exact value of the observable can be obtained by an appropriate resummation.

Examples in QFT where we know explicitly the detailed structure of the trans-series are scarce. In some two-dimensional asymptotically free theories, one can compute certain observables exactly, and then show that this exact expression can be re-expressed or "decoded" as a resummed trans-series. A beautiful realization of this idea was achieved in [1], building on previous work [2, 3], in the case of the two-point function of the non-linear sigma model at next-to-leading order in the $1/N$ expansion[1]. More recently, the free energy of integrable models coupled to a conserved charge was decoded as a resummed trans-series, even at finite $N$ [4–7].

In generic QFTs, in which no analytic answer is known for the observables, physicists have devised two ways of upgrading the perturbative series to a trans-series. The first one is to add instanton corrections, coming from non-trivial saddle-points of the path integral. Instanton

[1]The $1/N$ expansion, when used appropriately, gives a power series in $1/N$ where each term is an exact, non-perturbative function of the renormalized 't Hooft coupling, and not just a formal power series thereof. Sometimes we will use the expression "exact results at large $N$," or similar ones, to refer to this type of non-perturbative functions.

calculus is plagued with problems and it is fair to say that it is of limited use, except in supersymmetric theories or in very simple models. The second method to obtain trans-series can be applied to correlation functions in very general QFTs. It combines Wilson's operator product expansion (OPE) with some assumptions on the vacuum structure of the theory[2]. More precisely, this method assumes the existence of *condensates*, or non-zero vacuum expectation values (vevs), for the operators appearing in the OPE. It was used by Politzer in [10] to calculate the quark propagator in QCD beyond perturbation theory, and then extended and systematized in the famous QCD sum rules of Shifman, Vainshtein and Zakharov (SVZ) in [11, 12]

An obvious question is whether the method of OPE and vacuum condensates provides the correct trans-series representation of correlation functions. Although this might be obvious to most practitioners in the field, there are various reasons for a detailed inquiry. For example, it could be the case that the OPE provides only an approximate parametrization of non-perturbative corrections, rather than the real thing. There has also been some debate concerning which form of the OPE should be used to calculate trans-series. Most of the sum rule calculations done by physicists are based on a simplified or "practical" version of the OPE, in which the Wilson coefficients are calculated perturbatively, while the vevs of the operators contain the non-perturbative information, but it has been suggested [13, 14] that one might need more complicated procedures. For us, another important motivation to revisit these questions was the discovery, in some integrable models, of fractional power corrections which can not be easily accommodated in the current OPE/renormalon paradigm [4].

It is a good idea to ask foundational questions in simpler, solvable models where they can be answered with precision. In the case of the SVZ method this has been done in various papers, starting in the 1980s [1–3, 13–15]. These works usually focused on the two-dimensional non-linear sigma model at large $N$, and they extracted trans-series from exact results at leading and next-to-leading order in the $1/N$ expansion, as we mentioned above. However, and somewhat surprisingly, many of these studies essentially *assumed* that the trans-series obtained from the exact result could be obtained from an OPE calculation with condensates, but did not verify it explicitly. Some direct tests of the SVZ method were performed in [13–15] but in rather simple situations. In particular, the non-trivial perturbative series attached to the power corrections of [1] in the non-linear sigma model were never reproduced by a direct OPE calculation.

The goal of this paper is to provide such a direct comparison between an exact trans-series and a standard calculation of power corrections. The original example of the two-dimensional non-linear sigma model is not the simplest one to perform such a comparison, and we focus instead on a fermionic cousin, the Gross–Neveu (GN) model [16], where trans-series are of comparable complexity. The self-energy of elementary fermions (or "quarks") in this model can be calculated at the first non-trivial order in the $1/N$ expansion, as an exact function of the external momentum and the mass gap [17]. By using the Mellin transform techniques of [1] one can obtain an explicit trans-series representation of this function, involving an infinite series of power corrections. Schematically, we have

$$\Sigma(p) = \not{p}\Sigma_0(\lambda) + \Lambda\Sigma_1(\lambda) + \not{p}\frac{\Lambda^2}{p^2}\Sigma_2(\lambda) + \cdots , \tag{1.1}$$

where $\Lambda$ is the dynamically generated scale and $\lambda$ is the 't Hooft coupling. The series $\Sigma_0(\lambda)$ is the perturbative series, but each power correction involves a factorially divergent series $\Sigma_n(\lambda)$,

[2]The fact that the method of OPE with condensates leads to trans-series, in the sense of the theory of resurgence, was pointed out some time ago in [8]. See also [9] for related observations.

$n = 1, 2, \cdots$. If the OPE picture is correct, one should be able to reproduce these series by doing perturbation theory in the background of the appropriate vacuum condensates. This is precisely what we verify with complete success, and at *all* loops, for the first two power corrections, which are associated to the two-quark condensate and to the four-quark condensate (these are the terms with $n = 1, 2$ in the equation above, respectively). Once this is done, the values of the condensates at next-to-leading order (NLO) in $1/N$, which are unknown parameters in the sum rules, can be extracted from the large $N$ result.

Our calculation provides a precise and direct test of the SVZ method in the Gross–Neveu model, at the first non-trivial order in $1/N$ expansion, and it illustrates various conceptual and practical issues of the method. For example, it shows that the four-quark condensate is not ambiguous at leading order in the $1/N$ expansion, due to factorization, but it is indeed ambiguous at subleading order, as expected from the results of [2, 3]. Our calculation is done with the "practical" version of the OPE, which leads to the correct result in this example. An amusing spin-off result of this work is a diagrammatic derivation of the beta function of the model, at next-to-leading order in the $1/N$, which seems to be simpler than the approach usually followed in the literature [18, 19].

We should mention that the paper [20] already compared directly an OPE calculation with condensates to an exact $1/N$ result in the Gross–Neveu model, but in a substantially simpler situation in which the power series associated to the condensate has only one term.

This paper is structured as follows. In section 2 we review or derive various basic results for the GN model which will be needed for the paper. In section 3 we obtain the trans-series expression for the exact two-point function at the first non-trivial order in the $1/N$ expansion. In section 4 we calculate the two-point function in perturbation theory with condensates, where we include the first two power corrections, and we reproduce exactly the trans-series derived from the $1/N$ expansion. Finally, in 5 we conclude and present some questions and open problems. The Appendix collects some diagrammatic tools. It summarizes an important technique to calculate all-loop results in the $1/N$ expansion, due to Palanques-Mestre and Pascual [18, 21], which is used throughout the paper.

## 2 The Gross–Neveu model

The GN model is a two dimensional QFT, involving an $N$-uple of Dirac fermions with a quartic interaction, which was introduced in [16] as a toy model for various important physical phenomena. First of all, the GN model is asymptotically free. It can be solved in the large $N$ limit, where it can be shown that quantum effects lead to spontaneous symmetry breaking of a discrete $\mathbb{Z}_2$ symmetry and the formation of a bilinear fermion condensate. For these reasons, the GN model can be seen as a toy model for the quark sector of QCD. In addition, the model is integrable at the quantum level, its exact $S$-matrix has been conjectured in [22], and its spectrum is extremely rich. In this section we will review some aspects of the model which we will need in our precision test of the SVZ method.

We will work in Minkowski space, and our choice of Dirac algebra in two dimensions is:

$$\gamma^0 = \sigma^2, \qquad \gamma^1 = \mathrm{i}\sigma^1, \qquad \gamma^5 = \sigma^3. \tag{2.1}$$

The Lagrangian density describing the theory is

$$\mathcal{L} = \mathrm{i}\overline{\psi} \cdot \not{\partial}\psi + \frac{g_0}{2}\left(\overline{\psi} \cdot \psi\right)^2, \tag{2.2}$$

where $\boldsymbol{\psi} = (\psi_1, \cdots, \psi_N)$ is an $N$-uple of Dirac fermions. The model has a continuous $U(N)$ global symmetry, and a $\mathbb{Z}_2$ discrete symmetry

$$\boldsymbol{\psi} \to \gamma_5 \boldsymbol{\psi}. \tag{2.3}$$

In order to keep track of large $N$ counting, it is extremely useful to introduce an auxiliary scalar field $\sigma$ and write the GN Lagrangian as

$$\mathcal{L}_\sigma = \mathrm{i}\overline{\boldsymbol{\psi}} \cdot \partial\!\!\!/ \boldsymbol{\psi} - \frac{1}{2}\sigma^2 + \sqrt{g_0}\sigma\overline{\boldsymbol{\psi}} \cdot \boldsymbol{\psi}. \tag{2.4}$$

The original Lagrangian is obtained by integrating out $\sigma$. The symmetry (2.3) reads now

$$\boldsymbol{\psi} \to \gamma_5 \boldsymbol{\psi}, \qquad \sigma \to -\sigma. \tag{2.5}$$

One can also add a bare mass term for the fermions of the form

$$\mathcal{L}_m = -m_0 \overline{\boldsymbol{\psi}} \boldsymbol{\psi}, \tag{2.6}$$

although we will consider the massless theory with $m_0 = 0$ (as is well-known [16], even when $m_0 = 0$ a dynamical mass is generated at the quantum level and can be calculated in the $1/N$ expansion). From now one we will work with the Lagrangian (2.4). To write down the Feynman diagrams, we represent fermions by continuous lines and sigma particles by dashed lines. The fermion propagator in momentum space is given by

$$S_0(p)_{ij}^{\alpha\beta} = \left(\frac{\mathrm{i}}{p\!\!\!/ - m_0}\right)^{\alpha\beta} \delta_{ij}. \tag{2.7}$$

(Latin sub-indices are $U(N)$ indices, while Greek super-indices are spinor indices.) The propagator of the $\sigma$ field is $-\mathrm{i}$, and there is a single interaction vertex $\mathrm{i}\sqrt{g_0}$.

The GN model is renormalizable and asymptotically free. We will almost always adopt the $\overline{\mathrm{MS}}$ scheme and mostly work with bare fields, which we will simply denote by $\boldsymbol{\psi}$. Renormalized fields will be denoted by $\boldsymbol{\psi}_R$. The renormalization constants are defined as usual by

$$\boldsymbol{\psi} = Z_\psi^{1/2} \boldsymbol{\psi}_R, \qquad g_0 = (\nu^2)^{\epsilon/2} Z_g g, \qquad m_0 = Z_m m. \tag{2.8}$$

Our convention for $\epsilon$ is

$$d = 2 - \epsilon \tag{2.9}$$

and

$$\nu^2 = \mu^2 \mathrm{e}^{\gamma_E - \log(4\pi)}, \tag{2.10}$$

where $d$ is the number of space-time dimensions in dimensional regularization and $\mu$ is the scale parameter. The beta function is defined as

$$\beta_g(g;\epsilon) = -\frac{\epsilon g}{1 + g\partial_g \log Z_g} = -\epsilon g + \beta(g), \tag{2.11}$$

while

$$\gamma(g) = \beta(g;\epsilon)\frac{\partial \log Z_\psi}{\partial g}, \qquad \gamma_m(g) = \beta(g;\epsilon)\frac{\partial \log Z_m}{\partial g} \tag{2.12}$$

are the anomalous dimension of the field and the mass, respectively. The renormalization functions are known to four loops in conventional perturbation theory, see [23] for recent results and

references to the literature. However, we will work in the $1/N$ expansion. We define the 't Hooft parameter

$$\lambda = \frac{gN}{\pi} \tag{2.13}$$

whose beta function is

$$\beta_\lambda(\lambda;\epsilon) = \frac{N}{\pi}\beta(g;\epsilon) = -\epsilon\lambda + \beta_\lambda(\lambda). \tag{2.14}$$

The function $\beta_\lambda(\lambda)$ has a $1/N$ expansion at fixed 't Hooft coupling given by

$$\beta_\lambda(\lambda) = \sum_{j\geq 0} \beta_\lambda^{(j)}(\lambda) N^{-j} \tag{2.15}$$

and similarly for $\beta_\lambda(\lambda;\epsilon)$. We note that

$$\beta_\lambda^{(0)}(\lambda;\epsilon) = -\epsilon\lambda - \lambda^2, \qquad \beta_\lambda^{(j)}(\lambda;\epsilon) = \beta_\lambda^{(j)}(\lambda), \quad j \geq 1. \tag{2.16}$$

The first correction $\beta_\lambda^{(1)}(\lambda)$ is known in closed form [24] and is given by

$$\beta_\lambda^{(1)}(\lambda) = \lambda^2\left(1 + \int_0^\lambda \frac{\Gamma(2+u)}{(2+u)\Gamma^3\left(1+\frac{u}{2}\right)\Gamma\left(1-\frac{u}{2}\right)}\mathrm{d}u\right). \tag{2.17}$$

We have similar results for the anomalous dimensions. The mass anomalous dimension has a $1/N$ expansion of the form

$$\gamma_m(\lambda) = \sum_{j\geq 0} \gamma_m^{(j)}(\lambda) N^{-j}, \tag{2.18}$$

where

$$\gamma_m^{(0)}(\lambda) = \lambda, \tag{2.19}$$

and the first non-trivial correction is given by [25]

$$\gamma_m^{(1)}(\lambda) = \chi(\lambda) - \frac{\beta_\lambda^{(1)}(\lambda)}{\lambda} \tag{2.20}$$

where

$$\chi(\lambda) = \frac{\lambda\Gamma(2+\lambda)}{(2+\lambda)\Gamma^3\left(1+\frac{\lambda}{2}\right)\Gamma\left(1-\frac{\lambda}{2}\right)}. \tag{2.21}$$

Finally, the field anomalous dimension has the $1/N$ expansion

$$\gamma(\lambda) = \sum_{j\geq 1} \gamma^{(j)}(\lambda) N^{-j}, \tag{2.22}$$

where

$$\gamma^{(1)}(\lambda) = \frac{\lambda^2}{2}\frac{1}{2+\lambda}\frac{\Gamma(1+\lambda)}{\Gamma^3\left(1+\frac{\lambda}{2}\right)\Gamma\left(1-\frac{\lambda}{2}\right)}. \tag{2.23}$$

As we will see, the functions (2.23), (2.21) and (2.17) will be obtained as spin-offs of our trans-series calculation for the perturbative, the two-quark condensate, and the four-quark condensate sectors, respectively. The renormalization constants in (2.8) can be recovered from the renormalization functions. We have, for the coupling constant,

$$Z_g = \exp\left[-\int_0^g \frac{\mathrm{d}u}{u}\frac{\beta(u)}{\beta(u;\epsilon)}\right], \tag{2.24}$$

while, for the field and mass renormalization, one finds

$$Z_\psi = \exp\left[\int_0^g \mathrm{d}u \frac{\gamma(u)}{\beta(u;\epsilon)}\right], \qquad Z_m = \exp\left[\int_0^g \mathrm{d}u \frac{\gamma_m(u)}{\beta(u;\epsilon)}\right]. \tag{2.25}$$

The renormalization constants can also be obtained in a $1/N$ expansion by simply re-expressing everything in terms of the 't Hooft coupling. In particular, this coupling renormalizes as

$$\lambda_0 = (\nu^2)^{\epsilon/2} Z_\lambda \lambda, \tag{2.26}$$

where $Z_\lambda$ is the renormalization constant $Z_g$ expressed in terms of $\lambda$ and organized in a $1/N$ expansion. It is given explicitly as

$$Z_\lambda = \exp\left[-\int_0^\lambda \frac{\mathrm{d}u}{u} \frac{\beta_\lambda(u)}{\beta_\lambda(u;\epsilon)}\right]. \tag{2.27}$$

We will also need to renormalize the composite operators appearing in the OPE. We will denote renormalized composite operators by a bracket, $[\mathcal{O}]$. The renormalization constants are defined by

$$\mathcal{O}_i = \mathsf{Z}_{ij}[\mathcal{O}_j], \tag{2.28}$$

where repeated indices are summed, and we have assumed mixing between a set of operators $[\mathcal{O}_i]$, $i = 1, \cdots, n$. Our convention for the matrix of anomalous dimensions is

$$\boldsymbol{\gamma} = -\beta(g;\epsilon)\frac{\partial \mathsf{Z}^{-1}}{\partial g}\mathsf{Z}. \tag{2.29}$$

We will consider the operator of dimension 1

$$\overline{\boldsymbol{\psi}}(x)\boldsymbol{\psi}(x) \tag{2.30}$$

and the operators of dimension 2,

$$K = \mathrm{i}\overline{\boldsymbol{\psi}}(x)\cdot \partial\!\!\!/ \boldsymbol{\psi}(x), \qquad V = g_0\big(\overline{\boldsymbol{\psi}}(x)\boldsymbol{\psi}(x)\big)^2, \tag{2.31}$$

which appear in the Lagrangian. The renormalization of the fermion bilinear is straightforward, since it is the mass term appearing in the Lagrangian, and we have

$$\overline{\boldsymbol{\psi}}(x)\boldsymbol{\psi}(x) = Z_{\overline{\boldsymbol{\psi}}\boldsymbol{\psi}}[\overline{\boldsymbol{\psi}}(x)\boldsymbol{\psi}(x)] \tag{2.32}$$

where (see e.g. [26])

$$Z_{\overline{\boldsymbol{\psi}}\boldsymbol{\psi}} = Z_m^{-1}. \tag{2.33}$$

Let us now consider the operators $K$ and $V$. They mix under renormalization, and we can calculate the matrix $\mathsf{Z}_{ij}$ very easily by following the method of [27, 28]. In this method one starts with the bare and renormalized effective actions

$$\begin{aligned}
\Gamma^0 &= \int \mathrm{d}^d x \Big\{ A\mathrm{i}\overline{\boldsymbol{\psi}}\cdot \partial\!\!\!/ \boldsymbol{\psi} + B\frac{g_0}{2}\big(\overline{\boldsymbol{\psi}}\cdot\boldsymbol{\psi}\big)^2 + \cdots \Big\}, \\
\Gamma &= \int \mathrm{d}^d x \Big\{ AZ_\psi \mathrm{i}\overline{\boldsymbol{\psi}}_R \cdot \partial\!\!\!/ \boldsymbol{\psi}_R + BZ_g Z_\psi^2 \frac{g}{2}\big(\overline{\boldsymbol{\psi}}_R\cdot\boldsymbol{\psi}_R\big)^2 + \cdots \Big\}.
\end{aligned} \tag{2.34}$$

The renormalization constants $Z_\psi$ and $Z_g$ are chosen so that divergences are reabsorbed, and we will fix them in such a way that

$$AZ_\psi = BZ_g Z_\psi^2 = 1. \tag{2.35}$$

The bare integrated operators in (2.31) can be obtained by acting with appropriate (functional) derivatives with respect to bare quantities on the effective action $\Gamma^0$:

$$\begin{aligned} \Gamma^0_K &= \left[\frac{1}{2}\int \mathrm{d}^d x \left(\psi_i^\alpha \frac{\delta}{\delta\psi_i^\alpha(x)} + \psi_i^{\dagger\,\alpha}\frac{\delta}{\delta\psi_i^{\dagger\,\alpha}(x)}\right) - 2g_0\frac{\partial}{\partial g_0}\right]\Gamma^0 = \int \mathrm{d}^d x\, K, \\ \Gamma^0_V &= 2g_0\frac{\partial}{\partial g_0}\Gamma^0 = \int \mathrm{d}^d x\, V. \end{aligned} \tag{2.36}$$

Similarly, one can obtain the renormalized versions of the operators, $[K]$ and $[V]$, by taking derivatives of $\Gamma$ with respect to renormalized quantities. The renormalization matrix $\mathsf{Z}$ for the Lagrangian operators satisfies

$$\begin{pmatrix} \Gamma^0_K \\ \Gamma^0_V \end{pmatrix} = \mathsf{Z}\begin{pmatrix} \Gamma_K \\ \Gamma_V \end{pmatrix}, \tag{2.37}$$

and a simple calculation shows that

$$\mathsf{Z} = \begin{pmatrix} 1-2A' & -B' \\ 2A' & 1+B' \end{pmatrix}, \tag{2.38}$$

where we have denoted

$$A' = g_0\frac{\partial \log A}{\partial g_0} \tag{2.39}$$

and similarly for $B$. Explicit expressions for these quantities can be obtained from (2.24), (2.25), and one finds

$$A' = \frac{\gamma(g)}{\epsilon}, \qquad B' = \frac{1}{\epsilon}\left(2\gamma(g) - \frac{\beta(g)}{g}\right). \tag{2.40}$$

We conclude that

$$\begin{pmatrix} K \\ V \end{pmatrix} = \begin{pmatrix} 1-\frac{2\gamma}{\epsilon} & \frac{1}{\epsilon}\left(\frac{\beta}{g}-2\gamma\right) \\ \frac{2\gamma}{\epsilon} & 1+\frac{1}{\epsilon}\left(2\gamma-\frac{\beta}{g}\right) \end{pmatrix}\begin{pmatrix} [K] \\ [V] \end{pmatrix}, \tag{2.41}$$

and the matrix of anomalous dimensions is given by

$$\boldsymbol{\gamma} = \begin{pmatrix} 2g\gamma'(g) & 2g\gamma'(g) - \left(\beta'(g) - \frac{\beta(g)}{g}\right) \\ -2g\gamma'(g) & -2g\gamma'(g) + \beta'(g) - \frac{\beta(g)}{g} \end{pmatrix}. \tag{2.42}$$

One can verify explicitly from (2.41) that the operator $K+V$ does not renormalize. This is a consequence of the fact that $K+V$ can be written as the product of an operator times an equation of motion [26]. Indeed, we have

$$\overline{\psi}\frac{\delta S}{\delta\overline{\psi}} = K + V, \tag{2.43}$$

where $S$ is the action.

## 3 Trans-series from the $1/N$ expansion

### 3.1 An exact result for the self-energy

The Gross–Neveu model can be solved exactly at large $N$, and this means that one can calculate correlation functions as a systematic expansion in $1/N$ (see [29, 30] for an excellent presentation). In the large $N$ formulation, one integrates out the fermions in the action (2.4) and writes down the following effective action for the $\sigma$ field:

$$S_{\text{eff}} = -\int \mathrm{d}^2x \frac{\sigma^2}{2g_0} - \mathrm{i}N\,\mathrm{Tr}\log(\mathrm{i}S_0^{-1}), \tag{3.1}$$

where

$$S_0(\sigma) = \mathrm{i}\big(\mathrm{i}\partial\!\!\!/ - \sigma\big)^{-1} \tag{3.2}$$

is the free propagator for a Dirac fermion. This action has two saddle points at large $N$ in which $\sigma$ takes a constant value $\sigma_c = \pm m_0$, and the classical $\mathbb{Z}_2$ symmetry (2.5) is dynamically broken. The value of $m_0$ is determined by the gap equation

$$\frac{1}{Ng_0} = \frac{1}{m_0}\int \frac{\mathrm{d}^dk}{(2\pi)^d}\mathrm{Tr}\left[\frac{\mathrm{i}}{k\!\!\!/ - m_0}\right]. \tag{3.3}$$

The propagator for the fluctuations of the $\sigma$ field is defined as

$$\Delta^{-1}(x,y) = -\frac{\mathrm{i}}{N}\frac{\delta^2 S_{\text{eff}}}{\delta\sigma(x)\delta\sigma(y)}, \tag{3.4}$$

evaluated at the large $N$ saddle point $\sigma_c = m_0$. In momentum space it is given by

$$\Delta^{-1}(p;m_0) = \frac{\mathrm{i}}{2\pi}\xi\log\left[\frac{\xi+1}{\xi-1}\right], \tag{3.5}$$

where

$$\xi = \sqrt{1-\frac{4m_0^2}{p^2}}. \tag{3.6}$$

See Appendix A.1 for some ingredients in the derivation of this formula. The large $N$ theory describes $\sigma$ particles interacting with fermions, the coupling is of order $N^{-1/2}$, and correlation functions can be computed in terms of large $N$ Feynman diagrams. The fermion self-energy has the following form:

$$\Sigma(p) = m_0 + \frac{1}{N}\big(p\!\!\!/\,\Sigma_p + m_0\Sigma_m\big), \tag{3.7}$$

where $\Sigma_{p,m}$ have a $1/N$ expansion

$$\Sigma_{p,m} = \sum_{j\geq 1}\Sigma^j_{p,m}N^{-j+1}. \tag{3.8}$$

The leading order terms can be computed from the diagrams in Fig. 1, and they are given by [17, 29, 31]

$$\begin{aligned}\Sigma^1_p &= \frac{1}{p^2}\int\frac{\mathrm{d}^2k}{(2\pi)^2}\frac{p^2+k\cdot p}{(p+k)^2-m_0^2}\Delta(k;m_0),\\ \Sigma^1_m &= \int\frac{\mathrm{d}^2k}{(2\pi)^2}\left(\frac{\Delta(k;m_0)}{(p+k)^2-m_0^2} - \frac{2\pi\mathrm{i}}{k^2-4m_0^2}\right).\end{aligned} \tag{3.9}$$

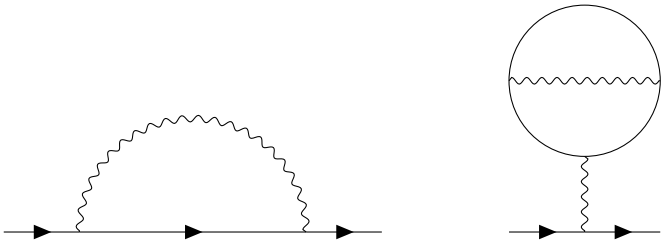


**Figure 1**: Fermion self-energy diagrams at large $N$. The wavy line corresponds to the propagator of the $\sigma$ field, defined in (3.5).

These integrals are divergent and they have to be regularized and renormalized. In [17], Campostrini and Rossi found explicit, finite expressions for them by using a sharp momentum cutoff (SM) regularization scheme. In the SM scheme, one first performs a Wick rotation to Euclidean space and computes the angular integral. Then, the resulting integrand is Taylor expanded at infinity. The terms which lead to a divergence are simply subtracted, but in order to avoid IR divergences in the subtracted pieces one has to introduce an IR cutoff $M$ which plays the role of the renormalization scale $\mu$ in dimensional regularization, see [32, 33] for more details. The renormalized self-energy in the SM scheme can be written in terms of the functions

$$\begin{aligned} A(x) &= \frac{1}{4x}\int_0^\infty \left(\xi_y \log\left[\frac{\xi_y+1}{\xi_y-1}\right]\right)^{-1}\left(1-\frac{1+y-x}{\sqrt{(x+y+1)^2-4xy}}\right)\mathrm{d}y, \\ B(x) &= \frac{1}{2}\int_0^\infty \left(\xi_y \log\left[\frac{\xi_y+1}{\xi_y-1}\right]\right)^{-1}\left(\frac{1}{\sqrt{(x+y+1)^2-4xy}}+\frac{1-\xi_y}{2}\right)\mathrm{d}y, \end{aligned} \tag{3.10}$$

where

$$\xi_y = \sqrt{1+\frac{4}{y}}. \tag{3.11}$$

Then the renormalized self-energy has the form,

$$\Sigma^{\mathrm{SM}}(p) = m + \frac{1}{N}\big(\not{p}\Sigma_p^{\mathrm{SM}} + m\Sigma_m^{\mathrm{SM}}\big), \tag{3.12}$$

where $\Sigma_{p,m}^{\mathrm{SM}}$ are given, at leading order in the $1/N$ expansion,

$$\Sigma_p^{\mathrm{SM}} = -A\left(-\frac{p^2}{m^2}\right) + \mathcal{O}(N^{-1}), \qquad \Sigma_m^{\mathrm{SM}} = -B\left(-\frac{p^2}{m^2}\right) + \mathcal{O}(N^{-1}). \tag{3.13}$$

In (3.12), $m$ is the mass gap, which differs from $m_0$ in $1/N$ corrections:

$$m = m_0 + \frac{m_1}{N} + \mathcal{O}\big(N^{-2}\big). \tag{3.14}$$

$m_1$ can be calculated in terms of the 't Hooft coupling in the SM scheme, see [17] for details. By relating the SM scheme to the $\overline{\mathrm{MS}}$ scheme, one finds [31, 34]

$$m = \left(1+\frac{1}{N}\left(\log 2 - \frac{\gamma_E}{2}+\frac{1}{2}\right)+\cdots\right)\Lambda, \tag{3.15}$$

where $\Lambda$ is the dynamically generated scale in the $\overline{\mathrm{MS}}$ scheme. Let us mention that, in the SM scheme, the beta function and the anomalous dimension of the field are given by [17]

$$\beta_{\mathrm{SM}}(\lambda) = -\lambda^2 + \frac{1}{2N}(2+\lambda)\lambda^2 + \mathcal{O}\big(N^{-2}\big), \qquad \gamma_{\mathrm{SM}}(\lambda) = \mathcal{O}\big(N^{-2}\big). \tag{3.16}$$

### 3.2 Trans-series expression

The functions $A(x)$, $B(x)$ give the exact, non-perturbative result for the self-energy of the GN model up to order $1/N$, in the SM scheme. We will now show that these functions can be written as Borel-resummed trans-series, by following the method developed in [1] for the self-energy of the non-linear sigma model (see also [35] for further details on the strategy and results of [1]). For background on trans-series, Borel resummation, and the theory of resurgence, we refer the reader to [36–39].

Let us first focus on $A(x)$. The first step is to write

$$\left[\log\left(\frac{\xi+1}{\xi-1}\right)\right]^{-1} = \int_0^\infty \mathrm{d}t \left[\frac{\xi-1}{\xi+1}\right]^t. \tag{3.17}$$

We define the following functions, given as Mellin transforms,

$$K_i(s,t) = \int_0^\infty \xi^i \left[\frac{\xi-1}{\xi+1}\right]^t y^{s-1}\mathrm{d}y, \qquad i\in\mathbb{Z}, \tag{3.18}$$

which can be computed as quotients of $\Gamma$ functions. We will use

$$K_{-1}(s,t) = \frac{\Gamma(2s+1)\Gamma(-s+t)}{\Gamma(s+t+1)}. \tag{3.19}$$

We also define

$$\begin{aligned} M(s) &= \int_0^\infty \frac{y^2}{\sqrt{(x+y+1)^2-4xy}} y^{s-1}\mathrm{d}y \\ &= (1+x)^{s+1} B(s+2,-1-s)\, {}_2F_1\left(s+2,-1-s;1;\frac{x}{1+x}\right). \end{aligned} \tag{3.20}$$

Then, by using the properties of the Mellin transform and its inverse, we obtain the following representation for the function $A(x)$,

$$\begin{aligned} A(x) = \frac{1}{4x}\int_0^\infty \Bigg[ &-\int_{C_1} \frac{\mathrm{d}s}{2\pi\mathrm{i}} K_{-1}(-s,t) M(s) \\ &+ (x-1)\int_{C_2} \frac{\mathrm{d}s}{2\pi\mathrm{i}} K_{-1}(-s,t) M(s-1) + K_{-1}(1,t)\Bigg]. \end{aligned} \tag{3.21}$$

In this equation, $C_1$ is the line $-c+\mathrm{i}\mathbb{R}$ with $c\in(-2,-1)$. For $C_2$, we must consider instead $c\in(-1,0)$[3]. We now expand the hypergeometric function in (3.20) at large $x$:

$$M(s) = x^{s+1}\sum_{k\geq 0} a_k(s) x^{-k}, \tag{3.22}$$

with coefficients

$$a_k(s) = \frac{(-1-s)_k \Gamma(s+2)}{k!^2\Gamma(s+2-k)}\Big[\log(x) + 2\psi(k+1) - \psi(-s+k-1) - \psi(s+2-k)\Big]. \tag{3.23}$$

[3]For $C_1$, the value of $c$ has to be in the region of convergence of the integral $M(s)$ appearing in the first line of (3.20). For $C_2$, we have to consider instead $M(s-1)$, so the region of convergence gets displaced by 1. The values of $c$ will be crucial to determine which singularities have to be included when we deform the contours to the left of the complex plane.

We must also consider the combination

$$\begin{aligned} -M(s)+(x-1)M(s-1) &= x^{s+1}\sum_{k\geq 0}\Big(-a_k(s)+a_k(s-1)-a_{k-1}(s-1)\Big)x^{-k} \\ &= x^{s+1}\sum_{k\geq 0} b_k(s)x^{-k}, \end{aligned} \tag{3.24}$$

with the convention that $a_{-1}(s)=0$. Explicitly, the coefficients $b_k(s)$ are given by

$$b_0(s) = 2/(1+s), \tag{3.25}$$

$$b_k(s) = \frac{2k\Gamma(s+1)\Gamma(k-s-1)}{k!^2\Gamma(-s)\Gamma(s+1-k)}\left[\log(x)+2\psi(k)+\frac{1}{k}-\psi(-s+k-1)-\psi(s-k+1)\right]. \tag{3.26}$$

Eq. (3.21) is then written as

$$\begin{aligned} A(x) = \frac{1}{4}\int_0^\infty \mathrm{d}t\Bigg[&\sum_{k\geq 0}\int_{C_1}\frac{\mathrm{d}s}{2\pi\mathrm{i}}K_{-1}(-s,t)b_k(s)x^{-k+s} \\ &+\int_{C_{-1}}\frac{\mathrm{d}s}{2\pi\mathrm{i}}K_{-1}(-s,t)a_k(s)x^{-k+s}+\frac{K_{-1}(1,t)}{x}\Bigg], \end{aligned} \tag{3.27}$$

where $C_{-1}$ is a small circle around $s=-1$. We compute the integral in $s$ by using the residue theorem. We have contributions from:

1. Poles of $K_{-1}(-s,t)$ at $s=-t-j$, $j\in\mathbb{N}$. We group the poles contributing with a factor $(-x)^{-n}x^{-t}$, with fixed $n=k+j$ ($0\leq k\leq n$, $0\leq j\leq n$). In this way, we define the functions

$$\mathcal{E}_n(t) = \frac{(-1)^n}{4}\sum_{j=0}^{n} b_{n-j}(-t-j)\mathrm{Res}_{s=-t-j}K_{-1}(-s,t). \tag{3.28}$$

2. Poles of $b_k(s)$ at $s=-j-1$, $j\in\mathbb{N}$, coming from the last digamma function in (3.26).[4] In this case, we group the poles contributing with a factor $-(-x)^{-n}$, with fixed $n=k+j+1$ ($0\leq k\leq n$, $0\leq j\leq n-1$). In this way, we define

$$\begin{aligned} \mathcal{H}_n(t) = -\frac{(-1)^n}{4}\Bigg[&K_{-1}(1,t)\mathrm{Res}_{s=-1}\big(b_{n-1}(s)+a_{n-1}(s)\big) \\ &+\sum_{j=1}^{n-1}K_{-1}(j+1,t)\mathrm{Res}_{s=-j-1}b_{n-j-1}(s)\Bigg], \qquad n\neq 1, \end{aligned} \tag{3.29}$$

   with the convention that $a_{-1}(s)=b_{-1}(s)=0$.

   For $n=1$, we have to add the contribution from the term

$$\tfrac{1}{4}K_{-1}(1,t) = \frac{1}{2t(t^2-1)} \tag{3.30}$$

   in (3.27).

[4] Naively, there are more poles coming from the digamma function, but they vanish when considering the rest of the expression for $b_k(s)$. The case $k=0$ is an exception in which we only have a single pole at $s=-1$.

With these conventions, we obtain the final result

$$A(x) = \sum_{n\geq 0}(-x)^{-n}\int_0^\infty \mathrm{d}t\big(x^{-t}\mathcal{E}_n(t) - \mathcal{H}_n(t)\big). \tag{3.31}$$

We now recall that, in order to obtain the self-energy $\Sigma_p^{\mathrm{SM}}$, we have to set $x = -p^2/m^2$. We introduce the variable $\lambda$ as

$$-\frac{m^2}{p^2} = \mathrm{e}^{-2/\lambda} \tag{3.32}$$

as well as the Borel variable $y = 2t$. Let us note the important fact that, at leading order in the $1/N$ expansion, $\lambda$ can be identified with the running 't Hooft parameter $\lambda(p)$ at the scale set by $p$ in the $\overline{\mathrm{MS}}$ scheme, since at this order one has $m \approx \Lambda$. We also define the functions $E_n(y)$, $F_n(y)$, $G_n(y)$ and $H_n(y)$ as

$$\begin{aligned} E_n(y) &= \frac{\mathcal{E}_n(y/2)}{2} = \frac{1}{\lambda}F_n(y) + G_n(y), \\ H_n(y) &= \frac{\mathcal{H}_n(y/2)}{2}. \end{aligned} \tag{3.33}$$

We can then write

$$A\Big(-\frac{p^2}{m^2}\Big) = \sum_{n\geq 0}\Big(\frac{m^2}{p^2}\Big)^n \int_0^\infty \mathrm{d}y\Big(\mathrm{e}^{-y/\lambda}E_n(y) - H_n(y)\Big). \tag{3.34}$$

One finds the explicit expressions, for $n = 0, 1$,

$$\begin{aligned} E_0(y) &= \frac{1}{2}\frac{1}{2-y}, & H_0(y) &= 0, \\ E_1(y) &= \frac{1}{\lambda}\frac{y}{4} - \frac{1}{y} - \frac{1}{2} + \frac{y}{8}\Big[1 - 2\gamma_E - \psi\Big(\frac{y}{2}\Big) - \psi\Big(-\frac{y}{2}\Big)\Big], & H_1(y) &= \frac{4}{y(y-2)(y+2)}, \end{aligned} \tag{3.35}$$

where $\psi(x)$ is the digamma function. It is possible to check, on a case by case basis, the equalities

$$\mathrm{Res}_{y=2k}G_n(y) = (-1)^k \mathrm{Res}_{y=2k}H_{n+k}(y), \quad k, n \in \mathbb{N}, \tag{3.36}$$

which are needed to have a complete cancellation of poles in the integral (3.34), after summing over all $n \geq 0$. However, the integrand in (3.34) is singular for a fixed $n$. This allows us to write a formal trans-series out of the above expression, as follows. Let us define

$$r_{n,k} = \mathrm{Res}_{y=2k}\, H_n(y). \tag{3.37}$$

The first step is to rearrange the integral for fixed $n$ as

$$\int_0^{\infty\mathrm{e}^{\mathrm{i}\theta}} \left\{\mathrm{e}^{-y/\lambda}\left(\frac{F_n(y)}{\lambda} + \widehat{G}_n(y)\right) - \left(H_n(y) - r_{n,0}\frac{\mathrm{e}^{-y/\lambda}}{y}\right)\right\}\mathrm{d}y, \tag{3.38}$$

where we have introduced the function

$$\widehat{G}_n(y) = G_n(y) - \frac{r_{n,0}}{y} \tag{3.39}$$

which is regular at the origin. Note that for fixed $n$ the integrand has singularities for positive values of $y$, and we have deformed the integration contour slightly above or below the positive real axis with a small angle $\theta$, to make sense of the integral. The integrand of (3.38)

$$F_n(y) + \lambda \widehat{G}_n(y) = \sum_{k\geq 0} \frac{y^k}{k!} \Big( F_n^{(k)}(0) + \lambda \widehat{G}_n^{(k)}(0) \Big), \tag{3.40}$$

can be regarded as the Borel transform of the factorially divergent series (our convention for the Borel transform is as in [37])

$$\varphi_n(\lambda) = \sum_{k\geq 0} \Big( \lambda^k F_n^{(k)}(0) + \lambda^{k+1} \widehat{G}_n^{(k)}(0) \Big) = F_n(0) + \sum_{k\geq 0} \Big( F_n^{(k+1)}(0) + \widehat{G}_n^{(k)}(0) \Big) \lambda^{k+1}, \tag{3.41}$$

and we can write

$$\int_0^{\infty \mathrm{e}^{\mathrm{i}\theta}} \mathrm{e}^{-y/\lambda} \left( \frac{F_n(y)}{\lambda} + G_n(y) - \frac{r_{n,0}}{y} \right) \mathrm{d}y = s_{\pm}(\varphi_n)(\lambda), \tag{3.42}$$

where $s_\pm$ are lateral Borel resummations (see e.g. [37] for a definition of these). In this and similar expressions in the following, the $\pm$ sign is correlated with the sign of $\theta$ in the contour deformation.

Let us now consider the second piece, involving $H_n(t)$. A simple calculation shows that

$$-\int_0^{\infty \mathrm{e}^{\mathrm{i}\theta}} \left( H_n(y) - r_{n,0} \frac{\mathrm{e}^{-y/\lambda}}{y} \right) \mathrm{d}y = r_{n,0} \log(\lambda) + c_n \pm \mathrm{i}\pi \sum_{k=1}^{n} r_{n,k}, \tag{3.43}$$

where the constant $c_n$ is defined as the principal value integral

$$c_n = -\mathrm{P} \int_0^{\infty} \left( H_n(y) - r_{n,0} \frac{\mathrm{e}^{-y}}{y} \right) \mathrm{d}y. \tag{3.44}$$

It is natural to include the logarithmic and constant pieces in the trans-series, so that the total, factorially divergent series for each $n$ is given by

$$\Phi_{2n}(\lambda) = r_{n,0} \log(\lambda) + c_n \pm \mathrm{i}\pi \sum_{k=1}^{n} r_{n,k} + \sum_{k\geq 0} \Big( \lambda^k F_n^{(k)}(0) + \lambda^{k+1} \widehat{G}_n^{(k)}(0) \Big), \qquad n \geq 0. \tag{3.45}$$

We have labelled these formal series in $\lambda$ with an even index, $2n$, since they multiply even powers of $m$, $m^{2n}$, in the trans-series, and as we will see the function $mB(x)$ leads to odd powers $m^{2n+1}$. For $n = 0, 1$ we use the expressions in (3.35) to obtain

$$\begin{aligned} \Phi_0(\lambda) &= \sum_{k\geq 1} \frac{(k-1)!}{2^{k+1}} \lambda^k, \\ \Phi_2(\lambda) &= \gamma_E + \log(2) \pm \frac{\mathrm{i}\pi}{2} - \log(\lambda) - \frac{\lambda}{4} + \frac{\lambda^2}{8} + \sum_{k\geq 1} \frac{(2k+1)!}{2^{2k+2}} \zeta(2k+1) \lambda^{2k+2}. \end{aligned} \tag{3.46}$$

The series $\Phi_0(\lambda)$ corresponds to the perturbative sector of the self-energy. It has appeared already in [17], since it gives the classical asymptotic expansion of the function $A(x)$. On the other hand,

$\Phi_2(\lambda)$ is the first trans-asymptotic correction to $A(x)$ and multiplies a power correction of order $m^2/p^2$. In addition, it is ambiguous. The ambiguity has a very simple interpretation, in view of the above analysis. The Borel transform of the perturbative series $\Phi_0(\lambda)$ has a singularity at $y = 2$, therefore its two lateral Borel resummations are different. However, the choice of the constant term in the trans-series $\Phi_2(\lambda)$ can be correlated with the choice of lateral resummation of $\Phi_0(\lambda)$, so that the final result is the same for both choices. This is an illustration of the general phenomenon noted by David in [2].

So far we have analyzed the trans-series expression for $A(x)$. In the case of $B(x)$, the trans-series structure can be obtained from the observation that [17]

$$B(x) = \frac{1}{2(1+3/x)}\left(4A(x) + \frac{S(x)}{x}\right), \tag{3.47}$$

where

$$S(x) = \int_0^\infty \mathrm{d}y \left(\log\left[\frac{\xi_y+1}{\xi_y-1}\right]\right)^{-1}\left[\frac{y\xi_y}{\sqrt{(x+y+1)^2-4xy}} - 1 + \frac{x+1}{2}\left(\frac{1}{\xi_y}-1\right)\right]. \tag{3.48}$$

The trans-series structure of $S(x)/x$ is known from [1, 35], and it involves corrections in even powers of $m$. Therefore the trans-series structure for $B(x)$ follows from the results for $A(x)$ and $S(x)/x$. For the purposes of this paper, only the classical asymptotic series of $B(x)$ will be needed. It is given by the formal series in the 't Hooft parameter

$$\Phi_1(\lambda) = \frac{1}{2} - \frac{\gamma_E}{2} - \frac{1}{2}\log(2) + \frac{1}{2}\log(\lambda) + \sum_{k\geq 1}\frac{(2k)!}{2^{2k+1}}\zeta(2k+1)\lambda^{2k+1}. \tag{3.49}$$

Let us summarize the results in this section. The fermion self-energy $\Sigma(p)$ in the SM scheme, at order $1/N$, is given by an exact expression which is a function of the external momentum $p$ and the mass gap $m$. By using the Mellin transform techniques of [1], this expression can be decoded as a trans-series. This involves a perturbative part, given by the series $\Phi_0(\lambda)$, a power correction proportional to $m \sim \Lambda$ and involving the series $\Phi_1(\lambda)$ in (3.49), and a power correction proportional to $m^2 \sim \Lambda^2$ and involving the series $\Phi_2(\lambda)$ in (3.46). More precisely, we have

$$\Sigma^{\mathrm{SM}}(p) \sim -\frac{1}{N}\not{p}\Phi_0(\lambda) + m\left(1 - \frac{1}{N}\Phi_1(\lambda)\right) - \frac{1}{N}\not{p}\frac{m^2}{p^2}\Phi_2(\lambda) + \cdots. \tag{3.50}$$

We note that the l.h.s. of this equation is a well-defined function of $p$, while on the r.h.s. we have a trans-series representation. The first term is the perturbative series, while the second and the third term are non-perturbative power corrections. The dots refer to higher order power corrections, and to higher order corrections in $1/N$. Thanks to the exact large $N$ analysis, we have precise, all-loop predictions for the power series attached to each of these power corrections.

We should be able to reproduce the first term in the r.h.s. of (3.50) from a conventional perturbative calculation in, say, the $\overline{\mathrm{MS}}$ scheme. This was essentially verified in [17], although we will present a more detailed calculation in the next section. In addition, if the method of OPE with vacuum condensates proposed in [10–12] gives the correct trans-series representation of observables, the power corrections in the r.h.s. of (3.50) should be calculable with the SVZ approach, up to unknown overall constants related to the values of the vacuum condensates. We will also verify this in the next section.

## 4 Trans-series from condensates

In this section we will calculate the series $\Phi_{0,1,2}(\lambda)$ in perturbation theory with condensates. We will always work in the $\overline{\text{MS}}$ scheme.

### 4.1 Perturbative series

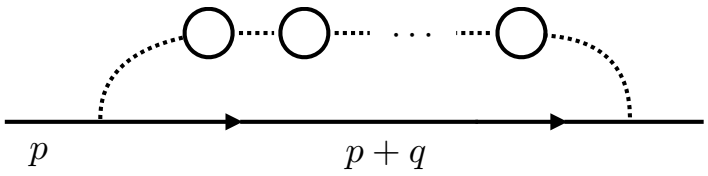



**Figure 2**: The $1/N$ correction to the fermion self-energy is given by a chain of $n$ bubbles.

The first step is to compute the leading $1/N$ correction to the fermion self-energy, at all orders in the 't Hooft coupling constant. This is a standard renormalon-like calculation, since the $n$-th order correction is due to a chain of $n$ fermion loops or "bubbles," linked by $\sigma$ propagators, as shown in Fig. 2. Using the Feynman rules, one finds that the contribution of $n$ loops is

$$(\mathrm{i}\sqrt{g_0})^{2n+2}(-\mathrm{i})^{n+1}N^n\Pi^n(q^2) = \frac{\mathrm{i}}{N}\big(\mathrm{i}\Pi(q^2)\big)^n(\pi\lambda_0)^{n+1}, \tag{4.1}$$

where we have introduced the bare 't Hooft parameter $\lambda_0$ as in (2.13), and $\Pi(q^2)$ is the fermion polarization loop (A.5). We now express $\lambda_0$ in terms of the renormalized coupling constant through (2.26), and we sum the geometric series of bubbles to

$$\frac{\mathrm{i}}{N}\frac{1}{(\pi\lambda(\nu^2)^{\epsilon/2})^{-1}Z_\lambda^{-1} - \mathrm{i}\Pi(q^2)}. \tag{4.2}$$

If we now take into account that

$$-\mathrm{i}\Pi(q^2) = -\frac{1}{\pi\epsilon} + \cdots, \tag{4.3}$$

we deduce that the renormalization constant $Z_\lambda$ is given, at large $N$, by

$$Z_\lambda^{-1} = 1 + \frac{\lambda}{\epsilon} + \cdots, \tag{4.4}$$

in agreement with the result (2.16) for the beta function at leading order in the $1/N$ expansion.

We will write the (bare) self-energy as in (3.7),

$$\Sigma(p) \sim \frac{1}{N}\not{p}\Sigma_p, \tag{4.5}$$

where the asymptotic sign $\sim$ emphasizes that our calculation will lead to a representation in terms of formal power series. If we define

$$\mathcal{I}_0(n) = \frac{1}{\not{p}}\pi^n\int\frac{\mathrm{d}^dq}{(2\pi)^d}\frac{(\not{p}+\not{q})}{(p+q)^2}(\mathrm{i}\Pi(q^2))^{n-1}, \tag{4.6}$$

it is easy to see that the leading term in the $1/N$ expansion of $\Sigma_p$ is

$$\Sigma_p^{\mathrm{P},1} = \sum_{n\geq 2} \mathrm{i}\mathcal{I}_0(n)\lambda_0^n, \tag{4.7}$$

where we have added a superscript P to indicate that this is the perturbative result, and the superindex 1 means that, as in (3.8), this is the term of order $1/N$ in the $1/N$ expansion. We can perform the integral (4.6) explicitly by using (A.3), and we find

$$\mathcal{I}_0(n) = \frac{\mathrm{i}}{4}\left(-\frac{p^2}{4\pi}\right)^{-n\epsilon/2} \frac{\Gamma\left(1-\frac{\epsilon}{2}\right)\Gamma\left(\frac{n\epsilon}{2}\right)\Gamma\left(1-\frac{n\epsilon}{2}\right)}{\Gamma\left(\frac{(n-1)\epsilon}{2}\right)\Gamma\left(2-\frac{(n+1)\epsilon}{2}\right)} \left[-\frac{1}{2}\frac{\Gamma\left(1-\frac{\epsilon}{2}\right)\Gamma\left(1+\frac{\epsilon}{2}\right)\Gamma\left(-\frac{\epsilon}{2}\right)}{\Gamma(1-\epsilon)}\right]^{n-1}. \tag{4.8}$$

We can now use the techniques of Appendix A.2 to write the sum (4.7) in terms of the structure function

$$F(x,y) = -\frac{1}{2}\left(-\frac{p^2}{4\pi\nu^2}\right)^{-y/2} \frac{\Gamma\left(1-\frac{x}{2}\right)\Gamma\left(1+\frac{y}{2}\right)\Gamma\left(1-\frac{y}{2}\right)}{\Gamma\left(\frac{y-x}{2}\right)\Gamma\left(2-\frac{y+x}{2}\right)} \left[\frac{\Gamma\left(1+\frac{x}{2}\right)\Gamma^2\left(1-\frac{x}{2}\right)}{\Gamma(1-x)}\right]^{y/x-1}. \tag{4.9}$$

This function contains both the divergent and the finite part of the bare self-energy. To renormalize the self-energy we need to introduce the renormalization of the field, which has the $1/N$ expansion

$$Z_\psi = 1 + \frac{1}{N} Z_\psi^{(1)} + \cdots. \tag{4.10}$$

The renormalized self-energy is then given by

$$\Sigma_{p,R}^{\mathrm{P},1} = \Sigma_p^{\mathrm{P},1} - Z_\psi^{(1)}. \tag{4.11}$$

By using the results of Appendix A.2 we find

$$Z_\psi^{(1)} = \left[\Sigma_p^{\mathrm{P},1}\right]_{\mathrm{div}} = \left[F_0(\epsilon)\log\left(1+\frac{\lambda}{\epsilon}\right)\right]_{\mathrm{div}}, \tag{4.12}$$

where the function $F_0(\epsilon) = F(\epsilon, 0)$ is given by

$$F_0(\epsilon) = \frac{\epsilon}{2}\frac{1}{2-\epsilon}\frac{\Gamma(1-\epsilon)}{\Gamma^3\left(1-\frac{\epsilon}{2}\right)\Gamma\left(1+\frac{\epsilon}{2}\right)} \tag{4.13}$$

and we have taken into account that $F_{0,0} = F(0,0) = 0$. We can now use (A.24) to obtain the anomalous dimension of the field, at the first non-trivial order in $1/N$, as

$$\gamma^{(1)}(\lambda) = -\lambda F_0(-\lambda) \tag{4.14}$$

which reproduces the result (2.23).

Let us now compute the renormalized self-energy, given by the finite part as $\epsilon \to 0$. The finite part can also be computed in terms of the structure function by using the general formula (A.23). We first note that, from (4.14) and $\beta_\lambda^{(0)}(\lambda) = -\lambda^2$, we have

$$-\int_0^\lambda \frac{F_0(-u)}{u}\mathrm{d}u = -\int_0^\lambda \frac{\gamma^{(1)}(u)}{\beta_\lambda^{(0)}(u)}\mathrm{d}u. \tag{4.15}$$

Therefore, in this case, (A.23) reads

$$\Sigma_{p,R}^{\mathrm{P},1} = -\int_0^\lambda \frac{\gamma^{(1)}(u)}{\beta_\lambda^{(0)}(u)}\mathrm{d}u - F_{0,1}\lambda + \sum_{m\geq 2}(m-1)!F_m(0)\lambda^m. \tag{4.16}$$

From the structure function (4.9), we find

$$-F_{0,1}\lambda = -\frac{\lambda}{4}, \qquad \sum_{m\geq 2} F_m(0)y^m = -\frac{y}{2(2-y)} + \frac{y}{4}, \tag{4.17}$$

where we have set $p^2 = -\mu^2$ in (4.9) and therefore the coupling constant appearing above is $\lambda = \lambda(p)$. Since $F_{0,1} = -F_{1,0}$, we obtain

$$\Sigma_{p,R}^{\mathrm{P},1} = -\int_0^{\lambda(p)} \frac{\gamma^{(1)}(u)}{\beta_\lambda^{(0)}(u)}\mathrm{d}u - \sum_{m\geq 1}\frac{(m-1)!}{2^{m+1}}(\lambda(p))^m. \tag{4.18}$$

To compare this result with (3.50) we have to take into account the change of scheme, from SM to $\overline{\mathrm{MS}}$. The coupling constants $\lambda$ appearing in both expressions are the same, up to this order in $1/N$. In a change of scheme characterized by a coupling $g$ to a new one characterized by $g'$, $\ell$-point Green functions pick a factor $\zeta^\ell(g)$, where $\zeta(g) = 1 + \mathcal{O}(g)$ is determined by the equation (see e.g. [26])

$$\gamma'(g') - \gamma(g) = -2\beta(g)\frac{\partial}{\partial g}\log\zeta(g), \tag{4.19}$$

and $\gamma'(g')$ is the anomalous dimension in the new scheme. In our case, using the SM anomalous dimension in (3.16), we find that

$$\zeta(\lambda) = 1 + \frac{1}{2N}\int_0^\lambda \frac{\gamma^{(1)}(u)}{\beta_\lambda^{(0)}(u)}\mathrm{d}u + \mathcal{O}\big(N^{-2}\big). \tag{4.20}$$

Therefore, the relation between the self-energy $\Sigma_p^{\mathrm{SM}}$ in the SM scheme and the self-energy $\Sigma_{p,R}$ in the $\overline{\mathrm{MS}}$ scheme, at leading order in the $1/N$ expansion, is given by

$$\Sigma_p^{\mathrm{SM}} = \Sigma_{p,R} + \int_0^\lambda \frac{\gamma^{(1)}(u)}{\beta_\lambda^{(0)}(u)}\mathrm{d}u + \mathcal{O}\big(N^{-1}\big). \tag{4.21}$$

Since the power series in the second term of the r.h.s. of (4.18) is nothing but $-\Phi_0(\lambda)$, we find that the perturbative calculation of the self-energy in the $\overline{\mathrm{MS}}$ scheme gives precisely the perturbative part of the trans-series (3.50), after taking into account the correction (4.21).

### 4.2 General aspects of perturbation theory with condensates

We would like now to calculate the power corrections appearing in (3.50) within the SVZ approach. The basic idea in this approach is to use the OPE for the operators appearing in the two-point function. In our case the OPE reads, schematically,

$$\overline{\boldsymbol{\psi}}(x)\boldsymbol{\psi}(0) = C(x)\mathbf{1} + C_{\overline{\psi}\psi}(x)[\overline{\boldsymbol{\psi}}(0)\boldsymbol{\psi}(0)] + C_K(x)[K(0)] + C_V(x)[V(0)] + \cdots, \tag{4.22}$$

where we have included the operators of dimension one and two. One further assumes that the different operators appearing in the OPE have non-vanishing vevs, also called vacuum condensates. In the GN model this is expected to be so, as it can seen for example in the large $N$

solution of the model. Indeed, it is elementary to show that, in the $\overline{\text{MS}}$ scheme, the solution of the gap equation (3.3) is

$$\sigma_c^2 = \Lambda^2. \tag{4.23}$$

On the other hand, $\sigma$ can be integrated out in (2.4) and is given by

$$\sigma = \sqrt{g_0}\overline{\psi}\psi, \tag{4.24}$$

therefore we have

$$\langle[\overline{\psi}\psi]\rangle_c \approx -\frac{N}{\pi\lambda}\Lambda \tag{4.25}$$

at large $N$. The operator $[V]$ appearing in the Lagrangian should also have a non-trivial vev, by large $N$ factorization, and we expect

$$\langle[V]\rangle_c \approx \frac{\pi\lambda}{N}\langle[\overline{\psi}\psi]\rangle_c^2 \tag{4.26}$$

at large $N$. In addition, due to (2.43), we also expect

$$\langle[K]\rangle_c = -\langle[V]\rangle_c. \tag{4.27}$$

Therefore, we will assume that all the operators appearing in the OPE (4.22) lead to non-trivial vacuum condensates, and we will assume that these condensates satisfy the properties (4.26), (4.27), as they follow from large $N$ factorization and basic principles.

In order to make contact with the trans-series (3.50) we need the precise relation between vacuum condensates and the dynamically generated scale $\Lambda$, which can be in turn related to the mass gap through (3.15). This relation follows from general principles, since the vevs of composite operators have to satisfy the Callan–Symanzik equation

$$\left[\delta_{ij}\left(\mu\frac{\partial}{\partial\mu}+\beta(g)\frac{\partial}{\partial g}\right)+\gamma_{ij}\right]\langle[\mathcal{O}_j]\rangle_c = 0, \tag{4.28}$$

in the general case of operator mixing. If there is a single operator $\mathcal{O}$ of dimension $d$, with anomalous dimension $\gamma_{\mathcal{O}}$, the solution of the equation (4.28) is

$$\langle[\mathcal{O}]\rangle_c = \xi\Lambda^d \exp\left(-\int^{g(\mu)}\frac{\gamma_{\mathcal{O}}(u)}{\beta(u)}\mathrm{d}u\right), \tag{4.29}$$

where $\xi$ is an overall constant which might depend on the parameters of the theory, like $N$. Eq. (4.29) applies in particular to the operator $\overline{\psi}\psi$, which does not mix. In this case we have $\gamma_{\overline{\psi}\psi} = -\gamma_m$. By using the explicit expressions (2.19), (2.20) for the mass anomalous dimension, one finds

$$\langle[\overline{\psi}\psi]\rangle_c = -Nc(N)\frac{\Lambda}{\pi\lambda}\exp\left(-\frac{1}{N}\int^{\lambda}\frac{\chi(u)}{u^2}\mathrm{d}u+\mathcal{O}\big(N^{-2}\big)\right), \tag{4.30}$$

where $\chi(u)$ is given in (2.21). Its expansion around $u=0$ has the form

$$\chi(u) = \chi_1 u + \mathcal{O}(u^2), \qquad \chi_1 = \frac{1}{2}, \tag{4.31}$$

and the integral appearing in (4.30) has to be understood as

$$\int^{\lambda}\frac{\chi(u)}{u^2}\mathrm{d}u = \chi_1\log(\lambda) + \int_0^{\lambda}\frac{\chi(u)-\chi_1 u}{u^2}\mathrm{d}u. \tag{4.32}$$

If we compare (4.30) with the result at large $N$ in (4.25) we find that $c(N)$ has the $1/N$ expansion

$$c(N) = 1 + \frac{c_1}{N} + \mathcal{O}\big(N^{-2}\big). \tag{4.33}$$

The choice of normalization in (4.30), with additional factors of $\pi$, $N$ was made so that $c(N) \approx 1$ at large $N$, as shown in (4.33). The subleading terms in the expansion of $c(N)$ could be obtained by calculating $1/N$ corrections to the effective potential.

In the case of the operators $K$, $V$, there is operator mixing and the matrix of anomalous dimensions is given in (2.42). We can reduce the system of equations to a single equation by taking into account (4.27), and solve it in the $1/N$ expansion. A simple calculation gives

$$\langle [V] \rangle_c = N d(N) \frac{\Lambda^2}{\pi\lambda} \exp\left( \frac{1}{N} \left( \frac{\beta^{(1)}(\lambda)}{\lambda^2} - 1 \right) + \mathcal{O}\big(N^{-2}\big) \right), \tag{4.34}$$

where

$$d(N) = 1 + \frac{d_1}{N} + \mathcal{O}\big(N^{-2}\big). \tag{4.35}$$

The fact that $d(N) \approx 1$ at large $N$ is a consequence of the large $N$ factorization (4.26).

With the above results, we can calculate the value of the condensates appearing in the OPE (4.22), up to the non-perturbative functions $c(N)$, $d(N)$. The Wilson coefficients can be determined in various ways. The most practical method is the following (see [40] for an excellent exposition in the context of QCD): one expands in series the interaction terms in the Lagrangian, as in standard perturbation theory. We then Wick-contract the elementary fields with the usual rules, except for the fields that will enter into the condensate. For example, to calculate $C_{\overline{\psi}\psi}$, corresponding to the two-quark condensate, two of the elementary fields shouldn't participate in the contraction, but form the vacuum condensate. The combinatorial possibilities for doing this expansion can be represented by Feynman diagrams in which the fields that form the condensate are represented by blobs. We will see plenty of examples of this procedure in the next section.

In forming the condensates, we will encounter vevs of fermion operators which are not Lorentz scalars or $U(N)$ singlets. These vevs can be determined easily by imposing Lorentz and $U(N)$ invariance. For example, for a general bilinear in fermions, we have

$$\langle \psi_i^\mu(0) \overline{\psi}_j^\nu(0) \rangle_c = -\frac{\delta_{ij}\delta^{\mu\nu}}{2N} \langle \overline{\boldsymbol{\psi}} \boldsymbol{\psi} \rangle_c, \tag{4.36}$$

while for the operators related to the Lagrangian operators $V$, $K$ we have

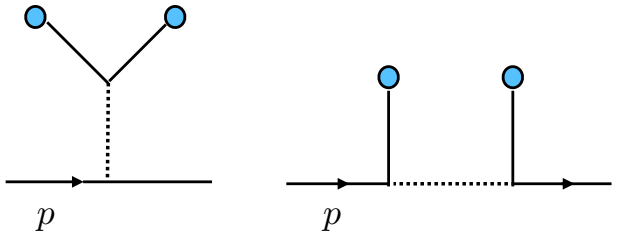
$$\begin{aligned} g_0 \langle \psi_i^\alpha(0) \overline{\psi}_j^\beta(0) \psi_k^\mu(0) \overline{\psi}_l^\nu(0) \rangle_c &= \Big( \delta^{\alpha\beta}\delta^{\mu\nu}\delta_{ij}\delta_{kl} - \delta^{\alpha\nu}\delta^{\mu\beta}\delta_{il}\delta_{jk} \Big) \frac{\langle V \rangle_c}{2N(2N-1)}, \\ \langle \partial_\mu \psi_i^\alpha(0) \overline{\psi}_j^\beta(0) \rangle_c &= \frac{\mathrm{i}}{2dN} (\gamma_\mu)^{\alpha\beta} \delta_{ij} \langle K \rangle_c. \end{aligned} \tag{4.37}$$

We note that we use a dimensional regularization in which the dimension of space-time is $d$, but the dimension of the Dirac spinor space is $2^{\lfloor d/2 \rfloor} = 2$. To relate the vevs of the bare operators to the finite vevs of the renormalized operators we use the renormalization constants in (2.32), (2.41). For the operators $K$, $V$, using (2.43), we find

$$\langle V \rangle_c = -\langle K \rangle_c = \left( 1 - \frac{\beta_\lambda(\lambda)}{\epsilon\lambda} \right) \langle [V] \rangle_c. \tag{4.38}$$

### 4.3 Trans-series for the two-quark condensate

Let us now calculate the contribution to the self-energy of the trans-series associated to the two-quark condensate $\langle\overline{\psi}\psi\rangle_c$. Similar calculations in QCD can be found in [10, 41, 42]. To illustrate the OPE method with condensates, we will first consider in some detail the contribution at leading order in the coupling constant. A pedagogical exposition of the method in the case of QCD can be found in [40].




(a) (b)

**Figure 3**: The two diagrams that contribute to the two-quark condensate correction to the two-point function, at leading order in $g_0$.

It is convenient to start the calculation in position space, and only later Fourier transform into momentum space. We recall the Wick contractions are defined as

$$\overbrace{\sigma(x)\sigma(y)} = -\mathrm{i}\delta(x-y) \tag{4.39}$$

and

$$\overbrace{\psi_i^\mu(x)\overline{\psi}_j^\nu(y)} = \delta_{ij}S_0^{\mu\nu}(x-y), \tag{4.40}$$

where

$$S_0^{\mu\nu}(x) = \int \frac{\mathrm{d}^d k}{(2\pi)^d}\mathrm{e}^{-\mathrm{i}kx}\left[\frac{\mathrm{i}}{\not{k}}\right]^{\mu\nu}. \tag{4.41}$$

To compute the two-point function at order $g_0$, we bring down two factors of the interaction term in the Lagrangian. This yields the following expression

$$-g_0\langle\psi_i^\mu(x)\overline{\psi}_j^\nu(0)\int \mathrm{d}^d y_1 \mathrm{d}^d y_2\, \overline{\psi}_m^\alpha(y_1)\psi_m^\alpha(y_1)\sigma(y_1)\overline{\psi}_n^\beta(y_2)\psi_n^\beta(y_2)\sigma(y_2)\rangle, \tag{4.42}$$

where we have to perform Wick contractions, but leaving a two-quark pair uncontracted to form the condensate. There are two ways of doing this. In the first way, we contract both external legs to the same vertex. One set of contractions is

$$\langle\overbrace{\psi_i^\mu(x)\overbrace{\overline{\psi}_j^\nu(0)\overline{\psi}_m^\alpha(y_1)}\psi_m^\alpha(y_1)}\overbrace{\sigma(y_1)\overline{\psi}_b^\nu(y_2)\psi_b^\nu(y_2)\sigma(y_2)}\rangle \tag{4.43}$$

and there is another, equivalent one obtained by exchange of the vertices $y_1$ and $y_2$. This type of contractions can be represented by the diagram in Fig. 3a. Note that the two-quark pair which leads to the vacuum condensate is traced over. After performing a Fourier transform into momentum space and doing the integrals in the spacetime variables, we obtain

$$-\frac{\mathrm{i}g_0}{p^2}\langle\overline{\psi}\psi\rangle_c. \tag{4.44}$$

This is a contribution of order one in the $1/N$ expansion and, after renormalization, it will give the term $m$ in the trans-series of the self-energy (3.50).

In the second type of contributions, we contract each external leg with a different vertex. This leads to contractions of the form

$$\langle \psi_i^\mu(x) \overline{\psi}_j^\nu(0) \overline{\psi}_m^\alpha(y_1) \psi_m^\alpha(y_1) \sigma(y_1) \overline{\psi}_b^\beta(y_2) \psi_b^\beta(y_2) \sigma(y_2) \rangle \tag{4.45}$$

as well as a similar one obtained by exchanging the vertices. They can be represented by the diagram in Fig. 3b. The vev of the product $\psi_m^\alpha(y_1)\overline{\psi}_b^\beta(y_2)$ leads to a condensate and, due to the delta function $\delta(y_1 - y_2)$ coming from (4.39), both fields are at the same point. After using (4.36), we find that the contribution of the diagram in Fig. 3b to the two-point function is simply

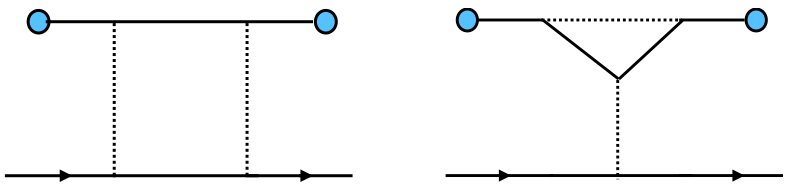
$$\frac{\mathrm{i}g_0}{2N} \frac{1}{p^2} \langle \overline{\boldsymbol{\psi}} \boldsymbol{\psi} \rangle_c. \tag{4.46}$$

A general principle to retain from this calculation is that condensates in which both quarks come from the same interaction vertex have a relative factor of $N$, as compared to condensates where the fermions are from different vertices. This is important when taking into account large $N$ counting.

After multiplying the diagrams in Fig. 3 by i and removing their external legs, we find that the total contribution of the two-quark trans-series to the self-energy, at first order in the coupling, is

$$-g_0 \left(1 - \frac{1}{2N}\right) \langle \overline{\boldsymbol{\psi}} \boldsymbol{\psi} \rangle_c. \tag{4.47}$$

Let us mention that the diagram in Fig. 3b has a counterpart in the calculation of the two-quark condensate correction to the quark propagator in QCD [10, 41], in which the $\sigma$ propagator is replaced with a gluon.



(a) (b)

**Figure 4**: These diagrams are of order $1/N$, but they do not contribute to the two-quark condensate trans-series. Diagram (a) vanishes since it is proportional to $\mathrm{Tr}\,\gamma^\mu = 0$, and (b) involves propagators at zero momentum.

After this pedagogical exercise, let us consider the diagrams which give the trans-series $\Phi_1(\lambda)$ in the self-energy (3.50). In order to proceed, we recall that in calculating the self-energy in conventional perturbation theory, one considers only one-particle irreducible diagrams, i.e. diagrams that can not be divided into subdiagrams by cutting one fermionic internal line. The underlying reason is that reducible diagrams factorize, therefore their contribution can be obtained from their subdiagrams. In the presence of condensates one has to be careful, since diagrams that look reducible in the conventional sense do not factorize. Let us then reconsider the relationship between the two-point function and the self-energy when one includes non-perturbative corrections coming from condensates.

As it is clear from (3.50), the self-energy up to order $1/N$ is a trans-series. The inverse, renormalized two-point function can then be written as

$$\mathrm{i}S_R^{-1}(p) = \not{p} - \frac{1}{N}\not{p}\Sigma_{p,R}^{\mathrm{P}} - \mathcal{C}\Sigma_{m,R}^{\mathrm{NP}} - \frac{\mathcal{C}^2}{N}\not{p}\Sigma_{p,R}^{\mathrm{NP}} + \mathcal{O}(\mathcal{C}^3). \tag{4.48}$$

In this equation, we have introduced a trans-series parameter $\mathcal{C}$ to keep track of the powers of $m$, the superscripts P, NP refer to perturbative and non-perturbative contributions, respectively, and the subscript $R$ stands for renormalized. The self-energies appearing in this equation have $1/N$ expansions with the structure

$$\begin{aligned}
\Sigma_{p,R}^{\mathrm{P}} &= \sum_{j\geq 1} \Sigma_{p,R}^{\mathrm{P},j} N^{-j+1}, \\
\Sigma_{m,R}^{\mathrm{NP}} &= \sum_{j\geq 0} \Sigma_{m,R}^{\mathrm{NP},j} N^{-j}, \\
\Sigma_{p,R}^{\mathrm{NP}} &= \sum_{j\geq 1} \Sigma_{p,R}^{\mathrm{NP},j} N^{-j+1}.
\end{aligned} \tag{4.49}$$

An important remark is that, in doing the $1/N$ expansions of the self-energies computed diagramatically, we keep the condensates themselves fixed, i.e. we do not expand them as in (4.30) or (4.34). This is the natural $1/N$ counting when working with diagrams with condensates. We now expand $S_R(p)$ in $\mathcal{C}$ and $1/N$ to obtain

$$S_R(p) = \frac{\mathrm{i}}{\not{p}}\left\{1 + \frac{1}{N}\Sigma_{p,R}^{\mathrm{P},1} + \frac{\mathcal{C}}{\not{p}}\left(\Sigma_{m,R}^{\mathrm{NP},0} + \frac{1}{N}\Sigma_{m,R}^{\mathrm{NP},1}\right) + \frac{2\mathcal{C}}{N\not{p}}\Sigma_{m,R}^{\mathrm{NP},0}\Sigma_{p,R}^{\mathrm{P},1} + \mathcal{O}(\mathcal{C}^2, N^{-2})\right\}. \tag{4.50}$$

The second term inside the brackets is the perturbative piece calculated in (4.18), and $\Sigma_{m,R}^{\mathrm{NP},0}$ can be obtained from (4.44). The fourth term appearing inside the brackets in the r.h.s. factorizes into a perturbative piece and a non-perturbative piece, and as we will see it corresponds to a reducible diagram.

Let us now consider the diagrams that contribute to (4.50). They have to be of order $1/N$, but incorporate all loops. In conventional perturbation theory, such diagrams are obtained by inserting chain bubbles, and the same principle holds in the case of perturbation theory with condensates. There are however various diagrams that have the right $1/N$ counting but vanish in dimensional regularization, or vanish because they involve a trace of an odd number of gamma matrices. An example is the diagram in Fig. 4a. There is another type of diagrams that do not contribute: condensates are essentially zero-momentum insertions and they can lead to diagrams in which we have propagators at zero momentum. These diagrams have to be discarded [40]. The diagram in Fig. 4b is an example of this. We note however that the diagram in Fig. 4a turns out to contribute to the four-quark condensate correction, as we will explain in section 4.4.

Among the non-vanishing diagrams, one finds Fig. 5. Its contribution factorizes into the contribution of the perturbative diagram of Fig. 2, which is of order $1/N$, and the contribution of the diagram in Fig. 3a. This is the reducible diagram that corresponds to the fourth term inside the bracket in (4.50), as we anticipated above. It does not contribute to the self-energy.

Let us then consider diagrams which do not factorize. We will call such diagrams irreducible. They are shown in Fig. 6 (in the drawings we show only insertions of two or three bubbles, but of course one should consider insertions of an arbitrary number of bubbles). The diagram in Fig. 6a

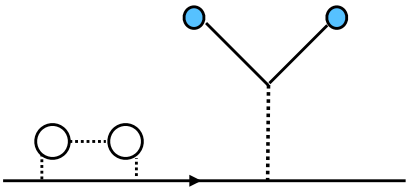
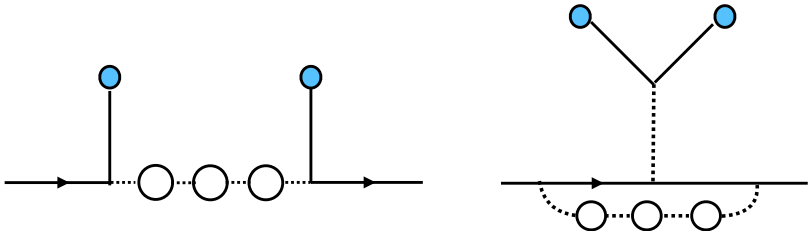


**Figure 5**: A reducible diagram.



(a) (b)

**Figure 6**: Irreducible diagrams that contribute to the two-quark condensate trans-series at order $1/N$, and all loops.

is obtained by inserting the bubble chain (4.1) inside in Fig. 3b, and no additional integration is needed. Its contribution to the self-energy is

$$\frac{\pi\lambda_0}{2N^2}\langle\overline{\psi}\psi\rangle_c \sum_{n\geq 1}\big(\mathrm{i}\Pi(p^2)\big)^n(\pi\lambda_0)^n, \tag{4.51}$$

where $n$ is the number of polarization loops inserted.[5] Notice that, when $n \geq 1$, the fermion and antifermion fields in the condensate are no longer at the same point. In this case, we have to expand

$$\begin{aligned}\langle\psi_m^\alpha(y_1)\overline{\psi}_n^\beta(y_2)\rangle_c &= \langle\psi_m^\alpha(0)\overline{\psi}_n^\beta(0)\rangle_c + (y_1-y_2)^\mu\langle\partial_\mu\psi_m^\alpha(0)\overline{\psi}_n^\beta(0)\rangle_c + \cdots \\ &= -\frac{1}{2N}\delta_{mn}\delta^{\alpha\beta}\langle\overline{\psi}\psi\rangle_c + \cdots .\end{aligned} \tag{4.52}$$

In calculating the contribution to the two-quark condensate we only retain the first term in the r.h.s. of the first line, but the derivative term will contribute to the four-quark condensate, as we will see in the next section.

Let us now consider the diagram in Fig. 6b. A straightforward calculation gives the following contribution to the self-energy:

$$\frac{\pi\lambda_0}{N^2}\langle\overline{\psi}\psi\rangle_c \sum_{n\geq 1}\mathrm{i}\mathcal{I}_1(n)\lambda_0^n, \tag{4.53}$$

where $\mathcal{I}_1(n)$ is the loop integral

$$\mathcal{I}_1(n) = \pi^n\int\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{\big(\mathrm{i}\Pi(q^2)\big)^{n-1}}{(p+q)^2}. \tag{4.54}$$

[5] In this equation, we start the sum at $n=1$, since the term $n=0$ will be already accounted for when we add (4.47).

This integral can be evaluated in dimensional regularization using (A.3):

$$\mathcal{I}_1(n) = -\frac{\mathrm{i}}{4}\left(-\frac{p^2}{4\pi}\right)^{-\epsilon n/2} \frac{\Gamma\left(-\frac{\epsilon}{2}\right)\Gamma\left(\frac{n\epsilon}{2}\right)\Gamma\left(1-\frac{n\epsilon}{2}\right)}{\Gamma\left(\frac{(n-1)\epsilon}{2}\right)\Gamma\left(1-\frac{(n+1)\epsilon}{2}\right)}\left[-\frac{1}{2}\frac{\Gamma\left(1-\frac{\epsilon}{2}\right)\Gamma\left(1+\frac{\epsilon}{2}\right)\Gamma\left(-\frac{\epsilon}{2}\right)}{\Gamma(1-\epsilon)}\right]^{n-1}. \tag{4.55}$$

We can now add the contributions of the two diagrams, together with (4.47), to obtain the bare, non-perturbative correction to the term $\Sigma_m$ in the self-energy:

$$\Sigma_m^{\mathrm{NP}} = -\frac{\pi\lambda_0}{N}\langle\overline{\boldsymbol{\psi}}\boldsymbol{\psi}\rangle_c\left\{1-\frac{1}{2N}-\frac{1}{N}\sum_{n\geq 1}\left(\mathrm{i}\mathcal{I}_1(n)+\frac{1}{2}(\mathrm{i}\pi\Pi)^n\right)\lambda_0^n+\mathcal{O}\left(N^{-2}\right)\right\}. \tag{4.56}$$

There are three sources of renormalization in this quantity: renormalization of the coupling constant, renormalization of the self-energy, and renormalization of the composite operator appearing in the condensate (see [43] for useful remarks on renormalization of fermion propagators in perturbation theory with condensates). The renormalization of the self-energy is done as in the perturbative case, and it just follows by multiplying the inverse two-point function by $Z_\psi$. The renormalization of the composite operator follows from (2.33):

$$\langle\overline{\boldsymbol{\psi}}\boldsymbol{\psi}\rangle_c = Z_m^{-1}\langle[\overline{\boldsymbol{\psi}}\boldsymbol{\psi}]\rangle_c. \tag{4.57}$$

The renormalized result is

$$\Sigma_{m,R}^{\mathrm{NP}} = -\frac{\pi\lambda}{N}\langle[\overline{\boldsymbol{\psi}}\boldsymbol{\psi}]\rangle_c\frac{Z_\lambda}{Z_m}\left\{1-\frac{1}{2N}+\frac{1}{N}Z_\psi^{(1)}-\frac{1}{N}\sum_{n\geq 1}\left(\mathrm{i}\mathcal{I}_1(n)+\frac{1}{2}(\mathrm{i}\pi\Pi)^n\right)\lambda_0^n+\mathcal{O}\left(N^{-2}\right)\right\}. \tag{4.58}$$

Let us note that the sign of the renormalization constant $Z_\psi^{(1)}$ is the opposite one to what is found for the perturbative part in (4.11). This can be seen by comparing (4.50) to the diagrammatic expansion, or simply by noting that $\not{p}$ and $m$ have opposite signs in the inverse propagator. The renormalization constants $Z_\lambda$, $Z_m$ have a $1/N$ expansion

$$\begin{aligned} Z_\lambda &= Z_\lambda^{(0)}\left(1+\frac{1}{N}\widehat{Z}_\lambda^{(1)}+\mathcal{O}\left(N^{-2}\right)\right), \\ Z_m &= Z_m^{(0)}\left(1+\frac{1}{N}\widehat{Z}_m^{(1)}+\mathcal{O}\left(N^{-2}\right)\right), \end{aligned} \tag{4.59}$$

and their first terms are equal:

$$Z_\lambda^{(0)} = Z_m^{(0)}. \tag{4.60}$$

A first consistency check of (4.58) is that it is finite, i.e. that the divergences in the diagrams of Fig. 6 cancel against the renormalization constants. To verify that, and to calculate the finite part, we will calculate the sum over $n$ appearing here by using the formalism explained in Appendix A.2. The structure function which calculates the sum

$$-\sum_{n\geq 1}\left(\mathrm{i}\mathcal{I}_1(n)+\frac{1}{2}(\mathrm{i}\pi\Pi)^n\right)\lambda_0^n \tag{4.61}$$

is given by

$$H(x,y) = -\frac{1}{2}\left(-\frac{p^2}{4\pi\nu^2}\right)^{-y/2}\left[\frac{\Gamma\left(1+\frac{x}{2}\right)\Gamma^2\left(1-\frac{x}{2}\right)}{\Gamma(1-x)}\right]^{y/x-1} \times\left[\frac{\Gamma\left(1+\frac{y}{2}\right)\Gamma\left(1-\frac{y}{2}\right)\Gamma\left(-\frac{x}{2}\right)}{\Gamma\left(\frac{y-x}{2}\right)\Gamma\left(1-\frac{y+x}{2}\right)} - y\frac{\Gamma\left(1+\frac{x}{2}\right)\Gamma\left(1-\frac{x}{2}\right)\Gamma\left(-\frac{x}{2}\right)}{2\Gamma(1-x)}\right]. \tag{4.62}$$

In particular, we have

$$H_0(\epsilon) = H(\epsilon,0) = \frac{\chi(-\epsilon)}{\epsilon} - F_0(\epsilon). \tag{4.63}$$

In this expression, $F_0(\epsilon)$ is the function defined in (4.13) and appearing in the calculation of the perturbative self-energy, and $\chi(\lambda)$ is the function (2.21) in the mass anomalous dimension. Moreover, the $1/N$ expansion of the renormalization constants gives

$$\widehat{Z}_\lambda^{(1)} - \widehat{Z}_m^{(1)} = \int_0^\lambda \mathrm{d}u \frac{\chi(u)}{u(u+\epsilon)}. \tag{4.64}$$

Using (A.19) and (4.12), we obtain

$$\left[H_0(\epsilon)\log\left(1+\frac{\lambda}{\epsilon}\right)\right]_{\mathrm{div}} = -\widehat{Z}_\lambda^{(1)} + \widehat{Z}_m^{(1)} - Z_\psi^{(1)}. \tag{4.65}$$

Now it is clear that all the divergent parts cancel in (4.58). In fact, one can use this calculation to determine the function $\chi(u)$, which essentially gives the mass anomalous dimension at NLO in the $1/N$ expansion.

Let us now evaluate the finite part of the sum. As shown in (A.23), it has two pieces. One of them can be read from the structure function evaluated at $x = 0$:

$$H(0,y) = \frac{y}{4}\left(2\gamma_E + \psi\left(1-\frac{y}{2}\right) + \psi\left(\frac{y}{2}\right)\right). \tag{4.66}$$

After expanding the above expression in powers of $y$, we can extract the coefficients $H_m(0)$ and obtain

$$-\frac{1}{2} + \sum_{m\geq 1}(m-1)!H_m(0)\lambda^m = -\frac{1}{2} - \sum_{k\geq 1}\frac{(2k)!}{2^{2k+1}}\zeta(2k+1)\lambda^{2k+1}, \tag{4.67}$$

where we have included the term $-1/2$ originating from the diagram in Fig. 3b. The other finite piece combines with the non-trivial power series in $\lambda$ which appears in (4.30), when we express the two-quark condensate in terms of $\Lambda$, to produce

$$-\int^\lambda \frac{\chi(u)}{u^2}\mathrm{d}u - \int_0^\lambda \frac{H_0(-u)-H_{0,0}}{u}\mathrm{d}u = -\frac{1}{2}\log(\lambda) + \int_0^\lambda \frac{\gamma^{(1)}(u)}{\beta_\lambda^{(0)}(u)}\mathrm{d}u. \tag{4.68}$$

We finally obtain

$$\Sigma_{m,R}^{\mathrm{NP}} = c(N)\Lambda\left\{1 - \frac{1}{N}\left(\frac{1}{2} + \frac{1}{2}\log(\lambda) + \sum_{k\geq 1}\frac{(2k)!}{2^{2k+1}}\zeta(2k+1)\lambda^{2k+1}\right) + \frac{1}{N}\int_0^\lambda \frac{\gamma^{(1)}(u)}{\beta_\lambda^{(0)}(\lambda)}\mathrm{d}u + \mathcal{O}\left(N^{-2}\right)\right\}. \tag{4.69}$$

The term in the second line is due to the change of scheme, since we have

$$m + \frac{m}{N}\Sigma_m^{\rm SM} = \Sigma_{m,R}\left(1 - \frac{1}{N}\int_0^\lambda \frac{\gamma^{(1)}(u)}{\beta_\lambda^{(0)}(u)}{\rm d}u + \mathcal{O}(N^{-2})\right). \tag{4.70}$$

Note that the correction involving the anomalous dimension of the field now comes with a minus sign, as compared to (4.21). We remove this scheme dependent term, write $\Lambda$ in terms of $m$ using the relation (3.15), and expand the factor $c(N)$ using (4.33). We finally obtain, from (4.69),

$$-\Sigma_m^{\rm SM} \sim 1 - \frac{\gamma_E}{2} + \log(2) - c_1 + \frac{1}{2}\log(\lambda) + \sum_{k\geq 1}\frac{(2k)!}{2^{2k+1}}\zeta(2k+1)\lambda^{2k+1} + \mathcal{O}\big(N^{-1}\big). \tag{4.71}$$

Due to the second equation in (3.13), the series in the r.h.s. should be equal to the series $\Phi_1(\lambda)$ in (3.49). This is indeed the case for the $\lambda$-dependent terms. Equality of the constant term fixes the value of $c_1$:

$$c_1 = \frac{1}{2} + \frac{3}{2}\log(2). \tag{4.72}$$

The calculation above also fixes that $c(N) \approx 1$ at large $N$. We have already incorporated this information by using the large $N$ calculation of the condensate, but we could have obtained it from the comparison of our result (4.69) with the explicit result for the trans-series.

Let us note that our calculation of the two-quark condensate correction to the fermion self-energy is conceptually very similar to what has been done for QCD in [10, 41, 42]. One of the goals of such a calculation in QCD is to dynamically generate a mass for the quarks out of the condensate. We know from the large $N$ analysis of [16] that this is indeed the case in the Gross–Neveu model: the non-zero vev of $\sigma$ gives simultaneously a vev to the two-quark condensate and a mass to the fermions. However, these two quantities are conceptually different (e.g. the first one is not RG invariant, while the second one is), and this difference makes itself manifest at next-to-leading order in the $1/N$ expansion. Our analysis of the self-energy shows in detail how the contribution of the two-quark condensate to the OPE of the fermion propagator generates a mass pole at large $N$ and an additional power correction at order $1/N$, in the manner intended in QCD.

### 4.4 Trans-series for the four-quark condensate

We will now consider the contribution of the four-quark condensate to the trans-series. In the calculation of the two-point function there are two possible sources for such condensate. First, we can leave two pairs of fermions uncontracted when expanding the action. This leads to factors of $\langle V\rangle_c$. Second, we can consider diagrams with a single pair of uncontracted fermions, like the ones we studied in the previous section, but at different locations. When we expand them as in (4.52) we will get factors of $\langle K\rangle_c$. Although this operator is strictly speaking not a four-quark operator, its vev gives the same contribution, but with an opposite sign, as we saw in (4.27). We will then refer to these contributions as also due to a four-quark condensate, by a slight abuse of language.

In order to proceed, we have to understand how to extract the self-energy from the two-point function, as we did in the case of the two-quark condensate. To see how this goes, let us consider the possible diagrams with a four-quark condensate which contribute to the two-point function. It is easy to see that there is a diagram, and only one, which gives a contribution of order one in the $1/N$ expansion, and shown in Fig. 7 (the condensates come from quark pairs in the same

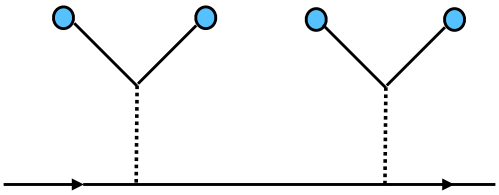


**Figure 7**: Diagram of order $g_0^2$ and order 1 in the $1/N$ expansion that contributes to the four-quark condensate correction to the two-point function.

vertices, so they give a factor of $N^2$, and the diagram goes like $g_0^2 N^2$). A precise evaluation of the diagram gives

$$\frac{g_0 \langle V \rangle_c}{p^2}. \tag{4.73}$$

Note that this does not factorize, so the diagram in Fig. 7 should not be regarded as a reducible diagram. The factorization only takes place at large $N$. In order to do the precise large $N$ counting, let us already renormalize the contribution of this diagram. For the composite operator $V$, we need the renormalization constant in (4.38):

$$Z_V = 1 - \frac{\beta_\lambda(\lambda)}{\lambda\epsilon} = Z_V^{(0)} \left( 1 + \frac{1}{N} \widehat{Z}_V^{(1)} + \mathcal{O}\big(N^{-2}\big) \right) \tag{4.74}$$

and we have

$$Z_V^{(0)} Z_\lambda^{(0)} = 1, \qquad \widehat{Z}_V^{(1)} = -\frac{\beta_\lambda^{(1)}(\lambda)}{\lambda(\lambda+\epsilon)}. \tag{4.75}$$

Using in addition the renormalization constant of the coupling and its $1/N$ expansion in (4.59), we find

$$Z_\lambda Z_V \frac{\pi\lambda}{N} \langle [V] \rangle_c = \frac{\pi\lambda}{N} \langle [V] \rangle_c \left( 1 + \frac{1}{N} \left( \widehat{Z}_\lambda^{(1)} + \widehat{Z}_V^{(1)} \right) + \cdots \right). \tag{4.76}$$

Note that we have the large $N$ scaling $\langle [V] \rangle_c \sim N$, so the front factor is of order 1. We will denote the renormalized two-quark and four-quark condensates by

$$\mathcal{T} = -\frac{\pi\lambda}{N} \langle [\overline{\psi}\psi] \rangle_c, \qquad \mathcal{F} = \frac{\pi\lambda}{N} \langle [V] \rangle_c. \tag{4.77}$$

By the large $N$ factorization of (4.26), one has

$$\mathcal{F} - \mathcal{T}^2 = \mathcal{O}\big(N^{-1}\big). \tag{4.78}$$

We also note that

$$\Sigma_{m,R}^{\mathrm{NP},0} = \mathcal{T}. \tag{4.79}$$

In addition to the diagram of Fig. 7, which is of order one at large $N$, we find many other diagrams of order $1/N$ in the calculation of the four-quark condensate corrections. Among these, there are the diagrams shown in Fig. 8 (together with their right-left reflections). These diagrams do not contribute to the self-energy, up to order $1/N$. To see this, we note that the term of order $\mathcal{C}^2$ in the expansion of the renormalized two-point function (4.50) is, up to order $1/N$,

$$\frac{1}{N} \Sigma_{p,R}^{\mathrm{NP},1} + \frac{1}{p^2} \left( \mathcal{T}^2 + \frac{2}{N} \mathcal{T} \Sigma_{m,R}^{\mathrm{NP},1} \right) + \frac{3}{p^2 N} \mathcal{T}^2 \Sigma_{p,R}^{\mathrm{P},1}. \tag{4.80}$$

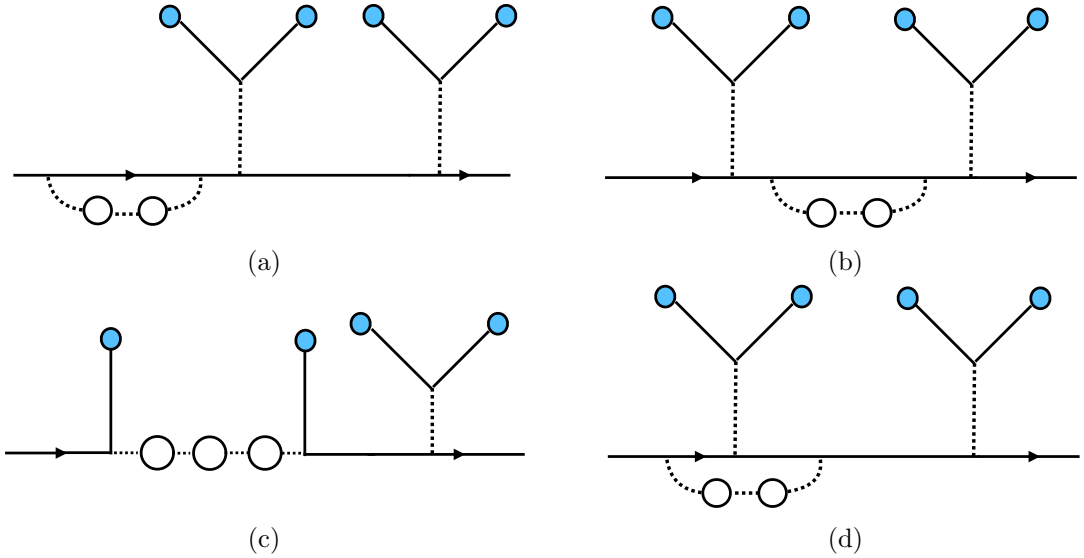



**Figure 8**: Diagrams with four-quark condensates that contribute to the two-point function but not to the self-energy, at order $1/N$. One should include as well the diagrams obtained by left-right reflection.

To compare this expression with the diagrammtic computation, we still need to include the field renormalization constant and write renormalized quantities together with their original divergences. This yields

$$\frac{1}{N}\Sigma_{p,R}^{\mathrm{NP},1}+\frac{1}{p^2}\mathcal{T}^2+\frac{2}{p^2N}\mathcal{T}^2\Big(\widehat{Z}_\lambda^{(1)}-\widehat{Z}_m^{(1)}\Big) \\ +\frac{3}{p^2N}\mathcal{T}^2\Big(\Sigma_{p,R}^{\mathrm{P},1}+Z_\psi^{(1)}\Big)+\frac{2}{p^2N}\mathcal{T}\Big(\Sigma_{m,R}^{\mathrm{NP},1}+\mathcal{T}\Big(-\widehat{Z}_\lambda^{(1)}+\widehat{Z}_m^{(1)}-Z_\psi^{(1)}\Big)\Big). \quad (4.81)$$

Due to the factorization (4.78), the first term in the second line corresponds to the diagrams in Fig. 8a–8b, while the second term in the second line corresponds to the diagrams in Fig. 8c–8d (in both cases, up to this order in $1/N$). Only the first line remains to be accounted for in (4.81), which has to be given by the diagram in Fig. 7 plus irreducible diagrams of order $1/N$, i.e. those that are not included in Fig. 8. Including the renormalization constants in (4.76), we obtain

$$\frac{1}{N}\Sigma_{p,R}^{\mathrm{NP}}=\frac{\mathcal{F}-\mathcal{T}^2}{p^2}+\frac{\mathcal{F}}{p^2N}\Big(\widehat{Z}_\lambda^{(1)}+\widehat{Z}_V^{(1)}-2\Big(\widehat{Z}_\lambda^{(1)}-\widehat{Z}_m^{(1)}\Big)\Big)-\mathrm{i}\not{p}\times\text{irreducible}+\mathcal{O}\big(N^{-2}\big). \quad (4.82)$$

The first term in (4.82) can be written more explicitly by reexpressing the condensates in terms of $\Lambda$, through (4.30) and (4.34). We find

$$\mathcal{F}-\mathcal{T}^2=\frac{\Lambda^2}{N}\left(d_1-2c_1+\log(\lambda)-1+\frac{\beta_\lambda^{(1)}(\lambda)}{\lambda^2}+2\int_0^\lambda\frac{\chi(u)-\chi_1 u}{u^2}\mathrm{d}u\right)+\mathcal{O}\big(N^{-2}\big), \quad (4.83)$$

where the coefficients $d_1$, $c_1$ were introduced in (4.33), (4.35). Based on the general arguments in [2, 3], we expect $d_1$, the $1/N$ correction to the four-quark condensate, to be ambiguous. We will see that this is indeed the case.

Let us note that the second term in the r.h.s. of (4.82) should cancel the divergences obtained from these diagrams. Let us find a more convenient form for this combination of renormalization constants. We first derive from (2.27) the integral form

$$\widehat{Z}_\lambda^{(1)} = \int_0^\lambda \frac{\epsilon\beta_1(u)}{u^2(u+\epsilon)^2}\mathrm{d}u. \tag{4.84}$$

After integration by parts we obtain the convenient expression

$$\widehat{Z}_\lambda^{(1)} + \widehat{Z}_V^{(1)} = \int_0^\lambda \frac{\rho(u)}{u(u+\epsilon)}\mathrm{d}u, \tag{4.85}$$

where

$$\rho(\lambda) = -\lambda^2 \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\beta_\lambda^{(1)}(\lambda)}{\lambda^2}\right). \tag{4.86}$$

The quantity $\widehat{Z}_\lambda^{(1)} - \widehat{Z}_m^{(1)}$ is known from the calculation of the two-quark condensate (4.64). Therefore, the singular part in the irreducible diagrams that contribute to the four-quark condensate determines the beta function at NLO in $1/N$. As we will see, this calculation is simpler than the one usually adopted for the calculation of the beta function at this order in [18, 19]. By using the known value (2.17), we have

$$\rho(\lambda) = -\frac{\lambda^2\Gamma(2+\lambda)}{(2+\lambda)\Gamma\left(1-\frac{\lambda}{2}\right)\Gamma^3\left(1+\frac{\lambda}{2}\right)} = -\lambda\chi(\lambda). \tag{4.87}$$

We can now write

$$\widehat{Z}_\lambda^{(1)} + \widehat{Z}_V^{(1)} - 2\left(\widehat{Z}_\lambda^{(1)} - \widehat{Z}_m^{(1)}\right) = \int_0^\lambda \frac{\upsilon(u)}{u(u+\epsilon)}\mathrm{d}u, \tag{4.88}$$

where

$$\upsilon(u) = -\frac{u\Gamma(2+u)}{\Gamma^3\left(1+\frac{u}{2}\right)\Gamma\left(1-\frac{u}{2}\right)}. \tag{4.89}$$

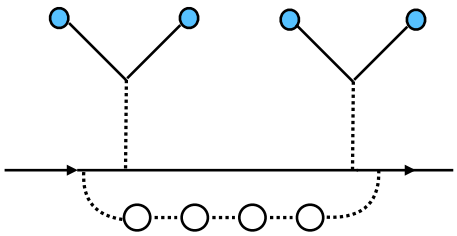


**Figure 9**: Irreducible diagram that contributes to the four-quark condensate.

We have then to find the irreducible diagrams, to order $1/N$ and all loops. There are four types of diagrams that contribute:

1. The first type of diagram is obtained by putting a four-quark condensate in the middle of the internal propagator line of the bubble chain in Fig. 2. This gives the diagram shown in Fig. 9.

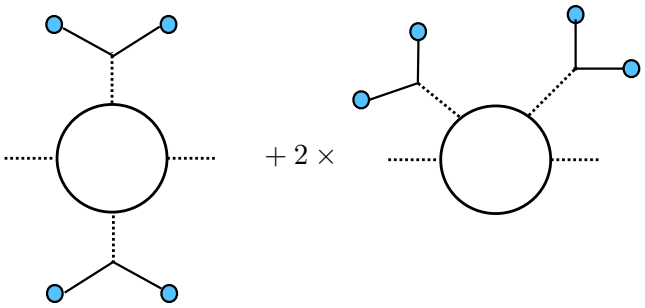



**Figure 10**: Two possible ways of inserting a four-quark condensate in a fermion polarization loop. The second diagram needs a symmetry factor 2.

2. The second type of diagram is obtained by noticing that one can insert a four-quark condensate in a fermion polarization loop, in two different ways, to obtain a "decorated" polarization loop, as shown in Fig. 10. This "decorated" loop can then be inserted at any point inside a bubble chain, and leads to diagrams like the one in Fig. 11.

3. The third type of diagram is obtained by noticing that one can insert a four-quark condensate in the sigma propagator, which can then be inserted in a bubble chain, as shown in Fig. 12.

4. Finally, there are contributions coming from diagrams involving a two-quark condensate, in which the quark and antiquark are at different points, and one has to expand. The diagrams that contribute to this are the diagram in Fig. 6a, and the diagram in Fig. 13. Note that the latter gives a vanishing contribution to the two-quark condensate, but not to the four-quark condensate.

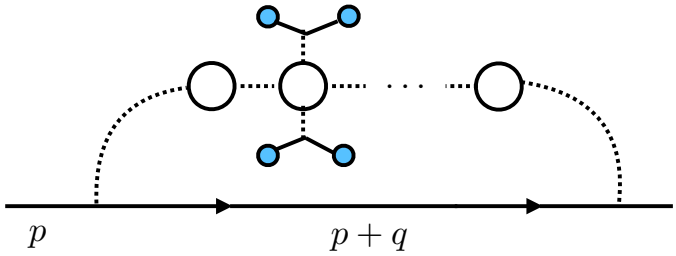



**Figure 11**: A chain of bubbles with an insertion of one of the "decorated" polarization loops.

Let us now compute the contribution of the irreducible diagrams. It will be convenient to group them in appropriate ways.

We first consider the irreducible diagram in Fig. 9, which we will combine with the diagram in Fig. 6a. The contribution of Fig. 9 to $\Sigma_p^{\mathrm{NP}}$ is of the form

$$-\frac{\pi\lambda_0}{N^2}\frac{\langle V\rangle_c}{p^2}\sum_{n\geq 1}\mathrm{i}\mathcal{I}_2(n)\lambda_0^n, \tag{4.90}$$

where

$$\mathcal{I}_2(n) = \frac{p^2}{\not{p}}\pi^n\int\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{\not{p}-\not{q}}{(p-q)^4}\big(\mathrm{i}\Pi(q^2)\big)^{n-1}. \tag{4.91}$$

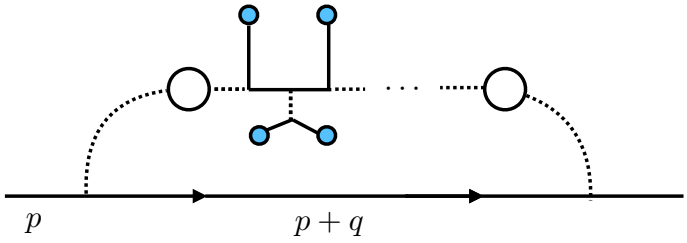



**Figure 12**: A chain of bubbles with an insertion of a four-quark condensate in the propagator of the sigma field.

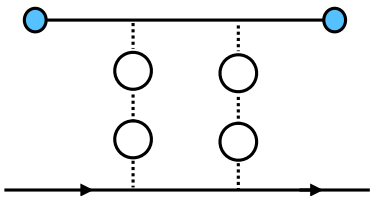


**Figure 13**: When we expand the two-quark condensate around the same point, this diagram gives a contribution to the four-quark condensate.

Let us now consider the diagram in Fig. 6a. In the previous section we calculated the first term in the expansion in the first line of (4.52), and we have to consider now the second term, which will produce a factor of $\langle K\rangle_c$. Its contribution to the diagram in position space involves an integral of the form

$$\int \mathrm{d}^d y_1 \mathrm{d}^d y_2 (y_1-y_2)^\rho \mathrm{e}^{\mathrm{i}(p-q)y_1+\mathrm{i}(q-k)y_2} \big(\Pi(q^2)\big)^n \tag{4.92}$$

where $p$ is the external momentum, $k$ and $q$ are internal momenta to be integrated over, and $n$ is the number of inserted bubbles. To calculate this integral, we write

$$(y_1-y_2)^\rho = -\frac{1}{\mathrm{i}}\frac{\partial}{\partial q_\rho}\mathrm{e}^{\mathrm{i}q(y_2-y_1)} \tag{4.93}$$

and we integrate by parts (see e.g. [40, 42] for similar calculations). After taking everything into account, we find that this diagram gives a contribution to $\Sigma_p^{\mathrm{NP}}$ of the form

$$-\frac{\pi\lambda_0}{dN^2}\langle K\rangle_c \sum_{n\geq 1}\frac{\partial(\mathrm{i}\pi\Pi(p^2))^n}{\partial p^2}\lambda_0^n, \tag{4.94}$$

which will combine with (4.90) into

$$\frac{\pi\lambda_0}{N^2}\frac{\langle V\rangle_c}{p^2}\sum_{n\geq 1}\left(\frac{p^2}{d}\frac{\partial(\mathrm{i}\pi\Pi(p^2))^n}{\partial p^2}-\mathrm{i}\mathcal{I}_2(n)\right)\lambda_0^n, \tag{4.95}$$

after using (4.27). The general formula (A.3) gives

$$\mathcal{I}_2(n)=\mathrm{i}\left(-\frac{p^2}{4\pi}\right)^{-n\epsilon/2}\frac{\Gamma\big(-\frac{\epsilon}{2}\big)\Gamma\big(1+\frac{n\epsilon}{2}\big)\Gamma\big(1-\frac{n\epsilon}{2}\big)}{\Gamma\Big(\frac{(n-1)\epsilon}{2}\Big)\Gamma\Big(1-\frac{(n+1)\epsilon}{2}\Big)}\left[-\frac{1}{2}\frac{\Gamma\big(1+\frac{\epsilon}{2}\big)\Gamma\big(1-\frac{\epsilon}{2}\big)\Gamma\big(-\frac{\epsilon}{2}\big)}{\Gamma(1-\epsilon)}\right]^{n-1}. \tag{4.96}$$

On the other hand,

$$\frac{p^2}{d}\frac{\partial(\mathrm{i}\pi\Pi(p^2))^n}{\partial p^2} = -\frac{1}{4}\Big(-\frac{p^2}{4\pi}\Big)^{-n\epsilon/2}\Big(-\frac{n\epsilon}{2}\Big)\frac{\Gamma\big(1+\frac{\epsilon}{2}\big)\Gamma\big(1-\frac{\epsilon}{2}\big)\Gamma\big(-\frac{\epsilon}{2}\big)}{\big(1-\frac{\epsilon}{2}\big)\Gamma(1-\epsilon)} \times\left[-\frac{1}{2}\frac{\Gamma\big(1+\frac{\epsilon}{2}\big)\Gamma\big(1-\frac{\epsilon}{2}\big)\Gamma\big(-\frac{\epsilon}{2}\big)}{\Gamma(1-\epsilon)}\right]^{n-1}. \tag{4.97}$$

Combining both, and using the technique of Appendix A.2, we find that the sum in (4.95) is governed by the structure function

$$M(x,y) = -\frac{1}{4}\Big(-\frac{p^2}{4\pi\nu^2}\Big)^{-y/2}\left[\frac{\Gamma\big(1+\frac{x}{2}\big)\Gamma^2\big(1-\frac{x}{2}\big)}{\Gamma(1-x)}\right]^{y/x-1} \times y\left[\frac{\Gamma\big(-\frac{x}{2}\big)\Gamma\big(1+\frac{y}{2}\big)\Gamma\big(1-\frac{y}{2}\big)}{\Gamma\big(\frac{y-x}{2}\big)\Gamma\big(1-\frac{y+x}{2}\big)} - \frac{y}{2}\frac{\Gamma\big(1+\frac{x}{2}\big)\Gamma\big(1-\frac{x}{2}\big)\Gamma\big(-\frac{x}{2}\big)}{\big(1-\frac{x}{2}\big)\Gamma(1-x)}\right]. \tag{4.98}$$

Note also that $M_0(x) = 0$. This means that the sum (4.95) can be made finite simply by renormalizing the coupling constant. We also find

$$M(0,y) = \frac{y^2}{8}\Big(-1+2\gamma_E+\psi\Big(1-\frac{y}{2}\Big)+\psi\Big(\frac{y}{2}\Big)\Big). \tag{4.99}$$

We conclude that the contribution of these two classes of diagrams to $\Sigma_p^{\mathrm{NP}}$ is

$$-\frac{m^2}{p^2}\left\{\frac{\lambda}{4}+\frac{\lambda^2}{8}+\sum_{k\geq 1}\frac{(2k+1)!}{2^{2k+2}}\zeta(2k+1)\lambda^{2k+2}\right\}. \tag{4.100}$$

We will now consider the combination of the diagrams in Fig. 11 and Fig. 12. First of all, we calculate the amplitude associated to the "decorated" loops in Fig. 10. We find

$$\Pi_{(\bar\psi\psi)^2}(p^2) = 2\int\frac{\mathrm{d}^d q}{(2\pi)^d}\left[\frac{3}{q^2(q-p)^2}-\frac{2q\cdot p}{q^4(q-p)^2}\right] = -\frac{2}{p^2}(1-\epsilon)\Pi(p^2). \tag{4.101}$$

Here we have included both the sign $-1$ due to the fermionic loop and the $-1=(-\mathrm{i})^2$ coming from the two $\sigma$ propagators. The symmetry factor of the diagram in Fig. 11, in which there are $n-2$ conventional polarization loops and one decorated loop, is $n-1$. The diagram in Fig. 12, in which there are $n-1$ polarization bubbles, has a symmetry factor $2n$. The contribution of these two classes of diagrams to the self-energy is

$$\frac{\pi\lambda_0}{N^2}\frac{\langle V\rangle_c}{p^2}\sum_{n\geq 1}(-2n\mathrm{i}\mathcal{I}_3(n)-(n-1)\mathrm{i}\mathcal{I}_4(n))\lambda_0^n, \tag{4.102}$$

where

$$\mathcal{I}_3(n) = \frac{p^2}{\not p}\pi^n\int\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{\not p-\not q}{q^2(p-q)^2}\big(\mathrm{i}\Pi(q^2)\big)^{n-1} \tag{4.103}$$

is the integral associated to the diagram in Fig. 12, while

$$\mathcal{I}_4(n) = \frac{p^2}{\not p}\pi^n\int\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{\not p-\not q}{(p-q)^2}\big(\mathrm{i}\Pi(q^2)\big)^{n-2}\Big(\mathrm{i}\Pi_{(\bar\psi\psi)^2}(q^2)\Big) \tag{4.104}$$

is the integral associated to the diagram in Fig. 11. In view of (4.101), the two integrals are related as

$$\mathcal{I}_4(n) = -2(1-\epsilon)\mathcal{I}_3(n). \tag{4.105}$$

The integral $\mathcal{I}_3(n)$ can be computed with the expression (A.3), and we find that the sum in (4.102) is governed by the structure function

$$R(x,y) = \left(-\frac{p^2}{4\pi\nu^2}\right)^{-y/2}(1+y-x)\frac{\Gamma\left(1-\frac{x}{2}\right)\Gamma\left(1-\frac{y}{2}\right)\Gamma\left(1+\frac{y}{2}\right)}{\Gamma\left(1+\frac{y-x}{2}\right)\Gamma\left(1-\frac{y+x}{2}\right)}\left[\frac{\Gamma\left(1+\frac{x}{2}\right)\Gamma^2\left(1-\frac{x}{2}\right)}{\Gamma(1-x)}\right]^{y/x-1}. \tag{4.106}$$

We note that

$$R_0(x) = \frac{\Gamma(2-x)}{\Gamma^3\left(1-\frac{x}{2}\right)\Gamma\left(1+\frac{x}{2}\right)} = \frac{\upsilon(-x)}{x}, \tag{4.107}$$

where $\upsilon(x)$ was introduced in (4.89). We also have

$$R(0,y) = 1+y \tag{4.108}$$

for $\mu^2 = -p^2$. Let us calculate the finite and divergent parts due to (4.107). We have

$$\left[R_0(\epsilon)\log\left(1+\frac{\lambda}{\epsilon}\right)\right]_{\text{div}} = \int_0^\lambda \frac{R_0(-u)}{u+\epsilon}\mathrm{d}u = -\int_0^\lambda \frac{\upsilon(u)}{u(u+\epsilon)}\mathrm{d}u. \tag{4.109}$$

This will cancel precisely the divergent part in (4.88). Therefore, the diagrams of Fig. 11 and Fig. 12 are the relevant ones to compute the anomalous dimension of the operator $(\overline{\psi}\psi)^2$ and, therefore, of the beta function at NLO. A related calculation of this anomalous dimension in [44] uses these diagrams in disguise.

The finite part due to $R_0(x)$ is given by

$$-\int_0^\lambda \frac{R_0(-u)-1}{u}\mathrm{d}u = 1 - \frac{\beta_\lambda^{(1)}(\lambda)}{\lambda^2} - 2\int_0^\lambda \frac{\chi(u)-\chi_1 u}{u^2}\mathrm{d}u, \tag{4.110}$$

and it cancels with part of the expression in (4.83), leaving us with the following contribution to $\Sigma_{p,R}^{\text{NP}}$:

$$\frac{m^2}{p^2}(d_1 - 2c_1 + \log(\lambda)). \tag{4.111}$$

This result combines with the finite part coming from (4.108), resulting in the contribution

$$\frac{m^2}{p^2}(d_1 - 2c_1 + \log(\lambda) + \lambda). \tag{4.112}$$

Let us finally consider the contribution of the family of diagrams in Fig. 13. When the two quarks are at the same point, this diagram is proportional to $\mathrm{Tr}(\gamma^\mu)$ and it vanishes. When we expand the two quark fields around the same point, we find a contribution to the four-quark condensate which involves an integral of the form

$$\int \mathrm{d}^d y_1 \mathrm{d}^d y_2\, \mathrm{e}^{\mathrm{i}(p-k-q)y_1 - \mathrm{i}(r-k-q)y_2}(y_1-y_2)^\rho \frac{\mathrm{i}}{\not{k}}\frac{q_\rho}{q^2}, \tag{4.113}$$

where $p$ is the external momentum and $k, q, r$ are internal momenta to be integrated over. In the integration by parts one obtains

$$\frac{\partial}{\partial q_\rho}\left(\frac{q_\rho}{q^2}\right) = \frac{d-2}{q^2}. \tag{4.114}$$

The contribution of the sum over all bubbles is

$$\frac{\pi\lambda_0}{N^2}\frac{\langle V\rangle_c}{p^2}\sum_{n\geq 1}\frac{d-2}{d}2ni\mathcal{I}_3(n)\lambda_0^n, \tag{4.115}$$

which can be calculated in terms of the structure function

$$S(x,y) = -\left(-\frac{p^2}{4\pi\nu^2}\right)^{-y/2}\frac{y}{2-x}\frac{\Gamma\left(1-\frac{x}{2}\right)\Gamma\left(1-\frac{y}{2}\right)\Gamma\left(1+\frac{y}{2}\right)}{\Gamma\left(1-\frac{y+x}{2}\right)\Gamma\left(1+\frac{y-x}{2}\right)}\left[\frac{\Gamma\left(1+\frac{x}{2}\right)\Gamma^2\left(1-\frac{x}{2}\right)}{\Gamma(1-x)}\right]^{y/x-1}. \tag{4.116}$$

We find $S_0(x) = 0$, and

$$S(0,y) = -\frac{y}{2}. \tag{4.117}$$

Including the factors in front of the sum in (4.115), we obtain a contribution to $\Sigma^{\rm NP}_{p,R}$ of the form

$$-\frac{m^2}{p^2}\frac{\lambda}{2}. \tag{4.118}$$

The results in (4.100), (4.112) and (4.118) combine into

$$\Sigma^{\rm NP}_{p,R} = -\frac{m^2}{p^2}\left\{2c_1 - d_1 - \log(\lambda) - \frac{\lambda}{4} + \frac{\lambda^2}{8} + \sum_{k\geq 1}\frac{(2k+1)!}{2^{2k+2}}\zeta(2k+1)\lambda^{2k+2}\right\} + \mathcal{O}\left(N^{-1}\right). \tag{4.119}$$

This agrees precisely with the result in (3.50), involving the series $\Phi_2(\lambda)$ of (3.46), except for the constant terms, which depend on the values of the condensates. Moreover, by comparing the two results we can read off the value of $d_1$, giving the $1/N$ correction to the four-quark condensate

$$d_1 = 1 - \gamma_E + 2\log(2) \mp \frac{i\pi}{2}. \tag{4.120}$$

As advertised, this is ambiguous due to the renormalon in the perturbative series, and the choice of sign in (4.120) should be correlated with a choice of resummation prescription for the series $\Phi_0(\lambda)$.

## 5 Conclusions and open problems

The combination of the OPE with vacuum condensates, as developed in the SVZ sum rules, produces formal trans-series for QFT observables, and it provides a method to systematically calculate exponentially small corrections to perturbative quantities. In this paper we have compared the result of an OPE calculation to an exact trans-series obtained in the $1/N$ expansion, and we have found complete agreement between the two, up to the unknown values of the condensates.

In our calculation we have used the "practical" version of the OPE. As we mentioned in the Introduction, it has been suggested that one should use a more complicated version, based on the introduction of an additional momentum scale. In this "Wilsonian" version, IR renormalons are

absent, since the additional scale provides an explicit IR cutoff for the Feynman integrals, and the condensates are defined unambiguously. The prize to pay is that each series in the trans-series depends on this scale, and the dependence only drops out in the total result. Our study of the Gross–Neveu model seems to confirm that there is nothing wrong with the "practical" version of the OPE, in agreement with the discussions in [3, 45]. The exact results for the two-point function can be decoded in terms of a trans-series in which condensates are ambiguous, but this ambiguity is due to the well-known Stokes phenomenon and does not lead to any inconsistency. In addition, the series in the 't Hooft parameter appearing in this trans-series can be reproduced exactly, and rather non-trivially, with the "practical" OPE.

A nice outcome of our calculation is the following. In the exact large $N$ answer, the trans-series emerges as a formal, algebraic object. The OPE calculation gives a concrete picture of this trans-series in terms of perturbation theory with condensates. In particular, the factorial divergence of the perturbative series appearing in the trans-series is the manifestation of a renormalon-like phenomenon, due to bubble chains attached to the condensates, as illustrated in e.g. Fig. 6 or Fig. 9.

Although this is probably well-known to many practitioners, it is worth noting that the OPE calculation reconstructs the trans-series more efficiently than a resurgent/renormalon analysis of the perturbative series. In particular, the two-quark condensate leads to a non-ambiguous power correction which is completely invisible in a resurgent analysis. The Borel singularity of the perturbative series detects the presence of a power correction in $m^2/p^2$ due to the four-quark condensate, but since this singularity is essentially a simple pole it misses the full series $\Phi_2(\lambda)$. The OPE calculation is in contrast able to reproduce the series $\Phi_{1,2}(\lambda)$ at all loops. It was suggested in [46] that part of the "blindness" of resurgent analysis in this situation might be due to the restriction to a given order in the $1/N$ expansion. It might happen that at finite $N$ one can have a better access to the four-quark condensate through the resurgent structure of the perturbative series, but the two-quark condensate series will remain undetected, since it transforms non-trivially under the $\mathbb{Z}_2$ chiral symmetry and does not mix with the identity operator [2]. In the language of [46], the trans-series for the self-energy at order $1/N$ does not satisfy the strong version of the resurgence program, since the resurgent analysis of the perturbative series does not make it possible to reconstruct the full trans-series. It does satisfy however the weak version of the program, since the exact result can be obtained by the (lateral) Borel resummation of the trans-series.

Our calculation can be extended in many ways. As we mentioned in the Introduction, the Gross–Neveu model turns out to be a simpler example to study than bosonic sigma models in two dimensions, since in the latter the fields are constrained and in addition there is an infinite number of operators with a fixed scaling dimension. It would be very interesting however to reproduce the large $N$ trans-series obtained in [1] (or the supersymmetric version studied more recently in [15]) by an OPE calculation with condensates similar to the one done here. The results of [14, 15] on non-linear sigma models might be a good starting point for this calculation.

Although going beyond the $1/N$ expansion is analytically difficult, we note that two-point functions can be computed non-perturbatively in integrable models by using form factors. In the case of the non-linear sigma model, it has been checked numerically that the form factor calculation reproduces asymptotically the perturbative series [47, 48]. It would be very interesting to see whether it is possible to detect as well condensate corrections to the perturbative result through form factors.

There have been claims that, in some two-dimensional models, exponentially small corrections associated with renormalons can be given a semiclassical interpretation after a twisted

compactification on the circle, see e.g. [49] for a review and references. The evidence usually given for this is based on an uncontrolled truncation to quantum mechanics, and in our view is not convincing. However, it would be interesting to test these claims by obtaining an exact trans-series for an observable in these models, like the self-energy in the $1/N$ expansion, and try to reproduce it both with a semi-classical approach and an OPE/condensate approach.

Finally, although the results of this paper vindicate the idea that the OPE with vacuum condensates leads to the correct exponentially small corrections to the perturbative series, it is still not clear how the fractional power corrections found in [4] can be reproduced by a first principles calculation. This remains in our view a sharp open problem for our understanding of non-perturbative effects in QFT.

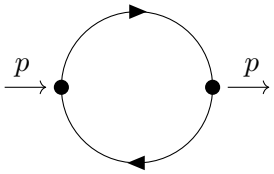

## Acknowledgements

We would like to thank Martin Beneke, Tomás Reis and Marco Serone for very useful discussions and a careful reading of the draft. The work of M.M. has been supported in part by the ERC-SyG project "Recursive and Exact New Quantum Theory" (ReNewQuantum), which received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, grant agreement No. 810573. The work of R.M. is supported by the NKFIH K134946 Grant.


## A Bubbles

### A.1 The fermion polarization loop

The building block of a bubble chain is the fermion polarization loop, given by the following diagram:




(A.1)

We define the corresponding amplitude, with massless fermions, as

$$\Pi(p^2) = \int \frac{\mathrm{d}^d q}{(2\pi)^d} \mathrm{Tr}\left[\frac{1}{\not{q}(\not{p}+\not{q})}\right] = 2\int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{q\cdot p}{q^2(p+q)^2}, \tag{A.2}$$

where we used dimensional regularization to drop a scale-less integral. This and other integrals appearing in this paper can be computed with the master formula

$$\int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{q^{(\mu_1\mu_2\dots\mu_n)}}{(q^2)^r[(p-q)^2]^s} = -\frac{\mathrm{i}}{4\pi}\left(-\frac{p^2}{4\pi}\right)^{-\epsilon/2} \frac{1}{(p^2)^{r+s-1}} p^{(\mu_1\mu_2\dots\mu_n)} \times \frac{\Gamma(1+n-r-\epsilon/2)\Gamma(1-s-\epsilon/2)\Gamma(r+s-1+\epsilon/2)}{\Gamma(r)\Gamma(s)\Gamma(2-r-s+n-\epsilon)}, \tag{A.3}$$

where $q^{(\mu_1\mu_2\dots\mu_n)}$ is the traceless symmetric tensor constructed from $q^{\mu_1}q^{\mu_2}\dots q^{\mu_n}$, see Appendix C in [40] for more details. For example,

$$q^{(\mu)} = q^\mu, \qquad q^{(\mu_1\mu_2)} = q^{\mu_1}q^{\mu_2} - \frac{1}{d} g^{\mu_1\mu_2} q^2, \tag{A.4}$$

but we will only need to compute integrals with one index at most. Using (A.3), we find

$$\Pi(q^2) = \frac{\mathrm{i}}{2\pi}\left(-\frac{q^2}{4\pi}\right)^{-\epsilon/2}\frac{\Gamma\left(1+\frac{\epsilon}{2}\right)\Gamma\left(1-\frac{\epsilon}{2}\right)\Gamma\left(-\frac{\epsilon}{2}\right)}{\Gamma(1-\epsilon)}. \tag{A.5}$$

Let us now briefly consider the massive case. The massive polarization loop is defined as

$$\Pi_{m_0}(p^2) = \int \frac{\mathrm{d}^d q}{(2\pi)^d}\mathrm{Tr}\left[\frac{1}{(\not{q}-m_0)(\not{p}+\not{q}-m_0)}\right]. \tag{A.6}$$

Using standard Feynman integral techniques, one finds in $d = 2-\epsilon$ dimensions the $\epsilon$ expansion (see e.g. [50] for additional details)

$$\Pi_{m_0}(p^2) = -\frac{\mathrm{i}}{\pi\epsilon} + \frac{\mathrm{i}}{2\pi}\log\left(\frac{m_0^2}{4\pi \mathrm{e}^{-\gamma_E}}\right) + \frac{\mathrm{i}}{2\pi}\xi\log\left[\frac{\xi+1}{\xi-1}\right] + \mathcal{O}(\epsilon), \tag{A.7}$$

where $\xi$ was introduced in (3.6). The $\sigma$ propagator in momentum space is closely related to the massive polarization loop:

$$\Delta^{-1}(p;m_0) = \Pi_{m_0}(p^2) - \frac{1}{m_0}\int\frac{\mathrm{d}^d q}{(2\pi)^d}\mathrm{Tr}\left[\frac{1}{\not{q}-m_0}\right], \tag{A.8}$$

and by using (A.7) one finds the expression (3.5).

### A.2 Summing over bubbles

In [18, 21], a powerful and simple technique was introduced to obtain the renormalized perturbative series associated to a chain of bubbles. We now summarize its main ingredients.

Let us consider a generic sum of corrections in the bare 't Hooft coupling $\lambda_0$, of order $1/N$:

$$A = \sum_{n\geq n_0} a_n(p^2)\lambda_0^n, \tag{A.9}$$

where $n_0 \geq 1$. The choice of $\lambda_0$ is such that, at leading order in the $1/N$ expansion, the renormalization function is given by (4.4). Since the quantity we are considering is already of order $1/N$, we only have to use the leading term $Z_\lambda^{(0)}$. Then the renormalized sum is of the form

$$A = \sum_{n\geq n_0}(\nu^2)^{n\epsilon/2}a_n(p^2)\left(Z_\lambda^{(0)}\right)^n\lambda^n. \tag{A.10}$$

Let us assume that one can find a "structure function" $F(x,y)$, which is analytic in both arguments at $x=0$, $y=0$, and satisfying

$$(\nu^2)^{n\epsilon/2}a_n(p^2) = \frac{F(\epsilon,n\epsilon)}{n\epsilon^n}. \tag{A.11}$$

Then the renormalized sum is of the form

$$A = \sum_{n\geq n_0} f_n\lambda^n\left(1+\frac{\lambda}{\epsilon}\right)^{-n}, \tag{A.12}$$

where we have abbreviated

$$f_n = \frac{F(\epsilon, n\epsilon)}{n\epsilon^n}. \tag{A.13}$$

We now expand the factor $(1+\lambda/\epsilon)^{-n}$ in powers of $\lambda$, using the binomial theorem, and we get

$$\begin{aligned} A &= \sum_{n\geq n_0} f_n \lambda^n \sum_{s\geq 0} \binom{n+s-1}{s} (-1)^s \left(\frac{\lambda}{\epsilon}\right)^s = \sum_{m\geq 1} \lambda^m \sum_{s=0}^{m-1} \binom{m-1}{s} \frac{(-1)^s}{\epsilon^s} f_{m-s} \\ &= \sum_{m\geq n_0} \left(\frac{\lambda}{\epsilon}\right)^m \sum_{s=0}^{m-1} \binom{m-1}{s} (-1)^s \sum_{j\geq 0} F_j(\epsilon)(m-s)^{j-1}\epsilon^j \\ &= \sum_{m\geq n_0} \lambda^m \sum_{j\geq 0} \frac{F_j(\epsilon)}{\epsilon^{m-j}} \sum_{s=0}^{m-1} \binom{m-1}{s} (-1)^s (m-s)^{j-1}, \end{aligned} \tag{A.14}$$

where we set $m = n+s$ and we performed the following expansion of the structure function:

$$F(x,y) = \sum_{j\geq 0} F_j(x) y^j. \tag{A.15}$$

We now use that

$$A = \sum_{s=0}^{m-1} \binom{m-1}{s} (-1)^s (m-s)^{j-1} = \begin{cases} \frac{(-1)^{m-1}}{m}, & \text{if } j=0, \\ 0, & \text{if } 1\leq j\leq m-1, \\ (m-1)!, & \text{if } j=m, \end{cases} \tag{A.16}$$

to write the renormalized sum as

$$A = F_0(\epsilon) \sum_{m\geq n_0} \frac{(-1)^{m-1}}{m} \left(\frac{\lambda}{\epsilon}\right)^m + \sum_{m\geq n_0} (m-1)! F_m(\epsilon) \lambda^m + \mathcal{O}(\epsilon). \tag{A.17}$$

The first series can also be expressed as

$$F_0(\epsilon) \sum_{m\geq n_0} \frac{(-1)^{m-1}}{m} \left(\frac{\lambda}{\epsilon}\right)^m = F_0(\epsilon) \log\left(1+\frac{\lambda}{\epsilon}\right) - F_0(\epsilon) \sum_{m=1}^{n_0-1} \frac{(-1)^{m-1}}{m} \left(\frac{\lambda}{\epsilon}\right)^m. \tag{A.18}$$

We now consider the behavior as $\epsilon \to 0$. There is a divergent part encoded in (A.18). We note the useful formula

$$\left[F_0(\epsilon) \log\left(1+\frac{\lambda}{\epsilon}\right)\right]_{\text{div}} = \int_0^\lambda \frac{F_0(-u)}{u+\epsilon} \mathrm{d}u, \tag{A.19}$$

and we can write the total divergent part in (A.18) as

$$[A]_{\text{div}} = \int_0^\lambda \frac{F_0(-u)}{u+\epsilon} \mathrm{d}u - \sum_{m=1}^{n_0-1} \frac{(-1)^{m-1}}{m} \lambda^m \sum_{k=0}^{m-1} F_{0,k} \epsilon^{k-m}, \tag{A.20}$$

where the coefficients $F_{0,k}$ are defined by

$$F_0(\epsilon) = \sum_{k\geq 0} F_{0,k} \epsilon^k. \tag{A.21}$$

The finite part of (A.17) arises from the second sum plus the terms of order $\epsilon^0$ in the first sum:

$$[A]_{\text{finite}} = \sum_{m\geq n_0} \frac{(-1)^{m-1}}{m} F_{0,m}\lambda^m + \sum_{m\geq n_0} (m-1)!F_m(\epsilon)\lambda^m. \tag{A.22}$$

It will be convenient to sum the first series into an integral, yielding

$$[A]_{\text{finite}} = -\int_0^\lambda \frac{F_0(-u) - F_{0,0}}{u}\mathrm{d}u - \sum_{m=1}^{n_0-1} \frac{(-1)^{m-1}}{m} F_{0,m}\lambda^m + \sum_{m\geq n_0} (m-1)!F_m(0)\lambda^m. \tag{A.23}$$

The last sum in the above expression has the form of an inverse Borel transform.

As we have seen, the structure function (A.11) is the relevant object in diagrammatic computations, as it contains all the necessary information to extract both the divergent part (A.20) and the finite part (A.23).

Another useful result for the calculation of renormalization functions is the following. Let $f(\epsilon)$ be analytic at $\epsilon = 0$. Then [17]

$$(\lambda+\epsilon)\frac{\partial}{\partial\lambda}\left[\log\left(1+\frac{\lambda}{\epsilon}\right)f(\epsilon)\right]_{\text{div}} = f(-\lambda). \tag{A.24}$$

This follows from a direct calculation:

$$\begin{aligned}(\lambda+\epsilon)\frac{\partial}{\partial\lambda}\left[\log\left(1+\frac{\lambda}{\epsilon}\right)f(\epsilon)\right]_{\text{div}} &= \sum_{m\geq 1}(-1)^{m-1}(\lambda^m + \epsilon\lambda^{m-1})\sum_{k=0}^{m-1} f_k\epsilon^{k-m} \\ &= \sum_{m\geq 0}(-1)^m f_m\lambda^m.\end{aligned} \tag{A.25}$$

## References


[1] M. Beneke, V. M. Braun and N. Kivel, *The Operator product expansion, nonperturbative couplings and the Landau pole: Lessons from the O(N) sigma model*, *Phys. Lett.* **B443** (1998) 308–316, [hep-ph/9809287].

[2] F. David, *On the Ambiguity of Composite Operators, IR Renormalons and the Status of the Operator Product Expansion*, *Nucl. Phys.* **B234** (1984) 237–251.

[3] F. David, *The Operator Product Expansion and Renormalons: A Comment*, *Nucl. Phys.* **B263** (1986) 637–648.

[4] M. Mariño, R. Miravitllas and T. Reis, *New renormalons from analytic trans-series*, *JHEP* **08** (2022) 279, [2111.11951].

[5] M. Mariño, R. Miravitllas and T. Reis, *Instantons, renormalons and the theta angle in integrable sigma models*, *SciPost Phys.* **15** (2023) 184, [2205.04495].

[6] M. C. Abbott, Z. Bajnok, J. Balog, A. Hegedús and S. Sadeghian, *Resurgence in the $O(4)$ sigma model*, *JHEP* **05** (2021) 253, [2011.12254].

[7] Z. Bajnok, J. Balog and I. Vona, *The full analytic trans-series in integrable field theories*, *Phys. Lett. B* **844** (2023) 138075, [2212.09416].

[8] M. Stingl, *Field theory amplitudes as resurgent functions*, hep-ph/0207349.

[9] M. Shifman, *New and Old about Renormalons: in Memoriam Kolya Uraltsev*, *Int. J. Mod. Phys. A* **30** (2015) 1543001, [1310.1966].

[10] H. D. Politzer, *Effective Quark Masses in the Chiral Limit*, *Nucl. Phys. B* **117** (1976) 397–406.

[11] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *QCD and Resonance Physics. Theoretical Foundations*, *Nucl. Phys. B* **147** (1979) 385–447.

[12] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *QCD and Resonance Physics: Applications*, *Nucl. Phys. B* **147** (1979) 448–518.

[13] V. A. Novikov, M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *Wilson's Operator Expansion: Can It Fail?*, *Nucl. Phys. B* **249** (1985) 445–471.

[14] V. A. Novikov, M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *Two-Dimensional Sigma Models: Modeling Nonperturbative Effects of Quantum Chromodynamics*, *Phys. Rept.* **116** (1984) 103.

[15] D. Schubring, C.-H. Sheu and M. Shifman, *Treating divergent perturbation theory: Lessons from exactly solvable 2D models at large N*, *Phys. Rev. D* **104** (2021) 085016, [2107.11017].

[16] D. J. Gross and A. Neveu, *Dynamical Symmetry Breaking in Asymptotically Free Field Theories*, *Phys. Rev.* **D10** (1974) 3235.

[17] M. Campostrini and P. Rossi, *Dimensional regularization in the 1/N expansion*, *Int. J. Mod. Phys. A* **7** (1992) 3265–3290.

[18] A. Palanques-Mestre and P. Pascual, *The $1/N_f$ Expansion of the $\gamma$ and beta Functions in QED*, *Commun. Math. Phys.* **95** (1984) 277.

[19] T. Alanne and S. Blasi, *The $\beta$-function for Yukawa theory at large $N_f$*, *JHEP* **08** (2018) 081, [1806.06954].

[20] K. Langfeld, L. von Smekal and H. Reinhardt, *Landau pole screening in the Gross-Neveu model: 1/N versus operator product expansion*, *Phys. Lett. B* **362** (1995) 128–133.

[21] A. Palanques-Mestre, *Renormalons in QED*, *Z. Phys. C* **32** (1986) 255.

[22] A. B. Zamolodchikov and A. B. Zamolodchikov, *Factorized S Matrices in Two-Dimensions as the Exact Solutions of Certain Relativistic Quantum Field Models*, *Annals Phys.* **120** (1979) 253–291.

[23] J. A. Gracey, T. Luthe and Y. Schroder, *Four loop renormalization of the Gross-Neveu model*, *Phys. Rev. D* **94** (2016) 125028, [1609.05071].

[24] J. A. Gracey, *Calculation of exponent eta to $O(1/N^2)$ in the $O(N)$ Gross-Neveu model*, *Int. J. Mod. Phys. A* **6** (1991) 395–408.

[25] J. A. Gracey, *Anomalous mass dimension at $O(1/N^2)$ in the $O(N)$ Gross-Neveu model*, *Phys. Lett. B* **297** (1992) 293–297.

[26] J. C. Collins, *Renormalization*, vol. 26 of *Cambridge Monographs on Mathematical Physics*. Cambridge University Press, Cambridge, 1986, 10.1017/CBO9780511622656.

[27] B. Grinstein and L. Randall, *The Renormalization of $G^2$*, *Phys. Lett. B* **217** (1989) 335–340.

[28] D. G. Robertson, *Composite operator renormalization and the trace anomaly*, *Phys. Lett. B* **253** (1991) 143–148.

[29] B. Berg, M. Karowski, V. Kurak and P. Weisz, "Higher order $1/N$ calculations in the Gross-Neveu and nonlinear $\sigma$ models." 1978.

[30] B. Berg, M. Karowski, V. Kurak and P. Weisz, *Scattering Amplitudes of the Gross-Neveu and Nonlinear $\sigma$ Models in Higher Orders of the 1/N Expansion*, *Phys. Lett. B* **76** (1978) 502–504.

[31] P. Forgacs, F. Niedermayer and P. Weisz, *The Exact mass gap of the Gross-Neveu model. 2. The 1/N expansion*, *Nucl. Phys.* **B367** (1991) 144–157.

[32] M. Campostrini and P. Rossi, *The 1/N expansion of two-dimensional spin models*, *Riv. Nuovo Cim.* **16** (1993) 1–111.

[33] P. Biscari, M. Campostrini and P. Rossi, *Quantitative Picture of the Scaling Behavior of Lattice Nonlinear $\sigma$ Models From the 1/N Expansion*, *Phys. Lett. B* **242** (1990) 225–233.

[34] P. Forgacs, F. Niedermayer and P. Weisz, *The Exact mass gap of the Gross-Neveu model. 1. The Thermodynamic Bethe ansatz*, *Nucl. Phys.* **B367** (1991) 123–143.

[35] M. Mariño, R. Miravitllas and T. Reis, *On the structure of trans-series in quantum field theory*, 2302.08363.

[36] M. Mariño, *Lectures on non-perturbative effects in large N gauge theories, matrix models and strings*, *Fortsch. Phys.* **62** (2014) 455–540, [1206.6272].

[37] M. Mariño, *Instantons and large N. An introduction to non-perturbative methods in quantum field theory*. Cambridge University Press, 2015.

[38] T. M. Seara and D. Sauzin, *Resumació de Borel i teoria de la ressurgencia*, *Butl. Soc. Catalana Mat.* **18** (2003) 131–153.

[39] I. Aniceto, G. Başar and R. Schiappa, *A primer on resurgent transseries and their asymptotics*, *Phys. Rep.* **809** (2019) 1–135.

[40] P. Pascual and R. Tarrach, *QCD: renormalization for the practicioner*, vol. 194 of *Lecture Notes in Physics*. Springer-Verlag, 1984.

[41] P. Pascual and E. de Rafael, *Gluonic Corrections to Quark Vacuum Condensate Contributions to Two Point Functions in QCD*, *Z. Phys. C* **12** (1982) 127.

[42] V. Elias, *Quantum Field Theory and the Dynamical Generation of Quark Masses*, *Can. J. Phys.* **64** (1986) 595–601.

[43] T. Huang and Z. Huang, *Quantum Chromodynamics in Background Fields*, *Phys. Rev. D* **39** (1989) 1213–1220.

[44] S. Giombi, V. Kirilin and E. Skvortsov, *Notes on Spinning Operators in Fermionic CFT*, *JHEP* **05** (2017) 041, [1701.06997].

[45] M. Beneke and V. M. Braun, *Heavy quark effective theory beyond perturbation theory: Renormalons, the pole mass and the residual mass term*, *Nucl. Phys. B* **426** (1994) 301–343, [hep-ph/9402364].

[46] L. Di Pietro, M. Mariño, G. Sberveglieri and M. Serone, *Resurgence and 1/N Expansion in Integrable Field Theories*, *JHEP* **10** (2021) 166, [2108.02647].

[47] J. Balog, *Form-factors and asymptotic freedom in the O(3) sigma model*, *Phys. Lett. B* **300** (1993) 145–151.

[48] J. Balog, M. Niedermaier and T. Hauer, *Perturbative versus nonperturbative QFT: Lessons from the O(3) NLS model*, *Phys. Lett. B* **386** (1996) 224–232, [hep-th/9604161].

[49] G. V. Dunne and M. Ünsal, *What is QFT? Resurgent trans-series, Lefschetz thimbles, and new exact saddles*, *PoS* **LATTICE2015** (2016) 010, [1511.05977].

[50] H. Kleinert, *Particles and quantum fields*. World Scientific, 2016.

# Infinite Order Hydrodynamics: an Analytical Example

L. Gavassino
*Department of Mathematics, Vanderbilt University, Nashville, TN, USA*

We construct a kinetic model for matter-radiation interactions where the hydrodynamic gradient expansion can be computed analytically up to infinite order in derivatives, in the fully non-linear regime, and for arbitrary flows. The frequency dependence of the opacity of matter is chosen to mimic the relaxation time of a self-interacting scalar field. In this way, the transient sector simulates that of a realistic quantum field theory. As expected, the gradient series is divergent for most flows. We identify the mechanism at the origin of the divergence, and we provide a successful regularization scheme. Additionally, we propose a universal qualitative framework for predicting the breakdown of the gradient expansion of an arbitrary microscopic system undergoing a given flow. This framework correctly predicts the factorial divergence of the gradient expansion in most non-linear flows and its breakdown due to stochastic fluctuations. It also predicts that jets may induce an ultraviolet divergence in the gradient expansion of quark matter hydrodynamics.

arXiv:2402.19343v1 [nucl-th] 29 Feb 2024

*Introduction* - The most pressing open question in relativistic fluid mechanics is: "How far can we push hydrodynamics before it breaks down?" [1–20]

Let us make this question precise. The Knudsen number $\mathrm{Kn} = \lambda/L$ is the ratio between the particles' mean free path $\lambda$ (defining the microscopic scale) and the characteristic length scale $L$ of the flow of interest [21, 22]. It is common knowledge [23] that, if $\mathrm{Kn} \to 0$, we can model the system as an ideal fluid, while when $\mathrm{Kn} \gtrsim 1$, we need to rely on microphysics. Thus, there is some value of $\mathrm{Kn} \in (0, 1]$ at which hydrodynamics stops working. Can we identify it precisely?

The answer depends on which hydrodynamic theory one is using. The ideal fluid breaks down at any finite value of Kn, since it is non-dissipative, and it predicts that waves survive forever. These issues are fixed in Navier-Stokes theory, whose breakdown is conventionally set at $\mathrm{Kn} = 0.1$ [24]. However, one may try to do even better. Most derivations of hydrodynamics from kinetic theory [25, 26] and holography [27, 28] express the stress-energy tensor $T^{\mu\nu}$ of a fluid as a formal power series in Kn, known as the "gradient expansion" [29],

$$T^{\mu\nu}(\mathrm{Kn}) = T^{\mu\nu}_{\mathrm{ID}} + T^{\mu\nu}_{1}\mathrm{Kn} + T^{\mu\nu}_{2}\mathrm{Kn}^2 + \dots\,, \qquad (1)$$

where the zeroth order is the ideal fluid, the first order is Navier-Stokes, the second order is the Burnett equation [30], and higher orders correspond to fluid theories with higher derivative corrections. One hopes that, by adding more and more powers of Kn, the regime of applicability of hydrodynamics will expand more and more, increasing the accuracy of hydrodynamics for all Kn up to the radius of convergence $\mathrm{Kn}^\star$ of the series (1). Following this line of thought, it would then be natural to conclude that such radius $\mathrm{Kn}^\star$ is what ultimately marks the rigorous breakdown of hydrodynamics [31].

Unfortunately, reality turns out to be more complicated. First, there are indications that, in most realistic scenarios, $\mathrm{Kn}^\star = 0$ [10, 13, 32]. This means that, if we keep adding higher and higher orders in Kn, at some point the region of applicability of (1) shrinks instead of expanding. Secondly, equation (1) makes sense only if the function $T^{\mu\nu}(\mathrm{Kn})$ is analytic in Kn. In principle, there may be non-smooth corrections like $\mathrm{Kn}^{3/2}$ [33], or non-perturbative corrections like $e^{-1/\mathrm{Kn}}$ [34]. Therefore, the breakdown scale of hydrodynamics remains unknown.

The dream of any fluid theorist would be to have a microscopic model where $T^{\mu\nu}(\mathrm{Kn})$ can be computed exactly for arbitrary flow, at arbitrary Kn, in the fully non-linear regime, and where all the terms $T^{\mu\nu}_n$ in the series (1) have exact analytical formulas. Such model should have realistic interactions, to ensure that the dependence of $T^{\mu\nu}$ on Kn mimics the behavior of some microscopic quantum field theory. Finally, one should be able to extract from it general lessons about the expansion (1). In this Letter, we provide a model that fulfills all these requirements.

Notation: We work in Minkowski space, with signature $(-,+,+,+)$, and adopt natural units: $c = \hbar = k_B = 1$.

*The kinetic model* - Our model setup is a radiation-hydrodynamic system [35–41], namely a fluid mixture comprised of two substances: a material medium $M$ with negligibly short mean free path (i.e. $\mathrm{Kn}_M \equiv 0$), and a diluted radiation gas $R$ with a finite, possibly large, mean free path (i.e. $\mathrm{Kn}_R > 0$) [42, 43]. The former is modeled as an ideal fluid, with a well-defined temperature $T(x^\alpha)$ and flow velocity $u^\mu(x^\alpha)$. The latter can exist in far-from-equilibrium states, and we must track its kinetic distribution function $f_p(x^\alpha)$ [44, 45] ($p$ = momentum). The total stress-energy tensor is the sum of a material non-viscous part and a radiation part:

$$T^{\mu\nu} = T^{\mu\nu}_M + \int \frac{d^3p}{(2\pi)^3 p^0}\, f_p\, p^\mu p^\nu\,. \qquad (2)$$

The radiation particles do not interact with each other, but are randomly absorbed and emitted by the medium (neglecting scattering). Thus, the Boltzmann equation is of relaxation type [35, 46]. Taking the relaxation time to be $\tau_p = -u_\mu p^\mu/g$ (with $g > 0 = \text{constant}$), we have

$$p^\mu \partial_\mu f_p = -u_\mu p^\mu\, \frac{f^{\mathrm{eq}}_p - f_p}{\tau_p} = g\,(f^{\mathrm{eq}}_p - f_p)\,. \qquad (3)$$

Since $\tau_p$ grows linearly with $-u_\mu p^\mu$, our medium is opaque at low frequencies, and transparent at high frequencies [35]. The function $f^{\mathrm{eq}}_p$ is the value that $f_p$ would

have if the radiation was in local equilibrium, with the temperature $T$ and flow velocity $u^\mu$ of the medium. Depending on the statistics, one has $f_p^{\rm eq} = (e^{-u_\mu p^\mu/T} - b)^{-1}$, $b \in \{0, \pm 1\}$ [47]. The backreaction of the radiation on the motion of the medium follows from the conservation law $\partial_\mu T^{\mu\nu} = 0$ which, combined with (2) and (3), gives

$$\partial_\mu T_M^{\mu\nu} = -g \int \frac{d^3p}{(2\pi)^3 p^0} \, (f_p^{\rm eq} - f_p) \, p^\nu \, . \qquad (4)$$

*Why this model?* - The relaxation time $\tau_p \propto$ "energy" mimics the transient behavior of a scalar field theory with $\phi^4$ interaction [48]. This choice ensures that the gradient expansion (1) will exhibit realistic scaling laws with powers of $\mathrm{Kn}_R = \max\{\mathrm{Kn}_M, \mathrm{Kn}_R\} = \mathrm{Kn}$.

Our main reason for relying on radiation hydrodynamics (rather than dealing with ideal gases) is that it allows the introduction of momentum-dependent relaxation times, as in (3), without violating any conservation law [49]. Furthermore, within radiation hydrodynamics, the relaxation-time approximation is *exact* (by Kirchhoff's law [50]), if the medium has negligibly short equilibration timescale [35, 36]. By contrast, in ideal gases, equations like (3) are a pretty crude approximation of the Boltzmann collision integral.

We remark that the fields $T$ and $u^\mu$ have unambiguous meanings in our model, as they characterize the actual state of the medium (which is in local equilibrium), establishing a natural "hydrodynamic frame" [29]. Indeed, providing initial data for the "matter+radiation" system requires specifying the initial states for both the radiation gas and the medium separately, so that $\{f_p, T, u^\mu\}$ are *independent* dynamical degrees of freedom. This is convenient, since we don't need to worry about matching conditions [25] (i.e., no integral constraint relates $T$ and $u^\mu$ to $f_p$ at a given time).

*Linear response properties* - Linearizing equations (2) and (3), and looking for solutions $\propto e^{ik_\mu x^\mu}$, we obtain

$$\delta T^{\mu\nu} = \delta T_M^{\mu\nu} + \int \frac{d^3p}{(2\pi)^3 p^0} \, \frac{(f_{p,T}^{\rm eq} \delta T + f_{p,u^\rho}^{\rm eq} \delta u^\rho) \, p^\mu p^\nu}{1 + i g^{-1} p^\alpha k_\alpha} \, . \qquad (5)$$

For $k^\alpha = (\omega, 0, 0, 0)$, the integrand is singular whenever $\omega = -ig/p^0$, which approaches $\omega = 0$ for $p^0 \to \infty$. Thus, the linear-response Green function of the stress-energy tensor has a branch cut that touches the origin of the complex $\omega-$plane. This same feature is expected to be shared by $\phi^4$ kinetic theory [51], confirming that the analytic structure of our model is "realistic" from a quantum field theory perspective.

*Analytic solution* - Equations like (3) can be solved analytically in the non-linear regime [35]. Fixed an event $x^\alpha$ and a momentum $p^\alpha$, consider the worldline

$$x^\alpha(\xi) = x^\alpha - \frac{p^\alpha}{g} \xi \, , \qquad (6)$$

where $\xi$ is the optical depth. Travelling along (6), at fixed momentum $p^\alpha$, equation (3) becomes an ordinary differential equation, $-\dot{f}_p(\xi) + f_p(\xi) = f_p^{\rm eq}(\xi)$, which has unique solution for initial data at past infinity:

$$f_p(x^\alpha) = \int_0^{+\infty} e^{-\xi} f_p^{\rm eq}\left(x^\alpha - \frac{p^\alpha}{g}\xi\right) d\xi \, . \qquad (7)$$

Plugging (7) in (2), we obtain an *exact* formula for the total stress-energy tensor:

$$T^{\mu\nu} = T_M^{\mu\nu} + \int \frac{d^3p}{(2\pi)^3 p^0} \, p^\mu p^\nu \int_0^{+\infty} f_p^{\rm eq}\left(x^\alpha - \frac{p^\alpha}{g}\xi\right) e^{-\xi} \, d\xi \, . \qquad (8)$$

This is the function "$T^{\mu\nu}(\mathrm{Kn})$" we were looking for. It remains valid at all orders in derivatives of the fluid fields $\{T, u^\mu\}$, which enter (8) through $f_p^{\rm eq} = (e^{-u_\mu p^\mu/T} - b)^{-1}$. We may interpret (8) as an "infinite-order hydrodynamic constitutive relation", where $T^{\mu\nu}$ is a non-local functional of $T$ and $u^\mu$, see figure 1.

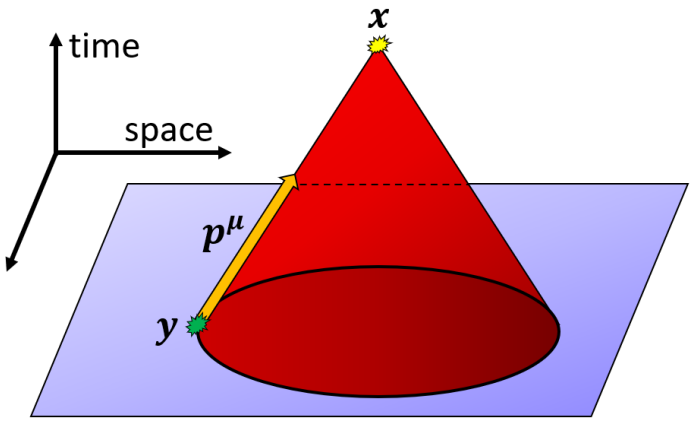



FIG. 1. Visual representation of equation (8), viewed as a non-local constitutive relation $T^{\mu\nu}[T, u^\alpha]$. The value of $T^{\mu\nu}(x)$ receives contributions from *all* events $y$ on the past lightcone of $x$ (although closer events contribute more). The information about $\{T(y), u^\alpha(y)\}$ is transported directly to $x$ by radiation particles emitted at $y$, which have probability $e^{-\xi} = e^{-g(x^0 - y^0)/p^0}$ to reach $x$ before being absorbed.

The ideal fluid is recovered in the infinite-opacity limit ($g \to +\infty$), since $x^\alpha(\xi) \to x^\alpha$, and the total stress-energy tensor $T^{\mu\nu}$ becomes that of a conglomerate fluid, where matter and radiation comove with velocity $u^\mu$ and both have temperature $T$:

$$\begin{aligned} T_{\rm ID}^{\mu\nu} &= T_M^{\mu\nu} + \int \frac{d^3p}{(2\pi)^3 p^0} \, f_p^{\rm eq} \, p^\mu p^\nu \\ &= (\varepsilon + P) u^\mu u^\nu + P \eta^{\mu\nu} \, . \end{aligned} \qquad (9)$$

The energy density $\varepsilon(T, ...)$ and the pressure $P(T, ...)$ are the sum of matter and radiation parts in thermal equilibrium, where the matter contributions may depend on additional state variables besides $T$, including, possibly, some chemical potential.

*Gradient series* - Let us evaluate the derivative expansion of (8). If $f_p^{\rm eq}(x^\mu)$ is an entire function of $x^\mu$, we can express the integrand of (7) as a Taylor series:

$$f_p^{\rm eq}(\xi) = \sum_{n=0}^{+\infty} \frac{\xi^n}{n!} \left(\frac{d}{d\xi}\right)^n f_p^{\rm eq}(0)\,, \tag{10}$$

where $d/d\xi = -g^{-1}p^\alpha\partial_\alpha$ is the derivative along the parametric line (6). We note that $d/d\xi \sim {\rm Kn}$, due to the following chain of estimates ($L$ is the hydrodynamic scale):

$$\frac{d}{d\xi} = -\frac{p^\alpha\partial_\alpha}{g} \sim \frac{T}{gL} \sim \frac{-u^\mu p_\mu}{gL} = \frac{\tau_p}{L} \sim \frac{\lambda}{L} = {\rm Kn}. \tag{11}$$

Hence, each $n^{\rm th}$ term in the series (10) corresponds to the power ${\rm Kn}^n$ in the Knudsen expansion. Plugging (10) into (8), and performing the integral in $\xi$, we can express the total stress-energy tensor as the sum of the ideal part (9) and a viscous part, namely $T^{\mu\nu} = T^{\mu\nu}_{\rm ID} + \Pi^{\mu\nu}$, with

$$\Pi^{\mu\nu} = \sum_{n=1}^{+\infty} (-g)^{-n} \partial_{\alpha_1}...\partial_{\alpha_n} \int \frac{d^3p}{(2\pi)^3 p^0}\, f_p^{\rm eq}\, p^\mu p^\nu p^{\alpha_1}...p^{\alpha_n}\,. \tag{12}$$

This is the gradient series we were looking for. Again, no approximation was made, besides the assumptions of the kinetic model. The integrals in (12) are the moments of the equilibrium distribution function $f_p^{\rm eq}$, all of which can be evaluated analytically.

Let us discuss some notable properties of (12):

- All the components of $\Pi^{\mu\nu}$ are in general non-zero, including $\Pi^{\mu\nu}u_\mu u_\nu$. Thus, the "material frame", where temperature and flow velocity are those of the medium, does not coincide with the Landau or the Eckart frame, as noted in [43].
- All the contributions to (12) contain factors of the form $\partial_{(\mu_1}\partial_{\mu_2}...\partial_{\mu_k}\beta_{\nu)}$, with $\beta_\nu = u_\nu/T$. It follows that $\Pi^{\mu\nu}$ vanishes whenever $\beta_\nu$ is a Killing vector, namely $\partial_{(\mu}\beta_{\nu)} = 0$, which is the condition of global thermodynamic equilibrium [52–55].
- If we truncate $\Pi^{\mu\nu}$ at finite order $n \geq 1$ (with Maxwell-Boltzmann statistics, i.e. $f_p^{\rm eq}{=}e^{\beta_\rho p^\rho}$), the principal part of the conservation law $\partial_\mu T^{\mu\nu} = 0$ (regarded as a dynamical equation for $\beta_\nu$) is

$$\frac{1}{(-g)^n} \int \frac{f_p^{\rm eq}\, d^3p}{(2\pi)^3 p^0}\; p^\mu p^\nu p^\rho p^{\alpha_1}...p^{\alpha_n}\partial_{\alpha_1}...\partial_{\alpha_n}\partial_\mu\beta_\rho\,. \tag{13}$$

The associated characteristic polynomial [56] is

$$\mathcal{P}(\zeta_\alpha) = \det\left[\int \frac{f_p^{\rm eq}\, d^3p}{(2\pi)^3 p^0}\, (p^\alpha\zeta_\alpha)^{n+1} p^\nu p^\rho\right]. \tag{14}$$

For $n$ odd, the matrix in the determinant is positive definite for all $\zeta_\alpha$. Thus, the evolution equation is elliptic, making all odd-order truncations acausal, unstable, and ill-posed [57–59].

*Divergence of the gradient series* - Let us show that (12) is usually divergent. We can estimate the order of magnitude of each term of the series by making the formal replacements $\partial_\alpha \to 1/L$ and $f_p^{\rm eq} \to e^{-p^0/T}$. Then, equation (12) becomes (ignoring constant factors)

$$\frac{\Pi^{\mu\nu}}{T^4} \sim \sum_n \left(\frac{T}{gL}\right)^n (n{+}3)!\,, \tag{15}$$

which is factorially divergent.

Let us now repeat the above calculation, assuming that the medium cannot produce radiation particles whose energy exceeds some large cutoff scale $\Lambda \gg T$, thereby restricting integrals in $d^3p$ to the sphere $p^0 \leq \Lambda$. Then, the series (15) becomes

$$\frac{\Pi^{\mu\nu}}{T^4} \sim \sum_n \left(\frac{T}{gL}\right)^n \gamma(n{+}4, \Lambda/T)\,, \tag{16}$$

where $\gamma$ is the lower incomplete Gamma function, which becomes $(n+3)!$ if we send $\Lambda \to +\infty$ at fixed $n$. Instead, let us evaluate the series for *finite* $\Lambda \gg T$. At large $n$, the addends scale like

$$\left(\frac{T}{gL}\right)^n \gamma(n{+}4, \Lambda/T) \sim \left(\frac{\Lambda}{gL}\right)^n \frac{e^{-\Lambda/T}(\Lambda/T)^4}{n}\,. \tag{17}$$

If $gL > \Lambda$, the series converges. Furthermore, the large-$n$ part of the series adds up to

$$-e^{-\Lambda/T}\left(\frac{\Lambda}{T}\right)^4 \ln\left(1 - \frac{\Lambda}{gL}\right), \tag{18}$$

which is negligible in the limit $\Lambda \gg T$.

The above argument shows that the divergence of the gradient expansion comes from fictitious particles that carry unreasonably large energy. We can prove that such particles are unphysical by showing that the exact formula (8) is unaffected by the cutoff $\Lambda$. Called $T^{\mu\nu}_\Lambda$ the regularised stress-energy tensor, we have

$$\begin{aligned}\frac{T^{\mu\nu}{-}T^{\mu\nu}_\Lambda}{T^4} &= \int_{p^0\geq\Lambda} \frac{d^3p}{(2\pi)^3 p^0}\, \frac{p^\mu p^\nu}{T^4} \int_0^{+\infty} f_p^{\rm eq}(\xi)\, e^{-\xi}\, d\xi \\ &\sim \int_{p^0\geq\Lambda} \frac{d^3p}{(2\pi)^3}\, \frac{p^0 e^{-p^0/T}}{T^4} \sim \left(\frac{\Lambda}{T}\right)^3 e^{-\Lambda/T}\,,\end{aligned} \tag{19}$$

which tends to zero when $\Lambda{\gg}T$. This estimate shows that the cutoff $\Lambda$ can be used as a practical tool, to regularize the gradient expansion (12), and restore agreement with (8). One only needs to make sure that $gL > \Lambda \gg T$. Unfortunately, this also requires ${\rm Kn} = T/(gL) \ll 1$. This suggests that adding higher derivative terms may not significantly expand the regime of applicability of hydrodynamics. Instead, it may just lead to a refinement of the accuracy of hydrodynamics, within the same interval $0 \leq {\rm Kn} \lesssim 0.1$ where Navier-Stokes is already valid. Indeed, several previous studies seem to support the present conclusion [16, 30, 60, 61].

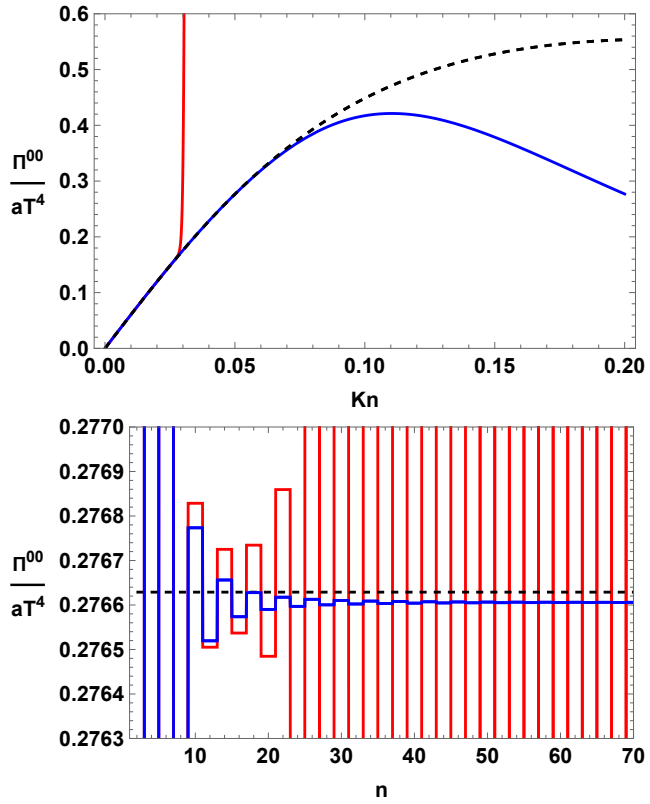



FIG. 2. Upper panel: Exact viscous density $\Pi^{00}(t{=}0)$ according to (20) (dashed) compared to the non-regularised gradient series (21) (red) truncated at $n{=}70$, and the regularised series (22) (blue) with $n\to\infty$ and cutoff scale $\Lambda/T=(\mathrm{Kn}{+}0.01)^{-1}$. Lower panel: Convergence test at Kn=0.05 (same colors).

*A quick example* - We can check the consistency of the estimates (18) and (19) with an example. Take $u^\mu=(1,0,0,0)$ and $T(t)\propto 1-a\sin(t/L)$, with $a\to 0$. Linearising in $a$, and defining $\mathrm{Kn}=T/(gL)$, the value of $\Pi^{00}(0)$ according to (8) and (12) is given by, respectively,

$$\text{Exact:}\qquad \frac{\Pi^{00}}{aT^4}=\int_0^{+\infty}\frac{dz}{2\pi^2}\frac{\mathrm{Kn}\,z^5e^{-z}}{1+\mathrm{Kn}^2\,z^2}\,, \tag{20}$$

$$\text{Expand:}\qquad \frac{\Pi^{00}}{aT^4}=\sum_{n\ \text{odd}}\frac{(-1)^{\frac{n-1}{2}}}{2\pi^2}\mathrm{Kn}^n(n{+}4)!\,, \tag{21}$$

$$\text{Cutoff:}\qquad \frac{\Pi^{00}}{aT^4}=\sum_{n\ \text{odd}}\frac{(-1)^{\frac{n-1}{2}}}{2\pi^2}\mathrm{Kn}^n\gamma\left(n{+}5,\frac{\Lambda}{T}\right), \tag{22}$$

where equation (22) is the same as (21), corrected with cutoff $\Lambda/T<\mathrm{Kn}^{-1}$ (for convergence of the series). We observe in figure 2 the breakdown of the gradient expansion at order 70 (higher orders explode at smaller Kn). The cutoff $\Lambda$ restores agreement with (8) at all orders, within an error that survives at large $n$ and increases with Kn. This fixes the breakdown of the regularised gradient series at $\mathrm{Kn}\sim 0.1$, which is precisely when Navier-Stokes ceases to hold [24].

*Origin of the divergence* - We now provide a general explanation for the divergence of the gradient expansion, which applies to the present model and to most other models in the literature.

Let $A$ be a density of interest, located at $x$. By causality, the event $x$ receives influence from all events $y$ in the causal past $J^-(x)$ of $x$ [62]. Information is transported from $y$ to $x$ by messengers (e.g. particles or excitations), each having an associated energy (or frequency) $E$. Thermalization erases the messenger's memory about $y$ over a timescale $\tau(E)$. Thus, we can schematically write

$$A=\int_0^{+\infty}dE\int_{J^-(x)}d^4y\;\sigma_A(E,y)e^{-(x^0-y^0)/\tau(E)}\,, \tag{23}$$

where $\sigma_A(E,y)$ quantifies how many messengers with energy $E$ are emitted from $y$, and "how much $A$" they carry. The Knudsen expansion (10) corresponds, in this general setting, to a Taylor expansion of $\sigma_A(E,y)$ around $x$. Since, in the lightcone, $|x^j-y^j|\le x^0-y^0$, we may write, schematically (recall: $\partial_\alpha\to 1/L$),

$$\sigma_A(E,y)\sim\sum_n\frac{\sigma_A(E,x)}{n!}\left(\frac{x^0-y^0}{L}\right)^n. \tag{24}$$

Plugging (24) into (23), and taking the series outside the volume integral, we obtain, up to $n-$independent factors,

$$A\sim\int_0^{+\infty}dE\,\sigma_A(E,x)\sum_n\left[\frac{\tau(E)}{L}\right]^{n+1}. \tag{25}$$

Now we see the problem: If $\tau(E)$ is unbounded above, fixed any $L$, there is some energy $E$ such that $\tau(E)>L$, and the geometric series in (25) diverges for such $E$. This implies that we were not allowed to exchange the series in $n$ with the integral in $d^4y$, and in general [63]

$$A\Big[\sum_{n=0}^{+\infty}\sigma_{A,n}\Big]\neq\sum_{n=0}^{+\infty}A[\sigma_{A,n}]\,. \tag{26}$$

This is precisely the issue with our analytical model. The relaxation time $\tau_p=-u_\mu p^\mu/g$ diverges at high energies, and we cannot take the series (10) outside the integral (8). The introduction of the energy cutoff resolves the problem because, if $\Lambda/T<\mathrm{Kn}^{-1}$, the series (25) converges for all $E\le\Lambda$, since (see figure 1)

$$\frac{\tau(E)}{L}=\frac{E}{gL}\le\frac{\Lambda}{gL}=\frac{\Lambda}{T}\,\mathrm{Kn}<1\,. \tag{27}$$

The same conclusions apply to all kinetic models of QCD plasmas [64–66], whose $\tau(E)$ ($\propto E^q$, with $q{>}0$) diverges at large $E$ (i.e. for jets). According to (25), the gradient series of all such quark matter models has a jet-induced divergence, since, fixed $L$, $\tau(E)\gg L$ when $E\to\infty$.

With similar reasoning, we can also predict the breakdown of all gradient expansions due to stochastic fluctuations [33]. In fact, long-wavelength hydrodynamic fluctuations can transport information from $y$ to $x$, and have decay timescale $\tau(E)\sim E^{-2}$ (with $E$=frequency), which diverges at small $E$. This generates an infrared (i.e. low $E$) divergence in (25), consistently with [33].

*General flows* - In the above estimates, we always assumed that $(\partial)^n \to 1/L^n$. This means that our results apply to sinusoidal and exponential flows, or finite superpositions thereof. However, different flows attribute different weights to each derivative. In Bjorken expansion [67], where everything scales like $(y^0)^{-1}$, the correct formal replacement is $(\partial)^n \to n!/L^n$, and (25) becomes

$$A \sim \int_0^{+\infty} dE\, \sigma_A(E,x) \sum_n \left[\frac{\tau(E)}{L}\right]^{n+1} n!\,, \tag{28}$$

which diverges independently from the value of $\tau(E)$. This explains why Israel-Stewart hydrodynamics has divergent gradient expansion in Bjorken flow [7], despite having only one relaxation time. We also understand why Borel summation gives correct predictions in Bjorken flow [3]. By Borel-summing the series (28), we obtain[1]

$$\sum_{n=0}^{+\infty} \left[\frac{\tau(E)}{L}\right]^{n+1} n! \;\overset{\mathcal{B}}{=}\; \text{P.V.}\int_0^{+\infty} \frac{e^{-t/\tau(E)}}{L-t} dt\,, \tag{29}$$

which has precisely the form of the spacetime integral in (23), with $\sigma_A(y)$ that decays like $1/y^0$. This suggests that Borel summation "undoes" the gradient expansion, whose existence wasn't justified in the first place, and it reconstructs the original (correct) integral (23) from which the gradient expansion was (improperly) derived.

The same reasoning can be generalized to arbitrary flows. For example, suppose that the Taylor expansion of $\sigma_A(y)$ has finite radius of convergence $\mathcal{R}$. Then, for large $n$, we have, schematically, $(\partial)^n \to n!/\mathcal{R}^n$ [68], and

$$A \sim \int_0^{+\infty} dE\, \sigma_A(E,x) \sum_n \left[\frac{\tau(E)}{\mathcal{R}}\right]^{n+1} n!\,. \tag{30}$$

This explains why, in most realistic flows, where $T(x^\alpha)$ and $u^\mu(x^\alpha)$ are not entire functions, the gradient expansion is factorially divergent [32].

*Conclusions* - Our analysis reveals what infinite-order hydrodynamics looks like. The stress-energy tensor is still a functional of the fluid fields $\{T, u^\alpha\}$, just like in finite-order hydrodynamics. However, such a functional $T^{\mu\nu}[T, u^\alpha]$ becomes *non-local*. By causality, $T^{\mu\nu}(x)$ depends solely on the restriction of $T$ and $u^\alpha$ within the past lightcone of $x$. Furthermore, by dissipation, it is mostly affected by past events that are close to $x$, as it loses memory of the far past.

Starting from these observations, we proposed a simple explanation of why the gradient series (1) often diverges: In a nutshell, the construction of the gradient series amounts to interchanging the Taylor series of the fields $\{T, u^\alpha\}$ with the spacetime integral defining $T^{\mu\nu}[T, u^\alpha]$. This delicate exchange of limits is often not allowed. In fact, the $n^{\text{th}}$ term of a Taylor expansion diverges like $\sim t^n$ at past infinity, and the usual theorems for exchanging limits and integrals (like the dominated convergence theorem [63]) do not apply. In some cases, Borel resummation "brings the series back inside $T^{\mu\nu}$", and it reconstructs the correct functional $T^{\mu\nu}[T, u^\alpha]$.

So, when does hydrodynamics break down? From our analytical example, one can give two answers, depending on the definition of the word "hydrodynamics".

**Definition 1 (Broad)** *Hydrodynamics is the ability to express the conserved currents in terms of the fluid variables alone, e.g.* $T^{\mu\nu} = T^{\mu\nu}[T, u^\alpha]$.

Using this definition, we found that hydrodynamics never breaks down, at any Kn, provided that we know the near past of the fluid ("near past" ≈ "3 mean free paths ago"). Unfortunately, according to definition 1, solving hydrodynamics amounts to solving $\partial_\mu T^{\mu\nu}[T, u^\alpha] = 0$, which is, in general, a convoluted integrodifferential equation.

**Definition 2 (Narrow)** *Hydrodynamics is the ability to express the conserved currents as local functions of the fluid fields and a finite number of derivatives thereof, e.g.* $T^{\mu\nu} = T^{\mu\nu}(T, u^\alpha, \partial T, \partial u^\alpha, ..., \partial^n T, \partial^n u^\alpha)$.

In our model, nature conspires to break the validity of the whole gradient expansion at the same breakdown scale of Navier-Stokes (Kn $\sim 0.1$). Thus, higher derivative corrections do not expand the applicability of the gradient expansion beyond first order.

## ACKNOWLEDGEMENTS

This work was partially supported by a Vanderbilt Seeding Success Grant.

[1] J. Noronha-Hostler, J. Noronha, and M. Gyulassy, Nucl. Phys. A **956**, 890 (2016), arXiv:1512.07135 [nucl-th].

[2] M. Strickland, Acta Phys. Polon. B **45**, 2355 (2014), arXiv:1410.5786 [nucl-th].

[3] M. P. Heller and M. Spaliński, Phys. Rev. Lett. **115**, 072501 (2015).

[4] M. Alqahtani, M. Nopoush, and M. Strickland, Prog. Part. Nucl. Phys. **101**, 204 (2018), arXiv:1712.03282 [nucl-th].

[5] P. Romatschke, JHEP **12**, 079 (2017), arXiv:1710.03234 [hep-th].

[1] The principal value guarantees that the (unphysical) singularity at $t = L$ gives a finite contribution to the integral. When $L \gg \tau(E)$, we can neglect it, and the integral can be effectively stopped at $L - \tau(E)$.

[6] P. Romatschke and U. Romatschke, arXiv e-prints , arXiv:1712.05815 (2017), arXiv:1712.05815 [nucl-th].
[7] P. Romatschke, Phys. Rev. Lett. **120**, 012301 (2018).
[8] W. Florkowski, M. P. Heller, and M. Spaliński, Reports on Progress in Physics **81**, 046001 (2018), arXiv:1707.02282 [hep-ph].
[9] G. S. Denicol and J. Noronha, Phys. Rev. Lett. **124**, 152301 (2020), arXiv:1908.09957 [nucl-th].
[10] M. P. Heller, A. Serantes, M. Spaliński, V. Svensson, and B. Withers, Phys. Rev. D **104**, 066002 (2021).
[11] E. R. Most, J. Noronha, and A. A. Philippov, Mon. Not. Roy. Astron. Soc. **514**, 4989 (2022), arXiv:2111.05752 [astro-ph.HE].
[12] E. R. Most, S. P. Harris, C. Plumberg, M. G. Alford, J. Noronha, J. Noronha-Hostler, F. Pretorius, H. Witek, and N. Yunes, Mon. Not. Roy. Astron. Soc. **509**, 1096 (2021), arXiv:2107.05094 [astro-ph.HE].
[13] M. P. Heller, A. Serantes, M. Spaliński, V. Svensson, and B. Withers, Phys. Rev. X **12**, 041010 (2022).
[14] C. Plumberg, D. Almaalol, T. Dore, J. Noronha, and J. Noronha-Hostler, Phys. Rev. C **105**, L061901 (2022).
[15] N. Mullins, G. S. Denicol, and J. Noronha, Phys. Rev. D **106**, 056024 (2022).
[16] D. Wagner and L. Gavassino, Phys. Rev. D **109**, 016019 (2024), arXiv:2309.14828 [nucl-th].
[17] L. Gavassino and J. Noronha, arXiv e-prints , arXiv:2305.04119 (2023), arXiv:2305.04119 [gr-qc].
[18] L. Gavassino, Entropy **26**, 147 (2024), arXiv:2312.13553 [nucl-th].
[19] I. Aniceto, J. Noronha, and M. Spaliński, (2024), arXiv:2401.06750 [nucl-th].
[20] J. Noronha, B. Schenke, C. Shen, and W. Zhao (2024) arXiv:2401.09208 [nucl-th].
[21] N. G. Hadjiconstantinou, Physics of Fluids **18**, 111301-111301-19 (2006).
[22] G. S. Denicol, H. Niemi, E. Molnár, and D. H. Rischke, Phys. Rev. D **85**, 114047 (2012).
[23] K. Huang, *Statistical Mechanics*, 2nd ed. (John Wiley & Sons, 1987).
[24] H. Struchtrup and P. Taheri, IMA Journal of Applied Mathematics **76**, 672 (2011), https://academic.oup.com/imamat/article-pdf/76/5/672/1835580/hxr004.pdf.
[25] G. S. Rocha, G. S. Denicol, and J. Noronha, Phys. Rev. D **106**, 036010 (2022), arXiv:2205.00078 [nucl-th].
[26] G. S. Rocha, D. Wagner, G. S. Denicol, J. Noronha, and D. H. Rischke, (2023), arXiv:2311.15063 [nucl-th].
[27] R. Baier, P. Romatschke, D. Thanh Son, A. O. Starinets, and M. A. Stephanov, Journal of High Energy Physics **2008**, 100 (2008), arXiv:0712.2451 [hep-th].
[28] S. c. v. Grozdanov, P. K. Kovtun, A. O. Starinets, and P. Tadić, Phys. Rev. Lett. **122**, 251601 (2019).
[29] P. Kovtun, Journal of High Energy Physics **2019**, 34 (2019), arXiv:1907.08191 [hep-th].
[30] H. Struchtrup, Continuum Mechanics and Thermodynamics **17**, 43 (2005).
[31] M. P. Heller, A. Serantes, M. Spaliński, and B. Withers, (2023), arXiv:2305.07703 [hep-th].
[32] M. P. Heller, A. Serantes, M. Spaliński, V. Svensson, and B. Withers, Phys. Rev. Lett. **128**, 122302 (2022).
[33] P. Kovtun, G. D. Moore, and P. Romatschke, Phys. Rev. D **84**, 025006 (2011), arXiv:1104.1586 [hep-ph].
[34] G. S. Denicol and J. Noronha, (2016), arXiv:1608.07869 [nucl-th].
[35] G. C. Pomraning, *The equations of radiation hydrodynamics* (Dover Publications, Mineola, 1973).
[36] D. Mihalas and B. Weibel Mihalas, *Foundations of radiation hydrodynamics* (Oxford University Press, Oxford, 1984).
[37] B. D. Farris, T. K. Li, Y. T. Liu, and S. L. Shapiro, Phys. Rev. D **78**, 024023 (2008), arXiv:0802.3210 [astro-ph].
[38] M. Shibata, K. Kiuchi, Y.-i. Sekiguchi, and Y. Suwa, Prog. Theor. Phys. **125**, 1255 (2011), arXiv:1104.3937 [astro-ph.HE].
[39] A. Sadowski, R. Narayan, A. Tchekhovskoy, and Y. Zhu, MNRAS **429**, 3533 (2013), arXiv:1212.5050 [astro-ph.HE].
[40] L. Gavassino, M. Antonelli, and B. Haskell, Symmetry **12**, 1543 (2020).
[41] D. Radice and S. Bernuzzi, Astrophys. J. **959**, 46 (2023), arXiv:2306.13709 [astro-ph.HE].
[42] L. H. Thomas, The Quarterly Journal of Mathematics **1**, 239 (1930).
[43] S. Weinberg, ApJ **168**, 175 (1971).
[44] N. Udey and W. Israel, MNRAS **199**, 1137 (1982).
[45] K. S. Thorne, MNRAS **194**, 439 (1981).
[46] J. L. Anderson and H. R. Witting, Physica **74**, 466 (1974).
[47] L. Rezzolla and O. Zanotti, *Relativistic Hydrodynamics, by L. Rezzolla and O. Zanotti. Oxford University Press, 2013. ISBN-10: 0198528906; ISBN-13: 978-0198528906* (2013).
[48] G. S. Denicol and J. Noronha, Phys. Lett. B **850**, 138487 (2024), arXiv:2209.10370 [nucl-th].
[49] G. S. Rocha, G. S. Denicol, and J. Noronha, Phys. Rev. Lett. **127**, 042301 (2021), arXiv:2103.07489 [nucl-th].
[50] D. D. Clayton, *Principles of stellar evolution and nucleosynthesis* (The University of Chicago Press (Chicago), 1983).
[51] G. D. Moore, JHEP **05**, 084 (2018), arXiv:1803.00736 [hep-ph].
[52] W. Israel and J. Stewart, Annals of Physics **118**, 341 (1979).
[53] F. Becattini, Acta Physica Polonica B **47**, 1819 (2016), arXiv:1606.06605 [gr-qc].
[54] L. Gavassino, Classical and Quantum Gravity **38**, 21LT02 (2021), arXiv:2104.09142 [gr-qc].
[55] L. Gavassino, Classical and Quantum Gravity **39**, 185008 (2022), arXiv:2202.06760 [gr-qc].
[56] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol 2: Partial Differential Equations* (John Wiley and Sons, New York, NY, 1989).
[57] J. Rauch, *Partial Differential Equations*, Graduate Texts in Mathematics (Springer, New York, NY, 1991).
[58] L. Gavassino, Phys. Rev. D **107**, 065013 (2023).
[59] L. Gavassino, Phys. Rev. X **12**, 041001 (2022).
[60] R. Geroch, arXiv e-prints , gr-qc/0103112 (2001), arXiv:gr-qc/0103112 [gr-qc].
[61] L. García-Colín, R. Velasco, and F. Uribe, Physics Reports **465**, 149 (2008).
[62] R. M. Wald, *General relativity* (Chicago Univ. Press, Chicago, IL, 1984).
[63] W. Rudin, *Principles of Mathematical Analysis, third edition* (McGraw-Hill, Inc., New York, NY, 1973).
[64] K. Dusling, G. D. Moore, and D. Teaney, Phys. Rev. C **81**, 034907 (2010), arXiv:0909.0754 [nucl-th].

[65] K. Dusling and T. Schäfer, Phys. Rev. C **85**, 044909 (2012), arXiv:1109.5181 [hep-ph].
[66] A. Kurkela and U. A. Wiedemann, Eur. Phys. J. C **79**, 776 (2019), arXiv:1712.04376 [hep-ph].
[67] J. D. Bjorken, Phys. Rev. D **27**, 140 (1983).
[68] H. Krants, S.; Parks, *A Primer of Real Analytic Functions* (Springer Basel, Basel, 1992).

arXiv:2402.19304v1 [cond-mat.mes-hall] 29 Feb 2024

# Viscoelastic response and anisotropic hydrodynamics in Weyl semimetals

A. A. Herasymchuk,[1] E. V. Gorbar,[1,2] and P. O. Sukhachov[3,4,5,*]

[1] *Faculty of Physics, Kyiv National Taras Shevchenko University, 64/13 Volodymyrska st., 01601 Kyiv, Ukraine*
[2] *Bogolyubov Institute for Theoretical Physics, 14-b Metrolohichna st., 03143 Kyiv, Ukraine*
[3] *Center for Quantum Spintronics, Department of Physics, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*
[4] *Department of Physics and Santa Cruz Materials Center, University of California Santa Cruz, Santa Cruz, CA 95064, USA*
[5] *Department of Physics, Yale University, New Haven, Connecticut 06520, USA*

(Dated: February 29, 2024)

We study viscoelastic response in Weyl semimetals with broken time-reversal symmetry. Topology and anisotropy of the Fermi surface are manifested in the viscoelasticity tensor of the electron fluid. In the dynamic (inter-band) part of this tensor, the anisotropy leads to a qualitatively different, compared to isotropic models, scaling with frequency and the Fermi energy. While components of the viscosity tensor determined by the Fermi surface properties agree in the Kubo and kinetic formalisms, the latter misses the anomalous Hall viscosity determined by filled states below the Fermi surface. The anisotropy of the dispersion relation is also manifested in the acceleration and relaxation terms of the hydrodynamic equations.

## I. INTRODUCTION

Experimental discovery of hydrodynamic transport in graphene [1–9] reignited the interest to electron hydrodynamics first observed in GaAs heterostructures [10, 11]. This is particularly fascinating since the foundation of electron hydrodynamics was laid down in the 1960s in the seminal works of R. Gurzhi [12, 13]. Being an interesting regime with strong electron-electron interactions, electron hydrodynamics has several manifestations including the Gurzhi effect [12], Poiseuille profile of electric current, formation of vortices of electron fluid, etc.; see Refs. [14–16] for a review.

A characteristic and very important property of any fluid is its viscosity. Viscosity is a rank-4 tensor that relates the fluid stress to coordinate derivatives of the fluid velocity. In continuum mechanics, the viscoelastic stress tensor combines the response to strain (elasticity) and the time derivative of strain (viscosity); in what follows, we will use viscoelastic and viscous tensors interchangeably. In systems with rotational invariance, such as most fluids, the structure of the shear viscosity tensor is fixed by symmetry, and viscous response is determined only by two coefficients [17]: bulk and shear viscosities. However, solid-state materials are rarely isotropic where the rotational symmetry is commonly reduced only to certain angles of rotation.

Manifestations of anisotropy of the band dispersion in electron hydrodynamics were considered in Refs. [18–25]. In anisotropic metals, the reduced symmetry can lead to unusual viscosity tensors with additional components. Using a lattice regularization proposed in Ref. [19], the nondissipative Hall viscosity was analyzed in [26], however, other components of the viscosity tensor and the dynamic viscosity received less attention.

In addition to the anisotropy of the band dispersion, certain materials allow for the nontrivial topology of their electron wave functions. Topological aspects of viscoelastic response were addressed in Refs. [19, 20, 23, 24, 27–32]. The main attention, however, was paid to nondissipative or Hall responses [19, 26, 33–35] in mainly 2D systems. Three-dimensional (3D) topological materials exemplified by Weyl semimetals may also demonstrate nontrivial hydrodynamic response related to their topology [36–43]; experimental signatures of the hydrodynamic behavior in the Weyl semimetal $WP_2$ were reported in Ref. [44]. As an example of the topology manifestation in the viscoelasticity tensor in time-reversal symmetry (TRS) broken Weyl materials, we mention the anomalous Hall viscosity determined by the separation of the Weyl nodes in the Brillouin zone [29]. However, the calculation of the complete viscoelastic response that goes beyond the anomalous Hall viscosity, includes both dissipative and nondissipative parts, and accounts for the Weyl node topology in 3D Weyl semimetals is still lacking.

In this paper, we fill this gap and calculate the viscoelastic response of TRS-broken Weyl semimetals. One of our main findings is that the separation of the Weyl nodes in the Brillouin zone is manifested in both nondissipative (i.e., anomalous Hall) and dissipative components of the viscoelasticity tensor. Using the Kubo framework [45], we calculate both static and dynamic viscoelastic responses in linearized and two-band models of Weyl semimetals. In our calculations, we use two formulations of the stress tensor, which plays the key role in the Kubo approach: the canonical stress tensor, which follows from the coupling to mechanical strains [28, 29], and the Belinfante-like tensor, which includes contributions connected with the internal angular momentum [19, 30] and arises due to coupling to the background geometry [19]. These approaches were recently used also in Ref. [24] for chiral magnetic materials, which realize the

* pavel.sukhachov@gmail.com

so-called "spin-1" fermions. We found that both definitions of the stress tensor can give the same result for the static part of the viscoelastic response but differ, in general, for the dynamic part. As to the kinetic approach, our calculations of the static electron viscosity agree with those for the canonical stress tensor. However, being sensitive only to the Fermi surface contributions, the kinetic approach excludes the anomalous Hall viscosity, which has a topological origin; the latter is reproduced in the Kubo approach. The results for the dynamic (inter-band) viscoelastic response show a more dramatic difference between the two definitions of the stress tensor: internal degrees of freedom are crucial in restoring the standard tensor structure of the viscoelasticity tensor [19, 30]. In addition to the viscoelasticity tensor, the anisotropy of the dispersion relation of Weyl materials quantified by the separation of the Weyl nodes is also manifested in acceleration and relaxation terms of the hydrodynamic equations. In view of their different dependence on the Fermi energy, these terms can be important in Weyl semimetals with well-separated nodes.

The viscosity tensor can be mapped out via transport measurements akin to those discussed in Refs. [18, 23, 46]. Such transport measurements are readily accessible in graphene [3, 4, 6, 7]; see also Refs. [14, 16] and references therein. As an example of such measurements, we mention the study of the current magnitude and its profile in crystals cut along different crystallographic directions. Experimentally accessible signatures of the anisotropy of the viscosity tensor are further discussed in Sec. III.

The paper is organized as follows. In Sec. II, we summarize the Kubo framework for viscoelastic response and apply it to linearized and two-band models of 3D Dirac and Weyl materials. We rederive the hydrodynamic equations and cross-verify the viscosity tensor in the kinetic framework in Sec. III. The results are discussed and summarized in Sec. IV. A few technical details are presented in Appendices A, B, C, D, and E. Throughout this paper, we use $\hbar = k_{\mathrm{B}} = 1$.

## II. KUBO APPROACH TO VISCOELASTIC RESPONSE

### A. Viscoelasticity tensor

Besides the linear response to external electromagnetic fields, the Kubo approach can be applied to calculating other responses of the system including the viscoelastic response to static and time-dependent strain. The viscoelasticity tensor is defined via the stress tensor $T_{\mu\nu}$ and the strain transformation generator $J_{\mu\nu}$ (see, Appendix B) as follows: [45]

$$\eta_{\mu\nu\alpha\beta}(\Omega) = \frac{1}{\Omega + i0}\left[\left\langle[T_{\mu\nu}(0), J_{\alpha\beta}(0)]\right\rangle + i\delta_{\alpha\beta}\left\langle T_{\mu\nu}\right\rangle - i\delta_{\mu\nu}\delta_{\alpha\beta}\kappa^{-1} - iC_{\mu\nu\alpha\beta}(\Omega)\right], \tag{1}$$

where $C_{\mu\nu\alpha\beta}(\Omega)$ is the Fourier transform of the stress-stress correlation function

$$C_{\mu\nu\alpha\beta}(t-t') = i\lim_{\varepsilon\to+0}\Theta(t-t')\left\langle[T_{\mu\nu}(t), T_{\alpha\beta}(t')]\right\rangle \times e^{-\varepsilon(t-t')}. \tag{2}$$

In writing the above expressions, we assumed the long-wavelength limit, i.e., the response to uniform strains, which allowed us to set the wave vector of perturbation to zero; this is explicitly manifested in Eq. (3) below. Unfortunately, the stress tensor and strain transformation generator are not uniquely defined and different expressions are used in the literature. The key issue is the form and structure of the pseudospin angular momentum which is connected to the internal degrees of freedom. Further, $\kappa^{-1} = -V\left(\partial P/\partial V\right)_N$ is the inverse isentropic compressibility defined as the derivative of pressure with respect to volume of fluid $V$ taken at fixed particle number $N$ [47] and $\Omega$ is frequency.

By using the spectral function $A(\omega;\mathbf{k})$, see Appendix A for its definition, in the stress-stress correlator, the viscoelasticity tensor can be rewritten as follows [45]:

$$\eta_{\mu\nu\alpha\beta}(\Omega) = \frac{1}{\Omega + i0}\Bigg\{\left[\left\langle[T_{\mu\nu}(0), J_{\alpha\beta}(0)]\right\rangle + i\delta_{\alpha\beta}\left\langle T_{\mu\nu}\right\rangle - i\delta_{\mu\nu}\delta_{\alpha\beta}\kappa^{-1}\right] - i\int_{-\infty}^{+\infty}d\omega\int_{-\infty}^{+\infty}d\omega'\frac{f(\omega)-f(\omega')}{\omega'-\omega-\Omega-i0}\int\frac{d^3k}{(2\pi)^3}\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right]\Bigg\}. \tag{3}$$

Here, $f(\omega) = 1/\left[e^{(\omega-\mu)/T}+1\right]$ is the Fermi-Dirac distribution function, $\mu$ is the chemical potential, $T$ is temperature. Henceforth, to simplify our calculations, we assume the zero-temperature limit $T \to 0$ unless otherwise stated.

In the following few Subsections, we calculate the viscoelasticity tensor (3) for linearized models of Dirac and Weyl semimetals as well as more a realistic two-model of Weyl semimetals; the corresponding electron energy dispersions for Weyl semimetals are shown in Fig. 1.

### B. Dirac semimetals

As a warm-up, let us start with Dirac semimetals whose linearized Hamiltonian for gapless electron quasi-

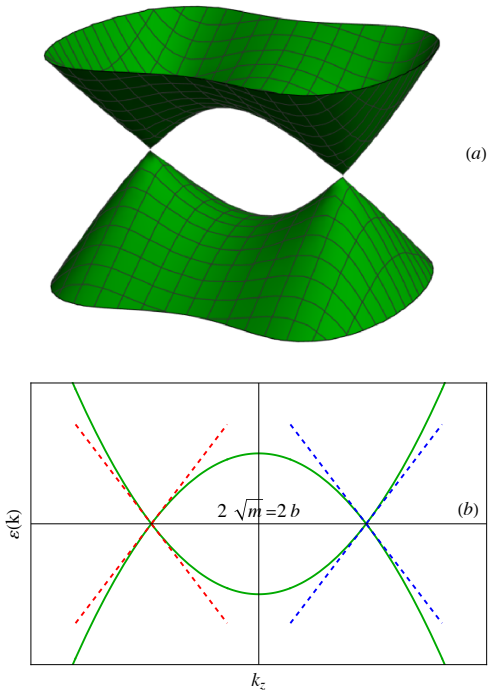



FIG. 1. Panel (a): The dispersion relation $\varepsilon = \varepsilon(k_x, 0, k_z)$ for the two-band model of Weyl semimetals (26). Panel (b) The dispersion relation $\varepsilon(0, 0, k_z)$ for the two-band model of Weyl semimetals shown by solid green lines with the dashed lines corresponding to the dispersion relation in the linearized model (13). The distance between the Weyl nodes is $2b = 2\sqrt{m}$.

particles reads

$$H = \mathrm{diag}\left(H_+, H_-\right), \quad H_\lambda = \lambda v_F \boldsymbol{\sigma}\cdot\mathbf{k}, \tag{4}$$

where $\lambda = \pm$ is chirality, $v_F$ is the Fermi velocity, and $\boldsymbol{\sigma}$ are the Pauli matrices in the pseudospin space. This Hamiltonian possesses full rotation symmetry, therefore, we use the following strain transformation generator and symmetric *Belinfante stress tensor* [48, 49]:

$$J_{\alpha\beta} = -\frac{1}{2}\left\{x_\beta, k_\alpha\right\} + \frac{i}{8}\left[\gamma_\beta, \gamma_\alpha\right], \tag{5}$$

$$T_{\mu\nu} = \frac{v_F}{2}\left(\gamma_0\gamma_\nu k_\mu + \gamma_0\gamma_\mu k_\nu\right), \tag{6}$$

respectively. Here, $\gamma_\alpha$ are the standard Dirac-$\gamma$ matrices, $x_\beta$ is the coordinate operator, $\{,\}$ is the anticommutator, and $[,]$ is the commutator. By calculating the real part of the viscoelasticity tensor in the Kubo approach, Eq. (3), we find

$$\mathrm{Re}\,\eta^{(\mathrm{B})}_{\mu\nu\alpha\beta}(\Omega) = \left[\frac{\Omega^3}{320\pi v_F^3}\Theta\left(\Omega - 2\mu\right) + \frac{\mu^4}{15\pi v_F^3}\delta(\Omega)\right] \times\left(\delta_{\mu\alpha}\delta_{\nu\beta} + \delta_{\nu\alpha}\delta_{\mu\beta} - \frac{2}{3}\delta_{\mu\nu}\delta_{\alpha\beta}\right). \tag{7}$$

For a rotationally-invariant $d$-dimensional system, the viscoelasticity tensor has only two independent components [17]

$$\eta_{\mu\nu\alpha\beta} = \zeta\delta_{\mu\nu}\delta_{\alpha\beta} + \eta^{\mathrm{sh}}\left(\delta_{\mu\alpha}\delta_{\nu\beta} + \delta_{\mu\beta}\delta_{\nu\alpha} - \frac{2}{d}\delta_{\mu\nu}\delta_{\alpha\beta}\right). \tag{8}$$

Comparing with Eq. (7), we obtain the trivial bulk viscosity $\zeta = 0$ and the following shear viscosity:

$$\eta^{\mathrm{sh}} = \frac{\Omega^3}{320\pi v_F^3}\Theta\left(\Omega - 2\mu\right) + \frac{\mu^4}{15\pi v_F^3}\delta(\Omega). \tag{9}$$

These findings agree with the results obtained in Refs. [32, 41].

By starting with a lattice model and studying the response to strain, one obtains a different strain transformation generator [28]

$$J_{\alpha\beta} = -\frac{1}{2}\left\{x_\beta, k_\alpha\right\}, \tag{10}$$

which, compared to Eq. (5), does not contain the pseudospin part. This strain transformation generator leads to a nonsymmetric *canonical stress tensor*

$$T_{\mu\nu} = v_F\gamma_0\gamma_\nu k_\mu, \tag{11}$$

and, as a result, to a different viscoelasticity tensor

$$\mathrm{Re}\,\eta^{(\mathrm{C})}_{\mu\nu\alpha\beta}(\Omega) = \frac{\Omega^3}{192\pi v_F^3}\Theta\left(\Omega - 2\mu\right)\left(\delta_{\mu\alpha}\delta_{\nu\beta} - \delta_{\mu\beta}\delta_{\nu\alpha}\right) + \mathrm{Re}\,\eta^{(\mathrm{B})}_{\mu\nu\alpha\beta}(\Omega), \tag{12}$$

cf. Eq. (7). Therefore, while the static (intra-band) part of the viscoelasticity tensor is the same for symmetric Belinfante and nonsymmetric canonical stress tensors, the dynamic (inter-band) part is sensitive to the choice of the stress tensor.

### C. Linear model of Weyl semimetals

Now we are ready to address the viscoelasticity tensor in TRS-broken Weyl semimetals. We start with a simple linearized model which describes two Weyl nodes separated in momentum space

$$H = \mathrm{diag}\left(H_+, H_-\right), \quad H_\lambda = \lambda v_F\boldsymbol{\sigma}\cdot\left(\mathbf{k} - \lambda\mathbf{b}\right), \tag{13}$$

where, without the loss of generality, we choose the momentum space separation to be along the $z$-direction, $\mathbf{b} = b_z\mathbf{e}_z$.

We first consider the canonical stress tensor. The viscoelasticity tensor can be separated into three parts: the static (intra-band) part, the dynamic (inter-band) part, and the Hall viscosity. The details of calculations and some intermediate results are given in Appendix C 1.

The real part of the static part of the viscoelasticity tensor reads as

$$\mathrm{Re}\,\eta^{\mathrm{stat}}_{\mu\nu\alpha\beta}(\Omega) = \frac{\mu^2}{3\pi v_F}\delta(\Omega)\delta_{\nu\beta}b_\mu b_\alpha + \frac{\mu^4}{15\pi v_F^3}\delta(\Omega) \times\left(\delta_{\mu\beta}\delta_{\nu\alpha} + \delta_{\mu\alpha}\delta_{\nu\beta} - \frac{2}{3}\delta_{\mu\nu}\delta_{\alpha\beta}\right). \tag{14}$$

The static part in the linearized model gets an additional term proportional to $\propto b_\mu b_\alpha$, see the first term in Eq. (14). Therefore, the separation vector $\mathbf{b}$ modifies the dissipative components of the viscoelasticity tensor. We cross-verify this finding in Sec. II D for the two-band model as well as in Sec. III in the kinetic approach.

The real part of the dynamic viscoelasticity tensor reads as

$$\mathrm{Re}\,\eta^{(\mathrm{C}),\mathrm{dyn}}_{\mu\nu\alpha\beta}(\Omega) = \Theta\left(\Omega-2\mu\right)\left[\frac{\Omega}{12\pi v_F}\delta_{\nu\beta}b_\mu b_\alpha + \frac{\Omega^3}{120\pi v_F^3} \times\left(\delta_{\mu\alpha}\delta_{\nu\beta}-\frac{1}{4}\delta_{\mu\beta}\delta_{\nu\alpha}-\frac{1}{4}\delta_{\mu\nu}\delta_{\alpha\beta}\right)\right]. \tag{15}$$

Like the static part, it also contains an additional term proportional to $\propto b_\mu b_\alpha$, see the first term in the square brackets in Eq. (15). As is easy to see, the limit $b\to 0$ reproduces the results for the Dirac semimetal obtained in the canonical stress tensor approach, cf. Eq. (12).

The Hall viscosity is determined by the imaginary part of trace $\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right]$ in Eq. (3); see also Eqs. (C1) and (C4). In the limit $\Omega\to 0$ and at $\mu=0$, we have

$$\mathrm{Re}\,\eta^{(\mathrm{C}),\mathrm{Hall}}_{\mu\nu\alpha\beta}(\Omega) = 4\sum_{j=1}^{3}\epsilon_{j\nu\beta}I^j_{\mu\alpha}, \tag{16}$$

where

$$I^{i=1,2}_{\mu\alpha} = \left(\delta_{\mu i}\delta_{\alpha 3}+\delta_{\mu 3}\delta_{\alpha i}\right)\frac{b_z}{4\pi^2}\left(\Lambda_z^2-\frac{b_z^2}{3}\right) \tag{17}$$

and

$$I^3_{\mu\alpha} = \delta_{\mu\alpha}\frac{b_z}{4\pi^2}\left[\left(\delta_{\mu 1}+\delta_{\mu 2}\right)\left(-\Lambda_\perp\Lambda_z+2\Lambda_z^2+\frac{2b_z^2}{3}\right) - \delta_{\mu 3}\frac{b_z^2}{6}\right]. \tag{18}$$

Evidently, all components are quadratically divergent except for $\eta_{zxzy}=-b_z^3/(6\pi^2)$. In the systems with axial symmetry, the other independent component of the Hall viscosity is given by $\eta_{xxxy}$. The obtained results agree with those in Ref. [29]. The expressions for $\Omega\neq 0$ are given in Eqs. (C8) and(C9), where the $\eta_{zxzy}$ component of the viscoelasticity tensor is also divergent albeit only logarithmically, $\propto \Omega^2 b_z\log\left(2v_F\Lambda_z/\Omega\right)$.

Now let us take into account the internal degrees of freedom, which lead to the pseudospin angular momentum, and consider the *Belinfante-like stress tensor* [50]. According to Eq. (13), rotational invariance is preserved in the plane perpendicular to the Weyl node separation $\mathbf{b}$. Therefore, the strain transformation generator and the Belinfante-like stress tensor read as

$$J_{\alpha\beta} = -\frac{1}{2}\left\{x_\beta,k_\alpha\right\}+\frac{i}{8}C_{\alpha\beta}\left[\gamma_\beta,\gamma_\alpha\right], \tag{19}$$

$$T_{\mu\nu} = v_F\gamma_0\gamma_\nu k_\mu + C_{\mu\nu}\frac{v_F}{2}\left(\gamma_0\gamma_\mu k_\nu-\gamma_0\gamma_\nu k_\mu\right) + C_{\mu\nu}\frac{v_F}{2}\left(\gamma_0\gamma_\mu\gamma_5 b_\nu-\gamma_0\gamma_\nu\gamma_5 b_\mu\right), \tag{20}$$

where, by choosing $\mathbf{b}=b\mathbf{e}_z$, the coefficients are $C_{\alpha 3}=C_{3\alpha}=0$ and $C_{\alpha\beta}=1$ with $\alpha,\beta=1,2$.

The real part of the static viscoelasticity tensor remains the same as in the case of the canonical stress tensor, see Eq. (14). On the other hand, the real part of the dynamic part of the viscoelasticity tensor is modified:

$$\mathrm{Re}\,\eta^{(\mathrm{B}),\mathrm{dyn}}_{\mu\nu\alpha\beta}(\Omega) = \left\{\delta_{\mu\alpha}\delta_{\nu\beta}\left[1-\frac{5}{8}\left(C_{\alpha\beta}+C_{\mu\nu}-C_{\alpha\beta}C_{\mu\nu}\right)\right] -\frac{1}{4}\delta_{\mu\beta}\delta_{\nu\alpha}\left[1-\frac{5}{2}\left(C_{\alpha\beta}+C_{\mu\nu}-C_{\alpha\beta}C_{\mu\nu}\right)\right] -\frac{1}{4}\delta_{\mu\nu}\delta_{\alpha\beta}\right\}\frac{\Omega^3}{120\pi v_F^3}\Theta\left(\Omega-2\mu\right) + \frac{\Omega}{12\pi v_F}\delta_{\nu\beta}b_\mu b_\alpha\Theta\left(\Omega-2\mu\right). \tag{21}$$

As one can see, compared to the viscoelasticity tensor of Dirac semimetals given in Eq. (7), there are a few terms related to the additional rotational symmetry of a Weyl semimetal.

The Hall viscosity reads as

$$\mathrm{Re}\,\eta^{(\mathrm{B}),\mathrm{Hall}}_{\mu\nu\alpha\beta}(\Omega) = \left(2-C_{\alpha\beta}\right)\left(2-C_{\mu\nu}\right)I_{\mu\nu\alpha\beta} + C_{\alpha\beta}\left(2-C_{\mu\nu}\right)I_{\mu\nu\beta\alpha} + C_{\mu\nu}\left(2-C_{\alpha\beta}\right)I_{\nu\mu\alpha\beta} + C_{\alpha\beta}C_{\mu\nu}I_{\nu\mu\beta\alpha}, \tag{22}$$

where $I_{\mu\nu\alpha\beta}=\sum_{j=1}^{3}\epsilon_{j\nu\beta}I^j_{\mu\alpha}$ with $I^j_{\mu\alpha}$ given in Eqs. (17) and (18). As in the case of the canonical stress tensor, see Eq. (16), only $\eta_{zxzy}$ is finite for $\Omega=0$; other nontrivial components are quadratically divergent. According to the structure of coefficients $C_{\mu\nu}$, Hall components $\eta_{zizj}$ and $\eta_{izzj}$ are the same in both approaches, but $\eta_{ijzz}$, $\eta_{ijii}$, and $\eta_{jiii}$ are different. Indeed, for the Belinfante-like stress tensor, we have the following relations for the Hall viscosity tensor components:

$$\mathrm{Re}\,\eta^{(\mathrm{B}),\mathrm{Hall}}_{ijii}(\Omega) = \frac{1}{2}\mathrm{Re}\,\eta^{(\mathrm{C}),\mathrm{Hall}}_{ijii}(\Omega), \tag{23}$$

$$\mathrm{Re}\,\eta^{(\mathrm{B}),\mathrm{Hall}}_{jiii}(\Omega) = \mathrm{Re}\,\eta^{(\mathrm{B}),\mathrm{Hall}}_{ijii}(\Omega) = \frac{1}{2}\mathrm{Re}\,\eta^{(\mathrm{C}),\mathrm{Hall}}_{ijii}(\Omega), \tag{24}$$

$$\mathrm{Re}\,\eta^{(\mathrm{B}),\mathrm{Hall}}_{ijzz}(\Omega) = 0, \tag{25}$$

while $\mathrm{Re}\,\eta^{(\mathrm{C}),\mathrm{Hall}}_{jiii}(\Omega)=0$ and $\mathrm{Re}\,\eta^{(\mathrm{C}),\mathrm{Hall}}_{ijzz}(\Omega)\neq 0$. For details of derivation, see Appendix C 1.

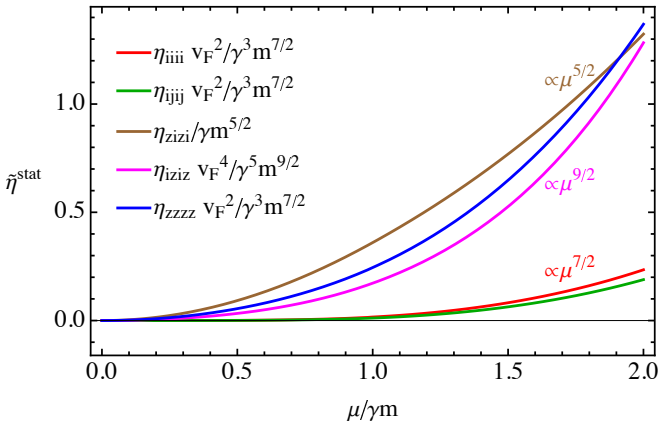



FIG. 2. The dependence of components of the real part of the static viscoelasticity tensor on $\mu/\gamma m$; see Eq. (C37) for its explicit form. The notation $\tilde{\eta}^{\rm stat}_{\mu\nu\alpha\beta}$ stands for dimensionless viscosity, see the legend for the definition for each of the components.

### D. Two-band model of Weyl semimetals

To cross-verify the obtained anomalous and anisotropic parts of the viscoelasticity tensor, we consider the following two-band model:

$$H = \boldsymbol{\sigma}\cdot\mathbf{d}(\mathbf{k}), \tag{26}$$

where $d_x(k_x) = v_F k_x$, $d_y(k_y) = v_F k_y$, and $d_z(k_z) = \gamma\left(k_z^2 - m\right)$. This model describes a TRS-breaking Weyl semimetal with two Weyl nodes separated by $2\sqrt{m}$ along the $\hat{z}$-direction.

Let us begin with the case of the canonical stress tensor. The complete result for the real part of the static viscoelasticity tensor is too bulky to be presented in the main text; it is given in Eq. (C37). We show the dependence of the few components of the viscoelasticity tensor in Fig. 2. As one can see from the figure, the static part of the viscoelasticity tensor $\mathrm{Re}\,\eta^{\rm stat}_{\mu\nu\alpha\beta}(\Omega)$ shows no kinklike features at $|\mu| = \gamma m$ where two pockets of the Fermi surface merge. This follows from the fact that parts of the viscoelasticity tensor with $\Theta\left(\gamma m - |\mu|\right)$ are $\propto \left(\gamma m - |\mu|\right)^{5/2}$, hence they have continuous derivatives.

For small Fermi level, $|\mu| \ll \gamma m$, the viscoelasticity tensor acquires a compact form:

$$\begin{aligned}\mathrm{Re}\,\eta^{\rm stat}_{\mu\nu\alpha\beta}(\Omega) \simeq \Bigg[&\frac{\mu^4}{15\pi v_F^3}\left(\delta_{\mu\beta}\delta_{\nu\alpha}+\delta_{\mu\alpha}\delta_{\nu\beta}-\frac{2}{3}\delta_{\mu\nu}\delta_{\alpha\beta}\right)\\ &+\frac{\mu^4}{10\pi v_F^3}\delta_{\mu 3}\left(\delta_{\nu 3}-\delta_{\nu 1}-\delta_{\nu 2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\\ &+\frac{\mu^2 m}{3\pi v_F}\delta_{\mu 3}\delta_{\mu\alpha}\delta_{\nu\beta}\Bigg]\delta(\Omega),\end{aligned} \tag{27}$$

where we also substituted $2\gamma\sqrt{m} \to v_F$. The above viscoelasticity tensor deviates from that in the linear model, cf. Eq. (14), because expansion in $|\mu| \ll \gamma m$ still allows for higher-order anisotropy corrections to contribute to $\mathrm{Re}\,\eta^{\rm stat}_{\mu\nu\alpha\beta}(\Omega)$. The static part of the viscoelasticity tensor is also reproduced in the kinetic approach, see Sec. III. Note that since $\sqrt{m}$ plays the same role as the momentum space separation $b_z$ in the linearized model, there is a correspondence between the third term in the square brackets in Eq. (27) and a term proportional to $\propto b_\mu b_\alpha$ in the linearized model, cf. Eqs. (14) and (27).

In the opposite limit of the Fermi energy away from the Weyl nodes, $|\mu| \gg \gamma m$, we obtain

$$\begin{aligned}\mathrm{Re}\,\eta^{\rm stat}_{\mu\nu\alpha\beta}(\Omega) \simeq\ \delta(\Omega)\Bigg\{&\frac{2|\mu|^{7/2}}{45\pi\gamma^{1/2}v_F^2}\Big[\delta_{\mu\beta}\delta_{\nu\alpha}-\frac{4}{5}\delta_{\mu\nu}\delta_{\alpha\beta}\\ &+\left(\delta_{\mu 1}+\delta_{\mu 2}\right)\left(\delta_{\nu 1}+\delta_{\nu 2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\Big]\\ &+\frac{4\gamma^{1/2}|\mu|^{9/2}}{77\pi v_F^4}\left(\delta_{\mu 1}+\delta_{\mu 2}\right)\delta_{\nu 3}\delta_{\mu\alpha}\delta_{\nu\beta}\\ &+\frac{|\mu|^{5/2}}{21\pi\gamma^{3/2}}\delta_{\mu 3}\left(\delta_{\nu 1}+\delta_{\nu 2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\\ &+\frac{2|\mu|^{7/2}}{15\pi\gamma^{1/2}v_F^2}\delta_{\mu 3}\delta_{\nu 3}\delta_{\mu\alpha}\delta_{\nu\beta}\Bigg\}.\end{aligned} \tag{28}$$

The inherent anisotropy of the electron energy spectrum of the two-band model is clearly imprinted in the viscoelasticity tensor leading to its highly anisotropic form.

The full expression for the real part of the dynamic viscoelasticity tensor is given in Eq. (C38). We plot a few of its components in Fig. 3. As one can see, there is no kinklike feature in the dynamical part of the viscoelasticity tensor at $\Omega = 2\gamma m$. Indeed, as follows from Eq. (C38), the inter-band parts of the viscoelasticity tensor are $\propto (2\gamma m - \Omega)^{a+1/2}$ with $a \geq 1$. Hence, the derivatives are continuous. The absence of a kinklike feature in the dynamical part of the viscoelasticity tensor differs from that of, e.g., optical conductivity studied in Ref. [51]. The latter shows a well-pronounced kink at $\Omega = 2\gamma m$ originating from the van Hove singularities of the density of states.

For $\Omega \ll 2\gamma m$, we have the following expression for the inter-band part of the viscoelasticity tensor:

$$\begin{aligned}\mathrm{Re}\,\eta^{\rm dyn}_{\mu\nu\alpha\beta}(\Omega) \simeq \Bigg[&\frac{\Omega^3}{120\pi v_F^3}\left(\delta_{\mu\alpha}\delta_{\nu\beta}-\frac{1}{4}\delta_{\mu\beta}\delta_{\nu\alpha}-\frac{1}{4}\delta_{\mu\nu}\delta_{\alpha\beta}\right)\\ &+\delta_{\mu 3}\delta_{\mu\alpha}\delta_{\nu\beta}\frac{\Omega m}{12\pi v_F}+\frac{\Omega^3}{480\pi v_F^3}\delta_{\mu 3}\delta_{\nu 3}\delta_{\mu\alpha}\delta_{\nu\beta}\\ &-\frac{\Omega^3}{80\pi v_F^3}\delta_{\mu 3}\left(\delta_{\nu 1}+\delta_{\nu 2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\Bigg]\Theta\left(\Omega-2\mu\right),\end{aligned} \tag{29}$$

where we also replaced $2\gamma\sqrt{m} \to v_F$. As with the static part, the viscoelasticity tensor expanded in $\Omega/(2\gamma m)$ in the two-band model is not equivalent to that in the linearized model even in the vicinity of the Weyl nodes.

In the opposite limit of large frequencies $\Omega/(2\gamma m) \gg 1$,

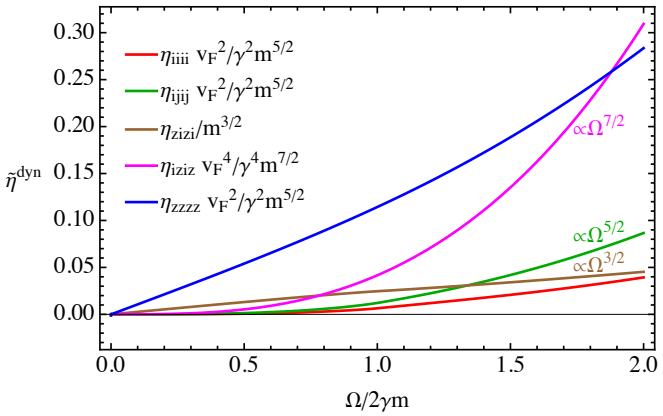



FIG. 3. The dependence of components of the real part of dynamic viscoelasticity tensor on $\Omega/2\gamma m$; see Eq. (C37) for its explicit form. The notation $\tilde{\eta}^{\rm dyn}_{\mu\nu\alpha\beta}$ stands for dimensionless viscosity, see the legend for the definition for each of the components. We used the canonical stress tensor.

we obtain

$$\begin{aligned}\mathrm{Re}\,\eta^{\rm dyn}_{\mu\nu\alpha\beta}(\Omega) \simeq \Bigg[&\frac{\gamma^{1/2}(\delta_{\mu1}+\delta_{\mu2})\delta_{\nu3}\delta_{\mu\alpha}\delta_{\nu\beta}}{231\sqrt{2}\pi v_F^4}\Omega^{7/2}\\ &+\frac{7(\delta_{\mu1}+\delta_{\mu2})(\delta_{\nu1}+\delta_{\nu2})\delta_{\mu\alpha}\delta_{\nu\beta}}{720\sqrt{2}\pi\gamma^{1/2}v_F^2}\Omega^{5/2}\\ &+\frac{5\delta_{\mu3}(\delta_{\nu1}+\delta_{\nu2})\delta_{\mu\alpha}\delta_{\nu\beta}}{336\sqrt{2}\pi\gamma^{3/2}}\Omega^{3/2}\\ &-\frac{(\delta_{\mu\beta}\delta_{\nu\alpha}+\delta_{\mu\nu}\delta_{\alpha\beta})}{360\sqrt{2}\pi\gamma^{1/2}v_F^2}\Omega^{5/2}\\ &+\frac{\delta_{\mu3}\delta_{\nu3}\delta_{\mu\alpha}\delta_{\nu\beta}}{60\sqrt{2}\pi\gamma^{1/2}v_F^2}\Omega^{5/2}\Bigg]\Theta(\Omega-2\mu).\end{aligned} \tag{30}$$

As one can see, the dynamical viscoelasticity tensor acquires dependence on fractional powers of $\Omega$; this is similar to the fractional powers of $\mu$ in Eq. (28). These results show that both static and dynamic parts of the viscoelasticity tensor can be used to probe the anisotropy of the Fermi surface.

For the Hall viscosity, we find

$$\mathrm{Re}\,\eta^{\rm Hall}_{\mu\nu\alpha\beta}(\Omega)=\sum_{j=1}^{3}\epsilon_{j\nu\beta}I^j_{\mu\alpha}, \tag{31}$$

where

$$\begin{aligned}I^{i=1,2}_{\mu\alpha} &= (\delta_{\mu3}\delta_{\alpha i}+\delta_{\mu i}\delta_{\alpha3})\\ &\times\left(-\frac{\gamma^2\Lambda_z^5}{10\pi^2v_F^2}+\frac{\gamma\Lambda_z^3\Lambda_\perp}{12\pi^2v_F}+\frac{\gamma^2m\Lambda_z^3}{6\pi^2v_F^2}-\frac{2\gamma^2m^{5/2}}{15\pi^2v_F^2}\right)\end{aligned} \tag{32}$$

and

$$\begin{aligned}I^3_{\mu\alpha} = \delta_{\mu\alpha}\Bigg[&(\delta_{\mu1}+\delta_{\mu2})\left(-\frac{\gamma^2\Lambda_z^5}{20\pi^2v_F^2}+\frac{\gamma^2m\Lambda_z^3}{6\pi^2v_F^2}\right.\\ &\left.+\frac{\gamma\Lambda_z^3\Lambda_\perp}{24\pi^2v_F}-\frac{\gamma^2m^2\Lambda_z}{4\pi^2v_F^2}-\frac{\gamma m\Lambda_z\Lambda_\perp}{8\pi^2v_F}+\frac{4\gamma^2m^{5/2}}{15\pi^2v_F^2}\right)\\ &+\delta_{\mu3}\frac{1}{12\pi^2}\left(\Lambda_z^3-2m^{3/2}\right)\Bigg]\end{aligned} \tag{33}$$

with $\Lambda_\perp$ and $\Lambda_z$ being momentum cutoffs.

A similar type of divergence is present in the Hall conductivity for the two-band model. Indeed, for $\Omega=\mu=0$, $\sigma^{\rm AHE}_{xy}$ reads as

$$\sigma^{\rm AHE}_{xy}=\int\frac{d^3k}{(2\pi)^3}\frac{v_F^2d_z}{2\varepsilon^3_{\mathbf{k}}}=\int_{-\Lambda_z}^{\Lambda_z}\frac{dk_z}{8\pi^2}\frac{d_z}{|d_z|}=\frac{\Lambda_z}{4\pi^2}-\frac{\sqrt{m}}{2\pi^2}. \tag{34}$$

This divergence originates from the flux of the Berry curvature at $k\to\infty$ [52]. The following subtraction:

$$\sigma^{\rm AHE}_{xy}\to\overline{\sigma}^{\rm AHE}_{xy}=\sigma^{\rm AHE}_{xy}-\sigma^{\rm AHE}_{xy}\Big|_{m=0} \tag{35}$$

allows us to cancel the divergent term and obtain the anomalous Hall conductivity consistent with that in the linearized model

$$\overline{\sigma}^{\rm AHE}_{xy}=-\frac{\sqrt{m}}{2\pi^2}. \tag{36}$$

Let us perform the same subtraction for the Hall viscoelasticity tensor (31). The corresponding coefficients $I^j_{\mu\alpha}$ in Eqs. (32) and (33) read

$$\begin{aligned}\overline{I}^{i=1,2}_{\mu\alpha} &= (\delta_{\mu3}\delta_{\alpha i}+\delta_{\mu i}\delta_{\alpha3})\left(\frac{\gamma^2m\Lambda_z^3}{6\pi^2v_F^2}-\frac{2\gamma^2m^{5/2}}{15\pi^2v_F^2}\right)\\ &= (\delta_{\mu3}\delta_{\alpha i}+\delta_{\mu i}\delta_{\alpha3})\left(\frac{\Lambda_z^3}{24\pi^2}-\frac{m^{3/2}}{30\pi^2}\right),\end{aligned} \tag{37}$$

$$\begin{aligned}\overline{I}^3_{\mu\alpha} &= \delta_{\mu\alpha}\Bigg[(\delta_{\mu1}+\delta_{\mu2})\left(\frac{\gamma^2m\Lambda_z^3}{6\pi^2v_F^2}-\frac{\gamma^2m^2\Lambda_z}{4\pi^2v_F^2}\right.\\ &\left.-\frac{\gamma m\Lambda_z\Lambda_\perp}{8\pi^2v_F}+\frac{4\gamma^2m^{5/2}}{15\pi^2v_F^2}\right)-\delta_{\mu3}\frac{m^{3/2}}{6\pi^2}\Bigg]\\ &= \delta_{\mu\alpha}\Bigg[(\delta_{\mu1}+\delta_{\mu2})\left(\frac{\Lambda_z^3}{24\pi^2}-\frac{m\Lambda_z}{16\pi^2}\right.\\ &\left.-\frac{m^{1/2}\Lambda_z\Lambda_\perp}{16\pi^2}+\frac{m^{3/2}}{15\pi^2}\right)-\delta_{\mu3}\frac{m^{3/2}}{6\pi^2}\Bigg],\end{aligned} \tag{38}$$

where we used $2\gamma\sqrt{m}\to v_F$ in the second expressions to compare with linearized model. Like in the linearized model, see Eq. (18) and the discussion after it, the only finite component is also $\eta_{zxzy}=-m^{3/2}/(6\pi^2)$. The other components are cubically divergent, unlike the quadratic divergence in the linearized model.

For the Belinfante-like stress tensor, the corrections to the dynamic and static parts of the viscoelasticity are

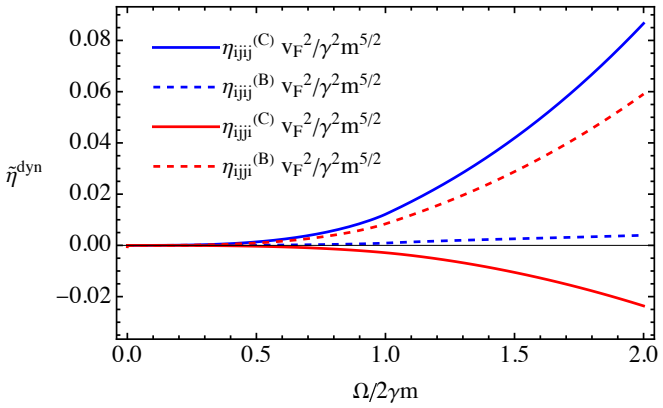



FIG. 4. The dependence of the components $\eta_{ijij}$ and $\eta_{ijji}$ of the real part of dimensionless dynamic viscoelasticity on $\Omega/(2\gamma m)$ for the canonical (solid lines) and Belinfante-like (dashed lines) stress tensors.

given in Eqs. (C48) and (C37), respectively. Like in the linearized model, the static part is the same for both definitions of the stress tensor, see Eq. (C54) for the discussion of the difference. As for the dynamic part, $\eta_{ijij}$ and $\eta_{ijji}$ components differ from the results of the canonical stress tensor case. For $\eta^{(\rm B)}_{ijij}$, the dependence on $\Omega/(2\gamma m)$ is modified, but the scaling for small and large frequencies $\Omega$ remains the same. In addition, $\eta^{(\rm B)}_{ijji}$ becomes positive. We compare the components of the dynamic viscoelasticity tensor for Belinfante-like and canonical stress tensors in Fig. 4.

The corrections to the Hall part of the viscoelasticity tensor are given in Eq. (C39). As in the linearized model, components $\eta_{zizj}$ and $\eta_{izzj}$ are the same for canonical and Belinfante-like stress tensors. On the other hand, components $\eta_{ijzz}$, $\eta_{ijii}$ are different

$$\mathrm{Re}\,\eta^{(\rm B)}_{ijii}(\Omega) = \frac{1}{2}\mathrm{Re}\,\eta^{(\rm C)}_{ijii}(\Omega), \tag{39}$$

$$\mathrm{Re}\,\eta^{(\rm B)}_{jiii}(\Omega) = \frac{1}{2}\mathrm{Re}\,\eta^{(\rm C)}_{ijii}(\Omega), \tag{40}$$

$$\mathrm{Re}\,\eta^{(\rm B)}_{ijzz}(\Omega) = 0, \tag{41}$$

while $\mathrm{Re}\,\eta^{(\rm C)}_{jiii}(\Omega)=0$ and $\mathrm{Re}\,\eta^{(\rm C)}_{ijzz}(\Omega)\neq 0$. For details of calculations, see Appendix C 2.

Thus, the nontrivial topology of the Weyl semimetals is manifested in the viscoelastic response leading to anomalous Hall terms. These results agree with the findings in Refs. [29, 53] extended here to the case of the two-band model.

## III. KINETIC APPROACH TO VISCOSITY AND HYDRODYNAMIC EQUATIONS

In this Section, we derive the hydrodynamic equations of motion, i.e., the continuity and Navier-Stokes equations, for the linearized and two-band models of Weyl semimetals focusing on the role of the Fermi surface anisotropy and band structure topology. The latter is quantified by the separation between the Weyl nodes. We show that the anomalous part of the viscoelasticity tensor does not appear in the standard approach based on the Boltzmann equation [54, 55]. Indeed, the stress-stress response used in Sec. II and current-current response are not guaranteed to give the same results for the fluid viscosity [56]; see Ref. [57] for the corresponding discussion in graphene. As we demonstrate below, while the "matter" terms, i.e., components of the viscosity tensor determined by the Fermi surface properties, agree in the Kubo and kinetic formalisms, the topological terms are absent in the latter. We trace this difference to the contribution of the filled states below the Fermi surface: while these states affect the viscoelastic response, only the Fermi surface properties determine the viscosity of the electron fluid flow in the kinetic approach. A similar situation is realized for the anomalous Hall effect or Chern-Simons contributions in the electric current; see, e.g., Ref. [58] for a detailed discussion.

Since the derivation of the hydrodynamic equation follows the standard scheme and is straightforward, we leave it to Appendices D and E. In what follows, we present the key results and discuss the physics behind them. Among our major findings are the manifestation of the shape (anisotropy) and structure (separation of the Weyl nodes) of the Fermi surface of Weyl semimetals in the hydrodynamic equations.

Using the linearized model of a Weyl semimetal (13), we first derive the following set of inviscid hydrodynamic equations:

$$\partial_t\rho(t,\mathbf{r})+\boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})\rho(t,\mathbf{r})+\frac{e^2}{2\pi^2}\left[\mathbf{b}\times\mathbf{E}\right]\right]=0, \tag{42}$$

$$\partial_t\left[w(t,\mathbf{r})\frac{\mathbf{u}(t,\mathbf{r})}{v_F^2}+\nu(t,\mathbf{r})\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{b}\right)\mathbf{b}\right]+\boldsymbol{\nabla}P\left(t,\mathbf{r}\right)$$
$$+\rho\left(t,\mathbf{r}\right)\mathbf{E}=-\frac{1}{\tau}\left[w(t,\mathbf{r})\frac{\mathbf{u}(t,\mathbf{r})}{v_F^2}+\nu(t,\mathbf{r})\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{b}\right)\mathbf{b}\right], \tag{43}$$

$$\partial_t\epsilon(t,\mathbf{r})+\boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})w(t,\mathbf{r})\right]=0, \tag{44}$$

where $w(t,\mathbf{r})=\epsilon(t,\mathbf{r})+P(t,\mathbf{r})=4\epsilon(t,\mathbf{r})/3$ is the enthalpy density, $P(t,\mathbf{r})=\epsilon(t,\mathbf{r})/3$ is the pressure. Assuming $T\to 0$, $\epsilon(t,\mathbf{r})=\mu^4(t,\mathbf{r})/(4\pi^2v_F^3)$ is the energy density and $\nu(t,\mathbf{r})=\mu^2(t,\mathbf{r})/(\pi^2v_F^3)$ is the density of states.

In the case of the two-band model (26), we have

$$\partial_t\rho(t,\mathbf{r})+\boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})\rho(t,\mathbf{r})+\mathbf{j}^{\rm AHE}\right]=0, \tag{45}$$

$$\partial_t\left[w(t,\mathbf{r})\frac{\mathbf{u}(t,\mathbf{r})}{v_F^2}+A(t,\mathbf{r})\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{e}_z\right)\mathbf{e}_z\right]+\boldsymbol{\nabla}P\left(t,\mathbf{r}\right)$$
$$+\rho\left(t,\mathbf{r}\right)\mathbf{E}=-\frac{1}{\tau}\left[w(t,\mathbf{r})\frac{\mathbf{u}(t,\mathbf{r})}{v_F^2}+A(t,\mathbf{r})\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{e}_z\right)\mathbf{e}_z\right], \tag{46}$$

$$\partial_t \epsilon(t,\mathbf{r}) + \boldsymbol{\nabla}\cdot[\mathbf{u}(t,\mathbf{r}) w(t,\mathbf{r})] = 0, \tag{47}$$

where the expressions for $\mathbf{j}^{\rm AHE}$ and $A(r,\mathbf{r})$ are given in Eqs. (D16) and (D19) respectively. Note that the Weyl nodes are separated along the $z$-axis in the two-band model (26); this corresponds to $\mathbf{b} = b\mathbf{e}_z$ in the linearized model.

The general structure of the zero-order hydrodynamic equations is preserved for both linearized and two-band models. However, the anisotropy of the Fermi surface leads to the appearance of terms $\propto \nu(t,\mathbf{r})\,(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{b})\,\mathbf{b}$ and $\propto A(t,\mathbf{r})\,(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{e}_z)\,\mathbf{e}_z$ under in time derivative (acceleration) and dissipative terms of the Euler equation. We note that, in view of weaker compared to $w(t,\mathbf{r})$ dependence of $\nu(t,\mathbf{r})$ and $A(t,\mathbf{r})$ on the chemical potential $\mu(t,\mathbf{r})$, these new terms are expected to play an important role for Fermi energies approaching Weyl nodes. In particular, at $T\to 0$, $w(t,\mathbf{r})/(\nu(t,\mathbf{r})b^2) \sim (p_F/b)^2$; for Weyl materials with well-separated Weyl nodes, $(p_F/b)^2 \lesssim 1$.

Let us now discuss the electron fluid viscosity, which affects the left-hand side of Eq.(46) via the term $-\eta_{ijkl}\nabla_j\nabla_l u_k(t,\mathbf{r})$. The viscosity tensor $\eta_{ijkl}$ in the linearized model reads

$$\begin{aligned}\eta_{ijkl} &= \frac{4\tau_{ee}}{15}\epsilon(t,\mathbf{r})\left(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}-\frac{2}{3}\delta_{ij}\delta_{kl}\right)\\ &+ \frac{\tau_{ee}v_F^2\nu(t,\mathbf{r})}{3} b_i b_k \delta_{jl},\end{aligned} \tag{48}$$

see Appendix E 1 for the derivation. In the case of the two-band model of Weyl semimetals, the viscosity tensor is given by Eq. (E12). By comparing terms at $\delta(\Omega)$ in Eq. (14) and at $\tau_{ee}$ in Eq. (48), we find the exact correspondence between the static (intra-band) parts of the dissipative viscoelastic response in the Kubo and kinetic approaches. However, the kinetic approach misses the topological Hall part of the viscoelasticity tensor and is insensitive to the internal degrees of freedom that may affect the Hall viscosity. These results underline important restrictions of the standard kinetic formalism.

Let us now discuss experimentally-relevant ways to probe the anisotropic viscosity and dissipation. A direct way would be to study the flow of electron liquid in channels cut along different crystalographic directions. For example, one can consider a flow in a Γ-shaped channel cut from a single crystal. Along one of the arms of the channel, the flow of electrons is directed along the crystallographic direction corresponding to the separation of the Weyl nodes along the flow velocity. The other arm of the channel should be prepared such that the flow velocity is perpendicular to the separation vector. In this case, according to Eqs. (42) and (48), the fluid viscosity and dissipation rate will be different in the different arms of the channel. This is manifested in the redistribution of the electric current density in the channel.

To show this, let us use Eq. (48) and obtain the steady-state Navier-Stokes equation for incompressible fluids at low velocities

$$\begin{aligned}&-\eta_0\Delta\mathbf{u}(\mathbf{r}) - \eta_1\Delta\left(\mathbf{u}(\mathbf{r})\cdot\mathbf{b}\right)\mathbf{b}\\ &+\frac{1}{\tau}\left[\frac{w}{v_F^2}\mathbf{u}(\mathbf{r}) + \nu\left(\mathbf{u}(\mathbf{r})\cdot\mathbf{b}\right)\mathbf{b}\right] = \rho\mathbf{E}.\end{aligned} \tag{49}$$

Here, $\eta_0 = \tau_{ee}w/5$, $\eta_1 = v_F^2\tau_{ee}\nu/3$, and, for simplicity, we used the linearized model (13). At the surfaces of the channel, we assume the no-slip boundary conditions with vanishing tangential components of velocity $\mathbf{u}_\parallel(r_\perp = 0, L) = \mathbf{0}$. The solution to Eq. (49) which captures the cases of flow with $\mathbf{u}\parallel\mathbf{b}$ and $\mathbf{u}\perp\mathbf{b}$ reads

$$u_\parallel(r_\perp) = \frac{v_F^2\tau}{w_b}\rho E\left[1-\frac{\cosh\left(\frac{L-2r_\perp}{2\lambda_{G,b}}\right)}{\cosh\left(\frac{L}{2\lambda_{G,b}}\right)}\right], \tag{50}$$

where $w_b = w + v_F^2\nu b_\parallel^2$, $\lambda_{G,b} = \sqrt{v_F^2\tau(\eta_0+b_\parallel^2\eta_1)/w_b} = \sqrt{15v_F^2\tau\tau_{ee}[3w+5v_F^2\nu b^2]/[w+v_F^2\nu b^2]}$ is the Gurzhi length, and $r_\perp$ is the coordinate normal to the channel. As one can see from Eq. (50), the chiral shift $b$ affects the profile of the fluid velocity by enhancing the Gurzhi length; the fluid velocity profile becomes more parabolic-like at larger $\lambda_{G,b}$, e.g., when $v_F^2\nu b_\parallel^2 \gg w$.

Another manifestation of the nontrivial Fermi surface structure of Weyl materials quantified by the distance between the Weyl nodes is in a suppression of the flow velocity along the separation direction $\mathbf{u}\parallel\mathbf{b}$. Indeed, as one can see from Eq. (50), the fluid velocity is suppressed as $\sim b_\parallel^2/p_F^2$, which can be a large factor for Weyl semimetals.

The electron fluid viscosity is also manifested in a dynamical response. For example, due to the interplay of viscous and inertial properties of the electron fluid as well as the boundary conditions, a nonuniform, double-peak, profile of the electric current appears in a channel [59, 60]. The current profile is manifested in a stray magnetic field surrounding the channel. In the case of anisotropic viscosity and Γ-shaped channel, the double-peak profile may be observed in one arm (smaller viscosity) but not in the other (larger viscosity). Another possible signature would be a different decay rate of hydrodynamic collective modes propagating along different crystallographic directions.

In passing, we note that high-quality channels of Weyl semimetals can be prepared via focused ion beam etching; see Ref. [61] for the use of this technique for $PdCoO_2$ and Ref. [62] for the Weyl semimetals $WP_2$.

## IV. DISCUSSION AND SUMMARY

Employing the Kubo formalism and chiral kinetic theory approach, we calculate the viscoelasticity tensor in Weyl semimetals with broken time-reversal symmetry paying special attention to the manifestation of topology

and the Fermi surface anisotropy. In the Kubo formalism, we found that, in addition to the anomalous Hall viscosity discussed in Ref. [29], the nontrivial topology and the Fermi surface anisotropy of Weyl semimetals affects the dissipative part of the viscoelasticity tensor. While the latter parts are in perfect agreement in Kubo and kinetic approaches, the anomalous contributions have a topological origin and are absent in the kinetic approach.

The nondissipative topological part of the viscoelasticity tensor, namely the anomalous Hall viscosity, see Eqs. (16) and (31), shares a similar origin with the Chern-Simons contributions to electric current in the anomalous Hall effect in Weyl semimetals and originates from the filled states below the Fermi surface; see, e.g., Ref. [58] for the corresponding discussion. Such contributions are captured in the Kubo approach and do not affect the responses solely determined by the Fermi surface properties, such as those obtained in the kinetic formalism.

Another advance of our work is in clarifying the impact of the nontrivial shape (anisotropy and Weyl node position) of the Fermi surface on viscoelastic properties of linearized and two-band models of Weyl semimetals, see Secs. II C and II D, respectively. The Fermi surface anisotropy is manifested in a rich structure of both static and dynamic parts of viscoelasticity tensors, see Eqs. (14), (15), and (21) for the linearized model and Eqs. (27), (28), (29), and (30) for the two-band model.

The static (intra-band) part of the viscoelasticity tensor contains additional terms determined by the momentum space separation of Weyl nodes, see Eqs. (14) and (27). Despite being related to the distance between the Weyl nodes, these terms originate from the shape of the Fermi surface rather than the nontrivial topology of electron wave functions. In particular, they are captured in both Kubo and kinetic approaches. The anisotropy of the two-band model is manifested in the unusual fractional scaling with the Fermi energy if the latter is large compared to the energy at which the Fermi surface pockets merge; see Eq. (28).

The dynamic (inter-band) part of viscosity contains linear and cubic in frequency terms in both linearized and two-band models, see Eqs. (21) and (29), respectively. The nonlinearity of the two-band model strongly modifies dynamic viscosity for larger frequencies exceeding the energy at which the Fermi surface pockets merge. In this case, the inter-band viscoelasticity tensor contains fractional powers of frequencies $\Omega^{3/2}$, $\Omega^{5/2}$, and $\Omega^{7/2}$; see Eq. (30). Such a scaling may be manifested in dynamic responses such as those involving time-dependent flows, see the end of Sec. III for the corresponding discussion.

In addition to viscosity, the anisotropy of the dispersion relation of Weyl materials is also manifested in acceleration and relaxation terms of the hydrodynamic equations, see Eqs. (43) and (46). The new terms are determined by the separation of the Weyl nodes and are parametrically larger compared to other terms while approaching the charge neutrality point. To the best of our knowledge, these terms were not discussed in the literature before. The proposed structure of the viscosity tensor and the new terms in the hydrodynamic equations can be investigated via transport measurements such as those in crystals cut along the different crystallographic directions or in a Γ-shaped channel discussed at the end of Sec. III. In particular, we expect a parametrically slower flow along the Weyl node separation direction with a more parabolic profile of the current.

Studies of the intra- and inter-band viscoelastic responses of the Fermi arc surface states or other topological materials such as multi-fold or spin-1 semimetals [63] are also interesting and will be reported elsewhere. Another possible extension of the present work is to investigate the viscoelastic response of the anisotropic spin-polarized Fermi surfaces of altermagnets [64].

## ACKNOWLEDGMENTS

The work of E.V.G. was supported by the Program "Dynamics of particles and collective excitations in high-energy physics, astrophysics and quantum macrosystems" of the Department of Physics and Astronomy of the NAS of Ukraine. P.O.S. acknowledges useful communications with A. Qaiumzadeh and S. Syzranov. P.O.S. acknowledges support through the Yale Prize Postdoctoral Fellowship in Condensed Matter Theory during his stay at Yale University as well as the Research Council of Norway through Grant No. 323766 and its Centres of Excellence funding scheme Grant No. 262633 "QuSpin" during his stay at NTNU.

## Appendix A: Spectral function

The retarded (+) and advanced (−) Green functions read

$$G\left(\omega \pm i0, \mathbf{k}\right) = \frac{i}{\omega + \mu \pm i0 - H(\mathbf{k})}, \tag{A1}$$

where $\omega$ is frequency (measured in energy units), $\mu$ is the chemical potential, $\mathbf{k}$ is momentum, and $H(\mathbf{k})$ is the Hamiltonian.

The spectral function is defined as the difference between the advanced and retarded Green functions at vanishing chemical potential

$$A\left(\omega;\mathbf{k}\right)=\frac{1}{2\pi}\left[G\left(\omega+i0,\mathbf{k}\right)-G\left(\omega-i0,\mathbf{k}\right)\right]_{\mu=0}. \tag{A2}$$

For the linearized model with broken TRS, see Eq. (4), the spectral function is given by $A(\omega;\mathbf{k})=\sum_{\lambda=\pm}\frac{1+\lambda\tau_z}{2}\otimes A_\lambda(\omega;\mathbf{k})$, where $\tau_z$ acts in the nodal space and

$$A_\lambda(\omega;\mathbf{k})=\frac{1}{2}\left[\delta(\omega-\varepsilon_{\mathbf{k},\lambda})+\delta(\omega+\varepsilon_{\mathbf{k},\lambda})\right]+\frac{\lambda v_F}{2\varepsilon_{\mathbf{k},\lambda}}\left[\boldsymbol{\sigma}\cdot(\mathbf{k}-\lambda\mathbf{b})\right]\left[\delta(\omega-\varepsilon_{\mathbf{k},\lambda})-\delta(\omega+\varepsilon_{\mathbf{k},\lambda})\right]. \tag{A3}$$

Here, $\varepsilon_{\mathbf{k},\lambda}=v_F|\mathbf{k}-\lambda\mathbf{b}|$ is the dispersion relation of quasiparticles with chirality $\lambda$.

Similarly, for the two-band model with $H(\mathbf{k})=(\boldsymbol{\sigma}\cdot\mathbf{d}(\mathbf{k}))$, we obtain

$$A(\omega;\mathbf{k})=\frac{1}{2}\left[\delta(\omega-\varepsilon_{\mathbf{k}})+\delta(\omega+\varepsilon_{\mathbf{k}})\right]+\frac{1}{2\varepsilon_{\mathbf{k}}}\left[\delta(\omega-\varepsilon_{\mathbf{k}})-\delta(\omega+\varepsilon_{\mathbf{k}})\right]H(\mathbf{k}), \tag{A4}$$

where $\varepsilon_{\mathbf{k}}=|\mathbf{d}(\mathbf{k})|$.

## Appendix B: Stress tensor

### 1. Stress tensor and strain transformation generators in continuum models

In this Section, we derive the stress tensor in the continuum model. Under deformations quantified by the displacement vector $\mathbf{u}(\mathbf{x},t)$, coordinates transform as $\mathbf{x}\to\mathbf{x}'=\mathbf{x}+\mathbf{u}(\mathbf{x},t)=\mathbf{x}+\mathbf{e}_\mu\frac{\partial u_\mu}{\partial x_\nu}x_\nu+o(|\mathbf{x}|)$, which can be rewritten in terms of homogeneous but time-dependent invertible matrix $\Lambda(t)$ with positive determinant

$$\mathbf{x}'=\Lambda(t)\mathbf{x}. \tag{B1}$$

The matrix $\Lambda(t)$ is expressed in terms of the strain tensor $\lambda(t)$ as $\Lambda(t)=\exp\left(\lambda(t)\right)=1+\lambda(t)+o(\lambda)$. This allows one to obtain the following transformed wave function $\psi(\mathbf{x},t)\to\psi_\lambda(\mathbf{x},t)$:

$$\psi_\lambda(\mathbf{x},t)=\sqrt{\det\Lambda}\,\psi\left(\Lambda\mathbf{x},t\right)=\psi(\mathbf{x},t)-i\lambda_{\mu\nu}\left(\frac{i}{2}\delta_{\mu\nu}+ix_\nu\frac{\partial}{\partial x_\mu}\right)\psi(\mathbf{x},t)+o(\lambda)=\psi(\mathbf{x},t)-i\lambda_{\mu\nu}L_{\mu\nu}\psi(\mathbf{x},t)+o(\lambda). \tag{B2}$$

If a system has additional symmetries, the strain transformation generator $J_{\mu\nu}$ is modified as [30]

$$J_{\mu\nu}=L_{\mu\nu}+L_{\mu\nu}^{(\mathrm{int})}. \tag{B3}$$

In other words, the generators of infinitesimal rotations are the operators of the total angular momentum which includes pseudospin accounting for internal degrees of freedom.

By using the strain transformation

$$S(t)=e^{-i\lambda_{\mu\nu}(t)J_{\mu\nu}}=I-i\lambda_{\mu\nu}(t)J_{\mu\nu}+o\left(\lambda\right), \tag{B4}$$

the transformed Hamiltonian can be represented as

$$H_\Lambda(t)=S(t)HS^{-1}(t)+i\frac{\partial S}{\partial t}S^{-1}=H-i\lambda_{\mu\nu}(t)\left[J_{\mu\nu},H\right]+\frac{\partial\lambda_{\mu\nu}}{\partial t}J_{\mu\nu}, \tag{B5}$$

where $[,]$ means commutator. The above relation leads to symmetric Belinfante-like stress tensor [65]

$$T_{\mu\nu}=-\frac{\partial H_\Lambda}{\partial\lambda_{\mu\nu}}=i\left[J_{\mu\nu},H\right]. \tag{B6}$$

### 2. Canonical stress tensor in a tight-binding model

Let us provide a different derivation of the stress tensor. We start with the following tight-binding Hamiltonian on cubic lattice: [28, 29]

$$\hat{H}_0 = \frac{1}{2}\sum_{i,j} c_i^\dagger \left(it\gamma_0\gamma_j - r\gamma_0\right) c_{i+j} + (m+3r)\sum_i c_i^\dagger\gamma_0 c_i + \sum_{i,l} b_l c_i^\dagger \gamma_0\gamma_l\gamma_5 c_i + \text{h.c.} \tag{B7}$$

Strains lead to the following changes in hopping parameters:

$$r \to r(1-\beta u_{jj}), \tag{B8}$$

$$t\gamma_0\gamma_j \to t\left(1 - \beta\sum_{j'}\frac{\partial u_j}{\partial x'_j}\gamma_0\gamma_{j'}\right), \tag{B9}$$

where $u_{\mu\nu} = \partial u_\mu/\partial x_\nu$. By using Eqs. (B8) and (B9) in Eq. (B7), we obtain

$$\begin{aligned}\hat{H}_u = \hat{H}_0 &+ \frac{1}{2}\sum_{i,j} c_i^\dagger it\Big( - \beta u_{jj}\gamma_0\gamma_j - \beta\sum_{j'\neq j} u_{jj'}\gamma_0\gamma_{j'}\Big) c_{i+j} + \frac{1}{2}\sum_{i,j} c_i^\dagger r\beta u_{jj}\gamma_0 c_{i+j} \\ &+ \left(m - r\beta\sum_j u_{jj}\right)\sum_i c_i^\dagger\gamma_0 c_i + \sum_{i,l} b_l c_i^\dagger\gamma_0\gamma_l\gamma_5 c_i + \text{h.c.}\end{aligned} \tag{B10}$$

Then, by using the standard definition of the stress tensor

$$\hat{T}_{\mu\nu} = \left.\frac{\delta\hat{H}_u}{\delta u_{\mu\nu}}\right|_{u_{\mu\nu}=0}, \tag{B11}$$

we find

$$\hat{T}_{\mu\nu} = -\frac{\beta}{2}\sum_i c_i^\dagger\left(it\gamma_0\gamma_\nu - r\delta_{\mu\nu}\gamma_0\right) c_{i+\mu} - r\beta\delta_{\mu\nu}\sum_i c_i^\dagger\gamma_0 c_i + \text{h.c.} \tag{B12}$$

By performing the Fourier transform, taking the continuum limit, i.e., setting the lattice constant $a$ to zero, $a \to 0$, and replacing $ta \to v$, we obtain canonical stress tensor for the linearized model

$$T_{\mu\nu} = \beta v k_\mu \gamma_0\gamma_\nu. \tag{B13}$$

It corresponds only to the orbital part of the strain transformation generator in (B3), i.e., $L_{\mu\nu}$.

By performing the Fourier transform and projecting out the high-energy states, Hamiltonian (B7) can be reduced to the Hamiltonian of the two-band model: $H = \boldsymbol{\sigma}\cdot\mathbf{d}(\mathbf{k})$, see Ref. [29] for details. The corresponding stress tensor is

$$T_{\mu\nu} = \beta k_\mu \partial_\nu\left(d_\nu\right)\sigma_\nu, \tag{B14}$$

where $d_x = vk_x$, $d_y = vk_y$, and $d_z = (b_z - m) - \frac{v^2}{m+b_z}k_z^2$ in the notation of [29]. Thus, we obtained the canonical stress tensor for the two-band model defined in Eq. (26) [66].

## Appendix C: Viscoelasticity tensor in the Kubo approach

The viscoelasticity tensor is defined in Eq. (3). Its real part reads as

$$\begin{aligned}
\mathrm{Re}\,\eta_{\mu\nu\alpha\beta}(\Omega) &= \frac{\mathrm{Im}\,C_{\mu\nu\alpha\beta}(\Omega)}{\Omega} + \pi\delta(\Omega)\left[\mathrm{Im}\left\langle[T_{\mu\nu}(0),J_{\alpha\beta}(0)]\right\rangle + \delta_{\alpha\beta}\left\langle T_{\mu\nu}\right\rangle - \delta_{\mu\nu}\delta_{\alpha\beta}\kappa^{-1} - \mathrm{Re}\,C_{\mu\nu\alpha\beta}(\Omega)\right] \\
&= \frac{1}{\Omega}\mathrm{v.p.}\int_{-\infty}^{+\infty}d\omega\int_{-\infty}^{+\infty}d\omega'\frac{f(\omega)-f(\omega')}{\omega'-\omega-\Omega}\int\frac{d^3k}{(2\pi)^3}\mathrm{Im}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right] \\
&+ \frac{\pi}{\Omega}\int_{-\infty}^{+\infty}d\omega\left[f(\omega)-f(\omega+\Omega)\right]\int\frac{d^3k}{(2\pi)^3}\mathrm{Re}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega+\Omega;\mathbf{k})\right] \\
&+ \pi\delta(\Omega)\,\mathrm{v.p.}\int_{-\infty}^{+\infty}d\omega\int_{-\infty}^{+\infty}d\omega'\frac{f(\omega)-f(\omega')}{\omega-\omega'}\int\frac{d^3k}{(2\pi)^3}\mathrm{Re}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right] \\
&+ \pi\delta(\Omega)\delta_{\alpha\beta}\int_{-\infty}^{+\infty}d\omega f(\omega)\int\frac{d^3k}{(2\pi)^3}\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})\right] - \pi\delta(\Omega)\kappa^{-1}\delta_{\mu\nu}\delta_{\alpha\beta} \\
&+ \pi\delta(\Omega)\int_{-\infty}^{+\infty}d\omega f(\omega)\int\frac{d^3k}{(2\pi)^3}\mathrm{Im}\,\mathrm{tr}\left\{[T_{\mu\nu},J_{\alpha\beta}]A(\omega;\mathbf{k})\right\} \\
&= \eta^{(1)}_{\mu\nu\alpha\beta}(\Omega) + \eta^{(2)}_{\mu\nu\alpha\beta}(\Omega) + \eta^{(3)}_{\mu\nu\alpha\beta}(\Omega) + \eta^{(4)}_{\mu\nu\alpha\beta}(\Omega) - \pi\delta(\Omega)\kappa^{-1}\delta_{\mu\nu}\delta_{\alpha\beta} + \eta^{(5)}_{\mu\nu\alpha\beta}(\Omega).
\end{aligned} \tag{C1}$$

In the subsequent Sections, we calculate $\eta^{(i)}_{\mu\nu\alpha\beta}(\Omega)$ in the linearized and two-band models of Weyl semimetals.

### 1. Linearized model of Weyl semimetals

In the linearized model, we start with the following strain transformation generator:

$$J_{\alpha\beta} = -\frac{1}{2}\left\{x_\beta,k_\alpha\right\} + \frac{i}{8}C_{\alpha\beta}\left[\gamma_\beta,\gamma_\alpha\right], \tag{C2}$$

whose second term includes the generic form of internal angular momentum. The corresponding stress tensor is

$$T_{\mu\nu} = v_F\gamma_0\gamma_\nu k_\mu + C_{\mu\nu}\frac{v_F}{2}\left(\gamma_0\gamma_\mu k_\nu - \gamma_0\gamma_\nu k_\mu\right) + C_{\mu\nu}\frac{v_F}{2}\left(\gamma_0\gamma_\mu\gamma_5 b_\nu - \gamma_0\gamma_\nu\gamma_5 b_\mu\right). \tag{C3}$$

The canonical stress tensor (B13) is reproduced by taking $C_{\alpha\beta}=0$.

By using the following expression for the imaginary part of the trace

$$\begin{aligned}
\mathrm{Im}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right] &= \sum_\lambda\lambda\frac{v_F^3}{4\varepsilon_{\mathbf{k},\lambda}}\left[\delta\left(\omega+\varepsilon_{\mathbf{k},\lambda}\right)\delta\left(\omega'-\varepsilon_{\mathbf{k},\lambda}\right) - \delta\left(\omega-\varepsilon_{\mathbf{k},\lambda}\right)\delta\left(\omega'+\varepsilon_{\mathbf{k},\lambda}\right)\right] \\
&\times\Big[C_{\alpha\beta}\left(2-C_{\mu\nu}\right)\left(k_j-\lambda b_j\right)k_\beta k_\mu\epsilon_{\nu\alpha j} + C_{\mu\nu}\left(2-C_{\alpha\beta}\right)\left(k_j-\lambda b_j\right)k_\alpha k_\nu\epsilon_{\mu\beta j} \\
&+ C_{\alpha\beta}C_{\mu\nu}\left(k_j-\lambda b_j\right)k_\beta k_\nu\epsilon_{\mu\alpha j} + \left(2-C_{\alpha\beta}\right)\left(2-C_{\mu\nu}\right)\left(k_j-\lambda b_j\right)k_\alpha k_\mu\epsilon_{\nu\beta j}\Big],
\end{aligned} \tag{C4}$$

we obtain the first term in Eq. (C1),

$$\begin{aligned}
\eta^{(1)}_{\mu\nu\alpha\beta}(\Omega) &= \frac{1}{\Omega}\mathrm{v.p.}\int_{-\infty}^{+\infty}d\omega\int_{-\infty}^{+\infty}d\omega'\frac{f(\omega)-f(\omega')}{\omega'-\omega-\Omega}\int\frac{d^3k}{(2\pi)^3}\mathrm{Im}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right] \\
&= \left(2-C_{\alpha\beta}\right)\left(2-C_{\mu\nu}\right)I_{\mu\nu\alpha\beta} + C_{\alpha\beta}\left(2-C_{\mu\nu}\right)I_{\mu\nu\beta\alpha} + C_{\mu\nu}\left(2-C_{\alpha\beta}\right)I_{\nu\mu\alpha\beta} + C_{\alpha\beta}C_{\mu\nu}I_{\nu\mu\beta\alpha}.
\end{aligned} \tag{C5}$$

Here

$$I_{\mu\nu\alpha\beta} = \sum_{j=1}^{3}\epsilon_{j\nu\beta}I^j_{\mu\alpha} \tag{C6}$$

and

$$I^j_{\mu\alpha} = \sum_\lambda \lambda \int \frac{d^3k}{(2\pi)^3} \left[f\left(-\varepsilon_{\mathbf{k},\lambda}\right) - f\left(\varepsilon_{\mathbf{k},\lambda}\right)\right] \frac{v_F^3 \left(k_j - \lambda b_j\right) k_\alpha k_\mu}{2\varepsilon_{\mathbf{k},\lambda}\left(4\varepsilon^2_{\mathbf{k},\lambda} - \Omega^2\right)}. \tag{C7}$$

For $\mu = 0$, we obtain

$$I^{i=1,2}_{\mu\alpha} = \left(\delta_{\mu i}\delta_{\alpha 3} + \delta_{\mu 3}\delta_{\alpha i}\right) \frac{b_z}{12\pi^2}\left[3\Lambda_z^2 - b_z^2 + \frac{\Omega^2}{2v_F^2}\left(\log\frac{2v_F\Lambda_z}{\Omega} + \frac{1}{3}\right)\right] \tag{C8}$$

and

$$\begin{aligned} I^3_{\mu\alpha} &= \delta_{\mu\alpha}\sum_\lambda \frac{v_F^3}{(2\pi)^3}\int_0^{2\pi} d\phi \int_{-\infty}^{+\infty} dk_z\, \lambda\left(k_z - \lambda b_z\right)\left\{\left(\delta_{\mu 1} + \delta_{\mu 2}\right)\left[\frac{\Lambda_\perp}{4} + \frac{\Omega^2 - 4v_F^2(k_z - \lambda b_z)^2}{8\Omega}\right.\right. \\ &\left.\left.\times\left[\frac{i\pi}{2} + \tanh^{-1}\left(\frac{2v_F\left|k_z - \lambda b_z\right|}{\Omega}\right)\right] - \frac{1}{8}v_F\left|k_z - \lambda b_z\right|\right] + \delta_{\mu 3}\frac{k_z^2}{4\Omega}\left[\frac{i\pi}{2} + \tanh^{-1}\left(\frac{2v_F\left|k_z - \lambda b_z\right|}{\Omega}\right)\right]\right\} \\ &= \delta_{\mu\alpha}\frac{b_z}{4\pi^2}\left\{\left(\delta_{\mu 1} + \delta_{\mu 2}\right)\left(-\Lambda_\perp\Lambda_z + 2\Lambda_z^2 + \frac{2b_z^2}{3} - \frac{\Omega^2}{6v_F^2}\right) + \delta_{\mu 3}\left[-\frac{b_z^2}{6} + \frac{\Omega^2}{12v_F^2}\left(\log\frac{2v_F\Lambda_z}{\Omega} - \frac{1}{6}\right)\right]\right\}, \end{aligned} \tag{C9}$$

where $\Lambda_\perp$ and $\Lambda_z$ are momentum cutoffs.

Combining the expressions for $\eta^{(2)}_{\mu\nu\alpha\beta}(\Omega)$ and $\eta^{(3)}_{\mu\nu\alpha\beta}(\Omega)$, and using the following expression for the real part of trace:

$$\begin{aligned} \mathrm{Re}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right] = \sum_\lambda \frac{v_F^2}{4}\Big\{ &\left[\delta(\omega - \varepsilon_{\mathbf{k},\lambda})\delta(\omega' + \varepsilon_{\mathbf{k},\lambda}) + \delta(\omega + \varepsilon_{\mathbf{k},\lambda})\delta(\omega' - \varepsilon_{\mathbf{k},\lambda})\right] \\ &\times\left[\delta_{\nu\beta}k_\alpha k_\mu\left(2 - C_{\alpha\beta}\right)\left(2 - C_{\mu\nu}\right) + \delta_{\nu\alpha}k_\beta k_\mu C_{\alpha\beta}\left(2 - C_{\mu\nu}\right)\right. \\ &\left.+\delta_{\mu\beta}k_\alpha k_\nu C_{\mu\nu}\left(2 - C_{\alpha\beta}\right) + \delta_{\mu\alpha}k_\beta k_\nu C_{\alpha\beta}C_{\mu\nu}\right] \\ &+ \frac{v_F^2}{\varepsilon^2_{\mathbf{k},\lambda}}\left[\delta(\omega - \varepsilon_{\mathbf{k},\lambda}) - \delta(\omega + \varepsilon_{\mathbf{k},\lambda})\right]\left[\delta(\omega' - \varepsilon_{\mathbf{k},\lambda}) - \delta(\omega' + \varepsilon_{\mathbf{k},\lambda})\right] \\ &\times\Big[\left(2 - C_{\alpha\beta}\right)\left(2 - C_{\mu\nu}\right)\left(k_\nu - \lambda b_\nu\right)\left(k_\beta - \lambda b_\beta\right)k_\alpha k_\mu \\ &+ C_{\alpha\beta}\left(2 - C_{\mu\nu}\right)\left(k_\nu - \lambda b_\nu\right)\left(k_\alpha - \lambda b_\alpha\right)k_\beta k_\mu \\ &+ \left(2 - C_{\alpha\beta}\right)C_{\mu\nu}\left(k_\mu - \lambda b_\mu\right)\left(k_\beta - \lambda b_\beta\right)k_\alpha k_\nu \\ &+ C_{\alpha\beta}C_{\mu\nu}\left(k_\mu - \lambda b_\mu\right)\left(k_\alpha - \lambda b_\alpha\right)k_\beta k_\nu\Big]\Big\}, \end{aligned} \tag{C10}$$

we derive

$$\begin{aligned} \eta^{(2)}_{\mu\nu\alpha\beta}(\Omega) + \eta^{(3)}_{\mu\nu\alpha\beta}(\Omega) &= \frac{\pi}{\Omega}\int_{-\infty}^{+\infty} d\omega\left[f(\omega) - f(\omega + \Omega)\right]\int\frac{d^3k}{(2\pi)^3}\mathrm{Re}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega + \Omega;\mathbf{k})\right] \\ &+ \pi\delta(\Omega)\,\mathrm{v.p.}\int_{-\infty}^{+\infty} d\omega\int_{-\infty}^{+\infty} d\omega'\frac{f(\omega) - f(\omega')}{\omega - \omega'}\int\frac{d^3k}{(2\pi)^3}\mathrm{Re}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right] \\ &= \Theta\left(\Omega - 2\mu\right)\Big\{\frac{\Omega^3}{120\pi v_F^3}\Big[\delta_{\mu\alpha}\delta_{\nu\beta}\left(1 - \frac{5}{8}\left(C_{\alpha\beta} + C_{\mu\nu} - C_{\alpha\beta}C_{\mu\nu}\right)\right) - \frac{1}{4}\delta_{\mu\beta}\delta_{\nu\alpha} \\ &\times\left(1 - \frac{5}{2}\left(C_{\alpha\beta} + C_{\mu\nu} - C_{\alpha\beta}C_{\mu\nu}\right)\right) - \frac{1}{4}\delta_{\mu\nu}\delta_{\alpha\beta}\Big] + \frac{\Omega}{48\pi v_F}\left[\left(2 - C_{\alpha\beta}\right)\left(2 - C_{\mu\nu}\right)\delta_{\nu\beta}b_\mu b_\alpha\right. \\ &\left.+C_{\alpha\beta}\left(2 - C_{\mu\nu}\right)\delta_{\nu\alpha}b_\mu b_\beta + \left(2 - C_{\alpha\beta}\right)C_{\mu\nu}\delta_{\mu\beta}b_\nu b_\alpha + C_{\alpha\beta}C_{\mu\nu}\delta_{\mu\alpha}b_\nu b_\beta\right]\Big\} \\ &+ \frac{\mu^2}{12\pi v_F}\delta(\Omega)\left[\left(2 - C_{\alpha\beta}\right)\left(2 - C_{\mu\nu}\right)\delta_{\nu\beta}b_\mu b_\alpha + C_{\alpha\beta}\left(2 - C_{\mu\nu}\right)\delta_{\nu\alpha}b_\mu b_\beta + \left(2 - C_{\alpha\beta}\right)C_{\mu\nu}\delta_{\mu\beta}b_\nu b_\alpha\right. \\ &\left.+C_{\alpha\beta}C_{\mu\nu}\delta_{\mu\alpha}b_\nu b_\beta\right] + \frac{\mu^4}{120\pi v_F^3}\delta(\Omega)\left(5C_{\mu\nu} + 5C_{\alpha\beta} - 5C_{\mu\nu}C_{\alpha\beta} - 2\right)\delta_{\mu\beta}\delta_{\nu\alpha} \\ &+ \frac{\mu^4}{120\pi v_F^3}\delta(\Omega)\left(8 - 5C_{\mu\nu} - 5C_{\alpha\beta} + 5C_{\mu\nu}C_{\alpha\beta}\right)\delta_{\mu\alpha}\delta_{\nu\beta} - \frac{\mu^4}{60\pi v_F^3}\delta(\Omega)\delta_{\mu\nu}\delta_{\alpha\beta}. \end{aligned} \tag{C11}$$

Here, we subtracted the contribution of completely filled states in terms with $\delta(\Omega)$.

The fourth $\eta^{(4)}_{\mu\nu\alpha\beta}(\Omega)$ and fifth $\kappa^{-1}$ terms in Eq. (C1) are

$$\begin{aligned}\eta^{(4)}_{\mu\nu\alpha\beta}(\Omega) &= \pi\delta(\Omega)\delta_{\alpha\beta}\int_{-\infty}^{+\infty} d\omega f(\omega)\int\frac{d^3k}{(2\pi)^3}\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})\right] = \delta(\Omega)\delta_{\mu\nu}\delta_{\alpha\beta}\frac{\mu^4}{12\pi v_F^3},\\ \kappa^{-1} &= -V\frac{\partial P}{\partial V} = (\epsilon+P)\frac{\partial P}{\partial\epsilon} = \frac{\mu^4}{9\pi^2 v_F^3},\end{aligned} \tag{C12}$$

where we also subtracted the contribution of completely filled states and used

$$\epsilon = \int_{-\infty}^{+\infty} d\omega f(\omega)\int\frac{d^3k}{(2\pi)^3}\mathrm{tr}\left[H(\mathbf{k})A(\omega;\mathbf{k})\right] = \frac{\mu^4}{4\pi^2 v_F^3}, \tag{C13}$$

$$\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})\right] = \sum_\lambda\frac{v_F^2}{\varepsilon_{\mathbf{k},\lambda}}\left[\delta(\omega-\varepsilon_{\mathbf{k},\lambda})-\delta(\omega+\varepsilon_{\mathbf{k},\lambda})\right]\left[k_\mu k_\nu + C_{\mu\nu}\lambda\left(k_\mu b_\nu - k_\nu b_\mu\right) - \lambda k_\mu b_\nu\right]. \tag{C14}$$

Here, $\kappa$ is the isentropic compressibility at fixed particle number, $V$ and $P$ are volume and pressure, and $\epsilon$ is the energy density.

The last term in the last line in Eq. (C1) reads

$$\begin{aligned}\eta^{(5)}_{\mu\nu\alpha\beta}(\Omega) &= \pi\delta(\Omega)\,\mathrm{Im}\int_{-\infty}^{+\infty} d\omega f(\omega)\int\frac{d^3k}{(2\pi)^3}\mathrm{tr}\left\{[T_{\mu\nu},J_{\alpha\beta}]A(\omega;\mathbf{k})\right\}\\ &= \delta(\Omega)\frac{\mu^4}{12\pi v_F^3}\left(\frac{C_{\mu\nu}-C_{\alpha\beta}+C_{\mu\nu}C_{\alpha\beta}}{2}\delta_{\mu\alpha}\delta_{\nu\beta}+\frac{2-C_{\mu\nu}+C_{\alpha\beta}-C_{\mu\nu}C_{\alpha\beta}}{2}\delta_{\mu\beta}\delta_{\nu\alpha}\right),\end{aligned} \tag{C15}$$

where we subtracted the contribution of completely filled states and used

$$\begin{aligned}\mathrm{tr}\left\{[T_{\mu\nu}(\mathbf{k}),J_{\alpha\beta}]\,A(\omega;\mathbf{k})\right\} &= \sum_\lambda\frac{iv_F^2}{\varepsilon_{\mathbf{k},\lambda}}\left[\delta(\omega-\varepsilon_{\mathbf{k},\lambda})-\delta(\omega+\varepsilon_{\mathbf{k},\lambda})\right]\\ &\times\Big\{\frac{C_{\mu\nu}}{2}(k_\mu-\lambda b_\mu)k_\alpha\delta_{\nu\beta}+\frac{C_{\alpha\beta}}{2}\left(1-\frac{C_{\mu\nu}}{2}\right)\left[(k_\beta-\lambda b_\beta)k_\mu\delta_{\nu\alpha}-(k_\alpha-\lambda b_\alpha)k_\mu\delta_{\nu\beta}\right]\\ &+\left(1-\frac{C_{\mu\nu}}{2}\right)(k_\nu-\lambda b_\nu)k_\alpha\delta_{\mu\beta}+\frac{C_{\alpha\beta}C_{\mu\nu}}{4}\left[(k_\beta-\lambda b_\beta)k_\nu\delta_{\mu\alpha}-(k_\alpha-\lambda b_\alpha)k_\nu\delta_{\mu\beta}\right]\\ &+\lambda\frac{C_{\alpha\beta}C_{\mu\nu}}{4}\left[(k_\beta-\lambda b_\beta)b_\nu\delta_{\mu\alpha}-(k_\alpha-\lambda b_\alpha)b_\nu\delta_{\mu\beta}-(k_\beta-\lambda b_\beta)b_\mu\delta_{\nu\alpha}+(k_\alpha-\lambda b_\alpha)b_\mu\delta_{\nu\beta}\right]\Big\}.\end{aligned} \tag{C16}$$

By combining terms in Eqs. (C5), (C11), (C12), and (C15), we obtain the viscoelasticity tensor whose different components are given in Eqs. (7), (12), (14), (15), (21), and (22).

### 2. Two-band model of Weyl semimetals

In this Section, we consider a two-band model of Weyl semimetals with the Hamiltonian $H(\mathbf{k}) = \boldsymbol{\sigma}\cdot\mathbf{d}(\mathbf{k})$, where $\mathbf{d} = \left\{v_F k_x, v_F k_y, \gamma(k_z^2-m)\right\}$. The strain transformation generator and the stress tensor are

$$J_{\alpha\beta} = -\frac{1}{2}\left\{x_\beta,k_\alpha\right\}+\frac{i}{8}C_{\alpha\beta}\left[\sigma_\beta,\sigma_\alpha\right] \tag{C17}$$

and

$$T_{\alpha\beta} = \sigma_i k_\alpha\partial_\beta\left(d_i\right)+\frac{1}{2}C_{\alpha\beta}\left(d_\beta\sigma_\alpha-d_\alpha\sigma_\beta\right), \tag{C18}$$

respectively. For the canonical stress tensor, $C_{\alpha\beta}=0$.

In what follows, we focus on the viscoelasticity tensor for the canonical stress tensor; terms related to the internal degrees of freedom $\propto C_{\alpha\beta}$ will be discussed at the end of this Section.

To calculate the viscoelasticity tensor (C1), we use

$$\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})\right]=\frac{1}{\varepsilon_{\mathbf{k}}}\left[\delta(\omega-\varepsilon_{\mathbf{k}})-\delta(\omega+\varepsilon_{\mathbf{k}})\right]d_i\partial_\nu\left(d_i\right)k_\mu, \tag{C19}$$

$$\begin{aligned}\mathrm{Im}\,\mathrm{tr}\left\{\left[T_{\mu\nu},J_{\alpha\beta}\right]A(\omega;\mathbf{k})\right\}=\frac{1}{\varepsilon_{\mathbf{k}}}\left[\delta(\omega-\varepsilon_{\mathbf{k}})-\delta(\omega+\varepsilon_{\mathbf{k}})\right]\Big\{&d_ik_\alpha\partial_\beta\left[k_\mu\partial_\nu\left(d_i\right)\right]+\frac{C_{\mu\nu}}{2}\left[k_\alpha\partial_\beta\left(d_\nu\right)d_\mu-k_\beta\partial_\alpha\left(d_\mu\right)d_\nu\right]\\&+\frac{C_{\alpha\beta}}{2}\left[k_\mu\partial_\nu\left(d_\beta\right)d_\alpha-k_\nu\partial_\mu\left(d_\alpha\right)d_\beta\right]+\frac{C_{\mu\nu}C_{\alpha\beta}}{4}\left(\delta_{\nu\beta}d_\alpha d_\mu-\delta_{\mu\beta}d_\alpha d_\nu+\delta_{\mu\alpha}d_\beta d_\nu-\delta_{\nu\alpha}d_\beta d_\mu\right)\Big\},\end{aligned} \tag{C20}$$

$$\begin{aligned}\mathrm{Re}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right]=&\left[\delta\left(\omega-\varepsilon_{\mathbf{k}}\right)\delta\left(\omega'+\varepsilon_{\mathbf{k}}\right)+\delta\left(\omega+\varepsilon_{\mathbf{k}}\right)\delta\left(\omega'-\varepsilon_{\mathbf{k}}\right)\right]\Big\{k_\alpha k_\mu\partial_\beta\left(d_i\right)\partial_\nu\left(d_i\right)\\&+\frac{C_{\alpha\beta}}{2}\left[k_\mu d_\beta\partial_\nu\left(d_\alpha\right)-k_\mu d_\alpha\partial_\nu\left(d_\beta\right)\right]+\frac{C_{\mu\nu}}{2}\left[k_\alpha d_\nu\partial_\beta\left(d_\mu\right)-k_\alpha d_\mu\partial_\beta\left(d_\nu\right)\right]\\&+\frac{C_{\alpha\beta}C_{\mu\nu}}{4}\left(\delta_{\mu\alpha}d_\nu d_\beta-\delta_{\nu\alpha}d_\mu d_\beta-\delta_{\mu\beta}d_\nu d_\alpha+\delta_{\nu\beta}d_\mu d_\alpha\right)\Big\}\\&+\frac{1}{\varepsilon_{\mathbf{k}}^2}k_\alpha k_\mu d_i\partial_\beta\left(d_i\right)d_j\partial_\nu\left(d_j\right)\left[\delta\left(\omega-\varepsilon_{\mathbf{k}}\right)-\delta\left(\omega+\varepsilon_{\mathbf{k}}\right)\right]\left[\delta\left(\omega'-\varepsilon_{\mathbf{k}}\right)-\delta\left(\omega'+\varepsilon_{\mathbf{k}}\right)\right],\end{aligned} \tag{C21}$$

and

$$\begin{aligned}\mathrm{Im}\,\mathrm{tr}\left[T_{\mu\nu}(\mathbf{k})A(\omega;\mathbf{k})T_{\alpha\beta}(\mathbf{k})A(\omega';\mathbf{k})\right]=&\frac{\delta\left(\omega-\varepsilon_{\mathbf{k}}\right)\delta\left(\omega'+\varepsilon_{\mathbf{k}}\right)-\delta\left(\omega+\varepsilon_{\mathbf{k}}\right)\delta\left(\omega'-\varepsilon_{\mathbf{k}}\right)}{\varepsilon_{\mathbf{k}}}\Big[d_ik_\alpha k_\mu\partial_\nu\left(d_k\right)\partial_\beta\left(d_l\right)\epsilon_{kil}\\&+\frac{C_{\alpha\beta}}{2}k_\mu d_i\partial_\nu\left(d_j\right)\left(d_\beta\epsilon_{ji\alpha}-d_\alpha\epsilon_{ji\beta}\right)+\frac{C_{\mu\nu}}{2}k_\alpha d_i\partial_\beta\left(d_j\right)\left(d_\nu\epsilon_{\mu ij}-d_\mu\epsilon_{\nu ij}\right)\\&+\frac{C_{\alpha\beta}C_{\mu\nu}}{4}d_i\left(\epsilon_{\mu i\alpha}d_\nu d_\beta-\epsilon_{\mu i\beta}d_\nu d_\alpha-\epsilon_{\nu i\alpha}d_\mu d_\beta+\epsilon_{\nu i\beta}d_\mu d_\alpha\right)\Big].\end{aligned} \tag{C22}$$

For the first term in the last line in Eq. (C1), we obtain

$$\eta^{(1)}_{\mu\nu\alpha\beta}(\Omega)=\mathrm{v.p.}\int\frac{d^3k}{(2\pi)^3}\frac{2}{\varepsilon_{\mathbf{k}}}\frac{1-\Theta\left(|\mu|-\varepsilon_{\mathbf{k}}\right)}{4\varepsilon_{\mathbf{k}}^2-\Omega^2}d_ik_\mu k_\alpha\partial_\nu\left(d_\nu\right)\partial_\beta\left(d_\beta\right)\epsilon_{i\nu\beta}=\sum_{j=1}^{3}\epsilon_{j\nu\beta}I^j_{\mu\alpha}, \tag{C23}$$

where

$$I^{i=1,2}_{\mu\alpha}=\int\frac{d^3k}{(2\pi)^3}\frac{2v_F^2}{\varepsilon_{\mathbf{k}}}\frac{\Theta\left(\varepsilon_{\mathbf{k}}-|\mu|\right)}{4\varepsilon_{\mathbf{k}}^2-\Omega^2}\partial_z\left(d_z\right)k_ik_\alpha k_\mu \tag{C24}$$

$$I^{3}_{\mu\alpha}=\int\frac{d^3k}{(2\pi)^3}\frac{2v_F^2}{\varepsilon_{\mathbf{k}}}\frac{\Theta\left(\varepsilon_{\mathbf{k}}-|\mu|\right)}{4\varepsilon_{\mathbf{k}}^2-\Omega^2}d_zk_\alpha k_\mu. \tag{C25}$$

At $\Omega=0$ and $\mu=0$, we find

$$\begin{aligned}I^{i=1,2}_{\mu\alpha}&=\left(\delta_{\mu3}\delta_{\alpha i}+\delta_{\mu i}\delta_{\alpha3}\right)\int\frac{d^3k}{(2\pi)^3}\frac{v_F^2\partial_z\left(d_z\right)}{2\varepsilon_{\mathbf{k}}^3}\frac{1}{2}k_\perp^2k_z\\&=\left(\delta_{\mu3}\delta_{\alpha i}+\delta_{\mu i}\delta_{\alpha3}\right)\left(-\frac{\gamma^2\Lambda_z^5}{10\pi^2v_F^2}+\frac{\gamma\Lambda_z^3\Lambda_\perp}{12\pi^2v_F}+\frac{\gamma^2m\Lambda_z^3}{6\pi^2v_F^2}-\frac{2\gamma^2m^{5/2}}{15\pi^2v_F^2}\right),\end{aligned} \tag{C26}$$

$$\begin{aligned}I^{3}_{\mu\alpha}=\delta_{\mu\alpha}\int\frac{d^3k}{(2\pi)^3}\frac{v_F^2d_z}{2\varepsilon_{\mathbf{k}}^3}\left[\left(\delta_{\mu1}+\delta_{\mu2}\right)\frac{1}{2}k_\perp^2+\delta_{\mu3}k_z^2\right]=\delta_{\mu\alpha}\Big[&\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(-\frac{\gamma^2\Lambda_z^5}{20\pi^2v_F^2}+\frac{\gamma^2m\Lambda_z^3}{6\pi^2v_F^2}+\frac{\gamma\Lambda_z^3\Lambda_\perp}{24\pi^2v_F}\right.\\&\left.-\frac{\gamma^2m^2\Lambda_z}{4\pi^2v_F^2}-\frac{\gamma m\Lambda_z\Lambda_\perp}{8\pi^2v_F}+\frac{4\gamma^2m^{5/2}}{15\pi^2v_F^2}\right)+\delta_{\mu3}\frac{1}{12\pi^2}\left(\Lambda_z^3-2m^{3/2}\right)\Big],\end{aligned} \tag{C27}$$

where $\Lambda_\perp$ and $\Lambda_z$ are momentum cutoffs.

In the inter-band part of $\eta^{(2)}_{\mu\nu\alpha\beta}(\Omega)+\eta^{(3)}_{\mu\nu\alpha\beta}(\Omega)$, we have

$$\begin{aligned}\eta^{(\mathrm{inter})}_{\mu\nu\alpha\beta}(\Omega) &= \frac{\pi}{\Omega}\Theta\left(\Omega-2\mu\right)\int\frac{d^3k}{(2\pi)^3}\delta\left(\Omega-2\varepsilon_{\mathbf{k}}\right)\left\{k_\alpha k_\mu\delta_{\nu\beta}\left[\partial_\nu\left(d_\nu\right)\right]^2-\frac{1}{\varepsilon^2_{\mathbf{k}}}k_\alpha k_\mu d_\beta\partial_\beta\left(d_\beta\right)d_\nu\partial_\nu\left(d_\nu\right)\right\}\\ &= \Theta\left(\Omega-2\mu\right)\left(I_1^{(\mathrm{inter})}-I_2^{(\mathrm{inter})}\right).\end{aligned} \tag{C28}$$

Here, we used the fact that in the model at hand, $d_i=d_i(k_i)$, and, therefore, $\partial_\mu(d_\nu)=\partial_\nu(d_\nu)\delta_{\mu\nu}$. The integrals $I_1^{(\mathrm{inter})}$ and $I_2^{(\mathrm{inter})}$ read

$$\begin{aligned}I_1^{(\mathrm{inter})} &= \frac{1}{32\pi^2v_F^2}\int_0^{2\pi}d\phi\int_{-\infty}^{+\infty}dk_z k_\alpha k_\mu\delta_{\nu\beta}\left[\partial_\nu\left(d_\nu\right)\right]^2\Theta\left(\Omega^2-4d_z^2\right)\Big|_{v_F^2k_\perp^2=\frac{\Omega^2}{4}-d_z^2}=\frac{\delta_{\mu\alpha}\delta_{\nu\beta}}{8\pi v_F^2}\Bigg\{\frac{\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(\delta_{\nu1}+\delta_{\nu2}\right)}{30}\\ &\times\left[\left(3\Omega^2+2\gamma m\Omega-8\gamma^2m^2\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\left(3\Omega^2-2\gamma m\Omega-8\gamma^2m^2\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]\\ &+\frac{\delta_{\mu3}\delta_{\nu3}}{5}\left[\left(2\gamma m+\Omega\right)^2\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\left(2\gamma m-\Omega\right)^2\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]\\ &-\frac{\gamma\delta_{\nu3}\left(\delta_{\mu1}+\delta_{\mu2}\right)}{105v_F^2}\left[\left(4\gamma m-5\Omega\right)\left(2\gamma m+\Omega\right)^2\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\left(4\gamma m+5\Omega\right)\left(2\gamma m-\Omega\right)^2\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]\\ &+\frac{v_F^2\delta_{\mu3}\left(\delta_{\nu1}+\delta_{\nu2}\right)}{6\gamma}\left[\left(2\gamma m+\Omega\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\left(2\gamma m-\Omega\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]\Bigg\}\end{aligned} \tag{C29}$$

and

$$\begin{aligned}I_2^{(\mathrm{inter})} &= \frac{1}{32\pi^2v_F^2}\int_0^{2\pi}d\phi\int_{-\infty}^{+\infty}dk_z\frac{1}{\varepsilon^2_{\mathbf{k}}}k_\alpha k_\mu d_\beta\partial_\beta\left(d_\beta\right)d_\nu\partial_\nu\left(d_\nu\right)\Theta\left(\Omega^2-4d_z^2\right)\Big|_{v_F^2k_\perp^2=\frac{\Omega^2}{4}-d_z^2}\\ &=\frac{1}{2\pi v_F^2\Omega^2}\Bigg\{\left[\delta_{\mu\beta}\delta_{\nu\alpha}+\delta_{\mu\nu}\delta_{\alpha\beta}+\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(\delta_{\nu1}+\delta_{\nu2}\right)\delta_{\mu\alpha}\delta_{\nu\alpha}\right]\frac{1}{1260}\Bigg[\left(2\gamma m+\Omega\right)^2\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}\\ &\times\left(16\gamma^2m^2-20\gamma m\Omega+7\Omega^2\right)-\Theta\left(2\gamma m-\Omega\right)\left(2\gamma m-\Omega\right)^2\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\left(16\gamma^2m^2+20\gamma m\Omega+7\Omega^2\right)\Bigg]\\ &+\delta_{\mu3}\delta_{\nu3}\delta_{\mu\alpha}\delta_{\nu\beta}\frac{\Omega^2}{60}\left[\left(2\gamma m+\Omega\right)^2\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\left(2\gamma m-\Omega\right)^2\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]\\ &-\left(\delta_{\mu1}+\delta_{\mu2}\right)\delta_{\nu3}\delta_{\mu\alpha}\delta_{\nu\beta}\frac{\gamma}{13860v_F^2}\Bigg[\left(256\gamma^3m^3-320\gamma^2m^2\Omega+148\gamma m\Omega^2-45\Omega^3\right)\left(2\gamma m+\Omega\right)^2\\ &\times\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\left(256\gamma^3m^3+320\gamma^2m^2\Omega+148\gamma m\Omega^2+45\Omega^3\right)\left(2\gamma m-\Omega\right)^2\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\Bigg]\\ &-\delta_{\mu3}\left(\delta_{\nu1}+\delta_{\nu2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\frac{v_F^2}{420\gamma}\Bigg[\left(2\gamma m+\Omega\right)^2\left(4\gamma m-5\Omega\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\left(2\gamma m-\Omega\right)^2\\ &\times\left(4\gamma m+5\Omega\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\Bigg]\Bigg\}.\end{aligned} \tag{C30}$$

Subtracting the divergent contribution of the filled states, the intra-band part of $\eta^{(2)}_{\mu\nu\alpha\beta}(\Omega)+\eta^{(3)}_{\mu\nu\alpha\beta}(\Omega)$ reads

$$\begin{aligned}\eta^{(\mathrm{intra})}_{\mu\nu\alpha\beta}(\Omega) &= \pi\delta(\Omega)\,\mathrm{v.p.}\int\frac{d^3k}{(2\pi)^3}\frac{f(\varepsilon_{\mathbf{k}})-f(-\varepsilon_{\mathbf{k}})}{\varepsilon_{\mathbf{k}}}\left\{k_\alpha k_\mu\delta_{\nu\beta}\left[\partial_\nu\left(d_\nu\right)\right]^2-\frac{1}{\varepsilon^2_{\mathbf{k}}}k_\alpha k_\mu d_\beta\partial_\beta\left(d_\beta\right)d_\nu\partial_\nu\left(d_\nu\right)\right\}\\ &= \delta(\Omega)\left(I_1^{(\mathrm{intra})}-I_2^{(\mathrm{intra})}\right),\end{aligned} \tag{C31}$$

where

$$
\begin{aligned}
I_1^{(\mathrm{intra})} &= \pi \int \frac{d^3k}{(2\pi)^3}\frac{1}{|\mathbf{d}(\mathbf{k})|}k_\mu k_\alpha \delta_{\nu\beta}\left(\partial_\nu(d_\nu)\right)^2 \Theta\left(|\mu|-|\mathbf{d}(\mathbf{k})|\right) = \frac{\delta_{\mu\alpha}\delta_{\nu\beta}}{\pi v_F^2}\Bigg\{\frac{2\left(\delta_{\mu 1}+\delta_{\mu 2}\right)\left(\delta_{\nu 1}+\delta_{\nu 2}\right)}{105}\Bigg[8\gamma^3 m^{7/2} \\
&+ (|\mu|+\gamma m)^2(3|\mu|-4\gamma m)\sqrt{\frac{\gamma m+|\mu|}{\gamma}} - \Theta\left(\gamma m-|\mu|\right)(|\mu|-\gamma m)^2(3|\mu|+4\gamma m)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg] \\
&+ \frac{v_F^2\delta_{\mu 3}\left(\delta_{\nu 1}+\delta_{\nu 2}\right)}{15\gamma}\left[-2\gamma^2 m^{5/2} + (|\mu|+\gamma m)^2\sqrt{\frac{\gamma m+|\mu|}{\gamma}} + \Theta\left(\gamma m-|\mu|\right)(|\mu|-\gamma m)^2\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\right] \\
&+ \frac{4\delta_{\mu 3}\delta_{\nu 3}}{35}\left[-2\gamma^3 m^{7/2} + (|\mu|+\gamma m)^3\sqrt{\frac{\gamma m+|\mu|}{\gamma}} - \Theta\left(\gamma m-|\mu|\right)(|\mu|-\gamma m)^3\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\right] \\
&- \frac{8\gamma\left(\delta_{\mu 1}+\delta_{\mu 2}\right)\delta_{\nu 3}}{945 v_F^2}\Bigg[-8\gamma^4 m^{9/2} + (|\mu|+\gamma m)^3(4\gamma m-5|\mu|)\sqrt{\frac{\gamma m+|\mu|}{\gamma}} - \Theta\left(\gamma m-|\mu|\right)(|\mu|-\gamma m)^3 \\
&\times(4\gamma m+5|\mu|)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg]\Bigg\}
\end{aligned}
\tag{C32}
$$

and

$$
\begin{aligned}
I_2^{(\mathrm{intra})} &= \pi\int\frac{d^3k}{(2\pi)^3}\frac{1}{|\mathbf{d}(\mathbf{k})|}k_\mu\frac{d_\nu\partial_\nu(d_\nu)}{|\mathbf{d}(\mathbf{k})|}k_\alpha\frac{d_\beta\partial_\beta(d_\beta)}{|\mathbf{d}(\mathbf{k})|}\Theta\left(|\mu|-|\mathbf{d}(\mathbf{k})|\right) = \\
&= \frac{1}{\pi v_F^2|\mu|}\Bigg\{\frac{4}{315}\left[\delta_{\mu\beta}\delta_{\nu\alpha}+\delta_{\mu\nu}\delta_{\alpha\beta}+\left(\delta_{\mu 1}+\delta_{\mu 2}\right)\left(\delta_{\nu 1}+\delta_{\nu 2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\right] \\
&\times\left[12|\mu|\gamma^3 m^{7/2} + (|\mu|+\gamma m)^3(|\mu|-2\gamma m)\sqrt{\frac{\gamma m+|\mu|}{\gamma}} - \Theta\left(\gamma m-|\mu|\right)(|\mu|-\gamma m)^3(|\mu|+2\gamma m)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\right] \\
&+ \frac{2v_F^2}{105\gamma}\delta_{\mu 3}\left(\delta_{\nu 1}+\delta_{\nu 2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\left[-7|\mu|\gamma^2 m^{5/2} + (|\mu|+\gamma m)^3\sqrt{\frac{\gamma m+|\mu|}{\gamma}} - \Theta\left(\gamma m-|\mu|\right)(\gamma m-|\mu|)^3\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\right] \\
&+ \frac{4}{315}\delta_{\mu 3}\delta_{\nu 3}\delta_{\mu\alpha}\delta_{\nu\beta}\left[-6|\mu|\gamma^3 m^{7/2} + 3|\mu|(|\mu|+\gamma m)^3\sqrt{\frac{\gamma m+|\mu|}{\gamma}} - \Theta\left(\gamma m-|\mu|\right)3|\mu|(|\mu|-\gamma m)^3\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\right] \\
&+ \frac{8\gamma}{3465 v_F^2}\left(\delta_{\mu 1}+\delta_{\mu 2}\right)\delta_{\nu 3}\delta_{\mu\alpha}\delta_{\nu\beta}\Bigg[-88\gamma^4 m^{9/2}|\mu| + \left(5\mu^2-12\gamma m|\mu|+16\gamma^2 m^2\right)(|\mu|+\gamma m)^3\sqrt{\frac{\gamma m+|\mu|}{\gamma}} \\
&+\Theta\left(\gamma m-|\mu|\right)\left(5\mu^2+12\gamma m|\mu|+16\gamma^2 m^2\right)(|\mu|-\gamma m)^3\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg]\Bigg\}.
\end{aligned}
\tag{C33}
$$

The fourth and fifth terms in the last line in Eq. (C1) are

$$
\begin{aligned}
\eta^{(4)}_{\mu\nu\alpha\beta}(\Omega) &= \pi\delta(\Omega)\delta_{\alpha\beta}\int\frac{d^3k}{(2\pi)^3}\frac{1}{\varepsilon_{\mathbf{k}}}\left[f\left(\varepsilon_{\mathbf{k}}\right)-f\left(-\varepsilon_{\mathbf{k}}\right)\right]d_i\partial_\nu\left(d_i\right)k_\mu \\
&= \frac{2}{105\pi v_F^2}\delta(\Omega)\delta_{\mu\nu}\delta_{\alpha\beta}\Bigg[8\gamma^3 m^{7/2} - (\gamma m+|\mu|)^2(4\gamma m-3|\mu|)\sqrt{\frac{\gamma m+|\mu|}{\gamma}} \\
&-\Theta(\gamma m-|\mu|)(\gamma m-|\mu|)^2(4\gamma m+3|\mu|)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg] = \pi\delta(\Omega)\delta_{\mu\nu}\delta_{\alpha\beta}P, \\
\kappa^{-1} &= P\left(1+\alpha\right),
\end{aligned}
\tag{C34}
$$

where $\alpha=\partial P/\partial\epsilon=\partial P/\partial\mu\left[\partial\epsilon/\partial\mu\right]^{-1}$ and

$$\epsilon=\int_{-\infty}^{+\infty}d\omega f(\omega)\int\frac{d^3k}{(2\pi)^3}\mathrm{tr}\left[H(\mathbf{k})A(\omega;\mathbf{k})\right]=\frac{1}{105\pi^2v_F^2}\Bigg[-16\gamma^3m^{7/2}+\left(8\gamma^3m^3-4\gamma^2m^2|\mu|+3\gamma m\mu^2+15|\mu|^3\right)\times\sqrt{\frac{\gamma m+|\mu|}{\gamma}}+\Theta(\gamma m-|\mu|)\left(8\gamma^3m^3+4\gamma^2m^2|\mu|+3\gamma m\mu^2-15|\mu|^3\right)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg]. \tag{C35}$$

For the last term in the last line in Eq. (C1), we have

$$\begin{aligned}\eta^{(5)}_{\mu\nu\alpha\beta}(\Omega)&=\pi\delta(\Omega)\int\frac{d^3k}{(2\pi)^3}\frac{f(\varepsilon_{\mathbf{k}})-f(-\varepsilon_{\mathbf{k}})}{\varepsilon_{\mathbf{k}}}d_ik_\alpha\partial_\beta\left[k_\mu\partial_\nu\left(d_i\right)\right]\\&=\pi\delta(\Omega)\int\frac{d^3k}{(2\pi)^3}\frac{f(\varepsilon)-f(-\varepsilon)}{\varepsilon}\left\{\left(\delta_{\nu1}+\delta_{\nu2}\right)v_F^2k_\nu k_\alpha\delta_{\mu\beta}+\delta_{\nu3}d_\nu k_\alpha\partial_\beta\left[k_\mu\partial_\nu\left(d_\nu\right)\right]\right\}\\&=\pi\delta(\Omega)\delta_{\mu\beta}\delta_{\nu\alpha}P+\delta_{\mu3}\delta_{\nu3}\pi\delta(\Omega)\delta_{\mu\alpha}\delta_{\nu\beta}P-\frac{4\gamma\left(\delta_{\mu1}+\delta_{\mu2}\right)\delta_{\nu3}}{945\pi v_F^4}\delta(\Omega)\delta_{\mu\alpha}\delta_{\nu\beta}\\&\times\Bigg[64\gamma^4m^{9/2}-\left(32\gamma^2m^2-26\gamma m|\mu|+5\mu^2\right)\left(\gamma m+|\mu|\right)^2\sqrt{\frac{\gamma m+|\mu|}{\gamma}}\\&-\Theta(\gamma m-|\mu|)\left(32\gamma^2m^2+26\gamma m|\mu|+5\mu^2\right)\left(\gamma m-|\mu|\right)^2\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg].\end{aligned} \tag{C36}$$

The complete expressions for the static and dynamic parts of the viscoelasticity tensor read

$$\begin{aligned}\mathrm{Re}\,\eta^{\mathrm{stat}}_{\mu\nu\alpha\beta}(\Omega)=\frac{\delta(\Omega)}{\pi v_F^2|\mu|}\Bigg\{&\left[\delta_{\mu\beta}\delta_{\nu\alpha}+\delta_{\mu\nu}\delta_{\alpha\beta}+\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(\delta_{\nu1}+\delta_{\nu2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\right]\frac{2}{315}\Bigg[\left(4\gamma^2m^2-10\gamma m|\mu|+7\mu^2\right)\\&\times\left(\gamma m+|\mu|\right)^2\sqrt{\frac{\gamma m+|\mu|}{\gamma}}-\Theta\left(\gamma m-|\mu|\right)\left(4\gamma^2m^2+10\gamma m|\mu|+7\mu^2\right)\left(\gamma m-|\mu|\right)^2\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg]\\&+\delta_{\mu3}\delta_{\nu3}\delta_{\mu\alpha}\delta_{\nu\beta}\frac{2\mu^2}{15}\left[\left(\gamma m+|\mu|\right)^2\sqrt{\frac{\gamma m+|\mu|}{\gamma}}-\Theta\left(\gamma m-|\mu|\right)\left(\gamma m-|\mu|\right)^2\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\right]\\&-\delta_{\mu3}\left(\delta_{\nu1}+\delta_{\nu2}\right)\delta_{\mu\alpha}\delta_{\nu\beta}\frac{v_F^2}{105\gamma}\Bigg[\left(\gamma m+|\mu|\right)^2\left(2\gamma m-5|\mu|\right)\sqrt{\frac{\gamma m+|\mu|}{\gamma}}-\Theta\left(\gamma m-|\mu|\right)\left(\gamma m-|\mu|\right)^2\\&\times\left(2\gamma m+5|\mu|\right)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg]-\left(\delta_{\mu1}+\delta_{\mu2}\right)\delta_{\nu3}\delta_{\mu\alpha}\delta_{\nu\beta}\frac{4\gamma}{3465v_F^2}\Bigg[\left(\gamma m+|\mu|\right)^2\sqrt{\frac{\gamma m+|\mu|}{\gamma}}\\&\times\left(32\gamma^3m^3-80\gamma^2m^2|\mu|+74\gamma m\mu^2-45|\mu|^3\right)-\Theta\left(\gamma m-|\mu|\right)\left(\gamma m-|\mu|\right)^2\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\\&\times\left(32\gamma^3m^3+80\gamma^2m^2|\mu|+74\gamma m\mu^2+45|\mu|^3\right)\Bigg]\Bigg\}-\pi\delta(\Omega)\delta_{\mu\nu}\delta_{\alpha\beta}\left(1+\alpha\right)P\end{aligned} \tag{C37}$$

and

$$\begin{aligned}
\mathrm{Re}\,\eta^{\mathrm{dyn}}_{\mu\nu\alpha\beta}(\Omega) = \Bigg\{ & \Bigg[ (2\gamma m+\Omega)\left(49\Omega^3-72\gamma m\Omega^2+48\gamma^2m^2\Omega-64\gamma^3m^3\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}} \\
& +\Theta\left(2\gamma m-\Omega\right)(2\gamma m-\Omega)\left(49\Omega^3+72\gamma m\Omega^2+48\gamma^2m^2\Omega+64\gamma^3m^3\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\Bigg] \\
& \times\frac{\delta_{\mu\alpha}\delta_{\nu\beta}\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(\delta_{\nu1}+\delta_{\nu2}\right)}{5040\pi\Omega^2v_F^2} - \Bigg[(2\gamma m+\Omega)^2\left(16\gamma^2m^2-20\gamma m\Omega+7\Omega^2\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}} \\
& -\Theta\left(2\gamma m-\Omega\right)(2\gamma m-\Omega)^2\left(16\gamma^2m^2+20\gamma m\Omega+7\Omega^2\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\Bigg]\frac{\delta_{\mu\beta}\delta_{\nu\alpha}+\delta_{\mu\nu}\delta_{\alpha\beta}}{2520\pi v_F^2\Omega^2} \\
& +\left[(2\gamma m+\Omega)^2\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)(2\gamma m-\Omega)^2\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]\frac{\delta_{\mu\alpha}\delta_{\nu\beta}\delta_{\mu3}\delta_{\nu3}}{60\pi v_F^2} \\
& +\Bigg[(2\gamma m+\Omega)\left(16\gamma^2m^2-12\gamma m\Omega+25\Omega^2\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right) \\
& \times(2\gamma m-\Omega)\left(16\gamma^2m^2+12\gamma m\Omega+25\Omega^2\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\Bigg]\frac{\delta_{\mu\alpha}\delta_{\nu\beta}\delta_{\mu3}\left(\delta_{\nu1}+\delta_{\nu2}\right)}{1680\pi\gamma\Omega^2} \\
& +\Bigg[\gamma(2\gamma m+\Omega)^3\left(16\gamma^2m^2-28\gamma m\Omega+15\Omega^2\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\gamma(2\gamma m-\Omega)^3 \\
& \times\left(16\gamma^2m^2+28\gamma m\Omega+15\Omega^2\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\Bigg]\frac{\delta_{\mu\alpha}\delta_{\nu\beta}\delta_{\nu3}\left(\delta_{\mu1}+\delta_{\mu2}\right)}{3465\pi\Omega^2v_F^4}\Bigg\}\Theta\left(\Omega-2\mu\right),
\end{aligned} \tag{C38}$$

respectively.

Let us calculate terms determined by the internal degrees of freedom, i.e., terms proportional to $\propto C_{\alpha\beta}$ with $\alpha,\beta=1,2$; we use the in-plane rotational symmetry of the model to fix $C_{\mu3}=C_{3\mu}=0$. At $\Omega=\mu=0$, the correction for the Hall viscosity reads

$$\mathrm{Re}\,\delta\eta^{\mathrm{Hall}}_{\mu\nu\alpha\beta}(\Omega) = \left(I^{(1)}_{\mu\nu\alpha\beta}-I^{(1)}_{\mu\nu\beta\alpha}\right)+\left(I^{(1)}_{\alpha\beta\mu\nu}-I^{(1)}_{\alpha\beta\nu\mu}\right)+\left(I^{(2)}_{\mu\nu\alpha\beta}-I^{(2)}_{\mu\nu\beta\alpha}-I^{(2)}_{\nu\mu\alpha\beta}+I^{(2)}_{\nu\mu\beta\alpha}\right), \tag{C39}$$

where

$$\begin{aligned}
I^{(1)}_{\mu\nu\alpha\beta} = C_{\alpha\beta}\int\frac{d^3k}{(2\pi)^3}\frac{1}{4\varepsilon^3_{\mathbf{k}}}k_\mu d_i d_\beta\partial_\nu\left(d_\nu\right)\epsilon_{\nu i\alpha} = C_{\alpha\beta}\Bigg[\delta_{\mu\beta}\left(\delta_{\beta1}+\delta_{\beta2}\right)\left(\delta_{\alpha1}\delta_{\nu2}-\delta_{\alpha2}\delta_{\nu1}\right) \\
\times\left(-\frac{\gamma^2\Lambda_z^5}{40\pi^2v_F^2}+\frac{\gamma\Lambda_z^3\Lambda_\perp}{48\pi^2v_F}+\frac{\gamma^2m\Lambda_z^3}{12\pi^2v_F^2}-\frac{\gamma m\Lambda_z\Lambda_\perp}{16\pi^2v_F}-\frac{\gamma^2m^2\Lambda_z}{8\pi^2v_F^2}+\frac{2\gamma^2m^{5/2}}{15\pi^2v_F^2}\right) \\
+\delta_{\mu\nu}\delta_{\nu3}\left(\delta_{\alpha2}\delta_{\beta1}-\delta_{\alpha1}\delta_{\beta2}\right)\left(-\frac{\gamma^2\Lambda_z^5}{20\pi^2v_F^2}+\frac{\gamma\Lambda_z^3\Lambda_\perp}{24\pi^2v_F}+\frac{\gamma^2m\Lambda_z^3}{12\pi^2v_F^2}-\frac{\gamma^2m^{5/2}}{15\pi^2v_F^2}\right)\Bigg],
\end{aligned} \tag{C40}$$

$$\begin{aligned}
I^{(2)}_{\mu\nu\alpha\beta} = C_{\alpha\beta}C_{\mu\nu}\int\frac{d^3k}{(2\pi)^3}\frac{1}{8\varepsilon^3_{\mathbf{k}}}\epsilon_{\mu i\alpha}d_i d_\beta d_\nu = C_{\alpha\beta}C_{\mu\nu}\delta_{\nu\beta}\left(\delta_{\beta1}+\delta_{\beta2}\right)\left(\delta_{\alpha1}\delta_{\mu2}-\delta_{\alpha2}\delta_{\mu1}\right) \\
\times\left(-\frac{\gamma^2\Lambda_z^5}{80\pi^2v_F^2}+\frac{\gamma\Lambda_z^3\Lambda_\perp}{96\pi^2v_F}+\frac{\gamma^2m\Lambda_z^3}{24\pi^2v_F^2}-\frac{\gamma m\Lambda_z\Lambda_\perp}{32\pi^2v_F}-\frac{\gamma^2m^2\Lambda_z}{16\pi^2v_F^2}+\frac{\gamma^2m^{5/2}}{15\pi^2v_F^2}\right).
\end{aligned} \tag{C41}$$

After subtraction discussed in the paragraph before Eq. (37), we have

$$\begin{aligned}
\overline{I}^{(1)}_{\mu\nu\alpha\beta} = C_{\alpha\beta}\Bigg[\delta_{\mu\beta}\left(\delta_{\beta1}+\delta_{\beta2}\right)\left(\delta_{\alpha1}\delta_{\nu2}-\delta_{\alpha2}\delta_{\nu1}\right)\left(-\frac{\gamma^2m\Lambda_z^3}{12\pi^2v_F^2}+\frac{\gamma m\Lambda_z\Lambda_\perp}{16\pi^2v_F}+\frac{\gamma^2m^2\Lambda_z}{8\pi^2v_F^2}-\frac{2\gamma^2m^{5/2}}{15\pi^2v_F^2}\right) \\
+\delta_{\mu\nu}\delta_{\nu3}\left(\delta_{\alpha2}\delta_{\beta1}-\delta_{\alpha1}\delta_{\beta2}\right)\left(-\frac{\gamma^2m\Lambda_z^3}{12\pi^2v_F^2}+\frac{\gamma^2m^{5/2}}{15\pi^2v_F^2}\right)\Bigg],
\end{aligned} \tag{C42}$$

$$\overline{I}^{(2)}_{\mu\nu\alpha\beta}=C_{\alpha\beta}C_{\mu\nu}\delta_{\nu\beta}\left(\delta_{\beta1}+\delta_{\beta2}\right)\left(\delta_{\alpha1}\delta_{\mu2}-\delta_{\alpha2}\delta_{\mu1}\right)\left(-\frac{\gamma^2m\Lambda_z^3}{24\pi^2v_F^2}+\frac{\gamma m\Lambda_z\Lambda_\perp}{32\pi^2v_F}+\frac{\gamma^2m^2\Lambda_z}{16\pi^2v_F^2}-\frac{\gamma^2m^{5/2}}{15\pi^2v_F^2}\right). \tag{C43}$$

The correction in the dynamic part reads

$$\begin{aligned}\delta\eta^{(2)}_{\mu\nu\alpha\beta}(\Omega)=\frac{\pi}{\Omega}\Theta\left(\Omega-2\mu\right)\int\frac{d^3k}{(2\pi)^3}\delta\left(\Omega-2\left|\mathbf{d}(\mathbf{k})\right|\right)\left\{\frac{C_{\alpha\beta}}{2}\left[k_\mu d_\beta\partial_\nu\left(d_\alpha\right)-k_\mu d_\alpha\partial_\nu\left(d_\beta\right)\right]\right.\\ \left.+\frac{C_{\mu\nu}}{2}\left[k_\alpha d_\nu\partial_\beta\left(d_\mu\right)-k_\alpha d_\mu\partial_\beta\left(d_\nu\right)\right]+\frac{C_{\alpha\beta}C_{\mu\nu}}{4}\left(\delta_{\mu\alpha}d_\nu d_\beta-\delta_{\nu\alpha}d_\mu d_\beta-\delta_{\mu\beta}d_\nu d_\alpha+\delta_{\nu\beta}d_\mu d_\alpha\right)\right\}.\end{aligned} \tag{C44}$$

Here, there are the following two types of integrals:

$$\begin{aligned}I_3^{(\mathrm{inter})}&=\frac{\pi}{\Omega}\int\frac{d^3k}{(2\pi)^3}k_\mu d_\alpha\partial_\nu\left(d_\beta\right)\delta\left(\Omega-2\left|\mathbf{d}(\mathbf{k})\right|\right)=\frac{\delta_{\mu\alpha}\delta_{\nu\beta}}{240\pi v_F^2}\left[\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(\delta_{\nu1}+\delta_{\nu2}\right)+\delta_{\mu3}\delta_{\nu3}\right]\\ &\times\left[\left(3\Omega^2+2\gamma m\Omega-8\gamma^2m^2\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}-\Theta\left(2\gamma m-\Omega\right)\left(3\Omega^2-2\gamma m\Omega-8\gamma^2m^2\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]\\ &\equiv\delta_{\mu\alpha}\delta_{\nu\beta}\left[\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(\delta_{\nu1}+\delta_{\nu2}\right)+\delta_{\mu3}\delta_{\nu3}\right]g_1(\Omega)\end{aligned} \tag{C45}$$

and

$$\begin{aligned}I_4^{(\mathrm{inter})}&=\frac{\pi}{\Omega}\int\frac{d^3k}{(2\pi)^3}d_\mu d_\alpha\delta_{\nu\beta}\delta\left(\Omega-2\left|\mathbf{d}(\mathbf{k})\right|\right)=\delta_{\mu\alpha}\delta_{\nu\beta}\left\{\frac{\left(\delta_{\mu1}+\delta_{\mu2}\right)}{240\pi v_F^2}\left[\left(3\Omega^2+2\gamma m\Omega-8\gamma^2m^2\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}\right.\right.\\ &\left.-\Theta\left(2\gamma m-\Omega\right)\left(3\Omega^2-2\gamma m\Omega-8\gamma^2m^2\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]+\frac{\delta_{\mu3}}{480\pi v_F^2}\left[\left(3\Omega^2-8\gamma m\Omega+32\gamma^2m^2\right)\sqrt{\frac{2\gamma m+\Omega}{2\gamma}}\right.\\ &\left.\left.-\Theta\left(2\gamma m-\Omega\right)\left(3\Omega^2+\gamma m\Omega+32\gamma^2m^2\right)\sqrt{\frac{2\gamma m-\Omega}{2\gamma}}\right]\right\}\equiv\delta_{\mu\alpha}\delta_{\nu\beta}\left[\left(\delta_{\mu1}+\delta_{\mu2}\right)g_1(\Omega)+\delta_{\mu3}g_2(\Omega)\right].\end{aligned} \tag{C46}$$

Hence

$$\begin{aligned}\delta\eta^{(2)}_{\mu\nu\alpha\beta}(\Omega)=\left(\delta_{\mu\beta}\delta_{\nu\alpha}-\delta_{\mu\alpha}\delta_{\nu\beta}\right)\Theta\left(\Omega-2\mu\right)\left\{\left(\frac{C_{\mu\nu}}{2}+\frac{C_{\alpha\beta}}{2}\right)\left[\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(\delta_{\nu1}+\delta_{\nu2}\right)+\delta_{\mu3}\delta_{\nu3}\right]g_1(\Omega)\right.\\ \left.-\frac{C_{\mu\nu}C_{\alpha\beta}}{4}\left[\left(\delta_{\mu1}+\delta_{\mu2}\right)g_1(\Omega)+\delta_{\mu3}g_2(\Omega)\right]-\frac{C_{\mu\nu}C_{\alpha\beta}}{4}\left[\left(\delta_{\nu1}+\delta_{\nu2}\right)g_1(\Omega)+\delta_{\nu3}g_2(\Omega)\right]\right\}.\end{aligned} \tag{C47}$$

The rotational symmetry with respect to the $z$-axis allows us to set $C_{3\mu}=C_{\mu3}=0$ and simplify the result,

$$\mathrm{Re}\,\delta\eta^{\mathrm{dyn}}_{\mu\nu\alpha\beta}(\Omega)=\left(\delta_{\mu\beta}\delta_{\nu\alpha}-\delta_{\mu\alpha}\delta_{\nu\beta}\right)\Theta\left(\Omega-2\mu\right)\left(\frac{C_{\mu\nu}}{2}+\frac{C_{\alpha\beta}}{2}-\frac{C_{\mu\nu}C_{\alpha\beta}}{4}\right)\left(\delta_{\mu1}+\delta_{\mu2}\right)\left(\delta_{\nu1}+\delta_{\nu2}\right)g_1(\Omega), \tag{C48}$$

where $g_1(\Omega)$ follows from Eq. (C45).

The corrections to the static part read

$$\begin{aligned}\delta\eta^{(3)}_{\mu\nu\alpha\beta}(\Omega)=\pi\delta(\Omega)\,\mathrm{v.p.}\int\frac{d^3k}{(2\pi)^3}\frac{\Theta\left(|\mu|-|\mathbf{d}(\mathbf{k})|\right)}{|\mathbf{d}(\mathbf{k})|}\left[\frac{C_{\mu\nu}}{2}\left(k_\alpha d_\nu\delta_{\mu\beta}-k_\alpha d_\mu\delta_{\nu\beta}\right)\partial_\beta\left(d_\beta\right)\right.\\ \left.+\frac{C_{\alpha\beta}}{2}\left(k_\mu d_\beta\delta_{\nu\alpha}-k_\mu d_\alpha\delta_{\nu\beta}\right)\partial_\nu\left(d_\nu\right)+\frac{C_{\mu\nu}C_{\alpha\beta}}{4}\left(\delta_{\mu\alpha}d_\nu d_\beta-\delta_{\nu\alpha}d_\mu d_\beta-\delta_{\mu\beta}d_\nu d_\alpha+\delta_{\nu\beta}d_\mu d_\alpha\right)\right]\end{aligned} \tag{C49}$$

and

$$\begin{aligned}\delta\eta^{(5)}_{\mu\nu\alpha\beta}(\Omega)=\pi\delta(\Omega)\int\frac{d^3k}{(2\pi)^3}\frac{\Theta\left(|\mu|-|\mathbf{d}(\mathbf{k})|\right)}{|\mathbf{d}(\mathbf{k})|}\left\{\frac{C_{\mu\nu}}{2}\left[k_\alpha\partial_\beta\left(d_\nu\right)d_\mu-k_\beta\partial_\alpha\left(d_\mu\right)d_\nu\right]\right.\\ \left.+\frac{C_{\alpha\beta}}{2}\left[k_\mu\partial_\nu\left(d_\beta\right)d_\alpha-k_\nu\partial_\mu\left(d_\alpha\right)d_\beta\right]+\frac{C_{\mu\nu}C_{\alpha\beta}}{4}\left(\delta_{\nu\beta}d_\alpha d_\mu-\delta_{\mu\beta}d_\alpha d_\nu+\delta_{\mu\alpha}d_\beta d_\nu-\delta_{\nu\alpha}d_\beta d_\mu\right)\right\}.\end{aligned} \tag{C50}$$

The integrals here are

$$I_3^{(\mathrm{intra})} = \pi \int \frac{d^3k}{(2\pi)^3} \frac{1}{|\mathbf{d}(\mathbf{k})|} k_\mu d_\alpha \partial_\nu \left(d_\beta\right) \Theta\left(|\mu| - |\mathbf{d}(\mathbf{k})|\right) = \frac{\delta_{\mu\alpha}\delta_{\nu\beta}}{4\pi v_F^2} \left[\left(\delta_{\mu 1} + \delta_{\mu 2}\right)\left(\delta_{\nu 1} + \delta_{\nu 2}\right) + \delta_{\mu 3}\delta_{\nu 3}\right]$$
$$\times \left\{ \frac{8}{105} \left[ 8\gamma^3 m^{7/2} - (|\mu| + \gamma m)^2 (4\gamma m - 3|\mu|) \sqrt{\frac{\gamma m + |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) (|\mu| - \gamma m)^2 \right.\right. \tag{C51}$$
$$\left.\left. \times (4\gamma m + 3|\mu|) \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \right] \right\} \equiv \delta_{\mu\alpha}\delta_{\nu\beta} \left[\left(\delta_{\mu 1} + \delta_{\mu 2}\right)\left(\delta_{\nu 1} + \delta_{\nu 2}\right) + \delta_{\mu 3}\delta_{\nu 3}\right] f_1(\mu)$$

and

$$I_4^{(\mathrm{intra})} = \pi \int \frac{d^3k}{(2\pi)^3} \frac{1}{|\mathbf{d}(\mathbf{k})|} d_\mu d_\alpha \delta_{\nu\beta} \Theta\left(|\mu| - |\mathbf{d}(\mathbf{k})|\right) = \frac{\delta_{\mu\alpha}\delta_{\nu\beta}}{4\pi v_F^2} \left\{ \left(\delta_{\mu 1} + \delta_{\mu 2}\right) \frac{8}{105} \left[ 8\gamma^3 m^{7/2} - (|\mu| + \gamma m)^2 \right.\right.$$
$$\times (4\gamma m - 3|\mu|) \sqrt{\frac{\gamma m - |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) (|\mu| - \gamma m)^2 (4\gamma m + 3|\mu|) \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \Big]$$
$$+ \delta_{\mu 3} \frac{4}{105} \left[ -48\gamma^3 m^{7/2} + \left(24\gamma^2 m^2 - 8\gamma m|\mu| + 3\mu^2\right) (|\mu| + \gamma m) \sqrt{\frac{\gamma m + |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) \right. \tag{C52}$$
$$\left.\left. \times \left(24\gamma^2 m^2 + 8\gamma m|\mu| + 3\mu^2\right) (|\mu| - \gamma m) \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \right] \right\} \equiv \delta_{\mu\alpha}\delta_{\nu\beta} \left[\left(\delta_{\mu 1} + \delta_{\mu 2}\right) f_1(\mu) + \delta_{\mu 3} f_2(\mu)\right].$$

Hence

$$\delta\eta^{(3)}_{\mu\nu\alpha\beta}(\Omega) + \delta\eta^{(5)}_{\mu\nu\alpha\beta}(\Omega) = \delta(\Omega) \left\{ \frac{C_{\alpha\beta}}{2} \left(\delta_{\mu\beta}\delta_{\nu\alpha} - \delta_{\nu\beta}\delta_{\mu\alpha}\right) \left[\left(\delta_{\mu 1} + \delta_{\mu 2}\right)\left(\delta_{\nu 1} + \delta_{\nu 2}\right) + \delta_{\mu 3}\delta_{\nu 3}\right] f_1(\mu) \right.$$
$$+ \frac{C_{\mu\nu}}{2} \left(\delta_{\mu\beta}\delta_{\nu\alpha} - \delta_{\nu\beta}\delta_{\mu\alpha}\right) \left[\left(\delta_{\mu 1} + \delta_{\mu 2}\right)\left(\delta_{\nu 1} + \delta_{\nu 2}\right) + \delta_{\mu 3}\delta_{\nu 3}\right] f_1(\mu)$$
$$+ \frac{C_{\mu\nu}C_{\alpha\beta}}{2} \left(\delta_{\mu\alpha}\delta_{\nu\beta} - \delta_{\nu\alpha}\delta_{\mu\beta}\right) \left[\left(\delta_{\nu 1} + \delta_{\nu 2}\right) f_1(\mu) + \delta_{\nu 3} f_2(\mu)\right] \tag{C53}$$
$$\left. + \frac{C_{\mu\nu}C_{\alpha\beta}}{2} \left(\delta_{\mu\alpha}\delta_{\nu\beta} - \delta_{\nu\alpha}\delta_{\mu\beta}\right) \left[\left(\delta_{\mu 1} + \delta_{\mu 2}\right) f_1(\mu) + \delta_{\mu 3} f_2(\mu)\right] \right\}.$$

Employing the rotational symmetry of the system to set $C_{3\mu} = 0$, the correction to the static viscoelasticity tensor due to the internal degrees of freedom reads

$$\mathrm{Re}\, \delta\eta^{\mathrm{stat}}_{\mu\nu\alpha\beta}(\Omega) = \delta(\Omega) \left( \frac{C_{\mu\nu}}{2} + \frac{C_{\alpha\beta}}{2} - C_{\mu\nu}C_{\alpha\beta} \right) \left(\delta_{\mu\beta}\delta_{\nu\alpha} - \delta_{\nu\beta}\delta_{\mu\alpha}\right) \left(\delta_{\mu 1} + \delta_{\mu 2}\right) \left(\delta_{\nu 1} + \delta_{\nu 2}\right) f_1(\mu). \tag{C54}$$

It is straightforward to see that for certain strain transformation generators, i.e., $C_{\mu\nu} = 1$ or $C_{\mu\nu} = 0$ at $\mu, \nu = 1, 2$, the internal degrees of freedom do not affect the viscoelasticity tensor, $\mathrm{Re}\, \delta\eta^{\mathrm{stat}}_{\mu\nu\alpha\beta}(\Omega) = 0$.

## Appendix D: Derivation of inviscid hydrodynamic equations

In this Section, we derive the inviscid hydrodynamic equations by using the chiral kinetic theory [67–70]. We follow the standard approach discussed in Refs. [54, 55]; see also Refs. [36, 41] for a similar derivation in Weyl semimetals. The viscosity is considered in Sec. E.

The Boltzmann kinetic equation has the form

$$\frac{\partial f}{\partial t} + \left( \dot{\mathbf{r}} \cdot \frac{\partial f}{\partial \mathbf{r}} \right) + \left( \dot{\mathbf{k}} \cdot \frac{\partial f}{\partial \mathbf{k}} \right) = I_{\mathrm{col}}\left\{f\right\}, \tag{D1}$$

where $f = f(t, \mathbf{r}, \mathbf{k})$ is the distribution function and $I_{\mathrm{col}}\{f\}$ is the collision integral. In the absence of magnetic fields, the semiclassical equations of motion are [68, 69]

$$\dot{\mathbf{r}} = \mathbf{v}_{\mathbf{k}} - e\left[\mathbf{E} \times \mathbf{\Omega}\right] \quad \text{and} \quad \dot{\mathbf{k}} = -e\mathbf{E}, \tag{D2}$$

where $\mathbf{v_k} = \boldsymbol{\partial}_\mathbf{k}\epsilon_\mathbf{k}$ is velocity and $\boldsymbol{\Omega}$ is the Berry curvature which determines the anomalous velocity.

We assume the hydrodynamic ansatz for the distribution function with the fluid velocity $\mathbf{u}(t,\mathbf{r})$ [71]

$$f_\eta^{(\mathbf{u})}(t,\mathbf{r},\mathbf{k}) = \frac{1}{e^{[\varepsilon_\mathbf{k} - \eta\mu - (\mathbf{u}(t,\mathbf{r})\cdot\mathbf{k})]/T} + 1}, \tag{D3}$$

where $\eta = \pm$ corresponds to electron ($\eta = +$) and hole ($\eta = -$) states and $T$ is temperature. In the linear response regime $|\mathbf{u}| \ll v_F$, we can expand the distribution function in $|\mathbf{u}|$

$$f_\eta^{(\mathbf{u})}(t,\mathbf{r},\mathbf{k}) \approx f_\eta^{(\mathbf{0})}(t,\mathbf{r},\mathbf{k}) - (\mathbf{u}(t,\mathbf{r})\cdot\mathbf{k})\,\partial_{\varepsilon_\mathbf{k}} f_\eta^{(\mathbf{0})}(t,\mathbf{r},\mathbf{k}). \tag{D4}$$

For simplicity, we consider the momentum-relaxing scattering in the relaxation time approximation,

$$I_{\mathrm{col}}\{f_\eta\} = -\frac{f_\eta(t,\mathbf{r},\mathbf{k}) - f_\eta^{(\mathbf{0})}(t,\mathbf{r},\mathbf{k})}{\tau}. \tag{D5}$$

### 1. Linearized model of Weyl semimetals

For the linearized model of Weyl semimetals, the kinetic equation for each Weyl node with chirality $\lambda = \pm$ reads

$$\partial_t f_{\lambda,\eta} + (\mathbf{v_k} - \eta e\,[\mathbf{E}\times\boldsymbol{\Omega}_{\lambda,\eta}])\cdot\boldsymbol{\nabla} f_{\lambda,\eta} - \eta e\,(\mathbf{E}\cdot\boldsymbol{\partial}_\mathbf{p})\,f_{\lambda,\eta} = I_{\mathrm{col}}\{f_{\lambda,\eta}\}, \tag{D6}$$

where $\varepsilon_{\mathbf{k},\lambda} = v_F\,|\mathbf{k} - \lambda\mathbf{b}|$, $\mathbf{v_k} = \partial_\mathbf{k}\varepsilon_{\mathbf{k},\lambda} = v_F\frac{\mathbf{k}-\lambda\mathbf{b}}{|\mathbf{k}-\lambda\mathbf{b}|}$, and $\boldsymbol{\Omega}_{\lambda,\eta} = \frac{\lambda\eta(\mathbf{k}-\lambda\mathbf{b})}{2|\mathbf{k}-\lambda\mathbf{b}|^3}$.

To obtain hydrodynamic equations, we calculate moments of the Boltzmann equation. For the charge continuity equation, we average Eq. (D6) over momentum, sum over Weyl nodes as well as particle and hole states. We obtain

$$\partial_t\rho(t,\mathbf{r}) + \boldsymbol{\nabla}\cdot\left\{\mathbf{u}(t,\mathbf{r})\rho(t,\mathbf{r}) + \frac{e^2}{2\pi^2}\,[\mathbf{b}\times\mathbf{E}]\right\} = 0, \tag{D7}$$

where the charge density is

$$\rho(t,\mathbf{r}) = \sum_{\lambda=\pm}\sum_\eta\int\frac{d^3k}{(2\pi)^3}\eta e f_{\lambda,\eta}^{(\mathbf{0})} = -\sum_{\lambda=\pm}\sum_\eta\frac{\eta e T^3}{\pi^2 v_F^3}\mathrm{Li}_3\left(-e^{\eta\mu/T}\right) = \frac{e}{3\pi^2 v_F^3}\mu\left(\mu^2 + \pi^2T^2\right). \tag{D8}$$

The Euler equation is obtained by averaging the Boltzmann equation with the quasiparticle momentum $\mathbf{k}$,

$$\partial_t\left[w(t,\mathbf{r})\frac{\mathbf{u}(t,\mathbf{r})}{v_F^2} + \nu(t,\mathbf{r})\,(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{b})\,\mathbf{b}\right] + \boldsymbol{\nabla}P(t,\mathbf{r}) + \rho(t,\mathbf{r})\,\mathbf{E} = -\frac{1}{\tau}\left[w(t,\mathbf{r})\frac{\mathbf{u}(t,\mathbf{r})}{v_F^2} + \nu(t,\mathbf{r})\,(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{b})\,\mathbf{b}\right], \tag{D9}$$

where $w(t,\mathbf{r}) = \epsilon(t,\mathbf{r}) + P(t,\mathbf{r}) = 4\epsilon(t,\mathbf{r})/3$, the energy density is

$$\epsilon = \sum_{\lambda=\pm}\sum_\eta\int\frac{d^3k}{(2\pi)^3}\varepsilon_{\mathbf{k},\lambda}f_{\lambda,\eta}^{\mathbf{0}} = -\sum_{\lambda=\pm}\sum_\eta\frac{\Gamma(4)T^4}{2\pi^2 v_F^3}\mathrm{Li}_4\left(-e^{\eta\mu/T}\right) = \frac{1}{4\pi^2 v_F^3}\left(\mu^4 + 2\pi^2\mu^2T^2 + \frac{7}{15}\pi^4T^4\right), \tag{D10}$$

and we introduced the shorthand notation

$$\nu = \frac{1}{\pi^2 v_F^3}\left(\mu^2 + \frac{1}{3}\pi^2T^2\right). \tag{D11}$$

Finally, for the energy continuity equation, we average the Boltzmann equation with the quasiparticle energy $\varepsilon_{\mathbf{k},\lambda}$,

$$\partial_t\epsilon(t,\mathbf{r}) + \frac{4}{3}\boldsymbol{\nabla}\cdot[\mathbf{u}(t,\mathbf{r})\epsilon(t,\mathbf{r})] = 0. \tag{D12}$$

### 2. Two-band model of Weyl semimetals

For the two-band model of Weyl semimetals, see Eq. (26), the Berry curvature is given by

$$\Omega_{i;\eta} = \eta\epsilon_{imn}\frac{\left(\mathbf{d}\left(\mathbf{k}\right)\cdot\left[\partial_{k_m}\mathbf{d}\left(\mathbf{k}\right)\times\partial_{k_n}\mathbf{d}\left(\mathbf{k}\right)\right]\right)}{4\left|\mathbf{d}\left(\mathbf{k}\right)\right|^3}. \tag{D13}$$

To simplify our results, in what follows, we consider the case of vanishing temperature $T = 0$. Then the charge continuity equation acquires the following form:

$$\partial_t\rho(t,\mathbf{r}) + \boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})\rho(t,\mathbf{r}) + \mathbf{j}^{\mathrm{AHE}}\right] = 0, \tag{D14}$$

where

$$\begin{aligned}\rho = \sum_\eta\int\frac{d^3k}{(2\pi)^3}\eta e f_\eta^{(\mathbf{0})} = \frac{e\,\mathrm{sign}(\mu)}{15\pi^2v_F^2}\Bigg[\left(3\mu^2+\gamma m|\mu|-2\gamma^2m^2\right)\sqrt{\frac{\gamma m+|\mu|}{\gamma}}\\ -\Theta\left(\gamma m-|\mu|\right)\left(3\mu^2-\gamma m|\mu|-2\gamma^2m^2\right)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg],\end{aligned} \tag{D15}$$

$$\begin{aligned}\mathbf{j}^{\mathrm{AHE}} = -e^2\sum_\eta\int\frac{d^3k}{(2\pi)^3}\left[\mathbf{E}\times\boldsymbol{\Omega}_\eta\right]f_\eta^{(\mathbf{0})} = \frac{e^2}{2\pi^2}\left(\mathbf{e}_xE_y-\mathbf{e}_yE_x\right)\Bigg[\left(\frac{\Lambda_z}{2}-\sqrt{m}\right)\\ +\mathrm{sign}\left(\mu\right)\left(\sqrt{m}-\frac{\gamma m+|\mu|}{3|\mu|}\sqrt{\frac{\gamma m+|\mu|}{\gamma}}+\Theta\left(\gamma m-|\mu|\right)\frac{\gamma m-|\mu|}{3|\mu|}\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\right)\Bigg],\end{aligned} \tag{D16}$$

and we used Faraday's law to construct the full derivative. The Euler equation reads

$$\begin{aligned}\partial_t\left[w(t,\mathbf{r})\frac{\mathbf{u}(t,\mathbf{r})}{v_F^2}+A(t,\mathbf{r})\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{e}_z\right)\mathbf{e}_z\right]+\boldsymbol{\nabla}P\left(t,\mathbf{r}\right)+\rho\left(t,\mathbf{r}\right)\mathbf{E}\\ =-\frac{1}{\tau}\left[w(t,\mathbf{r})\frac{\mathbf{u}(t,\mathbf{r})}{v_F^2}+A(t,\mathbf{r})\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{e}_z\right)\mathbf{e}_z\right],\end{aligned} \tag{D17}$$

where $w = \epsilon + P$ and

$$\begin{aligned}\epsilon = \sum_\eta\int\frac{d^3k}{(2\pi)^3}\varepsilon_{\mathbf{k}}f_\eta^{(\mathbf{0})} = \frac{1}{105\pi^2v_F^2}\Bigg[-16\gamma^3m^{7/2}+\left(8\gamma^3m^3-4\gamma^2m^2|\mu|+3\gamma m\mu^2+15|\mu|^3\right)\sqrt{\frac{\gamma m+|\mu|}{\gamma}}\\ +\Theta(\gamma m-|\mu|)\left(8\gamma^3m^3+4\gamma^2m^2|\mu|+3\gamma m\mu^2-15|\mu|^3\right)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg],\end{aligned} \tag{D18}$$

$$\begin{aligned}P = \sum_\eta\int\frac{d^3k}{(2\pi)^3}\frac{d_i\partial_i(d_i)k_i}{|\mathbf{d}(\mathbf{k})|}f_\eta^{(\mathbf{0})} = \frac{2}{105\pi^2v_F^2}\Bigg[8\gamma^3m^{7/2}-(\gamma m+|\mu|)^2(4\gamma m-3|\mu|)\sqrt{\frac{\gamma m+|\mu|}{\gamma}}\\ -\Theta(\gamma m-|\mu|)(\gamma m-|\mu|)^2(4\gamma m+3|\mu|)\sqrt{\frac{\gamma m-|\mu|}{\gamma}}\Bigg],\end{aligned} \tag{D19}$$

$$A = \frac{\gamma^2m|\mu|}{30\pi^2v_F^2}\left[\left(\frac{\gamma m+|\mu|}{\gamma}\right)^{3/2}\left(\frac{5}{\gamma^2m}+4-6\frac{|\mu|}{\gamma m}\right)-\Theta\left(\gamma m-|\mu|\right)\left(\frac{\gamma m-|\mu|}{\gamma}\right)^{3/2}\left(\frac{5}{\gamma^2m}+4+6\frac{|\mu|}{\gamma m}\right)\right]. \tag{D20}$$

Finally, the energy continuity equation reads

$$\partial_t\epsilon(t,\mathbf{r}) + \boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})w(t,\mathbf{r})\right] = 0. \tag{D21}$$

Equations (D14), (D17), and (D21) correspond to Eqs. (45), (46) and (47) in the main text, respectively.

### Appendix E: Derivation of viscosity in the kinetic approach

To determine the viscosity of electron fluid, one has to consider the first-order correction to the distribution function $f_\eta(t,\mathbf{r},\mathbf{k})$ due to the electron-electron scattering. By following the standard approach [54, 55] and using the relaxation time approximation, we obtain the following first-order in $\tau_{ee} \ll \tau$ correction to the distribution function:

$$\begin{aligned}\delta f_\eta(t,\mathbf{r},\mathbf{k}) = -\tau_{ee}\Big\{\partial_t f_\eta^{(\mathbf{u})} + \left(\mathbf{v_k} - \eta e\left[\mathbf{E}\times\mathbf{\Omega}_\eta\right]\cdot\boldsymbol{\nabla} f_\eta^{(\mathbf{0})}\right) - \eta e\left(\mathbf{E}\cdot\boldsymbol{\partial}_\mathbf{k} f_\eta^{(\mathbf{0})}\right) - \\ - \left(\mathbf{v_k} - \eta e\left[\mathbf{E}\times\mathbf{\Omega}_\eta\right]\cdot\boldsymbol{\nabla}\left[\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{k}\right)\partial_{\varepsilon_\mathbf{k}} f_\eta^{(\mathbf{0})}\right]\right) + \left(\eta e\mathbf{E}\cdot\boldsymbol{\partial}_\mathbf{k}\left[\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{k}\right)\partial_{\varepsilon_\mathbf{k}} f_\eta^{(\mathbf{0})}\right]\right)\Big\}.\end{aligned} \tag{E1}$$

Viscosity enters the Euler equation, see Eq. (D9) or (D17), as the following term on its left-hand side:

$$-\eta_{ijkl}\nabla_j\nabla_l u_k(t,\mathbf{r}). \tag{E2}$$

To obtain such a term, we focus only on terms with two spatial derivatives of $\mathbf{u}(t,\mathbf{r})$. Further, we rewrite the time derivative in Eq. (E1) as

$$\partial_t f_\eta^{(\mathbf{u})} = -\left(\mathbf{k}\cdot\partial_t\mathbf{u}(t,\mathbf{r})\right)\partial_{\varepsilon_\mathbf{k}} f_\eta^{(\mathbf{0})} + \left[\partial_t\epsilon_\eta^{\mathbf{0}}(t,\mathbf{r})\right]\partial_{\varepsilon_\eta^{(\mathbf{0})}} f_\eta \sim -\boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})w(t,\mathbf{r})\right]\partial_\epsilon f_\eta. \tag{E3}$$

Then, collecting the relevant terms in Eq. (E1), we obtain the following part of the distribution function responsible for the electron viscosity:

$$\delta\tilde{f}_\eta(t,\mathbf{r},\mathbf{k}) = \tau_{ee}\left\{\boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})w(t,\mathbf{r})\right]\partial_\epsilon f_\eta^{(\mathbf{0})} + \left(\mathbf{v_k}\cdot\boldsymbol{\nabla}\left[\mathbf{u}(t,\mathbf{r})\cdot\mathbf{k}\right]\right)\partial_{\varepsilon_\mathbf{k}} f_\eta^{(\mathbf{0})}\right\}. \tag{E4}$$

The subsequent derivations of the hydrodynamic equations follow the same steps as in Sec. D. Viscosity-related terms originate from the second term in the kinetic equation (D6). After performing the averaging over momenta, we have the following integrals:

$$I_{(\mathrm{visc})}^{(1)} = \tau_{ee}\sum_\eta\int\frac{d^3k}{(2\pi)^3}\mathbf{k}\left(\mathbf{v_k}\cdot\boldsymbol{\nabla}\right)\left\{\boldsymbol{\nabla}\cdot\left[\boldsymbol{u}(t,\mathbf{r})w(t,\mathbf{r})\right]\partial_\epsilon f_\eta^{(\mathbf{0})}\right\}, \tag{E5}$$

$$I_{(\mathrm{visc})}^{(2)} = \tau_{ee}\sum_\eta\int\frac{d^3k}{(2\pi)^3}\mathbf{k}\left(\mathbf{v_k}\cdot\boldsymbol{\nabla}\right)\left[\mathbf{v_k}\cdot\boldsymbol{\nabla}\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{k}\right)\partial_{\varepsilon_\mathbf{k}} f_\eta^{(\mathbf{0})}\right]. \tag{E6}$$

In the following two Sections, we calculate these integrals in linearized and two-band models of Weyl semimetals.

#### 1. Linearized model of Weyl semimetals

Let us start with the linearized model (13). The integrals in Eqs. (E5) and (E6) are

$$\begin{aligned}I_{(\mathrm{visc})} = I_{(\mathrm{visc})}^{(1)} + I_{(\mathrm{visc})}^{(2)} = \tau_{ee}\frac{4}{9}\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})\epsilon(t,\mathbf{r})\right]\right) - \tau_{ee}\frac{4}{15}\Delta\left[\mathbf{u}(t,\mathbf{r})\epsilon(t,\mathbf{r})\right] \\ - \tau_{ee}\frac{8}{15}\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\left[\mathbf{u}(t,\mathbf{r})\epsilon(t,\mathbf{r})\right]\right) - \tau_{ee}\mathbf{b}\frac{1}{3}\Delta\left[\left(\mathbf{u}(t,\mathbf{r})\cdot\mathbf{b}\right)v_F^2\nu(t,\mathbf{r})\right].\end{aligned} \tag{E7}$$

Therefore, linearizing in deviations from local equilibrium $\epsilon(t,\mathbf{r}) = \epsilon$ and $\nu(t,\mathbf{r}) = \nu$, we obtain

$$\eta_{ijkl} = \frac{4\tau_{ee}}{15}\epsilon\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \frac{2}{3}\delta_{ij}\delta_{kl}\right) + \frac{\tau_{ee}v_F^2\nu}{3}b_i b_k\delta_{jl}. \tag{E8}$$

### 2. Two-band model of Weyl semimetals

Let us consider the two-band model of Weyl semimetals defined in Eq. (26). Leaving only terms with the second-order derivatives of $\mathbf{u}$, integrals (E5) and (E6) read

$$I^{(1)}_{(\mathrm{visc})} = \tau_{ee} w \frac{\partial^2 u_i}{\partial x_i \partial x_j} \sum_\eta \frac{\partial}{\partial \epsilon} \int \frac{d^3k}{(2\pi)^3} \mathbf{k} v_j f^{(\mathbf{0})}_\eta = \tau_{ee} w \boldsymbol{\nabla} \left(\boldsymbol{\nabla} \cdot \mathbf{u}\right) \partial_\epsilon P, \tag{E9}$$

$$\begin{aligned} I^{(2)}_{(\mathrm{visc})} &= -\mathbf{e}_i \frac{\partial^2 u_k}{\partial x_j \partial x_l} \tau_{ee} \sum_\eta \int \frac{d^3k}{(2\pi)^3} k_i \, v_j k_k v_l \delta\left(|\mu| - |\mathbf{d}(\mathbf{k}|\right) \\ &= -\mathbf{e}_i \frac{\partial^2 u_k}{\partial x_j \partial x_l} \tau_{ee} \sum_\eta \int \frac{d^3k}{(2\pi)^3} k_i \, \frac{d_j \partial_j (d_j)}{|\mathbf{d}(\mathbf{k})|} k_k \frac{d_l \partial_l (d_l)}{|\mathbf{d}(\mathbf{k})|} \delta\left(|\mu| - |\mathbf{d}(\mathbf{k})|\right). \end{aligned} \tag{E10}$$

Evaluating the integral in Eq. (E10), we obtain

$$\begin{aligned} I^{(2)}_{(\mathrm{visc})} = -\mathbf{e}_i \frac{\partial^2 u_k}{\partial x_j \partial x_l} \frac{\tau_{ee}}{2\pi^2 v_F^2 |\mu|} \Bigg\{ & \left[\delta_{il}\delta_{jk} + \delta_{ij}\delta_{kl} + \left(\delta_{i1}+\delta_{i2}\right)\left(\delta_{j1}+\delta_{j2}\right)\delta_{ik}\delta_{jl}\right] \frac{4}{315} \Bigg[ \left(4\gamma^2 m^2 - 10\gamma m|\mu| + 7\mu^2\right) \\ & \times (\gamma m + |\mu|)^2 \sqrt{\frac{\gamma m + |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) \left(4\gamma^2 m^2 + 10\gamma m|\mu| + 7\mu^2\right) (\gamma m - |\mu|)^2 \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \Bigg] \\ & + \delta_{i3}\delta_{j3}\delta_{ik}\delta_{jl} \frac{4\mu^2}{15} \left[ (\gamma m + |\mu|)^2 \sqrt{\frac{\gamma m + |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) (\gamma m - |\mu|)^2 \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \right] \\ & - \delta_{i3}\left(\delta_{j1}+\delta_{j2}\right)\delta_{ik}\delta_{jl} \frac{2v_F^2}{105\gamma} \Bigg[ (\gamma m + |\mu|)^2 (2\gamma m - 5|\mu|) \sqrt{\frac{\gamma m + |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) (\gamma m - |\mu|)^2 \\ & \times (2\gamma m + 5|\mu|) \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \Bigg] - \left(\delta_{i1}+\delta_{i2}\right)\delta_{j3}\delta_{ik}\delta_{jl} \frac{8\gamma}{3465 v_F^2} \Bigg[ (\gamma m + |\mu|)^2 \sqrt{\frac{\gamma m + |\mu|}{\gamma}} \\ & \times \left(32\gamma^3 m^3 - 80\gamma^2 m^2 |\mu| + 74\gamma m \mu^2 - 45|\mu|^3\right) - \Theta\left(\gamma m - |\mu|\right) (\gamma m - |\mu|)^2 \\ & \times \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \left(32\gamma^3 m^3 + 80\gamma^2 m^2 |\mu| + 74\gamma m \mu^2 + 45|\mu|^3\right) \Bigg] \Bigg\}. \end{aligned} \tag{E11}$$

The final result for the viscosity tensor in the two-band model is given by

$$\begin{aligned} \eta_{ijkl} = \frac{\tau_{ee}}{\pi^2 v_F^2 |\mu|} \Bigg\{ & \left[\delta_{il}\delta_{jk} + \delta_{ij}\delta_{kl} + \left(\delta_{i1}+\delta_{i2}\right)\left(\delta_{j1}+\delta_{j2}\right)\delta_{ik}\delta_{jl}\right] \frac{2}{315} \Bigg[ \left(4\gamma^2 m^2 - 10\gamma m|\mu| + 7\mu^2\right) \\ & \times (\gamma m + |\mu|)^2 \sqrt{\frac{\gamma m + |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) \left(4\gamma^2 m^2 + 10\gamma m|\mu| + 7\mu^2\right) (\gamma m - |\mu|)^2 \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \Bigg] \\ & + \delta_{i3}\delta_{j3}\delta_{ik}\delta_{jl} \frac{2\mu^2}{15} \left[ (\gamma m + |\mu|)^2 \sqrt{\frac{\gamma m + |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) (\gamma m - |\mu|)^2 \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \right] \\ & - \delta_{i3}\left(\delta_{j1}+\delta_{j2}\right)\delta_{ik}\delta_{jl} \frac{v_F^2}{105\gamma} \Bigg[ (\gamma m + |\mu|)^2 (2\gamma m - 5|\mu|) \sqrt{\frac{\gamma m + |\mu|}{\gamma}} - \Theta\left(\gamma m - |\mu|\right) (\gamma m - |\mu|)^2 \\ & \times (2\gamma m + 5|\mu|) \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \Bigg] - \left(\delta_{i1}+\delta_{i2}\right)\delta_{j3}\delta_{ik}\delta_{jl} \frac{4\gamma}{3465 v_F^2} \Bigg[ (\gamma m + |\mu|)^2 \sqrt{\frac{\gamma m + |\mu|}{\gamma}} \\ & \times \left(32\gamma^3 m^3 - 80\gamma^2 m^2 |\mu| + 74\gamma m \mu^2 - 45|\mu|^3\right) - \Theta\left(\gamma m - |\mu|\right) (\gamma m - |\mu|)^2 \\ & \times \sqrt{\frac{\gamma m - |\mu|}{\gamma}} \left(32\gamma^3 m^3 + 80\gamma^2 m^2 |\mu| + 74\gamma m \mu^2 + 45|\mu|^3\right) \Bigg] \Bigg\} - \tau_{ee}\delta_{ij}\delta_{kl} P \left(1 + \partial_\epsilon P\right). \end{aligned} \tag{E12}$$

[1] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, and K. C. Fong, Observation of the Dirac fluid and the breakdown of the Wiedemann-Franz law in graphene, Science **351**, 1058 (2016), arXiv:1509.04713.

[2] F. Ghahari, H.-Y. Xie, T. Taniguchi, K. Watanabe, M. S. Foster, and P. Kim, Enhanced thermoelectric power in graphene: Violation of the Mott relation by inelastic scattering, Phys. Rev. Lett. **116**, 136802 (2016), arXiv:1601.05859.

[3] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, K. Watanabe, T. Taniguchi, I. V. Grigorieva, L. S. Levitov, M. Polini, and A. K. Geim, Superballistic flow of viscous electron fluid through graphene constrictions, Nat. Phys. **13**, 1182 (2017), arXiv:1703.06672.

[4] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. Ben Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, M. Polini, A. K. Geim, and D. A. Bandurin, Measuring Hall viscosity of graphene's electron fluid, Science **364**, 162 (2019), arXiv:1806.01606.

[5] D. A. Bandurin, A. V. Shytov, L. S. Levitov, R. K. Kumar, A. I. Berdyugin, M. Ben Shalom, I. V. Grigorieva, A. K. Geim, and G. Falkovich, Fluidity onset in graphene, Nat. Commun. **9**, 4533 (2018), arXiv:1806.03231.

[6] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, U. Vool, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, M. M. Fogler, P. Kim, A. Yacoby, and R. L. Walsworth, Imaging viscous flow of the Dirac fluid in graphene, Nature **583**, 537 (2020), arXiv:1905.10791.

[7] J. A. Sulpizio, L. Ella, A. Rozen, J. Birkbeck, D. J. Perello, D. Dutta, M. Ben-Shalom, T. Taniguchi, K. Watanabe, T. Holder, R. Queiroz, A. Principi, A. Stern, T. Scaffidi, A. K. Geim, and S. Ilani, Visualizing Poiseuille flow of hydrodynamic electrons, Nature **576**, 75 (2019), arXiv:1905.11662.

[8] S. Samaddar, J. Strasdas, K. Janßen, S. Just, T. Johnsen, Z. Wang, B. Uzlu, S. Li, D. Neumaier, M. Liebmann, and M. Morgenstern, Evidence for Local Spots of Viscous Electron Flow in Graphene at Moderate Mobility, Nano Lett. **21**, 9365 (2021), arXiv:2103.11466.

[9] C. Kumar, J. Birkbeck, J. A. Sulpizio, D. J. Perello, T. Taniguchi, K. Watanabe, O. Reuven, T. Scaffidi, A. Stern, A. K. Geim, and S. Ilani, Imaging Hydrodynamic Electrons Flowing Without Landauer-Sharvin Resistance (2021), arXiv:2111.06412.

[10] L. Molenkamp and M. de Jong, Observation of Knudsen and Gurzhi transport regimes in a two-dimensional wire, Solid. State. Electron. **37**, 551 (1994).

[11] M. J. De Jong and L. W. Molenkamp, Hydrodynamic electron flow in high-mobility wires, Phys. Rev. B **51**, 13389 (1995), arXiv:9411067 [cond-mat].

[12] R. Gurzhi, Minimum of resistance in impurity-free conductors, JETP **17**, 521 (1963).

[13] R. N. Gurzhi, Hydrodynamic effects in solids at low temperature, Sov. Phys. Uspekhi **11**, 255 (1968).

[14] A. Lucas and K. C. Fong, Hydrodynamics of electrons in graphene, J. Phys. Condens. Matter **30**, 053001 (2018), arXiv:1710.08425.

[15] B. N. Narozhny, Electronic hydrodynamics in graphene, Ann. Phys. (N. Y). **411**, 167979 (2019), arXiv:1905.09686.

[16] B. N. Narozhny, Hydrodynamic approach to two-dimensional electron systems, La Riv. del Nuovo Cim. **45**, 661 (2022), arXiv:2207.10004.

[17] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Butterworth-Heinemann, Oxford, 2013).

[18] J. M. Link, B. N. Narozhny, E. I. Kiselev, and J. Schmalian, Out-of-Bounds Hydrodynamics in Anisotropic Dirac Fluids, Phys. Rev. Lett. **120**, 196801 (2018), arXiv:1708.02759v5.

[19] P. Rao and B. Bradlyn, Hall Viscosity in Quantum Systems with Discrete Symmetry: Point Group and Lattice Anisotropy, Phys. Rev. X **10**, 021005 (2020), arXiv:1910.10727.

[20] B. Offertaler and B. Bradlyn, Viscoelastic response of quantum Hall fluids in a tilted field, Phys. Rev. B **99**, 035427 (2019), arXiv:1811.08443.

[21] C. Q. Cook and A. Lucas, Electron hydrodynamics with a polygonal Fermi surface, Phys. Rev. B **99**, 235148 (2019), arXiv:1903.05652.

[22] F. Peña-Benitez, K. Saha, and P. Surówka, Berry curvature and Hall viscosities in an anisotropic Dirac semimetal, Phys. Rev. B **99**, 045141 (2019), arXiv:1805.09827.

[23] G. Varnavides, A. S. Jermyn, P. Anikeeva, C. Felser, and P. Narang, Electron hydrodynamics in anisotropic materials, Nat. Commun. **11**, 4710 (2020), arXiv:2002.08976.

[24] I. Robredo, P. Rao, F. de Juan, A. Bergara, J. L. Mañes, A. Cortijo, M. G. Vergniory, and B. Bradlyn, Cubic Hall viscosity in three-dimensional topological semimetals, Phys. Rev. Res. **3**, L032068 (2021), arXiv:2102.02226.

[25] X. Huang and A. Lucas, Hydrodynamic effective field theories with discrete rotational symmetry, J. High Energy Phys. **2022** (3), 82, arXiv:2201.03565.

[26] J. E. Avron, R. Seiler, and P. G. Zograf, Viscosity of Quantum Hall Fluids, Phys. Rev. Lett. **75**, 697 (1995), arXiv:9502011 [cond-mat].

[27] L. Sun and S. Wan, Chiral viscoelastic response in Weyl semimetals, EPL (Europhysics Lett. **108**, 37007 (2014).

[28] H. Shapourian, T. L. Hughes, and S. Ryu, Viscoelastic response of topological tight-binding models in two and three dimensions, Phys. Rev. B **92**, 165131 (2015), arXiv:1505.03868.

[29] A. Cortijo, Y. Ferreirós, K. Landsteiner, and M. A. H. Vozmediano, Elastic Gauge Fields in Weyl Semimetals, Phys. Rev. Lett. **115**, 177202 (2015), arXiv:1603.02674.

[30] J. M. Link, D. E. Sheehy, B. N. Narozhny, and J. Schmalian, Elastic response of the electron fluid in intrinsic graphene: The collisionless regime, Phys. Rev. B **98**, 195103 (2018), arXiv:1808.04658.

[31] C. Copetti and K. Landsteiner, Anomalous Hall viscosity at the Weyl-semimetal–insulator transition, Phys. Rev. B **99**, 195146 (2019), arXiv:1901.11403.

[32] M. Moore, P. Surówka, V. Juričić, and B. Roy, Shear viscosity as a probe of nodal topology, Phys. Rev. B **101**, 161111 (2020), arXiv:1912.07611.

[33] N. Read, Non-Abelian adiabatic statistics and Hall vis-

cosity in quantum Hall states and px + ipy paired superfluids, Phys. Rev. B - Condens. Matter Mater. Phys. **79**, 1 (2009), arXiv:0805.2507.
[34] T. L. Hughes, R. G. Leigh, and E. Fradkin, Torsional Response and Dissipationless Viscosity in Topological Insulators, Phys. Rev. Lett. **107**, 075502 (2011), arXiv:1101.3541.
[35] C. Hoyos and D. T. Son, Hall Viscosity and Electromagnetic Response, Phys. Rev. Lett. **108**, 066805 (2012), arXiv:1109.2651.
[36] E. V. Gorbar, V. A. Miransky, I. A. Shovkovy, and P. O. Sukhachov, Consistent hydrodynamic theory of chiral electrons in Weyl semimetals, Phys. Rev. B **97**, 121105(R) (2018), arXiv:1712.01289.
[37] E. V. Gorbar, V. A. Miransky, I. A. Shovkovy, and P. O. Sukhachov, Hydrodynamic electron flow in a Weyl semimetal slab: Role of Chern-Simons terms, Phys. Rev. B **97**, 205119 (2018), arXiv:1802.07265.
[38] E. V. Gorbar, V. A. Miransky, I. A. Shovkovy, and P. O. Sukhachov, Nonlocal transport in Weyl semimetals in the hydrodynamic regime, Phys. Rev. B **98**, 035121 (2018), arXiv:1804.01550.
[39] P. O. Sukhachov, E. V. Gorbar, I. A. Shovkovy, and V. A. Miransky, Collective excitations in Weyl semimetals in the hydrodynamic regime, J. Phys. Condens. Matter **30**, 275601 (2018), arXiv:1802.10110.
[40] G. Bednik, K. S. Tikhonov, and S. V. Syzranov, Magnetotransport and internodal tunnelling in Weyl semimetals, Phys. Rev. Res. **2**, 023124 (2020), arXiv:1912.02284.
[41] P. O. Sukhachov and B. Trauzettel, Anomalous Gurzhi effect, Phys. Rev. B **105**, 085141 (2022), arXiv:2112.00781.
[42] S. Zhu, G. Bednik, and S. Syzranov, Weyl hydrodynamics in a strong magnetic field, Phys. Rev. B **105**, 125132 (2022), arXiv:2108.09313.
[43] J. Bernabeu and A. Cortijo, Bounds on phonon-mediated hydrodynamic transport in a type-I Weyl semimetal, Phys. Rev. B **107**, 235141 (2023), arXiv:2211.13245.
[44] J. Gooth, F. Menges, N. Kumar, V. Sü$\beta$, C. Shekhar, Y. Sun, U. Drechsler, R. Zierold, C. Felser, and B. Gotsmann, Thermal and electrical signatures of a hydrodynamic electron fluid in tungsten diphosphide, Nat. Commun. **9**, 4093 (2018), arXiv:1706.05925.
[45] B. Bradlyn, M. Goldstein, and N. Read, Kubo formulas for viscosity: Hall viscosity, Ward identities, and the relation with conductivity, Phys. Rev. B **86**, 245309 (2012), arXiv:1207.7021.
[46] T. Holder, R. Queiroz, and A. Stern, Unified Description of the Classical Hall Viscosity, Phys. Rev. Lett. **123**, 106801 (2019), arXiv:1903.05541.
[47] As is discussed in Ref. [45], it is crucial to subtract the inverse isentropic compressibility in order to obtain finite bulk viscosity.
[48] F. Belinfante, On the current and the density of the electric charge, the energy, the linear momentum and the angular momentum of arbitrary fields, Physica **7**, 449 (1940).
[49] L. Rosenfeld, On the energy-momentum tensor, Mémoires Acad. Roy. Belgique. **18**, 1 (1940).
[50] Unlike isotropic Dirac spectrum, see Eq. (6), rotational invariance is commonly broken in crystals. Therefore, one should consider the modification of the Belinfante symmetrization [19] or the Belinfante-like stress tensor, which takes into account reduced symmetries of a system.
[51] C. J. Tabert and J. P. Carbotte, Optical conductivity of Weyl semimetals and signatures of the gapped semimetal phase transition, Phys. Rev. B **93**, 085442 (2016), arXiv:1603.03722.
[52] In lattice models, the flux of the Berry curvature through the whole Brillouin zone should vanish. This is not the case for the two-band model given in Eq. (26).
[53] A. Cortijo, Y. Ferreirós, K. Landsteiner, and M. A. H. Vozmediano, Visco elasticity in 2D materials, 2D Mater. **3**, 011002 (2016), arXiv:1506.05136.
[54] K. Huang, *Statistical Mechanics* (Wiley, Hoboken, 1987).
[55] E. M. Lifshits and L. P. Pitaevskii, *Physical Kinetics* (Butterworth-Heinemann, Oxford, 2012).
[56] For Galilean-invariant systems, one can derive microscopic expressions for the viscosity tensor by equating the coefficients in the current response to a vector potential to the current-current response functions [72]. The connection to a more rigorous approach based on the stress-stress correlation function follows from the fact that the current density is proportional to the momentum density, therefore its time derivative is proportional to the divergence of the stress tensor.
[57] A. Principi, G. Vignale, M. Carrega, and M. Polini, Bulk and shear viscosities of the two-dimensional electron liquid in a doped graphene sheet, Phys. Rev. B **93**, 125410 (2016), arXiv:1506.06030.
[58] E. V. Gorbar, V. A. Miransky, I. A. Shovkovy, and P. O. Sukhachov, Origin of Bardeen-Zumino current in lattice models of Weyl semimetals, Phys. Rev. B **96**, 085130 (2017), arXiv:1706.02705.
[59] R. Moessner, P. Surówka, and P. Witkowski, Pulsating flow and boundary layers in viscous electronic hydrodynamics, Phys. Rev. B **97**, 161112 (2018), arXiv:1710.00354.
[60] P. O. Sukhachov and E. V. Gorbar, Stray magnetic field and stability of time-dependent viscous electron flow, Phys. Rev. B **104**, 195111 (2021), arXiv:2107.09740.
[61] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie, Evidence for hydrodynamic electron flow in $PdCoO_2$, Science **351**, 1061 (2016), arXiv:1509.05691.
[62] M. R. van Delft, Y. Wang, C. Putzke, J. Oswald, G. Varnavides, C. A. C. Garcia, C. Guo, H. Schmid, V. Süss, H. Borrmann, J. Diaz, Y. Sun, C. Felser, B. Gotsmann, P. Narang, and P. J. W. Moll, Sondheimer oscillations as a probe of non-ohmic flow in $WP_2$ crystals, Nat. Commun. **12**, 4799 (2021), arXiv:2012.08522.
[63] B. Bradlyn, J. Cano, Z. Wang, M. G. Vergniory, C. Felser, R. J. Cava, and B. A. Bernevig, Beyond Dirac and Weyl fermions: Unconventional quasiparticles in conventional crystals, Science **353**, aaf5037 (2016), arXiv:1603.03093.
[64] L. Šmejkal, J. Sinova, and T. Jungwirth, Emerging Research Landscape of Altermagnetism, Phys. Rev. X **12**, 040501 (2022), arXiv:2204.10844.
[65] For isotropic systems, it is equivalent to the Belinfante-Rosenfeld stress-energy tensor [48, 49].
[66] Compared to Ref. [29], we changed sign at $d_z$.
[67] D. Xiao, M. M.-C. C. Chang, and Q. Niu, Berry phase effects on electronic properties, Rev. Mod. Phys. **82**, 1959 (2010), arXiv:0907.2021.
[68] D. T. Son and N. Yamamoto, Kinetic theory with Berry curvature from quantum field theories, Phys. Rev. D **87**, 085016 (2013), arXiv:1210.8158.

[69] M. A. Stephanov and Y. Yin, Chiral kinetic theory, Phys. Rev. Lett. **109**, 162001 (2012), arXiv:1207.0747.

[70] D. T. Son and B. Z. Spivak, Chiral anomaly and classical negative magnetoresistance of Weyl metals, Phys. Rev. B **88**, 104412 (2013), arXiv:1206.1627.

[71] The distribution function (D3) satisfies the electron-electron collision integral. This form is also known as the Callaway ansatz [11, 73].

[72] S. Conti and G. Vignale, Elasticity of an electron liquid, Phys. Rev. B **60**, 7966 (1999), arXiv:9811214v2 [cond–mat].

[73] J. Callaway, Model for Lattice Thermal Conductivity at Low Temperatures, Phys. Rev. **113**, 1046 (1959).

# Fast neutrino flavor conversions in a supernova: emergence, evolution, and effects

Zewei Xiong,[1,*] Meng-Ru Wu,[2,3,4] Manu George,[2] Chun-Yu Lin,[5] Noshad Khosravi Largani,[6] Tobias Fischer,[6] and Gabriel Martínez-Pinedo[1,7]

[1] *GSI Helmholtzzentrum für Schwerionenforschung, Planckstraße 1, 64291 Darmstadt, Germany*
[2] *Institute of Physics, Academia Sinica, Taipei 11529, Taiwan*
[3] *Institute of Astronomy and Astrophysics, Academia Sinica, Taipei 10617, Taiwan*
[4] *Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan*
[5] *National Center for High-performance Computing, Hsinchu 30076, Taiwan*
[6] *Institute of Theoretical Physics, University of Wroclaw, Pl. M. Borna 9, 50-204 Wroclaw, Poland*
[7] *Institut für Kernphysik (Theoriezentrum), Fachbereich Physik, Technische Universität Darmstadt, Schlossgartenstraße 2, 64289 Darmstadt, Germany*

(Dated: March 1, 2024)

Fast flavor conversions (FFCs) of neutrinos, which can occur in core-collapse supernovae (CCSNe), are multiangle effects. They depend on the angular distribution of the neutrino's electron lepton number (ELN). In this work, we present a comprehensive study of the FFCs by solving the multi-energy and multiangle quantum kinetic equations with an extended set of collisional weak processes based on a static and spherically symmetric CCSN matter background profile. We investigate the emergence and evolution of FFCs in models featuring different ELN angular distributions, considering scenarios with two and three neutrino flavors. The spectrogram method is utilized to illustrate the small-scale spatial structure, and we show that this structure of neutrino flavor coherence and number densities in the nonlinear regime is qualitatively consistent with the dispersion relation analysis. On the coarse-grained level, we find that different asymptotic states can be achieved following the FFCs depending on the locations and shapes of the ELN distributions, despite sharing a common feature of the elimination of the ELN angular crossing. While equilibration among different neutrino flavors may be achieved immediately after the prompt FFCs, it is not a general outcome of the asymptotic state, as subsequent feedback effects from collisional neutrino-matter interactions come into play, particularly for cases where FFCs occur inside the neutrinosphere. The impacts of FFCs and the feedback effect on the net neutrino heating rates, the equilibrium electron fraction of CCSN matter, and the free-streaming neutrino energy spectra are quantitatively assessed. Other aspects including the impact of the vacuum term and the coexistence with other type of flavor instabilities are also discussed.

arXiv:2402.19252v1 [astro-ph.HE] 29 Feb 2024

## I. INTRODUCTION

Core-collapse supernovae (CCSNe) are cataclysmic events when massive stars reach the end of their lives. During the evolution of a CCSN, a nascent proto-neutron star (PNS) forms at the center as a prolific source of neutrinos. These neutrinos play pivotal roles on the CCSN dynamics and the evolution of chemical composition. They interact with the medium through both charged- and neutral-current weak interactions in the proximity of the PNS and deposit energy, facilitating the shock revival via reheating of material in the post-shocked layer and leading to the eventual mass ejection. Particularly, the charged-current interactions determine the proton-to-baryon ratio, denoted by the electron fraction $Y_e$, which is a crucial quantity for the nucleosynthesis results of CCSN explosions. A better knowledge of the flux intensities and flavor content of neutrinos is necessary to robustly model the inner dynamics of CCSNe, to predict the elemental compositions in CCSN ejecta, and to determine the neutrino signals for the detection of the next galactic supernova event (see e.g., Refs. [1, 2] for recent reviews).

The flavor oscillations of neutrinos among $\nu_e$, $\nu_\mu$, and $\nu_\tau$ in vacuum and in medium have been well studied and confirmed by various ground-based neutrino experiments [3]. When neutrino fluxes are sufficiently high in CCSNe, the forward scattering among neutrinos themselves leads to various collective phenomena of flavor instability. Particularly the fast flavor conversion (FFC) associated with the fast flavor instability (FFI; see e.g., Refs. [4–7] for reviews) attracts great interest in recent years owing to the vastly rapid conversion rate within nanoseconds and over a distance shorter than a coin. Studies based on results from the multi-dimensional CCSN simulations have shown a general occurrence of FFIs in certain regions ahead of the shock wave, near the neutrinosphere, or even deep inside the PNS [8–17].

Since the first proposal of FFI in Ref. [18], the advance of both theory and methodology throughout the past decade has improved our understanding of FFCs and guided the line of research toward the ultimate goal of implementing neutrino oscillations in CCSNe. A generic framework governing the coherent flavor evolution and collisional neutrino-matter interactions is prescribed by the neutrino quantum kinetic equation ($\nu$QKE) [19–22]. The linear stability analysis (LSA) of the $\nu$QKE provides a powerful tool to diagnose the existence of flavor instabilities [23–26]. It has been proved based on the

* Email: z.xiong@gsi.de

LSA that the FFI requires the presence of angular zero-crossings where the angular distribution of the neutrino lepton number between any two distinct flavors transitions from positive to negative values [27]. The dispersion relation based on LSA implies the break of spatial homogeneity, which results in small-scale spatial structure confirmed by several dynamical simulations that numerically solved $\nu$QKE within tiny boxes [28–42]. Quasi-steady states are found to be achieved asymptotically on the coarse-grained level after neutrinos undergo the kinematic decoherence [29, 43–46]. It was reported by these simulations that the FFCs lead to complete flavor equilibration or partial flavor conversion depending on the ratio of neutrino fluxes and the characteristics of the angular zero-crossings. Particularly when a complete flavor equilibration or even overconversion is achieved, the neutrino flavor content can be significantly affected, which was demonstrated to be able to exert substantial impact on the neutrino-driven explosion mechanism and the nucleosynthesis of CCSNe [47–52], or even in the context of neutron-star mergers [53–60].

Recently, a growing amount of studies solving $\nu$QKE in a large spatial domain ranging from the interior of the PNS to the region around the shock wave were performed [61–66]. These studies intended to capture the flavor instabilities and to evolve both collective flavor conversions as well as the neutrino transport simultaneously in a self-consistent manner. Despite all these triumphs in the progress of FFCs, there remain several unaddressed questions. (1) Does the small-scale structure observed in the local simulations survive when the advection and neutrino-matter interactions are considered? (2) What is the general outcome of FFCs as well as its impact on the neutrino flux intensities and their ratios? Is flavor equilibration the ultimate fate of FFCs? (3) What consequence could be anticipated on CCSNe and the associated neutrino signals?

In addition to the FFI, there are other flavor instabilities such as the slow (see Refs. [1, 67] for reviews), or collisional type (see Refs. [68–80]) that might coexist with the FFI. Particularly, the collisional flavor instability (CFI), which was discovered recently, can potentially trigger flavor conversions in regions close to or inside the neutrinosphere where the FFI is known to play a role. Certainly, there are still open issues regarding these topics. It is important to understand whether and how different flavor instabilities interplay with each other in more sophisticated supernova models.

In this work, we aim to provide answers to the above questions (1)–(3), and to investigate the coexistence as well as impact of other flavor instabilities. We extend our previous multi-energy and multi-angle neutrino flavor evolution code [71, 81] to solve the $\nu$QKE by including a more complete set of collisional weak interactions in static snapshots of background environments obtained with spherically symmetric CCSN hydro simulations. Notice that the multi-energy spherically symmetric neutrino transport in standard CCSNe typically finds no angular crossing [77, 82, 83] because the imposed spherical symmetry restricts the capability of capturing the multi-dimensional CCSN features such as the fluid convection [83] and lepton-emission self-sustained asymmetry (LESA) [84, 85]. These multi-dimensional features can lead to attenuated $Y_e$ compared to the spherically symmetric model and create conditions for FFIs. Therefore, we follow the spirit of Ref. [49] and introduce radial-dependent $Y_e$ attenuation schemes in order to mimic the suitable conditions for FFCs commonly observed in multi-dimensional simulations. We use schemes with different attenuation strengths to probe a variety of scenarios for the FFI close to and inside the neutrinosphere.

We adopt a similar two-step approach as in Ref. [71]. We first relax neutrino profiles into stationary states up to $\simeq 1$ ms without including coherent flavor oscillations. Those stationary states are then used as the initial condition to simulate the evolution of neutrinos including diffusion, collisions, and flavor oscillations. We then present a comprehensive analysis for the FFCs covering three main aspects: the small-scale structure after the FFI emerges from the linear to nonlinear regimes, the dynamical evolution of the neutrino flavor content, and the effects on CCSN physics as well as the neutrino signals.

This paper is organized as follows. We provide the descriptions and parameters of our models for the FFCs in Sec. II. We use the LSA and the spectrogram method introduced in Sec. III to characterize the small-scale features of FFI in Sec. IV. We present a detailed study on the evolution of FFCs in Sec. V and discuss their implications in the energy transport, chemical composition of the matter, and free-streaming neutrino spectra in Sec. VI. Further discussions and conclusions are given in Sec. VII. We adopt natural units with $\hbar = c = k_B = 1$ throughout the paper.

## II. MODELS

In this section we introduce our theoretical model for neutrino flavor conversions in CCSNe. We begin with a review about the supernova model and the simulations based on which the flavor study is conducted. We then revisit the neutrino quantum kinetic approach and introduce our computational implementation.

### A. Supernova model

The CCSN simulation considered in this work is launched from the stellar progenitor star with zero-age main sequence mass of 25 $M_\odot$ and solar metalicity, from the stellar evolution series of Ref. [86], denoted as `s25a28`. Our AGILE-BOLTZTRAN supernova model is based on general relativistic spherical symmetric neutrino radiation hydrodynamics in comoving coordinates [87–91], including six-species Boltzmann neutrino trans-

port assuming ultra-relativistic particles [92, 93]. The equations of radiation hydrodynamics are solved on an adaptive baryon mass mesh [94, 95], for which 102 radial grid points are used for the current simulations, including about 5 $M_\odot$ of the stellar core, which covers parts of the extended silicon-sulfur layer for this progenitor model. The neutrino distributions are solved by means of the Boltzmann neutrino transport equation in the discrete-ordinate method. They are discretized in terms of 6 propagation angles and 36 energy bins (from 0.5–300 MeV) following the setup of Refs. [89, 96].

For the collision integral of the Boltzmann equation, a complete set of weak processes is employed (see Table I in Ref. [97]). This includes the charged-current neutrino emission and absorption (EA) involving heavy nuclei [98] and unbound nucleons [97], isoenergetic neutrino scattering on nuclei and nucleons (IS) [89, 96], inelastic neutrino scattering on electrons and positrons (NES) [88, 99], and neutrino pair reactions (PR), including electron-positron annihilation [96], nucleon-nucleon bremsstrahlung [100, 101] and the annihilation of electron neutrino pairs [102, 103]. EA are implemented in the full kinematics approach at the mean-field level, taking weak-magnetism contributions self-consistently into account [104]. Note that muonic weak processes are omitted here, such that muon and tauon flavors are not distinguished for neutrinos and antineutrinos. For the equation of state (EOS) used in the present supernova simulation, see details in Appendix A.

During the early post-bounce evolution, on the order of 10–25 ms after the core bounce, the $\nu_e$ deleptonization burst is released. It is a standard feature of CCSN phenomenology that results in the rapid drop of the electron fraction $Y_e$, accordingly, from around $Y_e = 0.3$–$0.4$ at core bounce to $Y_e \simeq 0.1$, in the region where the neutrinos decouple from matter. Even though neutrino decoupling is an energy and momentum dependent process (c.f. Refs. [105, 106]), often the neutrino-phase-space-averaged neutrinospheres of last elastic and inelastic scatterings are used as references [107], inside of which the bulk part of the neutrino spectrum is trapped and outside of which it is freely streaming. The post-bounce evolution of this `s25a28` progenitor model corresponds in spherically symmetric supernova simulations to the failed branch, leading eventually to the PNS collapse and the formation of a black hole [103, 108, 109]. Before that happens, the post-bounce phase is determined by a steady state mass accretion from the gravitationally unstable layers above the stellar core. The infalling material falls onto the standing bounce shock. The latter is located at a radius of about 80–100 km, depending on the progenitor model and the nuclear EOS. Thereby, the infalling heavy nuclei are dissociated into bulk nuclear matter, composed of unbound nucleons and light nuclei, that accumulate at the PNS surface. This layer, illustrated in Fig. 1 via the shock at $r \simeq 85$ km and the PNS surface at $r \simeq 40$ km depicting a shallow rise in density (solid curve) and temperature (dotted curve), is subject to the

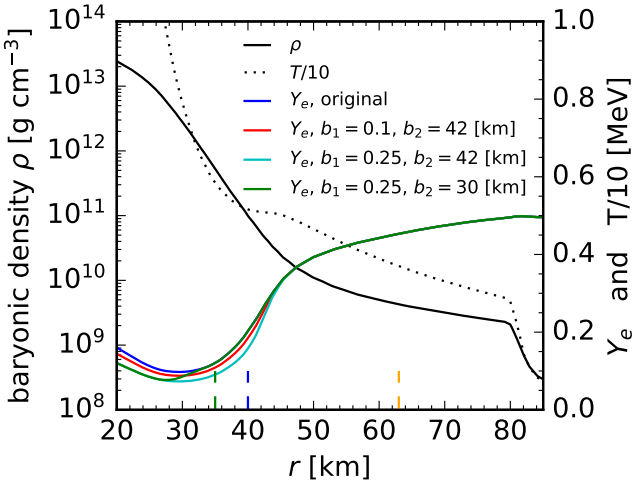



FIG. 1. Radial profiles of baryonic density, temperature $T/10$, and $Y_e$ (before and after attenuations) for a snapshot at the post-bounce time $t_{\rm pb} \approx 252$ ms. Orange, blue, and green vertical dashed lines indicate the gain radius, the radii of $\nu_e$- and $\bar{\nu}_e$-spheres at $r \approx 63$ km, 40 km, and 35 km, respectively.

neutrino decoupling, spanning a range of temperatures from up to $T \simeq 10$ MeV, at densities of a few times $10^{13}$ g cm$^{-3}$, down to $T \simeq 3$ MeV which corresponds to the shock location with densities around $10^9$ g cm$^{-3}$ where the temperature increases sharply due to the shock heating from $T \simeq 0.5$ MeV to $T \simeq 3$ MeV.

The timescale for the compression of the accumulated layer, settling at the PNS surface into the steepening gravitational potential, is determined by the high-density nuclear EOS, its temperature and isospin asymmetry dependence, and the enclosed PNS mass. As a consequence, neutrino luminosities $L_\nu$ are obtained on the order of several $10^{52}$ erg s$^{-1}$, after the $\nu_e$ deleptonization burst has been launched, and the average neutrino energies have values of $\langle E_\nu \rangle \simeq 10$–$20$ MeV. For a recent review of the role of neutrinos in the post-bounce phase of core-collapse supernovae, see Refs. [2, 110] and references therein.

Instead of surveying over different time stages of supernova, we will focus on one snapshot at a post-bounce time $t_{\rm pb} \approx 252$ ms in Fig. 1. Based on this snapshot, we explore different scenarios for FFCs by applying a radial-dependent attenuation to the original $Y_e$ profile to mimic suitable conditions for FFCs observed in multi-dimensional supernova simulations. A radial-dependent factor $b_{Y_e} = 1 - b_1/[1 + \exp(r - b_2[\text{km}])]$ is multiplied to the original $Y_e$ profile with $b_1$ denoting the fraction of attenuation and $b_2$ characterizing the radial range of attenuation. With the attenuated electron fraction $Y_e$ and smaller electron chemical potential $\mu_{e^-}$, the equilibrium neutrino chemical potential $\mu^{\rm eq}_{\nu_e} = \mu_{e^-} + \mu_p - \mu_n$ is reduced to create a condition less dominated by $\nu_e$'s, or even a condition where $\bar{\nu}_e$'s dominate over $\nu_e$'s. We vary $b_1$ from 10% to 25%, allowing different angular crossings dominated by $\nu_e$ or $\bar{\nu}_e$ to appear, which will be discussed later in Sec III A. We show

in Fig. 1 three attenuated $Y_e$ profiles with different values of $(b_1, b_2[\mathrm{km}]) = (0.1, 42), (0.25, 42)$, and $(0.25, 30)$, respectively. Note that taking $b_2 = 30$ km restricts the FFI solely within the PNS, but not extending to the free-streaming regime. Also noted is that we only attenuate $Y_e$ without modifying the neutron and proton abundances for simplification. When the latter abundances are modified accordingly based on the EOS, $\mu_p$ decreases, and $\mu_n$ increases. As a result, the attenuation factor, $b_1$, required to achieve a similar scenario for each model is reduced by approximately half. In addition, although the $Y_e$ profile with the largest attenuation factor adopted here might be hard to be obtained in typical supernovae, it may mimic other physical conditions such as the second collapse in a hadron-quark phase transition supernova featuring a unique burst-like neutrino signature dominated by $\bar{\nu}_e$ [111–115].

### B. Neutrino quantum kinetic equation

We solve the neutrino quantum kinetic equation ($\nu$QKE) in both two- and three-flavor schemes. The general equations governing the spatial and temporal evolution of density matrices $\varrho$ (for neutrinos) and $\bar{\varrho}$ (for anti-neutrinos) are given by

$$(\partial_t + v_r\partial_r + \frac{1-v_r^2}{r}\partial_{v_r})\varrho = -i[\mathbf{H}, \varrho] + \mathbf{C}, \tag{1}$$

and

$$(\partial_t + v_r\partial_r + \frac{1-v_r^2}{r}\partial_{v_r})\bar{\varrho} = -i[\bar{\mathbf{H}}, \bar{\varrho}] + \bar{\mathbf{C}}, \tag{2}$$

with

$$\varrho = \begin{bmatrix} \varrho_{ee} & \varrho_{e\mu} & \varrho_{e\tau} \\ \varrho_{e\mu}^* & \varrho_{\mu\mu} & \varrho_{\mu\tau} \\ \varrho_{e\tau}^* & \varrho_{\mu\tau}^* & \varrho_{\tau\tau} \end{bmatrix} \text{ and } \bar{\varrho} = \begin{bmatrix} \bar{\varrho}_{ee} & \bar{\varrho}_{e\mu} & \bar{\varrho}_{e\tau} \\ \bar{\varrho}_{e\mu}^* & \bar{\varrho}_{\mu\mu} & \bar{\varrho}_{\mu\tau} \\ \bar{\varrho}_{e\tau}^* & \bar{\varrho}_{\mu\tau}^* & \bar{\varrho}_{\tau\tau} \end{bmatrix} \tag{3}$$

in three flavor bases or

$$\varrho = \begin{bmatrix} \varrho_{ee} & \varrho_{e\mu} \\ \varrho_{e\mu}^* & \varrho_{\mu\mu} \end{bmatrix} \text{ and } \bar{\varrho} = \begin{bmatrix} \bar{\varrho}_{ee} & \bar{\varrho}_{e\mu} \\ \bar{\varrho}_{e\mu}^* & \bar{\varrho}_{\mu\mu} \end{bmatrix}, \tag{4}$$

in a simplified neutrino system of two flavors. Their diagonal elements are normalized to the neutrino number density $n_{\nu_i}(r,t) = \int dE\, dv_r\, \varrho_{ii}$ where $i = e, \mu, \tau$.

In general, the coherent propagation Hamiltonian $\mathbf{H}$ on the right-hand side has three contributions. They are the vacuum mixing term

$$\mathbf{H}_{\mathrm{vac}}(E) = \mathbf{U}\mathbf{M}\mathbf{U}^\dagger, \tag{5}$$

with the diagonal matrix in mass basis $\mathbf{M} = \mathrm{diag}[0, \delta m_{\mathrm{L}}^2, \delta m^2]/(2E)$ and the unitary mixing matrix $\mathbf{U}$ parametrized as in Eq. (14.34) of Ref. [3], the matter term

$$\mathbf{H}_{\mathrm{mat}} = \mathrm{diag}[V_{\mathrm{mat}}, 0, 0], \tag{6}$$

with the effective potential $V_{\mathrm{mat}}$ corresponding to neutrino forward scattering on ordinary matter, and the neutrino self-induced term

$$\mathbf{H}_{\nu\nu}(v_r) = \sqrt{2}G_F \int dE'\, dv_r'(1 - v_r v_r')(\varrho - \bar{\varrho}^*), \tag{7}$$

corresponding to neutrino forward scattering on other neutrinos, respectively. In the two-flavor scheme, the vacuum term is explicitly

$$\mathbf{H}_{\mathrm{vac}}(E) = \frac{\delta m^2}{4E}\begin{bmatrix} -\cos 2\theta_V & \sin 2\theta_V \\ \sin 2\theta_V & \cos 2\theta_V \end{bmatrix} \tag{8}$$

with one mass-square difference $\delta m^2$ and one mixing angle $\theta_V$. Although the matter term may affect the behaviors of flavor oscillations [71, 116], we will not include it explicitly in the following discussion of this work.

The collisional term $\mathbf{C}$ includes all four types of weak processes in Table I. We use the same scheme for the EA and IS processes as in Ref. [71] except that muonic EA reactions are not considered.

TABLE I. Set of weak processes considered in COSE$\nu$, where $\nu$ and $\bar{\nu}$ are for all neutrino flavors and $N = n, p$.

| Label | Weak process | Abbreviation |
|---|---|---|
| (1a) | $\nu_e + n \leftrightarrows p + e^-$ | EA |
| (1b) | $\bar{\nu}_e + p \leftrightarrows n + e^+$ | EA |
| (1c) | $\bar{\nu}_e + p + e^- \leftrightarrows n$ | EA |
| (2a) | $\nu + N \leftrightarrows \nu + N$ | IS |
| (2b) | $\bar{\nu} + N \leftrightarrows \bar{\nu} + N$ | IS |
| (3a) | $\nu + e^\pm \leftrightarrows \nu + e^\pm$ | NES |
| (3b) | $\bar{\nu} + e^\pm \leftrightarrows \bar{\nu} + e^\pm$ | NES |
| (4a) | $\nu + \bar{\nu} \leftrightarrows e^- + e^+$ | PR |
| (4b) | $\nu + \bar{\nu} + N + N \leftrightarrows N + N$ | PR |

Low-energy electron neutrinos and heavy-lepton neutrinos are thermalized by NES and PR. Because of the high computational cost of the scattering kernels associated with them, approximated schemes are employed. For PR, we integrate the scattering kernels in Refs. [96, 117] into the emissivity $j_{\mathrm{PR}}$ and opacity $\chi_{\mathrm{PR}}$ for all flavors in a similar way as in Ref. [118] except assuming that the accompanied neutrinos or antineutrinos in the pair processes obey the Fermi-Dirac distribution determined by the equilibrium neutrino chemical potential (e.g., $\mu_{\nu_e}^{\mathrm{eq}}$ for $\nu_e$), in order to take the blocking effect into account. For heavy-lepton neutrinos, the equilibrium neutrino chemical potential is assumed to be zero. The collisional term for neutrino can then be constructed in an effective way similar to the EA processes as follows,

$$\begin{aligned}\mathbf{C}_{\mathrm{PR}}(E) =& \frac{1}{2}\{\mathrm{diag}[j_{e,\mathrm{PR}}, j_{\mu,\mathrm{PR}}, j_{\tau,\mathrm{PR}}], \varrho_{\mathrm{FO}} - \varrho\} \\ & - \frac{1}{2}\{\mathrm{diag}[\chi_{e,\mathrm{PR}}, \chi_{\mu,\mathrm{PR}}, \chi_{\tau,\mathrm{PR}}], \varrho\}\end{aligned} \tag{9}$$

with the fully-occupied differential number density $\varrho_{\mathrm{FO}}$ and the curly bracket being the anti-commutator. The

full term of NES in $\nu$QKE involves channels of different Mandelstam variables [22], which increases the complication of the treatment. For simplicity, we take the classical treatment particularly for NES by only keeping the diagonal elements of $\mathbf{C}_{\rm NES}$ [88, 99, 119, 120]. The thermalization of the energy spectra of heavy-lepton neutrinos is guaranteed by the Kirchhoff-Planck relation ("detailed balance") between the in- and out-scattering kernels

$$R^{\rm in}_{\rm NES}(E, E', v_r, v'_r) = e^{\frac{E'-E}{T}} R^{\rm out}_{\rm NES}(E, E', v_r, v'_r). \quad (10)$$

In addition, for each flavor the NES scattering kernel is truncated at the zeroth moment following the method in [99, 121] (See Appendix B for more details).

### C. Numerical setup

The $\nu$QKE is solved by COSE$\nu$ under spherical symmetry within a radial range uniformly discretized by $N_r$ radial grids between $r_{\rm ib} = 20$ km and $r_{\rm ob} = 80$ km. At the inner boundary neutrinos of all species are in Fermi distribution functions with chemical potentials given by weak equilibrium with the processes of NES and PR included. The outer boundary is chosen to be around the position of the shock wave. We employ the free-streaming boundary condition there for forward propagating neutrinos with $v_r \geq 0$. Although neutrinos can be backscattered by heavy nuclei outside the shock through the neutrino-nucleus interactions, which leads to the presence of FFI around the shock [12], a recent work suggests that their effect on the dominant neutrino outflow is negligible due to the relatively diluted flux [35]. Therefore, we inject no neutrinos in the backward direction with $v_r < 0$.

Because the vacuum term $\mathbf{H}_{\rm vac}$ with a non-zero mixing angle generates flavor mixing automatically, we do not employ any artificial perturbation as the seed to trigger flavor conversions in the initial condition.

For consistency across all models, the neutrino spectral and angular distributions are all discretized into $N_E = 15$ energy grids and $N_{v_r} = 100$ angular grids for both $\varrho$ and $\bar{\varrho}$. The energy grids are taken between 1 MeV to 100 MeV spaced uniformly in logarithmic scale, and the angular grids are linearly uniform between $v_r = -1$ and 1.

Because both EA and IS rates that generally dominate over others strongly depend on the neutrino energy and nucleonic density, taking a smaller maximum neutrino energy and a larger radius of the inner boundary compared to our previous work [71] reduces the maximum collisional rate that can be well resolved within our simulation domain. Therefore, unlike in [71], no attenuation is made to any collisional rate for all models in this work.

In contrast to the collisional rates, a common challenge is the tremendous size of the neutrino self-induced term $\mathbf{H}_{\nu\nu}$ associated with subnanosecond timescales and subcentimeter length scales. They are much shorter than the scales related to the hydrodynamics (order of kilometers and milliseconds), which makes it computationally difficult for a simulation to well resolve both time and length scales with the unadjusted $\mathbf{H}_{\nu\nu}$. A strategical mitigation in recent works [63, 71] is to attenuate the values of $\mathbf{H}_{\nu\nu}$ while generally keeping the same ordering that the magnitude of $\mathbf{H}_{\nu\nu}$ dominates over collisional rates and vacuum oscillation frequency.[1] The oscillation term $\mathbf{H}$ on the right-hand side of Eq. (1) becomes $\mathbf{H}_{\rm vac} + a_{\nu\nu}(r)\mathbf{H}_{\nu\nu}$ with a radial-dependent factor $a_{\nu\nu}(r) = a_1/[1 + e^{(a_2-r)/a_3}]$ where $a_2 = 30$ km and $a_3 = 2.5$ km for all our models. This factor keeps almost constant $a_{\nu\nu}(r) \approx a_1$ for most of the radial range with $r > 35$ km. At inner radii where the neutrino number densities are much larger, more attenuation is provided by the denominator in the formula by around one more order of magnitude, i.e., the attenuation factor $a_{\nu\nu}$ takes a value of $0.1a_1$ at $r \approx 24.5$ km. For example, when $a_1$ is taken to be $10^{-3}$, the attenuation factor is always greater than $10^{-4}$ for $r > 24.5$ km and close to $10^{-3}$ for $r > 35$ km, which is a reasonable choice compared to other works [58, 66]. We utilize this attenuation strategy and explore different choices of the attenuation parameter $a_1$, with the radial resolutions $N_r$ adapted accordingly.

We summarize all models and their parameters of this paper in Table II. There is no flavor instability in Model I. In both two- and three-flavor models, all mixing angles in $\mathbf{H}_{\rm vac}$ are set to be $10^{-6}$ accounting for the effective suppression by the matter. In three-flavor models, the CP violation phase is omitted, and the lower mass-square difference is taken as $\delta m^2_{\rm L} = 7.39 \times 10^{-5}$ eV$^2$. In order to investigate the effects of other types of flavor instabilities on the FFCs, we run simulations with different values of $\delta m^2$ in the vacuum term and introduce an auxiliary parameter $c_{\rm off}$ that are multiplied to all off-diagonal elements of $\mathbf{C}$. The default value of $c_{\rm off}$ is 1. When $c_{\rm off}$ is set to be zero, the CFI is prohibited.

## III. ANALYSES

### A. Illustrative quantities

For a better illustration of our models, we define several quantities as follows. The energy- and angle-integrated density matrices are defined by $\langle\varrho\rangle_E(v_r) = \int dE\, \varrho(E, v_r)$ and $\langle\varrho\rangle_A(E) = \int dv_r\, \varrho(E, v_r)$, respectively. The neutrino mean energy of a given flavor is calculated as

$$\langle E_{\nu_i}\rangle = \frac{\int dE\, dv_r\, E\varrho_{ii}(E, v_r)}{\int dE\, dv_r\, \varrho_{ii}(E, v_r)}, \quad (11)$$

where $i = e, \mu, \tau$. A dimensionless ratio, $s_{e\mu} = |\langle\varrho_{e\mu}\rangle_E|/(|\langle\varrho_{ee}\rangle_E - \langle\varrho_{\mu\mu}\rangle_E|^2/4 + |\langle\varrho_{e\mu}\rangle_E|^2)^{1/2}$, is defined

[1] There are also other strategies, for example, assuming that neutrinos are homogeneous in radial blocks of a few hundred meters to reduce the needed number of the radial grids in computation and to allow for no attenuation on $\mathbf{H}_{\nu\nu}$ [64].

TABLE II. Parameters used in each model. For all models, $r_{\rm ib} = 20$ km to $r_{\rm ob} = 80$ km, $t_{\rm pb}$=252 ms, $a_2 = 30$ km, $a_3 = 2.5$ km, $N_{v_r} = 100$, $N_E = 15$.

| model | $b_1$ | $b_2$ [km] | no. of flavors | $a_1$ | $N_r$ | $c_{\rm off}$ | $\delta m^2$ [eV$^2$] |
|---|---|---|---|---|---|---|---|
| I | 0 | | | | | | |
| II | 0.1 | 42 | 2 | $10^{-3}$ | 25000 | 1 | $8\times10^{-5}$ |
| III | 0.15 | 42 | 2 | $10^{-3}$ | 25000 | 1 | $8\times10^{-5}$ |
| IV | 0.2 | 42 | 2 | $10^{-3}$ | 25000 | 1 | $8\times10^{-5}$ |
| V | 0.25 | 42 | 2 | $10^{-3}$ | 25000 | 1 | $8\times10^{-5}$ |
| VI | 0.25 | 32 | 2 | $10^{-3}$ | 25000 | 1 | $8\times10^{-5}$ |
| VII | 0.25 | 30 | 2 | $10^{-3}$ | 25000 | 1 | $8\times10^{-5}$ |
| Vc0 | 0.25 | 42 | 2 | $10^{-3}$ | 25000 | 0 | $8\times10^{-5}$ |
| IIf3 | 0.1 | 42 | 3 | $10^{-3}$ | 25000 | 1 | $8\times10^{-5}$ |
| IVf3 | 0.2 | 42 | 3 | $10^{-3}$ | 25000 | 1 | $8\times10^{-5}$ |
| IIa2 | 0.1 | 42 | 2 | $2\times10^{-3}$ | 50000 | 1 | $8\times10^{-5}$ |
| IIa4 | 0.1 | 42 | 2 | $4\times10^{-3}$ | 50000 | 1 | $8\times10^{-5}$ |
| IIa10 | 0.1 | 42 | 2 | $10^{-2}$ | 50000 | 1 | $8\times10^{-5}$ |
| IIIa2 | 0.15 | 42 | 2 | $2\times10^{-3}$ | 50000 | 1 | $8\times10^{-5}$ |
| IIIa4 | 0.15 | 42 | 2 | $4\times10^{-3}$ | 50000 | 1 | $8\times10^{-5}$ |
| IVa2 | 0.2 | 42 | 2 | $2\times10^{-3}$ | 50000 | 1 | $8\times10^{-5}$ |
| IVa4 | 0.2 | 42 | 2 | $4\times10^{-3}$ | 50000 | 1 | $8\times10^{-5}$ |
| IVv2 | 0.2 | 42 | 2 | $10^{-3}$ | 25000 | 1 | $2.4\times10^{-3}$ |

to quantify the nonlinearity of flavor conversions caused by FFI. The corresponding phase (angle in the complex plane) is defined as $\phi_{e\mu} = \arg[\langle\varrho_{e\mu}\rangle_E \,\mathrm{sgn}(\langle\varrho_{ee}\rangle_E - \langle\varrho_{\mu\mu}\rangle_E)]$, where sgn denotes the sign function. The electron-minus-muon lepton number of neutrinos (EMuLN, or shortened as ELN in the following discussion) is $G = \langle\varrho_{ee}\rangle_E - \langle\bar\varrho_{ee}\rangle_E - \langle\varrho_{\mu\mu}\rangle_E + \langle\bar\varrho_{\mu\mu}\rangle_E$. When there only exists one zero-crossing in the ELN angular distribution, we define two integrals $I_+ = \int_{-1}^{1} dv_r\, G\Theta(G)$ and $I_- = \int_{-1}^{1} dv_r\, G\Theta(-G)$, where $\Theta$ is the Heaviside step function. Those two quantities allow us to evaluate the "deepness ratio" of the ELN angular crossing based on the value of $\min(|I_-|, |I_+|)/\max(|I_-|, |I_+|)$.

Except the original Model I without $Y_e$ attenuation and angular crossings, Fig. 2 shows the initial ELN distributions for Models II–VII in the stationary state obtained by solving Eqs. (1) and (2) without including the commutators on the right-hand side. Clearly, these six models show a significant variation in the radial shapes of their ELN angular crossings due to different $Y_e$ attenuation, while each of them only contains at most one angular crossing at any specific radius. In particular, the ELN crossings in Models II–V extend all the way to the outer boundary where neutrinos freely stream outwards. With the least $Y_e$ attenuation imposed in Model II (corresponding to the red curve in Fig. 1), a relatively shallow angular crossing begins to appear with crossing radial velocity $v_{r,c} \lesssim 1$ at $r \approx 38$ km, and extends to smaller $v_{r,c}$ as low as $v_{r,c} \approx 0.8$ at $r \approx 41$ km. For the entire domain, the deepness ratios are less than 0.06 as shown in Fig. 3. When we apply more attenuation on $Y_e$, which results in more production of $\bar\nu_e$'s, the starting radii where the angular crossing appears shift to $r \approx 33$ km and 28 km in Model III and IV, respectively. The minimum values of $v_{r,c}$ are $v_{r,c} \approx 0.3$ at $r \approx 38$ km and $v_{r,c} \approx -0.4$ at $r \approx 32$ km accordingly. For Model V (corresponding to the cyan curve in Fig. 1), the inner boundary becomes so $\bar\nu_e$-dominated that the crossing starts from $v_{r,c} = -1$ at $\approx 35$ km and monotonically approaches $v_{r,c} \approx 0.85$ as the radius increases. The deepness ratios in Models III–V are generally greater than 0.3, particularly reaching $\approx 1$ at $r \simeq 30$ km and 45 km in Model IV. In the last two models (VI and VII), ELN crossings are only confined inside the neutrino spheres as they originate from the change of sign of $\mu^{\rm eq}_{\nu_e}$ due to the $Y_e$ attenuation. The crossing in Model VI spans from $r \approx 40$ km to $r \approx 42$ km, while it is narrower in Model VII (corresponding to the green curve in Fig. 1), ranging from $r \approx 38$ km to $r \approx 39$ km. These different radial shapes of ELN crossings lead to very different FFI and the evolution of FFC, which will be discussed in the following sections.

### B. Linear stability analysis

We perform the LSA for models with two flavors. The analyses on three-flavor models are expected to be similar when neutrinos of muon and tauon flavors are identical, although a nontrivial interplay among flavors may happen when difference in $\mu$ and $\tau$ neutrinos exist [122, 123]. Assuming $|\varrho_{e\mu}|/|\varrho_{ee} - \varrho_{\mu\mu}| \ll 1$ and $|\bar\varrho_{e\mu}|/|\bar\varrho_{ee} - \bar\varrho_{\mu\mu}| \ll 1$ before any instability leads to fla-

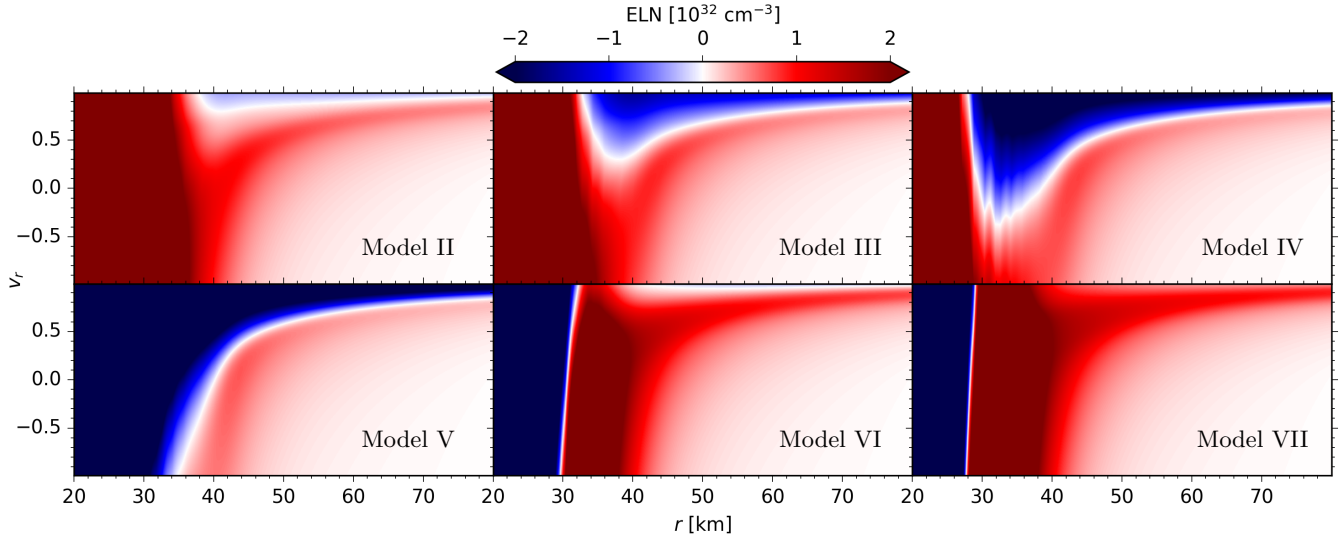



FIG. 2. Radial profiles of ELN distribution in Models II–VII that have different attenuated $Y_e$ profiles. The angular crossings of ELN are indicated by the location of curves with the white color. Model I (without $Y_e$ attenuation) is not shown as it contains no ELN angular crossing.

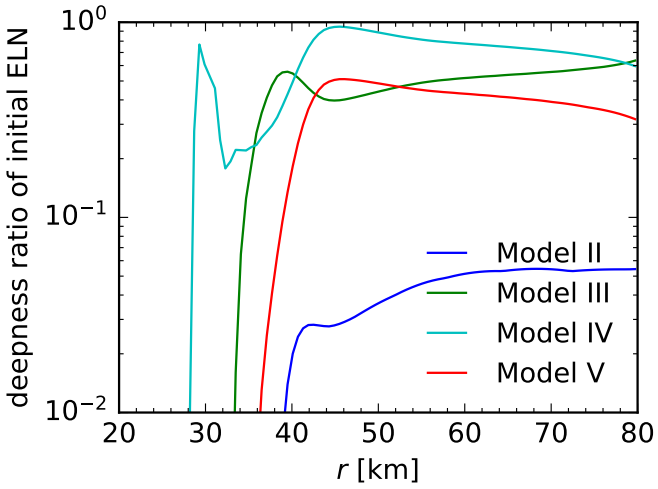



FIG. 3. Deepness ratios of the initial ELN distribution, $\min(|I_-|,|I_+|)/\max(|I_-|,|I_+|)$, in Models II–V (see text for the definition of $I_\pm$).

vor conversion, $\nu$QKE can be linearized to find unstable normal modes if flavor instability exists [23, 24].d We follow our previous treatment in [71], neglecting the terms associated with $\partial_{v_r}$ and IS but including the vacuum term. We further assume that a collective mode of the perturbation $\varrho_{e\mu} = Q(\Omega, K_r, E, r, v_r)e^{-i[\Omega t - K_r(r'-r)]}$ and $\bar{\varrho}^*_{e\mu} = \bar{Q}(\Omega, K_r, E, r, v_r)e^{-i[\Omega t - K_r(r'-r)]}$ can develop locally near $r$, so the off-diagonal parts of Eqs. (1) and (2) become[2]

$$\left[\Omega - K_r v_r - \Phi(v_r) + iC_{e\mu}(E) + \delta m^2/2E\right] Q(E, v_r)$$
$$= -\sqrt{2}G_F[\varrho_{ee}(E, v_r) - \varrho_{\mu\mu}(E, v_r)] \int dE'\, dv'_r \times$$
$$(1 - v_r v'_r)[Q(E', v'_r) - \bar{Q}(E', v'_r)], \tag{12}$$

and

$$\left[\Omega - K_r v_r - \Phi(v_r) + i\bar{C}_{e\mu}(E) - \delta m^2/2E\right] \bar{Q}(E, v_r)$$
$$= -\sqrt{2}G_F[\bar{\varrho}_{ee}(E, v_r) - \bar{\varrho}_{\mu\mu}(E, v_r)] \int dE'\, dv'_r \times$$
$$(1 - v_r v'_r)[Q(E', v'_r) - \bar{Q}(E', v'_r)], \tag{13}$$

where $\Phi(v_r) = \sqrt{2}G_F \int dE'\, dv'_r\, (1 - v_r v'_r)[\varrho_{ee}(E', v'_r) - \varrho_{\mu\mu}(E', v'_r) - \bar{\varrho}_{ee}(E', v'_r) + \bar{\varrho}_{\mu\mu}(E', v'_r)]$, $C_{e\mu}(E)$ and $\bar{C}_{e\mu}(E)$ are the damping rates attributed to EA and the approximate PR terms. For example, $C_{e\mu}(E) = (j_{e,\mathrm{EA}} + \chi_{e,\mathrm{EA}} + j_{e,\mathrm{PR}} + \chi_{e,\mathrm{PR}} + j_{\mu,\mathrm{PR}} + \chi_{\mu,\mathrm{PR}})/2$. To account for the attenuation scheme used for $\mathbf{H}_{\nu\nu}$, we modify Eqs. (12) and (13) by multiplying $G_F$ with the attenuation factor $a_{\nu\nu}$, which affects the right-hand side and the term $\Phi$ on the left-hand side.

For a given wave number $K_r$, the eigenvalues of $\Omega$ and their corresponding eigenvectors can be numerically solved. The contribution of $\Phi$ is often absorbed into $\Omega$ and $K_r$ in literature [24, 26], but we do not follow this procedure in order to keep the original meaning of $K_r$ as the wave number in the local frame. When $K_r = 0$, it

[2] The dependence of $Q$ and $\bar{Q}$ on $\Omega$, $K_r$, and $r$ is not shown explicitly.

represents the homogeneous mode. We highlight that we search for non-zero $K_r$ modes in additional to the homogeneous one to avoid missing the most unstable mode.

We note that there may exist some other aspects that cannot be captured by the linearized equations (12) and (13). For instance, the behaviors of flavor instability may be affected by the angular advection term associated with $\partial_{v_r}$ and the fact that neutrino number density is not homogeneous. Although the pure FFI without collisions and with vanishing vacuum oscillation frequency is free from the contamination of spurious modes [124], which are associated with the use of an insufficient number of angular grids in the discretization, those spurious modes may appear in multiangle slow instabilities when the vacuum term is taken into account [125–127] or in the CFI when collisions are included [74, 128]. Their growth rates should approach zero as the adopted angular grid number increases toward the continuous limit.

### C. Spectrogram

In addition to the LSA, the onset of specific unstable modes as well as the small-scale spatial structure can be simply analyzed by the Fourier transform of relevant quantities over the entire simulation domain in the periodic-box setup. In contrast, the inhomogeneous neutrino and matter profiles in the spherical symmetric supernova model do not allow such a simple approach. Nonetheless, the short-time Fourier transform (or spectrogram analysis) provides a powerful tool in signal processing to diagnose the characteristic features in the frequency domain over a short-time window for a time-varying system, which has been generically used in the engineering and science (see, e.g., [129–131]). Based on the same spirit, we can perform similar spectrogram analysis to the neutrino flavor evolution outcomes except that here we calculate the spectra in terms of wave number $K_r$ over a short-length window.

The spectrogram analysis can be performed in various ways. In order to mitigate the spectral leakage, we use the Tukey window function in the spectrogram analysis [132, 133]. The flat top of Tukey window also naturally provides less amplitude attenuation than in others such as a Gaussian window. We divide the whole radial range of our simulation domain into $N_r^{\rm sp}$ blocks with each block centered at a radius $r$, and calculate discretized Fourier transform for the neutrino off-diagonal mixing at a snapshot of simulation in each block

$$\mathcal{F}_{r,K_r} = \frac{1}{\Delta r^{\rm sp}} \int_{r-\frac{5}{8}\Delta r^{\rm sp}}^{r+\frac{5}{8}\Delta r^{\rm sp}} dr'\, e^{-iK_r r'} W(r'-r) \times \int dv_r \, \langle \varrho_{e\mu} \rangle_E(v_r, r') {\rm sgn}[\langle \varrho_{e\mu} \rangle_E(v_r, r') - \langle \varrho_{e\mu} \rangle_E(v_r, r')], \tag{14}$$

where $\Delta r^{\rm sp} = (r_{\rm ob} - r_{\rm ib})/N_r^{\rm sp}$ is the radial width of the nonoverlapping part of each block, $K_r$ takes discretized values $z\pi/\Delta r^{\rm sp}$ km$^{-1}$ with integers $z$, and $W(r)$ is the Tukey window function with a cosine lobe

$$W(r) = \begin{cases} 1 & \text{if } |r| \le \Delta r^{\rm sp}/2, \\ \cos^2\left(\frac{4\pi|r|}{\Delta r^{\rm sp}} - 2\pi\right) & \text{if } |r| > \Delta r^{\rm sp}/2. \end{cases} \tag{15}$$

Depending on $N_r$ used in the simulations, we take $N_r^{\rm sp}$ as 50 or 100 to ensure sufficient resolution in $K_r$ space.

## IV. EMERGENCE OF FLAVOR INSTABILITY

### A. Scalability of the dispersion relation

Perturbed by the vacuum flavor mixing (or potentially quantum fluctuation), the flavor waves of the neutrino fields can be launched and lead to flavor transformation driven by the dispersion relation and the unstable collective modes in LSA. Ignoring the contributions from other aspects such as advection, collisions and the vacuum term, the dispersion relation branches for the axisymmetric unstable mode can have distinct structures for different ELN distributions [26]. For a shallow crossing, the dispersion relation typically yields two separated unstable branches with respect to the wave number of the perturbed flavor wave $K_r$. For a deep crossing, only one unstable branch exists, and the maximum growth rate of the branch is enhanced.

All LSA results shown below are based on the stationary neutrino profiles obtained without including oscillations (see e.g., Fig. 2 for the energy-integrated profiles). The labels `v0`, `v1`, `v2`, and `v-2` represent results using four different values of $\delta m^2 = 0$ eV$^2$, $8\times10^{-5}$ eV$^2$, $2.4\times10^{-3}$ eV$^2$ and $-2.4\times10^{-3}$ eV$^2$, respectively. The labels `c0` and `c1` denote settings with $c_{\rm off} = 0$ and 1, respectively. Note that a positive (negative) $\delta m^2$ value represents the normal (inverted) neutrino mass ordering.

We begin with Model II that has the most shallow angular crossing compared to other models. Figure 4(a–c) shows the growth rates Im($\Omega$) of the unstable mode in Model II with different choices of $\delta m^2$, $c_{\rm off}$, and $a_{\nu\nu}$. From Eqs. (12) and (13), the dispersion relation of the FFI is clearly scalable when the vacuum and collision terms are omitted. In this case, applying the attenuation factor $a_{\nu\nu}$ simply reduces the magnitude of $\Omega$ and $K_r$ for the same amount but does not affect the shape of the dispersion relation branches.

The scalability approximately holds even if the collisional terms are included. Figure 4(a) compares the unstable branches of Im($\Omega$) at $r \approx 41.6$ km in Model II from the case without attenuation ($a_{\nu\nu} = 1$) and from cases with attenuation factors $a_1 = 10^{-2}$ and $a_1 = 10^{-3}$. For the later two with attenuation, their unstable branches are shown by multiplying factors of $10^2$ and $10^3$ to both $K_r$ and Im($\Omega$), respectively. Among those three cases, there is good agreement on the rescaled growth rate in the dominant parts of the two separated unstable branches.

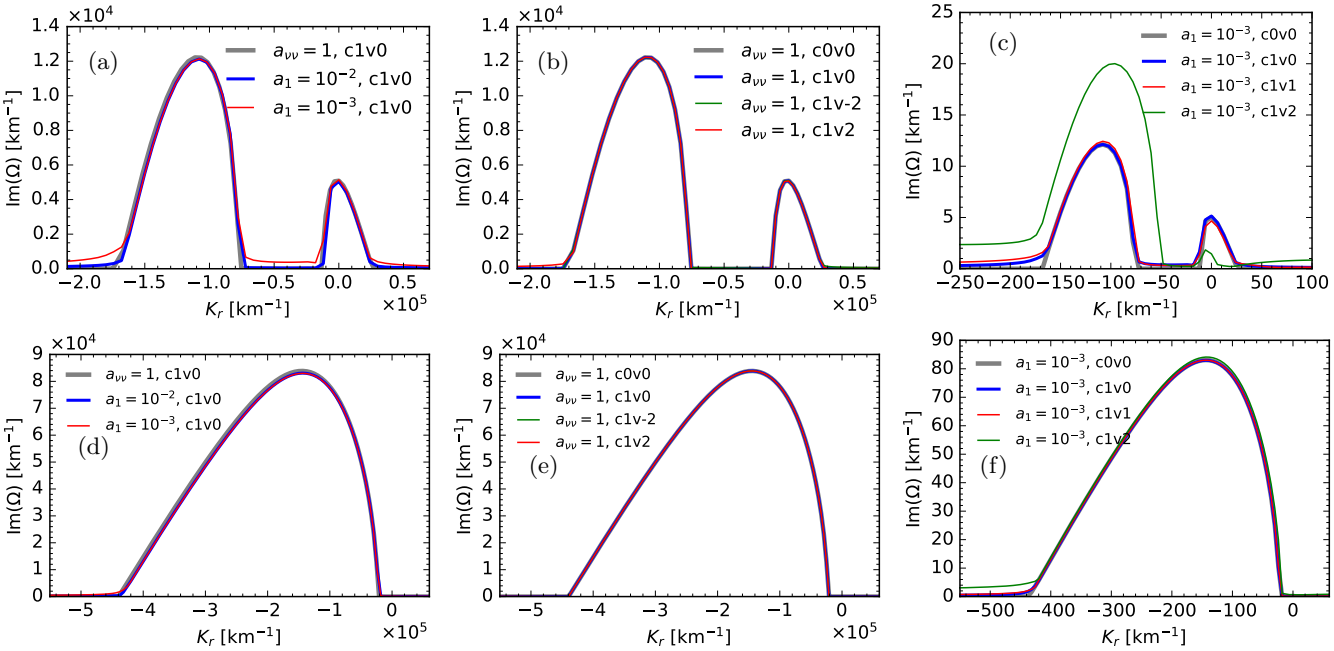



FIG. 4. Growth rates Im(Ω) as functions of $K_r$ at $r = 41.6$ km in Models II (a–c), and V (d–f). The values of $\mathrm{Im}(\Omega)_{\max}$ for the cases with $a_1 = 10^{-2}$ and $a_1 = 10^{-3}$ in (a) and (d) are multiplied by $10^2$ and $10^3$, respectively, to show the scalability of the dispersion relation discussed in the text.

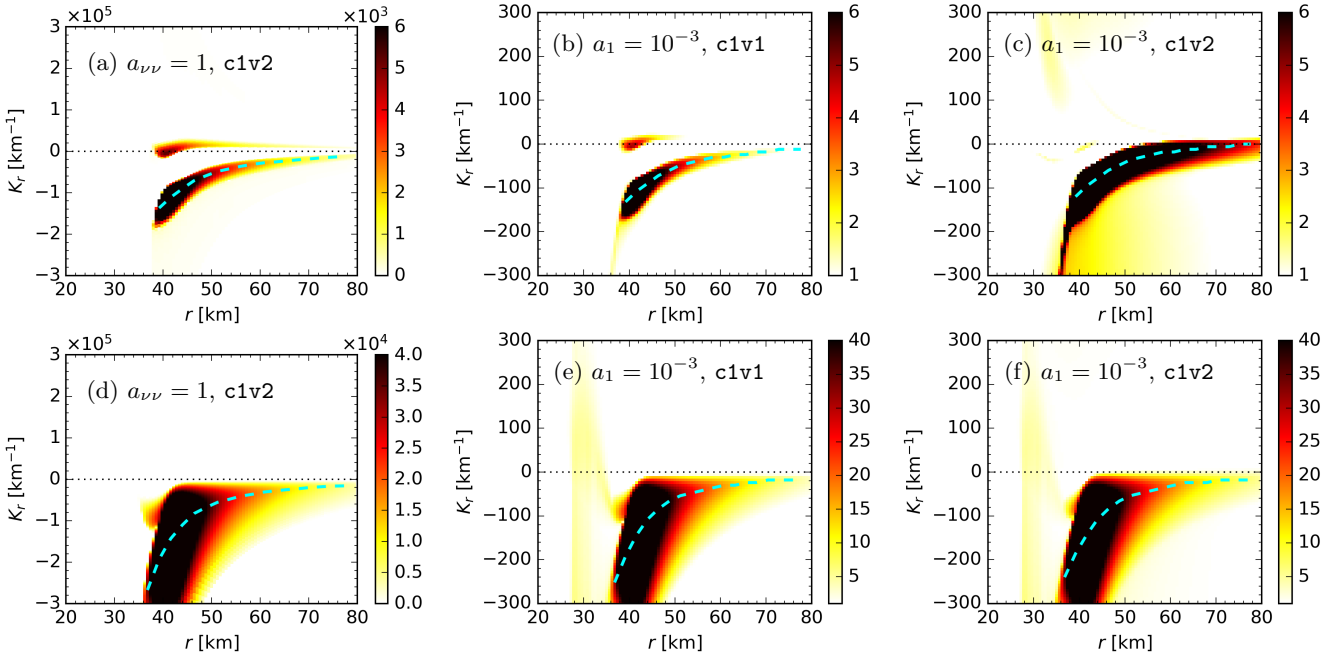



FIG. 5. Growth rates Im(Ω) [km$^{-1}$] as functions of $K_r$ and $r$ in Models II (a–c) and V (d–f). The cyan dashed curve marks $K_r$ with maximum growth rates.

We further examine the effects of collisions and vacuum terms on the dispersion relation in Model II for the case with $a_{\nu\nu} = 1$ and $a_1 = 10^{-3}$ in Fig. 4(b) and (c) separately. Figure 4(b) shows that without attenuation, the dispersion relation of FFI is hardly changed in the presence of collisions and the vacuum term. This suggests that at this radius, the neutrino number density is so high that other types of flavor instabilities are overwhelmed by the large growth rate from the fast mode.

When taking the attenuation factor $a_1 = 10^{-3}$ in Fig. 4(c), the shape of the dispersion relation remains quantitatively similar when including the collisions in

the case of `c1v0` or a vacuum term with smaller $\delta m^2$ in `c0v1`. A slight difference is that small nonzero growth rates $\mathrm{Im}(\Omega) \simeq 1$ km$^{-1}$, which connect the two otherwise separated branches of pure fast mode when only the attenuated $\mathbf{H}_{\nu\nu}$ is included (case `c0v0`), become visible in comparison to Fig. 4(b). This is because with the attenuation, $\mathbf{H}_{\nu\nu}$ becomes closer to the $\mathbf{H}_{\rm vac}$ in magnitude.[3] When we take a larger vacuum term with $\delta m^2 = 2.4\times10^{-3}$ eV$^2$, the slow flavor instability becomes so strong that it mixes with the FFI and drastically affects the structure of the dispersion relation branches. The growth rate of the left unstable branch is enhanced by a factor of $\simeq 2$, and that of the right one is suppressed.

For Model V with a much deeper ELN angular crossing at the same radius than that in Model II, we show the corresponding LSA results in Fig. 4(d-f). In this case, the dispersion relation has a single unstable branch with larger growth rates. With unattenuated $\mathbf{H}_{\nu\nu}$, the maximum growth rate in Model V shown in Fig. 4(d) is $\approx 8.4\times10^4$ km$^{-1}$, which is about seven times larger than that in Model II, and is nearly independent of the presence of the collisional and vacuum terms [Fig. 4(e)]. Moreover, due to the enhanced growth rate, even when an attenuation with $a_1 = 10^{-3}$ is taken, the shape of the dispersion relation branch remains hardly affected by the inclusion of collisional term and different sizes of vacuum terms as shown in Fig. 4(f).

The scalability of dispersion relation discussed above also holds for a wider radial range. Figure 5(a) shows the maximum growth rate $\mathrm{Im}(\Omega)$ as a function of $K_r$ and $r$ in Model II without attenuation for the case `c1v2`. Similaly, two separated unstable branches span over the whole radial region where there are ELN angular crossings. The maximum growth rate sits on the unstable branch with smaller (more negative) $K_r$ for all radii, with the corresponding $K_r$ value increasing from $K_r \simeq -1.5\times10^5$ km$^{-1}$ at $r = 38$ km to $K_r \simeq 0$ km$^{-1}$ at the outer boundary as the neutrino flux decreases. The other branch locates in the range of $0$ km$^{-1} \lesssim K_r \lesssim 2\times10^4$ km$^{-1}$. For the case with parameters of `c1v1` and with $a_1 = 10^{-3}$ shown in Fig. 5(b), one sees that the shapes of the unstable branches remain largely similar to those shown in Fig. 5(a), when both the $K_r$ range and the color scheme of $\mathrm{Im}(\Omega)$ are also rescaled by the same factor $a_1$. For the case `c1v2` with a larger $\delta m^2$ and with the attenuation parameter $a_1 = 10^{-3}$, Fig. 5(c) shows that the shape of the dispersion relation for the whole radial range is affected by the vacuum term. The values of $\mathrm{Im}(\Omega)$ in the lower branch are enhanced while those in the upper one are suppressed, consistent with the result discussed earlier at $r \approx 41.6$ km. Considering that this change is related to the attenuation of $\mathbf{H}_{\nu\nu}$ and does not appear in the original dispersion relation without attenuation, we mainly adopt $\delta m^2 = 8\times10^{-5}$ eV$^2$ in our simulation models to avoid this artifact.

Figure 5(d–f) compares the radial profiles of dispersion relation in Model V for the case without attenuation to the attenuated cases with `c1v1` and `c1v2`. For all cases, two unstable branches only appear at $r \simeq 37$ km where a very shallow angular crossing exists. Those two branches merge into one at larger radii. In the original case without attenuation in Fig. 5(d), the value of $K_r$ corresponding to the maximum growth rate increases from $K_r \simeq -2.5\times10^5$ km$^{-1}$ at $r = 37$ km to close to zero at the outer boundary. The other two cases in Fig. 5(e) and (f) also follow this trend with a factor of $10^3$ smaller in terms of $K_r$ and the growth rate, showing again very good scalability discussed earlier.

Quantitatively, we compare in Fig. 6(a) and (b) the maximum growth rates in Model II and V for cases with different attenuation factors as well as with different collision and vacuum terms. Once again, the values of $\mathrm{Im}(\Omega)_{\rm max}$ for the attenuated cases with $a_1 = 10^{-2}$ and $a_1 = 10^{-3}$ are multiplied by $10^2$ and $10^3$, respectively. For Model II, Fig. 6(a) clearly shows that a good scalability can be reached without attenuation or when the attenuated cases take the smaller $\delta m^2$ (case `c1v1`). When the larger $\delta m^2$ (case `c1v2`) is adopted for the attenuated case with $a_1 = 10^{-3}$, it artificially enhances the maximum growth rates for the whole region with FFI. For Model V shown in Fig. 6(b), because the angular crossing is deeper and the maximum growth rate is generally larger than in Model II, the impact due to the vacuum term is relatively smaller. Even for the case with the larger $\delta m^2$ (`c1v2`) and $a_1 = 10^{-3}$, the rescaled maximum growth rates closely follow the original ones, with deviations only up to $\lesssim 20\%$ at $r \gtrsim 60$ km.

Also shown in Fig. 6 are the maximum growth rates estimated by two previously proposed analytical formulas $\sqrt{2}G_F\sqrt{|I_-I_+|}$ and $\sqrt{2}G_F(2|I_-I_+|)/(|I_-|+|I_+|)$, which are proportional to the geometric and harmonic averages of $|I_-|$ and $|I_+|$, respectively [11, 134]. These estimated rates qualitatively agree with each other, as well as with the values computed by the LSA.

With all the discussions above, we conclude that the overall behavior of the dispersion relation in the region where the ELN angular crossings exist is predominantly determined by the FFI. The attenuated schemes can retrieve those structures of unstable modes to a good approximation by taking appropriate values of $\delta m^2$ for specific cases.

### B. Other type of flavor instability

For regions without angular crossings, it is curious to see whether other type of flavor instabilities such as the CFI or the slow type may exist. In Model IV where the angular crossing extends to more inner radius, the FFI essentially dominates all radii of concern. While for

[3] These extended nonzero growth rates may also originate from the spurious mode associated with insufficient number of grids. However, here we do not further clarify its exact origin, because the flavor conversion at this radius are dominated by the FFI.

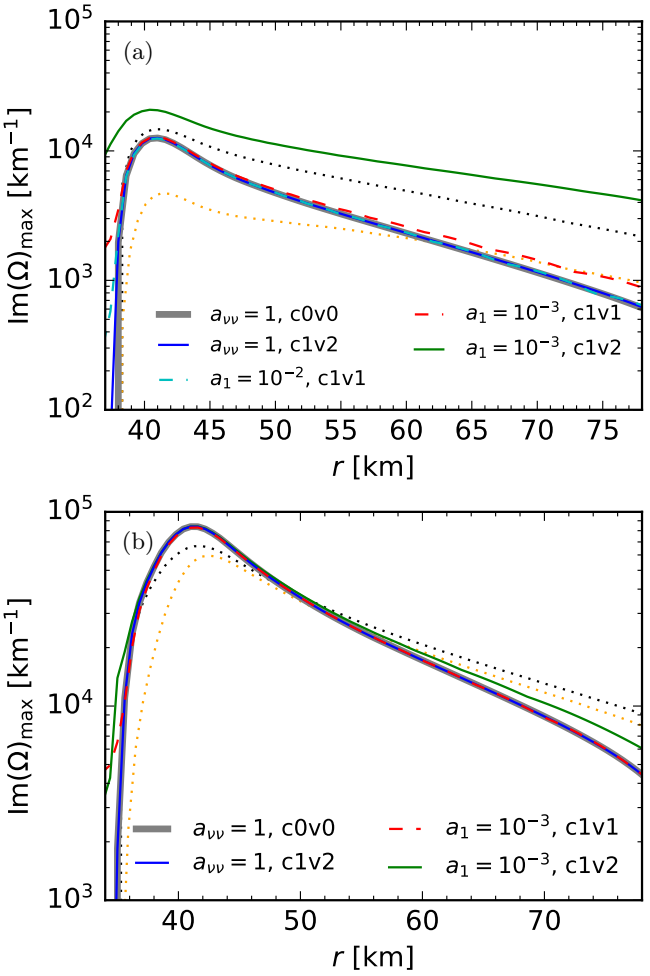



FIG. 6. Maximum growth rates $\mathrm{Im}(\Omega)_{\mathrm{max}}$ as functions of radius $r$ in Models II (a) and V (b). For the cases with $a_1 = 10^{-2}$ and $a_1 = 10^{-3}$, the values of $\mathrm{Im}(\Omega)_{\mathrm{max}}$ are multiplied by $10^2$ and $10^3$, respectively, for easier comparison. Black and orange dotted curves are obtained using the simple analytical formulas in Refs. [11] and [134], respectively.

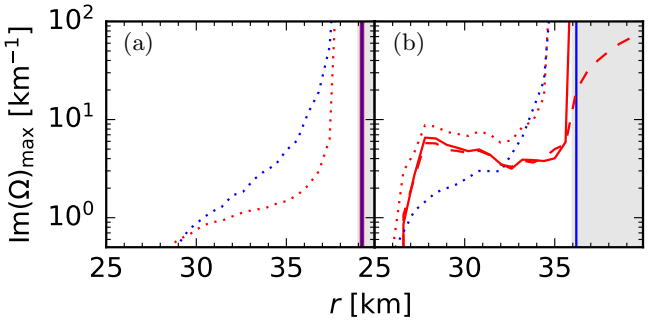



FIG. 7. Maximum growth rates $\mathrm{Im}(\Omega)_{\mathrm{max}}$ as functions of radius $r$ for in Models II (a) and V (b). The gray shaded areas indicate the radial regions where ELN angular crossings exist in the initial ELN distribution without oscillations. Blue and red curves represent results with parameters in the cases `c0v2` and `c1v0` respectively. Dotted and solid curves represent the cases using $a_{\nu\nu} = 1$ without and with excluding the spurious non-collective modes respectively. The dashed curve in (b) is with $a_1 = 10^{-3}$ and with all the spurious modes excluded.

other models, we find that CFI exists around the neutrinosphere in Model V with the strongest $Y_e$ attenuation leading to the dominance of $\bar{\nu}_e$'s over $\nu_e$'s, but it is suppressed in the $\nu_e$-dominated Models I–III. For Models VI and VII with FFI only appearing in limited range inside the neutrinosphere, their ELN angular profiles outside the neutrinosphere are similar to Model I and thus not discussed in detail below. We do not find any slow instabilities that are strong enough to be distinguished from spurious modes in any of the models.

As two representative instances, we show the maximum growth rates calculated from the LSA for Models II and V in Fig. 7(a) and (b) for $r < 40$ km by the dotted lines. Although in radii without angular crossings (non-shaded region), nonzero small growth rates are found for cases `c0v2` and `c1v0` in both Models II and V, some of them turn out to be spurious modes that contain non-collective features. When applying a practical selection rule detailed in Appendix C to preclude the spurious modes, none of those in Model II survive [see the solid lines in Fig. 7(a)]. In Model V, the unstable mode due to the nonzero vacuum term is also spurious, whereas the mode due to CFI persists [solid red line in Fig. 7(b)]. When the $\mathbf{H}_{\nu\nu}$ attenuation with $a_1 = 10^{-3}$ is applied, the growth rates due to CFI are nearly the same as the original case without attenuation in regions without ELN crossings [dashed red line in Fig. 7(b)]. Notice that this CFI mode also appears visibly around $r \simeq 30$ km over a wide range of $K_r$ in Fig. 5(e) and (f).

Some further comments regarding the spurious modes and the CFI are in order. First, the spurious modes in principle should also exist in simulations since the LSA results presented here are based on the same number of energy and angular grids as in simulations. However, we find that these noncollective spurious modes do not grow in the linear regime but are suppressed by the angular derivative in the advection term $(1 - v_r^2) r^{-1} \partial_{v_r}$. Thus, their presence does not affect the numerical outcome discussed in the next section.

Second, in contrast to our previous work [71] where CFI were found to be present in snapshots that are similar to the unattenuated Model I here, the CFI now only exists in Model V in which $\bar{\nu}_e$'s dominate over $\nu_e$'s due to the significant $Y_e$ attenuation. We find that the main distinction is related to the newly added NES process in this work. Without the NES process, the heavy-lepton neutrinos are thermalized less efficiently and have higher mean energy than the electron flavor neutrinos (see the comparison shown in Fig. 22 in Appendix B). Given that all flavors of neutrinos are still approximately isotropic at this radius, the distinct behaviors can be understood

by examining the quantity

$$\gamma_{\rm coll} = \frac{\int dE\, dv_r\, [C_{e\mu}(\varrho_{ee} - \varrho_{\mu\mu}) - \bar{C}_{e\mu}(\bar{\varrho}_{ee} - \bar{\varrho}_{\mu\mu})]}{\int dE\, dv_r\, (\varrho_{ee} - \varrho_{\mu\mu} - \bar{\varrho}_{ee} + \bar{\varrho}_{\mu\mu})}, \tag{16}$$

which is the imaginary part of the simple formula, Eq. (13) obtained in Ref. [73]. The denominator, the total ELN, is positive in Models I and II, as well as in all cases of the Ref. [71], but is negative in Model V. For the numerator, it is mainly dominated by the first term $C_{e\mu}(\varrho_{ee} - \varrho_{\mu\mu})$ due to the larger EA rate associated with $\nu_e$. In Ref. [71], the heavy-lepton neutrinos dominate over $\nu_e$'s at higher energy in the region where the CFI can develop. Weighted by the collisional rate $C_{e\mu}$, which is in general proportional to $E^2$, the high-energy part of the neutrinos contributes predominantly to the total integral of the numerator and leads to a negative $\gamma_{\rm coll}$. With a positive denominator, the associated CFI corresponds to the minus-type defined in Ref. [73], whose growth rate can be well approximated as $|\gamma_{\rm coll}|$ and as large as a few km$^{-1}$, mainly determined by the collisional rates of the corresponding high-energy part. On the other hand, in Models I and II considered here, the NES process thermalizes the spectra of heavy-lepton neutrinos so that the mean energy of $\nu_\mu$ is close to $\nu_e$. As a result, $\gamma_{\rm coll}$ is positive, which implies that the corresponding CFI would be the plus-type of Ref. [73] if it exists. However, unlike the minus-type solution, the CFI may not be assured for the plus-type. The actual growth rate can be less than $\gamma_{\rm coll}$ or even completely vanishes, depending on the neutrino spectral properties and the total ELN (see, e.g., Fig. 2 in Ref. [73]). In Model V, $\bar{\nu}_e$'s dominate over $\nu_e$'s with a negative total ELN. With the numerator being positive as in Models I and II, $\gamma_{\rm coll}$ is negative. As a result, the CFI is of the minus-type, for which $|\gamma_{\rm coll}|$ is a good approximation to indicate the corresponding growth rate.

The above discussion suggests the following intriguing fact. Although the EA processes in Table I only involve the electron flavor neutrinos from the classical viewpoint, it can damp the coherent state between $\nu_e$ and $\nu_\mu$ in $\nu$QKE so that the CFI does depend on the properties of heavy-lepton neutrinos. We also note that although CFI is suppressed in our models dominated by $\nu_e$'s, it does not imply that CFI is generally not present in all $\nu_e$-dominated conditions. Multidimensional CCSN simulations with convection may potentially result in conditions favoring the minus-type CFI when the dynamical evolving heavy-lepton neutrinos are less efficiently thermalized by the NES process.

### C. Spectrogram

In this subsection we perform the analysis based on the simulation results using the spectrogram defined in Sec. III C when flavor conversions reach the nonlinear regime, and compare the structures to the dispersion relation obtained by the LSA in the previous subsections. Note that here the model parameters are given in Table II, and the $\mathbf{H}_{\nu\nu}$ attenuation is applied to all the models.

Before presenting the spectrogram analysis, we first illustrate the evolution of FFC from the onset of FFI in Fig. 8 for Model II as an example. Since the maximum growth rate of FFI, $\mathrm{Im}(\Omega)_{\rm max}$, in this model peaks at $r \simeq 42$ km, FFI grows faster to reach the nonlinear regime around that radius first. The left panel of Fig. 8 shows that $s_{e\mu}$, the dimensionless ratio representing the off-diagonal flavor mixing introduced in Sec. III A, reaches the nonlinear regime over a wide range from $r \simeq 40$ km to $r \simeq 58$ km by the time of $t = 32$ $\mu$s. The range over which the nonlinear FFC takes place further widens to $r \simeq 70$ km at $t = 64$ $\mu$s and extends to the outer boundary at a much later time $t = 160$ $\mu$s.

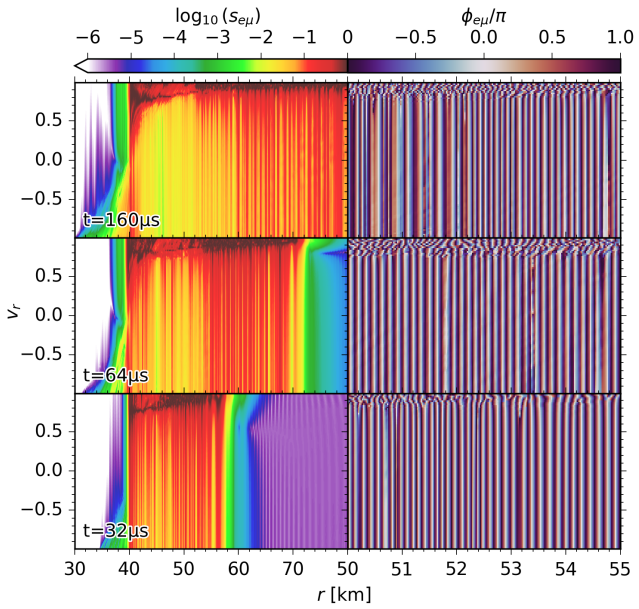



FIG. 8. Evolution of the dimensionless ratio of off-diagonal mixing $s_{e\mu}$ in 30 km $< r <$ 80 km (left panel), and its associated phase angle $\phi_{e\mu}$ within a smaller radial range 50 km $< r <$ 55 km (right panel) for Model II.

On the right panel in Fig. 8, we also show the radial shape of the complex quantum phase $\phi_{e\mu}$ defined in Sec. III A in 50 km $< r <$ 55 km to illustrate the coherent feature of FFC. The phase angle $\phi_{e\mu}$ not only oscillates rapidly in time during FFCs, but also develops coherent spatial-dependent oscillatory patterns. Moreover, the pattern is highly correlated for a wide range of $v_r$ due to the collective nature of FFC. The typical oscillation length scale of the synchronized phase angle is $\lesssim 0.2$ km, broadly consistent with the scale determined by the maximally unstable $K_r$ shown in Fig. 5(b) in the same radial range. This simple comparison suggests that the characteristic length scale determined by the dispersion relation in the linear regime correlates with the coherent length scale at late times when FFC are in the

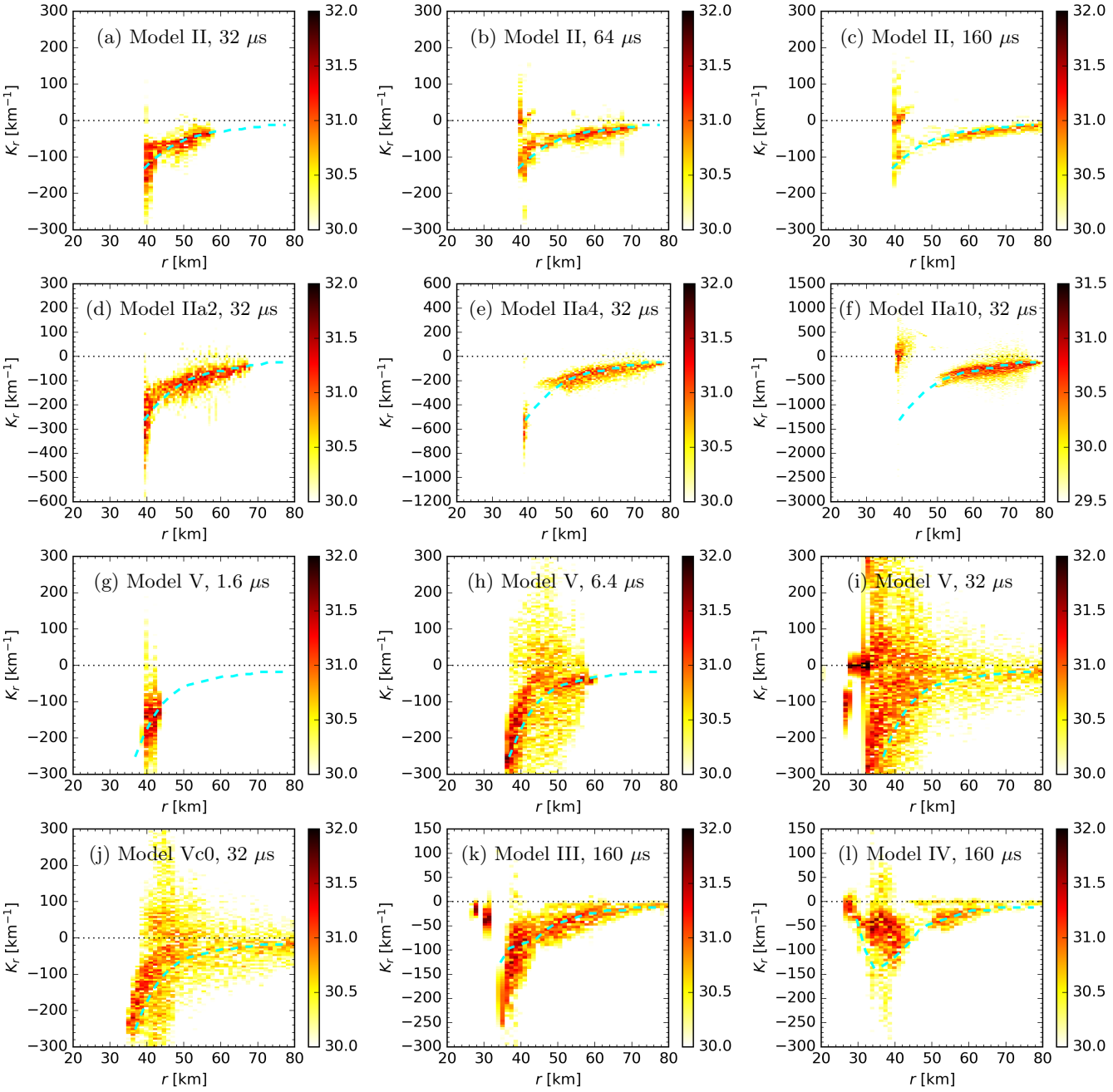



FIG. 9. Spectrograms with respect to $K_r$ and $r$ for different models at different times. The color bar is for the Fourier transformed magnitude $\log_{10}|\mathcal{F}_{r,K_r}[\mathrm{cm}^{-3}]|$. The cyan dashed curve shows the $K_r$ with maximum growth rates calculated from LSA.

nonlinear or asymptotic regime [33, 135], which will be further discussed below with the spectrograms. We also note that this spatial variation can lead to a substantial contribution to the flavor evolution equation through the advection term $v_r\partial_r$, and thus participate in determining the outcome of flavor evolution.

The spectrograms shown in Fig. 9(a–c) compare the Fourier transformed magnitudes $\mathcal{F}_{r,K_r}$ at the same three time frames in Fig. 8 with the curve of $K_r$ corresponding to the maximum growth rate from LSA. At $t = 32$ $\mu$s, the dominant wave number $K_r$ of the off-diagonal flavor mixing in $\mathcal{F}_{r,K_r}$ generally aligns well with the maximally unstable $K_r$ values derived with LSA. Relatively small differences of $\simeq$ 20–30 km$^{-1}$ that exist are mainly related to the change of ELN distribution due to flavor conversions in simulation, while the LSA results are based on unoscillated neutrino profiles. We note that near $r = 40$ km, the range of dominant $K_r$ widens over time to encloses the homogeneous modes with $K_r \approx 0$ km$^{-1}$. The amplitude of the off-diagonal flavor mixing between $r \approx 42$ km

and 50 km decreases at a later time $t = 160$ $\mu$s, with $\mathcal{F}_{r,K_r} \lesssim 10^{30}$ cm$^{-3}$. For $r \gtrsim 50$ km, similar to the case with a shallow angular crossing studied under the periodic boundary [36], our result suggests that a steady wave-like pattern along the radial direction can also survive in a more realistic setting without adopting the periodic boundary condition.

The radial-dependent coherent pattern of the flavor mixing remains qualitatively similar when different attenuation factors are used. Figure 9(d–f) compares the spectrograms for Models IIa2, IIa4, and IIa10 at the same time $t = 32$ $\mu$s. The features of the spectrograms in all three models again agree well with the results derived from the LSA. The patterns also show a good scalability as discussed before. Because models with less attenuation evolve faster, they reach the asymptotic pattern at earlier times as shown in Fig. 9(e) and (f).

In Model V, because the maximum growth rates are generally larger than those in Model II, flavor conversions occur faster. Figure 9(g) shows that a significant flavor mixing with $\mathcal{F}_{r,K_r} > 10^{31}$ cm$^{-3}$ is attained for $r \approx 42$ km at $t = 1.6$ $\mu$s. At $t = 6.4$ $\mu$s shown in Fig. 9(h), the FFC in the region from $\approx 36$ km to 60 km evolves to the nonlinear regime. The dominant wave numbers in $r \lesssim 40$ km and $r \gtrsim 50$ km still qualitatively follow the curve of maximum-growth-rate $K_r$ from the LSA. However, in the region between 40 and 50 km, the spectrogram covers a much wider wave number modes, spreading from $K_r \simeq -300$ km$^{-1}$ to 300 km$^{-1}$. This becomes even more evident at $t = 32$ $\mu$s shown in Fig. 9(i). Also at $t = 32$ $\mu$s, the emergence of CFI leads to flavor mixing between $r \approx 27$ km to 33 km with the unstable modes centered on $K_r \simeq 0$ km$^{-1}$.

We use Model Vc0 to disentangle the effect due to the interplay between the CFI and the FFI. Figure 9(j) shows that the asymptotic spectrogram of Model Vc0 at the same time $t = 32$ $\mu$s generally follows the trend of $K_r$ from LSA. The dominant $K_r$ of the spectrogram spans from $K_r \simeq -250$ km$^{-1}$ at $r \approx 36$ km to $K_r \simeq -30$ km$^{-1}$ at $r \approx 46$ km. Between $r \simeq 40$ km and 60 km, the range of $K_r$ in the spectrogram becomes wider, consistent with similar features shown in Fig 9(h) and (i) in the same radial range.

Finally, we also show in Fig. 9(k) and (l) the asymptotic spectrograms in Models III and IV, respectively. For both models, the radial shape of the dominant $K_r$ modes again agree qualitatively with the LSA results, although quantitative differences do exist. In any case, our spectrogram analysis reveals that asymmetric patterns with respect to $K_r$ generally exist, which suggests that the outcome of FFCs and the flavor waves are not necessarily homogeneous. The overall development of small-scale structures reported here also qualitatively agrees with the findings in local simulations adopting periodic boundary conditions.

Besides the off-diagonal flavor mixing, the diagonal elements of the density matrices associated with the neutrino number density also contain the same small-scale structures. In Fig. 10, we show the neutrino number densities of Models III and IIIa2 between $r = 38$ km and 40 km for illustration. Clearly, oscillatory patterns exist in general for all three different neutrino species. In Model III, its typical oscillatory wavelength is $\simeq 0.07$ km, corresponding to a $K_r \simeq 90$ km$^{-1}$, which is consistent with the dominant mode of spectrogram in Fig. 9(k). With a less attenuation in Model IIIa2, the oscillatory length scale become correspondingly smaller. The presence of small-scale structure may lead to implications associated with the refractive energy and the potential self-thermalization of the neutrino gas [136, 137].

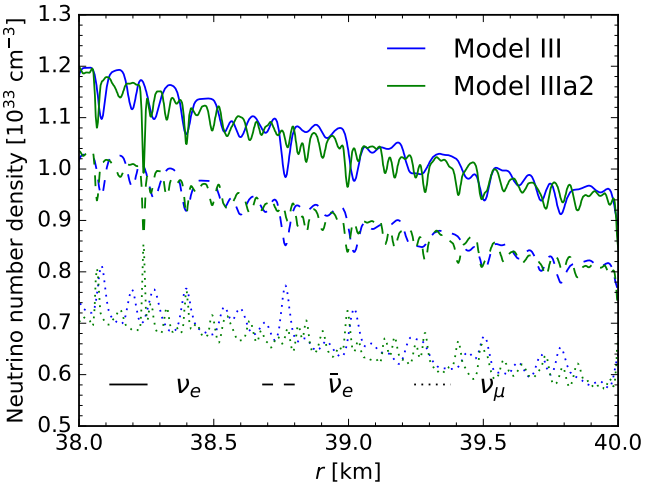



FIG. 10. Radial profiles of neutrino number densities between $r = 38$ km and 40 km at $t = 64$ $\mu$s in Models III and IIIa2.

## V. EVOLUTION OF FAST FLAVOR CONVERSIONS

In this section, we will present more details regarding the evolution of the FFC and its impact on key properties of neutrinos. Figures 11 and 12 show the neutrino number density and mean energy of different species for Models II–VII at different times. For the number densities, because the initial values of $\nu_\mu$ and $\bar{\nu}_\mu$ are rather close to each other, and given that the fast pairwise flavor conversion produces equal amount of $\nu_\mu$ and $\bar{\nu}_\mu$, only the $\nu_\mu$ ones are shown in Fig. 11. For the average energies, all fours species are shown explicitly in Fig. 12. Also noted is that the values shown in these two figures are coarse-grained ones, computed by averaging over a size of 0.6 km from the original simulation outputs, for the purpose of capturing the global trend. Below, we will first compare the similarities and differences between those models. Different adopted schemes such as the number of neutrino flavors, the amount of attenuation on $\mathbf{H}_{\nu\nu}$, and the size of the vacuum term will then be discussed.

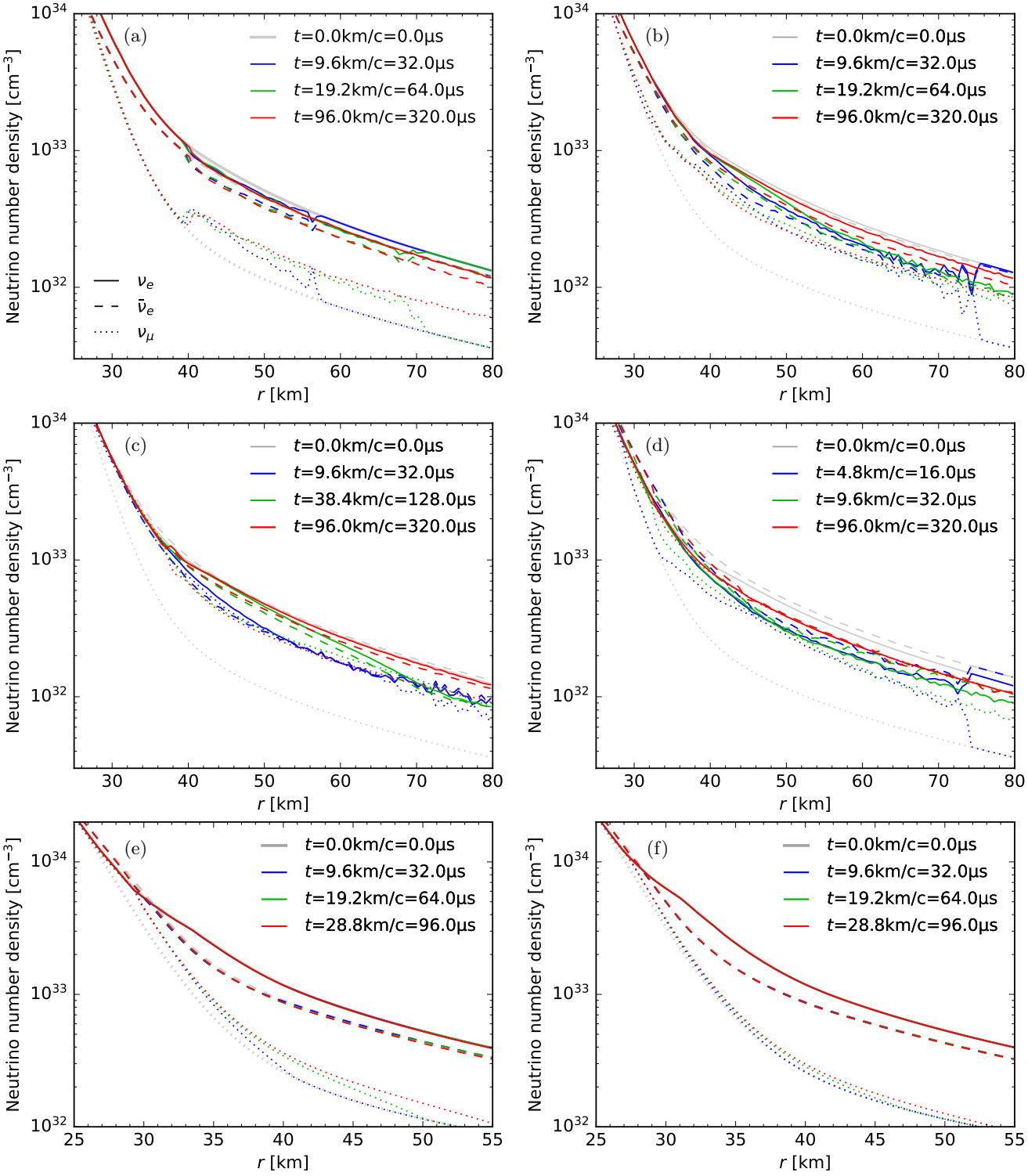



FIG. 11. Radial profiles of neutrino number densities for $\nu_e$, $\bar{\nu}_e$, and $\nu_\mu$ in Models II–VII (a–f).

### A. Model II

We start with Model II in which the angular crossing is the shallowest. Prior to the occurrence of the FFC, the number densities of all neutrino species decrease smoothly with radius, following a hierarchy of $n_{\nu_e} \gtrsim n_{\bar{\nu}_e} > n_{\nu_\mu}$ in the radial range shown in Fig. 11(a). Due to the FFI discussed earlier, the FFC takes place and leads to the pairwise conversion from $\nu_e$ and $\bar{\nu}_e$ to muon flavor neutrinos. At $t = 32$ $\mu$s, the region between

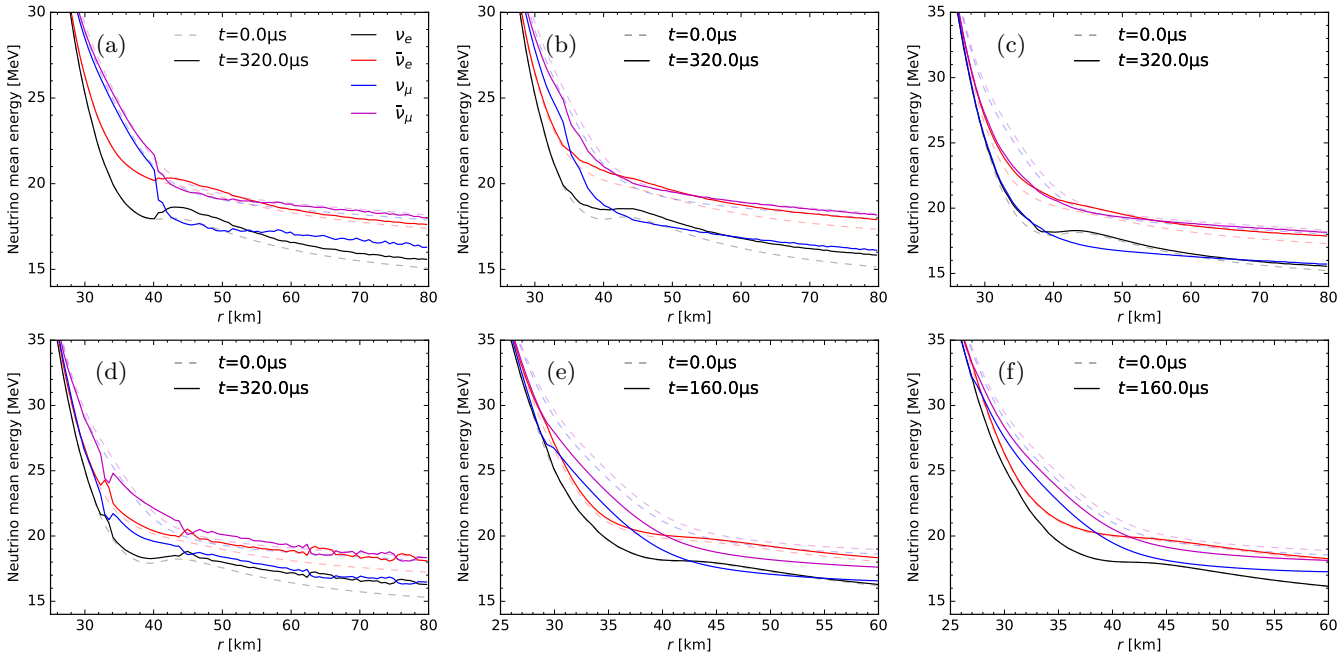



FIG. 12. Radial profiles of neutrino mean energies for $\nu_e$, $\bar{\nu}_e$, $\nu_\mu$, and $\bar{\nu}_\mu$ in Models II–VII (a–f).

$r \approx 40$ km and $r \approx 57$ km is influenced by FFC. The impact of FFC extends to $r \approx 70$ km at $t = 64$ $\mu$s and eventually reaches the outer boundary at a later time. Notice that there are "spiky" features just behind the outermost range affected by FFC, e.g., at $r \approx 56$ km at $t = 32$ $\mu$s and $r \approx 68$ km at $t = 64$ $\mu$s. Those are consistent with what reported in local simulations adopting periodic boundaries where a larger amount of flavor conversions can exist as a transient phenomenon before the system settles down to the asymptotic state.

In Model II, $n_{\nu_e}$ and $n_{\bar{\nu}_e}$ in the region affected by FFC are reduced by $\simeq$ 10% to 20%, while $n_{\nu_\mu}$ increases by $\simeq$ 50%–70%. Once the asymptotic state in a certain radial range is reached, the number densities do not vary much anymore, illustrated by comparing the curves at $t = 64$ $\mu$s and at $t = 320$ $\mu$s in 40 km$\lesssim r \lesssim$ 68 km. Clearly, flavor equilibration is not achieved in Model II. This is because the FFC here is restricted within a limited angular range (see e.g., Fig. 8 and later in Fig. 14).

For the mean energies of $\nu_\mu$ and $\bar{\nu}_\mu$, differently from the similar evolution of $n_{\nu_\mu}$ and $n_{\bar{\nu}_\mu}$, they can be affected very differently by FFCs. Figure 12(a) shows that $\langle E_{\nu_\mu} \rangle$ is reduced substantially to $\approx$ 16–17 MeV for $r > 40$ km, while $\langle E_{\bar{\nu}_\mu} \rangle$ remains around 18–20 MeV. For $\nu_e$ and $\bar{\nu}_e$, their mean energies increase by less than 1 MeV.

### B. Models III and IV

For Models III and IV that possess deeper angular crossings in terms of $v_r$, the ELN crossings also appear at radii deeper inside the neutrinospheres (see Fig. 2). We show in Fig. 11(b) and (c) the evolution of the neutrino number densities. At $t = 32$ $\mu$s, FFCs take place between $r \approx 32$ km to $r \approx 76$ km in Model III, and from $r \approx 28$ km till the outer boundary in Model IV. The deeper angular crossings in both models lead to an increased amount of the converted neutrinos from the electron to the muon flavor than that in Model II. Particularly, the nearly equal number densities in 50 km$\lesssim r \lesssim$ 70 km of all different flavors indicate that flavor equilibration is approximately achieved in that radial range in both models. The zig-zag pattern around $r = 70$–72 km in Model III is associated with the transient phenomenon from the linear to nonlinear regime aforementioned in Model II.

Differently from Model II, the existence of the ELN angular crossings at deeper radii inside the neutrino spheres of $\nu_e$ enables an intriguing feedback effect between the FFC and the collisional EA processes after the initial phase of FFC in both models. In Model III, the number densities of $\nu_e$ and $\bar{\nu}_e$ around $r \simeq 45$ km at $t = 64$ $\mu$s increase by a similar amount from those at $t = 32$ $\mu$s as shown in Fig. 11(b). This is because the EA collisions around the neutrinospheres of both $\nu_e$ and $\bar{\nu}_e$ act to restore their number densities to the stationary values prior to the occurrence of FFC. The increase of $n_{\nu_e}$ and $n_{\bar{\nu}_e}$ continues and also affects regions outside the neutrinosphere. For instance, at $t = 160$ $\mu$s, $n_{\nu_e}$ and $n_{\bar{\nu}_e}$ become closer to their initial values for $r \gtrsim 40$ km. Similar behaviors also appear in Model IV shown in Fig. 11(c). Nearly flavor equilibration is achieved at $t = 32$ $\mu$s initially, followed by the increase of $n_{\nu_e}$ and $n_{\bar{\nu}_e}$ locally in 40 km$\lesssim r \lesssim$ 66 km at $t = 128$ $\mu$s, and later at $t = 320$ $\mu$s the restoration of $n_{\nu_e}$ and $n_{\bar{\nu}_e}$ back to the initial values for $r \gtrsim 40$ km. The restoration is stronger than in Model III, particularly for $n_{\nu_e}$, which almost overlaps with the

unoscillated level. We note that because this feedback involves mostly the collisional processes and advection, the associated timescale of the number density increase is not related to the conversion rate of FFC. For instance, in Model IV, the range affected by this feedback effect extends from $r \simeq 40$ km at $t = 32$ $\mu$s to $r \approx 66$ km at $t = 128$ $\mu$s, consistent with the corresponding neutrino propagation time, which is a few tens of km/c.

In addition to the restoration of $\nu_e$ and $\bar{\nu}_e$ number densities discussed above, the feedback effect contains an interesting twist in reshaping the ELN angular distribution, which then leads to different local ELN angular distributions compared to those after the prompt flavor conversions. One major difference is related to the conservation of the neutrino lepton number. In the prompt FFCs, in addition to the elimination of the ELN angular crossings, the neutrino and antineutrino lepton numbers, as well as the total ELN are locally conserved to a good approximation. However, all these three numbers are no longer conserved quantities on a longer timescale when the collision and neutrino diffusion play a role. Although the FFC still works to erase the ELN angular crossings, the collisional weak processes such as EA can change those lepton numbers as well as the total ELN. Consequently, the angular distributions of the ELN and each neutrino species keep evolving on timescales determined by the collision rates, while the FFC functions to ensure that no new ELN angular crossings appear during the evolution. The system can then settles down to a quasi-steady state determined by the FFC, collisions, and advection all together.

To illustrate this effect, we first show in Fig. 13 the neutrino angular distributions for different times at $r \approx 36$ km, which is inside the $\nu_e$-sphere in both Model III and IV. In Model III, the initial ELN distribution transitions from positive to negative at $v_{r,c} \approx 0.37$ [Fig. 13(a)]. With the condition $|I_+| > |I_-|$, the prompt FFC between the electron and muon flavors erases the ELN crossing by making it positive in the entire range of $v_r$, accompanied with significant flavor swap for $v_r > 0$ at $t = 3.2$ $\mu$s [Fig. 13(b)]. Since this radius is inside the neutrinosphere, the EA processes of $\nu_e$ and $\bar{\nu}_e$ with matter tend to restore their angular distributions back to the initial state prior to FFC. Due to the higher opacity for $\nu_e$'s than $\bar{\nu}_e$'s, the angular distribution of $\nu_e$'s increases faster, which leads to a more positive ELN distributions at $t = 16$ $\mu$s [Fig. 13(a)]. On a longer timescale, the restoration of $\nu_e$'s slows down as its angular distribution becomes closer to the initial state. Due to the restoration of $\bar{\nu}_e$'s, the ELN distribution starts to decrease, which potentially leads to a new angular crossing. However, the FFC can smear any newly formed angular crossing and keep the ELN angular distribution positive during this process. Finally, the neutrino distributions settle into a quasi-steady state at $t \approx 64$ $\mu$s where the ELN angular distribution is nearly flat for $v_r \gtrsim 0.5$ [Fig. 13(a)]. Meanwhile, the $\nu_e$ and $\bar{\nu}_e$ distributions are restored to values closer to their initial states for $v_r \lesssim 0.5$ [Fig. 13(b)].

For Model IV, the neutrino distribution is dominated by $\bar{\nu}_e$'s with $|I_+| < |I_-|$ at the same radius $r \simeq 36$ km with an initial crossing at $v_r \simeq -0.2$. Thus, flavor equilibration takes place mainly for negative $v_r$ at $t = 3.2$ $\mu$s [Fig. 13(c)], and the ELN angular distribution becomes essentially negative for all $v_r$'s. However, the faster EA rates for $\nu_e$'s increase the ELN distribution and turn it into positive values at $t = 16$ $\mu$s [Fig. 13(c)]. Later, the effect of slower $\bar{\nu}_e$ rates take effects, and lowers the values of ELN distribution for a major $v_r$ range. At $t = 64$ $\mu$s, the distributions also settle into an approximately steady state, with nearly flat ELN distribution around zero for $v_r \gtrsim 0$ [Fig. 13(c)], as well as the $\langle \varrho_{\nu_e} \rangle_E \simeq \langle \varrho_{\bar{\nu}_e} \rangle_E$ and $\langle \varrho_{\nu_\mu} \rangle_E \simeq \langle \varrho_{\bar{\nu}_\mu} \rangle_E$ [Fig. 13(d)].

The collisional feedback effect inside the neutrinosphere also affects the flavor evolution in the outer part. Figure 14(c–f) shows that in both Model III and IV, the prompt FFC at $r = 50$ km and $t = 32$ $\mu$s turns the ELN distribution to the same sign, with significant amount of $\nu_e$'s and $\bar{\nu}_e$'s being converted. However, as neutrinos continuously stream out, what are observed at $r = 50$ km at later times reflect the state of neutrinos escaping from inside the neutrinosphere at earlier times. At $t = 160$ $\mu$s, the angular distributions are restored back to $\simeq 80\%$–$90\%$ of the initial values for $\nu_e$'s and $\bar{\nu}_e$'s in $v_r \gtrsim 0.8$ in Model III, or even beyond for $\nu_e$'s in $v_r \gtrsim 0.9$ in Model IV. The asymptotic ELN distributions are almost zero for $v_r \gtrsim 0.7$ in both Models, showing very different shapes from the prompt FFC.

For comparison, we also show in Fig. 14(a–b) the corresponding distributions in Model II. From $t = 32$ $\mu$s to $t = 160$ $\mu$s, the angular distributions of $\nu_e$'s and $\bar{\nu}_e$'s are only slightly changed. This is because in Model II, the FFCs only start at $r \approx 40$ km located just outside the $\nu_e$-sphere. No significant amount of $\nu_e$'s and $\bar{\nu}_e$'s are produced additionally from within the neutrinosphere.

In spite of the intriguing feedback effect on the neutrino distributions discussed above, the main impact of FFC on the radial profiles of the neutrino mean energies in Model III and IV are relatively similar to that in Model II, as shown in Fig. 12(b) and (c). FFC slightly enhances $\langle E_{\nu_e} \rangle$ and $\langle E_{\nu_e} \rangle$ by $\lesssim 1$ MeV in all radii, while it largely reduces $\langle E_{\nu_\mu} \rangle$ by $\simeq 2$ MeV in regions for $r \gtrsim 35$ km. One visible difference from Model II here is that for $r \lesssim 35$ km, both $\langle E_{\nu_\mu} \rangle$ and $\langle E_{\bar{\nu}_\mu} \rangle$ are reduced due to the backward-propagating neutrinos from outer radii.

### C. Model V

Model V is special because of the coexistence of CFI and FFI. Before discussing the combined effect originated from both flavor instabilities, we first discuss the results in the corresponding Model Vc0 where the the CFI is switched off. Figure 15 shows the radial profiles of the neutrino number densities and their mean energies in Model Vc0. The impact of FFCs here is qualitatively similar to those discussed earlier. At $t = 32$ $\mu$s, the

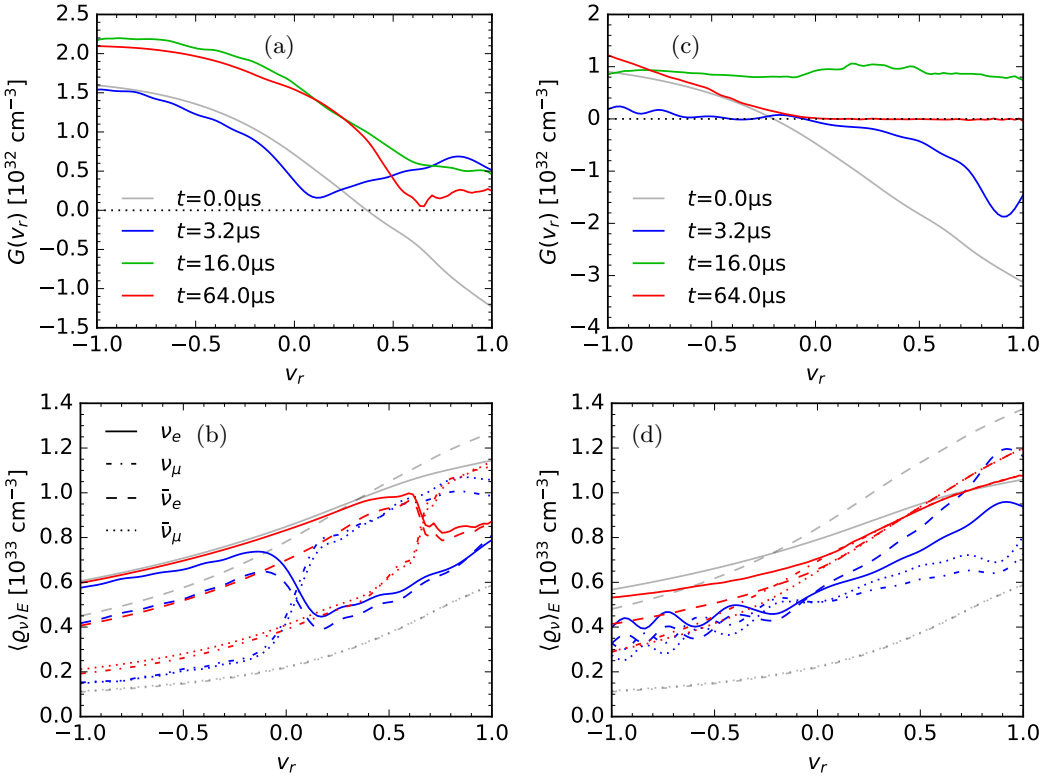



FIG. 13. Evolution of the angular distributions of the ELN, $G(v_r)$ (upper panels), and that of each neutrino species at $r \approx 36$ km (lower panels) in Models III (a–b) and IV (c–d).

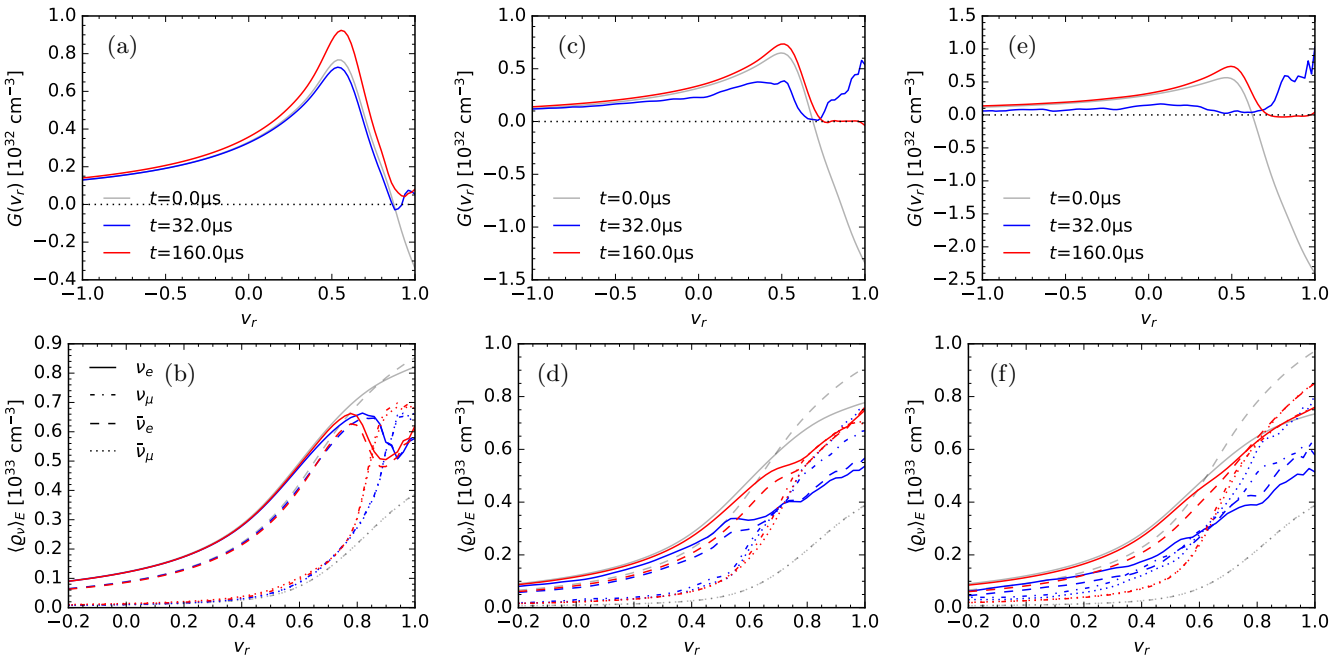



FIG. 14. Evolution of the angular distributions of the ELN, $G(v_r)$ (upper panels), and that of each neutrino species at $r \approx 50$ km (lower panels) in Models II (a–b), III (c–d), and IV (e–f). Note that $\langle \varrho_\nu \rangle_E$ are shown in $-0.2 \leq v_r \leq 1$ because most of neutrinos propagate outwards.

prompt FFC affects the radial range from $r \approx 35$ km to the outer boundary and leads to approximate flavor equilibration in 40 km$\lesssim r \lesssim$ 54 km where the deepness ratio of the initial ELN, rewritten as $|I_+/I_-|$, is larger (see

Fig. 3). As neutrinos continue propagating outwards, their flavor content is replaced by the emission and oscillation properties close to the neutrinosphere. Together with the feedback mechanism related to the EA processes, the number densities of $\nu_e$ and $\bar{\nu}_e$ are also enhanced for $r \gtrsim 40$ km at $t = 320$ $\mu$s as in Models III and IV. For $r \gtrsim 50$ km, the mean energies of $\nu_\mu$ and $\bar{\nu}_\mu$ become much closer to those of $\nu_e$ and $\bar{\nu}_e$, respectively. Because the angular crossings for radii inside the $\nu_e$ neutrinosphere in Model Vc0 is located at negative $v_r$, the FFC affects more the backward propagating $\nu_\mu$'s in that region, which leads to a minor enhancement of $n_{\nu_\mu}$ near $r = 32$ km as well as moderate reductions of the mean energies for heavy-lepton neutrinos there.

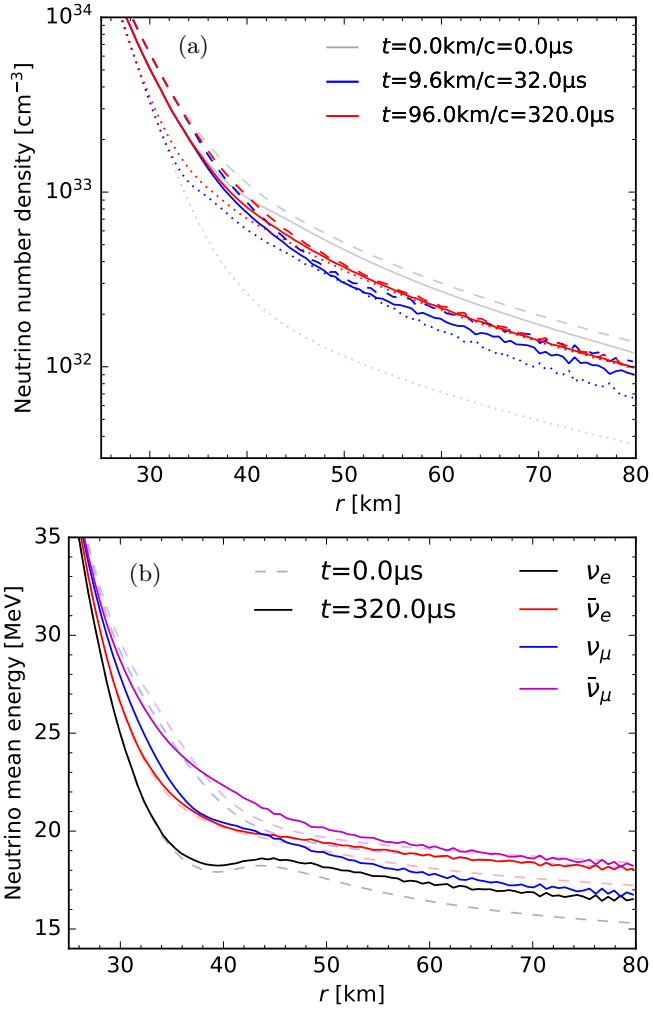



FIG. 15. Radial profiles of neutrino number densities (a) and mean energies (b) in Model Vc0.

For Model V where the collisional flavor conversions is also present, Fig. 11(d) shows that in addition to the same FFCs in $r \gtrsim 35$ km, the CFI also leads to flavor transformations and thus the enhanced $n_{\nu_\mu}$ in $r \simeq 28$–35 km inside the neutrinosphere. As a result, $n_{\nu_\mu}$ becomes almost the same as $n_{\nu_e}$ already at $t = 32$ $\mu$s and remains till $t = 320$ $\mu$s. For the mean energies, once again, the muon flavor neutrinos become closer to those of electron flavor as shown in Fig. 12(f).

We notice that in both Models V and Vc0 at $t = 320$ $\mu$s, the neutrino number densities in regions for $r \gtrsim 50$ km become $n_{\nu_e} \simeq n_{\bar{\nu}_e}$, different from the initial hierarchy of $n_{\nu_e} < n_{\bar{\nu}_e}$. Because the flavor content is still dominated by $\bar{\nu}_e$'s after the prompt conversions, this effect cannot be attributed to the pairwise FFC. It suggests that the collisional feedback effect inside the neutrinosphere leads to a larger net change of flavor content in $\bar{\nu}_e$'s than $\nu_e$'s.

### D. Models VI and VII

For Models VI and VII, the initial ELN angular crossings only reside in a limited radial range at 29 km $\lesssim r \lesssim$ 33 km and 27 km $\lesssim r \lesssim$ 29 km (see Fig. 2), respectively. Figure 11(e) and (f) shows that the FFC affects mostly the number density of heavy-lepton flavors outside the regions where the crossings exist initially. For $\nu_e$ and $\bar{\nu}_e$, their number densities remain nearly unchanged for all radii due to the following reasons. First of all, since these crossings are inside the neutrinospheres, the neutrino-matter interaction outside the FFI region plays an important role in resetting the properties of $\nu_e$ and $\bar{\nu}_e$. Secondly, as the neutrino angular distributions are more isotropic inside the neutrinosphere, neutrinos not affected by FFCs can also diffuse back into the region of FFC to mitigate the impact of flavor conversions. Moreover, even inside the region of FFC, the dynamical feedback from the collisions mentioned above in Models III and IV also acts to restore the distributions of $\nu_e$ and $\bar{\nu}_e$ back to the level close to those before oscillations. For the heavy-lepton neutrinos, their spectra at inner radii become closer to those of $\nu_e$ and $\bar{\nu}_e$ when the system reaches the quasi-steady state, which limits further flavor conversions. Since the FFI region in Model VI is closer to the neutrinosphere than in Model VII, the overall impact of flavor conversions is larger in Model VI.

Similar behaviors are found in terms of neutrino mean energies as well. In both models, the mean energy for heavy-lepton neutrinos are reduced significantly, while the mean energies of $\nu_e$ and $\bar{\nu}_e$ only increase very little. The impact in Model VI is larger than in Model VII.

### E. Two flavors v.s. three flavors

To illustrate the effects of adopting schemes with different numbers of neutrino flavors in the simulations, we choose Models II and IV as the representative examples for cases without and with the dynamical collisional feedback. The results from Models II and IV are compared with Models IIf3 and IVf3 in Fig. 16. For Models II and IIf3, the radial profile of the number density ratios $n_\nu(t = 320\ \mu\text{s})/n_\nu(t = 0)$ for all flavors, including $\nu_\tau$ and

$\bar{\nu}_e$ in the three-flavor scenario, are shown in Fig. 16(a). Note that we use different scales for the electron flavors (left axis) and the heavy-lepton flavors (right axis) for better readability. For Models IV and IVf3, the number density ratios at $t = 32$ $\mu$s and 320 $\mu$s are shown in Fig. 16(b) and (c), respectively, to illustrate the impact after the prompt FFC and that at the asymptotic state after the collisional feedback takes place.

These plots show that with three flavors, the enhancement of $\nu_\mu$ and $\bar{\nu}_\mu$ number densities is reduced, due to the additional flavor conversion channel from electron to tauon flavors. The enhancement in Model IIf3 is $\simeq 50\%$ of that in Model II. For Models IVf3 and Model IV, the enhancement of heavy-lepton flavor number densities in Model IVf3 is reduced by less than 50%, particularly at $t = 320$ $\mu$s. This implies that the flavor conversions result in a larger amount of the total heavy-lepton flavor number density in Model IVf3 than in Model IV.

For $\nu_e$ and $\bar{\nu}_e$, their number densities are generally more reduced by flavor conversions, while the amount of reduction in Models IIf3 and IVf3 from the corresponding two-flavor cases differs. In Model IIf3, the reduction is relatively minor and only appears at $r \gtrsim 50$ km. In Model IVf3, the reduction at $t = 32$ $\mu$s is much more significant than in Model IIf3, reflecting the shift of $\nu_e$ and $\bar{\nu}_e$ survival probability from $\simeq 1/2$ to $\simeq 1/3$ when reaching flavor equilibration (see, e.g. [33]). At $t = 320$ $\mu$s, although the $\nu_e$'s and $\bar{\nu}_e$'s number densities are restored back to values closer to those prior to FFCs due to the collisional feedback effect, the three-flavor case still shows more reduction than the two-flavor case. Overall, including the tauon flavor allows more flavor conversion of $\nu_e$'s and $\bar{\nu}_e$'s, which is consistent with the observed increase of the total heavy-lepton flavor content.

### F. Attenuation on the neutrino self-induced term

We discussed in Sec. IV the impact of $\mathbf{H}_{\nu\nu}$ attenuation on the dispersion relation and the small-scale structure. In this subsection, we examine its impact on the flavor evolution in macroscopic scale. Figure 17(a)–(c) shows the neutrino number density ratios in Models II–IV at $t = 64$ $\mu$s with different attenuation factors $a_1$. At this time, only the case in Model II with $a_1 = 10^{-3}$ shows substantial differences in the number density ratios at $r \gtrsim 70$ km. The reason is that the evolution in Model II is mainly determined by the prompt FFCs. Thus, cases with more attenuation (smaller $a_1$) evolve slower than those with less attenuation. Another slight difference is that the starting radius of FFCs with less attenuation is slightly shifted leftwards from $r \approx 39$ km to $\approx 38$ km.

For Models III and IV, the number density ratios in Fig. 17(b) and (c) exhibit a transition at $r \simeq 50$ km, within which the impact of collisional feedback appears. Since the timescale of the collisional feedback effect does not depend on the attenuation factor, provided that the FFC timescale remains short enough, the results shown in these two panels are therefore hardly affected by the choice of different attenuation factors.

### G. Vacuum term

We also discussed in Sec. IV that the general trend of the full dispersion relation is almost independent of the vacuum term when no $\mathbf{H}_{\nu\nu}$ attenuation is applied. When taking $a_1 = 10^{-3}$, the dispersion relation in Model II, which has the smallest instability growth rate, can be affected by the adopted value of $\delta m^2$, while the other models such as IV and V remain nearly unaffected. Here, we examine whether the simulation results at late times are affected by the strength of the vacuum term by comparing Model IV to Model IVv2 at $t = 64$ $\mu$s in Fig. 18. We find that taking a different $\delta m^2$ indeed only results in negligible effects, which is consistent with the findings reported in Refs. [66, 128, 138]. For both the radial regions affected by the prompt flavor conversion and that affected by the collisions feedback, a larger $\delta m^2$ only leads to insignificant deviations.

## VI. EFFECTS AND OBSERVABLES

After discussing the evolution of FFCs, we examine in this section their implications and feedback to key quantities relevant to CCSN physics, including the net neutrino heating rates and the equilibrium electron fraction. In addition, the impact of FFCs on the free-streaming neutrino energy spectra, which are important inputs for the expected supernova neutrino signals, will be discussed.

### A. Heating-cooling rates

The specific net heating rate dominated by the EA process of $\nu_e$ and $\bar{\nu}_e$, processes (1a–1c) in Table I, is given by

$$\dot{q} = \frac{1}{\rho}\int dE\, E[\chi_{e,\mathrm{EA}}\varrho_{ee} - j_{e,\mathrm{EA}}(\varrho_{ee,\mathrm{FO}} - \varrho_{ee}) + \bar{\chi}_{e,\mathrm{EA}}\bar{\varrho}_{ee} - \bar{j}_{e,\mathrm{EA}}(\bar{\varrho}_{ee,\mathrm{FO}} - \bar{\varrho}_{ee})]. \quad (17)$$

We show the above net heating rate in Models II–V for cases with and without FFCs in Fig. 19. Notice that the rates before FFCs (black curve) are practically indistinguishable on the plot for all models despite different attenuation strengths adopted to $Y_e$ profiles. This is because the reduction of $\nu_e$'s and enhancement of $\bar{\nu}_e$'s are complementary to each other. Without FFCs, the regions with net cooling ($\dot{q} < 0$) and heating ($\dot{q} > 0$) are separated by the gain radius at $r \approx 63$ km. In the cooling region, the maximum net cooling rate is as low as $\dot{q} \approx -8.4 \times 10^{21}$ erg g$^{-1}$ s$^{-1}$ at $r \simeq 46$ km.

With FFCs, other solid curves in Fig. 19(a) show the radial profiles of $\dot{q}$ at $t = 32$ $\mu$s for Models II–V. Since

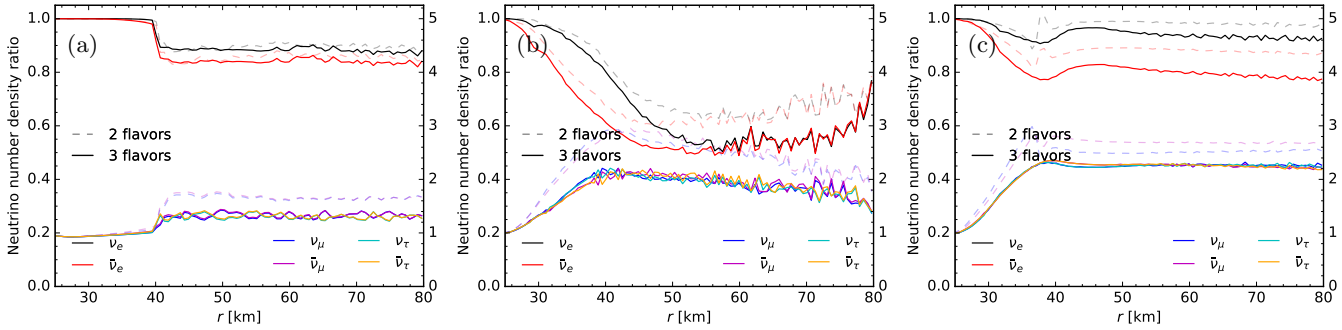



FIG. 16. Comparison of the radial profiles of neutrino number density ratios $n_\nu(t)/n_\nu(t=0)$ between two-flavor (dashed) and three-flavor (solid) schemes for Model II at 320 $\mu$s (a), Model IV at 32 $\mu$s (b) and at 320 $\mu$s (c). Note that the scales for heavy-lepton neutrinos are labeled on the right axes.

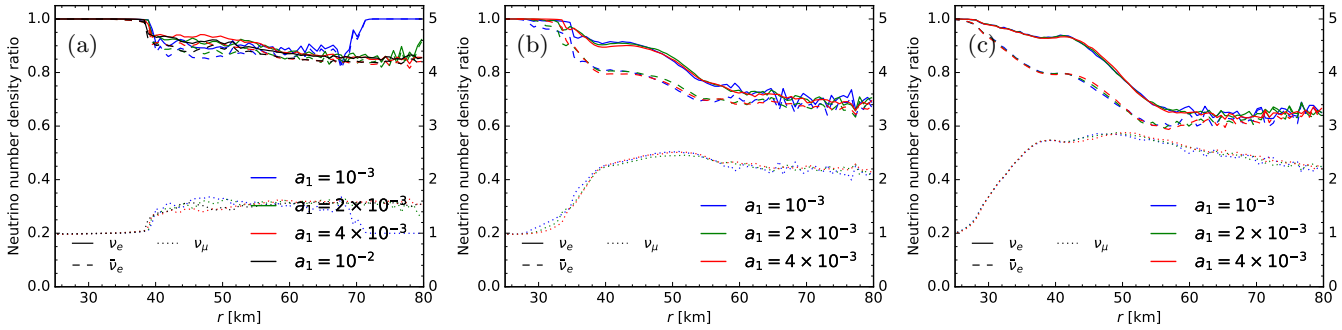



FIG. 17. Comparison of the radial profiles of neutrino number density ratios $n_\nu(t)/n_\nu(t=0)$ (blue) in Models II (a), III (b), and IV (c) at 64 $\mu$s with different attenuation factors $a_1 = 10^{-3}$ (blue), $2 \times 10^{-3}$ (green), $4 \times 10^{-3}$ (red), and $1 \times 10^{-2}$ (black in Model II). Note that the scales for the ratio of $\nu_\mu$ are labeled on the right axes.

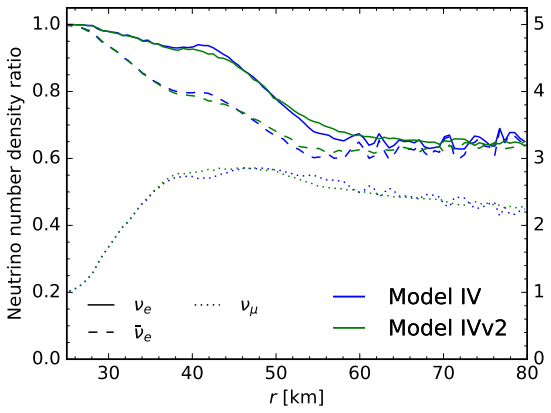



FIG. 18. Comparison of the radial profiles of neutrino number density ratios $n_\nu(t)/n_\nu(t=0)$ between Models IV and IVv2 at 64 $\mu$s. The scale for the ratio of $\nu_\mu$ is labeled on the right axis.

the pairwise FFCs from electron to heavy-lepton flavors lead to the reduction of heating rates in these models, $\dot{q}$ are reduced in the entire domain. As a result, the gain radius shifts to larger radii around $r \simeq 69$ km, consistent with what found in [49, 51]. At this time, $\dot{q}$ in models with more $Y_e$ attenuation that exhibit more flavor conversions of $\nu_e$ to $\nu_\mu$ shown in Fig. 11 are more negative, with Model V showing the most negative $\dot{q} \simeq -1.3 \times 10^{22}$ erg g$^{-1}$ s$^{-1}$.

It should also be highlighted that because the prompt flavor conversion takes place in different timescales among models and at different radii, the heating rate can be reduced more at an earlier time. For example, the cyan dashed curve in Fig. 19(a) demonstrates that the net cooling from $r \simeq 28$ km to $\simeq 46$ km are drastically enhanced at $t = 3.2$ $\mu$s in Model IV. The reduction of $\dot{q}$ at this time is particularly severe around $r \simeq 30$ km and $\simeq 44$ km, where the initial ELN deepness ratios are $\approx 1$ (Fig. 3), allowing nearly complete flavor equilibration or even overconversion after the prompt FFC. However, for models in which the feedback between FFCs and collisional weak processes inside neutrinospheres provides supply of additional $\nu_e$'s and $\bar{\nu}_e$'s, the reduction of $\dot{q}$ can be alleviated at later times. This can lead to the change of ordering in terms of $\dot{q}$ reduction from what was discussed above at $t = 32$ $\mu$s. Figure 19(b) shows that the reductions of $\dot{q}$ in Models III and IV are less than that in Model II around $r \simeq 46$ km at $t = 160$ $\mu$s when the systems approach toward asymptotic states. For Model V, although it remains to be the model affected by FFCs the most, the impact also becomes slightly smaller at this time. As for Models VI and VII, because the impact of

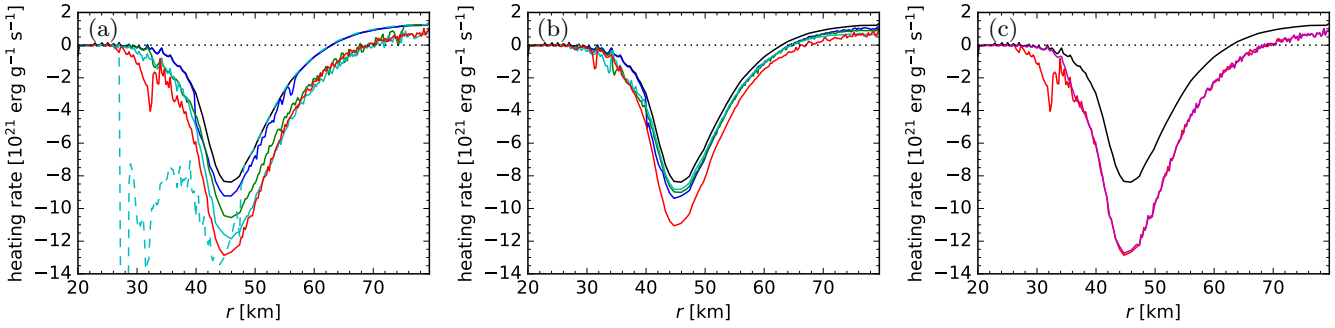



FIG. 19. Comparison of the specific heating-cooling rate after the prompt FFCs at $t = 32$ $\mu$s (a) and in the asymptotic state at $t = 160$ $\mu$s (b) for Models II (blue), III (green), IV (cyan), and V (red). The original heating rates without flavor conversions are shown by the black curves and are indistinguishable for all models. The blue dashed curve in (a) is for Model IV at $t = 3.2$ $\mu$s. The magenta curve in (c) is for Model Vc0 at $t = 32$ $\mu$s, compared to Model V (red).

FFC on $\nu_e$'s and $\bar{\nu}_e$'s flavor content are small, $\dot{q}$ in these two models are only affected insignificantly, and are thus not shown in Fig. 19.

Besides the FFCs, the CFI that is present in Model V also contributes to the reduction of $\dot{q}$. Figure 19(c) shows that an additional $\dot{q}$ reduction appears at $r \simeq 32$ km and $t = 32$ $\mu$s in Model V, which is otherwise not present in the corresponding model without the CFI (Model Vc0). The impact of this component also largely diminishes at later times as shown in Fig. 19(b).

We would like to caution that the impact of FFC on the net heating rate discussed here is under the assumption that background matter profiles are static. In realistic situations, even if the prompt flavor conversions can be reset by the collisional weak processes, its imprint on the matter background may introduce further feedback to neutrinos and lead to an outcome that sits in between the state after the prompt conversions and the asymptotic one shown above. This additional feedback effect may also depend on involved hydrodynamic timescales and will need to be evaluated in future work.

### B. Equilibrium electron fraction

Another pertinent implication of FFC on supernova physics is related to the nucleosynthesis of elements. Although our background matter profile is taken from a supernova model that does not explode, we expect that the analyses of the impact of FFC on the chemical composition in our models can still shed light on what may happen in successfully exploded supernovae with different progenitor masses as well as the neutrino-driven wind at a later epoch.

For this purpose, we examine the effect of FFC on the equilibrium $Y_e$, which is a key quantity relevant to supernova nucleosynthesis. For $T \lesssim 1$ MeV, the equilibrium $Y_e$ is dominated by the neutrino absorption reactions and can be approximated by

$$Y_e^{\nu,\mathrm{eq}} = \left(1 + \frac{L_{\bar{\nu}_e}}{L_{\nu_e}} \frac{\langle E_{\bar{\nu}_e}\rangle - 2\Delta_{np} + 1.2\Delta_{np}^2/\langle E_{\bar{\nu}_e}\rangle}{\langle E_{\bar{\nu}_e}\rangle + 2\Delta_{np} + 1.2\Delta_{np}^2/\langle E_{\nu_e}\rangle}\right)^{-1}, \tag{18}$$

where $\Delta_{np} = m_n - m_p$ is the neutron–to-proton mass difference [139]. Notice that although all neutrinos propagate radially outwards at the outer boundary of our simulation domain, there remains a fraction of neutrinos with non-negligible transverse velocities, i.e., $v_r \neq 1$. Assuming that the neutrino flux is conserved and no other types of flavor oscillations occur outside the outer boundary, the neutrino spectra beyond the simulation domain for $r > r_{\rm ob}$ can then be approximated by

$$\langle \varrho \rangle_A(E) = \int_0^1 dv_r\, \varrho(E, v_r, r_{\rm ob}) \times \left(\frac{r_{\rm ob}}{r}\right)^2 \frac{v_r}{\sqrt{1 - (\frac{r_{\rm ob}}{r})^2(1 - v_r^2)}}, \tag{19}$$

which can be further used to calculate the neutrino luminosity and mean energy relevant to nucleosynthesis.

The equilibrium $Y_e^{\nu,\mathrm{eq}}$ in Models II–V are shown from $r = 60$ km to 200 km in Fig. 20. The values asymptotically approach constants at larger radii where neutrinos propagate predominately along the radial direction. Unlike the heating rate, which is hardly affected by the $Y_e$ attenuation for cases without any flavor conversions, the $Y_e^{\nu,\mathrm{eq}}$ profiles depend on the ratio between $\nu_e$'s and $\bar{\nu}_e$'s and hence vary from Models II–IV without FFCs (transparent dashed curves). At $r = 200$ km, $Y_e^{\nu,\mathrm{eq}}$ are $\approx 0.527$, 0.511, 0.493, and 0.476 in Models II–V without FFCs, respectively, which reflects the increasing dominance of $\bar{\nu}_e$ capture rates that result in more neutron-rich conditions.

The impact of the prompt FFCs on $Y_e^{\nu,\mathrm{eq}}$ is shown in Fig. 20(a).[4] Since the prompt FFCs transform an approximately same amount of $\nu_e$ and $\bar{\nu}_e$ into the heavy-lepton flavors, it implies that for cases with an initial

[4] Because the timescale for prompt FFCs at $r_{\rm ob}$ varies from model

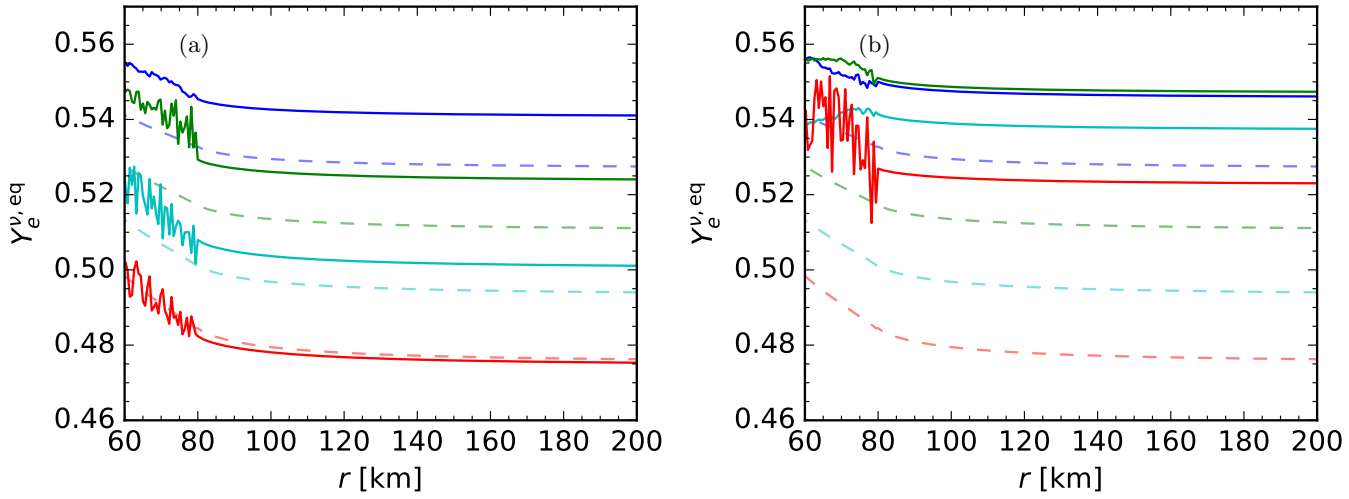



FIG. 20. Comparison of the equilibrium electron fraction $Y_e^{\nu,\mathrm{eq}}$ after the prompt FFCs (a) and at a later time $t = 160$ $\mu$s (b) for Models II (blue), III (green), IV (cyan), and V (red). The transparent dashed curves show the corresponding profiles without flavor conversions for comparisons.

asymmetry between $\nu_e$ and $\bar{\nu}_e$, the asymmetry gets further enhanced by FFCs. As a result, a $\nu_e$-dominated condition becomes even more dominated by $\nu_e$'s, and vice versa. In addition, given that the mean energy of $\nu_e$ is generally lower than those of other species, FFC enhances $\langle E_{\nu_e}\rangle$ relatively more than $\langle E_{\bar{\nu}_e}\rangle$, which also leads to the increased $Y_e^{\nu,\mathrm{eq}}$. Combining these two effects, Fig. 20(a) shows that $Y_e^{\nu,\mathrm{eq}}$ at $r = 200$ km are increased by $\approx 0.013$, 0.013, 0.008, and -0.001 in Models II–V, respectively. The impact is generally larger for models whose $Y_e^{\nu,\mathrm{eq}}$ deviate more from 0.5. Note that the impact in Model II is similar to that in Model III, due to the shallower ELN crossing in Model II that only allows a smaller amount of FFCs.

Similar to the impact on the heating rate, the impact of flavor conversions on $Y_e^{\nu,\mathrm{eq}}$ at the asymptotic times are reshaped by the feedback effect from the collisional weak processes in Models III–V. As discussed in Sec. V B, the feedback effect tends to bring the flavor content of $\nu_e$ back closer to the level prior to oscillations than for $\bar{\nu}_e$, this increases the ratio of $\nu_e$ number density over $\bar{\nu}_e$ by more than 10% compared to the ratio before conversions, which leads to a large change on $Y_e$. Figure 20(b) shows that for Model II without feedback effect, $Y_e^{\nu,\mathrm{eq}}$ at $t = 160$ $\mu$s is indeed similar to that shown in Fig. 20(a). In contrast, $Y_e^{\nu,\mathrm{eq}}$ in Models III–V are substantially increased by $\simeq 0.02$–$0.04$ at the asymptotic times compared to the values immediately after the FFCs, due to the increased $\nu_e$ to $\bar{\nu}_e$ number density ratios. These results suggest that with the feedback effect, a neutron-rich condition may even be turned into a proton-rich condition (Model V). Moreover, the feedback effect may further enhance the proton richness than the outcome of prompt FFCs. All these may facilitate the production of light $p$-nuclei through the $\nu p$-process [47, 140].

to model, we show the profiles at $t = 96$ $\mu$s for Model II, at $t = 64$ $\mu$s for Model III, and at $t = 32$ $\mu$s for Models IV and V, respectively.

### C. Free-streaming neutrino spectra

Finally, we examine the impact of the flavor conversions on the energy spectra of neutrinos at sufficiently large radii outside the simulation boundary. We follow the same extrapolation method in Sec. VI B to obtain the differential number luminosity

$$\frac{dL_{\mathrm{num}}}{dE}(E) = 4\pi \int_0^1 dv_r\, \varrho(E, v_r, r_{\mathrm{ob}}) r_{\mathrm{ob}}^2 v_r \qquad (20)$$

at sufficiently large radii. Note that no additional flavor conversions are included in order to focus on the impact of FFCs on the free-streaming neutrinos.

Figure 21(a–e) shows the free-streaming neutrino energy spectra at $t = 320$ $\mu$s in Models II–VI for cases with and without flavor conversions. In addition, Fig. 21(f) shows the spectra after the prompt FFCs at $t = 96$ $\mu$s in Model IV. A common feature is that the degeneracy between $\nu_\mu$'s and $\bar{\nu}_\mu$'s are all broken after the flavor conversions given that $\nu_e$'s and $\bar{\nu}_e$'s have different energy spectra prior to flavor conversions.

However, these plots also clearly show that flavor equilibration is not a general outcome of FFCs. Only the Model V [Fig. 21(d)] as well as the earlier phase after FFCs in Model III (not shown) and IV [Fig. 21(f)] obtain outcomes that are close to flavor equilibration. For the rest, equilibration is not achieved due to the initial shallow ELN angular crossings in Model II, the dynamical feedback from collisions after the FFCs in Model III and

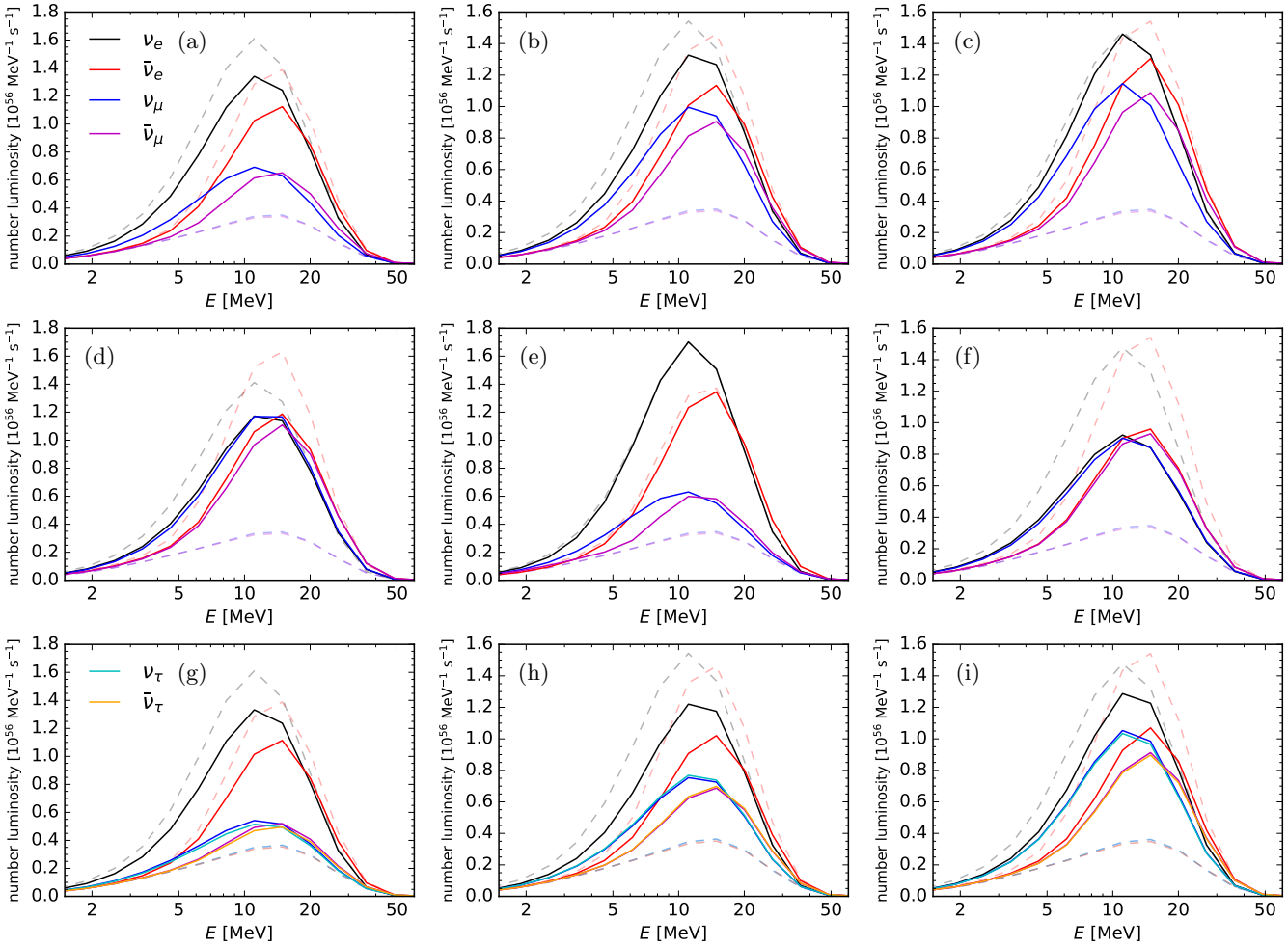



FIG. 21. Free-streaming neutrino energy spectra at $t = 320$ $\mu$s in Models II (a), III (b), IV (c), V (d), VI (e), IIf3 (g), IIIf3 (h), and IVf3 (i). Note that the panel (f) shows the spectrum immediately after the prompt FFC at $t = 96$ $\mu$s in Model IV for illustration.

IV, and a complete reset from the collisional weak processes when the angular crossing is inside neutrinospheres in Model VI. For cases that include three flavors shown in Fig. 21(g–i) for Model II–IV, the results are qualitatively similar to that in the corresponding two-flavor cases.

We also emphasize again that for models where the collisional feedback effect is important, it results in a net total amount of neutrino luminosity summing all flavors compared to the values without flavor conversions, as discussed earlier in Sec. V B. Even though the flavor equilibration is achieved in Model V, the energy spectra of all neutrino species lie above the average of the unoscillated level. Based on this observation, we expect that the FFC can lead to a significant enhancement for the energy loss rate due to neutrino emission, which will further affect the CCSN dynamics. This effect, however, needs to be fully evaluated via a CCSN hydrodynamic simulation that incorporates the FFCs.

## VII. DISCUSSION AND CONCLUSIONS

We have utilized our multi-group and discrete-ordinate collective neutrino oscillation simulation code based on COSE$\nu$ to comprehensively study the emergence, evolution, and effects of FFCs in a spherically symmetric supernova background obtained with AGILE-BOLTZTRAN. The collisional weak processes in this study have been extended from Ref. [71] to include neutrino-electron scatterings and neutrino pair reactions. We adopted radial dependent $Y_e$ attenuation schemes that allow us to probe various scenarios with the ELN angular crossings residing outside, inside, and across the neutrinosphere, as well as conditions with neutrino number densities dominated by $\nu_e$'s or $\bar{\nu}_e$'s, prior to the FFCs. We have also investigated the impact of assumed number of neutrino flavors, the choice of the $\mathbf{H}_{\nu\nu}$ attenuation factor, the size of the vacuum term, and the coexistent CFI. All models examined in this work are listed in Table II.

To analyze our simulation results, we introduced the

spectrogram method that is suitable for the inhomogeneous neutrino gas in Sec. III C to inspect the small-scale structure of the flavor coherence. The spectrograms are compared to the dispersion relation obtained from the LSA in Sec. IV. We examined in detail the evolution history in both two- and three-flavor models as well as the impact on FFCs from other flavor instabilities in Sec. V. Finally, we discussed the general effects of FFCs on CCSN physics and the free streaming neutrino energy spectra in Sec. VI.

We confirmed several characteristics of FFCs reported in a recent study [66]. After the FFCs take place, neutrino profiles eventually reach an asymptotic state on the coarse-grained level, with the angular crossings in the ELN distributions eliminated in general. The asymptotic state of neutrinos shows a strong dependence on the collisional neutrino-matter interactions, particularly when the ELN angular crossings appear inside the neutrinospheres.

On top of that, we summarize our new findings as follows:

1. The comparison of our LSA and spectrogram results shows that the emergence of FFI and the early evolution of FFCs qualitatively follow the prediction of dispersion relation from the LSA (Fig. 9). The unstable modes with larger growth rates develop predominantly, which corresponds to the emergence of small-scale structure in the off-diagonal flavor coherence. Although the dominant modes of the small-scale structure may be shifted to slightly different $K_r$ during the later evolution or even disappear for some radial ranges, they generally survive and stay rather close to the initially most unstable $K_r$ mode from LSA in Models II–IV and Vc0. In addition, we also observed that the small-scale structure can appear in the diagonal elements, i.e., neutrino number density, with a length scale consistent with that in the off-diagonal flavor coherence (Fig. 10).

2. We showed that a good scalability of the dispersion relation exists even when the attenuation on $\mathbf{H}_{\nu\nu}$ is employed (Figs. 4–6). Although this implies the typical length of the small-scale structure changes with the attenuation factors, the neutrino flavor evolution has rather weak dependence on the chosen attenuation factors so long as the dominant $K_r$ modes of the dispersion relation are well resolved (Fig. 17).

3. The evolution of the neutrino flavor content exhibits a large variation among different models (Fig. 11). The prompt flavor conversions immediately after the emergence of the local FFI can lead to nearly complete flavor equilibration for models with deep ELN angular crossings (Models III–V). Models with shallow crossings result in incomplete flavor conversions, demonstrated by Model II.

4. We found that after the prompt FFC phase in Models III–V, an interesting feedback effect due to the neutrino-matter interaction just below the neutrinosphere continues to restore the $\nu_e$'s and $\bar{\nu}_e$'s, and reshapes the local ELNs as well as the angular distribution of each neutrino species (Figs. 13 and 14). This effect can largely alter the flavor state of neutrinos around and outside the neutrinosphere without incurring an additional ELN angular crossing, leading to an asymptotic state much different from that obtained right after the prompt FFCs. Although heavy-lepton flavor neutrinos do not directly experience collisional restoration, their flavor contents are also enhanced through further flavor conversions from the restored $\nu_e$'s and $\bar{\nu}_e$'s. Detailed explanations on this dynamical feedback were discussed in Sec. V B.

5. When the ELN angular crossings only reside inside the neutrinosphere but not outside (Models VI and VII), the impact of FFCs can affect the heavy-lepton flavor neutrinos in the free-streaming regime, as they suffer less collisional processes on diffusing outwards. However, negligible effects are observed for $\nu_e$'s and $\bar{\nu}_e$'s owing to the large opacity that they experience (Fig. 11). Overall, the asymptotic outcome of FFCs in the free-streaming regime depends on the how deep the FFI region is inside the neutrinosphere.

6. Regarding the implication of FFCs on the CCSN dynamics, we found that with the prompt FFCs, it reduces the net heating rate behind the shock due to the reduced $\nu_e$ and $\bar{\nu}_e$ absorption rates, resulting in more cooling below as well as less heating above the gain radius, consistent with [66]. However, we also found that for Models III–V in which the collisional feedback plays an important role after the prompt FFCs, the restoration of $\nu_e$'s and $\bar{\nu}_e$'s at later times reduces the impact of FFCs on the enhanced cooling in the asymptotic state (Fig. 19).

7. The feedback mechanism also leads to another interesting consequence on $Y_e^{\nu,\mathrm{eq}}$. Right after the prompt FFCs, because similar amount of flavors is converted, the asymmetry between $\nu_e$ and $\bar{\nu}_e$ number densities is increased. Hence, it enhances the proton or neutron richness depending on the initial value of $Y_e$. However, the collisional feedback effect restores more $\nu_e$'s than $\bar{\nu}_e$'s due to the disparity between their collisional rates. It systematically leads to higher $Y_e^{\nu,\mathrm{eq}}$ and makes material more proton-rich in the asymptotic state (Fig. 20).

8. For the emerging neutrino energy spectra after the FFCs relevant to the neutrino detection, we found that flavor equilibration is not a general outcome. Significant differences between neutrinos of electron and heavy-lepton flavors can exist (Fig. 21). This

finding indicates that the subsequent slow flavor instability and Mikheyev–Smirnov–Wolfenstein matter effect can be relevant and will need to be considered for studying the neutrino signals.

9. Although the above conclusions are based on the two-flavor scenario, we found that the general features of flavor conversions remain qualitatively similar to that with three flavors. Nevertheless, including a third flavor can result in additional flavor conversions into the additional sector, which leads to less amount of $\nu_e$ and $\bar{\nu}_e$. For the heavy-lepton sectors, the enhanced amounts for muon and tauon flavors in the three-flavor scenarios are less than that for muon flavors in the two-flavor cases. However, the total amount of heavy-lepton flavors in three-flavor cases is in fact larger than that in the two-flavor cases (Fig. 16).

10. We showed that the dispersion relation in our simulation domain is barely affected by the inclusion of the vacuum term when no attenuation is applied to $\mathbf{H}_{\nu\nu}$. When attenuation is applied, one may pragmatically adopt a reduced $\delta m^2$ to maintain the scalability of the dispersion relation. In this case, the simulation results are largely insensitive to the presence of the vacuum term (Fig. 18).

11. We found that the occurrence and the strength of CFI, which is intrinsically a multienergy phenomenon due to the energy dependence of the EA collisional rates, can depend on the energy spectra of heavy-lepton neutrinos. When the NES processes are included, the condition for the CFI in this work becomes different from that in Ref. [71] and can be understood using the formula developed in Ref. [73] (Sec. IV B). Among all models considered in this paper, noticeable flavor conversion due to the CFI is only observed inside the FFC region in Model V when the initial condition is dominated by $\bar{\nu}_e$'s. While this affects less the properties of neutrinos than the FFC does, it can lead to a transient enhancement of the net cooling rate in the corresponding region (Fig. 19).

In summary, our comprehensive study of neutrino flavor conversions in the post-shocked region supports that it is important to include FFCs in supernova simulations, in order to fully understand the role of neutrinos in supernova dynamics, nucleosynthesis, as well as to predict the CCSN neutrino signals. Our work also highlights the need of using energy-dependent and realistic neutrino collision rates in relevant simulations, because the condition and evolution of FFC and CFI can be affected by those collision processes either directly or through feedback. Certain assumptions that were introduced in our work such as the spherical symmetry, static supernova background, and the missing coherent forward scattering with the matter remain to be further examined and will be studied in future.

## ACKNOWLEDGMENTS

We thank Ninoy Rahman and Gang Guo for discussions on collisional weak processes, Yi-Siou Wu on the comparison of two- and three-flavor results, and Sherwood Richers on the effective method of treating neutrino pair reactions. ZX and GMP acknowledge support of the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Advanced Grant KILONOVA No. 885281), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 279384907 – SFB 1245, and MA 4248/3-1. MRW and MG acknowledge supports from the National Science and Techonology Council, Taiwan under Grant No. 111-2628-M-001-003-MY4, and the Academia Sinica (Project No. AS-CDA-109-M11). MRW also acknowledges supports from the Physics Division of the National Center for Theoretical Sciences, Taiwan. CYL acknowledges support from the National Center for High-performance Computing (NCHC). TF and NKL acknowledge support from the Polish National Science Center (NCN) under Grant No. 2020/37/B/ST9/00691. The core-collapse supernova simulations were performed at the Wroclaw Center for Scientific Computing and Networking (WCSS). The neutrino quantum kinetic simulations were performed at the GSI Helmholtzzentrum für Schwerionenforschung. We acknowledge the following software: MATPLOTLIB [141], NUMPY [142], and SCIPY [143].

[1] A. Mirizzi, I. Tamborra, H.-T. Janka, N. Saviano, K. Scholberg, R. Bollig, L. Hüdepohl, and S. Chakraborty, Nuovo Cimento Rivista Serie **39**, 1 (2016), arXiv:1508.00785 [astro-ph.HE].
[2] T. Fischer, G. Guo, K. Langanke, G. Martinez-Pinedo, Y.-Z. Qian, and M.-R. Wu, (2023), arXiv:2308.03962 [astro-ph.HE].
[3] R. L. Workman, V. D. Burkert, V. Crede, E. Klempt, and et. al., Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).
[4] I. Tamborra and S. Shalgar, Annu. Rev. Nucl. Part. Sci. **71**, 165 (2021), arXiv:2011.01948.
[5] S. Richers and M. Sen, in *Handb. Nucl. Phys.* (Springer, 2023) pp. 3771–3787, arXiv:2207.03561.
[6] F. Capozzi and N. Saviano, Universe **8**, 94 (2022), arXiv:2202.02494.
[7] M. C. Volpe, arXiv e-prints (2023), arXiv:2301.11814.
[8] S. Abbar, H. Duan, K. Sumiyoshi, T. Takiwaki, and M. C. Volpe, Phys. Rev. D **100**, 043004 (2019), arXiv:1812.06883 [astro-ph.HE].
[9] M. Delfan Azari, S. Yamada, T. Morinaga, W. Iwakami, H. Okawa, H. Nagakura, and K. Sumiyoshi, Phys. Rev.

D **99**, 103011 (2019), arXiv:1902.07467.
[10] M. Delfan Azari, S. Yamada, T. Morinaga, H. Nagakura, S. Furusawa, A. Harada, H. Okawa, W. Iwakami, and K. Sumiyoshi, Phys. Rev. D **101**, 023018 (2020), arXiv:1910.06176 [astro-ph.HE].
[11] H. Nagakura, T. Morinaga, C. Kato, and S. Yamada, Astrophys. J. **886**, 139 (2019), arXiv:1910.04288.
[12] T. Morinaga, H. Nagakura, C. Kato, and S. Yamada, Phys. Rev. Res. **2**, 012046(R) (2020), arXiv:1909.13131 [astro-ph.HE].
[13] S. Abbar, H. Duan, K. Sumiyoshi, T. Takiwaki, and M. C. Volpe, Phys. Rev. D **101**, 043016 (2020), arXiv:1911.01983.
[14] R. Glas, H.-T. Janka, F. Capozzi, M. Sen, B. Dasgupta, A. Mirizzi, and G. Sigl, Phys. Rev. D **101**, 063001 (2020), arXiv:1912.00274.
[15] H. Nagakura, A. Burrows, L. Johns, and G. M. Fuller, Phys. Rev. D **104**, 083025 (2021), arXiv:2108.07281 [astro-ph.HE].
[16] S. Abbar, F. Capozzi, R. Glas, H. T. Janka, and I. Tamborra, Phys. Rev. D **103**, 063033 (2021), arXiv:2012.06594.
[17] A. Harada and H. Nagakura, Astrophys. J. **924**, 109 (2022), arXiv:2110.08291 [astro-ph.HE].
[18] R. F. Sawyer, Phys. Rev. D **79**, 105003 (2009), arXiv:arXiv:0803.4319.
[19] G. Sigl and G. Raffelt, Nuclear Physics B **406**, 423 (1993).
[20] A. Vlasenko, G. M. Fuller, and V. Cirigliano, Phys. Rev. D **89**, 105004 (2014), arXiv:1309.2628.
[21] C. Volpe, Int. J. Mod. Phys. E **24**, 1541009 (2015), arXiv:1506.06222.
[22] D. N. Blaschke and V. Cirigliano, Phys. Rev. D **94**, 033009 (2016), arXiv:1605.09383 [hep-ph].
[23] A. Banerjee, A. Dighe, and G. Raffelt, Phys. Rev. D **84**, 053013 (2011), arXiv:1107.2308.
[24] I. Izaguirre, G. Raffelt, and I. Tamborra, Phys. Rev. Lett. **118**, 021101 (2017), arXiv:1610.01612 [hep-ph].
[25] F. Capozzi, B. Dasgupta, E. Lisi, A. Marrone, and A. Mirizzi, Phys. Rev. D **96**, 043016 (2017), arXiv:1706.03360.
[26] C. Yi, L. Ma, J. D. Martin, and H. Duan, Phys. Rev. D **99**, 063005 (2019), arXiv:1901.01546 [hep-ph].
[27] T. Morinaga, Phys. Rev. D **105**, L101301 (2022), arXiv:2103.15267.
[28] J. D. Martin, C. Yi, and H. Duan, Phys. Lett. B **800**, 135088 (2020), arXiv:1909.05225 [hep-ph].
[29] S. Bhattacharyya and B. Dasgupta, Phys. Rev. Lett. **126**, 061302 (2021), arXiv:2009.03337 [hep-ph].
[30] S. Bhattacharyya and B. Dasgupta, Phys. Rev. D **102**, 063018 (2020), arXiv:2005.00459.
[31] M.-R. Wu, M. George, C.-Y. Lin, and Z. Xiong, Phys. Rev. D **104**, 103003 (2021), arXiv:2108.09886 [hep-ph].
[32] S. Richers, D. E. Willcox, N. M. Ford, and A. Myers, Phys. Rev. D **103**, 083013 (2021), arXiv:2101.02745.
[33] S. Richers, D. Willcox, and N. Ford, Phys. Rev. D **104**, 103023 (2021), arXiv:2109.08631.
[34] M. Zaizen and T. Morinaga, Phys. Rev. D **104**, 083035 (2021), arXiv:2104.10532 [hep-ph].
[35] S. Abbar and F. Capozzi, J. Cosmol. Astropart. Phys. **2022**, 051 (2022), arXiv:2111.14880 [astro-ph.HE].
[36] S. Richers, H. Duan, M.-R. Wu, S. Bhattacharyya, M. Zaizen, M. George, C.-Y. Lin, and Z. Xiong, Phys. Rev. D **106**, 043011 (2022), arXiv:2205.06282.
[37] S. Bhattacharyya and B. Dasgupta, Phys. Rev. D **106**, 103039 (2022), arXiv:2205.05129.
[38] E. Grohs, S. Richers, S. M. Couch, F. Foucart, J. P. Kneller, and G. C. McLaughlin, Phys. Lett. B **846**, 138210 (2023), arXiv:2207.02214.
[39] M. Zaizen and H. Nagakura, Phys. Rev. D **107**, 103022 (2023), arXiv:2211.09343 [astro-ph.HE].
[40] M. Zaizen and H. Nagakura, Phys. Rev. D **107**, 123021 (2023), arXiv:2304.05044 [astro-ph.HE].
[41] Z. Xiong, M.-R. Wu, S. Abbar, S. Bhattacharyya, M. George, and C.-Y. Lin, Phys. Rev. D **108**, 063003 (2023).
[42] H. Nagakura, L. Johns, and M. Zaizen, (2023), arXiv:2312.16285v1.
[43] G. G. Raffelt and G. Sigl, Phys. Rev. D **75**, 083002 (2007), arXiv:hep-ph/0701182 [hep-ph].
[44] S. Abbar and M. C. Volpe, Phys. Lett. B **790**, 545 (2019), arXiv:1811.04215.
[45] L. Johns, H. Nagakura, G. M. Fuller, and A. Burrows, Phys. Rev. D **102**, 103017 (2020).
[46] Z. Xiong, M.-R. Wu, and Y.-Z. Qian, Phys. Rev. D **108**, 043007 (2023), arXiv:2303.05906.
[47] Z. Xiong, A. Sieverding, M. Sen, and Y.-Z. Qian, Astrophys. J. **900**, 144 (2020), arXiv:2006.11414 [astro-ph.HE].
[48] S.-I. Fujimoto and H. Nagakura, Mon. Not. R. Astron. Soc. **519**, 2623 (2023), arXiv:2210.02106.
[49] H. Nagakura, Phys. Rev. Lett. **130**, 211401 (2023), arXiv:2301.10785 [astro-ph.HE].
[50] J. Ehring, S. Abbar, H. T. Janka, G. Raffelt, and I. Tamborra, Phys. Rev. D **107**, 103034 (2023), arXiv:2301.11938.
[51] J. Ehring, S. Abbar, H. T. Janka, G. Raffelt, and I. Tamborra, Phys. Rev. Lett. **131**, 061401 (2023), arXiv:2305.11207.
[52] A. B. Balantekin, M. J. Cervia, A. V. Patwardhan, R. Surman, and X. Wang, (2023), arXiv:2311.02562.
[53] M.-R. Wu, I. Tamborra, O. Just, and H.-T. Janka, Phys. Rev. D **96**, 123015 (2017), arXiv:1711.00477 [astro-ph.HE].
[54] M. George, M.-R. Wu, I. Tamborra, R. Ardevol-Pulpillo, and H.-T. Janka, Phys. Rev. D **102**, 103015 (2020), arXiv:2009.04046 [astro-ph.HE].
[55] X. Li and D. M. Siegel, Phys. Rev. Lett. **126**, 251101 (2021), arXiv:2103.02616 [astro-ph.HE].
[56] O. Just, S. Abbar, M. R. Wu, I. Tamborra, H. T. Janka, and F. Capozzi, Phys. Rev. D **105**, 083024 (2022), arXiv:2203.16559.
[57] R. Fernández, S. Richers, N. Mulyk, and S. Fahlman, Phys. Rev. D **106**, 103003 (2022), arXiv:2207.10680 [astro-ph.HE].
[58] H. Nagakura, Phys. Rev. D **108**, 103014 (2023), arXiv:2306.10108.
[59] J. Froustey, S. Richers, E. Grohs, S. Flynn, F. Foucart, J. P. Kneller, and G. C. McLaughlin, (2023), arXiv:2311.11968.
[60] E. Grohs, S. Richers, S. M. Couch, F. Foucart, J. Froustey, J. Kneller, and G. McLaughlin, (2023), arXiv:2309.00972.
[61] C. J. Stapleford, C. Fröhlich, and J. P. Kneller, Phys. Rev. D **102**, 081301(R) (2020), arXiv:1910.04172 [astro-ph.HE].
[62] S. Shalgar and I. Tamborra, Phys. Rev. D **108**, 043006 (2023), arXiv:2206.00676.

[63] H. Nagakura and M. Zaizen, Phys. Rev. Lett. **129**, 261101 (2022), arXiv:2206.04097v1.
[64] S. Shalgar and I. Tamborra, Phys. Rev. D **107**, 063025 (2023), arXiv:2207.04058 [astro-ph.HE].
[65] H. Nagakura and M. Zaizen, Phys. Rev. D **107**, 063033 (2023), arXiv:2211.01398 [astro-ph.HE].
[66] H. Nagakura and M. Zaizen, (2023), arXiv:2308.14800.
[67] H. Duan, G. M. Fuller, and Y.-Z. Qian, Annu. Rev. Nucl. Part. **60**, 569 (2010), arXiv:1001.2799 [hep-ph].
[68] L. Johns, Phys. Rev. Lett. **130**, 191001 (2023), arXiv:2104.11369.
[69] I. Padilla-Gay, I. Tamborra, and G. G. Raffelt, Phys. Rev. D **106**, 103031 (2022), arXiv:2209.11235.
[70] L. Johns and Z. Xiong, Phys. Rev. D **106**, 103029 (2022), arXiv:2208.11059 [hep-ph].
[71] Z. Xiong, M.-R. Wu, G. Martínez-Pinedo, T. Fischer, M. George, C.-Y. Lin, and L. Johns, Phys. Rev. D **107**, 083016 (2023), arXiv:2210.08254 [astro-ph.HE].
[72] Y.-C. Lin and H. Duan, Phys. Rev. D **107**, 083034 (2023), arXiv:2210.09218.
[73] Z. Xiong, L. Johns, M.-R. Wu, and H. Duan, Phys. Rev. D **108**, 083002 (2023), arXiv:2212.03750.
[74] C. Kato, H. Nagakura, and M. Zaizen, Phys. Rev. D **108**, 023006 (2023), arXiv:2303.16453 [astro-ph.HE].
[75] J. Liu, M. Zaizen, and S. Yamada, Phys. Rev. D **107**, 123011 (2023), arXiv:2302.06263.
[76] R. Akaho, J. Liu, H. Nagakura, M. Zaizen, and S. Yamada, **109**, 023012 (2024), arXiv:2311.11272.
[77] J. Liu, R. Akaho, A. Ito, H. Nagakura, M. Zaizen, and S. Yamada, Phys. Rev. D **108**, 123024 (2023), arXiv:2310.05050.
[78] C. Kato, H. Nagakura, and L. Johns, (2023), arXiv:2309.02619.
[79] D. F. G. Fiorillo, I. Padilla-Gay, and G. G. Raffelt, (2023), arXiv:2312.07612.
[80] S. Shalgar and I. Tamborra, (2023), arXiv:2307.10366.
[81] M. George, C.-Y. Lin, M.-R. Wu, T. G. Liu, and Z. Xiong, Comput. Phys. Commun. **283**, 108588 (2023), arXiv:2203.12866.
[82] I. Tamborra, L. Hüdepohl, G. G. Raffelt, and H.-T. Janka, Astrophys. J. **839**, 132 (2017), arXiv:1702.00060 [astro-ph.HE].
[83] H. Nagakura, A. Burrows, D. Radice, and D. Vartanyan, Mon. Not. R. Astron. Soc. **492**, 5764 (2020), arXiv:1912.07615.
[84] I. Tamborra, F. Hanke, H. T. Janka, B. Müller, G. G. Raffelt, and A. Marek, Astrophys. J. **792**, 96 (2014), arXiv:1402.5418.
[85] R. Glas, H.-T. Janka, T. Melson, G. Stockinger, and O. Just, Astrophys. J. **881**, 36 (2019), arXiv:1809.10150.
[86] T. Rauscher, A. Heger, R. D. Hoffman, and S. E. Woosley, Astrophys. J. **576**, 323 (2002), arXiv:0112478 [astro-ph].
[87] A. Mezzacappa and S. W. Bruenn, Astrophys. J. **405**, 637 (1993).
[88] A. Mezzacappa and S. W. Bruenn, Astrophys. J. **410**, 740 (1993).
[89] A. Mezzacappa and S. W. Bruenn, Astrophys. J. **405**, 669 (1993).
[90] M. Liebendörfer, A. Mezzacappa, and F.-K. Thielemann, Phys. Rev. D **63**, 104003 (2001), arXiv:astro-ph/0012201 [astro-ph].
[91] M. Liebendörfer, O. E. B. Messer, A. Mezzacappa, S. W. Bruenn, C. Y. Cardall, and F. K. Thielemann, Astrophys. J., Suppl. **150**, 263 (2004), arXiv:astro-ph/0207036 [astro-ph].
[92] R. W. Lindquist, Ann. Physics1 **37**, 487 (1966).
[93] T. Fischer, G. Guo, G. Martínez-Pinedo, M. Liebendörfer, and A. Mezzacappa, Phys. Rev. D **102**, 123001 (2020), arXiv:2008.13628 [astro-ph.HE].
[94] M. Liebendörfer, S. Rosswog, and F.-K. Thielemann, Astrophys. J., Suppl. **141**, 229 (2002), arXiv:astro-ph/0106539 [astro-ph].
[95] T. Fischer, S. C. Whitehouse, A. Mezzacappa, F.-K. K. Thielemann, and M. Liebendörfer, Astron. Astrophys. **517**, A80 (2010), arXiv:0908.1871.
[96] S. W. Bruenn, Astrophys. J., Suppl. **58**, 771 (1985).
[97] T. Fischer, G. Guo, A. A. Dzhioev, G. Martínez-Pinedo, M.-R. Wu, A. Lohs, and Y.-Z. Qian, Phys. Rev. C **101**, 025804 (2020).
[98] A. Juodagalvis, K. Langanke, W. R. Hix, G. Martínez-Pinedo, and J. M. Sampaio, Nucl. Phys. A. **848**, 454 (2010), arXiv:0909.0179.
[99] P. J. Schinder and S. L. Shapiro, Astrophys. J. Suppl. Ser. **50**, 23 (1982).
[100] T. Thompson and A. Burrows, Nucl. Phys. A **688**, 377 (2001).
[101] T. Fischer, Astron. Astrophys. **593**, 103 (2016), arXiv:1608.05004.
[102] R. Buras, H.-T. Janka, M. T. Keil, G. G. Raffelt, and M. Rampp, Astrophys. J. **587**, 320 (2003), arXiv:0205006 [astro-ph].
[103] T. Fischer, S. C. Whitehouse, A. Mezzacappa, F.-K. K. Thielemann, and M. Liebendörfer, Astron. Astrophys. **499**, 1 (2009), arXiv:0809.5129.
[104] G. Guo, G. Martínez-Pinedo, A. Lohs, and T. Fischer, Phys. Rev. D **102**, 023037 (2020), arXiv:2006.12051 [hep-ph].
[105] G. G. Raffelt, Astrophys. J. **561**, 890 (2001), arXiv:astro-ph/0105250 [astro-ph].
[106] M. T. Keil, G. G. Raffelt, and H.-T. Janka, Astrophys. J. **590**, 971 (2003), arXiv:astro-ph/0208035 [astro-ph].
[107] T. Fischer, G. Martínez-Pinedo, M. Hempel, and M. Liebendörfer, Phys. Rev. D **85**, 083003 (2012), arXiv:1112.3842 [astro-ph.HE].
[108] K. Sumiyoshi, S. Yamada, H. Suzuki, and S. Chiba, Phys. Rev. Lett. **97**, 091101 (2006), arXiv:1010.5550.
[109] E. O'Connor and C. D. Ott, Astrophys. J. **730**, 70 (2011), arXiv:1010.5550.
[110] T. Fischer, N.-U. Bastian, D. Blaschke, M. Cierniak, M. Hempel, T. Klähn, G. Martínez-Pinedo, W. G. Newton, G. Röpke, and S. Typel, Publications of the Astronomical Society of Australia **34**, e067 (2017), arXiv:1711.07411 [astro-ph.HE].
[111] T. Fischer, N.-U. F. Bastian, M.-R. Wu, P. Baklanov, E. Sorokina, S. Blinnikov, S. Typel, T. Klähn, and D. B. Blaschke, Nature Astronomy **2**, 980 (2018), arXiv:1712.08788 [astro-ph.HE].
[112] N. K. Largani, T. Fischer, and N. U. F. Bastian, (2023), arXiv:2304.12316.
[113] P. Jakobus, B. Müller, A. Heger, A. Motornenko, J. Steinheimer, and H. Stoecker, Mon. Not. R. Astron. Soc. **516**, 2554 (2022), arXiv:2204.10397 [astro-ph.HE].
[114] T. Kuroda, T. Fischer, T. Takiwaki, and K. Kotake, Astrophys. J. **924**, 38 (2022), arXiv:2109.01508 [astro-ph.HE].
[115] S. Zha and E. O'Connor, Phys. Rev. D **106**, 123037 (2022), arXiv:2209.12418 [astro-ph.HE].

[116] G. Sigl, Phys. Rev. D **105**, 043005 (2022), arXiv:2109.00091.

[117] S. Hannestad and G. Raffelt, Astrophys. J. **507**, 339 (1998), arXiv:9711132 [astro-ph].

[118] E. O'Connor, Astrophys. Journal, Suppl. Ser. **219**, 24 (2015).

[119] W. R. Yueh and J. R. Buchler, Astrophysic and Space Science **39**, 429 (1976).

[120] W. R. Yueh and J. R. Buchler, Astrophysics and Space Science **41**, 221 (1976).

[121] M. Rampp and H. T. Janka, Astronomy and Astrophysics **396**, 361 (2002), arXiv:astro-ph/0203101 [astro-ph].

[122] F. Capozzi, M. Chakraborty, S. Chakraborty, and M. Sen, Phys. Rev. Lett. **125**, 251801 (2020).

[123] F. Capozzi, M. Chakraborty, S. Chakraborty, and M. Sen, Phys. Rev. D **106**, 083011 (2022), arXiv:2205.06272.

[124] F. Capozzi, G. Raffelt, and T. Stirner, J. Cosmol. Astropart. Phys. **2019**, 002 (2019), arXiv:1906.08794.

[125] S. Sarikas, D. De Sousa Seixas, and G. Raffelt, Phys. Rev. D **86**, 12520 (2012).

[126] S. Abbar, H. Duan, and S. Shalgar, Phys. Rev. D **92**, 065019 (2015), arXiv:1507.08992.

[127] T. Morinaga and S. Yamada, Phys. Rev. D **97**, 023024 (2018), arXiv:1803.05913v1.

[128] J. D. Martin, J. Carlson, V. Cirigliano, and H. Duan, Phys. Rev. D **103**, 063001 (2021), arXiv:2101.01278 [hep-ph].

[129] J. L. Flanagan, *Speech Analysis Synthesis and Perception*, Vol. 3 (Springer Berlin, Heidelberg, 1972).

[130] J. Shen, R. Pang, R. J. Weiss, M. Schuster, N. Jaitly, Z. Yang, Z. Chen, Y. Zhang, Y. Wang, R. Skerrv-Ryan, R. A. Saurous, Y. Agiomvrgiannakis, and Y. Wu, in *IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, Vol. 2018-April (2018) pp. 4779–4783, arXiv:1712.05884.

[131] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, K. Agatsuma, N. Aggarwal, *et al.*, Phys. Rev. Lett. **116**, 61102 (2016), arXiv:1602.03837.

[132] F. Harris, Proceedings of the IEEE **66**, 51 (1978).

[133] P. Bloomfield, *Fourier analysis of time series: an introduction* (John Wiley & Sons, 2004).

[134] I. Padilla-Gay, S. Shalgar, and I. Tamborra, J. Cosmol. Astropart. Phys. **2021**, 017 (2021), arXiv:2009.01843.

[135] H. Duan, J. D. Martin, and S. Omanakuttan, Phys. Rev. D **104**, 123026 (2021), arXiv:2110.02286.

[136] L. Johns, (2023), arXiv:2306.14982.

[137] D. F. G. Fiorillo, G. G. Raffelt, and G. Sigl, (2024), arXiv:2401.05278.

[138] S. Abbar, M.-R. Wu, and Z. Xiong, (2024), arXiv:2401.17424v1.

[139] Y.-Z. Qian and S. E. Woosley, Astrophys. J. **471**, 331 (1996), arXiv:9611094 [astro-ph].

[140] C. Fröhlich, G. Martínez-Pinedo, M. Liebendörfer, F. K. Thielemann, E. Bravo, W. R. Hix, K. Langanke, and N. T. Zinner, Phys. Rev. Lett. **96**, 142502 (2006), arXiv:astro-ph/0511376 [astro-ph].

[141] J. D. Hunter, Computing in Science & Engineering **9**, 90 (2007).

[142] S. van der Walt, S. C. Colbert, and G. Varoquaux, Computing in Science & Engineering **13**, 22 (2011).

[143] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, *et al.*, Nature Methods **17**, 261 (2020).

[144] S. Typel, G. Röpke, T. Klähn, D. Blaschke, and H. H. Wolter, Phys. Rev. C **81**, 015803 (2010), arXiv:0908.2344 [nucl-th].

[145] M. Hempel and J. Schaffner-Bielich, Nucl. Phys. A. **837**, 210 (2010), arXiv:0911.4073 [nucl-th].

[146] M. Hempel, T. Fischer, J. Schaffner-Bielich, and M. Liebendörfer, Astrophys. J. **748**, 70 (2012), arXiv:1108.0848 [astro-ph.HE].

[147] T. Fischer, M. Hempel, I. Sagert, Y. Suwa, and J. Schaffner-Bielich, Eur. Phys. J. A **50**, 1 (2014), arXiv:1307.6190.

[148] I. Tews, T. Krüger, K. Hebeler, and A. Schwenk, Phys. Rev. Lett. **110**, 032504 (2013), arXiv:1206.0025 [nucl-th].

[149] J. Antoniadis, P. C. C. Freire, N. Wex, T. M. Tauris, R. S. Lynch, M. H. van Kerkwijk, M. Kramer, C. Bassa, V. S. Dhillon, T. Driebe, J. W. T. Hessels, V. M. Kaspi, V. I. Kondratiev, N. Langer, T. R. Marsh, M. A. McLaughlin, T. T. Pennucci, S. M. Ransom, I. H. Stairs, J. van Leeuwen, J. P. W. Verbiest, and D. G. Whelan, Science **340**, 448 (2013), arXiv:1304.6875 [astro-ph.HE].

[150] H. T. Cromartie, E. Fonseca, S. M. Ransom, P. B. Demorest, Z. Arzoumanian, H. Blumer, P. R. Brook, M. E. DeCesar, T. Dolch, J. A. Ellis, R. D. Ferdman, E. C. Ferrara, N. Garver-Daniels, P. A. Gentile, M. L. Jones, M. T. Lam, D. R. Lorimer, R. S. Lynch, M. A. McLaughlin, C. Ng, D. J. Nice, T. T. Pennucci, R. Spiewak, I. H. Stairs, K. Stovall, J. K. Swiggum, and W. W. Zhu, Nature Astronomy **4**, 72 (2020), arXiv:1904.06759 [astro-ph.HE].

[151] M. C. Miller, F. K. Lamb, A. J. Dittmann, S. Bogdanov, Z. Arzoumanian, K. C. Gendreau, S. Guillot, A. K. Harding, W. C. G. Ho, J. M. Lattimer, R. M. Ludlam, S. Mahmoodifar, S. M. Morsink, P. S. Ray, T. E. Strohmayer, K. S. Wood, T. Enoto, R. Foster, T. Okajima, G. Prigozhin, and Y. Soong, Astrophys. J. Lett. **887**, L24 (2019), arXiv:1912.05705 [astro-ph.HE].

[152] A. V. Bilous, A. L. Watts, A. K. Harding, T. E. Riley, Z. Arzoumanian, S. Bogdanov, K. C. Gendreau, P. S. Ray, S. Guillot, W. C. G. Ho, and D. Chakrabarty, Astrophys. J. Lett. **887**, L23 (2019), arXiv:1912.05704 [astro-ph.HE].

[153] M. C. Miller, F. K. Lamb, A. J. Dittmann, S. Bogdanov, Z. Arzoumanian, K. C. Gendreau, S. Guillot, W. C. G. Ho, J. M. Lattimer, M. Loewenstein, S. M. Morsink, P. S. Ray, M. T. Wolff, C. L. Baker, T. Cazeau, S. Manthripragada, C. B. Markwardt, T. Okajima, S. Pollard, I. Cognard, H. T. Cromartie, E. Fonseca, L. Guillemot, M. Kerr, A. Parthasarathy, T. T. Pennucci, S. Ransom, and I. Stairs, Astrophys. J. Lett. **918**, L28 (2021), arXiv:2105.06979 [astro-ph.HE].

[154] T. E. Riley, A. L. Watts, P. S. Ray, S. Bogdanov, S. Guillot, S. M. Morsink, A. V. Bilous, Z. Arzoumanian, D. Choudhury, J. S. Deneva, K. C. Gendreau, A. K. Harding, W. C. G. Ho, J. M. Lattimer, M. Loewenstein, R. M. Ludlam, C. B. Markwardt, T. Okajima, C. Prescod-Weinstein, R. A. Remillard, M. T. Wolff, E. Fonseca, H. T. Cromartie, M. Kerr, T. T. Pennucci, A. Parthasarathy, S. Ransom, I. Stairs,

L. Guillemot, and I. Cognard, Astrophys. J. Lett. **918**, L27 (2021), arXiv:2105.06980 [astro-ph.HE].
[155] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Afrough, B. Agarwal, M. Agathos, K. Agatsuma, N. Aggarwal, *et al.*, Phys. Rev. Lett. **119**, 161101 (2017).
[156] S. De, D. Finstad, J. M. Lattimer, D. A. Brown, E. Berger, and C. M. Biwer, Phys. Rev. Lett. **121**, 091102 (2018), arXiv:1804.08583 [astro-ph.HE].
[157] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt, B. Agarwal, M. Agathos, K. Agatsuma, N. Aggarwal, O. D. Aguiar, L. Aiello, A. Ain, *et al.*, Physical Review X **9**, 011001 (2019), arXiv:1805.11579.
[158] T. Fischer, S. Typel, G. Röpke, N.-U. F. Bastian, and G. Martínez-Pinedo, Phys. Rev. C **102**, 55807 (2020).
[159] F. X. Timmes and F. D. Swesty, Astrophys. J., Suppl. **126**, 501 (2000).

## Appendix A: Equation of state used in supernova simulation

For the present supernova simulation, the DD2 relativistic mean field (RMF) nuclear equation of state (EOS) is employed [144–146]. Its bulk properties agree quantitatively with nuclear physics experiments in the vicinity of nuclear saturation density (see Table 1 in Ref. [147] and references therein) as well as with chiral perturbation calculation for pure neutron matter [148]. Furthermore, the DD2 EOS is in agreement with current astrophysical constraints derived from pulsar observations [149–154] and from the GW170817 gravitational wave observation associated with a binary neutron star merger event [155–157]. At densities below nuclear saturation density, the DD2 RMF EOS is coupled with the modified nuclear statistical equilibrium approach of Ref. [145], featuring several 1,000 nuclear species with partly measured as well as theoretical nuclear masses. The transition to uniform nuclear matter at the saturation density is modeled via the excluded volume approach in the DD2 EOS, suppressing nuclear clusters (see Ref. [158] for a detailed discussion about this transition in the context of supernova simulations). At temperatures below $T \simeq 0.45$ MeV, which corresponds to low densities in the CCSN simulation, the RMF EOS is matched with the ideal silicon-sulfur gas approximation. This matches the iron-core silicon-sulfur layer transition of the stellar progenitor [86]. Furthermore, the nuclear EOS is coupled with electron, positron and photons EOS of Ref. [159] at all densities. Note that both, RMF and silicon-sulfur gas EOS treat Coulomb contributions differently, however, the differences remain negligible for the matching conditions here.

## Appendix B: Neutrino-electron scattering

We only include the diagonal elements in $\mathbf{C}_{\rm NES}$ for the heavy-lepton neutrinos as they are mainly thermalized by the NES process. The general formula is given as follows [87, 88],

$$\begin{aligned}\frac{4\pi^2}{E^2}\mathbf{C}_{ii,\rm NES}(E,v_r) \;=\; & [1-f_{\nu_i}(E,v_r)]\int E'^2dE'\,dv_r'\,R^{\rm in}_{\rm NES}(E,E',v_r,v_r')f_{\nu_i}(E',v_r') \\ & -f_{\nu_i}(E,v_r)\int E'^2dE'\,dv_r'\,R^{\rm out}_{\rm NES}(E,E',v_r,v_r')\,[1-f_{\nu_i}(E',v_r')]\;, \end{aligned} \tag{B1}$$

where $f_{\nu_i}$ is the neutrino distribution function, $i=\mu,\tau$, and $R^{\rm out/in}_{\rm NES}(E,E',v_r,v_r')$ are the in- and out-scattering kernels. Unlike the IS, which only has contributions from the first two moments in $v_r$ and $v_r'$, the angular dependence of the NES scattering kernel can be rather complicated and computationally challenging, depending on the kinematics of the target electron and position as well as on the numerical realization of the neutrino phase space. Nonetheless, it is common for neutrino transport codes to introduce the moment approximation. Here, we choose to retain only the zeroth moment, $\Phi_{\rm NES,0}(E,E')$, obtained by the appropriate integration of $R_{\rm NES}$ following Refs. [99, 121],

$$\Phi^{\rm out}_{\rm NES,0}(E,E') = \int_{-1}^{1} d\omega\, R^{\rm out}_{\rm NES}(E,E',\omega) \tag{B2}$$

where $\omega$ is the cosine of the angle between the incident and outgoing neutrinos. With the zeroth moment approximation, the scattering kernel is given as

$$R^{\rm out}_{\rm NES}(E,E',v_r,v_r') = \pi\Phi^{\rm out}_{\rm NES,0}(E,E'), \tag{B3}$$

denoted as the angle-independent $R^{\rm out}_{\rm NES}(E,E')$. Notice that the detailed balance ensures the validity of Eq. (10), see also Sec. IIa of Ref. [107], and hence

$$R^{\rm in}_{\rm NES}(E,E') = \exp\left(\frac{E'-E}{T}\right) R^{\rm out}_{\rm NES}(E,E'). \tag{B4}$$

Consequently, Eq. (B1) becomes

$$\mathbf{C}_{ii,\mathrm{NES}}(E,v_r) = -2\varrho_{ii}(E,v_r)\int dE'\, R^{\mathrm{out}}_{\mathrm{NES}}(E,E')E'^2 + \int dE'\, R^{\mathrm{in}}_{\mathrm{NES}}(E,E')\left\{E^2+4\pi^2\varrho_{ii}(E,v_r)\left[\exp\left(\frac{E-E'}{T}\right)-1\right]\right\}\int dv'_r\,\varrho_{ii}(E',v'_r). \tag{B5}$$

The integration over $v'_r$ can be calculated separately, which significantly speeds up the simulation.

For the implementation in the code, the neutrino density matrix is discretized, and so is the kernel $\varrho_{ii}(E',v'_r)$. Given that NES is mostly efficient in exchanging energy between neutrinos with similar energies, in practice we keep the elements sufficiently close to the diagonal in $R_{\mathrm{NES}}$, i.e., $R^{i_E,i_E+\Delta i_E}_{\mathrm{NES}}$ where $i_E$ is the index of the energy grid. We choose the range of energy exchange to be within $|\Delta i_E| \leq 3$. This choice not only helps accelerate the computation but also ensures the thermalization of neutrinos at a given temperature without breaking the detailed balance.

We compare the radial profiles of the neutrino number density and mean energy, obtained from COSE$\nu$ before any neutrino oscillations with the result based on the supernova simulations, denoted as BOLTZTRAN in Fig. 22. Both profiles of $\nu_e$ and $\bar{\nu}_e$ are highly consistent despite the different codes used. For the heavy-lepton neutrinos, the most significant effect by including the NES is the reduction of the mean energy for $r \gtrsim 25$ km. It decreases from $\simeq 40$ MeV to $\simeq 29$ MeV at $r = 30$ km, and from $\simeq 22$ MeV to $\simeq 19$ MeV at $r = 80$ km, due to the efficient thermalization through the NES [96, 105, 106]. After including the NES, the deviation of $\langle E_{\nu_\mu}\rangle$ profiles between COSE$\nu$ and BOLTZTRAN is strongly suppressed within $\simeq 2$ MeV. It is expected that the deviation should become even smaller with a convergence between the two codes when a complete NES scheme is used. Although the number density profile is also affected by the NES, the change is relatively small at the level of about 10%. This is because the scattering process does not produce extra neutrinos unlike the processes of EA and PR.

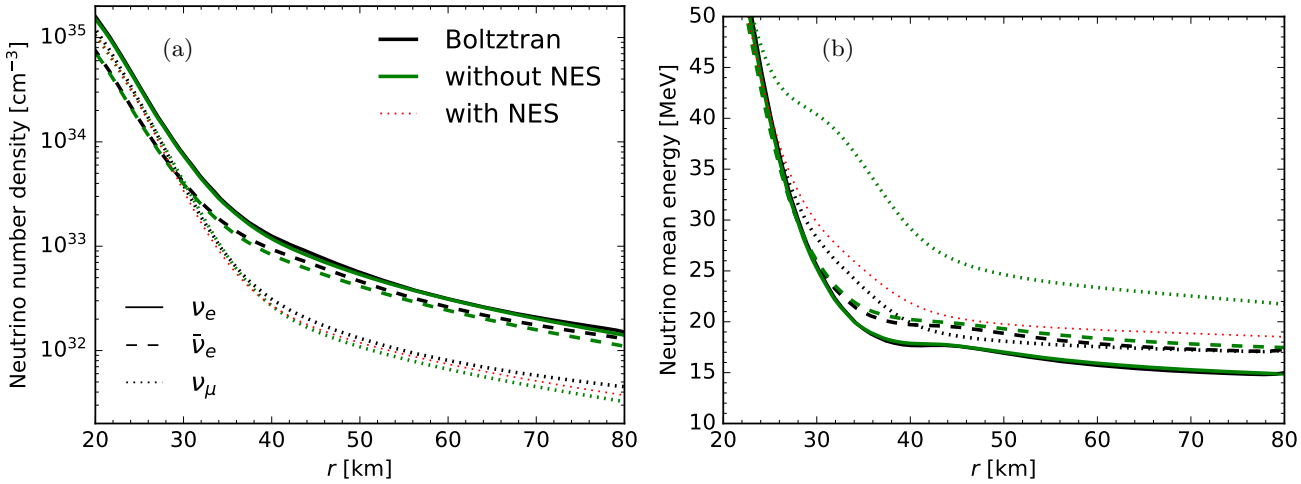



FIG. 22. Radial profiles of neutrino number density (a) and mean energy (b) based on BOLTZTRAN (black), COSE$\nu$ without NES (green), and COSE$\nu$ with NES (red) in Model I.

## Appendix C: Practical method to select collective non-spurious modes

To preclude those spurious modes, we take the following practical selection rule based on the non-collectiveness of a given mode. After the diagonalization of Eqs. (12) and (13), we sort the normalized modulus square of eigenmode $|Q_{i_E,i_{v_r}}|^2$ and $|\bar{Q}_{i_E,i_{v_r}}|^2$ for each eigenvalue $\Omega$. Considering the number of grids that we used, there are 3000 components in each eigenmode. To exclude those non-collective modes that mainly concentrates on one or a few angular grid, we then sum the first 30 largest components. If the sum is more than 0.1, it implies that on average these components dominate the averaged contributions from the rest. As a result, the corresponding eigenmode is likely of non-collective nature and is thus ruled out.

arXiv:2402.19239v1 [cond-mat.quant-gas] 29 Feb 2024

# Bogoliubov phonons in a Bose-Einstein condensate from the one-loop perturbative renormalization group

Niklas Rasch,[1,*] Aleksandr N. Mikheev,[1,2,†] and Thomas Gasenzer[1,2]

[1] *Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*
[2] *Institut für Theoretische Physik, Ruprecht-Karls-Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany*

(Dated: March 1, 2024)

Wilson's renormalization-group approach to the weakly-interacting single-component Bose gas is discussed within the symmetry-broken, condensate phase. Extending upon the work by Bijlsma and Stoof [1], wave-function renormalization of the temporal derivative contributions to the effective action is included in order to capture sound-like quasiparticle excitations with wave lengths larger than the healing-length scale. By means of a suitable rescaling scheme we achieve convergence of the coupling flows, which serve as a means to determine the condensate depletion in accordance with Bogoliubov theory, as well as the interaction-induced shift of the critical temperature.

## I. INTRODUCTION

Since the experimental realization of Bose-Einstein condensation in dilute interacting atomic gases [2, 3] there has been great interest in further exploring these systems both experimentally and theoretically. The basic theoretical approach is based on the Gross-Pitaevskii model, which, at the relevant low energies, describes the weak interparticle interactions within a contact-potential approximation. To capture the condensate phase and its fundamental excitations, Bogoliubov's ansatz builds on the assumption of a spontaneously broken U(1) symmetry and on an expansion about the field expectation value [4]. A subsequent truncation at Gaussian order represents an approximation, in which Bogoliubov-de Gennes diagonalization of the propagator yields the spectrum of quasiparticle excitations, which have collective, sound-wave character at long wavelengths. It allows determining leading-order perturbative corrections to thermodynamic observables such as the depletion of the condensate density and the associated shift of the chemical potential [5]. In three spatial dimensions, the approximation is valid in the dilute-gas regime, where the characteristic $s$-wave scattering length is much smaller than the interparticle spacing. Beyond this order of approximation, perturbation theory allows computing further corrections. At next-to-leading order, Beliaev's approximation involves all one-loop corrections to the Bogoliubov quasiparticle effective action [6–8]. See [9] for a review of related approaches to the weakly interacting Bose gas.

In the aforementioned perturbative extensions of the self-consistent framework, the sound-wave energy of excitations, linear in momentum $k$ for $k \to 0$, typically gives rise to logarithmic infrared (IR) divergences in $d = 3$, which makes it difficult to apply self-consistent diagrammatic renormalization methods [10]. It was shown that these divergences cause the anomalous self-energy to vanish at zero momentum [11, 12], which implies that self-consistent higher-order approaches require the corresponding normal contribution to the zero-momentum self-energy to equal the chemical potential in order for the Hugenholtz-Pines or Goldstone theorem to be fulfilled and thus the dispersion to be gapless [13]. Different approaches have been explored to obtain, in this context, perturbative corrections, including self-consistent diagrammatic perturbation theory [11, 12, 14, 15], in density-phase representation [16], within a number-conserving perturbative approach [17, 18], or in a both, conserving and gapless Hartree-Fock-Bogoliubov approximation [19]. In the strong-coupling limit a resummation scheme of IR divergences has been presented in [20]. IR divergences can in general be handled by means of renormalization group (RG) theory [21–26]. Within the RG approach, one iteratively integrates out fluctuations starting from the ultraviolet (UV) end of the spectrum and reabsorbs their effect in redefined coupling constants. This is achieved by means of RG flow equations governing the couplings, which parametrize an effective description at a given maximum spatial resolution scale.

The weakly interacting Bose gas in $d = 3$ and thermal equilibrium has been extensively studied using RG methods [1, 27–50]: Near criticality, the interaction-induced shift in the critical temperature has been determined using flow equations in perturbative approximation [1, 28, 29], within a non-perturbative functional RG approach [30, 31], and in the high temperature limit [32–35, 51]. The results can be compared with those of numerical approaches [52–57]. Furthermore, RG treatments have served to obtain critical exponents such as the exponent $\nu$ governing the scaling of the correlation length near criticality. Thereby, different techniques and approximation methods were employed, including derivative expansions [36], Wilson's momentum-shell method [1], and expansions in powers of $\epsilon = 4 - d$ in $d$ spatial dimensions [37], considering homogeneous as well as confined [28, 38, 39] systems. In [1], the critical exponent was determined as $\nu = 0.685$. Given that the thermal condensation phase transition of the interacting Bose gas belongs to the O(2) universality class, this value is to be compared with high-precision results from both theoretical and experimental studies. An expansion of anomalous dimensions up to six loops in perturbation theory in three spatial dimensions gave $\nu = 0.6703(15)$, while the $\epsilon$-expansion, which avoids the problem of IR divergences, yields $\nu = 0.6680(35)$ [58]. Measuring the superfluid phase transi-

* niklas.rasch@kip.uni-heidelberg.de
† current address: Institut für Physik, Johannes Gutenberg-Universität Mainz, Staudingerweg 7, 55128 Mainz, Germany

tion in liquid $^4$He gave a value of $\nu = 0.6717(4)$ [59], cf. [60] for a review.

The RG treatment presented here is not focused on computing universal properties near criticality. It is rather motivated by the divergent flows of the chemical potential $\mu$ and of the interaction coupling $g$ described in [1], which were obtained from a Taylor expansion with respect to the slowly varying modes of the one-loop effective action in leading-order derivative expansion, i.e., in local-potential approximation. In fact, these divergences can be traced to a violation of the U(1) symmetry at a given order of approximation, and imposing the relevant Ward identities to be satisfied can be avoided [40, 41]. At order of approximation chosen in [1], the RG flow does not account for the system entering the linear, phononic regime of the Bogoliubov dispersion, such that finite coupling flows are achieved only within the short-wave-length regime $\Lambda_{\rm th}^2 \ll (na_0)^{-1}$, where the linear part of the dispersion relation is negligible. Here, $n$ is the particle density, $a_0$ the $s$-wave scattering length and $\Lambda_{\rm th}^2 = 2\pi/(mT)$ the thermal de Broglie wavelength.

So far, the description of linear, phononic excitations has been achieved within a functional RG approach both at $T = 0$ [30, 42–46] and $T \neq 0$ [30, 31, 47–49], within a next-to-leading order gradient expansion $O(\nabla^2, \partial_\tau, \partial_\tau^2)$ of the effective action. The vanishing anomalous self energy has been linked to the transition to an effective SO($d+1$) space-time symmetry in the IR [42, 43]. Approaching the low temperature limit from the thermal phase transition [30, 31, 47] the thermodynamic bulk quantities approach the Bogoliubov expressions in the limit of low temperatures [31, 49]. Beyond the above approaches, using an interpolation between the cartesian and amplitude-phase representations of the Bose field, converging flows and the hydrodynamic effective action valid in the IR were determined in [48, 49].

Extending upon [1] and using the framework developed by Wilson [21–24], we here derive perturbative, one-loop RG flow equations at next-to-leading order in the derivative expansion, that take into account renormalization of the first- and second-order temporal derivative terms in the action functional, neglecting, for the first, the renormalization of the kinetic energy term. This allows us to obtain a description of the transition of a dilute Bose gas in $d = 3$ dimensions to the IR phononic regime with a linear sound-wave dispersion. In this broken-symmetry regime, we make use of the Hugenholtz-Pines relation, which results from a Ward identity reflecting the U(1) symmetry of the model. In this respect, our approach is similar to that of [40, 41] for a $T = 0$ system. As we remain within the perturbative, one-loop approximation and neglect wave-function renormalization of the spatial derivative terms, the scope of our analysis can not be to improve upon precision beyond results from numerical analyses or functional RG approaches, but to present a simple perturbative RG approach to sound excitations in a Bose condensate. Besides giving results consistent with the Bogoliubov-de Gennes theory, this approach can be considered as a first step towards treating more complex problems such as non-universal as well as universal properties in a multi-component dilute Bose gas [61, 62].

Our paper is organized as follows. In Sect. II, we introduce the Wilsonian effective action of the Gross-Pitaevskii model in next-to-leading order gradient expansion, for which we derive, in Sect. III, the RG flow equations both, in the symmetric, 'thermal', and the symmetry-broken, condensate phases. Our numerical results are presented in Sect. IV, including a comparison of the flows with and without wave-function renormalization, and a determination of the depletion of the condensate and the interaction-induced shift of the critical temperature. We summarize and draw our conclusions in Sect. V.

## II. WEAKLY INTERACTING BOSE GAS

### A. Gross-Pitaevskii model

A single-component gas of weakly interacting bosons of mass $m$, exposed to an external (e.g. trapping) potential $V_{\rm ext}$, is described by a complex field $\Psi(\mathbf{x})$, the classical dynamics of which is governed by the Gross-Pitaevskii (GP) Hamiltonian

$$H = \int \mathrm{d}^d x \left( -\Psi(\mathbf{x})^* \frac{\nabla^2}{2m} \Psi(\mathbf{x}) + V_{\rm ext}(\mathbf{x}) |\Psi(\mathbf{x})|^2 + \frac{g}{2} |\Psi(\mathbf{x})|^4 \right). \tag{1}$$

Throughout this paper we use natural units where $\hbar = k_{\rm B} = 1$. The quartic coupling term accounts for two-body scattering, which can be described, at low energies, as a contact interaction. In three spatial dimensions, $d = 3$, the coupling $g$ is related to the $s$-wave scattering length $a_0$ by $g = 4\pi a_0/m$. Genuine three-body and higher interactions can be neglected in the dilute limit, sufficiently far from a Feshbach resonance, as their effect on the later renormalization is negligible [1, 63]. Since in ultracold gases, the total particle number in a closed volume $V$,

$$N = \int_V \mathrm{d}^d x \, |\Psi(\mathbf{x})|^2 = \int \mathrm{d}^d x \, n(\mathbf{x}) \,, \tag{2}$$

is conserved, one usually resorts to a grand canonical ensemble, in which the chemical potential $\mu$ acts as a Lagrange multiplier in the 'Kamiltonian' $K = H - \mu N$. Within the Bogoliubov mean-field approach, the classical field configuration in the ground-state of (1) is then determined, in the absence of an external trapping potential, by the stationary GP equation for the field expectation value $\psi = \langle \Psi \rangle$:

$$\left( -\frac{\nabla^2}{2m} - \mu + g \, |\psi(\mathbf{x})|^2 \right) \psi(\mathbf{x}) = 0 \,. \tag{3}$$

Assuming positive $g$, i.e., repulsive interactions, the mean-field ground state, at non-zero density $n = |\psi|^2$, is uniform, $n(\mathbf{x}) \equiv n_0 = n_{\rm c} = \text{const}$, and thus the chemical potential $\mu$, in the condensed phase, becomes a positive energy, resulting, in this mean-field approximation, as $\mu = g n_{\rm c}$. The density enters the non-vanishing field expectation value $\psi = \sqrt{n_{\rm c}}\, \mathrm{e}^{\mathrm{i}\phi}$, with a constant but irrelevant phase angle $\phi$, which constitutes spontaneous breaking of the U(1) symmetry of the model Hamiltonian (1).

### B. Effective action in next-to-leading order gradient expansion

In view of the renormalization group (RG) approach to the quantized system in thermal equilibrium, one introduces the (Wilsonian) effective action $S$,

$$S[\Psi] = \int_x \left( \Psi^*(Z_\tau \partial_\tau - V\partial_\tau^2 - Z_x \frac{\nabla^2}{2m} - \mu)\Psi + \frac{g}{2}|\Psi|^4 \right), \quad (4)$$

which follows from the coherent-state path integral of (1). The fluctuating fields also depend on the imaginary time $\tau$, which is integrated, within $\int_x \equiv \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^d x$, from 0 to $\tau = \beta = 1/T$, with periodic boundary conditions imposed along the time direction on the bosonic fields, $\Psi(0, \mathbf{x}) = \Psi(\beta, \mathbf{x})$. As a consequence, the action can be expanded in terms of discrete Matsubara frequency modes $\omega_n = 2\pi n T$, cf. (A1) in App. A. The Gaussian part of (4) yields

$$S_0[\Psi_k] = \sum\!\!\!\!\!\!\int_k \Psi_k^*( -\mathrm{i}\, Z_\tau \omega_n + V\omega_n^2 + Z_x \varepsilon_k - \mu)\Psi_k \,, \quad (5)$$

where the shorthands $\sum\!\!\!\!\!\!\int_k = \beta^{-1} \sum_{\omega_n} \int \mathrm{d}^d k\,/(2\pi)^d$ and $\Psi(\omega_n, \mathbf{k}) = \Psi(k) \equiv \Psi_k$ are used and the single-particle energy $\varepsilon_k = \mathbf{k}^2/(2m)$ is introduced. For the interaction part of (4) one finds

$$S_{\mathrm{int}}[\Psi_k] = \frac{g}{2} \sum\!\!\!\!\!\!\int_{k_i} \delta(k_2 + k_4 - k_1 - k_3)\, \Psi^*_{k_1} \Psi_{k_2} \Psi^*_{k_3} \Psi_{k_4} \,, \quad (6)$$

where the $\delta$-distribution $\delta(k) = (2\pi)^d \beta\, \delta_{0,\omega_n} \delta(\mathbf{k})$ ensures energy-momentum conservation.

In analogy to existing schemes, the Berry-phase term $\sim \Psi^* \partial_\tau \Psi$ as well as the kinetic energy receive couplings $Z_\tau$ and $Z_x$, respectively, commonly termed wave-function renormalization factors. Within the gradient expansion, we furthermore introduce a second-order time derivative term $\sim \Psi^* \partial_\tau^2 \Psi$ with the coupling $V$. The RG flow is initialized by the leading order action in the gradient expansion, where the couplings take their microscopic values in the UV, i.e., $Z_{\tau,\mathrm{in}} = 1$, $V_{\mathrm{in}} = 0$, and $Z_{x,\mathrm{in}} = 1$.

While renormalization gives rise to and alters all couplings that are allowed by symmetry, our choice is motivated by the aim to explicitly account for a linear sound-wave dispersion in the low-temperature condensate. In this regime, the first-order time derivative gets replaced by a second-order one, i.e., $Z_\tau \to 0$, while $V$ grows with the RG flow; and, neglecting residual sound-wave interactions, the system assumes an approximate Euclidean SO($d+1$) symmetry in space-time, as previously described in the functional RG framework [30, 42, 43]. The coupling $Z_x$, away from criticality in $d \geq 2$ dimensions, is anticipated to neither qualitatively nor quantitatively modify our results [30, 49]. Nonetheless, it will be kept in our formulation for bookkeeping purposes as it gets rescaled during the flow, cf. Sect. III C.

### C. Symmetry-broken phase

Before we move on to the WRG formulation, we introduce and discuss a parametrization of the effective action, which takes into account spontaneous symmetry breaking in the Bose-condensed phase. The quasiparticle excitations can be approximated in a Gaussian fashion by means of a saddle-point expansion of the action (4). In Bogoliubov theory, one expresses the field in terms of fluctuations around its expectation value, $\Psi(x) = \sqrt{n_\mathrm{c}} + \delta\Psi(x)$, and only retains terms up to quadratic order and solves the resulting action exactly. Setting, for the first, $V = 0$ and $Z_x = Z_\tau = 1$, one finds an excitation spectrum with eigenfrequencies of gapless Bogoliubov modes $\omega_k = \sqrt{\varepsilon_k(\varepsilon_k + 2gn_\mathrm{c})}$ that also follows directly from a linearization of the GP equation (3). Taking into account the possibility of $V \neq 0$, a second Bogoliubov mode emerges, which will be discussed in the next section.

Within the WRG formulation, one similarly needs to break the U(1) symmetry by hand, choosing an ansatz which assumes the presence of a condensate. The U(1) symmetry of the action (4) is broken by splitting the field $\Psi_k = \delta(k)\,\psi_\mathrm{c} + \psi_k$ into its condensate part $\psi_\mathrm{c}$ and thermal fluctuations $\psi_k \equiv \psi(k)$ with vanishing expectation values $\langle \psi_k \rangle = 0$. Without external potential, the condensate field is just a uniform background $\psi_\mathrm{c} = \sqrt{n_\mathrm{c}}$, with $n_\mathrm{c}$ being the condensate density. As before, we dropped the irrelevant phase of the condensate. This expansion gives rise to a modified Gaussian part of the action,

$$\widetilde{S}_0 = \sum\!\!\!\!\!\!\int_k \Big[ \psi_k^* \left( -\mathrm{i} Z_\tau \omega_n + V\omega_n^2 + Z_x \varepsilon_k + \mu_1 \right) \psi_k + \frac{\mu_2}{2} \left( \psi_k \psi_{-k} + \psi_k^* \psi_{-k}^* \right) \Big] + \sqrt{n_\mathrm{c}}\,(n_\mathrm{c} g - \mu) \left( \psi_0 + \psi_0^* \right) . \quad (7)$$

We have also dropped constant terms as they correspond to irrelevant shifts of the energy zero. To satisfy the constraint $\langle \psi_k \rangle = 0$ for any $k \geq 0$, the linear term must vanish. This yields the basic mean-field definition of the chemical potential $\mu = n_\mathrm{c} g$, which will remain valid under the renormalization procedure.

Through the lens of renormalization, couplings are defined for each field term; thus, two generalized two-point couplings $\mu_1$ and $\mu_2$, with initial values

$$\mu_{1,\mathrm{in}} = 2n_\mathrm{c} g - \mu \,, \qquad \mu_{2,\mathrm{in}} = n_\mathrm{c} g \,, \quad (8)$$

appear in the quadratic part of the action. These couplings are equal initially, $\mu_{1,\mathrm{in}} = \mu_{2,\mathrm{in}}$, and remain so throughout the RG flow, thereby ensuring a gapless dispersion relation in line with Goldstone's and, equivalently, the Hugenholtz-Pines theorem [13]. This will need to be proven explicitly (cf. Sect. III B 3 below), before which we treat them as in general independent. The terms proportional to $\mu_2$ in (7) give rise to non-vanishing anomalous correlators $\langle \psi\psi \rangle_0$.

In the symmetry-broken phase, the interacting part of the action results as

$$\widetilde{S}_{\mathrm{int}} = \sqrt{n_\mathrm{c}} g \sum\!\!\!\!\!\!\int_{k_1,k_2,k_3} \delta(k_1 - k_2 - k_3)(\psi_1^* \psi_2 \psi_3 + \text{ c.c. }) + \frac{g}{2} \sum\!\!\!\!\!\!\int_{k_1,k_2,k_3,k_4} \delta(k_2 + k_4 - k_1 - k_3) \psi_1^* \psi_2 \psi_3^* \psi_4 \,. \quad (9)$$

Hence, three-point interactions emerge between one condensed and three thermal modes. This term is responsible for the appearance of wave-function renormalization at the one-loop level as discussed in the following.

## III. WILSONIAN RG FLOWS

In this section, we summarize the WRG approach [21–24] as applied to Bose gases described by the action (4) and extend it to the symmetry-broken regime, cf. (7), (9). We derive the flow equations in both the thermal and the condensed phase up to one-loop order. Wave-function renormalization in terms of $Z_\tau$ and $V$ is included straight from the beginning. Our notation follows Ref. [64].

The procedure of WRG starts with defining the action $S[\mathbf{g}]$ on a range of momenta $\lambda_0 \leq |\mathbf{k}| \leq \Lambda_0$. The IR cut-off $\lambda_0$ refers, e.g., to the inverse linear size of the considered system. We can, in principle, take it to $\lambda_0 \to 0$, which we will do during the numerical evaluation of the RG equations.

The UV cut-off $\Lambda_0$ reflects the range of validity of the action on microscopic scales. To ensure the applicability of the $s$-wave interactions we choose it to be $\Lambda_0 \lesssim 1/a_0$. Note that, in the following, we are interested in the effective renormalization of the action to obtain a low-energy effective theory of the thermodynamic properties of a dilute ultracold Bose gas in the symmetry-broken, Bose-Einstein condensed phase. Hence, we will use WRG to study the flow of the effective couplings to momentum scales farther in the IR than $\Lambda_0$, rather than studying their renormalization in the farther UV. While the latter is used to express the couplings in terms of the microscopic scattering properties of, e.g., the alkali atoms, our aim here is to determine the effective couplings and field correlations as functions of thermodynamic bulk quantities such as particle density and temperature.

The Wilsonian renormalization flow is the result of many infinitesimal steps each consisting of three sub-steps:

(A) The first one is mode elimination, in which an infinitesimal momentum shell $\Lambda \leq |\mathbf{k}| \leq \Lambda_0$, with

$$b = \Lambda_0/\Lambda \to 1^+ , \tag{10}$$

is integrated out from the partition function. The effective couplings $\mathbf{g} = \left(Z_x, Z_\tau, V, \mu, g\right)$ defining the action on the remaining momentum volume $\lambda_0 \leq |\mathbf{k}| \leq \Lambda$ are thereby being modified.

(B) In the subsequent rescaling step, the renormalized couplings of the new effective action defined on the remaining momentum range are obtained by rescaling all couplings by powers of $b$. This is necessary to arrive at a new, coarse-grained effective description of the system valid at correspondingly larger length scales.

(C) In the final step, one combines both, steps (A) and (B) to update the couplings.

The three steps are then repeated in an iterative manner to determine a flow of the couplings with the continuously decreasing cut-off scale $\Lambda$.

### A. Mode elimination: symmetric phase

In the following, the above steps are described in more detail and specific to the Bose gas in the symmetric, 'thermal' phase. The mode elimination step starts with splitting the field $\Psi_k$ into a long-wavelength (IR) $\Psi_k^<$ and a short-wavelength (UV) $\Psi_k^>$ part:

$$\Psi_k = \Psi_k^< + \Psi_k^> = \Theta(\Lambda - |\mathbf{k}|)\,\Psi_k + \Theta(|\mathbf{k}| - \Lambda)\,\Psi_k . \tag{11}$$

They are separated by the cut-off $\Lambda$ above which the short wavelengths will be subsequently integrated out. Inserting this into the action $S[\Psi_k]$ one derives a part below the cut-off $\Lambda$, $S^<[\Psi_k^<]$, one above the cut-off, $S^>[\Psi_k^>]$, as well as a mixed term $S_{\text{mix}}^{<>}[\Psi_k^<, \Psi_k^>]$. Due to the lacking overlap, $\Psi_k^<\Psi_k^> = 0$, the Gaussian part of the action $S_0$ splits into a sum $S_0 = S_0^< + S_0^>$. Only the interaction part contains mixing of the IR and UV scales,

$$\begin{aligned} S_{\text{int}} = S_{\text{int}}^< + S_{\text{int}}^> + g \sum_{k_i}\!\!\!\!\!\!\int \delta(k_2 + k_4 - k_1 - k_3) \\ \times \Big(\Psi_1^{*<}\Psi_2^<\Psi_3^{*<}\Psi_4^> + \frac{1}{2}\Psi_1^{*<}\Psi_2^>\Psi_3^{*<}\Psi_4^> \\ + \Psi_1^{*<}\Psi_2^<\Psi_3^{*>}\Psi_4^> + \Psi_1^{*<}\Psi_2^>\Psi_3^{*>}\Psi_4^> + \text{ c.c.}\Big), \end{aligned} \tag{12}$$

where we use the shorthand $\Psi_{k_i} \equiv \Psi_i$. In the next step, the UV modes are integrated out explicitly from the partition function, which determines all thermodynamic bulk quantities:

$$\mathcal{Z} = \int \mathcal{D}\Psi^*\mathcal{D}\Psi\, \mathrm{e}^{-S[\Psi;\mathbf{g}]} = \int \mathcal{D}\Psi^{*<}\mathcal{D}\Psi^<\, \mathrm{e}^{-S^<[\Psi^<;\mathbf{g}^<]} . \tag{13}$$

We thus rewrite the partition function in terms of the macroscopic IR fields $\Psi^<$ only. The coarse-grained action $S^<[\Psi^<;\mathbf{g}^<]$ entering the partition function after mode elimination hence only depends on the modes $|\mathbf{k}| < \Lambda$. Since the effective action we aim to derive must exhibit the same thermodynamic behavior as the microscopic action, the partition function must not alter, and therefore the coupling constants $\mathbf{g}^<$ are suitably modified as compared to the original ones. While, so far, all the steps have in principle been exact, to determine the modified couplings one needs to solve the path integral in the UV over an action that contains higher-order than quadratic terms in $\Psi$ and $\Psi^*$. Since such integrals can in general not be computed analytically, we will make use of perturbation theory and write $S^<[\Psi^<;\mathbf{g}^<]$ as an expansion around $S_0^>[\Psi^>;\mathbf{g}]$ to second-order, which corresponds to a one-loop approximation:

$$\begin{aligned} S^<[\mathbf{g}^<] = S^<[\mathbf{g}] + \left\langle S_{\text{mix}}^{<>}[\mathbf{g}] + S_{\text{int}}^>[\mathbf{g}]\right\rangle_{0,>} \\ + \frac{1}{2}\left[\left\langle S_{\text{mix}}^{<>}[\mathbf{g}] + S_{\text{int}}^>[\mathbf{g}]\right\rangle_{0,>}^2 - \left\langle\left(S_{\text{mix}}^{<>}[\mathbf{g}] + S_{\text{int}}^>[\mathbf{g}]\right)^2\right\rangle_{0,>}\right] . \end{aligned} \tag{14}$$

Here, a constant term has been dropped, which corresponds to an irrelevant shift of the energy zero. $\langle O\rangle_{0,>}$ denotes an expectation value with respect to the UV free action $S_0^>[\mathbf{g}]$, and we have dropped the explicit notion of the field arguments

$\Psi^{\gtrless}$. To determine $\mathbf{g}^<$, the expectation values in (14) need to be expressed in powers of the IR field. The explicit change can be found after evaluating the free two-point correlators of $\Psi_k^>$. Note that the structure of the quadratic terms in (14) is such that only connected Feynman diagrams contribute to $\mathbf{g}^<$ [25]. This is reminiscent of the expansion of the free energy in terms of connected diagrams.

Free correlators are in general obtained by functional derivatives,

$$\left\langle \Psi_k^* \Psi_{k'} \right\rangle_0 = \frac{\delta}{\delta J_k} \frac{\delta}{\delta J_{k'}^*} \mathcal{Z}_0[J] \Big|_{J=0} , \tag{15}$$

of the Gaussian generating functional

$$\begin{aligned} \mathcal{Z}_0[J] &= \int \mathcal{D}\Psi^* \mathcal{D}\Psi \exp\left( -S_0[\mathbf{g}] + \sum\!\!\!\!\!\!\int_k \left[ J_k^* \Psi_k + \text{ c.c.} \right] \right) \\ &= \exp\left( \sum\!\!\!\!\!\!\int_k \mathcal{J}_{-k}^T \mathcal{M}_k^{-1} \mathcal{J}_k \right) \end{aligned} \tag{16}$$

with respect to the current $J_k$. The vectorial current $\mathcal{J}_k^T$ is defined in terms of the real and imaginary parts of the current expressed in spatial coordinates, i.e.,

$$\mathcal{J}_k^T = \left( [J_k + J_{-k}^*]/2, \ [J_k - J_{-k}^*]/2\mathrm{i} \right) , \tag{17}$$

and the corresponding kernel matrix $\mathcal{M}_k$ is given by

$$\mathcal{M}_k = \begin{pmatrix} Z_x \varepsilon_k + V\omega_n^2 - \mu & -Z_\tau \omega_n \\ Z_\tau \omega_n & Z_x \varepsilon_k + V\omega_n^2 - \mu \end{pmatrix} . \tag{18}$$

The two-point normal and anomalous correlators are found using its inverse,

$$\mathcal{M}_k^{-1} = \frac{1}{\det(\mathcal{M}_k)} \begin{pmatrix} Z_x \varepsilon_k + V\omega_n^2 - \mu & Z_\tau \omega_n \\ -Z_\tau \omega_n & Z_x \varepsilon_k + V\omega_n^2 - \mu \end{pmatrix} , \tag{19}$$

and taking functional derivatives (15) of $\mathcal{Z}_0[J]$:

$$\left\langle \Psi_k^* \Psi_{k'} \right\rangle_0 = \frac{\delta(k-k')}{Z_x \varepsilon_k + V\omega_n^2 - \mu + \mathrm{i} Z_\tau \omega_n} \equiv \delta(k-k') \, G(k) , \tag{20}$$

$$\langle \Psi_k \Psi_{k'} \rangle_0 = 0 . \tag{21}$$

Eq. (21) reflects the U(1) symmetry of the action (4), which ensures particle number conservation in the thermal phase. All higher-order expectation values in (14) can be decomposed by means of Wick's theorem into products of the two-point correlators (20).

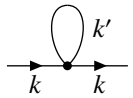
In the following, we express the expansion (14) in Feynman diagrammatic form in terms of the free normal correlator $G(k)$, (20), and couplings $\mathbf{g}$. The expectation value $\left\langle S_{\text{int}}^>[\mathbf{g}] \right\rangle_{0,>}$ does not contain macroscopic fields $\Psi_k^<$ and thus consists of vacuum diagrams which only shift the energy zero. Furthermore, since all disconnected Feynman diagrams cancel in (14), we can discard $\left\langle S_{\text{mix}}^{<>}[\mathbf{g}] + S_{\text{int}}^>[\mathbf{g}] \right\rangle_{0,>}^2$.

As we neglect all two-loop diagrams, the remaining contributions to be evaluated are $\left\langle S_{\text{mix}}^{<>}[\mathbf{g}] \right\rangle_{0,>}$ and $\left\langle \left( S_{\text{mix}}^{<>}[\mathbf{g}] \right)^2 \right\rangle_{0,>}$.

Also diagrams contributing to the renormalization of the six-point coupling, i.e., three-particle scattering, are neglected since we are in the weakly-interacting, dilute regime.

The first of these terms gives the tadpole correction

$$\begin{aligned} \left\langle S_{\text{mix}}^{<>}[\mathbf{g}] \right\rangle_{0,>} &= 2g \sum\!\!\!\!\!\!\int_{\substack{k,k' \\ \Lambda<|\mathbf{k}'|<\Lambda_0}} \Psi_k^{*<} G(k') \Psi_k^< \\ &\propto \sum\!\!\!\!\!\!\int_{\substack{k,k' \\ |\mathbf{k}|<\Lambda<|\mathbf{k}'|<\Lambda_0}} \end{aligned} \tag{22}$$

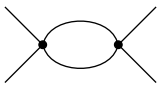



Note that, in the first integral, we do not need to explicitly restrict the range of momenta $k$, as this is achieved by the $\Theta$-functions in $\Psi_k^<$, cf. (11). It is worth mentioning, that this diagram does not contain contributions to wave-function renormalization, since the correlator $G(k')$ is independent of the external momentum $k$. Hence, this term does only renormalize the coupling attributed to $\Psi_k^{*<}\Psi_k^<$, i.e., the chemical potential $\mu$. Generally speaking, at one-loop order in the symmetric phase no wave-function renormalization in $Z_\tau$ and $V$ emerges. Using the mode elimination equation (14), one can directly read off the new chemical potential entering $S^<[\mathbf{g}^<]$:

$$\mu^< = \mu - 2g \sum\!\!\!\!\!\!\int_{\substack{k \\ \Lambda<|\mathbf{k}|<\Lambda_0}} G(k) . \tag{23}$$

The remaining expectation value in (14) contributes to the renormalization of the two-particle scattering:

$$\left\langle \left( S_{\text{mix}}^{<>}[\mathbf{g}] \right)^2 \right\rangle_{0,>} \propto$$

 . (24)

Labeling the external momenta in (24) by $k_1, \ldots, k_4$ we obtain the integral

$$\begin{aligned} \left\langle \left( S_{\text{mix}}^{<>}[\mathbf{g}] \right)^2 \right\rangle_{0,>} = \sum\!\!\!\!\!\!\int_{\substack{k_i,k \\ \Lambda<|\mathbf{k}|<\Lambda_0}} & \delta(k_2 + k_4 - k_1 - k_3) \Psi_1^{*<} \Psi_2^< \Psi_3^{*<} \Psi_4^< \\ & \times g^2 \, G(k) \left[ 4\, G(k_4 - k_3 + k) + G(k_4 + k_2 - k) \right] . \end{aligned} \tag{25}$$

Note that, here, loop momentum depends on the external momenta, which generates momentum- and frequency-dependent interaction terms such as $\mathbf{k}^2\Psi^4$ and $\omega_n \Psi^4$. However, these couplings are neglected as they exhibit irrelevant scaling according to their engineering dimensions (cf. Sect. III C) and we thus only take the zeroth-order Taylor expansion around $k$ into account. Inserting (25) into (14) one reads off the new four-point coupling

$$g^< = g - \frac{g^2}{2} \sum\!\!\!\!\!\!\int_{\substack{k \\ \Lambda<|\mathbf{k}|<\Lambda_0}} \left[ 4\, G(k)\, G(k) + G(k)\, G(-k) \right] . \tag{26}$$

The opposite-sign momenta in the last term correspond to the twofold scattering of two incoming particles with each other.

### B. Mode elimination: symmetry-broken phase

In this section, mode elimination is extended to the symmetry-broken, i.e., condensate phase. Mode splitting and integrating out the momentum shell between $\Lambda$ and $\Lambda_0$ proceeds as before, but now using the ansatz (7) and (9) for the effective action in evaluating $S^<[\mathbf{g}^<]$ as given in (14).

#### 1. Normal and anomalous correlators

We begin again by determining the free correlator (15), as well as the anomalous one,

$$\langle\psi_k\psi_{k'}\rangle_0 = \frac{\delta}{\delta J_k^*}\frac{\delta}{\delta J_{k'}^*}\exp\left(\sum\!\!\!\!\!\!\int_k \mathcal{J}_{-k}^T\mathcal{M}_k^{-1}\mathcal{J}_k\right)\Bigg|_{\mathcal{J}=0}. \tag{27}$$

In the symmetry-broken phase, the kernel matrix $\mathcal{M}_k$ takes the from

$$\mathcal{M}_k = \begin{pmatrix} Z_x\varepsilon_k + V\omega_n^2 + \mu_1 + \mu_2 & -Z_\tau\omega_n \\ Z_\tau\omega_n & Z_x\varepsilon_k + V\omega_n^2 + \mu_1 - \mu_2 \end{pmatrix}. \tag{28}$$

The diagonals of the normal, $\langle\psi_k^*\psi_{k'}\rangle_0 = \delta(k-k')\,G_k$, and anomalous correlator, $\langle\psi_k\psi_{k'}\rangle_0 = \delta(k+k')\,\widetilde{G}_k$, result as

$$\begin{aligned} G_k &= \frac{Z_x\varepsilon_k + \mu_1 + \mathrm{i}Z_\tau\omega_n + V\omega_n^2}{V^2\omega_n^4 + A_k\omega_n^2 + \omega_k^2}, \quad \text{and} \\ \widetilde{G}_k &= -\frac{\mu_2}{V^2\omega_n^4 + A_k\omega_n^2 + \omega_k^2}, \end{aligned} \tag{29}$$

respectively. Here, we introduced $A_k = Z_\tau^2 + 2V(Z_x\varepsilon_k + \mu_1)$ and the dispersion relation at $V = 0$:

$$\omega_k^2 = (Z_x\varepsilon_k + \mu_1)^2 - (\mu_2)^2\,. \tag{30}$$

In the case of $\mu_1 = \mu_2$ this reduces to the well-known gapless Bogoliubov dispersion. If $V \neq 0$, factorizing the denominator into $V^2[\omega_n^2 + (\omega_k^+)^2][\omega_n^2 + (\omega_k^-)^2]$ results in the quasiparticle frequencies as the poles of the correlators,

$$(\omega_k^\pm)^2 = \frac{A_k \pm \sqrt{A_k^2 - 4V^2\omega_k^2}}{2V^2}\,. \tag{31}$$

In the limit $V \to 0$ one has $(\omega_k^-)^2 \to \omega_k^2/Z_\tau^2$ and $V^2(\omega_k^+)^2 \to Z_\tau^2$, and with the initial $Z_\tau = 1$ one recovers (30). The gapped dispersion relation $\omega_k^+$ is initially decoupled at $V = 0$.

The asymptotic behavior for $V \gg 1$ of $(\omega_k^+)^2 \to (Z_x\varepsilon_k + \mu_1 + \mu_2)/V$ and $(\omega_k^-)^2 \to (Z_x\varepsilon_k + \mu_1 - \mu_2)/V$ yields two relativistic dispersion relations. The same result is found for $Z_\tau \to 0$, which makes the action (4) quasi-relativistic and is equivalent to $V \gg 1$.

Hence, for $\mu_i \neq 0$, the dispersion relations describe the transition from the Bogoliubov dispersion at UV scales, $V \ll 1$, into the phononic regime, with a gapped and, for $\mu_1 = \mu_2$, a gapless mode in the $k \to 0$ limit.

#### 2. Mode decomposition of the action

As in the symmetric phase, we need to compute the expectation values in (14). However, a bigger variety of Feynman diagrams appears owing to the anomalous correlator and the three-point vertex. Splitting the fields into $\psi = \psi^< + \psi^>$ leads, again, to $\widetilde{S}_0 = \widetilde{S}_0^< + \widetilde{S}_0^>$. The quartic interactions $\widetilde{S}_{\mathrm{int}}^{(4)}$ decompose as in the thermal phase (12), while the cubic interactions decompose according to

$$\begin{aligned} \widetilde{S}_{\mathrm{int}}^{(3)} = \sqrt{n_{\mathrm{c}}}g \sum\!\!\!\!\!\!\int_{k_1,k_2,k_3} \delta(k_1 - k_2 - k_3)\Big[2\psi_1^{*<}\psi_2^>\psi_3^< + 2\psi_1^{*>}\psi_2^>\psi_3^< \\ + \psi_1^{*>}\psi_2^<\psi_3^< + \psi_1^{*<}\psi_2^>\psi_3^> + \text{ c.c. }\Big] + \widetilde{S}_{\mathrm{int}}^{(3)<} + \widetilde{S}_{\mathrm{int}}^{(3)>}\,. \end{aligned} \tag{32}$$

Collecting all mixed contributions from (12) and (32) yields $\widetilde{S}_{\mathrm{mix}}^{<>}$. Its expectation value reads

$$\begin{aligned} \left\langle \widetilde{S}_{\mathrm{mix}}^{<>}\right\rangle_{0,>} = \frac{g}{2} \sum\!\!\!\!\!\!\int_{\substack{k,k' \\ \Lambda<|\mathbf{k}'|<\Lambda_0}} \left[4\psi_k^{*<}\psi_k^< G_{k'} + (\psi_k^<\psi_{-k}^< + \psi_k^{*<}\psi_{-k}^{*<})\widetilde{G}_{k'}\right] \\ + \sqrt{n_{\mathrm{c}}}g\left(\psi_0^< + \psi_0^{*<}\right) \sum\!\!\!\!\!\!\int_{\substack{k \\ \Lambda<|\mathbf{k}|<\Lambda_0}} \left(2G_k + \widetilde{G}_k\right). \end{aligned} \tag{33}$$

In contrast to the situation in the thermal phase, the above expectation value does not suffice in determining the change in the two-point couplings because additional two-vertex diagrams from the second-order term must be taken into account. The above result does not feature wave-function renormalization since all loop correlators are independent of the external momentum, as it was the case in (22).

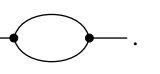
Note furthermore that the last term in (33) represents a tadpole contribution, which introduces a term linear in the fields to the action,

$$S_d[h] \propto h(\psi_0^< + \psi_0^{*<})\,. \tag{34}$$

The coupling constant $h$ vanishes in the broken-symmetry action (7) where $\mu = gn_{\mathrm{c}} = gn$ is chosen initially. A renormalization of $h$ will account for a depletion of the condensate density due to zero-point and thermal excitations.

When calculating $\langle(\widetilde{S}_{\mathrm{mix}}^{<>})^2\rangle_{0,>}$, all disconnected diagrams can be dropped, since, as before, they mutually cancel in (14). We keep restricting our approximation to one-loop diagrams and up to quartic interactions, i.e., two-particle scattering. Following the approach employed in Ref. [1], we can restrict ourselves to determining the change in the two-point couplings, while the four-point coupling can be obtained from the constraint $\mu = n_{\mathrm{c}}g$. This allows us to only keep terms that are quadratic in the IR fields, i.e., diagrams with two external legs:

$$\left\langle\left(\widetilde{S}_{\mathrm{mix}}^{<>}\right)^2\right\rangle_{0,>} = \left\langle\left(\widetilde{S}_{\mathrm{mix}}^{(3)<>}\right)^2\right\rangle_{0,>} \propto$$



. (35)

Here, energy-momentum conservation causes the loop momentum to depend on the external momentum. Hence, in the condensed phase, wave-function renormalization emerges already at one-loop order. One finds

$$\frac{1}{2}\left\langle\left(\widetilde{S}^{(3)<>}_{\text{mix}}\right)^2\right\rangle_{0,>} = n_c g^2 \sum_{\substack{k,k' \\ \Lambda<|\mathbf{k}'|<\Lambda_0}}\Big\{\psi^<_k\psi^{*<}_k\left[4\widetilde{G}_{k'}\widetilde{G}_{k'-k}+2G_{k'}\left(4\widetilde{G}_{k'-k}+2G_{k'-k}+G_{k-k'}\right)\right] + \left(\psi^<_k\psi^<_{-k}+\psi^{*<}_k\psi^{*<}_{-k}\right)\left(2G_{k'-k}G_{k'}+3\widetilde{G}_{k'+k}\widetilde{G}_{k'}+4G_{k'}\widetilde{G}_{k+k'}\right)\Big\}. \tag{36}$$

### 3. *Coupling renormalization*

We begin by considering the change in the local-in-momentum part of the quadratic part of the action. To that end, we Taylor expand the expressions (33) and (35) to the zeroth order in $k$. Terms beyond this order will be needed for the flow equations of the anomalous couplings $Z_\tau$ and $V$, which we will derive in III B 5 below.

Once both expectation values, (33) and (36), have been evaluated, we can read off $\delta h$ and $\delta\mu_i$. For this, one inserts these expectation values into (14) and, e.g., for $\mu_1$, collects all terms proportional to $\psi^<_k\psi^{*<}_k$, which then represent the modification $\delta\mu_1$. Doing so yields

$$\begin{aligned}
\delta h &= \sqrt{n_c}g \sum_{\substack{k \\ \Lambda<|\mathbf{k}|<\Lambda_0}}(2G_k+\widetilde{G}_k),\\
\delta\mu_2 &= \sum_{\substack{k \\ \Lambda<|\mathbf{k}|<\Lambda_0}}\left[g\widetilde{G}_k - n_c g^2\left(4G_kG_k+6\widetilde{G}_k\widetilde{G}_k+8G_k\widetilde{G}_k\right)\right],\\
\delta\mu_1 &= \sum_{\substack{k \\ \Lambda<|\mathbf{k}|<\Lambda_0}}\Big[2gG_k - n_c g^2\left(4G_kG_k+2G_kG_{-k}\right.\\
&\qquad\left.+8G_k\widetilde{G}_k+4\widetilde{G}_k\widetilde{G}_k\right)\Big].
\end{aligned}\tag{37}$$

The change in the coupling $h$ in the linear term (34) through mode elimination corresponds to a depletion of the condensate mode due to zero-point and thermal excitations [1]. This depletion thus appears as a shift of the zero-mode field and is accounted for by splitting the $<$ contribution of the field after mode elimination into a zero-mode contribution with non-vanishing expectation value and one for $k>0$, with vanishing mean, $\psi^<_k = \delta(k)\delta\sqrt{n_c}+\psi'^<_k$. Inserting this into (7) gives rise to a contribution of the linear term,

$$S_d = \left[\delta h+\delta\sqrt{n_c}\,(\mu_1+\mu_2)\right]\left(\psi'^<_0+\psi'^{*<}_0\right), \tag{38}$$

neglecting higher-order corrections in the infinitesimal shift $\delta\sqrt{n_c}$. As the action after mode elimination is to be expanded about the new minimum, this linear term must vanish, which fixes the condensate amplitude shift as

$$\delta\sqrt{n_c} = -\delta h/(\mu_1+\mu_2). \tag{39}$$

In the same fashion, the shift in the condensate amplitude gives an additional contribution to the quadratic couplings:

$$\begin{aligned}
\mu^<_2 &= \mu_2+\delta\mu_2+2\sqrt{n_c}g\,\delta\sqrt{n_c} = \mu_2+\Delta\mu_2,\\
\mu^<_1 &= \mu_1+\delta\mu_1+4\sqrt{n_c}g\,\delta\sqrt{n_c} = \mu_1+\Delta\mu_1.
\end{aligned}\tag{40}$$

As mentioned earlier, U(1) symmetry requires the two-point couplings $\mu_1$ and $\mu_2$ to be equal, to ensure a gapless quasiparticle mode frequency (30). This relation is a consequence of a U(1) Ward identity [65], which ensures the absence of approximation-related IR divergences in the course of the flow. As a result, also the 'hole' dispersion $\omega^-_k$ is gapless. To explicitly demonstrate that the identity $\mu_1=\mu_2$ holds at one-loop order, we compute the difference between the renormalized couplings using the relation (C16) between different Matsubara sums:

$$\begin{aligned}
\mu^<_1-\mu^<_2 = \mu_1-\mu_2+g\sum_{\substack{k \\ \Lambda<|\mathbf{k}|<\Lambda_0}}&\left[2G_k\left(1-\frac{2n_cg}{\mu_1+\mu_2}\right)\right.\\
&\left.+\widetilde{G}_k\left(\frac{2n_cg}{\mu_2}-\frac{2n_cg}{\mu_1+\mu_2}-1\right)\right].
\end{aligned}\tag{41}$$

Upon inserting the initial value $\mu_{2,\text{in}} = n_c g$ and using the initial equality of both two-point couplings (8) this expression vanishes, which proves the invariance of $\mu_1=\mu_2=\mu$ in the RG flow. The new chemical potential $\mu^<$ after mode elimination thus follows from (40) to be

$$\mu^< = \mu-2g\sum_{\substack{k \\ \Lambda<|\mathbf{k}|<\Lambda_0}}\left[G_k+\mu\left(2G_kG_k+3\widetilde{G}_k\widetilde{G}_k+4G_k\widetilde{G}_k\right)\right]. \tag{42}$$

Finally, the renormalized four-point coupling is obtained from the relation $\mu = n_c g$, which is equivalent to the Hugenholtz-Pines relation, i.e., the equality of the couplings, $\mu_1=\mu_2$, both initialized in (8), which holds throughout the renormalization flow as demonstrated above.

As a result, the new four-point coupling reads

$$g^< = g+\Delta g = g+\frac{\Delta\mu_2}{n_c}-\frac{\mu_2\delta n_c}{n_c^2} = g+\frac{\delta\mu_2}{n_c}. \tag{43}$$

Hence, while we determined the renormalization of the two-point couplings explicitly, we neglect all newly emerging coupling constants at the order of quartic terms in the field. This will later be relevant when considering the scaling behavior of $g$ which we thus merely infer from the renormalization of the chemical potential.

#### 4. Particle density

Besides the flow equations for both couplings, we need to determine the flow of the total particle density $n = n_c + n_T$, consisting of a condensate and a thermal fraction. The renormalized density is first expanded around the condensate density

$$\begin{aligned} n^< &= \frac{1}{\beta V} \sum_{|\mathbf{k}|<\Lambda}' \left\langle \Psi_k^* \Psi_k \right\rangle \\ &= n_c + \frac{\sqrt{n_c}}{\beta V} \left( \left\langle \psi_0^< \right\rangle + \left\langle \psi_0^{*<} \right\rangle \right) + \frac{1}{\beta V} \sum_{|\mathbf{k}|<\Lambda}' \left\langle \psi_k^{*<} \psi_k^< \right\rangle , \end{aligned} \tag{44}$$

using $\Psi_k = \delta(k)\sqrt{n_c} + \psi_k^<$. Since the condensate density changes under a mode elimination step, the field $\psi_k^<$ acquires a contribution in the zero mode with non-vanishing expectation value. In order to express (44) in terms of 'thermal' fields $\psi_k'$ with vanishing expectation value, we expand around the change in condensate density $\psi_k^< = \delta(k)\delta\sqrt{n_c} + \psi_k'$. This results in the shift of the field expectation value, as by definition $\left\langle \psi_k'^< \right\rangle = 0$ in all $k$ modes,

$$\left\langle \psi_k^< \right\rangle = \delta(k)\delta\sqrt{n_c} + \left\langle \psi_k'^< \right\rangle = \delta(k)\delta\sqrt{n_c} \,. \tag{45}$$

The change in the condensate amplitude $\delta\sqrt{n_c}$ is obtained from the shift of the coupling $h$ as obtained in (39). Thus, using $\delta(0) = \beta V$, the second term in (44) corresponds to the change in the condensate density. The third term in (44) is expanded to

$$\left\langle \psi_k^{*<} \psi_k^< \right\rangle = \left\langle \psi_k'^{*<} \psi_k'^< \right\rangle + \mathcal{O}\left( (\delta\sqrt{n_c})^2 \right) , \tag{46}$$

where we can neglect the terms of order $(\delta\sqrt{n_c})^2$. As a result, it yields the renormalized density of thermal particles within the remaining range of modes $k < \Lambda$,

$$n_T^< = \frac{1}{\beta V} \sum_{|\mathbf{k}|<\Lambda}' \left\langle \psi_k'^{*<} \psi_k'^< \right\rangle . \tag{47}$$

The change of the thermal particle density $\delta n_T = n_T^< - n_T$ can be obtained by performing the momentum shell integration

$$\delta n_T = \frac{1}{\beta V} \sum_{\Lambda<|\mathbf{k}|<\Lambda_0}' \left\langle \psi_k^{*>} \psi_k^> \right\rangle . \tag{48}$$

The expectation value $\left\langle \psi_k^{*>} \psi_k^> \right\rangle$ lacks any zero-mode contribution and accounts for an infinitesimal change of the total density. Here, we restrict ourselves to its Gaussian approximation, which is equivalent to the 1-loop approximation we chose in (14),

$$\left\langle \psi_k^{*>} \psi_k^> \right\rangle \simeq \left\langle \psi_k^{*>} \psi_k^> \right\rangle_0 = \delta(0)\,\Theta(|\mathbf{k}| - \Lambda)\, G_k \,. \tag{49}$$

Inserting the above results into (44) using $n_T^< = n_T + \delta n_T$ and (39), with $\mu_1 = \mu_2$, yields

$$n^< = n_c + 2\sqrt{n_c}\delta\sqrt{n_c} + n_T + \delta n_T = n - \sum_{\Lambda<|\mathbf{k}|<\Lambda_0}' \left( G_k + \widetilde{G}_k \right) . \tag{50}$$

The momentum integration ranges only from $\Lambda$ to $\Lambda_0$; hence, for $l = \ln(\Lambda_0/\Lambda) \to 0$ we obtain the classical expression for the total density.

#### 5. Wave-function renormalization

Wave-function renormalization appears in the symmetry-broken phase already at one-loop order due to the emerging three-point interaction in (9). This gives rise to loop terms in the effective action, of the form (35). As a result, renormalization of the derivative terms in the action appears and gives rise to new terms, at leading order $Z_\tau \partial_\tau$, $V\partial_\tau^2$ and $Z_x \nabla^2$.

In the following, we will account for $Z_\tau \neq 1$ and $V \neq 0$ only and neglect the renormalization of the spatial derivative as well as of derivative terms quadratic in the fields, $\sim \psi^2$ and $\sim \psi^{*2}$. The renormalization of the spatial derivative term, $Z_x\nabla^2$, is expected to remain at the percent level for temperatures sufficiently far away from criticality and thus does not play a role in the quasiparticle regime of sound excitations, where the action shows an approximate SO($d+1$) symmetry [30, 42, 49]. In the vicinity of the critical point, however, renormalization of $Z_x$ would be relevant and thus gives rise to a non-zero anomalous dimension $\eta_x$.

To derive the renormalization of $Z_\tau$ and $V$, we first extract the linear and quadratic contributions in Matsubara frequencies by performing a Taylor expansion of the correlator

$$G_{k'-k} = G_{k'} - \omega_n \dot{G}_{k'} + \frac{\omega_n^2}{2} \ddot{G}_{k'} + \mathcal{O}(\omega_n^3) , \tag{51}$$

with frequency derivatives written as

$$\dot{G}_{k'} = \left. \frac{\partial G_k}{\partial \omega_n} \right|_{k=k'} , \qquad \ddot{G}_{k'} = \left. \frac{\partial^2 G_k}{\partial \omega_n^2} \right|_{k=k'} . \tag{52}$$

They are given, together with the corresponding derivatives of $\widetilde{G}_k$, in the appendix, (C17) and (C18). In the mode-elimination step, we insert the expansion (51) into (36). $\delta Z_\tau^<$ then follows from the coupling multiplying $-\mathrm{i}\omega_n \psi_k^< \psi_k^{*<}$,

$$\delta Z_\tau^< = 2\mathrm{i}\mu g \sum_{\Lambda<|\mathbf{k}|<\Lambda_0}' \left[ 2\left(2G_k + \widetilde{G}_k\right)\dot{\widetilde{G}}_k + (2G_k - G_{-k})\,\dot{G}_k \right] , \tag{53}$$

while $\delta V^<$ is read off the term proportional to $\omega_n^2 \psi_k^< \psi_k^{*<}$, cf. (36),

$$\delta V^< = -\mu g \sum_{\Lambda<|\mathbf{k}|<\Lambda_0}' \left[ 2\left(2G_k + \widetilde{G}_k\right)\ddot{\widetilde{G}}_k + (2G_k + G_{-k})\,\ddot{G}_k \right] . \tag{54}$$

### C. Rescaling

Typically, within the WRG procedure, the next step is to rescale all coordinates as well as couplings modified by mode

elimination, $\mathbf{g}^< = \mathbf{g} + \Delta\mathbf{g}$, by powers of the scale factor $b = \Lambda_0/\Lambda$, which yields the renormalized couplings $\mathbf{g}'$. The goal of this standard procedure, which is outlined, for comparison, in App. B, is to remove all explicit dependencies on the scale factor $b$ and obtain an effective description which, close to criticality, exhibits self-similar scaling behavior.

In order to obtain, however, effective values for the couplings at a fixed length scale – as set by the scattering length – and at a fixed temperature $T$ one rather works without rescaling the coordinates explicitly. This approach allows us to extract bulk quantities like the condensate density in terms of extrinsic units. Thus, only the fields in momentum space are rescaled as

$$\Psi'(\omega_n, \mathbf{k}) = \zeta_b^{-1}\Psi^<(\omega_n, \mathbf{k})\,, \tag{55}$$

with $\zeta_b$ being the field rescaling factor. Inserting the above scaling relation into $S^<[\mathbf{g}^<]$, cf. (5), and demanding that the action does not rescale leads to the scaling of the wave-function renormalization factors,

$$Z'_\tau = \zeta_b^2 Z^<_\tau\,, \qquad Z'_x = \zeta_b^2 Z^<_x\,, \tag{56}$$

where, in this work, we will set and keep $Z^<_x = 1$ throughout. In order to conserve the canonical commutation relations of the Bose field operators $\hat{\Psi}$ we demand $Z'_\tau \stackrel{!}{=} 1$. Hence, we read off that

$$\zeta_b^2 = 1/Z^<_\tau\,, \qquad Z'_x = 1/Z^<_\tau\,. \tag{57}$$

With these relations, the rescaling of the chemical potential and of the wave-function renormalization factor $V$ result as

$$\mu' = \frac{\mu^<}{Z^<_\tau}\,, \qquad V' = \frac{V^<}{Z^<_\tau}\,. \tag{58}$$

Finally, without spatio-temporal rescaling, the total and condensate densities remain invariant,

$$n' = n^<\,, \qquad n'_\mathrm{c} = n^<_\mathrm{c}\,. \tag{59}$$

The rescaling of the interaction coupling

$$g' = \frac{g^<}{Z^<_\tau}\,, \tag{60}$$

is then inferred from $\mu = n_\mathrm{c} g$ and (58).

From $1 = Z'_\tau = b^{[Z_\tau]} Z^<_\tau$, it follows that the scaling dimension of $Z_\tau$ is given by

$$[Z_\tau] = -\partial_l \ln Z^<_\tau \equiv \eta_\tau\,, \tag{61}$$

expressed in terms of the scale parameter

$$l = \ln(\Lambda_0/\Lambda) = \ln(b)\,. \tag{62}$$

Hence, the emergence of a wave-function renormalization $Z_\tau \neq 1$ in the time derivative leads to the emergence of a temporal anomalous dimension $\eta_\tau$, which is closely related to the dynamical scaling exponent $z = \partial \ln \omega(\mathbf{k})/\partial \ln k$. The latter characterizes the momentum scaling of the quasiparticle frequencies and thus of the density of states, and is in general different from the engineering dimension $[\omega]_\mathrm{eng} = -[\tau]_\mathrm{eng} = 1$ of frequency as an inverse of time. In the course of the renormalization procedure, within the symmetry-broken phase, $z$ is expected to flow, from the case of free particles ($z = 2$) to sound waves ($z = 1$). In the thermal phase away from criticality, no sound waves appear and $z = 2$ throughout.

The derivative terms in the action imply that, for $Z^<_x \equiv 1$, the kinetic term $\sim \nabla^2$ must scale as the temporal term $Z^<_\tau \partial_\tau$ and thus

$$2 = [Z_\tau] + z = \eta_\tau + z\,, \tag{63}$$

where we used the scaling dimensions of space $[\mathbf{x}] = -1$ and time $[\tau] = -z$.

### D. Flow equations: symmetric phase

Combining the mode elimination and rescaling steps and choosing the scale factor $b = \Lambda_0/\Lambda$ to be infinitesimally close to 1, one derives the differential RG equations that govern the flow of the couplings in $l = \ln b$. First of all, the Matsubara sums in (23) and (26) must be evaluated, as described in detail in App. C 1.

Recall that, according to (22), no wave-function renormalization of the two-point couplings appears in the symmetric phase up to one-loop order, which implies that $Z^<_\tau = 1$ and $V^< = 0$ and that the dynamical scaling exponent therefore remains $z = 2$. This will change, however, in the symmetry broken phase described in the next subsection, where $z$ will flow to $z = 1$ in the IR limit.

The free correlator (20) therefore takes the form $G(k) = 1/(\mathrm{i}\omega_n + \omega_k)$ with Matsubara frequencies $\omega_n = 2\pi n T$ and the (shifted) free dispersion relation $\omega_k = \varepsilon_k - \mu$, such that the sum in (23),

$$\frac{1}{\beta}\sum_{\omega_n} G(k) = n_\mathrm{B}(\omega_k)\,, \tag{64}$$

yields the Bose-Einstein distribution

$$n_\mathrm{B}(\omega_k) = \frac{1}{\exp(\beta\omega_k) - 1}\,. \tag{65}$$

In the following we drop the argument, thereby implying that $n_k \equiv n_\mathrm{B}(\omega_k)$. The sums in (26) are evaluated in a similar manner, cf. (C4) in App. C 1. Inserting these results into (23) and (26), and rescaling according to (58) and (60), one obtains, in the limit $\delta l = 1 - \Lambda/\Lambda_0 \to 0$, the differential flows. Expressed in terms of dimensionless couplings and dispersion,

$$\bar{g} = \frac{S_d \Lambda_0^d}{(2\pi)^d}\frac{g'}{\varepsilon_{\Lambda_0}}\,, \qquad \bar{\mu} = \frac{\mu'}{\varepsilon_{\Lambda_0}}\,, \qquad \bar{T} = \frac{T}{\varepsilon_{\Lambda_0}}\,, \qquad \bar{\omega}_k = \frac{\omega'_k}{\varepsilon_{\Lambda_0}}\,, \tag{66}$$

with $S_d$ denoting the surface of a $d$-dimensional sphere, the flow equations result as [1]

$$\begin{aligned} \partial_l \mu &= -2b^{-d} g n_\Lambda\,, \\ \partial_l g &= -b^{-d} g^2 \left[4\beta n_\Lambda(1 + n_\Lambda) + \frac{1 + 2n_\Lambda}{2(b^{-2} - \mu)}\right], \end{aligned} \tag{67}$$

where we have dropped the overbars for notational simplicity. Note that the dimensionless dispersion in the second term in square brackets is rescaled to $\omega_\Lambda = b^{-2} - \mu$.

### E. Flow equations: symmetry-broken phase

The flow equations (67) are valid only in the symmetric phase, i.e., for a given density, sufficiently far above the critical temperature, where the field expectation value vanishes and wave-function renormalization of the spatial derivative term can be neglected. In the following, we derive the respective flow equations in the symmetry-broken phase, again sufficiently far away from criticality, keeping $Z_x^< = 1$, but taking into account the flow of $Z_\tau$ and thus the anomalous dimension $\eta_\tau$.

In addition to the dimensionless couplings (66), we need to define the dimensionless densities, quasiparticle frequencies (31) and wave-function renormalization $V$,

$$\bar{n} = \frac{(2\pi)^d\, n'}{S_d \Lambda_0^d}\,, \qquad \bar{n}_\mathrm{c} = \frac{(2\pi)^d\, n'_\mathrm{c}}{S_d \Lambda_0^d}\,,$$
$$\bar{\omega}^\pm_{\Lambda_0} = \frac{\omega'^\pm_{\Lambda_0}}{\varepsilon_{\Lambda_0}}\,, \qquad \bar{V} = \varepsilon_{\Lambda_0} V'\,, \tag{68}$$

where $S_d$ denotes again the surface of a $d$-dimensional sphere. Performing the Matsubara sums as summarized in App. C 2, we obtain the flow equations in the symmetry-broken phase for the dimensionless couplings, dropping again the overbars,

$$\begin{aligned}\partial_l \mu &= \eta_\tau \mu - 2 b^{-d} g \left[\left(Z_x b^{-2} + 3\mu\right) K_{1,0} + V K_{1,1}\right.\\ &\quad \left. + \mu^2 \left(\mu - 4 Z_x b^{-2}\right) K_{2,0} - 4\mu\,(1+\mu V)\, K_{2,1} - 1\right]\,,\\ \partial_l g &= \eta_\tau g - b^{-d} g^2 \left[5 K_{1,0} + 2\mu\left(\mu - 4 Z_x b^{-2}\right) K_{2,0}\right.\\ &\quad \left. - 8\,(1+\mu V)\, K_{2,1}\right]\,,\end{aligned} \tag{69}$$

and densities,

$$\begin{aligned}\partial_l n &= b^{-d}\left(-Z_x b^{-2} K_{1,0} - V K_{1,1} + 1\right)\,,\\ \partial_l n_\mathrm{c} &= b^{-d}\left[-\left(2 Z_x b^{-2} + \mu\right) K_{1,0} - 2 V K_{1,1} + 2\right]\,.\end{aligned} \tag{70}$$

The auxiliary functions $K_{i,j}$ are defined in App. C 2 in their dimensionless form. The flow equation for the condensate density can be obtained from $n_\mathrm{c}(l) = \mu(l)/g(l)$.

Note that in the limit of vanishing wave-function renormalization, $Z_\tau^< \equiv 1$ and $V^< \equiv 0$, and thus $z = 2$ and $\eta_\tau = 0$, the flow equations (69) and (70) reduce to those derived in [1].

Extending upon this description, we find, from (53), the flow equation for $Z_\tau^<(l)$ and, using (54), the flow of $V$. With the Matsubara sums given in (C19) and (C20) of App. C 2, expressed in terms of rescaled and dimensionless couplings, the temporal anomalous dimension $\eta_\tau = -\partial_l \ln Z_\tau^<$ results as

$$\begin{aligned}\eta_\tau = 2 b^{-d} Z_x \mu g \left[\left(Z_x b^{-2} + \mu\right) K_{2,0} - 9 V K_{2,1}\right.\\ \left. + 8 Z_x (2\mu V - 1) K_{3,1} + 8 V (2\mu V + 1) K_{3,2}\right]\,.\end{aligned} \tag{71}$$

As can be seen from Eqs. (69) and (70), it affects the flows of both, the couplings and the densities.

The flow of the coupling $V$ is determined analogously, using the Matsubara sums (C21) and (C22), and results as

$$\begin{aligned}\partial_l V = \ \eta_\tau V - b^{-d} Z_x \mu g \Big\{ &-3(1+A_k) K_{2,0} + 18 V^2 K_{2,1} + 2\mu A_k \left(4 Z_x b^{-2} - \mu\right) K_{3,0}\\ &- 4\left[2 + V^2\left(4 Z_x^2 b^{-4} + 28 Z_x b^{-2}\mu - 13\mu^2\right) - A_k\left(11 + 2 V \omega^2\right)\right] K_{3,1} - 16 V^2 (3 + 5\mu V) K_{3,2}\\ &+ 8\left(4 V^2 \omega^2 - A_k^2\right)\left[\mu\left(4 Z_x b^{-2} - \mu\right) K_{4,1} + 4(1+\mu V) K_{4,2}\right]\Big\}\,,\end{aligned} \tag{72}$$

with $A_k = 1 + 2V\left(Z_x b^{-2} + \mu\right)$ in its rescaled form. The flow equations (69)–(72) constitute a closed set of equations defining the RG evolution including wave-function renormalization of the first- and second-order temporal derivative terms.

## IV. SOLVING THE FLOW EQUATIONS

### A. Initialization of the flow equations

We can now use the flow equations (69) and (70) to determine the evolution of the couplings in the condensed phase. Before doing so, we need to initialize them at the UV cut-off $\Lambda_0$. We start with the quartic coupling and thereafter define the initialization of the remaining quantities.

#### 1. *Initialization of the coupling at a chosen microscopic scale*

Solving the flow equations (69) and (70), starting at the scale set by the UV cut-off $\Lambda_0$, yields solutions, which in principle depend on this initial scale $\Lambda_0$. However, the cut-off scale should not have any physical significance for the IR physics since it merely defines the limit of validity of the $s$-wave approximation. A cut-off independent flow in the IR can be achieved by considering the experimentally measured scattering length $a_0$ as a renormalized quantity measured in vacuum. This is reasonable since the experimental value is

already an effective quantity containing quantum corrections. Hence, we initialize the flow such that, in vacuum, the flowing scattering length eventually reaches the experimental value of $a_0$ [1]. The vacuum flow equations can be obtained from those in the symmetric phase, Eqs. (67), by setting $n_k = 0$. Assuming a small chemical potential, $\mu \ll 1$, the equation can be analytically solved to yield

$$g(l) = \frac{2\epsilon g_{\text{in}}}{2\epsilon - g_{\text{in}} + \mathrm{e}^{\epsilon l} g_{\text{in}}}\,, \tag{73}$$

where we introduced the scaling dimension $\epsilon = 2 - d$ for the interaction coupling. We determine, from this, the bare coupling $g_{\text{in}}$ in $d = 3$ dimensions, where $\epsilon = -1$, since no wave-function renormalization is present in the thermal phase. The dimensionless four-point coupling, cf. (66), in vacuum asymptotically approaches the $s$-wave result $g(l \gg 1) = (m\Lambda_0/\pi^2) \times (4\pi a_0/m) = 4\Lambda_0 a_0/\pi$ for large flow parameter $l = \ln(b) = \ln(\Lambda_0/\Lambda)$. Thus, the initial value for $g$ from (73) is

$$g_{\text{in}} = \frac{4a_0\Lambda_0}{\pi - 2a_0\Lambda_0}\,. \tag{74}$$

We will employ this initial, bare coupling in the non-vacuum flow when determining the low-energy effective four-point coupling in the macroscopic regime. By making, in this way, the initial value cut-off dependent we achieve cut-off independent outcomes and fix the flow equations to a particular physical vacuum scattering length.

#### 2. *Initialization of the chemical potential, temperature, and density*

We begin solving the flow equations by initializing all particles in the symmetry-broken phase, $n_{\text{in}} = \mu_{\text{in}}/g_{\text{in}}$. The initial chemical potential $\mu_{\text{in}}$ is determined by means of the UV cut-off dependent initial coupling $g_{\text{in}}$ (74) and the set initial density. We set the UV cut-off to $\Lambda_0 = \gamma/a_0$, with the free parameter $\gamma$ set to $\gamma = 1$ unless stated otherwise. The temperature $T$ is set to the physical temperature.

For a fixed initial total particle density $n_{\text{in}}$, the density in the IR would depend on the chosen temperature. Therefore, to obtain a desired physical density $n(l \to \infty) = n_{\text{phys}}$, one needs to correctly specify the initial condition $n(l = 0) = n_{\text{in}}(T)$ for each given temperature $T$. Thus, the initial density $n_{\text{in}}$ can be found as the root of $n(n_{\text{in}}, l \to \infty) - n_{\text{phys}}$ using the secant method. The initial condensate density is set to the initial total particle density, i.e., all particles are initialized in the condensate.

During the RG flow, the chemical potential will remain positive but will decrease due to the redistribution of condensed particles into the excited states. For temperatures at and above $T_{\text{c}}$, the chemical potential vanishes at some point during the flow. This moment indicates the restoration of the thermal phase as no particles are condensed anymore. Beyond this point, the couplings evolve according to the thermal flow equations (67). A smooth transition between the phases is guaranteed since the flow equations for both couplings are identical for $\mu = 0$.

### B. Numerical results without wave-function renormalization

Given the initial values for the couplings and the densities, the flow equations are numerically solved and the renormalized couplings are read out at some finite maximum value $l_{\text{max}}$. In the remainder of this subsection, we neglect wave-function renormalization by setting the respective factors to $Z_\tau^< = 1$ and $V = 0$ and demonstrate how this leads to a diverging flow of physical quantities and the necessity to take wave-function renormalization into account in the IR regime where collective quasiparticle excitations prevail.

#### 1. *Flows of the couplings and densities*

In order to discuss the RG flows at a chosen momentum and a temperature scale, we work within the rescaling scheme defined in Eqs. (58)–(60). Since $Z_\tau^< = 1$ and $V = 0$, this means that no rescaling is applied, $\mu'(l) = \mu^<(l)$, $g'(l) = g^<(l)$, and $n'(l) = n^<(l)$.

All the numerical evaluations are performed for parameters describing the example of a dilute Bose gas of $^{23}$Na atoms with the $s$-wave scattering length $a_0 = 51.12\,a_{\text{B}}$ in the $(F, m_F)+(F, m_F) = (1,0) + (1,0)$ hyperfine channel [66][67] and a particle density $n_{\text{phys}} = 10^{19}/\text{m}^3$. All RG flows were evolved from a UV cut-off $\Lambda_0 = 1/a_0$ up to a maximum flow parameter $l_{\text{max}} = 20$ situated deep in the IR.

Fig. 1 shows the flows of the chemical potential $\mu(l)$ and the interaction coupling $g(l)$, both normalized to their respective initial values. The vertical gray line indicates the healing-length scale, $\Lambda_\xi^{-1} = \mathrm{e}^{l_\xi}/\Lambda_0 = \xi = (8\pi a_0 n_{\text{phys}})^{-1/2}$. For the chosen density and scattering length, we obtain a healing-length scale parameter $l_\xi \approx 6.1$. Sufficiently far below this scale one observes qualitatively similar flows for all three temperatures chosen, $T = 0$, 10, and 100 nK. Beyond the healing-length scale, the flows run into asymptotic solutions which exhibit scaling in $b = \mathrm{e}^l$. As a result, there is no convergence even for $T = 0$, for which the scaling is strongly suppressed, see the insets.

Fig. 2 shows the flows of $n_{\text{tot}}(l)$ (dashed) and $n_{\text{c}}(l)$ (solid lines) normalized to the chosen asymptotic total density $n_{\text{phys}} = 10^{19}\text{m}^{-3}$, for the same three temperatures as above. As anticipated, the condensate density decreases with increasing temperature, whereas the total density flows towards the chosen final total density. In contrast to the couplings, we observe convergence in the flows of both densities. As a result, the asymptotic behavior of the condensate and total densities results as $n_{\text{c}}^*(l) = n_{\text{c,phys}}$ and $n^*(l) = n_{\text{phys}}$, respectively. The $T$-dependence of $n_{\text{c,phys}}$ will be examined in more details below.

#### 2. *Asymptotic scaling*

The observed asymptotic scaling behavior for non-zero temperatures can be read off the approximate asymptotic forms of the flow equations (69) and (70). For $l \gg 1$ these

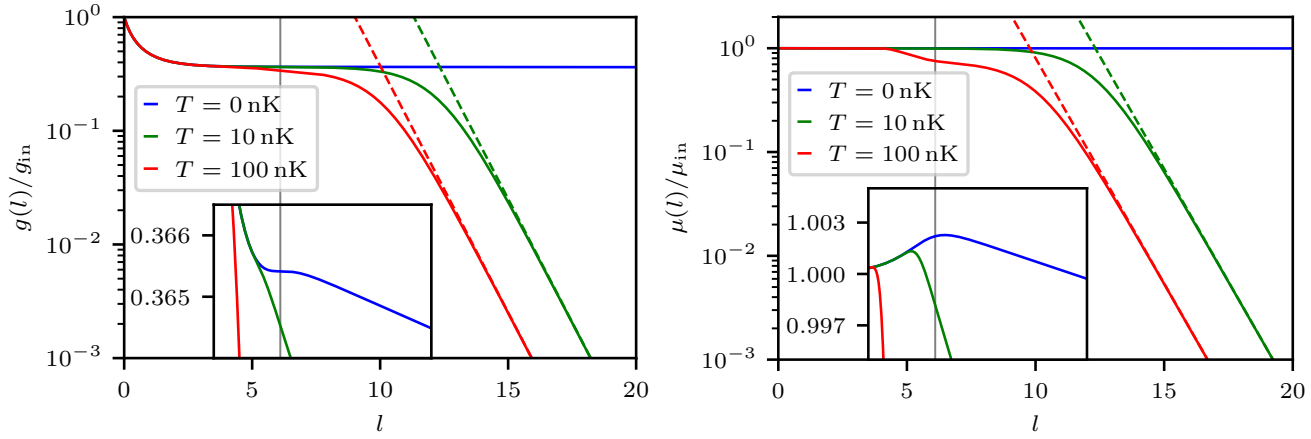



FIG. 1. RG flow of the quartic coupling $g(l)$ (left) and the chemical potential $\mu(l)$ (right) normalized by their initial values $\mu_{\rm in}$ and $g_{\rm in}$, respectively. The couplings are shown as functions of the flow parameter $l = \ln(\Lambda_0/\Lambda)$ encoding the flowing cut-off scale $\Lambda$, for three different temperatures below the non-interacting critical temperature $T_{\rm c}^0 \approx 324\,\mathrm{nK}$ of Bose-Einstein condensation in a 3-dimensional gas of $^{23}$Na atoms of density $n = 10^{19}\mathrm{m}^{-3}$. No wave-function renormalization has been taken into account here. The initial UV cut-off is chosen to be $\Lambda_0 = 1/a_0$, with the $s$-wave scattering length $a_0$. The healing-length scale $l_\xi = \ln(\xi\Lambda_0)$ is marked by the gray line. The insets show the flow in the vicinity of the healing-length scale on an enlarged $y$-scale and with the $x$-axis shared with the main plots. In both graphs, the dashed lines represent the asymptotic solutions $g^*(l) = 2b^{-1}\beta$ and $\mu^*(l) = n_{\rm c,phys}g^*(l)$, respectively, which exhibit anomalous scaling dimensions $[\mu]_{\rm scaling} = 1$ and $[g]_{\rm scaling} = -2$ compared to the engineering dimensions $[\mu]_{\rm eng} = 2$ and $[g]_{\rm eng} = -1$.

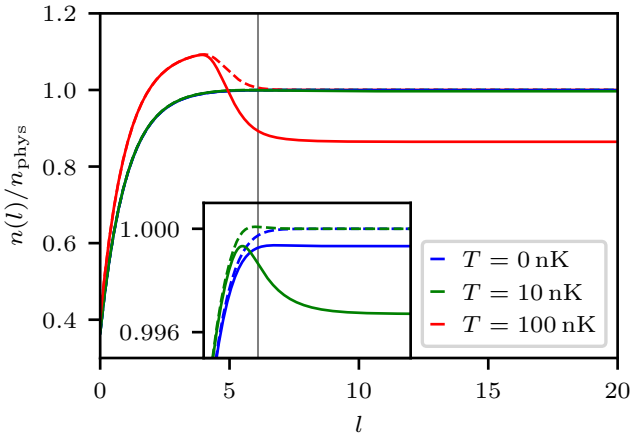



FIG. 2. RG flow of the condensate density $n_{\rm c}(l)$ (solid) and the total particle density $n_{\rm tot}(l)$ (dashed) normalized by the final density $n_{\rm phys} = 10^{19}\mathrm{m}^{-3}$. The densities are plotted against the flow parameter $l$ without wave-function renormalization below the critical temperature. As before, the UV cut-off is $\Lambda_0 = 1/a_0$ and the healing length scale is indicated by the gray line. The subplot, which shares the $x$-axis with the main plot, shows the flow around the healing scale enlarged and exhibits a condensate depletion at $T = 0$. For large $l$, i.e., in the IR limit, one observes convergence to constant values.

read

$$\partial_l g = -b^{4-d}\frac{g^2}{2\beta}\,, \qquad \partial_l \mu = -b^{4-d}\frac{g\mu}{2\beta}\,,$$
$$\partial_l n = 0\,, \qquad \partial_l n_{\rm c} = 0\,. \tag{75}$$

For $n$ and $n_{\rm c}$, the flows converge as found in Fig. 2. The coupling approaches the asymptotic scaling behavior $g(l) \sim g^*(l) = 2b^{-1}\beta$ satisfying $\partial_l b g^*(l) = b(1-\partial_l)g^*(l) = 0$, as can be inferred from the flow equation (75) for the coupling $g$ in $d = 3$ dimensions. This behavior is reflected by the exponential decay for $l \gtrsim 15$ in the left panel of Fig. 1, depicting the normalized $g(l)/g_{\rm in}$. For comparison, we also show, as dashed lines, the respective asymptotic scaling $2b^{-1}\beta/g_{\rm in}$.

Analogously, the flow equation for $\mu$ gives the asymptotic scaling $\mu(l) \sim \mu^*(l) \sim b^{-1}$ at large $l$, as can be seen in the right panel of Fig. 1. Combining the asymptotic scaling of the coupling and the density confirms that the chemical potential scales as $\mu^*(l) = n_{\rm c}^*(l)g^*(l) = 2b^{-1}n_{\rm c,phys}\beta$.

If the couplings were scaling with their engineering dimensions, $[\mu]_{\rm eng} = 2$ and $[g]_{\rm eng} = 2 - d$, cf. App. B, then $\mu(l)$ and $g(l)$ would approach constants asymptotically. However, both scale as $b^{-1}$, exhibiting scaling dimensions $[\mu]_{\rm scaling} = [\mu]_{\rm eng} - 1 = 1$ and $[g]_{\rm scaling} = [g]_{\rm eng} - 1 = -2$ in three dimensions. This also implies the absence of cut-off independence, as exhibited by the diverging flows in Fig. 1. Since both couplings deviate in their observed scaling by the same single power in $b$, the engineering and scaling dimensions of the density agree with each other, $[n]_{\rm scaling} = [n]_{\rm eng} = d = 3$.

While anomalous scaling, i.e., a deviation between engineering and scaling dimensions is a generic feature of RG flows close to critical points, it is in general not expected to occur away from criticality, in particular close to zero temperature. The appearance of anomalous scaling is caused by the renormalization of operators such as $\nabla^2$ and $\partial_\tau$ which emerges from the Taylor expansion of the two-point diagram (35), cf. the discussion in [1]. To be able to avoid taking into account the flow of $Z_\tau$ and $Z_x$, one needs to restrict the description to the regime $na_0\Lambda_{\rm th}^2 \ll 1$, in which the thermal energy is much larger than the chemical potential and thus the linear part of the Bogoliubov dispersion becomes negligible and IR divergent flows are absent. At more macroscopic scales, in

the sound-wave regime, though, one needs to go beyond this approximation. This regime, in the RG framework, so far has been investigated by means of non-perturbative, functional approaches, cf. Refs. [30, 42, 68]. There, flow equations have been derived, in which the second-order temporal derivative $\partial_\tau^2$ becomes dominant as compared with the first-order derivative $\partial_\tau$ [68]. In the following, we will demonstrate how this transition can be qualitatively described also at the perturbative level.

### C. Numerical results including wave-function renormalization

In the following, we derive the flow equations taking into account the wave-function renormalization factors $Z_\tau$ and $V$. Following the same procedure as without this renormalization, the system is prepared in the fully condensed state, where UV cut-off dependent initial values for the couplings are chosen and the density flows to a pre-determined value of the physical density $n_{\rm phys}$.

#### 1. Asymptotic scaling

We start by analyzing the asymptotic flow equations in order to determine the significance of the wave-function renormalization for the asymptotic scaling. Including now the anomalous dimension $\eta_\tau$, the leading terms in the flow equations (69) for the couplings $\mu$ and $g$ read:

$$\partial_l g = \eta_\tau g - b^{4-d}\frac{Z_\tau g^2}{2\beta}\,, \qquad \partial_l \mu = \eta_\tau \mu - b^{4-d}\frac{Z_\tau \mu g}{2\beta}\,. \tag{76}$$

The asymptotic anomalous dimension is given by

$$\eta_\tau = b^{4-d}\frac{Z_\tau g}{2\beta}\,, \tag{77}$$

cf. (71). Thus, it follows from Eqs. (76) that

$$\partial_l g = 0\,, \qquad \partial_l \mu = 0\,, \tag{78}$$

asymptotically. Hence, the second-order one-loop contribution to the changes arising from mode-elimination decreases the scaling dimension of both couplings by the temporal anomalous dimension. This shows that the rescaled couplings $g'(l) = g^<(l)/Z_\tau^<(l)$ and $\mu'(l) = \mu^<(l)/Z_\tau^<(l)$, as defined in (58) and (60), each have a converging flow, cf. (78).

The asymptotic flow equation for $g$, Eq. (78), has the solution $g(l) = g^*$ with $g^*$ being the fixed-point value of the coupling. The flow of the scale-dependent anomalous dimension, cf. (61), is therefore determined by the ordinary differential equation

$$\partial_l \eta_\tau = (4 - d - \eta_\tau)\,\eta_\tau\,, \tag{79}$$

as follows from (77) using $\partial_l g = 0$. It possesses a fixed point at $\eta_\tau^* = 0$, which corresponds to the free gas with dynamical scaling exponent $z^* = 2$, cf. (63). The second, non-vanishing fixed-point of the anomalous dimension depends on the dimensionality,

$$\eta_\tau^* = 4 - d\,. \tag{80}$$

Thus, in $d = 3$ dimensions, the anomalous dimension flows to $\eta_\tau(l) \to 1$, which, according to (63), corresponds to $z(l) \to 1$, as it is expected for sound waves.

#### 2. Flows of the couplings and densities

When numerically solving the flow equations we need to take care of singularities appearing due to the additional coupling $V$. The derivative couplings are initialized as $Z_{\tau,\rm in} = 1$ and $V_{\rm in} = 0$. In this limit, the "anti-particle" modes formally decouple: $\omega_k^+ \to \infty$. For $V > 0$, there are two additional modes, which, for $\mu_1 = \mu_2 = m + \mu_{\rm NR}$, with rest mass $m$ and non-relativistic chemical potential $\mu_{\rm NR} \simeq gn$, reflect the emergent SO($d+1$) symmetry of the model in the sound-wave regime. These modes account for anti-particles, which are equivalent to the particles due to the gapless dispersion. Note that, while the description is similar to that of relativistic massless bosons, we here describe the transition from thermal energies in the single-particle limit to the phononic regime of collective excitations and therefore choose $\mu_1 = \mu_2 = \mu$.

Fig. 3 shows the flows of the quartic coupling $g(l)$ and the chemical potential $\mu(l)$, for the same three temperatures as in Fig. 1. While the unrescaled (dashed) couplings $\mu^<(l)$ and $g^<(l)$ vanish asymptotically at large $l$, as was the case without wave-function renormalization, they reach a constant value once rescaled with $Z_\tau^<(l)$ (solid). As discussed in Sect. III C, this rescaling corresponds to fixing the spatial and temporal scales to the experimental scattering length and the physical temperature, respectively. Within this rescaling scheme, only the fields need to be rescaled by $1/Z_\tau^<$, cf. (58) and (60).

In Fig. 4, we display the wave-function renormalization $Z_\tau^<(l)$ (solid), the dynamical scaling exponent $z(l)$ (dashed), and the temporal anomalous dimension $\eta_\tau(l)$ (dotted). For $T \neq 0$, they demonstrate the predicted behavior (80). Beyond the healing scale, the transition into the quasi-relativistic regime leads to a decrease of $Z_\tau^<(l)$, which eventually asymptotically vanishes as $Z_\tau^<(l) \sim 1/b$. Meanwhile, the second-order temporal derivative term grows up leading to a quasi-relativistic form of the action. The temporal anomalous dimension reaches the expected limit $\eta_\tau(l) \to 1$, see (80). Since the dynamical scaling exponent is directly related to the wave-function renormalization via the temporal anomalous dimension (63), one finds $z(l) \to 1$ (cf. Fig. 4). While we could reproduce the above scaling of the wave-function renormalization $Z_\tau^<(l)$ down to $T = 0.1$nK, we find, in the case of $T = 0$, that $Z_\tau^<(l)$ reaches a fixed-point value at $Z_\tau^{<*} \approx 0.9984$ as the temporal anomalous dimension $\eta_\tau$ stays at zero.

We finally discuss the relevance of the flow of the wave-function renormalization factor $V$. During an initial period of the flow, as long as $l \lesssim l_1 < l_\xi$, the value of $V$ remains sufficiently small such that it does not account for a significant change of the right-hand sides of the flow equations compared to $V = 0$. This, however, changes once the flow reaches the

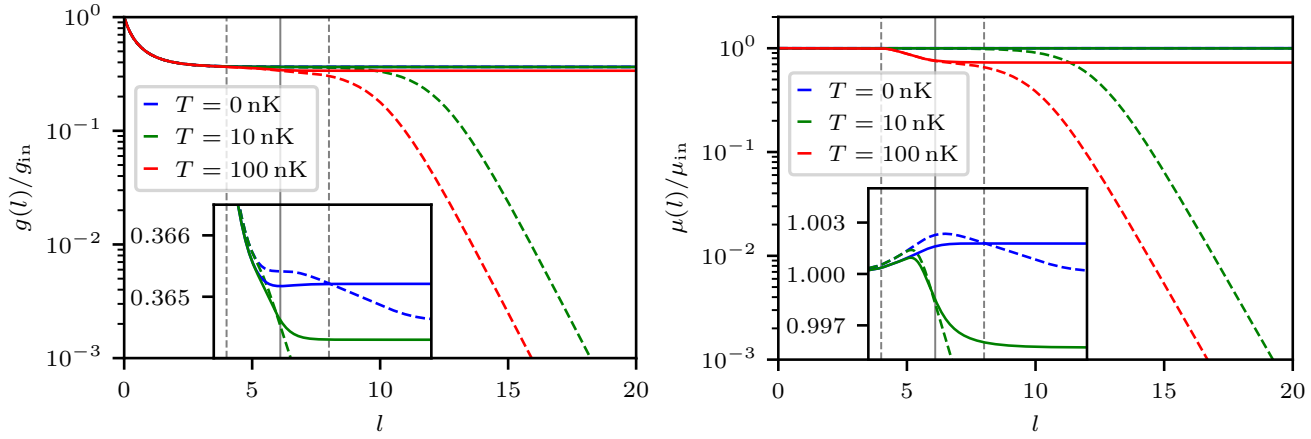



FIG. 3. Quartic coupling $g(l)$ (left) and chemical potential $\mu(l)$ (right) normalized by their initial values $\mu_{\mathrm{in}}$ and $g_{\mathrm{in}}$, respectively. The couplings are plotted against the flow parameter $l$ including wave-function renormalization within the symmetry-broken phase. Solid lines correspond to rescaled couplings according to (60) and (58) and show IR convergent flows, whereas dashed lines correspond to the unrescaled couplings $g^<(l)$ and $\mu^<(l)$. The solid, vertical, gray line indicates again the healing scale, whereas the two dashed lines separate the three distinct intervals of the RG flows, see the discussion in the main text. The insets, which share the $x$-axis with the main plots, show the flow on an enlarged $y$-scale.

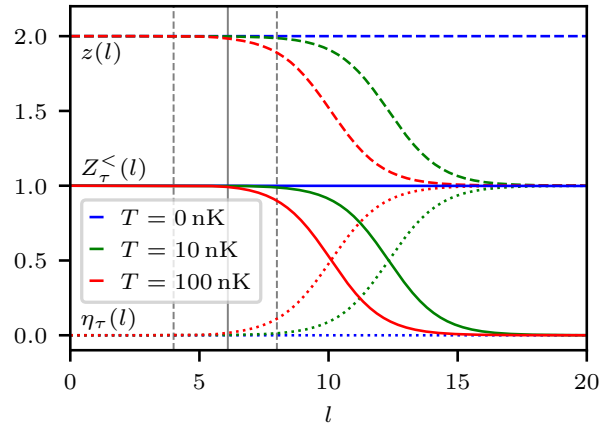



FIG. 4. The wave-function renormalization $Z_\tau^<(l)$ (solid), the dynamical scaling exponent $z(l)$ (dashed), and the temporal anomalous dimension $\eta_\tau(l)$ (dotted) plotted against the flow parameter $l$. In the case of non-zero temperature, the transition towards a relativistic regime with $Z_\tau^<(l) \to 0$, $z(l) \to 1$, and $\eta_\tau(l) \to 1$ beyond the healing scale is observed. At $T = 0$, no such transition occurs.

healing-length scale $l \approx l_\xi \approx 6$, cf. the discussion in IV B 1, where $V(l)$ becomes relevant [30]. Finally, in the asymptotic limit $l \to \infty$, the coupling $V$ decouples again from the flow equations, as can be inferred from the explicit forms of the functions $K_{i,j}$ quoted in App. C 2.

For these reasons, when numerically solving the flow equations we employ different schemes depending on the value of the flow parameter $l$. At small values $l \leq l_1$, we evolve the flow equations with the quadratic derivative term left out, i.e., $V = 0$, while we compute the growth of $V$ using (72). In the intermediate region $l_1 < l \leq l_2$, we evolve the full flow equations including the feedback of $V$. Finally, at large $l > l_2$, the couplings are evolved using the asymptotic flow equations for $V \gg 1$, obtained from (69) and (71), which ultimately approach (76) and (77), respectively. The values of $l_1$ and $l_2$ are chosen such that a smooth transition is warranted, i.e., that the effect of $V \neq 0$ is small and the feedback of $V$ on the flow equations becomes negligible. In our computations, we chose $l_1 = 4$ and $l_2 = 8$, which we indicate by the gray dashed lines in Figs. 3 and 4. The flow of $V$ is shown in Fig. 7 in App. D.

### 3. *Temperature dependence of the couplings*

Having achieved converging RG flows we can investigate the temperature-dependence of various physical quantities. The left panel of Fig. 5 depicts the chemical potential $\mu$ (dashed) and the quartic coupling $g$ (dotted) normalized to their zero-temperature mean-field values $\mu_{\mathrm{MF}}$ and $g_{\mathrm{MF}}$, respectively, as functions of temperature. One observes a phase transition at around $T_{\mathrm{c}} \approx 340\,\mathrm{nK}$. As expected, the presence of interactions shifts the critical temperature to a larger value compared to the non-interacting result $T_{\mathrm{c}}^0 \approx 324\,\mathrm{nK}$. This shift does not change qualitatively when setting $V = 0$, cf. App. D. The interaction coupling $g$ remains close to its mean-field value for most temperatures and falls significantly only in the vicinity of the critical temperature. Thus, the steady decrease of the chemical potential stems from the depletion of the condensate density. As can be seen in the right panel of Fig. 5, this decrease of the condensate density (dashed) with growing temperature is qualitatively similar to that of the non-interacting gas (dashed, gray). The dotted horizontal line demonstrates that we chose the total density constant over the whole temperature range.

Fig. 5 furthermore confirms the desired weak dependence of the chemical potential, coupling, and densities as functions of temperature on the flowing cut-off scale. This is reflected in the alternating color scheme indicating overlapping

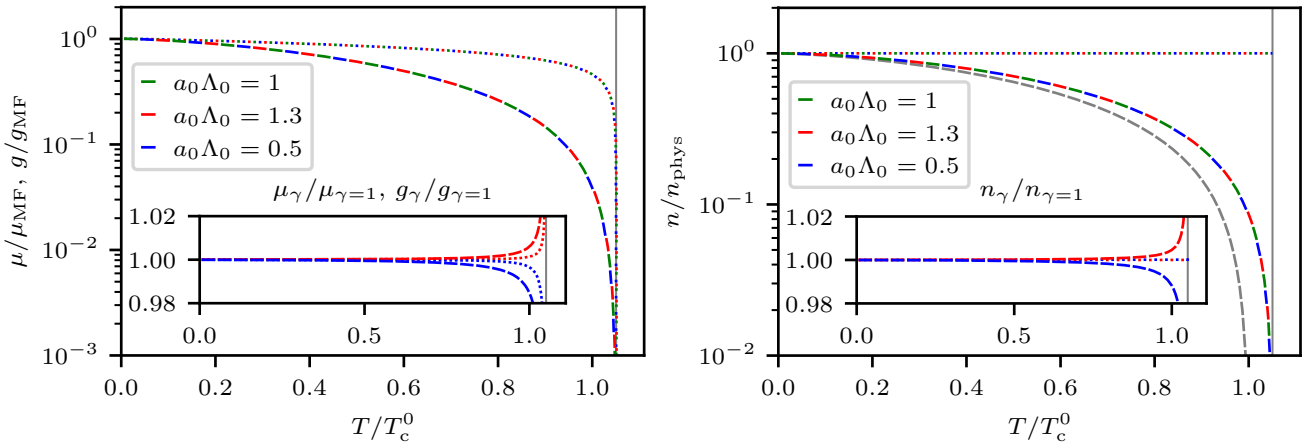



FIG. 5. (Left panel) The chemical potential $\mu$ (dashed) and the quartic coupling $g$ (dotted) normalized by their respective zero-$T$ mean-field value $\mu_{\rm MF}$ and $g_{\rm MF}$ plotted against temperature $T$. Both couplings are rescaled according to (58) and evaluated at the maximal flow parameter $l_{\rm max} = 20$. The couplings are plotted for three different UV cut-offs and show no cut-off dependence, which is reflected in the alternating color scheme indicating overlapping results. The inset shows the relative deviation of the $\gamma = a_0\Lambda_0 = 1.3$ (red) and $\gamma = 0.5$ (blue) results from the $\gamma = 1$ case for both couplings. (Right panel) The condensate density $n_{\rm c}$ (dashed) and the total density $n_{\rm tot}$ (dotted) normalized by the physical particle density $n_{\rm phys} = 10^{19}\,{\rm m}^{-3}$ plotted against temperature $T$ for three different values of the UV cut-off. The densities are rescaled according to (59) and evaluated at the maximal flow parameter $l_{\rm max} = 20$. The condensate density of the non-interacting result (dashed, gray) exhibits the shift to a larger critical temperature. The inset shows the relative deviation of the $\gamma = a_0\Lambda_0 = 1.3$ (red) and $\gamma = 0.5$ (blue) results from the $\gamma = 1$ case for both densities. The vertical gray line indicates the critical temperature.

results for various UV cut-offs. The relative deviations of the $\gamma = a_0\Lambda_0 = 1.3$ and $\gamma = 0.5$ functions from those at $\gamma = 1$ are shown in the insets and remain below 1% for all temperatures sufficiently far away from the critical point. This is the result of both, the UV cut-off independent initialization described in Sect. IV A 1 and of the procedure, which ensures convergence, cf. Eqs. (58) and (60). Hence, including wave-function renormalization allows us making quantitative predictions in the long-wave-length regime $\Lambda_{\rm th}^2 \ll (na_0)^{-1}$, where the excitations have phononic character.

To benchmark our approach, we determine the condensate depletion from the flow as depicted in Fig. 2. At zero temperature the flow equations for the condensate density are solved for a range of total densities $n_{\rm tot}$. The results are shown, as functions of the diluteness parameter $\eta = \sqrt{n_{\rm tot}a_0^3}$, in the left panel of Fig. 6, where we observe good agreement with the leading-order Bogoliubov result

$$1 - \frac{n_{\rm c}}{n_{\rm tot}} = \frac{8}{3\sqrt{\pi}}\sqrt{n_{\rm tot}a_0^3}\,. \tag{81}$$

The relative deviation from this result, shown in the inset, demonstrates corrections on the order of a few percent for large densities. Furthermore, the deviation is not cut-off independent anymore showcasing limitations of our initialization scheme that assumes small chemical potential, see the discussion in Sect. IV A.

We finally consider the interaction-induced shift of the critical temperature $\Delta T_{\rm c}/T_{\rm c}^0 = (T_{\rm c} - T_{\rm c}^0)/T_{\rm c}^0$ relative to the ideal-gas value $T_{\rm c}^0 = 2\pi/m\,(n/\zeta(3/2))^{2/3}$, with $\zeta(3/2) \approx 2.61$ being the Riemann zeta function. This relative shift is known to obey $\Delta T_{\rm c}/T_{\rm c}^0 = \kappa a_0 n^{1/3}$ at leading order $\eta^{2/3}$ in the diluteness parameter. A proportionality constant of $\kappa \approx 1.3$ has been found in different types of Monte-Carlo simulations and from variational perturbation theory, see [9] for a review, as well as [30, 51, 54, 56]. For an RG calculation on the basis of the thermal phase flow equations, cf. (67), together with non-perturbative input see [29]. Here we compute critical temperatures from the symmetry-broken flow equations (69) for a range of different densities as shown in the right panel of Fig. 6. At small densities, the critical temperature approaches its non-interacting value $T_{\rm c}^0$. Increasing the density leads to a shift to larger critical temperatures in accordance with known results [9]. In terms of the critical degeneracy parameter $n\Lambda_{\rm th}^3$ our results agree with those in [1]. The linear dependence of the shift on $a_0 n^{1/3}$ is only approached for small diluteness parameter with $\kappa$ being an order of magnitude larger than the standard result, as can be seen in the right panel of Fig. 6. This deviation can be attributed to the perturbative nature of the approach employed here, cf. [30].

## V. CONCLUSIONS

We applied perturbative Wilsonian renormalization-group theory to computing observables in the symmetry-broken, condensate phase of a weakly-interacting Bose gas in three spatial dimensions. Extending upon the approach of [1] by taking into account, to one-loop order, wave-function renormalization of both the first- and second-order derivative terms in the effective action with respect to imaginary time, we could follow the system into the regime where sound-wave collective quasiparticle excitations prevail. Without this wave-function renormalization included, the flows exhibited unphysical scalings of certain couplings, which thus vanished at

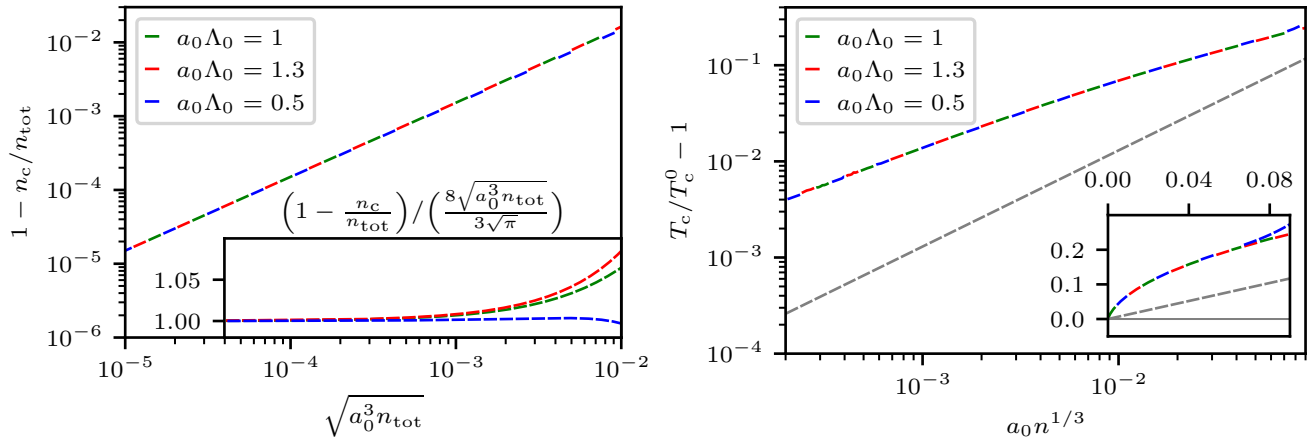



FIG. 6. (Left panel) The condensate depletion $1 - n_c/n_{tot}$ for three different values of the UV cut-off evaluated at $l_{max} = 20$ and rescaled according to (59). The alternating color scheme indicates overlapping results and thus cut-off independence. The subplot, which shares the $x$-axis with the main plot, displays the relative deviation from the Bogoliubov prediction (81) showing a good agreement over many orders of magnitude. (Right panel) The critical temperature $T_c$ normalized by the non-interacting critical temperature $T_c^0$ for three different values of the UV cut-off. The alternating color scheme indicates again overlapping results and thus cut-off independence within most of the range of couplings and densities shown. We observe a shift to larger critical temperatures as expected but the linear dependence on $a_0 n^{1/3}$ predicted as $T_c/T_c^0 - 1 = \kappa a_0 n^{1/3}$, with $\kappa = 1.3$ (gray, dashed), appears only for much smaller $a_0 n^{1/3}$, with a larger prefactor $\kappa \approx 19$ (as also seen in the inset on a linear scale).

wave lengths beyond the healing length scale. Including temporal wave-function renormalization, in contrast, and adopting a rescaling scheme which fixes both the spatial and temporal scales, we were able to achieve converging flows. This allowed us to describe the transition of the system from the particle-like into the phononic regime. Hence, we found the dynamical scaling exponent flowing towards $z \to 1$ marking the emergence of a linear sound dispersion in line with the non-perturbative Bogoliubov-de Gennes description. We used the RG flow equations to compute the depletion of the condensate density due to thermal fluctuations and determine the shift of the critical temperature. We found the depletion to be in good agreement with Bogoliubov-de Gennes theory up to linear order in the diluteness parameter $\eta = \sqrt{na^3}$. For the critical temperature a positive shift was found, which, however, did not linearly scale in $a_0 n^{1/3}$. A further improvement to the perturbative Wilsonian approach adopted here requires including the spatial wave-function renormalization, which is expected to play a significant role close to the critical point. Our approach can be considered as a first step in extending perturbative RG treatments to describe non-universal properties also in multicomponent Bose gases.

## ACKNOWLEDGMENTS

The authors thank A. Baum, P. Heinen, M. K. Oberthaler, H. Köper, J. Mayr, J. M. Pawlowski, A. Pelster, I. Siovitz, and H. T. C. Stoof for discussions and collaboration on related topics. The authors acknowledge support by the German Research Foundation (DFG), through SFB 1225 ISOQUANT (Project-ID 273811115), grant GA677/10-1, and under Germany's Excellence Strategy – EXC 2181/1 – 390900948 (the Heidelberg STRUCTURES Excellence Cluster). N. R. acknowledges support from the Studienstiftung des deutschen Volkes. A. N. M. acknowledges financial support by the IMPRS-QD (International Max Planck Research School for Quantum Dynamics).

### Appendix A: Fourier transformations

Fourier transforms in the Euclidean $d + 1$-dimensional space-time are defined as

$$\Psi(\tau, \mathbf{x}) = \frac{1}{\beta} \sum_{\omega_n} \int \frac{\mathrm{d}\mathbf{k}}{(2\pi)^d} \Psi(\omega_n, \mathbf{k})\, \mathrm{e}^{-\mathrm{i}(\omega_n \tau + \mathbf{k}\mathbf{x})}\,,$$
$$\Psi(\omega_n, \mathbf{k}) = \int_0^\beta \mathrm{d}\tau \int \mathrm{d}\mathbf{x}\, \Psi(\tau, \mathbf{x})\, \mathrm{e}^{+\mathrm{i}(\omega_n \tau + \mathbf{k}\mathbf{x})}\,. \tag{A1}$$

### Appendix B: Rescaling including coordinates

In this appendix, we outline the standard rescaling procedure, which can be used to discuss critical exponents in the vicinity of fixed-points of the RG flow equations. First, momenta, Matsubara frequencies, coordinates, and time are rescaled with the scale factor $b = \Lambda_0/\Lambda$ according to their scaling dimensions as

$$\mathbf{k}' = b^{[\mathbf{k}]}\mathbf{k}\,, \qquad \mathbf{x}' = b^{-[\mathbf{k}]}\mathbf{x}\,,$$
$$\omega_n' = b^{[\omega_n]}\omega_n\,, \qquad \tau' = b^{-[\omega_n]}\tau\,, \tag{B1}$$

with $[\mathbf{k}] = 1$, and $[\omega_n] = z$. The rescaling of the fields in momentum space is given as

$$\Psi'(\omega_n', \mathbf{k}') = \zeta_b^{-1} \Psi^<(\omega_n, \mathbf{k}) \,, \tag{B2}$$

with $\zeta_b$ being the field rescaling factor. In contrast to (55), the coordinates are being explicitly rescaled here. As the dimensionless action does not rescale, the scaling of the wave-function renormalization factors is

$$Z_\tau' = b^{-d-2z} \zeta_b^2 Z_\tau^< \,, \qquad Z_x' = b^{-d-z-2} \zeta_b^2 Z_x^< \,, \tag{B3}$$

with $Z_x^< = 1$ as elsewhere in this work. The constraint $Z_\tau' \stackrel{!}{=} 1$ ensures conservation of the canonical commutation relations, such that

$$\zeta_b^2 = \frac{b^{d+2z}}{Z_\tau^<} \,, \qquad Z_x' = \frac{b^{z-2}}{Z_\tau^<} \,, \tag{B4}$$

with $l = \ln b$. Thus, the rescaling of the chemical potential and of the wave-function renormalization factor $V$ are

$$\mu' = \frac{b^z \mu^<}{Z_\tau^<} \,, \qquad V' = \frac{b^{-z} V^<}{Z_\tau^<} \,. \tag{B5}$$

The temporal anomalous dimension $\eta_\tau$ is defined as in (61) and relates to the dynamical scaling exponent $z$ via (63). The total and condensate densities are rescaled as

$$n' = b^d \, n^< \,, \qquad n_\mathrm{c}' = b^d \, n_\mathrm{c}^< \,, \tag{B6}$$

with the scaling dimension being equivalent to the engineering dimension $[n]_\mathrm{eng} = [n_\mathrm{c}]_\mathrm{eng} = d$.

The rescaled interaction coupling is

$$g' = \frac{b^\epsilon g^<}{Z_\tau^<} \,, \tag{B7}$$

with $\epsilon = z - d$, which is here inferred from the relation $\mu = n_\mathrm{c} g$. The scale factors in (B5) and (B7) reflect the scaling dimensions of the couplings, which are in general different from the respective engineering dimensions,

$$[\mu] = z + \eta_\tau \,, \qquad [g] = \epsilon + \eta_\tau \,, \qquad [V] = -z + \eta_\tau \,. \tag{B8}$$

## Appendix C: Matsubara summation

In this appendix, we outline how the Matsubara sums encountered in this work can be performed analytically. To this end, we first introduce the notations for the Bose-Einstein distributions

$$n_\mathrm{B}(\omega_k) = \frac{1}{\exp(\beta \omega_k / Z_\tau) - 1} \,, \qquad n_\mathrm{B}^\pm(\omega_k^\pm) = \frac{1}{\exp\left(\beta \omega_k^\pm\right) - 1} \tag{C1}$$

where $\omega_k$ is the dispersion relation for $V = 0$ within the respective phase and $\omega_k^\pm$ is the dispersion in the case of $V \neq 0$ in the symmetry-broken phase (31). In the following, we drop the argument of the Bose-Einstein distributions and employ $n_k \equiv n_\mathrm{B}(\omega_k)$ and $n_k^\pm \equiv n_\mathrm{B}^\pm(\omega_k^\pm)$.

### 1. Symmetric phase

In the symmetric phase, no wave-function renormalization occurs at one-loop order and thus $Z_\tau^< = 1$ and $V^< = 0$. The propagator thus takes the form

$$G(k) = \frac{1}{\mathrm{i}\omega_n + \omega_k} \tag{C2}$$

with the free dispersion relation $\omega_k = \varepsilon_k - \mu$, cf. (20).

The sum over a single propagator, which determines the change of $\mu$ (23), results in the Bose-Einstein contribution:

$$\frac{1}{\beta} \sum_{\omega_n} G(k) = n_k \,. \tag{C3}$$

Here and in the following, summation over the Matsubara frequencies $\omega_n = 2\pi n T$ always goes from $n = -\infty$ to $n = \infty$. The one-loop correction to the four-point coupling $g$ (26) involves Matsubara sums over two propagators, which read

$$\begin{aligned} \frac{1}{\beta} \sum_{\omega_n} G(k) G(k) &= \beta n_k \left(1 + n_k\right) \,, \\ \frac{1}{\beta} \sum_{\omega_n} G(k) G(-k) &= \frac{1 + 2n_k}{2\omega_k} \,. \end{aligned} \tag{C4}$$

The main difference between these two sums is that the second one does not vanish in the zero temperature limit. This accounts for the fact that at vanishing temperatures only ladder diagrams, i.e., repeated scatterings between two initial particles, must be taken into account.

### 2. Condensed phase

In the symmetry-broken phase with $V \neq 0$, besides the normal, also an anomalous propagator must be taken into account:

$$\begin{aligned} G(k) &= \frac{Z_x \varepsilon_k + \mu - \mathrm{i} Z_\tau \omega_n + V \omega_n^2}{V^2 \left[\omega_n^2 + (\omega_k^-)^2\right] \left[\omega_n^2 + (\omega_k^+)^2\right]} \,, \\ \widetilde{G}(k) &= -\frac{\mu}{V^2 \left[\omega_n^2 + (\omega_k^-)^2\right] \left[\omega_n^2 + (\omega_k^+)^2\right]} \,. \end{aligned} \tag{C5}$$

The dispersion relations are defined in (30) and (31), where we already employed $\mu_1 = \mu_2 \equiv \mu$ as detailed in Sect. III B 3.

To shorten the forthcoming equations, let us first introduce a set of auxiliary functions. We start by defining

$$\Omega_k = \frac{1}{V^2 \left[\omega_n^2 + (\omega_k^-)^2\right] \left[\omega_n^2 + (\omega_k^+)^2\right]} \tag{C6}$$

and the general auxiliary function

$$K_{i,j} = \frac{1}{\beta} \sum_{\omega_n} \omega_n^{2j} \Omega_k^i \,. \tag{C7}$$

The auxiliary functions $K_{i,j}$ satisfy the relations

$$\begin{aligned} K_{i,j} &= V^2 K_{i+1,j+2} + A_k K_{i+1,j+1} + \omega_k^2 K_{i+1,j}\,, \\ \frac{\partial K_{i,j}}{\partial \ln\beta} &= -(1+2j)\,K_{i,j} + 2i\left(A_k K_{i+1,j+1} + 2V^2 K_{i+1,j+2}\right). \end{aligned} \tag{C8}$$

with $A_k = Z_\tau^2 + 2V(Z_x\varepsilon_k + \mu)$. The second relation can straightforwardly be used to express higher-order derivatives of $K_{i,j}$ in terms of the auxiliary functions (C7). Using the relations (C8), every auxiliary function with $i > 1$ can be expressed in terms of auxiliary functions and respective derivatives at $i = 1$. For our purposes, only two auxiliary function will be required explicitly:

$$\begin{aligned} K_{1,0} &= \frac{\omega_k^-\left(1+2n_k^+\right) - \omega_k^+\left(1+2n_k^-\right)}{2V^2\omega_k^-\omega_k^+\left[(\omega_k^-)^2 - (\omega_k^+)^2\right]}\,, \\ K_{1,1} &= \frac{\omega_k^-\left(1+2n_k^-\right) - \omega_k^+\left(1+2n_k^+\right)}{2V^2\left[(\omega_k^-)^2 - (\omega_k^+)^2\right]}\,. \end{aligned} \tag{C9}$$

Their first derivatives read

$$\begin{aligned} \frac{\partial K_{1,0}}{\partial \ln\beta} &= \beta\frac{n_k^-\left(1+n_k^-\right) - n_k^+\left(1+n_k^+\right)}{V^2\left[(\omega_k^-)^2 - (\omega_k^+)^2\right]}\,, \\ \frac{\partial K_{1,1}}{\partial \ln\beta} &= \beta\frac{(\omega_k^+)^2 n_k^+\left(1+n_k^+\right) - (\omega_k^-)^2 n_k^-\left(1+n_k^-\right)}{V^2\left[(\omega_k^-)^2 - (\omega_k^+)^2\right]}\,, \end{aligned} \tag{C10}$$

and likewise the second derivatives:

$$\begin{aligned} \frac{\partial^2 K_{1,0}}{\partial(\ln\beta)^2} &= \frac{\partial K_{1,0}}{\partial \ln\beta} + \beta^2\frac{\omega_k^+ n_k^+\left(1+n_k^+\right)\left(1+2n_k^+\right) - (+\leftrightarrow -)}{V^2\left[(\omega_k^-)^2 - (\omega_k^+)^2\right]}\,, \\ \frac{\partial^2 K_{1,1}}{\partial(\ln\beta)^2} &= \frac{\partial K_{1,1}}{\partial \ln\beta} + \beta^2\frac{(\omega_k^-)^3 n_k^-\left(1+n_k^-\right)\left(1+2n_k^-\right) - (+\leftrightarrow -)}{V^2\left[(\omega_k^-)^2 - (\omega_k^+)^2\right]}\,. \end{aligned} \tag{C11}$$

Here and in the following, for brevity, we indicate by $(+\leftrightarrow -)$ the respective exchange of signs of the dispersion relations and the Bose-Einstein distributions. In order to derive all auxiliary functions required in this work, we further derive the third derivatives which are

$$\begin{aligned} \frac{\partial^3 K_{1,0}}{\partial(\ln\beta)^3} &= 3\frac{\partial^2 K_{1,0}}{\partial(\ln\beta)^2} - 2\frac{\partial K_{1,0}}{\partial \ln\beta} \\ &\quad + \beta^3\frac{(\omega_k^-)^2 n_k^-\left(1+n_k^-\right)\left(1+6n_k^- + 6n_k^{-2}\right) - (+\leftrightarrow -)}{V^2\left[(\omega_k^-)^2 - (\omega_k^+)^2\right]}\,, \\ \frac{\partial^3 K_{1,1}}{\partial(\ln\beta)^3} &= 3\frac{\partial^2 K_{1,1}}{\partial(\ln\beta)^2} - 2\frac{\partial K_{1,1}}{\partial \ln\beta} \\ &\quad + \beta^3\frac{(\omega_k^+)^4 n_k^+\left(1+n_k^+\right)\left(1+6n_k^+ + 6n_k^{+2}\right) - (+\leftrightarrow -)}{V^2\left[(\omega_k^-)^2 - (\omega_k^+)^2\right]}\,. \end{aligned} \tag{C12}$$

Together with the relations (C8) this allows us to express all the relevant auxiliary functions in terms of Bose-Einstein distributions.

At this point, we can finally evaluate the Matsubara sums, starting with the single-propagator cases:

$$\begin{aligned} \frac{1}{\beta}\sum_{\omega_n} G(k) &= (Z_x\varepsilon_k + \mu)\,K_{1,0} + VK_{1,1} - \frac{1}{Z_\tau}\,, \\ \frac{1}{\beta}\sum_{\omega_n} \widetilde{G}(k) &= -\mu K_{1,0}\,. \end{aligned} \tag{C13}$$

In the case of two propagators, the Matsubara sums with only normal propagators are

$$\begin{aligned} \frac{1}{\beta}\sum_{\omega_n} G(k)G(k) &= K_{1,0} + \mu^2 K_{2,0} - 2Z_\tau^2 K_{2,1}\,, \\ \frac{1}{\beta}\sum_{\omega_n} G(k)G(-k) &= K_{1,0} + \mu^2 K_{2,0}\,, \end{aligned} \tag{C14}$$

and the ones with at least one anomalous propagator read

$$\begin{aligned} \frac{1}{\beta}\sum_{\omega_n} G(k)\widetilde{G}(k) &= -\mu\left(Z_x\varepsilon_k + \mu\right)K_{2,0} - \mu V K_{2,1}\,, \\ \frac{1}{\beta}\sum_{\omega_n} \widetilde{G}(k)\widetilde{G}(k) &= \mu^2 K_{2,0}\,. \end{aligned} \tag{C15}$$

Together, Eqs. (C13)–(C15) yield the relation

$$\frac{\mu}{\beta}\sum_{\omega_n}\left(\widetilde{G}(k)\widetilde{G}(k) - G(k)G(-k)\right) = \frac{1}{\beta}\sum_{\omega_n}\widetilde{G}(k)\,, \tag{C16}$$

which proves useful in showing the Hugenholtz-Pines theorem, $\mu_1 = \mu_2$, in (41).

Corrections to the wave-function renormalizations involve first- and second-order frequency-derivatives of the propagators, cf. (52). In terms of the auxiliary functions the latter read

$$\begin{aligned} \dot{G}(k) &= \left(\mathrm{i}Z_\tau + 2V\omega_n\right)\Omega_k - 2\omega_n\left(2V^2\omega_n^2 + A_k\right)\Omega_k G(k)\,, \\ \dot{\widetilde{G}}(k) &= -2\omega_n\left(2V^2\omega_n^2 + A_k\right)\Omega_k\widetilde{G}(k)\,, \end{aligned} \tag{C17}$$

and

$$\begin{aligned} \ddot{G}(k) &= 2\Omega_k\left[V - \left(6V^2\omega_n^2 + A_k\right)G(k)\right. \\ &\qquad \left. - 2\omega_n\left(2V^2\omega_n^2 + A_k\right)\dot{G}(k)\right], \\ \ddot{\widetilde{G}}(k) &= -2\Omega_k\left[\left(6V^2\omega_n^2 + A_k\right)\widetilde{G}(k)\right. \\ &\qquad \left. + 2\omega_n\left(2V^2\omega_n^2 + A_k\right)\dot{\widetilde{G}}(k)\right], \end{aligned} \tag{C18}$$

respectively. These derivatives can now be used when computing the Matsubara sums, starting with the first derivatives of the normal propagator

$$\begin{aligned} \frac{1}{\beta}\sum_{\omega_n} G(k)\dot{G}(k) &= \mathrm{i}Z_\tau\left[\left(Z_x\varepsilon_k + \mu\right)K_{2,0} - 5VK_{2,1}\right] \\ &\quad - 4\mathrm{i}Z_\tau\left\{\left[Z_\tau^2\left(Z_x\varepsilon_k + \mu\right) + 2V\mu^2\right]K_{3,1} - Z_\tau^2 V K_{3,2}\right\}, \\ \frac{1}{\beta}\sum_{\omega_n} G(k)\dot{G}(-k) &= \mathrm{i}Z_\tau\left[VK_{2,1} - \left(Z_x\varepsilon_k + \mu\right)K_{2,0}\right], \end{aligned} \tag{C19}$$

and continuing with those of the anomalous propagator

$$\frac{1}{\beta}\sum_{\omega_n} G(k)\dot{\widetilde{G}}(k) = 2\mathrm{i}Z_\tau\mu\left(A_k K_{3,1} + 2V^2 K_{3,2}\right),$$
$$\frac{1}{\beta}\sum_{\omega_n} \widetilde{G}(k)\dot{\widetilde{G}}(k) = 0\,. \tag{C20}$$

We furthermore need the Matsubara sums containing second derivatives in order to determine the flow equation for $V$. For the second derivative of the normal propagator, we find

$$\begin{aligned}\frac{1}{\beta}\sum_{\omega_n} G(k)\ddot{G}(k) &= -\left(Z_\tau^2 + A_k\right)K_{2,0} + 6V^2 K_{2,1} - 2A_k\mu^2 K_{3,0}\\ &\quad + 4\left(5Z_\tau^2 A_k + 9V^2\mu^2 - Z_\tau^4\right)K_{3,1} - 24V^2 Z_\tau^2 K_{3,2}\\ &\quad + 8\left(A_k^2 - 4V^2\omega_k^2\right)\left(\mu^2 K_{4,1} - 2Z_\tau^2 K_{4,2}\right),\\ \frac{1}{\beta}\sum_{\omega_n} G(k)\ddot{G}(-k) &= -\left(Z_\tau^2 + A_k\right)K_{2,0} + 6V^2 K_{2,1}\\ &\quad - 2A_k\mu^2 K_{3,0} + 4\left(A_k^2 - 4V^2\omega_k^2 + 5V^2\mu^2\right)K_{3,1}\\ &\quad + 8\mu^2\left(A_k^2 - 4V^2\omega_k^2\right)K_{4,1}\,,\end{aligned} \tag{C21}$$

and for the anomalous propagator

$$\begin{aligned}\frac{1}{\beta}\sum_{\omega_n} G(k)\dddot{\widetilde{G}}(k) &= 2\mu A_k\left[\left(Z_x\varepsilon_k + \mu\right)K_{3,0} + VK_{3,1}\right]\\ &\quad - 20V^2\mu\left[\left(Z_x\varepsilon_k + \mu\right)K_{3,1} + VK_{3,2}\right]\\ &\quad + 8\mu\left(4V^2\omega_k^2 - A_k^2\right)\left[\left(Z_x\varepsilon_k + \mu\right)K_{4,1} + VK_{4,2}\right],\\ \frac{1}{\beta}\sum_{\omega_n} \widetilde{G}(k)\dddot{\widetilde{G}}(k) &= -2\mu^2\left(A_k K_{3,0} - 10V^2 K_{3,1}\right)\\ &\quad + 8\mu^2\left(A_k^2 - 4V^2\omega_k^2\right)K_{4,1}\,.\end{aligned} \tag{C22}$$

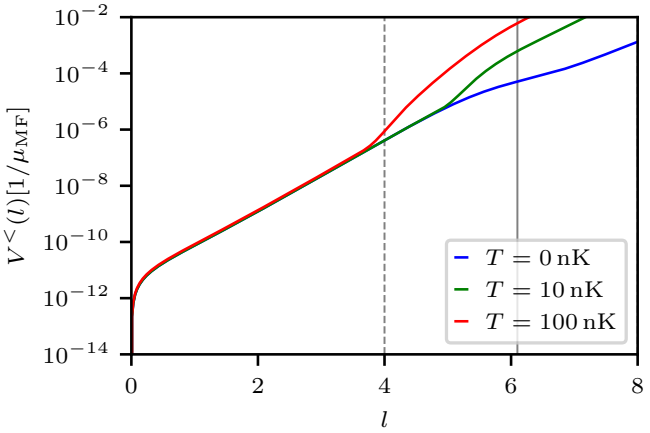



FIG. 7. The wave-function renormalization $V^<(l)$ in units of the inverse mean-field chemical potential is plotted against the flow parameter $l$ for three different temperatures. The healing scale is shown by the solid, vertical, gray line, whereas the dashed line separates the first two distinct intervals of the RG flow.

### Appendix D: Flow of $V$

In Fig. 7 the flow of the wave-function renormalization $V^<(l)$ is shown for three different temperatures in units of the inverse mean-field chemical potential. Within the first distinct interval below $l_1 < 4$, where the feedback of $V$ on the other couplings is neglected, we observe a temperature independent growth. This is succeeded by the second interval $l_1 \leq l < l_2 = 8$ around the healing scale in which the full feedback of $V$ is taken into account and a temperature dependent flow is found. In the asymptotic regime $l \geq l_2$ the wave-function renormalization $V$ continues growing, while having a decreasing $Z_\tau^<$ for $T \neq 0$, cf. Fig. 4. Thus, the action (7) turns approximately SO($d+1$) symmetric for $l \gg l_2$ [30, 42]. Similarly to [30], we do not observe qualitative modifications between $V = 0$ and $V \neq 0$ for the couplings, the condensate depletion, and the shift in critical temperature, cf. Fig. 5 and Fig. 6, respectively. This relates to the fact that in $d = 3$ the IR divergences in our theory only scale logarithmically for $k \to 0$.


[1] M. Bijlsma and H. T. C. Stoof, Phys. Rev. A **54**, 5085 (1996).
[2] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, Science **269**, 198 (1995).
[3] K. B. Davis, M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, Phys. Rev. Lett. **75**, 3969 (1995).
[4] N. N. Bogoliubov, J. Phys. (USSR) **11**, 23 (1947).
[5] T. D. Lee, K. Huang, and C. N. Yang, Phys. Rev. **106**, 1135 (1957).
[6] S. T. Beliaev, Sov. Phys. JETP **7**, 299 (1958).
[7] S. T. Beliaev, Sov. Phys. JETP **7**, 289 (1958).
[8] B. Capogrosso-Sansone, S. Giorgini, S. Pilati, L. Pollet, N. Prokof'ev, B. Svistunov, and M. Troyer, New J. Phys. **12**, 043010 (2010).
[9] J. O. Andersen, Rev. Mod. Phys. **76**, 599 (2004).
[10] J. Gavoret and P. Nozières, Ann. Phys. (NY) **28**, 349 (1964).
[11] A. A. Nepomnyashchii and Y. A. Nepomnyashchii, JETP Lett. **21**, 1 (1975).
[12] Y. A. Nepomnyashchii and A. A. Nepomnyashchii, Sov. Phys. JETP **48**, 493 (1978).
[13] N. M. Hugenholtz and D. Pines, Phys. Rev. **116**, 489 (1959).
[14] P. B. Weichman, Phys. Rev. B **38**, 8739 (1988).
[15] T. Haugset, H. Haugerud, and F. Ravndal, Ann. Phys. (NY) **266**, 27 (1998).
[16] V. N. Popov, *Functional integrals and collective excitations* (CUP, Cambridge, UK, 1987).
[17] S. A. Morgan, J. Phys. B **33**, 3847 (2000).

[18] S. A. Gardiner and T. P. Billam, "Number-conserving approaches for atomic Bose-Einstein condensates: An overview," in *Quantum Gases* (Imperial College Press, 2013) pp. 133–145.
[19] V. I. Yukalov and H. Kleinert, Phys. Rev. A **73**, 063612 (2006).
[20] H. T. C. Stoof and J. J. R. M. van Heugten, J. Low Temp. Phys. **174**, 159 (2014).
[21] K. G. Wilson, Phys. Rev. B **4**, 3184 (1971).
[22] K. G. Wilson, Phys. Rev. B **4**, 3174 (1971).
[23] K. G. Wilson and J. Kogut, Phys. Rep. **12**, 75 (1974).
[24] K. G. Wilson, Rev. Mod. Phys. **55**, 583 (1983).
[25] J. Polchinski, Nucl. Phys. B **231**, 269 (1984).
[26] C. Wetterich, Phys. Lett. **B301**, 90 (1993).
[27] G. Benfatto, in *Constructive Physics Results in Field Theory, Statistical Mechanics and Condensed Matter Physics*, edited by V. Rivasseau (Springer Berlin Heidelberg, Berlin, Heidelberg, 1995) pp. 219–247.
[28] G. Metikas, O. Zobay, and G. Alber, Phys. Rev. A **69**, 043614 (2004).
[29] O. Zobay, Phys. Rev. A **73**, 023616 (2006).
[30] S. Floerchinger and C. Wetterich, Phys. Rev. A **77**, 053603 (2008).
[31] S. Floerchinger and C. Wetterich, Phys. Rev. A **79**, 063602 (2009).
[32] S. Ledowski, N. Hasselmann, and P. Kopietz, Phys. Rev. A **69**, 061601 (2004).
[33] N. Hasselmann, S. Ledowski, and P. Kopietz, Phys. Rev. A **70**, 063621 (2004).
[34] J.-P. Blaizot, R. Méndez-Galain, and N. Wschebor, Phys. Rev. E **74**, 051116 (2006).
[35] J.-P. Blaizot, R. Méndez-Galain, and N. Wschebor, Phys. Rev. E **74**, 051117 (2006).
[36] J. O. Andersen and M. Strickland, Phys. Rev. A **60**, 1442 (1999).
[37] G. Metikas, O. Zobay, and G. Alber, J. Phys. B: At. Mol. Opt. Phys. **36**, 4595 (2003).
[38] G. Alber and G. Metikas, Appl. Phys. B **73**, 773 (2001).
[39] G. Alber, Phys. Rev. A **63**, 023613 (2001).
[40] C. Castellani, C. Di Castro, F. Pistolesi, and G. C. Strinati, Phys. Rev. Lett. **78**, 1612 (1997).
[41] F. Pistolesi, C. Castellani, C. Di Castro, and G. C. Strinati, Phys. Rev. B **69**, 024513 (2004).
[42] C. Wetterich, Phys. Rev. B **77**, 064504 (2008).
[43] N. Dupuis and K. Sengupta, Europhys. Lett. **80**, 50007 (2007).
[44] N. Dupuis, Phys. Rev. Lett. **102**, 190401 (2009).
[45] N. Dupuis, Phys. Rev. A **80**, 043627 (2009).
[46] A. Sinner, N. Hasselmann, and P. Kopietz, Phys. Rev. A **82**, 063632 (2010).
[47] C. Eichler, N. Hasselmann, and P. Kopietz, Phys. Rev. E **80**, 051129 (2009).
[48] F. Isaule, M. C. Birse, and N. R. Walet, Phys. Rev. B **98**, 144502 (2018).
[49] F. Isaule, M. C. Birse, and N. R. Walet, Ann. Phys. (NY) **412**, 168006 (2020).
[50] M. Crisan, D. Bodea, I. Grosu, and I. Tifrea, J. Phys. A **35**, 239 (2002).
[51] G. Baym, J.-P. Blaizot, M. Holzmann, F. Laloë, and D. Vautherin, Phys. Rev. Lett. **83**, 1703 (1999).
[52] P. Grüter, D. Ceperley, and F. Laloë, Phys. Rev. Lett. **79**, 3549 (1997).
[53] M. Holzmann and W. Krauth, Phys. Rev. Lett. **83**, 2687 (1999).
[54] V. A. Kashurnikov, N. V. Prokof'ev, and B. V. Svistunov, Phys. Rev. Lett. **87**, 120402 (2001).
[55] P. Arnold and G. Moore, Phys. Rev. Lett. **87**, 120401 (2001).
[56] K. Nho and D. P. Landau, Phys. Rev. A **70**, 053614 (2004).
[57] P. Heinen and T. Gasenzer, Phys. Rev. A **106**, 063308 (2022).
[58] R. Guida and J. Zinn-Justin, J. Phys. A **31**, 8103 (1998).
[59] A. Singasaas and G. Ahlers, Phys. Rev. B **30**, 5103 (1984).
[60] A. Pelissetto and E. Vicari, Phys. Rep. **368**, 549 (2002).
[61] Y. Kawaguchi and M. Ueda, Phys. Rep. **520**, 253 (2012).
[62] M. Ueda, Nature Rev. Phys. **2**, 669 (2020).
[63] T. Köhler, K. Góral, and P. S. Julienne, Rev. Mod. Phys. **78**, 1311 (2006).
[64] P. Kopietz, L. Bartosch, and F. Schütz, *Introduction to the functional renormalization group*, Vol. 798 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010).
[65] D. Boyanovsky, S.-Y. Wang, D.-S. Lee, H.-L. Yu, and S. Alamoudi, Ann. Phys. (NY) **300**, 1 (2002).
[66] C. Samuelis, E. Tiesinga, T. Laue, M. Elbs, H. Knöckel, and E. Tiemann, Phys. Rev. A **63**, 012710 (2000).
[67] See [69] for more recent data on scattering lengths, which would, however, not change our results here.
[68] I. Boettcher, J. M. Pawlowski, and S. Diehl, Nucl. Phys. Proc. Suppl. **228**, 63 (2012).
[69] S. Knoop, T. Schuster, R. Scelle, A. Trautmann, J. Appmeier, M. K. Oberthaler, E. Tiesinga, and E. Tiemann, Phys. Rev. A **83**, 042704 (2011).

arXiv:2402.19238v1 [astro-ph.HE] 29 Feb 2024

# Bosonic Dark Matter Dynamics in Hybrid Neutron Stars

Zakary Buras-Stubbs and Ilídio Lopes
*Centro de Astrofísica e Gravitação - CENTRA,*
*Departamento de Física,*
*Instituto Superior Técnico - IST,*
*Universidade de Lisboa - UL,*
*Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal**

This research studies the intricate interplay between dark and baryonic matter within hybrid neutron stars enriched by anisotropic bosonic dark matter halos. Our modelling, guided by the equation of state with a free parameter, reveals diverse mass-radius correlations for these astronomical objects. A pivotal result is the influence of dark matter characteristics—whether condensed or dispersed—on the observable attributes of neutron stars based on their masses. Our investigation into anisotropic models, which offer a notably authentic representation of dark matter anisotropy, reveals a unique low-density core halo profile, distinguishing it from alternative approaches. Insights gleaned from galactic clusters have further refined our understanding of the bosonic dark matter paradigm. Observational constraints derived from the dynamics of galaxy clusters have been fundamental in defining the dark matter particle mass to lie between 0.05 GeV and 0.5 GeV and the scattering length to range from 0.9 fm to 3 fm. Using terrestrial Bose-Einstein condensate experiments, we have narrowed down the properties of bosonic dark matter, especially in the often overlooked 3 to 30 GeV mass range. Our findings fortify the understanding of dark and baryonic matter synergies in hybrid neutron stars, establishing a robust foundation for future astrophysical pursuits.

## I. INTRODUCTION

Dark matter, one of the most intriguing and mysterious phenomena in astrophysics, remains a subject of ongoing investigation [e.g., 1, 2]. This invisible form of matter, which does not emit, absorb, or reflect electromagnetic radiation, accounts for roughly 27% of the Universe's total mass-energy content [e.g., 3]. The presence of dark matter has been deduced from its gravitational influence on galaxy movements, the vast cosmic structures, and the cosmic microwave background radiation [e.g., 4].

It is widely accepted that dark matter cannot be composed of fundamental particles within the Standard Model of particle physics, and it is more likely that dark matter is made up of particles that have not yet been discovered. Various hypothetical particle candidates have been proposed as potential solutions to the dark matter problem, and these particles can generally be classified into several categories based on their mass [5–8]. These categories include fuzzy dark matter or axion-like particles, axions, sterile neutrinos, and weakly interacting massive particles (WIMPs).

Our focus in this work is to investigate how the structure of neutron stars can be affected by dark matter particles within the WIMPs and axion mass ranges. In particular, we will be focusing on particles in the $10^{-2}$ – $10^{2}$ GeV range.

There are several papers in the literature that discuss the relationship between dark matter and neutron stars. Some recent examples include [9, 10], which focus on general self-annihilating WIMP models; [11–21], which focus on specific fermionic dark matter models; [22, 23], that consider Bose-Einstein Condensate (BEC) dark matter; and [24] where both fermionic and BEC models are studied. These papers explore the possible effects of dark matter on neutron star properties and the potential for dark matter detection through observations of neutron stars. Additionally, there are several reviews that provide a more comprehensive overview of the field, such as Lattimer [25] and Del Popolo et al. [26].

Anisotropic matter exhibits different properties or behaviours depending on the direction [for a review see 27, 28]. Such behaviour can lead to pressure differences in various directions within neutron stars or other astronomical objects [e.g., 29–31]. This characteristic influences how matter interacts with surrounding objects, affecting their formation, structure, and evolution, shaping our understanding of dark matter. For instance, dark matter clouds are expected to exhibit local anisotropy, similar to any collisionless system of particles. Researchers have extensively studied these systems, especially in the context of galaxy dynamics [e.g., 32–34].

The theory of anisotropic fluids in General Relativity is well-established. Past research has demonstrated that anisotropic fluids could be geodesic in general relativity [35]. A comprehensive study of spherically symmetric dissipative anisotropic fluids has been presented, and exact static spherically symmetric anisotropic solutions of field equations have been obtained and analyzed by Herrera et al. [36], Bayin [37], and Herrera et al. [38]. Furthermore, calculations of anisotropic stars in general relativity and their mass-radius relations have been conducted Mak and Harko [39]. A detailed collection of

* zburasstubbs@tecnico.ulisboa.pt

recent articles on this topic is available on Kumar and Bharti [28].

This study introduces a novel astrophysical object: a hybrid compact star composed of a neutron star within an anisotropic dark matter halo. We will investigate the unique characteristics of these models to gain a deeper understanding of their formation. By examining the properties of dark matter, we propose that anisotropic dark matter may play a vital role in forming these hybrid compact stars. Our analysis focuses on a static, spherically symmetric configuration featuring a neutron star enveloped by dark matter. Analogous to the baryonic matter within the neutron star, the properties and distribution of the dark matter, positioned at a specific distance from the star's centre, are described using the Tolman-Oppenheimer-Volkoff (TOV) equations.

We assume that the dark matter envelope does not interact with the neutron star's baryonic matter within the envelope. We model how dark matter affects neutron star observations using the theory of compact stars. We focus on two main assumptions: first, dark matter with a nonzero radial pressure term in the TOV equations, $P_{r\chi}(r) = P_{\chi}(r)$, and second, an anisotropic energy-momentum tensor leading to a nonzero tangential pressure, $P_{\perp\chi}(r)$. We justify these choices and describe the dark matter distribution by calculating $\Pi_{\chi}(r) = P_{\perp\chi}(r) - P_{r\chi}(r)$.

In this study, we concentrate on anisotropic dark matter's effects on neutron stars within the dark matter halo, examining the mass-radius relation. Furthermore, we assume that the dark matter particles are bosons. We compare the dark matter density profiles and mass-radius relations to those with isotropic dark matter and a neutron star in a vacuum.

In this article, we structure our presentation as follows. We begin with a summary of current research topics on neutron stars and dark matter in this section. In Section II, we provide an overview of the hybrid model employed in this study. Subsequently, in Section III, we discuss the properties of the equation of state for neutron stars and anisotropic dark matter. In Section IV, we investigate the properties of dark matter halos formed around neutron stars. In Section V, we examine constraints on the properties of dark matter particles. Finally, we summarize our findings and present our conclusions in Section VI.

Here, unless specified otherwise, we utilize natural units, where the speed of light ($c$), the Boltzmann constant ($k_B$) and the reduced Planck constant ($\hbar$) are set to 1. Through standard conversion rules, all conventional units can be expressed in terms of GeV. The most frequently used units in this work are: 1 centimetre = $5.05 \times 10^{13}$ GeV$^{-1}$, 1 gram = $5.610 \times 10^{23}$ GeV and 1 second = $1.519 \times 10^{24}$ GeV$^{-1}$.

## II. PROPERTIES OF NEUTRON STARS AND ANISOTROPIC DARK MATTER

Here we start by briefly summarising the equations governing the components of our admixed neutron star.

### A. Tolman–Oppenheimer–Volkoff Equations

We identify that the action governing our system will essentially take the form

$$S = S_G + S_M, \tag{1}$$

for which $S_G$ defines the gravitational component, and hence is given by the usual Einstein-Hilbert action, and $S_M$ encapsulates matter terms associated with a perfect fluid whose properties are fully determined by its energy density $\rho$, its radial pressure $P_r$, its tangential pressure $P_\perp$, as well as an equation of state defined such that $F(\rho, P_r) = 0$ [e.g., 40].

From this derivation, by varying with respect to the metric, we obtain the standard Einstein field equations without the inclusion of a cosmological constant:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi T_{\mu\nu} \tag{2}$$

Here, Newton's constant G is set equal to unity. The total stress-energy tensor can be expressed as:

$$T^{\mu}_{\nu} = \mathrm{Diag}\left(-\rho, P_r, P_\perp, P_\perp\right). \tag{3}$$

In the comoving frame, the physical matter content consists of an anisotropic fluid with energy density $\rho$, radial pressure $P_r$, and tangential pressure $P_\perp$.

We define a general metric tensor as follows:

$$ds^2 = -e^{\nu} dt^2 + e^{\lambda} dr^2 + r^2 d\Omega^2, \tag{4}$$

where $\nu(r)$ and $\lambda(r)$ are the metric potentials, depending only on the radial coordinate, and $d\Omega^2 \equiv (d\theta^2 + \sin^2\theta\, d\phi^2)$ represents the solid angle element [e.g., 41]. By solving Einstein's equations, we find that:

$$e^{\lambda(r)} = \frac{1}{1 - \frac{2m}{r}}, \tag{5}$$

and

$$\nu'(r) = \frac{2m + 8\pi r^3 P_r}{r^2 - 2mr}. \tag{6}$$

Moreover, we obtain the hydrostatic equilibrium equation [42, 43], also referred to as the generalized Tolman–Oppenheimer–Volkoff (TOV) equation:

$$\frac{dP_r}{dr} = -\frac{(\rho + P_r)(m + 4\pi r^3 P_r)}{r(r - 2m)} + \frac{2\Pi}{r}, \tag{7}$$

where $\Pi \equiv P_\perp - P_r$. It is important to note that for an isotropic substance, we have $\Pi = 0$ or $P_\perp = P_r$. In our calculations for the baryonic component, we will enforce this condition, causing the final term to vanish and providing the hydrostatic equilibrium equation, also known as the generalised TOV equation.
Within this a natural interpretation of the calculations we have performed becomes manifest. Recognising that hydrostatic forces are described by $F_h = -\frac{dP_r}{dr}$, gravitational forces by $F_g = -\frac{(\rho+P_r)(m+4\pi r^3 P_r)}{r(r-2m)}$, and anisotropic forces are given by $F_a = \frac{2\Pi}{r}$, then it becomes clear that equation (7) is just the statement that

$$F_g + F_h + F_a = 0. \tag{8}$$

The total mass, $m(r)$, is computed using the standard differential equation relating mass and density:

$$\frac{dm}{dr} = 4\pi r^2 \rho. \tag{9}$$

In accordance with a prevalent approach in the literature, as seen in the works of Panotopoulos and Lopes [44], Ciarcelluti and Sandin [45], Sandin and Ciarcelluti [46], we assume that the interaction between dark matter and the baryonic matter is so weak that it can be reasonably neglected. Therefore, we consider the fluid to have two components: a baryonic component ($B$) representing neutron star matter and a non-baryonic dark matter ($\chi$) component. For instance, we define $\rho = \rho_B + \rho_\chi$, where $\rho_B$ is the energy density associated with standard baryonic matter ($B$) and $\rho_\chi$ is related to dark matter ($\chi$). Consequently, we assume that baryonic matter ($B$) and dark matter ($\chi$) are coupled solely through gravity, with their energy-momentum tensors conserved separately. As a result, the previous system of equations (7) and (9) is split into two components:

$$\frac{dP_{ri}}{dr} = -\frac{(\rho_i + P_{ri})(m + 4\pi r^3 P_r)}{r(r-2m)} + \frac{2\Pi_i}{r}, \tag{10}$$

where $\Pi_i = P_{ri} - P_{\perp i}$, and $P_r = P_{rB} + P_{r\chi}$, accompanied by the additional equation:

$$\frac{dm_i}{dr} = 4\pi r^2 \rho_i. \tag{11}$$

We see that our model reproduces the TOV equations found in the literature, [e.g., 12, 47], where the effects of anisotropies only appears as an additive term, as one would intuit from equation (8).

In the aforementioned equations, the subscript index $i = B$ and $i = \chi$ denote baryonic matter ($B$) and dark matter ($\chi$), respectively. Therefore, $\rho_i$, $P_{ri}$, and $P_{\perp i}$ correspond to the density and pressure terms of the $i-$th component, where $r$ represents a purely radial component, and $\perp$ a tangential one. However, in our study, we assume that the baryonic matter is isotropic, meaning that $P_{\perp B}(r) = P_{rB}(r)$ and, as a result, $\Pi_B(r) = 0$. The gravitational mass, $m$, is the sum of the masses of both components, that is, $m(r) = m_B(r) + m_\chi(r)$. The system of equations (10), along with equation (11), represents a generalization of equations (7) and (9).
We determine the radial density profiles of baryonic matter ($B$) and dark matter ($\chi$) using Equations (10) and (11) by applying two central and two boundary conditions for the star. The first set of conditions relates to the central densities of $B$ and $\chi$, while the second set ensures hydrostatic equilibrium for each matter component at their respective boundaries, forming spheres with radii $R_B$ and $R_\chi$ containing all the $B$ and $\chi$ mass components, respectively. Generally, $R_B \neq R_\chi$. Hydrostatic equilibrium is independently maintained for $B$ and $\chi$ by the conditions:

$$P_i(R_i) = 0. \tag{12}$$

The $\chi$ component cannot be directly observed, so it is reasonable to identify the neutron star radius $R$ with $R_B$, i.e., $R_B = R$, where $R$ represents the observed radius of the neutron star. We choose to define our radius like this so that our analysis may be compared with observational data, as the neutron stars will have a visible radius in the usual sense, with an additional hidden mass. The total gravitational mass and the fraction of $\chi$ inside the neutron star are defined as:

$$M_T = M_B(R_B) + M_\chi(R_\chi) \tag{13}$$

and

$$f_\chi = \frac{M_\chi(R_\chi)}{M_T}. \tag{14}$$

It is evident that by varying the central densities of baryonic matter ($B$) and dark matter ($\chi$), we can obtain different values of the total gravitational mass $M_T$ and neutron star radius $R$ for a given $\chi$ fraction $f_\chi$.

## III. EQUATIONS OF STATE AND ANISOTROPIC DARK MATTER

### A. Microphysics Within Neutron Stars

In our study of hybrid stars, we aim to accurately depict the complex physics inside neutron stars at various densities and temperatures. Our chosen microscopic model is based on a comprehensive framework suitable for the entire high-energy Quantum Chromodynamics (QCD) range. The model covers various temperatures and densities, considering different levels of isospin and strangeness. Notably, this model is consistent with known requirements for compact stars [e.g., 48–50] and findings from collider experiments [e.g., 51],

like those based on Lattice QCD. We mainly focus on the requirements of cold, dense stars, the behaviour of evenly distributed nuclear matter, and the conditions of high-temperature QCD at very low or no density. Moreover, this model also clearly explains how chiral symmetry is restored and how quarks are deconfined, supporting data from lattice QCD [e.g., 52], perturbative QCD [e.g., 53] , and high-energy collider tests.

Addressing all aspects above, we have chosen an equation of state computed within the framework of the Chiral Mean Field (CMF) model. This model is grounded in the three-flavour chiral Lagrangian for hadronic matter expanded to encompass neutron stars [54, 55]. In our study, we utilise a version of the CMF model augmented by Motornenko *et. al.*, as detailed in their article [56, 57]. This enhanced model encompasses a broad spectrum of QCD degrees of freedom. The model includes, beyond protons and neutrons, hyperons and their parity partners, as well as a comprehensive array of hadronic resonances (encompassing both strange and non-strange baryons and mesons), as catalogued in the Particle Data Book [e.g., 58]. Additionally, this model incorporate the thermal effects of deconfined quarks and gluons as per the PNJL model [59]. When combined with electrons, this represents the most comprehensive range of QCD degrees of freedom for the high-density equation of state. Furthermore, the model enables the depiction of nuclear matter in a low-density regime, specifically for densities below $10^{-2}$ $n_{\rm sat}$, which are prevalent in binary neutron star mergers [60].

This comprehensive approach enables the calculation of a unique equation of state for nuclear matter valid simultaneously in heavy ion collisions and binary neutron star mergers without the need for additional adjustments that might arise from using different and potentially inconsistent equations of state for distinct nuclear matter scenarios.

The equation of state version used in our study effectively simulates a crossover transition for deconfinement at finite and zero density, as indicated by lattice QCD [57, 61]. It accurately represents hadrons in medium, nuclei, nuclear matter, and neutron stars. This model generates a realistic nuclear ground state with these properties: saturation at a baryon number density of $\mathrm{n_{sat}} = 0.15$ fm$^{-3}$, binding energy per nucleon of $E_0/B = -15.2$ MeV, symmetry energy of $S_0 = 31.9$ MeV, symmetry energy slope of $L = 57$ MeV, and incompressibility of $K_0 = 267$ MeV [e.g., 56, 62]. The model can reproduce neutron stars heavier than $2M_\odot$ and stars with sizes within the limits established by LIGO-Virgo [63] and NICER [64].

### B. Neutron Star equation of state

The neutron star equation of state (EOS) we use is based on the Chiral Mean-Field (CMF) model presented in [56]. A crust is simulated using [65] by finding the intersection of the EOSs and merging them to determine the surface. This model has been chosen to describe our baryonic matter primarily because it is effective in producing neutron stars with masses and radii close to those observed by the LIGO/VIRGO collaborations, such as [63], and by NICER, [64]. Before the introduction of dark matter, the maximum neutron star mass with this model is $2.15\mathrm{M}_\odot$ with a radius of 12.0km.
This, however, is not the only choice of EOS that could have been made, as countless viable models exist in the literature, for example the induced surface tension (IST) model, [66], or indeed other relativistic mean field approaches, e.g. [67].

### C. Dark Matter equation of state

Here, we introduce a relatively simple bosonic dark matter model, with an effective Lagrangian of the form

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_\chi^2\phi^2 - \lambda\phi^4. \tag{15}$$

where $\lambda$ relates to the scattering length, $l_\chi$, by $\lambda = \frac{4\pi l_\chi}{m_\chi}$
A coherent wavefunction for the condensate can then be produced by considering the following ansatz,

$$\psi = \phi e^{im_\chi t}, \tag{16}$$

from which the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{2}g_{\mu\nu}\partial^\mu\psi^*\partial^\nu\psi - \frac{m_\chi^2}{2}|\psi|^2 - \lambda|\psi|^4. \tag{17}$$

By considering that the object rotates sufficiently slowly that the rotational motion of the condensate may be neglected, then we receive the following equations of motion,

$$i\frac{\partial\psi}{\partial t} = -\frac{1}{2m_\chi}\nabla^2\psi + m_\chi V\psi + \frac{4\pi l_\chi}{m_\chi^2}|\psi|^2\psi \tag{18}$$

$$\nabla^2 V = 4\pi G|\psi|^2 \tag{19}$$

known generally as the Gross-Pitaevskii-Poisson Equations, where $V$ is the gravitational potential. Following the methodology in Böhmer and Harko [68], one sees that for an interaction term of the form found in equation (17), and by identifying that $|\psi|^2 \sim \rho$ ,the dynamics of the condensate reduce to that of a poltrope with the following equation of state:

$$P_{r\chi} = A_\chi\,\rho_\chi^2, \tag{20}$$

where

$$A_\chi = \frac{\lambda}{2m_\chi^2} = \frac{2\pi l_\chi}{m_\chi^3} = 1.25\ 10^5\left(\frac{l_\chi}{1\ \mathrm{fm}}\right)\left(\frac{1\ \mathrm{GeV}}{m_\chi}\right)^3, \tag{21}$$

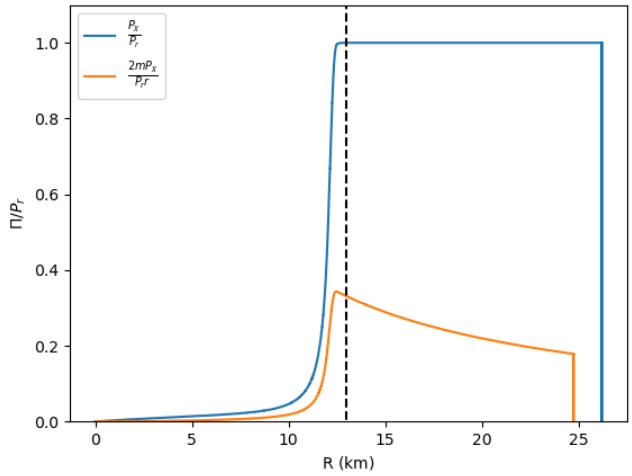



Figure 1. The relative anisotropy of the system $\frac{2\Pi_\chi}{P_r}(r)$ (where $\Pi_\chi$ is given by equation (22)) as a function of radius for $\alpha_\chi = 1$, $A_\chi = 10^7$, and for an object of total mass $M \sim 1.5 M_\odot$. $\Pi_\chi(r)$ is computed from the center of the neutron star to the total radius $R_\chi$ of the halo for the two forms of anistotropy considered in equations (23) and (24). The dashed line indicates the very edge of the neutron star, located at approximately 13.0 km, with $\frac{2\Pi_\chi}{P_r}(r)$ peaking for both models at around 12.6 km.

for which $m_\chi$ represents the mass of a dark matter particle, and $l_\chi$ denotes its scattering length. Throughout this work, including all tables, figures, and the main text, the units for $A_\chi$ are presented in the cgs system unless specified otherwise: $\mathrm{g}^{-1}\ \mathrm{cm}^5\ \mathrm{s}^{-2}$. In this case, the dark matter condensate's equation of state is given by $P(\rho)$, as provided by equation (20). From this equation of state, it follows that $P_{r\chi} \propto \rho_\chi^2$. A general polytropic equation of state can be written as $P \propto \rho_\chi^{1+1/n}$, where $n = 1$ is the polytropic index. For the time being we only concern ourselves with the particular values of $A_\chi$ and the impact that this will have on neutron star structure.
Later we will consider the implications of this on the possible range of masses that dark matter may take on these regimes. More detailed information about this equation of state in the context of dark matter may be found in Li et al. [69], Harko and Lobo [70], Panotopoulos and Lopes [44], and Crăciun and Harko [71].

### D. Anisotropic Dark Matter

Anisotropy in stars can originate from various sources. For instance, boson stars are inherently anisotropic astronomical objects [72], and the energy-momentum tensors of both electromagnetic and fermionic fields are intrinsically anisotropic [73]. Local anisotropy could arise from the presence of viscosity [e.g., 74] or from a two-fluid mixture [e.g., 37, 75, 76]. In this context, several authors have recently developed stellar models incorporating anisotropic matter [47, 77, 78].

This study aims to investigate the consequences of introducing anisotropy into a dark matter model. As mentioned earlier, the motivation for including such an effect can be attributed to the bosonic nature of dark matter particles, which inherently exhibit anisotropy.

A viable approach to incorporating anisotropy in the dark matter model has been explored by Horvat et al. [79] and others [80, 81]. In this approach, the anisotropy function $\Pi_\chi(r)$ is defined as:

$$\Pi_\chi = P_{\perp\chi} - P_{r\chi} = \frac{\alpha_\chi \mu_\chi}{2} \tag{22}$$

Here, $\alpha_\chi$ denotes a dimensionless constant characterizing the strength of the anisotropy, while $\mu_\chi$ defines the relationship between system anisotropies and their local properties. We explore two different models for $\mu_\chi(r)$:

1. In the first model, commonly adopted in many articles, we assume that $\Pi_\chi$ is directly proportional to $P_{\chi r}$, resulting in the equation:

$$\mu_\chi(r) = P_{\chi r} \tag{23}$$

2. In a more recent description found in the literature, the anisotropy is also proportional to the total enclosed mass $m(r)$ inside a sphere of radius $r$. Thus, we have:

$$\mu_\chi(r) = \frac{2m(r)}{r} P_{\chi r}. \tag{24}$$

Here, $m(r)$ represents the combined mass, including both dark matter and baryonic matter, within a sphere of radius $r$. The choice of equation (24) for the anisotropy factor has several appealing features. First, the compactness $\mu_\chi$ scales approximately as $r^2$, while $m(r)$ scales roughly as $r^3$ as $r$ approaches zero. This leads to a vanishing anisotropy at the centre, guaranteeing the regularity of the right-hand side of equation (10) when $i = \chi$. Second, the anisotropy factor becomes significant only for highly relativistic configurations where $\mu_\chi \sim \mathcal{O}(1)$. This is consistent with the consensus that fluid anisotropy increases at higher matter densities.

The realisation of both models may be seen in Fig. 1, where we consider the size of the anisotropy relative to the total radial pressure $P_r$. For equation (23) we see that the relative anisotropy, $\frac{2\Pi_\chi}{P_r}$, increases smoothly until it reaches the edge of the neutron star, after which the anisotropy is equivalent to the total radial pressure, as there is no baryonic pressure outside of the neutron star.
In the regime governed by equation (24), the relative

anisotropy increases at a lower rate, and once outside of the neutron star it instead now slowly declines until it reaches the edge of the dark matter halo.
One can see that despite both objects having identical initial conditions the latter model produces halos of smaller extent.

## IV. MODELLING DARK MATTER ADMIXED NEUTRON STARS

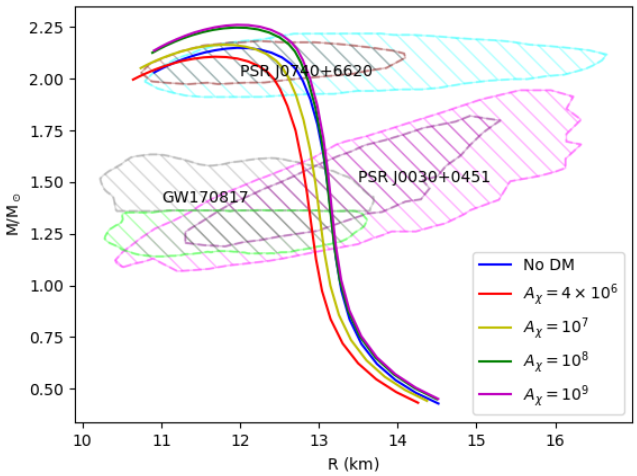



Figure 2. Results of hybrid neutron star simulations for all choices of $A_\chi$ before the inclusion of anisotropy, $\alpha_\chi = 0$. Included also is the contour data pertaining to neutron star mass and radii as measured by NICER and LIGO. For PSR J0740+6620, a 68% credibility region is highlighted in maroon, and 95% credibility region in turquoise,[82] and [83].For PSR J0030+0451, the purple region indicates 68% credibility and the fuchsia region 95%, [64]. For GW170817, the upper grey region and lower green region indicate the major and minor component respectively of the event, [84].

### A. Hybrid Neutron Star Simulations

In the computation of these hybrid stars, which consist of a neutron star and a dark matter halo, we have made several numerical adjustments to ensure the accuracy and reliability of our results. Specifically, we highlight the following considerations:

1. The equation of state for a neutron star exhibits a hard edge, meaning that numerical tables do not smoothly approach zero pressure or density. Therefore, determining the edge of the object accurately from a numerical perspective is crucial, as selecting a region where pressure or density is zero lies outside the scope of an equation of state interpolated from pressure-density data.
   As such, the surface of the neutron star is determined as the radius at which the increase in mass due to baryonic components produced by increasing the radius by one metre is less than a certain threshold.

2. We must similarly be careful in considering the properties and size of the dark matter halo. The equation of state is smooth all the way to zero pressure and density, and as such we consider the dark matter halo ending at a radius $R_\chi$ where $P_{r\chi}(R_\chi) = 0$.
   Depending on the value of $A_\chi$, the dark matter halo has the capacity to become extremely large, meaning that we may be concerned with its interaction with other astrophysical objects. The maximum extent of a Bose-Einstein Condensate (BEC) halo of this kind will essentially be its Lane-Emden radius, as governed by the Lane-Emden equations for a polytropic fluid which may be seen in [71] in a similar context.
   From this we see that the Lane-Emden radius is described by

$$R_{LE} = \sqrt{\frac{\pi A_\chi}{2}}, \tag{25}$$

   which for $A_\chi = 10^{10}$ would produce a radius of around 4800km, and therefore the domains probed in this article should be free from any of the aforementioned concerns.

For each parameter used, we simulate a set of 30 hybrid neutron stars with masses ranging from $0.4M_\odot$ to $2.4M_\odot$. These simulations are specified by defining separate central pressures and densities for both the dark and baryonic matter. Subsequently, numerical integration is performed outward to a specified radius, chosen to ensure that the full extent of the dark matter halo is completely encompassed.
All models produced have a dark matter fraction of 5% of the neutron star mass, and this parameter will be kept fixed across all simulations so that relevant differences produced by other properties may be effectively identified. We choose this fiducial value as it is in line with quantities seen in the wider literature, e.g. [12], as well as that the effects and differences between models are felt to be sufficiently visible for this choice.

### B. Properties of the Dark Matter Halo

In this study, we explore the behaviour and properties of dark matter halos described by a variety of choices of $A_\chi$, and effected by several levels of anisotropy as dictated by $\alpha_\chi$. We will therefore be able to see the effects of different choices of dark matter equation of state properties, the effects of introducing anisotropy, as well as how

| $A_\chi$ $\frac{cm^5}{gs^2}$ | $M_{max}$ $M_\odot$ | $R_B(M_{max})$ km | $R_B^{min}$ km | $R_B^{max}$ km | $R_\chi^{min}$ km | $R_\chi^{max}$ km |
|---|---|---|---|---|---|---|
| 0 | 2.15 | 12.0 | 10.9 | 14.5 | | |
| 4 $10^6$ | 2.11 | 11.7 | 10.6 | 14.3 | 12.5 | 17.8 |
| $10^7$ | 2.17 | 11.8 | 10.7 | 14.4 | 20.2 | 26.9 |
| $10^8$ | 2.25 | 11.9 | 10.9 | 14.5 | 77.4 | 84.2 |
| $10^9$ | 2.26 | 12.0 | 10.9 | 14.5 | 262.1 | 268.2 |

Table I. Mass-Radius data for hybrid compact objects containing isotropic BEC dark matter for various values of $A_\chi$, for which $\alpha_\chi = 0$. Within we observe the maximum mass seen in figure 2, the associated baryonic radius, as well as the range of baryonic and dark radii observed'

significant the manifestation of an anisotropy is in each regime.

### 1. *The Equation of State Parameter $A_\chi$*

We now begin by discussing the effect of varying $A_\chi$, the constant controlling our polytropic equation of state, equation (20). As one may see in equation (21), these choices will also be indicative of the range of masses that these models will represent, depending on the choice of scattering length that is made.

The nuances of these $A_\chi$ values, before factoring in anisotropy, are illustrated in Figure 2 and enumerated in Table I. The table delineates mass-radius data for hybrid compact objects embedded with isotropic Bose-Einstein Condensate (BEC) dark matter. The data span a range of $A_\chi$ from 4 $10^6$ to $10^9$, with all observations adhering to the $\alpha_\chi = 0$ stipulation. For completeness we include the smallest and largest radii simulated, denoted $R^{min}$ and $R^{max}$ respectively. It is worth noting the smallest dark radius belongs to the smallest baryonic radius neutron star, and the largest dark radius to the largest baryonic radius object.

When juxtaposing the data from Table I with Figure 2, two distinct behaviours in the mass-radius relations of these hybrid objects emerge based on the type of dark matter: (i) objects with denser dark matter (indicated by smaller $A_\chi$) appear more compact at lower masses and resemble characteristics of non-dark matter objects at higher masses; (ii) on the other hand, with more diffuse dark matter (larger $A_\chi$), neutron stars at lower masses are akin to those without dark matter, but display notably larger masses at reduced radii as their masses increase.

Delving deeper into Table I, we note that when $A_\chi$ = 4 $10^6$, it records the smallest radius across all parameters and also the least maximum mass. As the value of $A_\chi$ escalates, there is a noticeable surge in the dark matter halo radius, with the most larger dark radii observed at $A_\chi = 10^9$. Meanwhile, the baryonic radius seems to stabilize for values of $A_\chi > 10^8$, showcasing a range of radii akin to a conventional neutron star. The masses of all the examined neutron stars peak between 11.5 km and 12 km. Both $A_\chi = 10^8$ and $10^9$ possess strikingly similar maximum masses, notably around 5% more than $A_\chi = 0$.

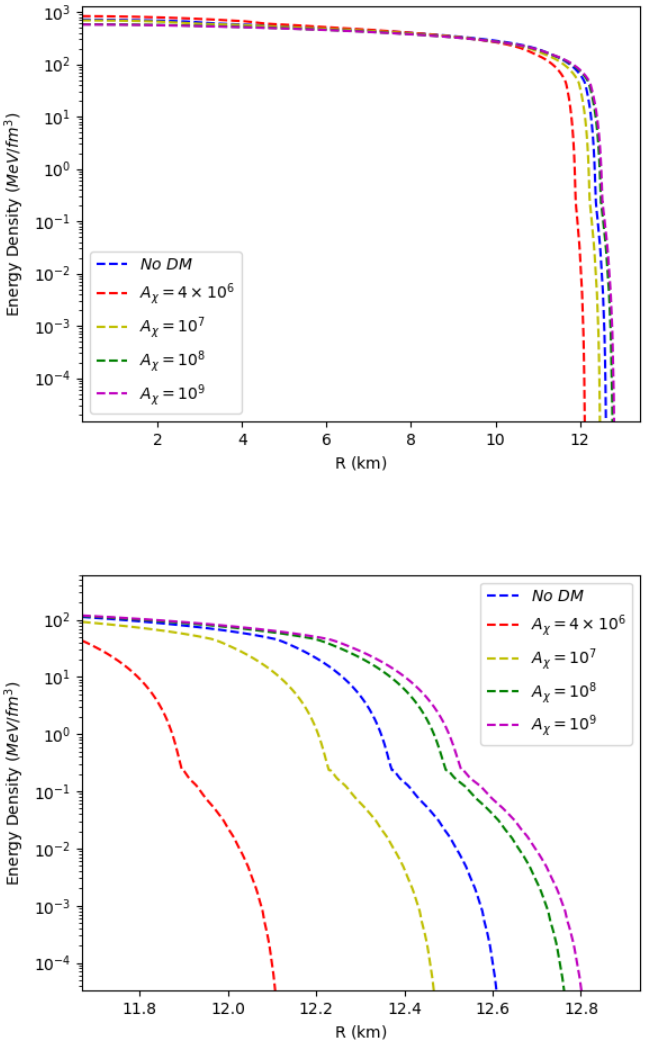



Figure 3. The figure shows the Energy Density-Radius curves for the baryonic matter present in neutron stars with a mass of 2.083$M_\odot$ for different values of $A_\chi$. For diminished $A_\chi$ values, the baryonic radius is notably smaller than in cases devoid of dark matter, revealing pronounced differences in radii. In contrast, for augmented $A_\chi$ values, the baryonic radii increase and converge. The lower panel provides a detailed view of the Density-Radius profiles near the neutron star's radius.

In Figure 3, we further explore this analysis, where we showcase the density-baryonic radius relationship of five compact objects of similar mass but distinct compositions. For $A_\chi$ = 4 $10^6$ and $10^7$, the object is more compact than a standard neutron star, evidenced by a denser

inner core and reduced radius. Objects with $A_\chi = 10^8$ and $10^9$ have radii that align more with conventional neutron stars, leading to the speculation that heavier dark matter might cause the baryonic matter to be more compact compared to a neutron star. Meanwhile, larger-radius, lighter dark matter seems to expand the baryonic radius beyond its typical dimensions.

However Figure 3 does not tell the full story. Each hybrid object has a baryonic mass approximately equal to 1.98 $M_\odot$, as opposed to its total combined mass of 2.083 $M_\odot$, a value that for a standard neutron star generally equates to a radius of 12.81 km. This suggests that in all objects mixed with dark matter, the baryonic matter is more densely packed. For greater values of $A_\chi$, the baryonic radius is in close proximity to a non-hybrid object of comparable baryonic mass. It suggests that lighter dark matter has a more limited impact on the baryonic structure of the neutron star and mainly contributes to the overall dark mass, and in essence a 1.98 $M_\odot$ neutron star appears to have a dark matter cloud superimposed on top of it.

#### 2. *Variations in Anisotropy Models*

We now proceed to examine the influence of introducing anisotropy into our dark matter. Furthermore, we contrast the disparities between the anisotropy models described by equations (23) and (24) and observe their behaviours for each dark matter variant under consideration. In Figure 4, the mass-radius relations corresponding to all $A_\chi$ values are presented: both in the absence of anisotropy and when anisotropies are governed by equations (23) and (24), with $\alpha_\chi = 1$. It is evident that the mass-radius curves generated by equation (24) (represented by dashed lines) possess higher masses for a specific radius than those from equation (23). This distinction becomes most pronounced when considering lower values of $A_\chi$.

The contrast between models with and without anisotropy diminishes with increasing $A_\chi$ values, to the extent that all three curves become virtually indistinguishable at $A_\chi = 10^9$. Recalling prior discussions, this trend can likely be attributed to the need for a decreased central pressure to form a dark matter halo constituting 5% of the aggregate mass for elevated $A_\chi$ values. Consequently, the radial contribution to pressure within the neutron star from dark matter is substantially overshadowed by that of the baryonic matter. This renders the TOV equations largely unresponsive to the anisotropies inherent in the dark matter. For $A_\chi = 10^9$, the curve closely parallels the one devoid of dark matter, with each mass incrementally increased by roughly 5%. This implies that, in this scenario, the dark matter primarily serves to augment the object's mass without significantly influencing the baryonic radius. As a result, the anisotropic nature of the matter exerts negligible influence the mass-radius relation.

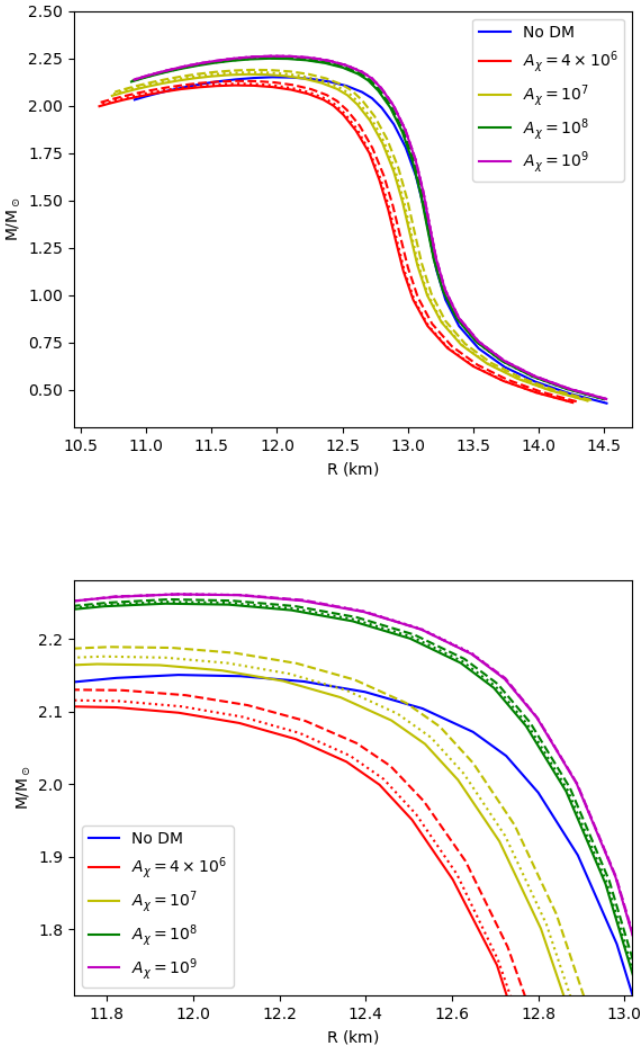



Figure 4. Plots containing mass-radius relations for all choices of $A_\chi$ and $\mu_\chi$, where we also have $\alpha_\chi = 1$. Solid lines indicate curves without anisotropy, dotted lines indicate the presence of anisotropy governed by Eq. 23, and dashed lines an anisotropy governed by Eq. 24.

### C. Impact of Anisotropy on Halo Structure

We delve into the influence of anisotropies on the structural characteristics of the dark matter halo and investigate the resultant shifts in mass-radius curves as $\alpha_\chi$ transitions from 0 to 1. Figure 5 illustrates the repercussions of each anisotropy model on the halo's structure for $A_\chi = 10^7$ and $A_\chi = 10^9$. The impact of the former can be discerned in the upper panel of figure 5. The halo

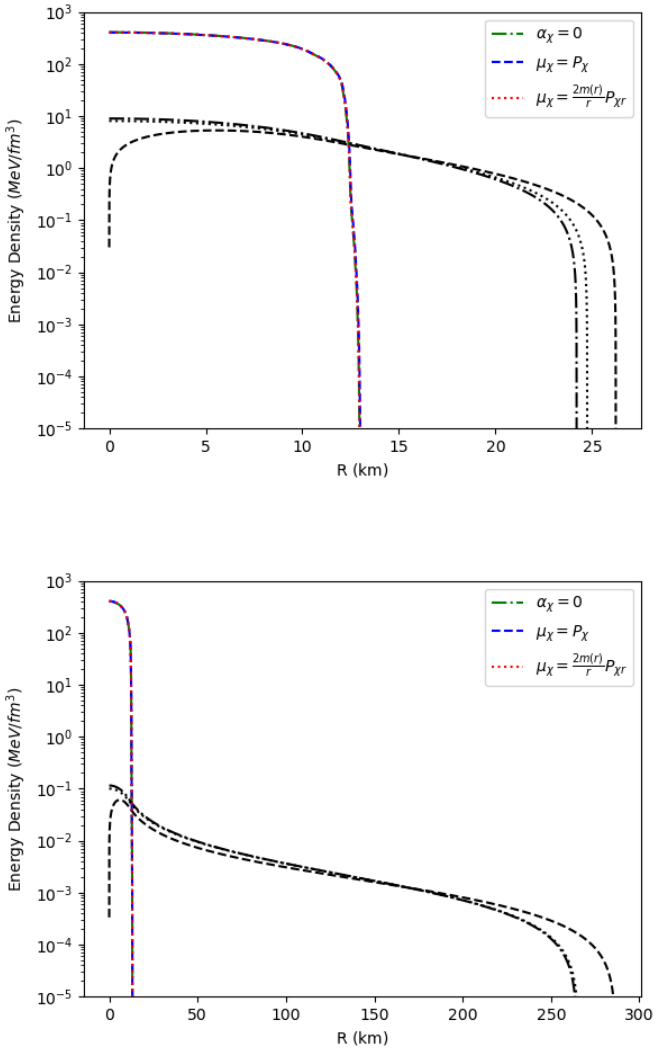



Figure 5. The Energy Density-Radius curves for dark matter admixed neutron stars with and without anisotropy. For $A_\chi = 10^7$ (upper) we consider $1.50M_\odot$ objects, and for $A_\chi = 10^9$ (lower) $1.54M_\odot$ objects. The anisotropies are governed by equations (23) (dashed) and (24) (dotted) where $\alpha_\chi = 1$, as well as including $\alpha_\chi = 0$ (dot-dash) as a reference. Baryonic matter is indicated by a coloured line, and the black line of the respective line style indicates the associated dark matter curve.

with the smallest size arises in the absence of anisotropy, while the dynamics governed by the equation (23) mould the largest halo. Notably, the behaviour steered by the anisotropies of equation (23) deviates notably from the patterns seen in equation (24) and the non-anisotropic case near $r = 0$. Even though anisotropies are often modelled in this manner (as found in the literature), the resultant attributes at the core of the star might be deemed unrealistic, differentiating it from the alternative model (equation 24) presented in this analysis.

All dark matter halos have uniform masses; therefore, halos governed by anisotropies directly proportional to pressure inevitably exhibit larger sizes to offset their less dense cores. For anisotropies governed by equation (24), the changes are more subtle. Within the neutron star, the dark matter distribution exhibits a slightly reduced density compared to its isotropic counterpart but spans an additional few hundred meters in radius. This pattern persists in the lower panel of Figure 5, with the disparities between the two models amplifying. Remarkably, while equation (24) engenders a halo only marginally larger than the isotropic case, equation (23) continues to feature its hallmark low-density core. Moreover, the halo influenced by equation (23) now stretches several tens of kilometers further than the other two scenarios.

## V. CONSTRAINTS ON THE PROPERTIES OF DARK MATTER PARTICLES

In light of our investigation into these bosonic dark matter models, we will now discuss potential constraints on the dark matter properties in the context of astrophysical observations. While this study primarily focuses on a hybrid dark matter-neutron star model considering only gravitational interactions, we also explore dark matter models wherein the self-interaction of dark matter and its interaction with baryons are so negligible that they can be approximated as non-interactive at a first approximation.

As our dark matter EOS is fully specified by a choice of $A_\chi$, we have been free to simulate hybrid neutron stars without the need to concern ourselves with more specific properties of the dark matter particles. With a known value of $A_\chi$, equation 21 can be used to fully specify the particle mass if its scattering length is known, and vice versa. As such we will now consider possible constraints that may be set on our system based on both astrophysical data, and terrestrial experiments on BEC gases.

### A. Galactic Cluster Constraints on Sub-GeV Bosonic Dark Matter Particle Masses

In this study, we refine our bosonic dark matter model using contemporary astronomical observations. Notably, collision data from galaxy clusters, such as the Bullet Cluster [1E 0657-56, 85] and the Baby Bullet [MACSJ0025-12, 86], provide essential insights into the properties of dark matter. Further observations of galaxy cluster collisions, notably MACS J0025.4-1222 [86], reveal mismatches between the centers of visible matter and gravitational mass. Clowe et al. [87] determined that the observed absence of dark matter deceleration offers a constraint on its self-interaction strength. Specifically,

| $A_\chi$ $\frac{\text{cm}^5}{\text{gs}^2}$ | $l_\chi$ fm | $m_\chi$ GeV |
|---|---|---|
| 4 $10^6$ | 2.79 | 0.442 |
| $10^7$ | 2.33 | 0.306 |
| $10^8$ | 1.47 | 0.122 |
| $10^9$ | 0.93 | 0.0485 |

Table II. Properties of dark matter corresponding to $\sigma_m^{bc} = 1.25\ \text{cm}^2/\text{g}$, tabulated against various values of the polytropic constant $A_\chi$.

this can be expressed in terms of the self-interaction cross section for long-range forces [e.g., 4]. This observational data provides constraints on boson dark matter particles, represented by the scattering cross section $\sigma_\chi$. This cross section is directly related to the scattering length according to $\sigma_\chi = 4\pi l_\chi^2$.

Under the ΛCDM cosmological model, based on galaxy merger observations, $\sigma_m$ should not exceed a critical value $\sigma_m^{bc}$: $\sigma_m \leq \sigma_m^{bc}$. Studies, such as Crăciun and Harko [71] and Thompson et al. [88], cite the Bullet Cluster and Baby Bullet as evidence for this upper limit on the ratio $\sigma_m$. Randall et al. [89] found that $\sigma_m^{bc} = 1.25\ \text{cm}^2/\text{g}$ denoting the maximal value consistent with these findings, we can deduce constraints on the dark matter particle mass and scattering length for each of our dark matter proposed model, ensuring their coherence with this astronomical observation.

We initiate our analysis with the ratio:

$$\sigma_m = \frac{\sigma_\chi}{m_\chi} = \frac{4\pi l_\chi^2}{m_\chi}. \tag{26}$$

We can derive constraints on $m_\chi$ and $l_\chi$. Utilizing $\sigma_m^{bc} = 1.25\ \text{cm}^2/\text{g}$:

$$\left(\frac{m_\chi}{1\ \text{GeV}}\right) = 0.0565 \left(\frac{l_\chi}{1\ \text{fm}}\right)^2. \tag{27}$$

Here, we aim to find parameters to our the validity of our model. Considering that observational constraints predominantly hinge on the scattering properties of galactic dark matter, and in alignment with this study's objectives, we rewrite equation (21) in terms of the mass of the dark matter particle:

$$A_\chi = \frac{\sqrt{\pi \sigma_c^{bc}}}{m_\chi^{5/2}}. \tag{28}$$

By setting this and considering $\sigma_m = \sigma_m^{bc}$, we can deduce the parameters associated with the dark matter's mass and scattering length. Table II presents the properties of dark matter, with each row detailing the associated values of the polytropic constant $A_\chi$, scattering length $l_\chi$, and mass $m_\chi$ for a given dark matter configuration corresponding to $\sigma_m^{bc} = 1.25\ \text{cm}^2/\text{g}$. Using the algorithm based on $A_\chi$, we determine $m_\chi$ and $l_\chi$ with equations (28) and (27). The results indicate that, within the scope of $A_\chi$ addressed in this study, $m_\chi$ ranges from 0.05 GeV to 0.5 GeV, while $l_\chi$ lies between 0.9 fm and 3 fm. These constraints are particularly intriguing because this mass range of bosonic dark matter has not been extensively discussed in the literature. In fact, much of the literature [e.g., 71] primarily emphasizes lightweight BEC dark matter species with significantly smaller masses, associated with dark matter halos spanning kilo-parsec radii.

### B. Experimental Constraints on Sub-GeV Bosonic Dark Matter Particle Masses

| $A_\chi$ $\frac{\text{cm}^5}{\text{gs}^2}$ | $m_\chi$ GeV |
|---|---|
| 4 $10^6$ | 31.5 |
| $10^7$ | 23.2 |
| $10^8$ | 10.8 |
| $10^9$ | 5.0 |

Table III. Properties of dark matter corresponding to $l_\chi^{ex} = 10^6$ fm, tabulated against various values of the polytropic constant $A_\chi$.

A Bose-Einstein Condensate (BEC) typically forms when the temperature of a system drops beneath the critical temperature, represented as:

$$T_c = \frac{2\pi}{m_\chi}\left(\frac{n_\chi}{\zeta(3/2)}\right)^{2/3}, \tag{29}$$

where $n_\chi = \frac{\rho_\chi}{m_\chi}$ and $\zeta(3/2)$ is the Riemann zeta function. This critical temperature, $T_c$, is characterized by the thermal de Broglie wavelength equating to the mean interparticle distance, facilitating the synchronization and overlap of individual particle wave functions [e.g., 90]. A coherent state emerges when particle density is sufficiently high or the temperature is adequately low [e.g., 91]. Given the properties of BEC dark matter, its interactions with the observable universe are predominantly gravitational. Hence, we hypothesized that the temperature of the condensate will experience negligible alterations upon accretion onto a neutron star. Consequently, we assume that the dark matter temperature stays below the critical temperature $T_c$.

Bose-Einstein Condensation (BEC) is a well-established phenomenon in terrestrial experiments [e.g., 91]. Therefore, the idea that similar condensation could occur elsewhere in the universe is not implausible [e.g., 71, 90, 92]. Specifically, if the temperature of a boson gas falls below its critical temperature, BEC can manifest at some point in the universe's cosmic history. Indeed, terrestrial experiments with cold gases have already verified

the presence of this unique state of matter for bosonic particles in ultra-cold, low-density conditions [93, 94].

Now, we turn our focus to the scattering lengths observed in terrestrial Bose-Einstein Condensate (BEC) gas experiments, as documented in [94] and [95]. For instance, experimentally measured scattering lengths are typically 2.75 nm for $^{23}Na$ [96] and 5.77 nm for $^{87}Rb$ [97]. The use of constraints set by considering ordinary matter provides a counterpoint to those set by astronomical observations. As the precise nature of dark matter is not yet understood, it is valuable for us to consider the implications of bosonic dark matter with properties more similar to condensates encountered terrestrially.
If there exists a zoo of dark matter particles, much like with the standard model, then it can be justifiable to explore the possibility that the accumulated dark matter is distinct from that make up galactic halos.
For our calculations, we adopt a benchmark scattering length, $l_\chi^{ex} = 10^6$ fm, which aligns well with contemporary laboratory findings. Aligned with the objectives of this study, we recast equation (21) to be articulated in terms of the dark matter particle's mass:

$$A_\chi = \frac{2\pi l_\chi^{ex}}{m_\chi^3} = 1.25\ 10^{11} \left(\frac{1\ \text{GeV}}{m_\chi}\right)^3, \qquad (30)$$

taking into consideration our predefined value for $l_\chi^{ex}$. Using this formula, we can determine the dark matter's mass parameters. Table III lists these values, showing the corresponding polytropic constant $A_\chi$ and mass $m_\chi$.

## VI. SUMMARY AND CONCLUSION

In this study, we examined the intricacies surrounding the properties of dark matter particles using experimental data and astrophysical observations, with a special emphasis on the hybrid dark matter-neutron star model.

We initiated our analysis by modelling a hybrid neutron star embedded within an anisotropic bosonic dark matter halo without pre-defined constraints on the dark matter particles' mass range or scattering length. Varying the equation of state parameter $A_\chi$, we derived distinct mass-radius correlations for these hybrid constructs. Notably, for our fiducial dark matter model, the parameter $A_\chi$ needs to lie between 4 $10^6$ and $10^9$ g$^{-1}$ cm$^5$ s$^{-2}$ to profoundly influence the empirically observed mass-radius relationship of the neutron star. It became evident that the dark matter's nature, whether condensed or dispersed, has a significant bearing on the characteristics of neutron stars correlated with their masses. Introducing anisotropies further diversified the mass-radius trajectories. Among the models assessed, equation (24) emerged as an authentic representation of anisotropy. This model is distinct in its generation of halos with a unique low-density core profile, distinguishing it from alternative model forecasts. These anisotropic nuances influence the dark matter halo's intrinsic properties, especially in size, density gradient, and the neutron star's baryonic component.

Galactic clusters, specifically the Bullet Cluster, have been instrumental in refining our bosonic dark matter model. From this analysis, we have inferred a relationship for bosonic dark matter particles, characterized by the scattering cross section $\sigma_\chi$, with the scattering length described as $\sigma_\chi = 4\pi l_\chi^2$.This insight sharpens our understanding of the parameter $m_\chi$, which we have now restricted between 0.05 GeV to 0.5 GeV and $l_\chi$ in the range of 0.9 fm to 3 fm. Grounded in terrestrial Bose-Einstein Condensate (BEC) experiments and using a benchmark scattering length $l_\chi \approx 1$ nm, we delineated the constraints on dark matter properties. With a fiducial scattering length set at $l_\chi = 10^6$ fm, our model suggests dark matter particle masses to fall between 3 GeV and 30 GeV.

Notably, the dark matter particle mass range identified in the preceding sub-sections, specifically between $10^2$ to $10^{-3}$ GeV, presents notable challenges for direct dark matter detection experiments [98]. Specifically, in the mass range from $10^{-3}$ to 0.5 GeV, current detectors face technical limitations exacerbated by an anticipated substantial neutrino background [99].

In conclusion, our study illuminates the intricate relationship between dark and baryonic matter in hybrid neutron stars. The variability in outcomes, influenced by different equations of state and anisotropic models, underscores the imperative for more detailed investigations. Given the insights we have gleaned from galactic clusters and terrestrial BEC experiments, this work is a relevant reference point, guiding subsequent research into the enigmatic realm of dark matter and its astronomical manifestations.

## ACKNOWLEDGMENTS

I.L. and Z.B.S. gratefully acknowledge the financial support provided to the Center for Astrophysics and Gravitation (CENTRA/IST/ULisboa) through the Grant Project No. UIDB/00099/2020 and Grant No. PTDC/FIS-AST/28920/2017 by the Fundação para a Ciência e Tecnologia (FCT), Portugal.
We also gratefully thank the anonymous referee for their positive insights and suggestions, which have surely improved the quality of this work.

[1] M. Battaglieri, A. Belloni, A. Chou, P. Cushman, B. Echenard, R. Essig, J. Estrada, J. L. Feng, B. Flaugher, P. J. Fox, et al., arXiv e-prints arXiv:1707.04591 (2017), 1707.04591.

[2] J. Billard, M. Boulay, S. Cebrián, L. Covi, G. Fiorillo, A. Green, J. Kopp, B. Majorovits, K. Palladino, F. Petricca, et al., Reports on Progress in Physics **85**, 056201 (2022), 2104.07634.

[3] A. Del Popolo, International Journal of Modern Physics D **23**, 1430005 (2014), 1305.0456.

[4] B.-L. Young, Frontiers of Physics **12**, 121201 (2017).

[5] J. L. Feng, ARA&A **48**, 495 (2010), 1003.0904.

[6] H. Baer, K.-Y. Choi, J. E. Kim, and L. Roszkowski, Phys. Rep. **555**, 1 (2015), 1407.0017.

[7] L. Roszkowski, E. M. Sessolo, and S. Trojanowski, Reports on Progress in Physics **81**, 066201 (2018), 1707.06277.

[8] S. Profumo, L. Giani, and O. F. Piattella, Universe **5**, 213 (2019), 1910.05610.

[9] C. Kouvaris, Phys. Rev. D **77**, 023006 (2008), 0708.2362.

[10] C. Kouvaris and P. Tinyakov, Phys. Rev. D **82**, 063531 (2010), 1004.0586.

[11] G. Panotopoulos and I. Lopes, Phys. Rev. D **96**, 083004 (2017), 1709.06312.

[12] O. Ivanytskyi, V. Sagun, and I. Lopes, Phys. Rev. D **102**, 063028 (2020), 1910.09925.

[13] Q.-F. Xiang, W.-Z. Jiang, D.-R. Zhang, and R.-Y. Yang, Phys. Rev. C **89**, 025803 (2014), URL https://link.aps.org/doi/10.1103/PhysRevC.89.025803.

[14] A. Das, T. Malik, and A. C. Nayak, Phys. Rev. D **99**, 043016 (2019), URL https://link.aps.org/doi/10.1103/PhysRevD.99.043016.

[15] A. Quddus, G. Panotopoulos, B. Kumar, S. Ahmad, and S. K. Patra, Journal of Physics G Nuclear Physics **47**, 095202 (2020), 1902.00929.

[16] H. C. Das, A. Kumar, and S. K. Patra, Phys. Rev. D **104**, 063028 (2021), URL https://link.aps.org/doi/10.1103/PhysRevD.104.063028.

[17] H. C. Das, A. Kumar, S. K. Biswal, and S. K. Patra, Phys. Rev. D **104**, 123006 (2021), URL https://link.aps.org/doi/10.1103/PhysRevD.104.123006.

[18] O. Lourenço, T. Frederico, and M. Dutra, Phys. Rev. D **105**, 023008 (2022), URL https://link.aps.org/doi/10.1103/PhysRevD.105.023008.

[19] Y. Dengler, J. Schaffner-Bielich, and L. Tolos, Phys. Rev. D **105**, 043013 (2022), URL https://link.aps.org/doi/10.1103/PhysRevD.105.043013.

[20] O. Lourenço, C. H. Lenzi, T. Frederico, and M. Dutra, Phys. Rev. D **106**, 043010 (2022), URL https://link.aps.org/doi/10.1103/PhysRevD.106.043010.

[21] A. Das, T. Malik, and A. C. Nayak, Phys. Rev. D **105**, 123034 (2022).

[22] G. Panotopoulos and I. Lopes, International Journal of Modern Physics D **27**, 1850093 (2018), 1804.05023.

[23] D. Rafiei Karkevandi, S. Shakeri, V. Sagun, and O. Ivanytskyi, Phys. Rev. D **105**, 023001 (2022), URL https://link.aps.org/doi/10.1103/PhysRevD.105.023001.

[24] J. Ellis, G. Hütsi, K. Kannike, L. Marzola, M. Raidal, and V. Vaskonen, Phys. Rev. D **97**, 123007 (2018), URL https://link.aps.org/doi/10.1103/PhysRevD.97.123007.

[25] J. M. Lattimer, Annual Review of Nuclear and Particle Science **71**, 433 (2021).

[26] A. Del Popolo, M. Le Delliou, and M. Deliyergiyev, Universe **6**, 222 (2020).

[27] L. Herrera and N. O. Santos, Phys. Rep. **286**, 53 (1997).

[28] J. Kumar and P. Bharti, NAR **95**, 101662 (2022).

[29] T. Grammenos, F. Rahamn, S. Ray, D. Deb, and S. R. Chowdhury, arXiv e-prints arXiv:1611.02253 (2016), 1611.02253.

[30] A. Rahmansyah and A. Sulaksono, in *Journal of Physics Conference Series* (2021), vol. 1816 of *Journal of Physics Conference Series*, p. 012025.

[31] P. H. R. S. Moraes, G. Panotopoulos, and I. Lopes, Phys. Rev. D **103**, 084023 (2021), 2101.02207.

[32] R. W. Michie, MNRAS **125**, 127 (1963).

[33] J. Binney and S. Tremaine, *Galactic dynamics* (1987).

[34] P. Cuddeford, MNRAS **253**, 414 (1991).

[35] L. Herrera, J. Martin, and J. Ospino, Journal of Mathematical Physics **43**, 4889 (2002), gr-qc/0207040.

[36] L. Herrera, A. di Prisco, J. Martin, J. Ospino, N. O. Santos, and O. Troconis, Phys. Rev. D **69**, 084026 (2004), gr-qc/0403006.

[37] S. S. Bayin, Phys. Rev. D **26**, 1262 (1982).

[38] L. Herrera, J. Ospino, and A. di Prisco, Phys. Rev. D **77**, 027502 (2008), 0712.0713.

[39] M. K. Mak and T. Harko, Proceedings of the Royal Society of London Series A **459**, 393 (2003), gr-qc/0110103.

[40] I. Lopes, G. Panotopoulos, and Á. Rincón, European Physical Journal Plus **134**, 454 (2019), 1907.03549.

[41] R. Sharma and S. D. Maharaj, MNRAS **375**, 1265 (2007), gr-qc/0702046.

[42] J. R. Oppenheimer and G. M. Volkoff, Physical Review **55**, 374 (1939).

[43] R. C. Tolman, Physical Review **55**, 364 (1939).

[44] G. Panotopoulos and I. Lopes, Phys. Rev. D **96**, 023002 (2017), 1706.07272.

[45] P. Ciarcelluti and F. Sandin, Physics Letters B **695**, 19 (2011), 1005.0857.

[46] F. Sandin and P. Ciarcelluti, Astroparticle Physics **32**, 278 (2009), 0809.2942.

[47] H. Heintzmann and W. Hillebrandt, A&A **38**, 51 (1975).

[48] V. Sagun, I. Lopes, and A. Ivanytskyi, Nucl. Phys. A **982**, 883 (2019), 1810.04245.

[49] M. Al-Mamun, A. W. Steiner, J. Nättilä, J. Lange, R. O'Shaughnessy, I. Tews, S. Gandolfi, C. Heinke, and S. Han, Phys. Rev. Lett. **126**, 061101 (2021), 2008.12817.

[50] G. Raaijmakers, S. K. Greif, K. Hebeler, T. Hinderer, S. Nissanke, A. Schwenk, T. E. Riley, A. L. Watts, J. M. Lattimer, and W. C. G. Ho, ApJ **918**, L29 (2021), 2105.06981.

[51] S. Pratt, E. Sangaline, P. Sorensen, and H. Wang, Phys. Rev. Lett. **114**, 202301 (2015), 1501.04042.

[52] A. Bazavov, H. T. Ding, P. Hegde, O. Kaczmarek, F. Karsch, E. Laermann, Y. Maezawa, S. Mukherjee, H. Ohno, P. Petreczky, et al., Phys. Rev. D **95**, 054504 (2017), 1701.04325.

[53] N. Haque, A. Bandyopadhyay, J. O. Andersen, M. G. Mustafa, M. Strickland, and N. Su, Journal of High Energy Physics **2014**, 27 (2014), 1402.6907.

[54] P. Papazoglou, D. Zschiesche, S. Schramm, J. Schaffner-

Bielich, H. Stoecker, and W. Greiner, Phys. Rev. C **59**, 411 (1999), nucl-th/9806087.
[55] V. Dexheimer and S. Schramm, Astrophys. J. **683**, 943 (2008), 0802.1999.
[56] A. Motornenko, J. Steinheimer, V. Vovchenko, S. Schramm, and H. Stoecker, Phys. Rev. C **101**, 034904 (2020), URL https://link.aps.org/doi/10.1103/PhysRevC.101.034904.
[57] A. Motornenko, S. Pal, A. Bhattacharyya, J. Steinheimer, and H. Stoecker, Phys. Rev. C **103**, 054908 (2021), 2009.10848.
[58] P. D. Group, R. L. Workman, V. D. Burkert, V. Crede, E. Klempt, U. Thoma, L. Tiator, K. Agashe, G. Aielli, B. C. Allanach, et al., Progress of Theoretical and Experimental Physics **2022**, 083C01 (2022), ISSN 2050-3911, https://academic.oup.com/ptep/article-pdf/2022/8/083C01/49175539/ptac097.pdf, URL https://doi.org/10.1093/ptep/ptac097.
[59] P. A. Zyla et al. (Particle Data Group), PTEP **2020**, 083C01 (2020).
[60] A. S. Schneider, L. F. Roberts, and C. D. Ott, Phys. Rev. C **96**, 065802 (2017), 1707.01527.
[61] A. Motornenko, V. Vovchenko, J. Steinheimer, S. Schramm, and H. Stoecker, Nucl. Phys. A **982**, 891 (2019), 1809.02000.
[62] E. R. Most, A. Motornenko, J. Steinheimer, V. Dexheimer, M. Hanauske, L. Rezzolla, and H. Stoecker, Phys. Rev. D **107**, 043034 (2023), 2201.13150.
[63] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, et al. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. X **9**, 011001 (2019), URL https://link.aps.org/doi/10.1103/PhysRevX.9.011001.
[64] M. C. Miller, F. K. Lamb, A. J. Dittmann, S. Bogdanov, Z. Arzoumanian, K. C. Gendreau, S. Guillot, A. K. Harding, W. C. G. Ho, J. M. Lattimer, et al., ApJ **887**, L24 (2019), 1912.05705.
[65] G. Baym, C. Pethick, and P. Sutherland, ApJ **170**, 299 (1971).
[66] V. V. Sagun, I. Lopes, and A. I. Ivanytskyi, ApJ **871**, 157 (2019), 1805.04976.
[67] B. D. Serot and J. D. Walecka, International Journal of Modern Physics E **6**, 515 (1997), nucl-th/9701058.
[68] C. G. Böhmer and T. Harko, JCAP **2007**, 025 (2007), 0705.4158.
[69] X. Y. Li, T. Harko, and K. S. Cheng, JCAP **2012**, 001 (2012), 1205.2932.
[70] T. Harko and F. S. N. Lobo, Phys. Rev. D **92**, 043011 (2015), 1505.00944.
[71] M. Crăciun and T. Harko, European Physical Journal C **80**, 735 (2020), 2007.12222.
[72] F. E. Schunck and E. W. Mielke, Classical and Quantum Gravity **20**, R301 (2003), 0801.0307.
[73] M. Gleiser and K. Dev, International Journal of Modern Physics D **13**, 1389 (2004), astro-ph/0401546.
[74] W. Barreto, Ap&SS **201**, 191 (1993).
[75] P. S. Letelier, Phys. Rev. D **22**, 807 (1980).
[76] N. F. Naidu, S. Carloni, and P. Dunsby, Phys. Rev. D **106**, 124023 (2022), 2210.06867.
[77] P. León, E. Fuenmayor, and E. Contreras, Phys. Rev. D **107**, 064010 (2023), 2303.04211.
[78] B. S. Ratanpal and R. Patel, Ap&SS **368**, 21 (2023).
[79] D. Horvat, S. Ilijić, and A. Marunović, Classical and Quantum Gravity **28**, 025009 (2011), 1010.0878.
[80] R. L. Bowers and E. P. T. Liang, ApJ **188**, 657 (1974).
[81] V. Folomeev and V. Dzhunushaliev, Phys. Rev. D **91**, 044040 (2015), 1501.06275.
[82] T. E. Riley, A. L. Watts, P. S. Ray, S. Bogdanov, S. Guillot, S. M. Morsink, A. V. Bilous, Z. Arzoumanian, D. Choudhury, J. S. Deneva, et al., ApJ **918**, L27 (2021), 2105.06980.
[83] M. C. Miller, F. K. Lamb, A. J. Dittmann, S. Bogdanov, Z. Arzoumanian, K. C. Gendreau, S. Guillot, W. C. G. Ho, J. M. Lattimer, M. Loewenstein, et al., ApJ **918**, L28 (2021), 2105.06979.
[84] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, et al., Phys. Rev. Lett. **121**, 161101 (2018), 1805.11581.
[85] D. Clowe, A. Gonzalez, and M. Markevitch, ApJ **604**, 596 (2004), astro-ph/0312273.
[86] M. Bradač, S. W. Allen, T. Treu, H. Ebeling, R. Massey, R. G. Morris, A. von der Linden, and D. Applegate, ApJ **687**, 959 (2008), 0806.2320.
[87] D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, ApJ **648**, L109 (2006), astro-ph/0608407.
[88] R. Thompson, R. Davé, and K. Nagamine, MNRAS **452**, 3030 (2015), 1410.7438.
[89] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradač, ApJ **679**, 1173 (2008), 0704.0261.
[90] C. Gruber and A. Pelster, European Physical Journal D **68**, 341 (2014), 1403.3812.
[91] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, Reviews of Modern Physics **71**, 463 (1999), cond-mat/9806038.
[92] P.-H. Chavanis and T. Harko, Phys. Rev. D **86**, 064011 (2012), 1108.3986.
[93] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, Science **269**, 198 (1995).
[94] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, Phys. Rev. Lett. **75**, 3969 (1995), URL https://link.aps.org/doi/10.1103/PhysRevLett.75.3969.
[95] K. B. Davis, M.-O. Mewes, M. A. Joffe, M. R. Andrews, and W. Ketterle, Phys. Rev. Lett. **74**, 5202 (1995), URL https://link.aps.org/doi/10.1103/PhysRevLett.74.5202.
[96] E. Tiesinga, C. Williams, P. Julienne, K. Jones, P. Lett, and W. Phillips, *A spectroscopic determination of scattering lengths for sodium atom collisions* (1996).
[97] H. M. J. M. Boesten, C. C. Tsai, J. R. Gardner, D. J. Heinzen, and B. J. Verhaar, Phys. Rev. A **55**, 636 (1997).
[98] B. Dutta and L. E. Strigari, Annual Review of Nuclear and Particle Science **69**, 137 (2019), 1901.08876.
[99] F. Ruppin, J. Billard, E. Figueroa-Feliciano, and L. Strigari, Phys. Rev. D **90**, 083510 (2014), 1408.3581.

arXiv:2402.18965v1 [astro-ph.CO] 29 Feb 2024

# Induced Gravitational Wave interpretation of PTA data: complete study for general equation of state

Guillem Domènech[a,b],* Shi Pi[c,d,e],† Ao Wang[c,f],‡ and Jianing Wang[c,f,g]§

[a] *Institute for Theoretical Physics, Leibniz University Hannover, Appelstraße 2, 30167 Hannover, Germany.*

[b] *Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, 30167 Hannover, Germany*

[c] *CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*

[d] *School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

[e] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*

[f] *Center for High Energy Physics, Peking University, Beijing 100871, China and*

[g] *International Center for Quantum-field Measurement Systems for Studies of the Universe and Particles (QUP, WPI), High Energy Accelerator Research Organization (KEK), Oho 1-1, Tsukuba, Ibaraki 305-0801, Japan*

(Dated: March 1, 2024)

We thoroughly study the induced gravitational wave interpretation of the possible gravitational wave background reported by PTA collaborations, considering the unknown equation of state $w$ of the early universe. We perform a Bayesian analysis of the NANOGrav data using the publicly available PTArcade code together with SIGWfast for the numerical integration of the induced gravitational wave spectrum. We focus on two cases: a monochromatic and a log-normal primordial spectrum of fluctuations. For the log-normal spectrum, we show that, while the results are not very sensitive to $w$ when the GW peak is close to the PTA window, radiation domination is out of the $2\sigma$ contours when only the infra-red power-law tail contributes. For the monochromatic spectrum, the $2\sigma$ bounds yield $0.1 \lesssim w \lesssim 0.9$ so that radiation domination is close to the central value. We also investigate the primordial black hole (PBH) counterpart using the peak formalism. We show that, in general terms, a larger width and stiffer equation of state alleviates the overproduction of PBHs. No PBH overproduction requires $w \gtrsim 0.42$ up to 2-$\sigma$ level for the monochromatic spectrum. Furthermore, including bounds from the cosmic microwave background, we find in general that the mass range of the PBH counterpart is bounded by $10^{-5} M_\odot \lesssim M_{\rm PBH} \lesssim 10^{-1} M_\odot$. Lastly, we find that the PTA signal can explain the microlensing events reported by OGLE for $0.42 \lesssim w \lesssim 0.50$. Our work showcases a complete treatment of induced gravitational waves and primordial black holes for general $w$ for future data analysis.

* guillem.domenech@itp.uni-hannover.de

† shi.pi@itp.ac.cn

‡ wangao@itp.ac.cn

§ wangjianing@itp.ac.cn

## I. INTRODUCTION

The world-wide pulsar timing array (PTA) collaborations [1–10] reported tentative evidence of a low-frequency Gravitational Wave (GW) background. Although an astrophysical GW background is expected from inspiraling supermassive black hole binaries [11],[1] the simplest predictions seem to be in mild tension with the data [15]. Other plausible explanations of the GW background have a cosmological origin [16–18]. For instance, one of the cosmic interpretations that seems to behave better in the Bayesian analysis of NANOGrav [15] are GWs induced by primordial curvature fluctuations (see also Refs. [19–55] for subsequent works). If that is the case, these results hint at new physics during cosmic inflation and possibly a significant abundance of Primordial Black Holes (PBHs).

There is strong evidence of cosmic inflation from observations of the Cosmic Microwave Background (CMB) [56], with a measured amplitude of the power spectrum of primordial curvature fluctuations of about $10^{-9}$ at Mpc – Gpc scales. Yet, to reproduce the reported amplitude of the GW background, the spectrum of primordial fluctuations should have a peak amplitude of about $10^{-2}$ at roughly parsec scales [15]. Nevertheless, there are plenty of inflationary models capable of enhancing the primordial spectrum of curvature fluctuations on small scales, such as ultra-slow-roll (USR) [57–62], sound speed resonance (SSR) [63, 64], multi-field inflation [65–67], and the curvaton scenario [68–71] among others (see *e.g.* Ref. [72] for a review and references therein). Most interestingly, though, such a large variance for primordial fluctuations implies that some fluctuations may be large enough to end up forming PBHs [73–80] (for recent reviews on PBHs see also Refs. [81–85]).

PBHs associated with the induced GW interpretation of the PTA data (assuming a radiation dominated universe) apparently have a mass around $10^{-4}M_\odot - 10^{-2}M_\odot$ [20, 22], where $M_\odot$ is a solar mass. However, despite important uncertainties in the calculations of the PBH abundance, one of the issues pointed out in Ref. [15] is that the amplitude of the primordial spectrum given by the $2\sigma$ contours from PTA analysis tends to predict too many PBHs in that mass range (assuming Gaussian primordial fluctuations and a radiation dominated very early universe). However, while the induced GW spectrum depends quadratically on the power spectrum, PBH abundance is exponentially sensitive to it. This means, on the one hand, that finding induced GWs might not necessarily imply a significant fraction of PBHs. For example, Ref. [22] argues that a small fraction of the available parameter space could explain the planet-mass microlensing events seen by OGLE [86, 87]. This is an interesting possibility as it may also give a detectable GW background from the mergers of PBH binaries [22]. On the other hand, it also means that relatively small changes in the statistical nature of primordial fluctuations could significantly suppress the PBH abundance, as is the case of negative non-Guassianity of primordial fluctuations[2] [20, 24, 25, 27, 31, 47] (which might not be straightforward to obtain in single field models of inflation [94, 95]) and a period of the universe deviating from standard radiation domination [33, 35, 36, 41, 96, 97].

In this paper, we take an agnostic approach. Since the induced GWs and the PBHs correspond-

[1] See though Refs. [12–14] interpreting the GW background as the merger of supermassive primordial black holes.
[2] Induced GWs are not very sensitive to the amount of non-Gaussianity, see *e.g.* Ref. [51, 88–93].

ing to PTA frequencies were generated and formed when the temperatures in the universe were above 100 MeV (a period of the universe which has not been probed directly), we do not assume any particular value of the equation of state (EoS) of the universe, commonly denoted $w$. For comparison, note that a successful Big Bang Nucleosynthesis (BBN) requires the universe to be radiation dominated ($w = 1/3$) and thermalized for $T \gtrsim 4\,\mathrm{MeV}$ [98–101] (or at redshifts greater than $10^9$). Other values of the EoS are tied to different models of reheating (see, *e.g.*, Ref. [102] for a review). For example, the coherent oscillations of a scalar field yield $w \sim 0$, and in quintessential inflation scenarios, one has a period with $w = 1$ after inflation [103–106]. In this way, since the induced GW spectrum depends on $w$, it serves as a probe of the content of the primordial universe [107–124].

We would like to emphasize that while this approach has also been considered for PTAs in Refs. [35, 36, 41, 96, 97], we improve previous analysis by performing the most thorough study on both the induced GWs and PBHs related to PTAs. Requiring no PBH overproduction, we obtain that $0.42 \lesssim w \lesssim 0.87$. And, using CMB bounds on the effective number of relativistic species (which sets upper bounds on the peak GW frequency), we conclude that the allowed range of the typical PBH masses is given by $2 \times 10^{-5} M_\odot \lesssim M_{\rm PBH} \lesssim 0.09 M_\odot$. Interestingly, the mass range allowed by OGLE for PBHs, namely $[1.6 \times 10^{-6} M_\odot, 6.9 \times 10^{-5} M_\odot]$ [87], overlaps slightly with the range we obtained for $0.42 \lesssim w \lesssim 0.50$. There are also candidate merger events in the LVK data of compact objects with a mass of about $O(0.1) M_\odot$ [125–128]. For simplicity, we do not include non-Gaussianity of primordial fluctuations. But, as future work, it would be interesting to study both effects simultaneously.

This paper is organized as follows. In § II, we review the induced GW spectrum generated during the general EoS stage for a log-normal primordial spectrum. In § III, we present the results of our Bayesian analysis for a log-normal and a Dirac delta primordial spectra. We conclude our work in § IV with further discussions. Lastly, we provide details on the PBH calculations in App. A and on the CMB bounds on extra relativistic species in B.

## II. INDUCED GRAVITATIONAL WAVES IN A GENERAL $w$ STAGE

Whenever there are fluctuations in the universe, there is inevitably a secondary production of GWs [129–131]. The spectrum of such induced GWs [132–137] (see Refs. [138, 139] for recent reviews) depends on the spectrum of such fluctuations as well as on the EoS of the universe at the time of generation [117, 138, 140]. Here, as is common practice, we assume that around a specific wavenumber $k_{\rm p}$, there is a peak in the primordial spectrum of curvature fluctuations $\mathcal{R}$. For analytical simplicity and inspired by inflationary models [57–67], we parametrize such a peaked (dimensionless) spectrum with a log-normal shape given by [141]

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{\mathcal{A}_{\mathcal{R}}}{\sqrt{2\pi}\Delta} \exp\left[-\frac{\ln^2\left(k/k_{\rm p}\right)}{2\Delta^2}\right], \tag{2.1}$$

where $\mathcal{A}_{\mathcal{R}}$ is the amplitude, $\Delta$ the logarithmic width and $k_{\rm p}$ the scale where the power spectrum peaks. Quite often, one also considers a Dirac delta power spectrum, which corresponds to the zero

width limit of the log-normal, as an exact analytical case for the induced GW spectrum.

Since induced GWs (and PBHs) were produced during an era dominated by a perfect fluid with a constant $w$, there is another important scale besides the peak scale $k_{\rm p}$, namely the scale of reheating $k_{\rm rh}$. For simplicity, we take an almost instantaneous transition from a general $w$ to the standard radiation domination with $w = 1/3$.[3] In this case, the reheating scale corresponds to the scale that entered the horizon at the moment of reheating, that is $k_{\rm rh} = a_{\rm rh}H_{\rm rh}$ where $a_{\rm rh}$ and $H_{\rm rh}$ respectively are the scale factor and the Hubble parameter at reheating. For further simplicity, we also assume that the frequency corresponding to the reheating scale evaluated today, namely $f_{\rm rh} = k_{\rm rh}/(2\pi a_0)$, is smaller than the lowest frequency bin of PTA data. In other words, we impose $f_{\rm rh} < f_{\rm rh,max} \approx 2 \times 10^{-9}\,{\rm Hz}$. By doing so, we do not need the explicit expression of the induced GW spectrum for $f < f_{\rm rh}$ as it never enters the observational range. Also, it is important to note that for $w = 1/3$, the "reheating" scale is meaningless, as there is no reheating. In that case, $f_{\rm rh}$ can be considered an arbitrary pivot scale.

If we let $f_{\rm rh}$ vary, we would find two possibilities: either $f_{\rm rh}$ is well inside or below the PTA frequencies. The former is not so interesting as one mostly recovers the results of standard radiation domination $w = 1/3$, already studied in [15]. The latter is bounded from below by BBN constraints on the reheating temperature, $T \gtrsim 4\,{\rm MeV}$ [98–101]. From now on, we consider a fixed $f_{\rm rh}$ because our main focus is only to understand the effects of different values of $w$. For practical purposes, we set

$$f_{\rm rh,\star} \approx 4 \times 10^{-10}\ {\rm Hz}\,, \tag{2.2}$$

which corresponds to $T_{\rm rh,\star} = 20\,{\rm MeV}$, assuming the standard model of particle physics. Such an arbitrary choice is roughly in the middle of the lowest frequency of PTAs ($\sim 10^{-9}\,{\rm Hz}$) and the frequency corresponding to the lowest reheating temperature ($\sim 10^{-10}\,{\rm Hz}$). Our main results do not depend significantly on the choice of the reheating scale, yielding at most some $\mathcal{O}(1)$ factor variations, as we show later. For future reference, the temperature after reheating $T_{\rm rh}$ is given by

$$T_{\rm rh} = 110\,{\rm MeV}\,\left(\frac{f_{\rm rh}}{f_{\rm rh,max}}\right)\,\left(\frac{g_{\rho,{\rm rh}}}{10.75}\right)^{-1/2}\left(\frac{g_{s,{\rm rh}}}{10.75}\right)^{1/3}\,, \tag{2.3}$$

where $g_{\rho,{\rm rh}}$ and $g_{s,{\rm rh}}$ respectively are the effective degrees of freedom in the energy density $\rho$ and entropy density $s$ of the radiation fluid evaluated at the reheating temperature, $T_{\rm rh}$. For the values of $g_{\rho,{\rm rh}}$ and $g_{s,{\rm rh}}$ in the standard model of particles see, *e.g.*, Ref. [143]. Note that in Eq. (2.3) there is an implicit dependence on $T_{\rm rh}$ inside $g_{\rho,{\rm rh}}$ and $g_{s,{\rm rh}}$, which we neglected as we are considering the small range $4\,{\rm MeV} < T_{\rm rh} < 110\,{\rm MeV}$.

The spectral energy density of the induced GWs produced during an era dominated by a perfect fluid with a constant $w > 0$ is given by [117, 121, 138, 140]

$$\Omega_{\rm GW,rh}(k > k_{\rm rh}) = \left(\frac{k}{k_{\rm rh}}\right)^{-2b}\int_0^1 {\rm d}d\int_1^\infty {\rm d}s\ \mathcal{T}_w(d,s)\,\mathcal{P}_\mathcal{R}\left(\frac{k}{2}(s+d)\right)\mathcal{P}_\mathcal{R}\left(\frac{k}{2}(s-d)\right), \tag{2.4}$$

[3] Note that an exact instantaneous transition leaves some high frequency oscillatory patterns in the induced GW spectrum, as shown in Ref. [142]. However, we believe these are an artifact of the instantaneous transition and would quickly disappear as one considers a finite duration. Also, when the instantaneous transition goes from $c_s^2 = w = 0$ to $w = 1/3$, there is an enhanced production of the induced GWs [108, 109]. We do not consider this case here.

where

$$b = \frac{1-3w}{1+3w}\,, \tag{2.5}$$

and $\mathcal{T}_w(d,s)$ is the general kernel for modes reentering horizon before reheating, given, *e.g.*, in Refs. [121, 138]. In Eq. (2.4), $d$ and $s$ are related to the internal momenta of the scalar modes. For the first derivation of semi-analytical formulas in radiation domination, see Refs. [144, 145]. Note that in (2.4), the factor $(k/k_{\rm rh})^{-2b}$ comes from the dilution of the energy density of subhorizon GW modes, which decay as radiation (*i.e.* $\rho_{\rm GW} \propto a^{-4}$), as compared to background energy density (that is $\rho_w \propto a^{-3(1+w)}$). Using entropy conservation, the current spectral density of induced GWs reads [146]

$$\begin{aligned}\Omega_{\rm GW,0}(k)h^2 &= \frac{g_{\rho,\rm rh}}{g_{\rho,0}}\left(\frac{g_{s,0}}{g_{s,\rm rh}}\right)^{4/3}\Omega_{\rm GW,rh}(k)\Omega_{r,0}h^2 \\ &= 3.40\times10^{-5}\left(\frac{g_{s,\rm rh}}{10.75}\right)^{-4/3}\left(\frac{g_{\rho,\rm rh}}{10.75}\right)\left(\frac{\Omega_{r,0}h^2}{2.47\times10^{-5}}\right)\Omega_{\rm GW,rh}(k)\,,\end{aligned} \tag{2.6}$$

where $\Omega_{r,0}h^2$ is the reduced density ratio of radiation today given by Planck [147]. From Eq. (2.4) we see that the amplitude of the induced GW spectrum depends on the amplitude $\mathcal{A}_\mathcal{R}$ of the power spectrum (2.1), the equation of state $w$ and the ratio of scales $k_{\rm p}/k_{\rm rh}$.

Before we get into the details and results of the Bayesian analysis, it is useful to have an analytic understanding of the spectrum of induced GWs in the low frequency tail, also called infra-red (IR) tail, which corresponds to $k_{\rm rh} \ll k \ll k_{\rm p}$. This is because, in the end, the power-law behavior of the IR tail mostly fits the PTA data. We now discuss the case of a Dirac delta and a log-normal primordial spectrum separately.

### A. Dirac delta peak

We start with a Dirac delta primordial spectrum given by

$$\mathcal{P}_\mathcal{R}(k) = \mathcal{A}_\mathcal{R}\delta(\ln(k/k_{\rm p}))\,. \tag{2.7}$$

In this case, the spectrum of induced GW after inserting Eq. (2.7) into Eq.(2.4) exactly reads

$$\Omega^{(\delta)}_{{\rm GW},w}(k) = 2\,\mathcal{A}^2_\mathcal{R}\left(\frac{k_{\rm p}}{k_{\rm rh}}\right)^{-2b}\left[\left(\frac{k}{k_{\rm p}}\right)^{-2-2b}\mathcal{T}_w(s=2k/k_{\rm p},d=0)\Theta(2-k/k_{\rm p})\right]. \tag{2.8}$$

We are interested in the IR tail of the induced GWs, which corresponds to the limit $k \ll k_{\rm p}$ of Eq. (2.8). After a Taylor expansion, one finds [118, 138]

$$\Omega^{(\delta)}_{{\rm GW},w}(k\ll k_{\rm p}) \approx F[b,c_s]\,\frac{\mathcal{A}^2_\mathcal{R}}{c_s^4}\left(\frac{k_{\rm p}}{k_{\rm rh}}\right)^{-2b}\left(\frac{k}{k_{\rm p}}\right)^{2-2|b|}, \tag{2.9}$$

where $c_s^2 = w$ is the sound speed of fluctuations, and $F[b, c_s]$ is a function of $b$ and $c_s$, whose concrete form is irrelevant for the following discussion.[4]

Using Eq. (2.9) we may study degeneracies among model parameters, which are relevant for the subsequent data analysis. First, we note that the function $F[b, c_s]$ yields at most $\mathcal{O}(1-10)$ variations for $1 > b > -1/2$ (or $0 < w < 1$) and $c_s^2 = w$. For instance, for $b = 1$ and $b = -1/2$ one respectively finds $F[b, c_s] \approx 0.06$ and $F[b, c_s] \approx 0.95$. This shows that the IR tail has a mild dependence on $w$ through the function $F[b, c_s]$. But, since the factor $c_s^{-4}$ in Eq. (2.9) yields a much larger change for varying $w$, we neglect the dependence on $F[b, c_s]$ in the following discussion. Thus, the amplitude of the IR tail is then mostly dependent on the combination (for a fixed $k$) given by

$$\frac{\mathcal{A}_{\mathcal{R}}}{c_s^2}\left(\frac{k_{\rm p}}{k_{\rm rh}}\right)^{-b}\left(\frac{k}{k_{\rm p}}\right)^{1-|b|} \approx \text{constant}. \tag{2.11}$$

For $b = 0$ ($w = 1/3$), one recovers the results of Ref. [33].

Since the reported amplitude of the GW background is fairly large, the peak of the induced GW spectrum cannot lie too far from the PTA window; otherwise, it might conflict with CMB constraints on the effective number of relativistic species. For this reason, it is also instructive to study the amplitude of the induced GW spectrum at the higher end of the IR tail, or the "bump" of the GW spectrum. Based on the results for $w = 1/3$ of Ref. [33], we find that the bump is located to a good approximation at

$$k_{\rm bump} \approx \frac{1}{\sqrt{2}} c_s k_{\rm p}\,, \tag{2.12}$$

with an amplitude given by

$$\Omega_{{\rm GW},w}^{(\delta),{\rm bump}} \approx \frac{C[b]}{c_s^{2(1+b)}}\left(1-\frac{c_s^2}{8}\right)^2 \mathcal{A}_{\mathcal{R}}^2 \left(\frac{k_{\rm p}}{k_{\rm rh}}\right)^{-2b} \tag{2.13}$$

where we introduced for compactness the function $C[b]$.[5] Such a function is bounded by $0.5 > C[b] > 0.1$ for $1 > b > -1/2$ ($0 < w < 1$). Thus, the induced GW spectrum has a similar amplitude and position of the bump (for a fixed $k$) when

$$\frac{\mathcal{A}_{\mathcal{R}}}{c_s^{1+b}}\left(\frac{k_{\rm p}}{k_{\rm rh}}\right)^{-b} \approx \text{constant} \qquad \text{and} \qquad \frac{1}{\sqrt{2}} c_s k_{\rm p} \approx \text{constant}\,. \tag{2.15}$$

From Eqs. (2.11) and (2.15), we see that increasing the value of $w$ allows for a decrease in the value of $\mathcal{A}_{\mathcal{R}}$. For $b < 0$ ($b > 0$), we see that increasing the value of $k_{\rm p}$ (or lowering the value of $k_{\rm rh}$) leads to a decrease (increase) to the necessary value of $\mathcal{A}_{\mathcal{R}}$. We will encounter such behavior in the Bayesian analysis of § III.

[4] The function $F[b, c_s]$ in Eq. (2.9) is given by

$$F[b,c_s] \equiv \frac{4^b(2+b)^2}{3(1+b)^{2(1+b)}}\Gamma^2[\tfrac{3}{2}+b]\begin{cases}\dfrac{\Gamma^2[\frac{3}{2}+b]}{c_s^{4b}\sin^2(b\pi)\Gamma^2[2+b]} & (b<0\ ;\ w>1/3)\\[2ex] \dfrac{\left(1+b+b^2\right)^2}{4\pi b^2(1+b)^4} & (b>0\ ;\ w<1/3)\end{cases}. \tag{2.10}$$

[5] The function $C[b]$ in Eq. (2.13) is given by

$$\begin{aligned} C[b] =& \frac{7^{2b}(1+b)^{-2(3+b)}(2+b)^2(18+25b)^2\,\Gamma^4[b+\frac{3}{2}]}{3\times 2^{11+5b}\,\Gamma^2[1+b]} \\ &\times\left\{1+\left(\cot(b\pi)+\frac{2^b\sqrt{\pi}\csc(\pi b)}{\Gamma[1-b]\Gamma[b+\frac{1}{2}]}\left(\frac{3\times 8^{1+b}b(2+b)}{7^b(1+2b)(18+25b)} - {}_2F_1\left[-b,1+b;1-b;\tfrac{1}{8}\right]\right)\right)^2\right\}. \end{aligned} \tag{2.14}$$

### B. Log-normal peak

The case of the log-normal peak (2.1) has not been studied analytically for general $w$. Analytical approximations for the case of radiation domination have been derived in Ref. [141]. Nevertheless, we may get some intuition if we focus on the far IR tail of the induced GW spectrum. In this case, most of the results of the Dirac delta carry over to the log-normal case with few changes. In particular, in the far IR, there is a change of slope from the $k^{2-2|b|}$ of the Dirac delta to $k^{3-2|b|}$ for the log-normal [118, 148]. We follow Ref. [141] to provide more precise estimates to derive some approximations in the far IR limit. Concretely, we perform the integral Eq. (2.4) in the limit $k \ll k_{\rm p}\Delta e^{-\Delta^2}$, which corresponds to the regime where $s \gg 1$ and $d \ll 1$ and where the log-normal power spectrum may be considered as a very sharp peak. By doing so, we obtain that

$$\Omega^{(\Delta)}_{{\rm GW},w}(k \ll k_{\rm p}\Delta e^{-\Delta^2}) \approx \Omega^{(\delta)}_{{\rm GW},w}(k \ll k_{\rm p}) \times \left(\frac{k}{k_{\rm p}}\right) \frac{1}{2\sqrt{\pi}\Delta} \begin{cases} e^{\left(\frac{3}{2}+2b\right)^2\Delta^2} & (b<0\ ;\ w>1/3) \\ e^{\frac{9}{4}\Delta^2} & (b\geq 0\ ;\ w\leq 1/3) \end{cases}, \tag{2.16}$$

where $\Omega^{(\delta)}_{{\rm GW},w}(k \ll k_{\rm p})$ is given in Eq. (2.9). It is worth mentioning that, while performing numerical integrations, we noticed that the approximations of Ref. [141] for $w = 1/3$ present some variations in the case of general $w$. In particular, the precise position of the change of slope from $k^{2-2|b|}$ to $k^{3-2|b|}$ depends on the value of $w$.[6] While interesting, this is out of the scope of this paper. We will present detailed analytical approximations in subsequent works.

## III. BAYESIAN ANALYSIS AND RESULTS

We perform the Bayesian analysis with the package PTArcade using the ceffyl model [149]. Also, as in the NANOGrav analysis [15], we only consider the first 14 frequency bins as signals for the GW background [15]. At the same time, the pulsar-intrinsic red noise is obtained considering all frequency bins. To derive the likelihood of the amplitude of the GW background signal, say $P(\Theta \mid \mathcal{D}, \mathcal{M})$, we need to obtain the induced GW spectra in various target models $\mathcal{M}$ with different model parameters $\Theta$ for analyzing the data $\mathcal{D}$. We do so as follows.

We use the analytical formula, Eq. (2.8), for the Dirac delta case. For the log-normal, we use the SIGWfast [140] package for an efficient numerical calculation of the energy spectra of induced GWs for general $w$. Since the role of sound speed has already been discussed in Ref. [33] (although for fixed $w = 1/3$), we only consider the adiabatic perfect fluid case with $c_s^2 = w$ in our sampling. We checked that the case of $c_s^2 = 1$ yields similar results for the GW spectrum. In our analysis, we also consider *a posteriori* the PBH counterpart and CMB constraints on the effective number

[6] For example, for a narrow log-normal peak ($\Delta < 0.1$) there is a good approximation to the log-normal case for $k < \Delta\, k_{\rm p}$ and $w = 1/3$ which is given by [138, 141]

$$\Omega^{(\Delta)}_{\rm GW} = {\rm Erf}\left[\frac{1}{\Delta}\sinh^{-1}\frac{k}{2k_{\rm p}}\right]\Omega^{(\delta)}_{\rm GW}. \tag{2.17}$$

This equation should be derived more precisely for general values of $w$ by including the dependence on $w$ in the function in front of $\Omega^{(\delta)}_{\rm GW}$.

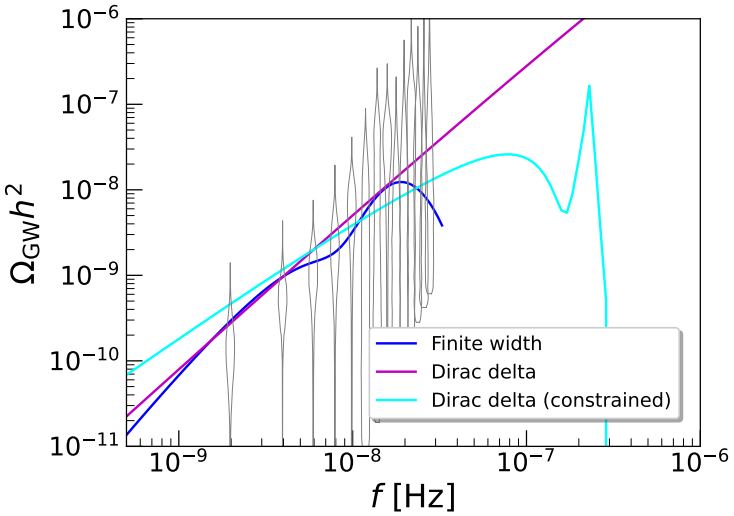



FIG. 1. Best fit induced GW spectra for different models using the first 14 frequency bins of NANOGrav data (the gray violins). The best fit parameters for the Dirac delta (2.7) read $[w, \log_{10}\mathcal{A}, \log_{10} f_{\rm p}] = [0.35, 0.98, -3.79]$, shown in purple. Including the bound from no PBH overproduction the best fit moves to $[w, \log_{10}\mathcal{A}, \log_{10} f_{\rm p}] = [0.59, -2.02, -6.81]$ which we show with a cyan line. For the log-normal (2.1) we find $[w, \log_{10}\mathcal{A}, \log_{10} f_{\rm p}, \Delta] = [0.31, -1.60, -7.79, 0.49]$.

of relativistic species. For details on the calculations of the PBH abundance and CMB constraints, we refer the reader respectively to Apps. A and B. We show the GW spectrum for the best fit parameters of different models in Fig. 1. The results of the Bayesian analysis are shown in Figs. 2 and 4, and our priors and 95% posterior confidence intervals are presented in Tabs. I and II.

For clarity, we summarize below the parametrization used in our analysis. We compute the induced GW spectrum (2.4) as

$$\Omega_{\rm GW,rh}(f) = \mathcal{A}^2 \left(\frac{f}{f_{\rm rh,\star}}\right)^{-2b} \int_0^1 {\rm d}d \int_1^\infty {\rm d}s\, \mathcal{T}_w(d,s)\, G_\Delta\left(\frac{f(s+d)}{2f_{\rm p}}\right) G_\Delta\left(\frac{f(s-d)}{2f_{\rm p}}\right) , \tag{3.1}$$

where $f_{\rm rh,\star} \approx 4 \times 10^{-10}$ Hz (as explained around Eq. (2.2)),

$$G_\Delta(x) = \frac{1}{\sqrt{2\pi}\Delta} \exp\left[-\frac{\ln^2 x}{2\Delta^2}\right] . \tag{3.2}$$

and we re-defined

$$\mathcal{A} \equiv \mathcal{A}_\mathcal{R} \left(\frac{f_{\rm rh}}{f_{\rm rh,\star}}\right)^b . \tag{3.3}$$

With the above redefinition, Eq. (3.3), the constraints we find in the variable $\mathcal{A}$ can be easily translated to constraints in $\mathcal{A}_\mathcal{R}$ for a given value of $f_{\rm rh}$. For instance, for $f_{\rm rh,min} \sim 10^{-10}$ (roughly corresponding to $T_{\rm rh} \sim 4\,$MeV) and $f_{\rm rh,min} \sim 10^{-9}$ (roughly corresponding to $T_{\rm rh} \sim 100\,$MeV) we respectively have that $f_{\rm rh,min}/f_{\rm rh,\star} \sim 4/20 \sim 1/5$ and $f_{\rm rh,max}/f_{\rm rh,\star} \sim 100/20 \sim 5$. In fact, since

| $c_s^2 = w$ | $w$ | $\log_{10}\mathcal{A}$ | $\log_{10} f_{\rm p}$ |
|---|---|---|---|
| Prior | Uniform [0.01,1] | Uniform [-5,3] | Uniform [-8.5,-5] |
| 95%-HDPI | $[0.134, 0.867]$ | $[-2.64, 1.49]$ | $[-6.99, \mathrm{NaN}]$ |

TABLE I. Priors and 95% HPDI of the posterior marginalized distributions of the Dirac delta case (2.7).

for the Dirac delta case we eventually find that no PBH overproduction requires $w > 0.42$, the variation in reheating frequency can only amount at most to a factor $\sim 2$ for $f_{\rm rh,min}$ (or $\sim 1/2$ for $f_{\rm rh,max}$) taking the maximum value of $|b|$ in that range, namely $b = -1/2$. Thus, our results for a fixed $f_{\rm rh} = f_{\rm rh,\star}$ already provide good intuition for the general case of $f_{\rm rh,min} < f_{\rm rh} < f_{\rm rh,max}$. We proceed to discuss our results for the Dirac delta and log-normal cases.

### A. Results for Dirac delta peak

We present the priors used and the 95% highest probability distribution interval (HPDI) of the marginalized distributions in the Dirac delta case in Tab. I. In Fig. 2, we show the posterior distributions with the NANOGrav dataset. In the same figure, we present with an orange dashed line the upper bound on the amplitude of primordial fluctuations from the requirement that PBHs constitute 100% of the dark matter (*i.e.*, above the orange line PBHs are overproduced and inconsistent with observations). The gray dashed vertical line then shows the lower bound on the equation of state $w$ with no PBH overproduction at the 2-$\sigma$ level. With a purple dashed line, we also give the upper bounds on the GW energy density from CMB observations, as explained in App. B. Dashed black lines illustrate the PBH mass range compatible with the OGLE microlensing events, which could be explained by planet-mass PBHs. We provide a more detailed plot of the typical PBH mass in terms of $f_{\rm p}$ and $w$ in Fig. 3. With a red and an orange star, we respectively point the best fit values before and after, including the bound from no PBH overproduction.

From Fig. 2, we also see the degeneracies discussed in § II A, Eqs. (2.11) and (2.15). Let us start with the $\mathcal{A}$-$w$ plane. There, we see how increasing the value of $b$ (or, equivalently, lowering the value of $w$) increases the negative power in the ratio $f_{\rm p}/f_{\rm rh}$ and leads to a bigger value of $\mathcal{A}$. For low enough $w$ ($w < 1/3$), we also see that there is a spread in the values of $\mathcal{A}$ that comes from the factor $c_s^{-4} = w^{-2}$, such that lowering $w$ also allows for smaller $\mathcal{A}$. Next, in the $f_{\rm p}$-$\mathcal{A}$ plane, we see a degeneracy along the $\mathcal{A}_{\mathcal{R}} \propto f_{\rm p}^{1+b-|b|}$ direction. Increasing $f_{\rm p}$ requires larger values of $\mathcal{A}$. Note that because of the term $|b|$ in the exponent, there are two degeneracy lines, one for $b < 0$ and one for $b > 0$. We do not see them in Fig. 2 because the range of $w$ is all connected. However, we will see the two degeneracy lines in the log-normal case. Lastly, we see no particular dependence of $f_{\rm p}$ on $w$. This is expected as the $f_{\rm p}$ and $w$ directions are somewhat unrelated.

We also find interesting implications from our analysis. First, from the IR tail, that is $\Omega_{\rm GW} \propto f^{2-2|b|}$, we may constrain $w$. In fact, from the $2\sigma$ bounds, we see that $0.13 < w < 0.87$. This is consistent with analytical scaling arguments, as also pointed out in Ref. [22]. Since the slope of the power law fitting for NANOGrav roughly satisfies $5/2 > n_{\rm PTA} > 1$, it yields $|b| < 1/2$, or equivalently $1/9 < w < 1$. We note that, although our results are also qualitatively consistent

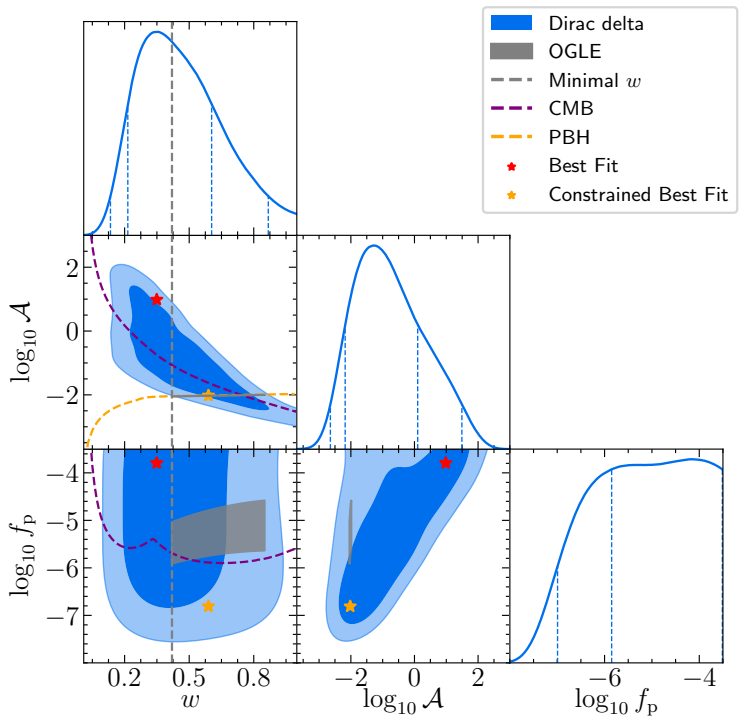



FIG. 2. Posterior distribution inferred from the NANOGrav spectrum probability distribution function for the Dirac delta primordial spectrum (2.7). In orange and purple dashed lines, we respectively show the constraints from no PBH overproduction (see App. A) and the effective number of relativistic species in the CMB (see App. B). The regions above these lines are excluded from the parameter space. The gray shaded region shows the corresponding parameters space for the 95% CL mass range in OGLE 5-year [87], including the PBH abundance. The stars show the best fit parameters before (red) and after (orange) considering PBH and CMB bound. We see that there is still a small parameter space to explain OGLE microlensing events and PTA signal simultaneously. The blue dashed lines in the marginalized distributions show the 68% and 95% HDPI in the marginalized distributions.

with the results of Refs. [35, 36], we find it difficult to compare with their results due to some simplifications in these works. Ref. [36] also includes $f_{\rm rh}$ as a free parameter.

The results become more restrictive once we include information from the PBH counterpart and CMB constraints. However, it is important to be aware that the PBH abundance is exponentially sensitive to several factors, such as the precise formulation one uses, how one computes the threshold or the amount of non-Gaussianity (for the latter see Refs. [20, 24, 25, 27, 31, 47]). Thus, any implication from the PBH counterpart should be taken with a grain of salt, and it may differ from the analysis of Refs. [35, 36, 41] by some $\mathcal{O}(1-10)$. Here, we use the formalism developed in Refs. [150, 151] based on peak theory, which we believe is the currently most accurate method. Including the PBH counterpart, we first find that no PBH overproduction excludes $w = 1/3$ at $2\sigma$. To be more precise, we find $w \gtrsim 0.42$. We also note that for $w > 0.8$ there is no PBH overproduction for the 2-$\sigma$ contour. This is consistent with [33, 35, 36, 41] where increasing the equation of state

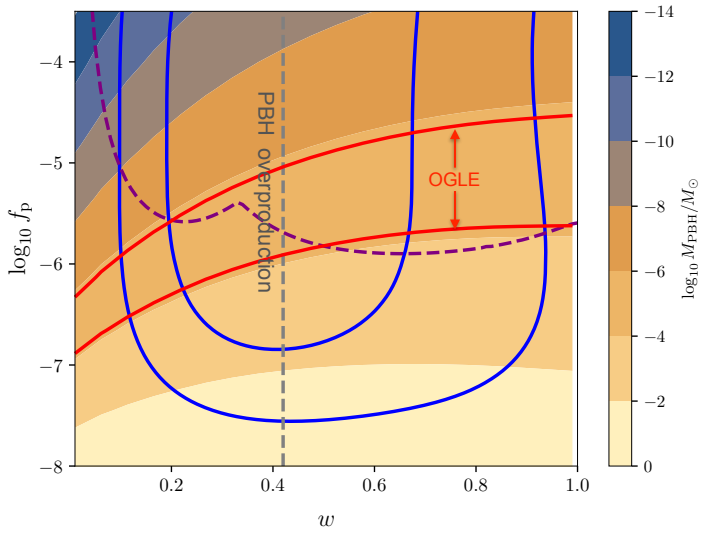



FIG. 3. Peak of the mass spectrum, $\log_{10} M/M_\odot$, of the PBH counterpart to the induced GWs with respect to different EoS $w$ and peak frequency $\log_{10} f_{\rm p}$. Red lines show the minimum and the maximum PBH mass (respectively $M_{\rm min} = 1.6 \times 10^{-6} M_\odot$ and $M_{\rm max} = 6.9 \times 10^{-5} M_\odot$) compatible with the microlensing events seen by OGLE [87]. Gray lines indicate the EoS $w$ in which there are too many PBHs (left of the dashed line) and too few PBHs to be observed (right of the dot-dashed line). The blue lines show the 1-$\sigma$ and 2-$\sigma$ of NANOGrav data, and the purple line shows the CMB bound as Fig. 2.

(or the sound speed) suppresses PBH formation.

Including bounds from CMB, the parameter space is further bounded. We can zoom in on the $f_{\rm p}$-$w$ plane in the Bayesian contour plot to discuss the allowed PBH mass considering constraints imposed by our CMB and PBH results in Fig. 3. We conclude that the available parameter space compatible with PTAs, without PBH overproduction, and obeys the CMB constraint is roughly given by $0.42 \lesssim w \lesssim 0.87$ and predicts a typical PBH mass range between $2.5 \times 10^{-5} M_\odot \lesssim M_{\rm PBH} \lesssim 0.09 M_\odot$. If the planet-mass PBHs can further explain OGLE microlensing events, the overlapped parameter space gives $0.42 \lesssim w \lesssim 0.50$. It is remarkable that by combining PTAs, CMB, and PBHs, one can considerably narrow down the value of $w$. Although our conclusions align with the results of Ref. [22], it is worth noting that because Press-Schechter significantly underestimates PBH production, we have excluded the case of $w = 1/3$ in this paper at a 2-$\sigma$ level. Interestingly, the upper bound of $\mathcal{O}(0.1) M_\odot$ could be related to candidate merger events in the LVK data with sub-solar mass compact objects [125–128, 152].

### B. Results for Finite width peak

We now include a finite width $\Delta$ in the primordial spectrum using the log-normal peak (2.1). We use the code SIGWFAST to obtain more accurate induced GW spectra for different widths, especially for scales around the peak. This is an important feature since a broader primordial

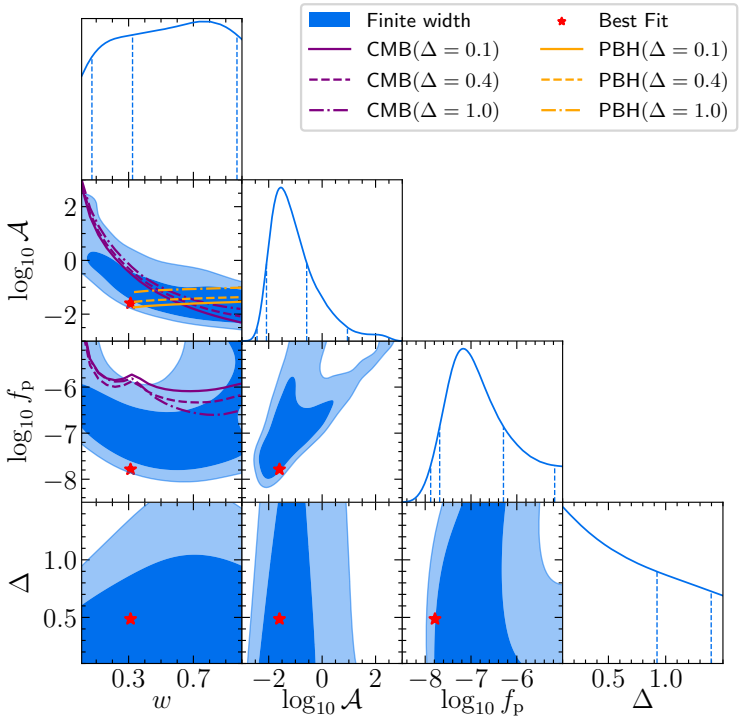



FIG. 4. Posterior distribution inferred from the NANOGrav spectrum probability distribution function for the log-normal delta primordial spectrum (2.1). In orange and purple, we respectively show the constraints from no PBH overproduction (see App. A) and the effective number of relativistic species in the CMB (see App. B). The solid, dashed, and dotted-dashed lines, respectively, denote the constraints for $\Delta = 0.1$, $\Delta = 0.4$, and $\Delta = 1$. For the PBH constraints, we only show $w \geq 1/3$ as the threshold values for $w < 1/3$ are unavailable for the broad peak case. The red stars show the best fit parameters. The blue dashed lines in the marginalized distributions show the 68% and 95% HDPI in the marginalized distributions.

spectrum leads to a broader peak in the induced GW spectrum, which may also fit the PTA data. This was not possible for the Dirac delta case as the resonant peak in the induced GW spectrum is very sharp. We also find from the formulas in App. A that a broader primordial spectrum leads to softer bounds on the amplitude from no PBH overproduction. Our results for the prior and the 68% HPDI of the marginalized distributions are displayed in Tab. II. In Fig. 4, we present the posterior distribution with the NANOGrav dataset. In the same figure, we show the constraints from no PBH overproduction and CMB in orange and purple, respectively. The solid, dashed, and dotted-dashed lines, respectively, denote the constraints for $\Delta = 0.1$, $\Delta = 0.4$, and $\Delta = 1$. The total power is normalized to be $\mathcal{A}_\mathcal{R}$, so increasing $\Delta$ will decrease the peak value of the spectrum ($\sim \mathcal{A}_\mathcal{R}/(\sqrt{2\pi}\Delta)$), thus the PBHs generated are suppressed significantly. Therefore, we expect that for broad peaks, the constraints from no PBH overproduction become looser, as is shown clearly in Fig. 4. Note that the orange lines (no PBH overproduction) stop at $w = 1/3$ as the threshold values for $w < 1/3$ are not available to us for the log-normal case. The red star shows the best fit values.

| $c_s^2 = w$ | $w$ | $\log_{10}\mathcal{A}$ | $\log_{10} f_{\rm p}$ | $\Delta$ |
|---|---|---|---|---|
| Prior | Uniform [0.01,1] | Uniform [-5,3] | Uniform [-8.5,-5] | Uniform [0.1,1.5] |
| 95%-HDPI | $[0.073, 0.979]$ | $[-2.43, 0.95]$ | $[-7.88, -5.17]$ | [NaN, 1.40] |

TABLE II. Priors and 95% HPDI of the posterior marginalized distributions for the log-normal case (2.1).

As we include one more parameter $\Delta$, the 1- and 2-$\sigma$ contours in Fig. 4 are more complicated than in the Dirac delta case in Fig. 2. However, we may understand the main effects of $\Delta$ as follows. On the one hand, since the log-normal peak (2.1) is normalized, the amplitude of the resulting induced GW spectrum around the peak does not change significantly with $\Delta$. Although we see from Eq. (2.16) that the far IR tail is sensitive to large values of $\Delta$, it still does not matter much with the current data. These are the reasons why in Fig. 4 the values of $\mathcal{A}$, $f_{\rm p}$ and $w$ are not sensitive to the precise value of $\Delta$. On the other hand, and most importantly, the presence of $\Delta$ introduces two distinct regimes roughly separated at $f_{\rm p} \sim 10^{-6}\,{\rm Hz}$ where the PTA data is fitted either by $(i)$ the peak in the induced GW spectrum for $f_{\rm p} < 10^{-6}\,{\rm Hz}$ or $(ii)$ the IR tail for $f_{\rm p} > 10^{-6}\,{\rm Hz}$. This provides different features in the $f_{\rm p}$-$w$ and the $f_{\rm p}$-$\mathcal{A}$ planes with respect to the Dirac delta case, which we explain below.

First, look at the $f_{\rm p}$-$w$ plane. It is clear that for $f_{\rm p} < 10^{-6}\,{\rm Hz}$ there is barely any degeneracy. This is because the shape of the peak of the induced GW spectrum is not very sensitive to $\Delta$ for $\Delta > 0.1$. On the other hand, when $f_{\rm p} > 10^{-6}\,{\rm Hz}$ the IR tail enters the PTA window. This time, because the primordial spectrum is broad, the IR tail goes as $\Omega_{\rm GW} \propto f^{3-2|b|}$. Then, from the power law fitting $5/2 > n_{\rm PTA} > 1$ we conclude that there are two disconnected ranges for $b$, namely $-1/4 > b > -1/2$ $(5/9 < w < 1)$ and $1 > b > 1/4$ $(0 < w < 1/5)$. Because of this, the $f_{\rm p}$-$w$ plane has a half doughnut shape. Remarkably, the value of $w = 1/3$ is excluded at more than $2\sigma$ for $f_{\rm p} > 10^{-6}\,{\rm Hz}$. We also notice that the doughnut shape of the $f_{\rm p}$-$w$ plane is partly carried over to the $f_{\rm p}$-$\mathcal{A}$ plane. From Eq. (2.16), the degeneracy in the $f_{\rm p}$-$\mathcal{A}$ is now roughly given by $\mathcal{A}_{\mathcal{R}} \propto f_{\rm p}^{3/2+b-|b|}$. Since for $f_{\rm p} > 10^{-6}\,{\rm Hz}$ the allowed range for $b$ is disconnected, we can see the two degeneracy lines. Unfortunately, we also find that such an interesting regime is mostly incompatible with CMB constraints, shown in purple lines in Fig. 4.

## IV. CONCLUSION AND DISCUSSION

There is tentative evidence of a GW background at nHz frequencies reported by PTAs [1–10], which could be explained by GWs induced by primordial fluctuations. Since these GWs were generated in a period of the early universe that has not been directly probed by other means, we included the EoS $w$ of the early universe as a free parameter in our analysis. For the primordial spectrum of fluctuations, we considered a Dirac delta and a log-normal peak (with a free logarithmic width $\Delta$), inspired by several models of inflation [57–67]. We then analyzed the NANOGrav data set with PTArcade for the Bayesian analysis and SIGWfast for the numerical calculations of the induced GWs. Our results for the posterior distributions for the Dirac delta and log-normal can be respectively found in Figs. 2 and 4.

For the Dirac delta spectrum, we found that $0.13 \lesssim w \lesssim 0.87$ at $2\sigma$, which is consistent with previous analytical estimations [22] (namely $1/9 < w < 1$). Most interestingly, while standard radiation domination ($w = 1/3$) is in the middle of the $1\sigma$ contours using the GW data only, it is excluded at $2\sigma$ when including the upper bounds from no PBH overproduction. Thus, our results on induced GWs and PBHs imply $0.42 \lesssim w \lesssim 0.87$ when the universe had temperatures above 100 MeV. We also investigated the typical mass range of the PBHs, combining the PBH and CMB bounds. We found that, including the $w$ dependence, it is given by $2.5 \times 10^{-5} M_\odot \lesssim M_{\rm PBH} \lesssim 0.09 M_\odot$, which can still explain the OGLE microlensing events [87] as well for a narrow range of the EoS, namely $0.42 \lesssim w \lesssim 0.5$. Also, the upper bound of $O(0.1) M_\odot$ could explain the candidate merger events in the LVK data with sub-solar mass compact objects [125–128, 152, 153]. As this is out of the scope of this paper, we leave a detailed study for future work. Note that our conclusions slightly differ from Ref. [22] because of the different formalism adopted for estimating the PBH abundance. Nevertheless, based on peak theory other than the simplest Press-Schechter formalism, our analysis gives the most accurate estimate for the PBH bounds.

For the log-normal spectrum, we find that, in general terms, the GW data cannot constrain the EoS $w$. The main reason is that while the Dirac delta case has a precise prediction for the low frequency tail of the induced GWs that enters the PTA window, there are different possibilities for the log-normal depending on the width $\Delta$. For example, for sharp peaks ($\Delta \lesssim \mathcal{O}(0.1)$) there is the near-IR tail and the far-IR tail which respectively scale as $f^{2-2|b|}$ ($f_p \gg f \gg \Delta \times f_{\rm p}$) and $f^{3-2|b|}$ ($f \ll \Delta \times f_{\rm p}$). For broad peaks ($\Delta \gtrsim \mathcal{O}(0.1)$), the induced GW is also a broad peak, the shape of which is not very sensitive to the value of $w$. Thus, with the current precision, we cannot constrain the value of $w$. We only find a faint preference for small $\Delta$ in the NANOGrav data set and that the best-fit point fits the PTA data with the peak of the induced GW. To explore this interpretation further, namely the log-normal peak, one would need an experiment with a much larger frequency range. Or, one could also discuss specific models to remove the uncertainty in $\Delta$. Regarding the PBH counterpart, we find that while a stiffer EoS $w$ alleviates the possible PBH overproduction, a broader log-normal peak also reduces the abundance of the PBH counterpart.

In the future, we will be able to combine the likelihood from PTAs, micro-lensing observations as well as ground and space-based detectors such as $\mu$-Ares [154], binary resonances [155, 156] and the Nancy Roman telescope [157] (in the $\mu$Hz regime), LISA and Taiji [158, 159] (for mHz frequencies), TianQin [160] and DECIGO [161, 162] (for frequencies around Hz), Einstein Telescope [163], Cosmic Explorer [164] and Voyager [165, 166] (for GWs with Hz-kHz frequencies). In this way, we will have access to the induced GW spectrum in a much broader frequency range. At the same time, the PBH counterpart and the GW from the mergers of such PBH binaries [22] will allow us to do a precise Bayesian analysis to probe primordial fluctuations on small scales as well as the content of the early universe. And, perhaps, confirm (or rule out) the induced GW interpretation of the PTA data.

Our work could also be improved in several ways. First, the calculations of the PBH abundance for a finite-width peak require refinements, mostly due to the high momentum tail of the primordial scalar perturbation and the complexity of the threshold in peak theory. Second, we could include the position of the reheating frequency $f_{\rm rh}$ as a free parameter. To do so, one should improve the

kernels for $f < f_{\rm rh}$ of Refs. [117, 118]. One could also consider a gradual reheating instead of an instantaneous one. Lastly, one may consider including primordial non-Gaussianity of primordial fluctuations.

## ACKNOWLEDGMENTS

We thank Albert Escrivà for useful discussions and for providing the numerical data for the collapse threshold for a Dirac delta primordial spectrum and $w < 1/3$. This work is supported in part by the National Key Research and Development Program of China Grant No. 2021YFC2203004. G.D. is supported by the DFG under the Emmy-Noether program, project number 496592360. S.P. is supported by Project No. 12047503 of the National Natural Science Foundation of China, by the JSPS Grant-in-Aid for Early-Career Scientists No. JP20K14461, and by the World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan.

## Appendix A: The Primordial Black Hole counterpart

In this section, we will discuss the PBH produced by the scalar perturbation in the case that the scalar-induced GW can interpret the SGWB detected by PTAs. We use the peak theory to calculate the PBH formation for a monochromatic and log-normal scalar perturbation. Due to the complexity of the calculation and the uncertainty of the threshold in the finite-width case, the following discussion will only focus on interpreting the OGLE ultrashort-timescale microlensing events by the PBH counterpart of the nHz SIGW. We draw the $f_{\rm PBH} = 1$ lines for different widths in Fig.4 only with a fixed frequency $f_{\rm p} = 10^{-7}$ Hz. The PBH abundance is calculated following [151], an extension to the finite-width case of the peak theory approach used in [150]. Since the amplitude ($\sim \exp(-\tilde{\mu}^2_{2,\rm th}/(2\mathcal{A}_\mathcal{R}))$), where $\tilde{\mu}_{2,\rm th}$ is the critical peak amplitude value depending on the peak shape, affects the PBH abundance far more than the central frequency ($\sim f_{\rm p}^{\frac{6w}{1+3w}}$), it's reasonable to fix the peak frequency when calculating bound on $\mathcal{A}_\mathcal{R}$ for different $w$'s.

In peak theory, whether a PBH will form around a peak of the density fluctuation depends on whether the compaction function exceeds its threshold, which depends on the profile of the peak. We follow the definitions and calculations in Ref.[150], but generate them to a stiff EoS $w$ case. The horizon mass during that period at the horizon reentry of a comoving wavenumber $k$ ($aH/k = 1$)

$$M(k) = M_{\rm rh}\left(\frac{k}{k_{\rm rh}}\right)^{-\frac{3(1+w)}{1+3w}}, \tag{A1}$$

$$M_{\rm rh} = 5.33\times 10^{35}\left(\frac{g_*}{10.75}\right)^{-1/6}\left(\frac{T_{\rm rh}}{20{\rm MeV}}\right)^{-2}{\rm g}, \tag{A2}$$

where $M_{\rm rh}$ stands for the horizon mass at the end of reheating. When the physical radius $R_m = ae^{\hat{\zeta}(r_m)}r_m$ reenter the horizon, $ae^{\hat{\zeta}(r_m)}r_m H = 1$ , the horizon mass

$$M_H \equiv M(a/R_m) = M(k_{\rm p})\left(x_m e^{\hat{\zeta}_m(\tilde{\mu}_2)}\right)^{\frac{3(1+w)}{1+3w}}, \tag{A3}$$

where $x_m = r_m k_{\rm p}$.

The relation between the PBH mass and the horizon mass follows a scaling relation around the threshold

$$M\left(\mu_2, \tilde{k}_3, w\right) = \mathcal{K}\left(\tilde{k}_3, w\right)\left(\tilde{\mu}_2 - \tilde{\mu}_{2,\rm th}\left(\tilde{k}_3, w\right)\right)^{\gamma(w)} M_H\left(\tilde{\mu}_2, \tilde{k}_3, w\right) \tag{A4}$$

where $\tilde{\mu}_2$ and $\tilde{k}_3$ describe the height and the width of the peak. The probability distribution of them is given by the power spectrum of the random field, and $\tilde{k}_3$ is restricted to 1 in monochromatic case [150]. $\gamma(w)$ can be expressed semi-analytically, given in [167, 168], which has been confirmed numerically in [85, 169]. $\mathcal{K}$ has a $O(1)$ value which depends on shape as well as $w$ and we choose $\mathcal{K} = 1$ in our calculations. We only need the threshold for different EoS parameters $w$. For $w > 1/3$, the threshold almost only depends upon the compaction function

$$\hat{\mathcal{C}}(r) = \frac{2}{3}\left[1 - \left(1 + r\hat{\zeta}'\right)^2\right] \tag{A5}$$

around its peak value, which can be expressed by an analytical formula up to the second order derivative of the compaction function [85, 170]. The numerical results are used for the monochromatic case when $w < 1/3$. Thus, the mass function of PBH can be written as

$$f_{\rm PBH}(M) = \left|\frac{d\ln M}{d\mu_2}\right|^{-1} n_{\rm pk}(\tilde{\mu}_2(M))\ M \propto {k_{\rm p}}^{\frac{6w}{1+3w}}, \tag{A6}$$

where

$$n_{\rm pk}(\tilde{\mu}_2) = \left(\frac{3}{2\pi}\right)^{3/2} k_{\rm p}^3 f(\tilde{\mu}_2/\sqrt{\mathcal{A}_\mathcal{R}})\frac{1}{\sqrt{2\pi\mathcal{A}_\mathcal{R}}}\exp\left(-\frac{\tilde{\mu}_2^2}{2\mathcal{A}_\mathcal{R}}\right), \tag{A7}$$

$$\left|\frac{d\ln M}{d\mu_2}\right| = \left|\frac{3(1+w)}{1+3w}\hat{\zeta}'_m(\mu_2) + \frac{\gamma(w)}{\mu_2 - \mu_{2,\rm th}}\right|, \tag{A8}$$

and $\hat{\zeta}'(\mu_2) \equiv d\hat{\zeta}(x_m)/d\tilde{\mu_2}$. The $f$ function is the number density of peaks and $f(x) \sim x^3$ at high tails when $\tilde{\mu}_2/\sqrt{\mathcal{A}_\mathcal{R}} \gg 1$.[7] The total mass factor can be given by integrating over Eq.(A6),

$$f_{\rm tot} = \int f_{\rm PBH}(M) d\ln M. \tag{A10}$$

The PBH abundance is mainly determined by the $\exp(-\tilde{\mu}_2^2/(2\mathcal{A}_\mathcal{R}))$ in Eq. (A6) when $\tilde{\mu}_{2,\rm th} \gg \sqrt{\mathcal{A}_\mathcal{R}}$, which indicates the PBH bound almost has $\mathcal{A}_\mathcal{R}(f_{\rm PBH} = 1) \propto \tilde{\mu}_{2,\rm th}^2$ with a slight dependence on $f_{\rm p}$ from the $k_{\rm p}^{\frac{6w}{(1+3w)}}$ term in Eq.(A6). So we directly fix $f_{\rm p} = 10^{-7}$ Hz and plot the corresponding upper bound for the amplitude of the scalar perturbation for different EoM $w$ in Dirac-delta peak case in

[7] The concrete form of $f$ in Eq. (A7) is given by

$$\begin{aligned} f(\xi) =& \frac{1}{2}\xi\left(\xi^2 - 3\right)\left(\mathrm{erf}\left[\frac{1}{2}\sqrt{\frac{5}{2}}\xi\right] + \mathrm{erf}\left[\sqrt{\frac{5}{2}}\xi\right]\right) \\ &+ \sqrt{\frac{2}{5\pi}}\left\{\left(\frac{8}{5} + \frac{31}{4}\xi^2\right)\exp\left[-\frac{5}{8}\xi^2\right] + \left(-\frac{8}{5} + \frac{1}{2}\xi^2\right)\exp\left[-\frac{5}{2}\xi^2\right]\right\} \end{aligned} \tag{A9}$$

Fig.2. We use the extension in our unpublished work [151] to give a conservative estimation for the upper bound for the peak with different widths, shown in Fig.4. Note that the PBH constraints are with respect to $\mathcal{A}_{\mathcal{R}}$, but as we pointed out earlier, this at most contributes a factor $\sim 2$.

After obtaining the mass function by Eq. (A6), we can check whether the planet-mass PBH can explain the ultrashort-timescale microlensing events observed by OGLE, firstly proposed in [96]. In [87], the PBH mass and abundance needed to explain these events are given at 95% C.L. for the monochromatic mass function, which is about 1% to 10% of the dark matter at around $10^{-4}$ to $10^{-6}\ M_\odot$. Reducing the PBH abundance from $f_{\rm PBH} = 1$ to $f_{\rm PBH} = 0.01$ only reduces the amplitude of the perturbation very slightly, so we directly substitute the upper bound amplitude for $f_{\rm PBH} = 1$ to calculate peak mass for different parameters, which is displayed in Fig.3. It is a zoom-in of the $f_{\rm p}$-$w$ subfigure in Fig.2. We can see that there is still an allowed region in the parameter space, which means that the OGLE ultrashort-timescale microlensing events and the nHz SGWB can be simultaneously explained.

## Appendix B: Constraints from CMB Anisotropy

The GW background, acting as an additional radiation component, may affect Big Bang Nucleosynthesis (BBN) and the anisotropy of the CMB. The additional energy density in the form of radiation is conventionally parameterized by $N_{\rm eff}$, the effective number of neutrino species after electron-positron annihilation. One may translate the effective number of relativistic species $\Delta N_{\rm eff}$ to an upper bound on the integrated energy density of GW background, namely [171]

$$\Omega_{\rm GW,0}h^2 \leq 5.6\times10^{-6}\left(\frac{\Omega_{r,0}h^2}{2.47\times10^{-5}}\right)\left(\frac{g_S\left(T_0\right)}{3.91}\right)^{4/3}\left(\frac{g_S\left(T_0\right)}{3.91}\right)^{-4/3}\Delta N_{\rm eff}. \tag{B1}$$

Since BBN and CMB yield similar results, we choose the constraints from the CMB, $\Delta N_{\rm eff} < 0.30$ (i.e $(h^2\rho_{\rm GW}/\rho_c)_0 \leq 1.7\times10^{-6}$) at 95 % [147], as our conservative estimation.

We compute the total energy density from the induced GWs as follows. First, we parametrize the primordial spectrum as $\mathcal{P}_{\mathcal{R},k_{\rm p}}(k) = \mathcal{A}_{\mathcal{R}}F(k/k_{\rm p})$ (which could be the Dirac delta (2.7) or the log-normal peak (2.1)). Then, we integrate the induced GW spectrum over all frequencies as

$$\Omega^{\rm tot}_{\rm GW,rh} = \mathcal{A}^2\left(\frac{f_{\rm p}}{f_{\rm rh,*}}\right)^{-2b}\int_0^\infty d\ln\frac{f}{f_{\rm p}}\,Q_w\left(\frac{f}{f_{\rm p}}\right), \tag{B2}$$

where we defined

$$Q_w(x) = x^{-2b}\int_0^1 {\rm d}d\int_1^\infty {\rm d}s\,\mathcal{T}_w(d,s)\,F\left(x(s+d)\right)F\left(x(s-d)\right). \tag{B3}$$

In the Dirac delta case, Eq. (B2) has four free parameters: $f_{\rm p}$, $w$ and $\mathcal{A}$. In the log-normal case, we have an additional parameter $\Delta$. Then, for any given $w$ and $\Delta$, we find the relation $\mathcal{A}(f_{\rm p})$ such that the induced GW spectrum goes through the geometric mean of the energy spectral density with the highest probability in different frequency bins, called $\Omega_{\rm GW,PTA}$. This provides a fixed point in the $(f,\Omega_{\rm GW})$ domain. Further assuming that $f_{\rm p} \gg f_{\rm PTA}$, *i.e.* the peak frequency is much

larger than the PTA frequencies, we may use the IR tail of the induced GW spectrum to find the relation $\mathcal{A}(f_p)$. Explicitly, we substitute the IR limit of Eq. (3.1) into Eq. (2.6),

$$\Omega_{\rm GW,PTA} = \Omega_{\rm GW,rh}(f \ll f_{\rm p})\Big|_{f=f_{\rm PTA}} \approx \mathcal{A}_{\mathcal{R}}^2 \begin{cases} C^{\delta}(b) \dfrac{f_{\rm p}}{f_{\rm PTA}}^{2(1-|b|+b)} & \text{(Dirac-Delta)} \\ C^{\Delta}(b) \dfrac{f_{\rm p}}{f_{\rm PTA}}^{2(3/2-|b|+b)} & \text{(Log-normal)} \end{cases} , \tag{B4}$$

where the coefficients $C^{\delta}(b)$ and $C^{\Delta}(b)$ can be calculated numerically (although one could also use the analytical formulas given in (2.9) and (2.16)). Eq. (B4) should be understood as a relation $\mathcal{A}(f_{\rm p})$ when fixing ($f_{\rm PTA}$, $\Omega_{\rm GW,PTA}$, $w$, $\Delta$). With such relations, we compute Eq. (B2) and impose the CMB bound (B1). The results are shown in purple lines Figs. 2 and 4.


[1] J. Antoniadis et al. (EPTA, InPTA:), Astron. Astrophys. **678**, A50 (2023), arXiv:2306.16214 [astro-ph.HE].
[2] J. Antoniadis et al. (EPTA), Astron. Astrophys. **678**, A48 (2023), arXiv:2306.16224 [astro-ph.HE].
[3] J. Antoniadis et al. (EPTA), (2023), arXiv:2306.16227 [astro-ph.CO].
[4] A. Zic et al., (2023), arXiv:2306.16230 [astro-ph.HE].
[5] D. J. Reardon et al., Astrophys. J. Lett. **951**, L6 (2023), arXiv:2306.16215 [astro-ph.HE].
[6] D. J. Reardon et al., Astrophys. J. Lett. **951**, L7 (2023), arXiv:2306.16229 [astro-ph.HE].
[7] G. Agazie et al. (NANOGrav), Astrophys. J. Lett. **951**, L9 (2023), arXiv:2306.16217 [astro-ph.HE].
[8] G. Agazie et al. (NANOGrav), Astrophys. J. Lett. **951**, L8 (2023), arXiv:2306.16213 [astro-ph.HE].
[9] G. Agazie et al. (International Pulsar Timing Array), (2023), arXiv:2309.00693 [astro-ph.HE].
[10] H. Xu et al., Res. Astron. Astrophys. **23**, 075024 (2023), arXiv:2306.16216 [astro-ph.HE].
[11] G. Agazie et al. (NANOGrav), Astrophys. J. Lett. **952**, L37 (2023), arXiv:2306.16220 [astro-ph.HE].
[12] H.-L. Huang, Y. Cai, J.-Q. Jiang, J. Zhang, and Y.-S. Piao, (2023), arXiv:2306.17577 [gr-qc].
[13] Y. Gouttenoire, S. Trifinopoulos, G. Valogiannis, and M. Vanvlasselaer, (2023), arXiv:2307.01457 [astro-ph.CO].
[14] P. F. Depta, K. Schmidt-Hoberg, P. Schwaller, and C. Tasillo, (2023), arXiv:2306.17836 [astro-ph.CO].
[15] A. Afzal et al. (NANOGrav), Astrophys. J. Lett. **951**, L11 (2023), arXiv:2306.16219 [astro-ph.HE].
[16] M. V. Sazhin, Vestn. Mosk. Univ. Fiz. Astron. **18**, 82 (1977).
[17] C. Caprini and D. G. Figueroa, Class. Quant. Grav. **35**, 163001 (2018), arXiv:1801.04268 [astro-ph.CO].
[18] R. Roshan and G. White, (2024), arXiv:2401.04388 [hep-ph].
[19] V. Dandoy, V. Domcke, and F. Rompineve, SciPost Phys. Core **6**, 060 (2023), arXiv:2302.07901 [astro-ph.CO].
[20] G. Franciolini, A. Iovino, Junior., V. Vaskonen, and H. Veermae, (2023), arXiv:2306.17149 [astro-ph.CO].
[21] G. Franciolini, D. Racco, and F. Rompineve, (2023), arXiv:2306.17136 [astro-ph.CO].
[22] K. Inomata, K. Kohri, and T. Terada, (2023), arXiv:2306.17834 [astro-ph.CO].
[23] Y.-F. Cai, X.-C. He, X. Ma, S.-F. Yan, and G.-W. Yuan, (2023), arXiv:2306.17822 [gr-qc].
[24] S. Wang, Z.-C. Zhao, J.-P. Li, and Q.-H. Zhu, (2023), arXiv:2307.00572 [astro-ph.CO].
[25] L. Liu, Z.-C. Chen, and Q.-G. Huang, (2023), arXiv:2307.01102 [astro-ph.CO].
[26] C. Unal, A. Papageorgiou, and I. Obata, (2023), arXiv:2307.02322 [astro-ph.CO].

[27] D. G. Figueroa, M. Pieroni, A. Ricciardone, and P. Simakachorn, (2023), arXiv:2307.02399 [astro-ph.CO].
[28] Z. Yi, Q. Gao, Y. Gong, Y. Wang, and F. Zhang, (2023), arXiv:2307.02467 [gr-qc].
[29] Q.-H. Zhu, Z.-C. Zhao, and S. Wang, (2023), arXiv:2307.03095 [astro-ph.CO].
[30] H. Firouzjahi and A. Talebian, JCAP **10**, 032 (2023), arXiv:2307.03164 [gr-qc].
[31] J.-P. Li, S. Wang, Z.-C. Zhao, and K. Kohri, JCAP **10**, 056 (2023), arXiv:2305.19950 [astro-ph.CO].
[32] Z.-Q. You, Z. Yi, and Y. Wu, JCAP **11**, 065 (2023), arXiv:2307.04419 [gr-qc].
[33] S. Balaji, G. Domènech, and G. Franciolini, (2023), arXiv:2307.08552 [gr-qc].
[34] S. A. Hosseini Mansoori, F. Felegray, A. Talebian, and M. Sami, JCAP **08**, 067 (2023), arXiv:2307.06757 [astro-ph.CO].
[35] Z.-C. Zhao, Q.-H. Zhu, S. Wang, and X. Zhang, (2023), arXiv:2307.13574 [astro-ph.CO].
[36] L. Liu, Z.-C. Chen, and Q.-G. Huang, JCAP **11**, 071 (2023), arXiv:2307.14911 [astro-ph.CO].
[37] Z. Yi, Z.-Q. You, and Y. Wu, (2023), arXiv:2308.05632 [astro-ph.CO].
[38] N. Bhaumik, R. K. Jain, and M. Lewicki, (2023), arXiv:2308.07912 [astro-ph.CO].
[39] S. Choudhury, A. Karde, S. Panda, and M. Sami, (2023), arXiv:2308.09273 [astro-ph.CO].
[40] Z. Yi, Z.-Q. You, Y. Wu, Z.-C. Chen, and L. Liu, (2023), arXiv:2308.14688 [astro-ph.CO].
[41] K. Harigaya, K. Inomata, and T. Terada, (2023), arXiv:2309.00228 [astro-ph.CO].
[42] S. Basilakos, D. V. Nanopoulos, T. Papanikolaou, E. N. Saridakis, and C. Tzerefos, (2023), arXiv:2307.08601 [hep-th].
[43] J.-H. Jin, Z.-C. Chen, Z. Yi, Z.-Q. You, L. Liu, and Y. Wu, JCAP **09**, 016 (2023), arXiv:2307.08687 [astro-ph.CO].
[44] E. Cannizzaro, G. Franciolini, and P. Pani, (2023), arXiv:2307.11665 [gr-qc].
[45] Z. Zhang, C. Cai, Y.-H. Su, S. Wang, Z.-H. Yu, and H.-H. Zhang, Phys. Rev. D **108**, 095037 (2023), arXiv:2307.11495 [hep-ph].
[46] L. Liu, Y. Wu, and Z.-C. Chen, (2023), arXiv:2310.16500 [astro-ph.CO].
[47] S. Choudhury, K. Dey, A. Karde, S. Panda, and M. Sami, (2023), arXiv:2310.11034 [astro-ph.CO].
[48] M. Tagliazucchi, M. Braglia, F. Finelli, and M. Pieroni, (2023), arXiv:2310.08527 [astro-ph.CO].
[49] S. Basilakos, D. V. Nanopoulos, T. Papanikolaou, E. N. Saridakis, and C. Tzerefos, (2023), arXiv:2309.15820 [astro-ph.CO].
[50] K. Inomata, M. Kawasaki, K. Mukaida, and T. T. Yanagida, (2023), arXiv:2309.11398 [astro-ph.CO].
[51] J.-P. Li, S. Wang, Z.-C. Zhao, and K. Kohri, (2023), arXiv:2309.07792 [astro-ph.CO].
[52] G. Domènech, G. Vargas, and T. Vargas, (2023), arXiv:2309.05750 [astro-ph.CO].
[53] M. R. Gangopadhyay, V. V. Godithi, K. Ichiki, R. Inui, T. Kajino, A. Manusankar, G. J. Mathews, and Yogesh, (2023), arXiv:2309.03101 [astro-ph.CO].
[54] B. Cyr, T. Kite, J. Chluba, J. C. Hill, D. Jeong, S. K. Acharya, B. Bolliet, and S. P. Patil, (2023), arXiv:2309.02366 [astro-ph.CO].
[55] K. D. Lozanov, S. Pi, M. Sasaki, V. Takhistov, and A. Wang, (2023), arXiv:2310.03594 [astro-ph.CO].
[56] Y. Akrami et al. (Planck), Astron. Astrophys. **641**, A10 (2020), arXiv:1807.06211 [astro-ph.CO].
[57] W. H. Kinney, Phys. Rev. D **72**, 023515 (2005), arXiv:gr-qc/0503017.
[58] J. Martin, H. Motohashi, and T. Suyama, Phys. Rev. D **87**, 023514 (2013), arXiv:1211.0083 [astro-ph.CO].
[59] S.-L. Cheng, W. Lee, and K.-W. Ng, Phys. Rev. D **99**, 063524 (2019), arXiv:1811.10108 [astro-ph.CO].
[60] C. T. Byrnes, P. S. Cole, and S. P. Patil, JCAP **06**, 028 (2019), arXiv:1811.11158 [astro-ph.CO].
[61] S. Pi and J. Wang, JCAP **06**, 018 (2023), arXiv:2209.14183 [astro-ph.CO].
[62] S.-L. Cheng, D.-S. Lee, and K.-W. Ng, (2023), arXiv:2305.16810 [astro-ph.CO].
[63] Y.-F. Cai, X. Tong, D.-G. Wang, and S.-F. Yan, Phys. Rev. Lett. **121**, 081306 (2018), arXiv:1805.03639

[astro-ph.CO].
[64] C. Chen and Y.-F. Cai, JCAP **10**, 068 (2019), arXiv:1908.03942 [astro-ph.CO].
[65] M. Kawasaki, N. Sugiyama, and T. Yanagida, Phys. Rev. D **57**, 6050 (1998), arXiv:hep-ph/9710259.
[66] M. Kawasaki, N. Kitajima, and T. T. Yanagida, Phys. Rev. D **87**, 063519 (2013), arXiv:1207.2550 [hep-ph].
[67] S. Pi, Y.-l. Zhang, Q.-G. Huang, and M. Sasaki, JCAP **05**, 042 (2018), arXiv:1712.09896 [astro-ph.CO].
[68] S. Pi and M. Sasaki, Phys. Rev. D **108**, L101301 (2023), arXiv:2112.12680 [astro-ph.CO].
[69] G. Ferrante, G. Franciolini, A. Iovino, Junior., and A. Urbano, Phys. Rev. D **107**, 043520 (2023), arXiv:2211.01728 [astro-ph.CO].
[70] D.-S. Meng, C. Yuan, and Q.-G. Huang, Sci. China Phys. Mech. Astron. **66**, 280411 (2023), arXiv:2212.03577 [astro-ph.CO].
[71] C. Chen, A. Ghoshal, Z. Lalak, Y. Luo, and A. Naskar, JCAP **08**, 041 (2023), arXiv:2305.12325 [astro-ph.CO].
[72] O. Özsoy and G. Tasinato, Universe **9**, 203 (2023), arXiv:2301.03600 [astro-ph.CO].
[73] I. D. Zel'dovich, Ya.B.; Novikov, Soviet Astron. AJ (Engl. Transl. ), **10**, 602 (1967).
[74] S. Hawking, Mon. Not. Roy. Astron. Soc. **152**, 75 (1971).
[75] S. W. Hawking, Nature **248**, 30 (1974).
[76] B. J. Carr and S. Hawking, Mon. Not. Roy. Astron. Soc. **168**, 399 (1974).
[77] P. Meszaros, Astron. Astrophys. **37**, 225 (1974).
[78] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975), [Erratum: Commun.Math.Phys. 46, 206 (1976)].
[79] B. J. Carr, Astrophys. J. **201**, 1 (1975).
[80] M. Khlopov, B. Malomed, and I. Zeldovich, Mon. Not. Roy. Astron. Soc. **215**, 575 (1985).
[81] M. Y. Khlopov, Res. Astron. Astrophys. **10**, 495 (2010), arXiv:0801.0116 [astro-ph].
[82] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, Class. Quant. Grav. **35**, 063001 (2018), arXiv:1801.05235 [astro-ph.CO].
[83] B. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, Rept. Prog. Phys. **84**, 116902 (2021), arXiv:2002.12778 [astro-ph.CO].
[84] A. M. Green and B. J. Kavanagh, J. Phys. G **48**, 4 (2021), arXiv:2007.10722 [astro-ph.CO].
[85] A. Escrivà, F. Kuhnel, and Y. Tada, (2022), arXiv:2211.05767 [astro-ph.CO].
[86] P. Mroz et al., Nature **548**, 183 (2017), arXiv:1707.07634 [astro-ph.EP].
[87] H. Niikura, M. Takada, S. Yokoyama, T. Sumi, and S. Masaki, Phys. Rev. D **99**, 083503 (2019), arXiv:1901.07120 [astro-ph.CO].
[88] R.-g. Cai, S. Pi, and M. Sasaki, Phys. Rev. Lett. **122**, 201101 (2019), arXiv:1810.11000 [astro-ph.CO].
[89] C. Unal, Phys. Rev. D **99**, 041301 (2019), arXiv:1811.09151 [astro-ph.CO].
[90] V. Atal and G. Domènech, JCAP **06**, 001 (2021), arXiv:2103.01056 [astro-ph.CO].
[91] P. Adshead, K. D. Lozanov, and Z. J. Weiner, JCAP **10**, 080 (2021), arXiv:2105.01659 [astro-ph.CO].
[92] K. T. Abe, R. Inui, Y. Tada, and S. Yokoyama, JCAP **05**, 044 (2023), arXiv:2209.13891 [astro-ph.CO].
[93] C. Yuan, D.-S. Meng, and Q.-G. Huang, (2023), arXiv:2308.07155 [astro-ph.CO].
[94] S. Pi and M. Sasaki, Phys. Rev. Lett. **131**, 011002 (2023), arXiv:2211.13932 [astro-ph.CO].
[95] H. Firouzjahi and A. Riotto, Phys. Rev. D **108**, 123504 (2023), arXiv:2309.10536 [astro-ph.CO].
[96] G. Domènech and S. Pi, Sci. China Phys. Mech. Astron. **65**, 230411 (2022), arXiv:2010.03976 [astro-ph.CO].
[97] S. Choudhury, K. Dey, and A. Karde, (2023), arXiv:2311.15065 [astro-ph.CO].
[98] M. Kawasaki, K. Kohri, and N. Sugiyama, Phys. Rev. Lett. **82**, 4168 (1999), arXiv:astro-ph/9811437.
[99] M. Kawasaki, K. Kohri, and N. Sugiyama, Phys. Rev. D **62**, 023506 (2000), arXiv:astro-ph/0002127.
[100] S. Hannestad, Phys. Rev. D **70**, 043506 (2004), arXiv:astro-ph/0403291.

[101] T. Hasegawa, N. Hiroshima, K. Kohri, R. S. L. Hansen, T. Tram, and S. Hannestad, JCAP **12**, 012 (2019), arXiv:1908.10189 [hep-ph].

[102] R. Allahverdi et al., Open J. Astrophys. **4** (2021), 10.21105/astro.2006.16182, arXiv:2006.16182 [astro-ph.CO].

[103] B. Spokoiny, Phys. Lett. B **315**, 40 (1993), arXiv:gr-qc/9306008.

[104] P. J. E. Peebles and A. Vilenkin, Phys. Rev. D **59**, 063505 (1999), arXiv:astro-ph/9810509.

[105] P. Brax and J. Martin, Phys. Rev. D **71**, 063530 (2005), arXiv:astro-ph/0502069.

[106] M. W. Hossain, R. Myrzakulov, M. Sami, and E. N. Saridakis, Phys. Rev. D **90**, 023512 (2014), arXiv:1402.6661 [gr-qc].

[107] H. Assadullahi and D. Wands, Phys. Rev. D **79**, 083511 (2009), arXiv:0901.0989 [astro-ph.CO].

[108] K. Inomata, K. Kohri, T. Nakama, and T. Terada, JCAP **10**, 071 (2019), arXiv:1904.12878 [astro-ph.CO].

[109] K. Inomata, K. Kohri, T. Nakama, and T. Terada, Phys. Rev. D **100**, 043532 (2019), arXiv:1904.12879 [astro-ph.CO].

[110] K. Inomata, M. Kawasaki, K. Mukaida, T. Terada, and T. T. Yanagida, Phys. Rev. D **101**, 123533 (2020), arXiv:2003.10455 [astro-ph.CO].

[111] T. Papanikolaou, V. Vennin, and D. Langlois, JCAP **03**, 053 (2021), arXiv:2010.11573 [astro-ph.CO].

[112] G. Domènech, C. Lin, and M. Sasaki, JCAP **04**, 062 (2021), [Erratum: JCAP 11, E01 (2021)], arXiv:2012.08151 [gr-qc].

[113] G. Domènech, V. Takhistov, and M. Sasaki, Phys. Lett. B **823**, 136722 (2021), arXiv:2105.06816 [astro-ph.CO].

[114] I. Dalianis and C. Kouvaris, JCAP **07**, 046 (2021), arXiv:2012.09255 [astro-ph.CO].

[115] F. Hajkarim and J. Schaffner-Bielich, Phys. Rev. D **101**, 043522 (2020), arXiv:1910.12357 [hep-ph].

[116] S. Bhattacharya, S. Mohanty, and P. Parashari, Phys. Rev. D **102**, 043522 (2020), arXiv:1912.01653 [astro-ph.CO].

[117] G. Domènech, Int. J. Mod. Phys. D **29**, 2050028 (2020), arXiv:1912.05583 [gr-qc].

[118] G. Domènech, S. Pi, and M. Sasaki, JCAP **08**, 017 (2020), arXiv:2005.12314 [gr-qc].

[119] I. Dalianis and K. Kritos, Phys. Rev. D **103**, 023505 (2021), arXiv:2007.07915 [astro-ph.CO].

[120] K. T. Abe, Y. Tada, and I. Ueda, JCAP **06**, 048 (2021), arXiv:2010.06193 [astro-ph.CO].

[121] L. T. Witkowski, G. Domènech, J. Fumagalli, and S. Renaux-Petel, JCAP **05**, 028 (2022), arXiv:2110.09480 [astro-ph.CO].

[122] S. Balaji, G. Domenech, and J. Silk, JCAP **09**, 016 (2022), arXiv:2205.01696 [astro-ph.CO].

[123] K. D. Lozanov, M. Sasaki, and V. Takhistov, (2023), arXiv:2304.06709 [astro-ph.CO].

[124] K. D. Lozanov, M. Sasaki, and V. Takhistov, Phys. Lett. B **848**, 138392 (2024), arXiv:2309.14193 [astro-ph.CO].

[125] S. Clesse and J. Garcia-Bellido, Phys. Dark Univ. **38**, 101111 (2022), arXiv:2007.06481 [astro-ph.CO].

[126] G. Morras et al., Phys. Dark Univ. **42**, 101285 (2023), arXiv:2301.11619 [gr-qc].

[127] R. Abbott et al. (LIGO Scientific, VIRGO, KAGRA), Mon. Not. Roy. Astron. Soc. **524**, 5984 (2023), [Erratum: Mon.Not.Roy.Astron.Soc. 526, 6234 (2023)], arXiv:2212.01477 [astro-ph.HE].

[128] M. Prunier, G. Morrás, J. F. N. n. Siles, S. Clesse, J. García-Bellido, and E. Ruiz Morales, (2023), arXiv:2311.16085 [gr-qc].

[129] K. Tomita, Progress of Theoretical Physics **37**, 831 (1967), https://academic.oup.com/ptp/article-pdf/37/5/831/5234391/37-5-831.pdf.

[130] S. Matarrese, O. Pantano, and D. Saez, Phys. Rev. D **47**, 1311 (1993).

[131] S. Matarrese, O. Pantano, and D. Saez, Phys. Rev. Lett. **72**, 320 (1994), arXiv:astro-ph/9310036.

[132] K. N. Ananda, C. Clarkson, and D. Wands, Phys. Rev. D **75**, 123518 (2007), arXiv:gr-qc/0612013.

[133] D. Baumann, P. J. Steinhardt, K. Takahashi, and K. Ichiki, Phys. Rev. D **76**, 084019 (2007), arXiv:hep-th/0703290.

[134] R. Saito and J. Yokoyama, Phys. Rev. Lett. **102**, 161101 (2009), [Erratum: Phys.Rev.Lett. 107, 069901 (2011)], arXiv:0812.4339 [astro-ph].

[135] R. Saito and J. Yokoyama, Prog. Theor. Phys. **123**, 867 (2010), [Erratum: Prog.Theor.Phys. 126, 351–352 (2011)], arXiv:0912.5317 [astro-ph.CO].

[136] E. Bugaev and P. Klimai, Phys. Rev. D **81**, 023517 (2010), arXiv:0908.0664 [astro-ph.CO].

[137] E. Bugaev and P. Klimai, Phys. Rev. D **83**, 083521 (2011), arXiv:1012.4697 [astro-ph.CO].

[138] G. Domènech, Universe **7**, 398 (2021), arXiv:2109.01398 [gr-qc].

[139] G. Domènech, (2023), arXiv:2311.02065 [gr-qc].

[140] L. T. Witkowski, (2022), arXiv:2209.05296 [astro-ph.CO].

[141] S. Pi and M. Sasaki, JCAP **09**, 037 (2020), arXiv:2005.12306 [gr-qc].

[142] C. Altavista and J. Rey, (2023), arXiv:2309.14993 [astro-ph.CO].

[143] K. Saikawa and S. Shirai, JCAP **05**, 035 (2018), arXiv:1803.01038 [hep-ph].

[144] J. Espinosa, D. Racco, and A. Riotto, Journal of Cosmology and Astroparticle Physics **2018**, 012 (2018).

[145] K. Kohri and T. Terada, Phys. Rev. D **97**, 123532 (2018), arXiv:1804.08577 [gr-qc].

[146] K. Inomata and T. Nakama, Phys. Rev. D **99**, 043511 (2019), arXiv:1812.00674 [astro-ph.CO].

[147] N. Aghanim et al. (Planck), Astron. Astrophys. **641**, A6 (2020), [Erratum: Astron.Astrophys. 652, C4 (2021)], arXiv:1807.06209 [astro-ph.CO].

[148] R.-G. Cai, S. Pi, and M. Sasaki, Phys. Rev. D **102**, 083528 (2020), arXiv:1909.13728 [astro-ph.CO].

[149] W. G. Lamb, S. R. Taylor, and R. van Haasteren, "The need for speed: Rapid refitting techniques for bayesian spectral characterization of the gravitational wave background using ptas," (2023), arXiv:2303.15442 [astro-ph.HE].

[150] N. Kitajima, Y. Tada, S. Yokoyama, and C.-M. Yoo, JCAP **10**, 053 (2021), arXiv:2109.00791 [astro-ph.CO].

[151] S. Pi, M. Sasaki, V. Takhistov, and J. Wang, Eprint = "2305.12325", archivePrefix = "arXiv", primaryClass = "astro-ph.CO", doi = "10.1088/1475-7516/2023/08/041", journal = "JCAP", volume = "08", pages = "041", year = "2023".

[152] B. Carr, S. Clesse, J. Garcia-Bellido, M. Hawkins, and F. Kuhnel, (2023), arXiv:2306.03903 [astro-ph.CO].

[153] A. H. Nitz and Y.-F. Wang, Phys. Rev. D **106**, 023024 (2022), arXiv:2202.11024 [astro-ph.HE].

[154] A. Sesana et al., Exper. Astron. **51**, 1333 (2021), arXiv:1908.11391 [astro-ph.IM].

[155] D. Blas and A. C. Jenkins, Phys. Rev. D **105**, 064021 (2022), arXiv:2107.04063 [gr-qc].

[156] D. Blas and A. C. Jenkins, Phys. Rev. Lett. **128**, 101103 (2022), arXiv:2107.04601 [astro-ph.CO].

[157] Y. Wang, K. Pardo, T.-C. Chang, and O. Doré, Phys. Rev. D **106**, 084006 (2022), arXiv:2205.07962 [gr-qc].

[158] S. Barke, Y. Wang, J. J. Esteban Delgado, M. Tröbs, G. Heinzel, and K. Danzmann, Class. Quant. Grav. **32**, 095004 (2015), arXiv:1411.1260 [physics.ins-det].

[159] W.-H. Ruan, Z.-K. Guo, R.-G. Cai, and Y.-Z. Zhang, Int. J. Mod. Phys. A **35**, 2050075 (2020), arXiv:1807.09495 [gr-qc].

[160] Y. Gong, J. Luo, and B. Wang, Nature Astron. **5**, 881 (2021), arXiv:2109.07442 [astro-ph.IM].

[161] K. Yagi and N. Seto, Phys. Rev. **D83**, 044011 (2011), [Erratum: Phys. Rev.D95,no.10,109901(2017)], arXiv:1101.3940 [astro-ph.CO].

[162] S. Kawamura et al., (2020), arXiv:2006.13545 [gr-qc].

[163] M. Maggiore et al., JCAP **03**, 050 (2020), arXiv:1912.02622 [astro-ph.CO].

[164] "Cosmic explorer sensitivity curve," https://cosmicexplorer.org/sensitivity.html, [Online; accessed 05-May-2023].
[165] "The A+ design curve," https://dcc.ligo.org/LIGO-T1800042/public, [Online; accessed 05-May-2023].
[166] "Ligo unofficial sensitivity curves," https://dcc.ligo.org/LIGO-T1500293/public, [Online; accessed 05-May-2023].
[167] D. Maison, Phys. Lett. B **366**, 82 (1996), arXiv:gr-qc/9504008.
[168] T. Koike, T. Hara, and S. Adachi, Phys. Rev. D **59**, 104008 (1999).
[169] I. Musco and J. C. Miller, Class. Quant. Grav. **30**, 145009 (2013), arXiv:1201.2379 [gr-qc].
[170] A. Escrivà, C. Germani, and R. K. Sheth, JCAP **01**, 030 (2021), arXiv:2007.05564 [gr-qc].
[171] P. Binetruy, A. Bohe, C. Caprini, and J.-F. Dufaux, JCAP **06**, 027 (2012), arXiv:1201.0983 [gr-qc].

arXiv:2402.18912v1 [nucl-th] 29 Feb 2024

# Hadron-quark phase transition in the neutron star with vector MIT bag model and Korea-IBS-Daegu-SKKU functional

Debashree Sen,[1] Hana Gil,[1] and Chang Ho Hyun[2,*]

[1] *Center of Extreme Nuclear Matters, Korea University, Seoul 02841, Korea*

[2] *Department of Physics Education, Daegu University, Gyeonsan 38453, Korea*

(Dated: March 1, 2024)

## Abstract

Employing the Korea-IBS-Daegu-SKKU (KIDS) density functional for the hadron phase and the MIT bag model with vector (vBag) model for the quark phase, we obtain hadron-quark phase transition in neutron stars considering Maxwell construction. The structural properties of the resultant hybrid stars are computed for three different values of bag constant ($B$) in the range $B^{1/4}$ =(145−160 MeV). We studied the effects of symmetry energy ($J$) on the hybrid star properties with the different KIDS model and found that $J$ has important influence not only on the transition properties like the transition mass, transition radius and jump in density due to phase transition, but also on the stability of the hybrid stars. The vector repulsion of the quark phase via the parameter $G_V$ has profound influence in obtaining reasonable hybrid star configurations, consistent with the recent astrophysical constraints on the structural properties of compact stars. Within the aforesaid range of $B$, the value of $G_V$ is constrained to be $0.3 \lesssim G_V \lesssim 0.4$ in order to obtain reasonable hybrid star configurations.

* hch@daegu.ac.kr

## I. INTRODUCTION

Mass-radius ($M-R$) diagram of the neutron star is unprecedentedly crowded with modern astronomical data that covers the neutron star mass from about $0.5M_\odot$ to more than $2.0M_\odot$. They consist of GW170817 by the LIGO/Vigro Collaboration [1], PSR J0030+0451 by NICER [2, 3] and PSR J0740+6620 [4–6] and HESS J1731-347 [7]. Simultaneous measurement of the mass and radius for a single object helps greatly deepen our understanding of the state of matter at densities above the nuclear saturation.

A naive but quite certain scenario is the phase transition from hadronic matter to quark matter. Initially the matter is composed of nucleons at low densities, where they are far apart. Since the nucleon has a finite size, they start to overlap to each other at a density higher than the saturation. As the density evolves furthermore, confinement of quarks in the initial nucleon becomes uncertain, and finally quarks are completely released from the nucleons and form a deconfined quark matter. The key points to the hadron-quark phase transition are the physical properties of the nucleon and quarks, and their interactions at finite densities. The main concern of the present work is to examine the effect of the uncertainties in the interactions of quarks and nucleons to the equation of state (EoS) and the phase transition in the core of neutron stars, forming hybrid stars. For the calculation of the EoS, we use the MIT bag model with vector repulsion (vBag) for the quark matter, and the Korea-IBS-Daegu-SKKU (KIDS) density functional for the hadron phase.

In the vBag model, stiffness of the quark matter is controlled by the bag constant and the vector repulsion [8–11]. The two quantities have different physical origins and implications. Bag constant is introduced in the original bag model. In the simplest description of the baryon bags, quarks are treated as free particles within a bag. In order to confine the quarks within the baryon, inward pressure is assumed to be exerted on the quarks. The inward pressure is described in terms of the bag constant, so it is a natural way to determine the bag constant from the baryon masses in free space. In this work, as a lower limit of the bag constant, we adopt $B^{1/4} = 145$ MeV from Ref. [12]. As far as the baryon spectrum in free space is concerned, there is no upper limit to the bag constant. However, if we adopt the Bodmer-Witten conjecture which states that the quark matter composed of $u$, $d$ and $s$ quarks is the most stable state of matter in the deconfined quark phase, an upper limit value can be determined. We use $B^{1/4} = 160$ MeV, which is determined by applying the

Bodmer-Witten conjecture to the original bag model. In this work, uncertainty arising from the bag constant is accounted by considering three values, the lower limit 145 MeV, a middle value 155 MeV and the upper limit 160 MeV.

While the bag constant is inherent in the bag model, there is no repulsive force in the original bag model. The reason might be that it is sufficient to produce the baryon spectrum accurately with the kinetic energy of quarks and the bag constant. With these two ingredients, however, it is hard to satisfy the observation of large mass stars, $M \geq 2M_{\odot}$ because the EoS is too soft to sustain the strong gravity of super dense matter [8]. Request for more stiff EoS demands a repulsive interaction in terms of the vector field. Nowadays, it is common to include vector repulsion in the widely used quark models such as Nambu-Jona-Lasino model as well as the bag model. An issue related to the vector force is how strong the repulsion is. Thanks to the improvement in the precision and diversity of the neutron star observational data, it becomes feasible to constrain the strength of vector repulsion more precisely. Combined with the uncertainties of the bag constant and the symmetry energy, we can propose values of the vector coupling constant that are compatible with the observed neutron star properties.

In the description of neutron-rich hadronic matter, the largest uncertainty comes from the symmetry energy. Its value at the saturation density is relatively accurately determined, most conservatively in the range $30-34$ MeV, but its density dependence is much more uncertain. The average density of a neutron star whose mass is larger than the canonical mass is well above the saturation density, so the neutron star provides a dependable way to determine the density dependence of the symmetry energy. In recent works, symmetry energy has been constrained by using the neutron star data [13–15], and the uncertainty could be reduced substantially. Nevertheless, the density dependence still has sizable ambiguity, and it can have critical impact to the transition to deconfined quark matter. In order to investigate the influence of the symmetry energy, we adopt the KIDS-A, B, C, D models [14], which are determined to satisfy the nuclear properties, neutron star data and gravitational wave measurements simultaneously, and have evidently different stiffness of the symmetry energy.

In the calculation of the EoS, we treat the bag constant as a free parameter, and determine the range of vector coupling constants to satisfy the data of GW170817, NICER and HESS simultaneously for the given models of the symmetry energy. In the result, we find that the

critical density at which the phase transition occurs is sensitive the bag constant, vector coupling constant and the symmetry energy. However, maximum mass depends on the bag constant and the symmetry energy weakly, so we can obtain appreciably model-independent range of the vector coupling constant.

We organize the work in the following order. In Sect. II, models for the quark and hadronic matters are described. Section III shows the results and we present the discussions on them. The work is summarized in Sect. IV.

## II. MODEL

### A. MIT bag with vector repulsion

We consider quark matter with $u$, $d$ and $s$ quarks and the electrons as a lepton. The Lagrangian for the bag model with vector repulsion is given by

$$\begin{aligned}\mathcal{L} &= \sum_{f=u,d,s} \left[\bar{\psi}_f\{\gamma^\mu(i\partial_\mu - g_{qqV}V_\mu) - m_f\} - B\right]\Theta(\bar{\psi}_f\psi_f) \\ &+ \frac{1}{2}m_V^2 V_\mu V^\mu - \frac{1}{4}V_{\mu\nu}V^{\mu\nu} + \overline{\psi}_e(i\gamma_\mu\partial^\mu - m_e)\psi_e, \end{aligned} \tag{1}$$

where $V_\mu$ denotes the vector field, $g_{qqV}$ is the quark-vector meson coupling constant, $B$ is the bag constant, $V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu$ and the Heaviside function $\Theta$=1 inside the bag. For the vector coupling constant, we assume $g_{uuV} = g_{ddV}$ and $g_{ssV} = 0.4\ g_{uuV}$. It is shown that the variation on the ratio $g_{ssV}/g_{uuV}$ does not affect the EoS as much as the bag constant or the vector coupling constant [16]. Since the vector meson is an isoscalar particle, it corresponds to the $\omega$ meson, so we use the $\omega$ meson mass 783 MeV for $m_V$. In the quark matter EoS, vector coupling constant contributes in the form $(g_{qqV}/m_V)^2$, so we use $G_V = (g_{uuV}/m_V)^2$ as a parameter to be constrained from the neutron star data. The bag constant is considered in the range $B^{1/4}$ = 145–160 MeV, so three values 145, 155 and 160 MeV are used in the calculation.

### B. KIDS functional

In the KIDS density functional framework, energy of a nucleon in nuclear matter is expanded in powers of $\rho^{1/3}$ where $\rho$ is the matter density as

$$\mathcal{E} = \mathcal{T} + \sum_{i=0} \rho^{1+i/3}(\alpha_i + \beta_i \delta^2). \tag{2}$$

$\mathcal{T}$ is the kinetic energy, and $\delta = (\rho_n - \rho_p)/\rho$ where $\rho_n$ and $\rho_p$ are the neutron and the proton density, respectively. Coefficients $\alpha_i$ are determined from the properties of symmetric nuclear matter, and $\beta_i$ are constrained by the neutron star data in the KIDS-A, B, C, D models. Incompressibility of the symmetric matter $K_0$ and the symmetry energy parameters are summarized in Tab. I.

TABLE I: Incompressibility $K_0$ of the symmetric nuclear matter and the symmetry energy parameters $J$, $L$ and $K_{\rm sym}$ in units of MeV for the KIDS-A, B, C, D models.

| | KIDS-A | KIDS-B | KIDS-C | KIDS-D |
|---|---|---|---|---|
| $K_0$ | 230 | 240 | 250 | 260 |
| $J$ | 33 | 32 | 31 | 30 |
| $L$ | 66 | 58 | 58 | 47 |
| $K_{\rm sym}$ | −139.5 | −162.1 | −91.5 | −134.5 |

In the present work, phase transition is achieved via Maxwell construction assuming that the surface tension at the hadron-quark interface is sufficiently large [17] and the resultant hybrid EoS is employed to obtain the mass $M$ and the radius $R$ of hybrid stars using the Tolman-Oppenheimer-Volkoff (TOV) equations [18, 19] while dimensionless tidal deformability $\Lambda$ is then calculated following [20, 21]. For the outer crust region, we adopt the Baym-Pethick-Sutherland EoS [22] and for the inner crust, we have considered the EoS including the pasta phases [23].

## III. RESULT

We first examine the dependence of symmetry energy on the properties of hadronic stars with the four different KIDS models with different symmetry energy as tabulated in Tab. I.

For the hadronic model, as the symmetry energy ($J$) decreases and compression modulus $K_0$ increases from KIDS-A to KIDS-D in Tab. I, the EoS stiffens and the hence the maximum mass increases. For example, the maximum mass for KIDS-A model ($J$=33 MeV, $K_0$=230 MeV) is $2.10M_\odot$ while it is $2.26M_\odot$ for the KIDS-D model ($J$=30 MeV, $K_0$=260 MeV). From Fig. 1 it can be seen that the hadronic star configurations with all the four KIDS model satisfy the present day astrophysical constraints on the mass-radius relation of compact stars viz., the maximum mass constraint from PSR J0740+6620 [4] with corresponding radius constraint [5, 6], the constraints from GW170817 [1], NICER experiment for PSR J0030+0451 [2, 3] and HESS J1731-347 [7].

In order to study the influence of symmetry energy on the phase transition and the hybrid star structure, we consider the vBag model with $B^{1/4}$=155 MeV. We find that phase transition is quite early for low values of $G_V$ but the crossover shifts abruptly to high chemical potential and pressure (density) at $G_V$=0.4. This is consistent with the result obtained by Lopes *et al.* [8]. As a result we observe in Fig. 1 that for any particular hadronic model, the maximum mass and corresponding radius of the hybrid stars increase with increasing repulsion via $G_V$. The value of $R_{1.4}$ and the radius of low mass hybrid stars are greatly affected by $G_V$. Fig. 1 also confirms that all the present day astrophysical constraints are satisfied by the hybrid star configurations with all the four KIDS model with the chosen values of $G_V$ except for 0.25, for which the NICER data for PSR J0030+0451 are not satisfied. Also for any particular hadronic model, the recently obtained HESS J1731-347 data is better satisfied with lower repulsion i.e., lower values of $G_V$. Both transition mass ($M_t$) and transition radius ($R_t$) increase with increasing values of $G_V$, and for $G_V$=0.4 there is an abrupt jump in the value of $M_t$ for all the hadronic models as a consequence of the delayed transition. Such delayed transition leads to unstable hybrid star configurations obtained with both KIDS-A and KIDS-B hadronic models as seen from Figs. 1a and 1b.

We show the dependence of transition properties like $M_t$, $R_t$ and the energy difference ($\Delta\varepsilon = \varepsilon_t^Q - \varepsilon_t^H$) or the jump in density from hadronic to quark phase as functions of symmetry energy $J$ in Figs. 2 and 3. For any particular value of $G_V$, the relationships (fitted) between $M_t$ and $R_t$ with $J$ show that symmetry energy has a considerable influence on the transition properties of hybrid stars. From Fig. 2 we find that for increasing (decreasing) values of $J$ ($K_0$) the transition mass decreases considerably (Fig. 2a) while the transition radius follows the reverse trend (Fig. 2b). The relation between $\Delta\varepsilon$ and $J$ in Fig. 3 shows that $\Delta\varepsilon$ also

increases with increasing $J$ or decreasing $K_0$ for any particular value of $G_V$. The dependence of $\Delta\varepsilon$ on $J$ is almost linear and the slope is almost independent of the value of $G_V$.

We also study the variation of the dimensionless tidal deformability with the canonical mass of hadronic and hybrid stars in Fig. 4, where it is seen that the values of $\Lambda_{1.4}$ for all the hadronic and hybrid star configurations are in good agreement with the constraint from GW170817 [1]. However, dependence on the $G_V$ value is evident, so $\Lambda_{1.4} \simeq 260$ with $G_V = 0.25$ but $\Lambda_{1.4} \simeq 380$ with $G_V = 0.35$. Accurate measurement of the tidal deformability will certain help determine $G_V$ value less ambiguously.

We next consider a lower value of $B$ as $B^{1/4}$ =145 MeV. We have already seen from Fig. 1 that for $B^{1/4}$ =155 MeV, none of the hybrid star configurations obtained with $G_V$=0.25 and the four KIDS model could satisfy the NICER data for PSR J0030+0451 due to early transition. Therefore, it can be expected that further low value of $B$ will lead to further early or no transition at all. For $B^{1/4}$ =145 MeV we do not consider $G_V$=0.25. We, however, consider $G_V$=0.4 to check if early transition for lower value of $B$ can give rise to stable hybrid stars. For the purpose we consider the softest (KIDS-A) and the stiffest (KIDS-D) among the four KIDS models. As expected, phase transition is early in the case of $B^{1/4}$ =145 MeV. This is also reflected in Fig. 5a and 5b where it can be seen that the transition mass is lower than that in case of $B^{1/4}$ =155 MeV for a particular value of $G_V$. For example, considering the KIDS-A model ($J$=33 MeV) and $G_V$=0.35, $M_t$=0.422 $M_\odot$ when $B^{1/4}$ =155 MeV as seen from Figs. 2a and 1a while $M_t$=0.389 $M_\odot$ when $B^{1/4}$ =145 MeV as seen from Fig. 5a. Considering KIDS-D model ($J$=30 MeV) and $G_V$=0.35, $M_t$=0.539 $M_\odot$ as seen from Figs. 2a and 1d for $B^{1/4}$ =155 MeV while $M_t$=0.397 $M_\odot$ in case of $B^{1/4}$ =145 MeV as seen from Fig. 5b. Simultaneously, the transition radius is also affected by $B$. For example, in case of KIDS-A and $G_V$=0.35, $R_t$=13.081 km for $B^{1/4}$ =155 MeV as seen from Figs. 2b and 1a while $R_t$=13.255 km in case of $B^{1/4}$ =145 MeV as seen from Fig. 5a. Considering the KIDS-D model and $G_V$=0.35, $R_t$=11.924 km as seen from Figs. 2b and 1d for $B^{1/4}$ =155 MeV while $R_t$=12.171 km in case of $B^{1/4}$ =145 MeV as seen from Fig. 5b. Therefore transition mass is lowered while transition radius is increased for lower value of $B$. Similar to the case of higher value of $B^{1/4}$ =155 MeV, the hybrid star configurations obtained for $G_V$=0.4 with $B^{1/4}$ =145 MeV is unstable for the KIDS-A model as seen from Figs. 1a and 5a. It is, however, stable with both the values of $B$ and $G_V$=0.4 in case of KIDS-D model having value of $J$ lower than that of the KIDS-A model. So, it can be said

that symmetry energy also affects the stability of hybrid stars and EoS with lower value of symmetry energy is more suitable for obtaining stable hybrid star configurations with large quark repulsion $G_V > 0.35$. We also find that for $B^{1/4}$ =145 MeV, all the present day astrophysical constraints are satisfied with $G_V$=0.3, 0.35 and 0.4 in terms of mass, radius (Figs. 5a and 5b) and tidal deformability (Figs. 5c and 5d). Another notable result is, similar to the result of $B^{1/4} = 155$ MeV, there is a huge gap in the transition mass and transition density between $G_V = 0.35$ and $G_V = 0.4$. The transition mass is close to the maximum mass for $G_V = 0.4$, which means that even though the trainstion to quark matter happens in the core, its effect is not significant to the bulk properties of the neutron star.

We finally consider a higher value of $B$ as $B^{1/4}$ =160 MeV which we consider as the upper limit of $B$ to obtain hybrid star configurations with the KIDS-A and KIDS-D hadronic models. With this choice of $B$, it is expected that for $G_V$=0.4, transition will be further delayed and since we already obtained unstable hybrid star configurations with lower values of $B$, we do not consider $G_V$=0.4 in case of the maximum value of $B$. However, with low values of $G_V$ in case of $B^{1/4}$ =160 MeV, transition density may be suitable to achieve reasonable hybrid star configurations. Therefore we also consider $G_V$=0.25 along with 0.3 and 0.35 in this case. The results are displayed in Fig. 6. We find from Figs. 6a and 6b that even low quark repulsion ($G_V$=0.25) can satisfactorily satisfy all the astrophysical and observational constraints including the NICER data for PSR J0030+0451 which was not satisfied with the lower values of bag constant. Considering $G_V$=0.35 and $B^{1/4}$ =160 MeV, $M_t = 0.798 M_\odot$ and $0.963 M_\odot$ while $R_t$=12.479 km and 11.905 km for the KIDS-A and KIDS-D models, respectively. It can thus be concluded that for a particular value of $G_V$ and $J$, $M_t$ increases with $B$ while $R_t$ follows the reverse trend. Also, for a particular value of $G_V$ and $B$, $M_t(R_t)$ decreases(increases) with a larger $J$. Summarizing the results up to now, depending on the value of the bag constant and the symmetry energy, the hybrid star configurations for $G_V$ values in the range $0.25 \lesssim G_V \lesssim 0.4$ are consistent with the observational data.

In Tab. II we display the transition densities and the maximum mass of the various hybrid star configurations. At the saturation density $\rho_0 = 0.16$ fm$^{-3}$, the volume occupied by a nucleon is $V = 1/\rho_0 = 6.25$ fm$^3$. Assuming that the nucleon is a sphere located at the center of a cube whose volume is 6.25 fm$^3$, the distance between the center of two nucleons is 1.84 fm. The radius of a nucleon is about 0.8 fm, so there is no overlapping between the

nucleons at the saturation density. When the distance between the center of two nucleons is 1.6 fm, the surface of a nucleon touches the surface of other nucleons. The volume of a cube with the length of one side 1.6 fm is 4.10 fm$^3$, and it gives a density $\rho = 0.244$ fm$^{-3} = 1.53\rho_0$. Taking into account the swelling of nucleons in nuclear medium, onset of the overlapping could happen at densities smaller than $1.5\rho_0$, but it must be larger than the saturation density. Considering this mechanical condition for the phase transition, $G_V = 0.25$ is ruled out in any combinations of $B$ and the symmetry energy. Even $G_V = 0.3$ is not acceptable for some cases. Therefore, combining the consistency with the astronomical data and the mechanical condition for the phase transition, acceptable range of $G_V$ is $0.3 \lesssim G_V \lesssim 0.4$. Since $G_V = 0.4$ has minor effects to the properties of the neutron star, it could be regarded as an upper limit of $G_V$ value if one assumes the phase transition in the neutron star.

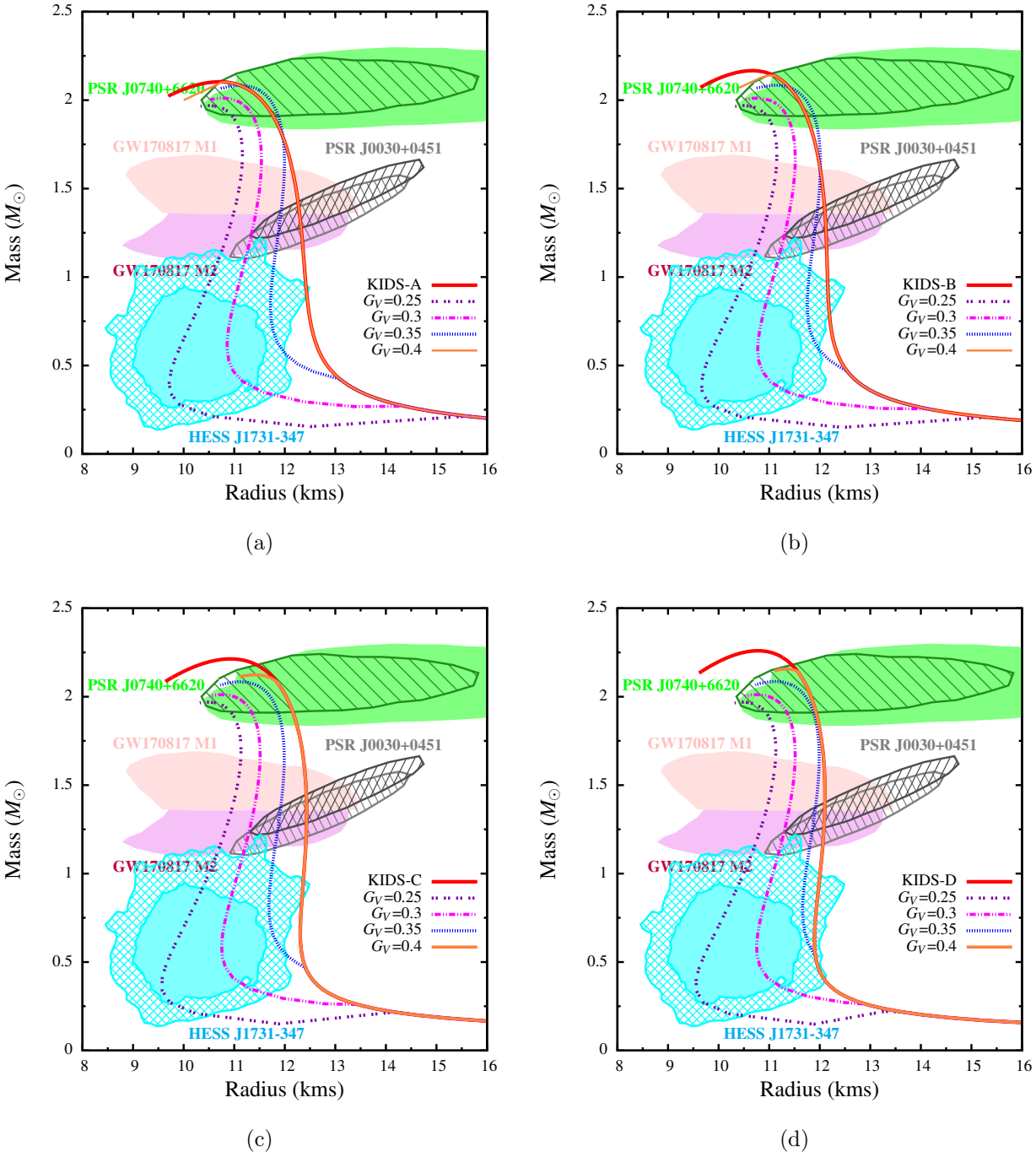
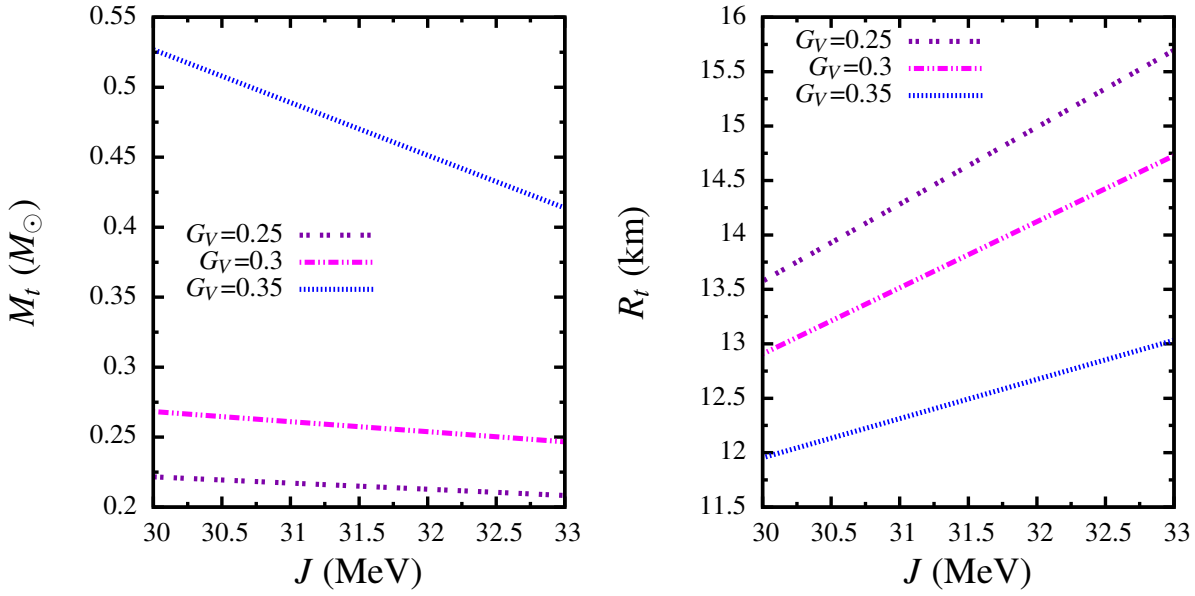



FIG. 1: *Variation of mass with radius of hadronic star (solid lines) and hybrid stars (dashed lines) with hadronic models (a) KIDS-A, (b) KIDS-B, (c) KIDS-C and (d) KIDS-D for different values of $G_V$ and $B^{1/4}$ =155 MeV. Observational limits imposed from the most massive pulsar PSR J0740+6620 on maximum mass [4] and corresponding radius [5, 6] are also indicated. The constraints on $M-R$ plane prescribed from GW170817 [1], NICER experiment for PSR J0030+0451 [2, 3] and HESS J1731-347 [7] are also compared.*




(a) (b)

FIG. 2: *Variation of (a) transition mass and (b) transition radius of hybrid stars with respect to symmetry energy for different values of* $G_V$ *with* $X_V$ *=0.4 and* $B^{1/4}$ *=155 MeV.*

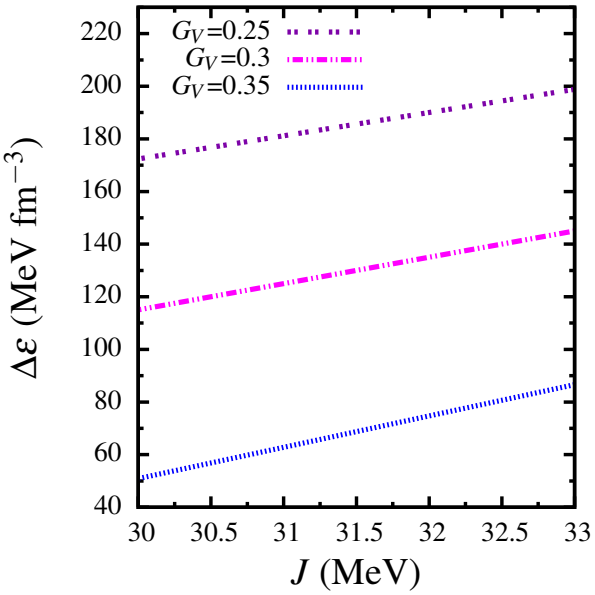



FIG. 3: *Variation of jump in energy density of hybrid stars with respect to symmetry energy for different values of* $G_V$ *with* $X_V$ *=0.4 and* $B^{1/4}$ *=155 MeV.*

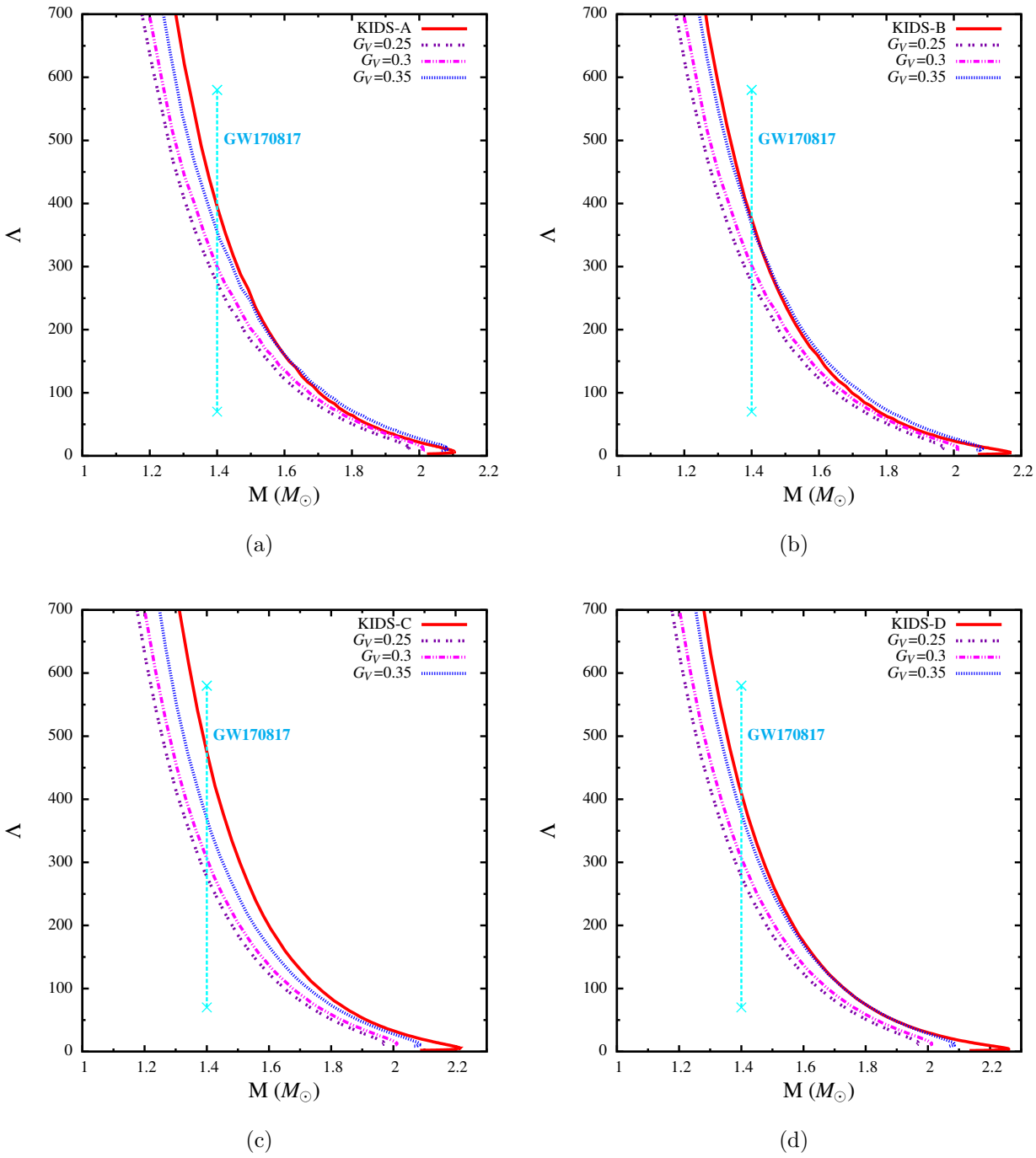



FIG. 4: *Variation of tidal deformability with mass of hadronic star (solid lines) and hybrid stars (dashed lines) with hadronic models (a) KIDS-A, (b) KIDS-B, (c) KIDS-C and (d) KIDS-D for different values of* $G_V$ *and* $B^{1/4}$ *=155 MeV.*

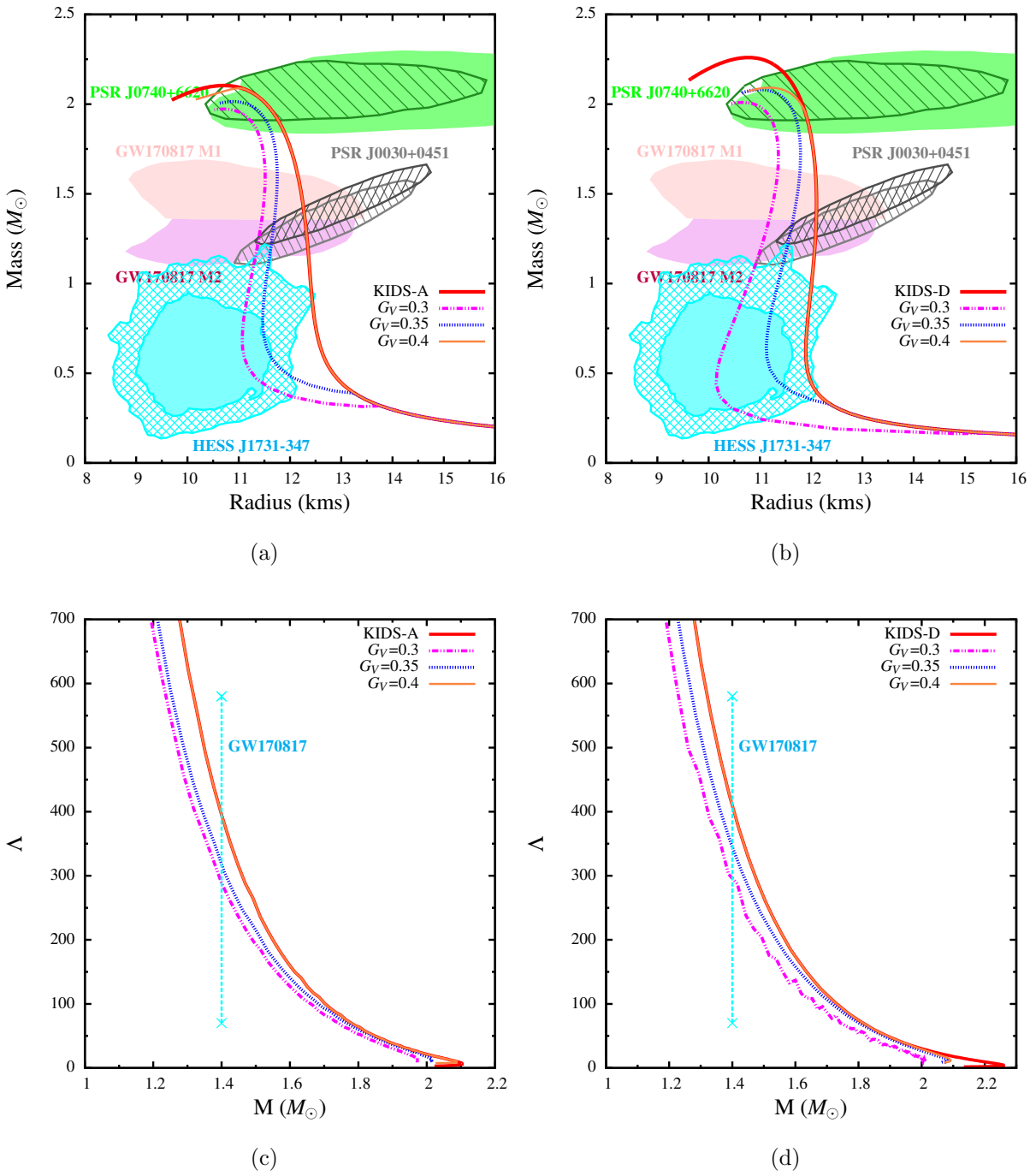



FIG. 5: *Variation of mass with radius (a, b) and tidal deformability with mass (c, d) of hadronic star (solid lines) and hybrid stars (dashed lines) with hadronic models KIDS-A and KIDS-D for different values of $G_V$ and $B^{1/4}$ =145 MeV.*

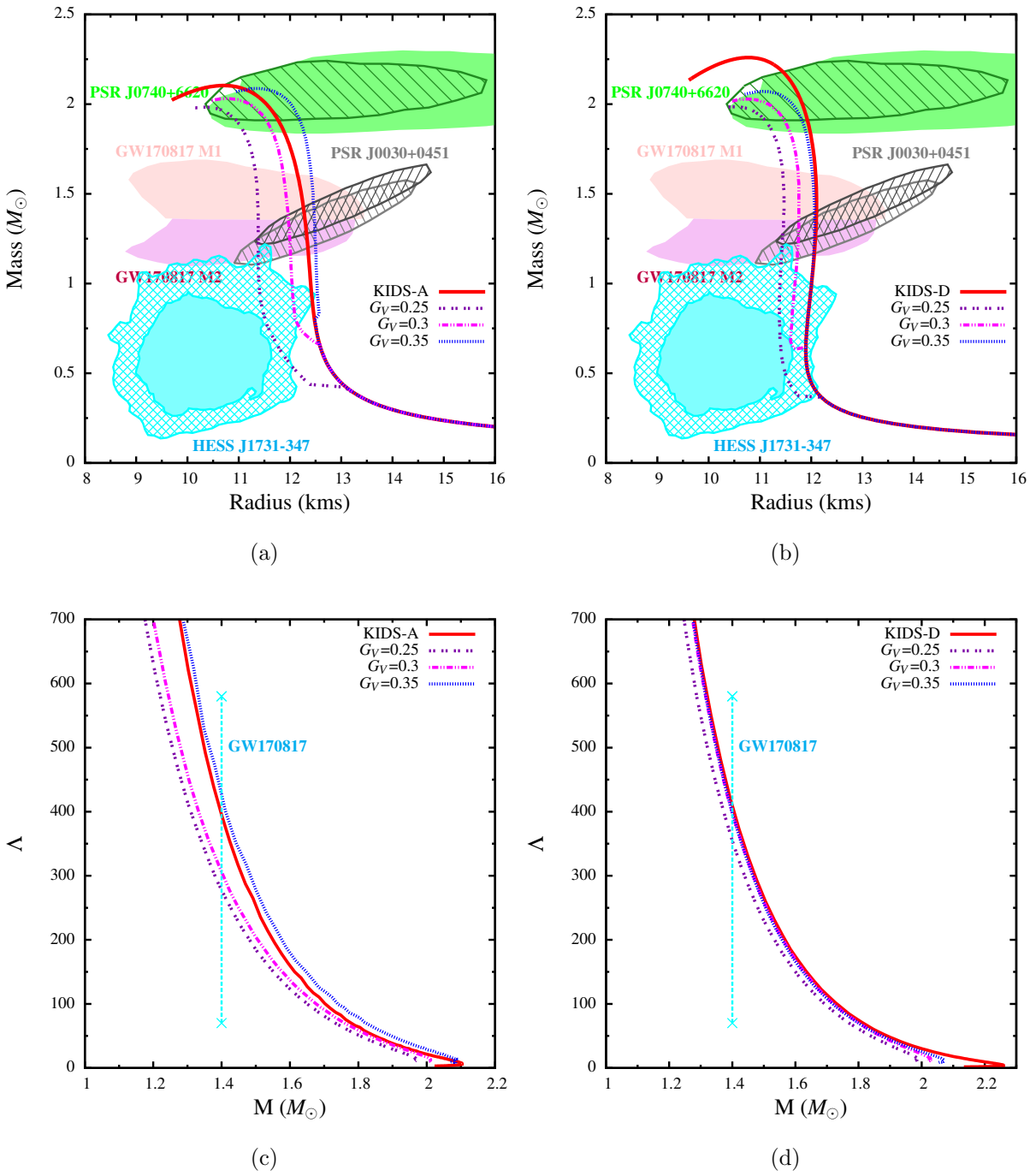



FIG. 6: *Variation of mass with radius (a, b) and tidal deformability with mass (c, d) of hadronic star (solid lines) and hybrid stars (dashed lines) with hadronic models KIDS-A and KIDS-D for different values of* $G_V$ *and* $B^{1/4}$ *=160 MeV.*

TABLE II: Comparison of hadron-quark transition densities, and maximum mass for different $G_V$ with different values of $B$.

| $B^{1/4}$ (MeV) | Hadronic Model | $G_V$ | $\rho_t^H/\rho_0$ | $\rho_t^Q/\rho_0$ | $M_{max}$ ($M_\odot$) |
|---|---|---|---|---|---|
| 145 | KIDS-A | 0.3 | 0.821 | 1.875 | 1.972 |
| | | 0.35 | 1.273 | 1.963 | 2.014 |
| | | 0.4 | 5.800 | 6.238 | 2.088 (unstable) |
| | KIDS-D | 0.3 | 1.006 | 1.900 | 2.009 |
| | | 0.35 | 1.533 | 1.983 | 2.081 |
| | | 0.4 | 3.845 | 4.638 | 2.092 |
| 155 | KIDS-A | 0.25 | 0.737 | 1.955 | 1.969 |
| | | 0.3 | 1.093 | 2.218 | 2.011 |
| | | 0.35 | 1.444 | 2.468 | 2.083 |
| | | 0.4 | 6.500 | 7.262 | 2.103 (unstable) |
| | KIDS-D | 0.25 | 0.606 | 1.937 | 1.969 |
| | | 0.3 | 1.219 | 2.401 | 2.011 |
| | | 0.35 | 1.756 | 2.897 | 2.084 |
| | | 0.4 | 4.563 | 5.137 | 2.156 |
| 160 | KIDS-A | 0.25 | 0.794 | 1.973 | 1.983 |
| | | 0.3 | 1.146 | 2.344 | 2.031 |
| | | 0.35 | 1.500 | 2.508 | 2.087 |
| | KIDS-D | 0.25 | 0.671 | 1.997 | 1.985 |
| | | 0.3 | 1.269 | 2.088 | 2.028 |
| | | 0.35 | 1.803 | 2.179 | 2.087 |

## IV. SUMMARY

The present work is dedicated to study hadron-quark phase transition and the structural properties of hybrid stars. They depend on the properties of both the hadronic and quark phases like the symmetry energy of the former and the bag constant and the strength of quark repulsion of the latter. In order to investigate the effects of symmetry energy on the hybrid star properties we consider four KIDS models with different symmetry energy while the vBag model is adopted for the quark phase to study the influence of quark repulsion via the different values of vector coupling $G_V$ at three individual values of bag constant $B$.

We found that for any particular value of $G_V$ and $J$, higher bag constant leads to delayed transition in terms of both density and mass. Therefore, the maximum mass $M_{max}$ of hybrid stars also increases with increasing value of $B$. Moreover, the transition mass $M_t$ increases with $B$ while $R_t$ follows the reverse trend.

For any particular value of $B$ and $J$, $M_{max}$ increases with increasing value of $G_V$. The value of $R_{1.4}$ is also greatly affected by $G_V$. The radius of low mass hybrid stars are seen to increase with increasing values of $G_V$. Both $M_t$ and $R_t$ are found to increase with $G_V$.

For fixed values of $G_V$ and $B$, $M_{max}$ increases with decreasing value of $J$. $M_t(R_t)$ decreases(increases) with increasing $J$. The jump in energy density $\Delta\varepsilon$ from hadronic to quark phase also increases with increasing value of $J$ for any particular value of $G_V$. The dependence (fitted) of $\Delta\varepsilon$ on $J$ is linear and the slope is almost independent of the value of $G_V$. Considering a reasonable and moderately wide range of $B^{1/4}$=(145−160) MeV and symmetry energy $J$=(30−33) MeV, we find that the most suitable value of $G_V$ is approximately $0.3 \lesssim G_V \lesssim 0.4$ for the hybrid stars to satisfy all the present day astrophysical constraints on the structural properties of compact stars. $G_V \leq 0.25$ is allowed only when the bag constant is as high as $B^{1/4} > 155$ MeV in the light of the NICER data for PSR J0030+0451. The stability of hybrid stars configurations with large quark repulsion $G_V > 0.35$ is also affected by symmetry energy. We find that lower values of symmetry energy ($J < 32$ MeV) are more suitable for obtaining stable hybrid star configurations with large quark repulsion $G_V > 0.35$. Accurate measurement of the tidal deformability can make a significant contribution to reducing the uncertainty of $G_V$ in a narrower range.

**ACKNOWLEDGEMENTS**

This work was supported by the National Research Foundation of Korea (Grant Nos. 2018R1A5A1025563 and 2023R1A2C1003177).

---

[1] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **121**, 161101 (2018), arXiv:1805.11581 [gr-qc].

[2] T. E. Riley *et al.*, Astrophys. J. Lett. **887**, L21 (2019), arXiv:1912.05702 [astro-ph.HE].

[3] M. C. Miller *et al.*, Astrophys. J. Lett. **887**, L24 (2019), arXiv:1912.05705 [astro-ph.HE].

[4] E. Fonseca *et al.*, Astrophys. J. Lett. *915*, L12 (2021), arXiv:2104.00880 [astro-ph.HE].

[5] M. C. Miller *et al.*, Astrophys. J. Lett. **918**, L28 (2021), arXiv:2105.06979 [astro-ph.HE].

[6] T. E. Riley *et al.*, Astrophys. J. Lett. **918**, L27 (2021), arXiv:2105.06980 [astro-ph.HE].

[7] V. Doroshenko, V. Suleimanov, G. Puhlhofer, and A. Santangelo, Nature Astronomy **6**, 1444 (2022).

[8] L. L. Lopes, C. Biesdorf, and D. P. Menezes, Mon. Not. Roy. Astron. Soc. **512**, 5110 (2022), arXiv:2111.13732 [hep-ph].

[9] A. Kumar, V. B. Thapa, and M. Sinha, Mon. Not. Roy. Astron. Soc. **513**, 3788 (2022), arXiv:2204.11034 [astro-ph.HE].

[10] P. Laskos-Patkos, P. S. Koliogiannis, and C. C. Moustakidis, (2023), arXiv:2312.07113 [astro-ph.HE].

[11] A. Kumar, V. B. Thapa, and M. Sinha, Phys. Rev. D **107**, 063024 (2023), arXiv:2303.06387 [astro-ph.HE].

[12] K. Johnson, Acta Phys. Polon. B **6**, 865 (1975).

[13] H. Gil, Y.-M. Kim, P. Papakonstantinou, and C. H. Hyun, Phys. Rev. C **103**, 034330 (2021), arXiv:2010.13354 [nucl-th].

[14] H. Gil and C. H. Hyun, New Phys. Sae Mulli **71**, 242 (2021), arXiv:2012.00930 [nucl-th].

[15] H. Gil, P. Papakonstantinou, and C. H. Hyun, Int. J. Mod. Phys. E **31**, 2250013 (2022), arXiv:2110.09802 [nucl-th].

[16] S. Pal, S. Podder, D. Sen, and G. Chaudhuri, Phys. Rev. D **107**, 063019 (2023), arXiv:2303.04653 [nucl-th].

[17] T. Maruyama, S. Chiba, H.-J. Schulze, and T. Tatsumi, Phys. Lett. B **659**, 192 (2008), arXiv:nucl-th/0702088.

[18] R. C. Tolman, Phys. Rev. **55**, 364 (1939).

[19] J. R. Oppenheimer and G. M. Volkoff, Phys. Rev. **55**, 374 (1939).

[20] T. Hinderer, Astrophys. J. **677**, 1216 (2008), arXiv:0711.2420 [astro-ph].

[21] T. Hinderer, B. D. Lackey, R. N. Lang, and J. S. Read, Phys. Rev. D **81**, 123016 (2010), arXiv:0911.3535 [astro-ph.HE].

[22] G. Baym, C. Pethick, and P. Sutherland, Astrophys. J. **170**, 299 (1971).

[23] F. Grill, H. Pais, C. Providência, I. Vidaña, and S. S. Avancini, Phys. Rev. C **90**, 045803 (2014), arXiv:1404.2753 [nucl-th].

# Weak Lensing Constraints on Dark Matter-Baryon Interactions with $N$-Body Simulations and Machine Learning

Chi Zhang,[1,2] Lei Zu,[1,*] Hou-Zun Chen,[3] Yue-Lin Sming Tsai,[1,2,†] and Yi-Zhong Fan[1,2]

[1] *Key Laboratory of Dark Matter and Space Astronomy, Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, China*

[2] *School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, China*

[3] *Institute for Astronomy, the School of Physics, Zhejiang University, Hangzhou 310027, China*

We investigate the elastic scattering cross section between dark matter and protons using the DES Year 3 weak lensing data. This scattering induces a dark acoustic oscillation structure in the matter power spectra. To address non-linear effects at low redshift, we utilize principal component analysis alongside a limited set of $N$-body simulations, improving the reliability of our matter power spectrum prediction. We further perform a robust Markov Chain Monte Carlo analysis to derive the upper bounds on the DM-proton elastic scattering cross-section, assuming different velocity dependencies. Our results, presented as the first Frequentist upper limits, are compared with the ones obtained by Bayesian approach. Compared with the upper limits derived from the Planck cosmic microwave background data, our findings from DES Year 3 data exhibit improvements of up to a factor of five. In addition, we forecast the future sensitivities of the China Space Station Telescope, the upcoming capabilities of this telescope could improve the current limits by approximately one order of magnitude.

arXiv:2402.18880v1 [astro-ph.CO] 29 Feb 2024

## I. INTRODUCTION

Dark Matter (DM) is one of the most fundamental mysteries in modern physics, even though its gravitational effects are well understood. Besides gravity, the interactions between DM and baryons are also of great interest, as explored by experiments such as PandaX [1] and XENONnT [2]. However, no signal has been detected for DM masses above about 1 GeV. Therefore, the focus of experiments has shifted to the sub-GeV mass range, for example, CDEX [3], SENSEI [4], etc.

The interaction between DM and baryons can change the matter distribution in our universe, creating a dark acoustic oscillation (DAO) feature in the matter power spectra. DAO affects the Cosmic microwave background (CMB) anisotropies and the structure formation in the early universe, and constrains the DM-proton elastic scattering cross-section [5–8]. However, the DAO suppression, similar to that of warm DM, is more sensitive to small-scale observations. For velocity independent case, one obtains a stronger limit on this cross-section ($\sigma_{\chi p} < 2.8\times10^{-28}$ cm$^3$ s$^{-1}$ for DM mass around 10 MeV) from the Milky Way satellite abundance [8–11], and the tightest limit ($\sigma_{\chi p} < 1.7\times10^{-29}$ cm$^3$ s$^{-1}$ for DM mass around 10 MeV) from the Lyman-$\alpha$-forest [8, 12, 13]. Nevertheless, the theoretical predictions of these small-scale observations are affected by the non-linear evolution of power spectra and the baryonic feedback. A recent research have explored such baryonic feedback in the galaxies affected by the DM-baryon interactions through hydrodynamical simulation [14]. However, the systematic uncertainties of these predictions remain unclear.

Weak Gravitational Lensing (WL) is a good tool for probing the late-time Large Scale Structure (LSS) of the universe. Through statistical analyses of shape distortions in numerous galaxies induced by foreground matter fields, it can directly map the LSS of the universe. We can mask the small-scale WL data to reduce the uncertainties associated with baryonic feedback, which are more pronounced in observations of the Lyman-$\alpha$-forest at the small scales. In addition, WL data is expected to be more sensitive than the CMB anisotropies. Many recent and upcoming surveys, including Dark Energy Survey (DES) [15], Kilo-Degree Survey (KiDS) [16, 17], Subaru Hyper Suprime-Cam (HSC) [18, 19], Euclid [20], the Vera C. Rubin Observatory [21], the Nancy Grace Roman Space Telescope [22], the Wide Filed Survey Telescope (WFST) [23, 24], the Mephisto Telescope [25, 26] and China Space Station Telescope (CSST) [27–29], greatly improve our understanding of the matter distribution in the late universe. They, in turn, have the potential to reveal the fundamental physics of the interaction between DM and baryonic matter. In this work, we use the data from DES three-year (DES Y3) '`3×2pt`' WL observations along with the CMB and baryonic acoustic oscillation (BAO) observations data. Furthermore, we generate the mock data for CSST and present a forecast of the power of CSST.

Because the photometric galaxy surveys can cover the red-shift with the range $0 < z < 5$, the non-linear effects on the matter power spectrum are essential for the theoretical prediction of WL signal. In this study, we conducted a series of DM-only $N$-body simulations to accurately account for the non-linear effects on the matter power spectrum. The matter power spectrum can be modified by the elastic scattering of DM particle $\chi$ with proton $p$. The scattering cross-section, denoted as $\sigma_{\chi\mathrm{p}} \equiv \sigma_n v_{\mathrm{rel.}}^n$, is parameterized by a power-law index $n$ and the relative velocity between DM and protons $v_{\mathrm{rel.}}$.

* zulei@pmo.ac.cn

† smingtsai@pmo.ac.cn

These scattering interactions induce perturbations in the CMB power spectrum and lead to suppression and oscillations in the matter power spectrum. Taking into account the nonlinear effects from gravitation and scattering, especially at the suppression scale $k \gtrsim 0.1\,h/\mathrm{Mpc}$, is crucial. However, it is a great challenge to incorporate DM-proton scattering in cosmological hydrodynamic simulations, particularly for negative values of $n$ that may govern the nonlinear effects in the late-time Universe. In the case of $n \geqslant 0$, because $v_{\rm rel}$ in the late-time Universe is also suppressed, the DM-proton scattering cross-section can be ignored in cosmological hydrodynamic simulations. Hence, for simplicity, we only discuss the $n \geqslant 0$ scenario to address the nonlinear effects at low redshift. To fast compute our matter power spectrum prediction at low redshift, we employ principal component analysis (PCA) together with a limited set of $N$-body simulations.

The structure of this paper is as follows: We first summarize the linear evolution in DM-proton scattering in Sec. II. In Sec. III, we present our approach to a fast computation of non-linear correction including DM-proton scattering. In Sec. IV, we conduct a high-dimensional Markov Chain Monte Carlo (MCMC) scan to explore the parameter space. Our results are presented and discussed in Sec. V. Finally, we summarize and conclude in Sec. VI.

## II. LINEAR EVOLUTION

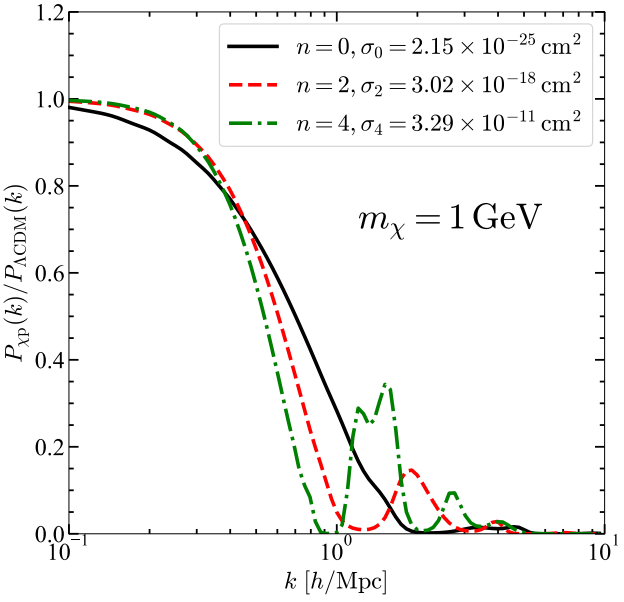



FIG. 1. The ratio of the linear matter power spectrum in DM-proton scattering scenario $P_{\chi\mathrm{p}}(k)$ to that in ΛCDM scenario $P_{\Lambda\mathrm{CDM}}(k)$. The black solid line, red dashed line, and green dashed-dotted line correspond to $n = 0$, $n = 2$, and $n = 4$ at $m_\chi = 1\,\mathrm{GeV}$, respectively. The cross sections $\sigma_{n=0,2,4}$ are the 95% upper limits obtained by P18 + BAO + DES `3×2pt` likelihood in Sec. V.

Considering the elastic scattering interactions between DM ($\chi$) and baryon ($b$), a collision term emerges within the standard Boltzmann equation [30–32],

$$\begin{aligned}
\dot{\delta}_b &= -\theta_b - \frac{\dot{h}}{2},\\
\dot{\theta}_b &= -\frac{\dot{a}}{a}\theta_b + c_b^2 k^2 \delta_b + R_\gamma\left(\theta_\gamma - \theta_b\right) + \frac{\rho_\chi}{\rho_b} R_\chi\left(\theta_\chi - \theta_b\right),\\
\dot{\delta}_\chi &= -\theta_\chi - \frac{\dot{h}}{2},\\
\dot{\theta}_\chi &= -\frac{\dot{a}}{a}\theta_\chi + c_\chi^2 k^2 \delta_\chi + R_\chi\left(\theta_b - \theta_\chi\right).
\end{aligned} \tag{1}$$

The density fluctuations, the velocity divergences in Fourier space, the energy density, and the speeds of sound in each fluid are denoted as $\delta_i$, $\theta_i$, $\rho_i$, and $c_i$, respectively, where subscript $i$ is $b$ for baryon but $\chi$ for DM. The over-dot represents a derivative for conformal time. The quantity $k$ is the wave number, $a$ is the scale factor, and $h$ is the trace of the scalar metric perturbation. The coefficients $R_\chi$ and $R_\gamma$ represent the momentum-transfer rate of DM-proton scattering and the standard Compton scattering. The temperatures of baryon and DM are

$$\begin{aligned}
\dot{T}_b + 2\frac{\dot{a}}{a}T_b &= 2\frac{\mu_b}{m_e}R_\gamma\left(T_\gamma - T_b\right) + 2\frac{\mu_b}{m_\chi}\frac{\rho_\chi}{\rho_b}R'_\chi\left(T_\chi - T_b\right)\\
\dot{T}_\chi + 2\frac{\dot{a}}{a}T_\chi &= 2R'_\chi\left(T_b - T_\chi\right),
\end{aligned} \tag{2}$$

where the temperature of baryon, photon, and DM are $T_b$, $T_\gamma$, and $T_\chi$. The quantity $m_e$ is the mass of the electron, $m_\chi$ is the mass of DM particle, $\mu_b$ is the mean molecular weight of the baryons. The heat transfer rate coefficient $R'_\chi = R_\chi m_\chi/(m_\chi + m_p)$. The DM-proton scattering cross section $\sigma_{\chi p}$ can be simply parameterized as a power law of the DM-proton relative velocity $v_{\rm rel.}$,

$$\sigma_{\chi p} = \sigma_n v_{\rm rel.}^n, \tag{3}$$

where $\sigma_n$ is a dimensional constant factor and the index $n$ can be either a positive or a negative integer. In terms of velocity dependence, the momentum transfer rate coefficient $R_\chi$ can be written as

$$R_\chi = a\rho_b\frac{Y_{\rm H}\mathcal{N}_n\sigma_n}{m_\chi + m_p}\left(\frac{T_\chi}{m_\chi} + \frac{T_b}{m_p}\right)^{(n+1)/2}, \tag{4}$$

where the mass fraction of hydrogen is $Y_{\rm H}$ and the mass of a proton is $m_p$. The numerical factor $\mathcal{N}_n \equiv 2^{(5+n)/2}\Gamma(3+n/2)/(3\sqrt{\pi})$. More details are shown in [31]. Note that we only consider the positive cases in this work, namely $n = 0, 2, 4$ [30, 33, 34], but will return to $n < 0$ cases in the future. Notably, for the case of $n \geq 0$, the bulk relative velocities are much smaller in comparison to the thermal velocities, as mentioned in [31, 35]. We neglect the impact of bulk relative velocities here and thus the calculation in [35] does not affect our analysis.

To model the effects of DM-proton interactions on both the linear matter power spectra and the CMB power

spectra, we use a modified version of the cosmological Boltzmann solver `CLASS` [32, 36]. In Fig. 1, we present the ratios of the linear matter power spectra in the DM-proton scattering scenario $P_{\chi\rm p}$ to that in the ΛCDM $P_{\Lambda\rm CDM}$ for a DM mass $m_\chi = 1\,\rm GeV$. We present three velocity dependencies: $n = 0$ (the black solid line), $n = 2$ (the red dashed line), and $n = 4$ (the green dashed-dotted line). Despite varying cross-sections, the power spectra are all suppressed at $k \sim 0.3\ h/\rm Mpc$. These similar suppression are because we take the cross-sections from their 95% upper limits obtained by P18 + BAO + DES `3×2pt` likelihood, whose detailed description will be given in Sec. V. Although the velocity dependencies show distinct DAOs at $k \sim 1\ h/\rm Mpc$, the tested likelihood is less sensitive to these differences.

## III. NON-LINEAR MATTER POWER SPECTRA

The linear perturbation theory relies on the assumption that the density fluctuation $\delta(k)$ remains significantly smaller than unity. However, during cosmic structure formation, the rapid growth of matter density causes $\delta(k)$ to exceed the mean matter density, rendering it no longer a perturbation. As a consequence, the linear perturbation theory becomes inadequate for describing the intricate structural evolution in regions with wave numbers $k \gtrsim 0.1\ h/\rm Mpc$, commonly referred to as the non-linear region. In addition, DM-proton interactions with a cross-section characterized by $n > 0$ cease to contribute to non-linear evolution, as their interaction rate falls below the Hubble expansion rate.

The easiest but computationally expansive way to explore the non-linear region is to conduct $N$-body simulations. To efficiently include accurate non-linear corrections, several research groups have developed varied analytical methodologies to rapidly compute non-linear corrections. For example, widely used tools `HMCode` [37] and `Halofit` [38] rely on classical CDM simulation outcomes, as well as `WarmAndFuzzy` [39] rely on warm dark matter (WDM) and fuzzy dark matter (FDM) simulations. In this work, we develop a new approach to speed up our non-linear matter power spectrum computations, which allows us to reuse $\mathcal{O}(100)$ $N$-body simulations for a global scan.

First, we define the matter power spectrum ratio $\mathcal{R}$ between $P_{\Lambda\rm CDM}$ and $P_{\chi\rm p}$

$$\mathcal{R}^i(k,z) \equiv \frac{P^i_{\chi\rm p}(k,z)}{P^i_{\Lambda\rm CDM}(k,z)}, \tag{5}$$

where $i$ is either "nl" or "lin" for the cases with and without non-linear effect. To simplify the expression, here we omit the standard arguments (ΛCDM cosmological parameters, $m_\chi$, and $\sigma_n$) in Eq. 5. We compute $\mathcal{R}^{\rm lin}$ by using the Boltzmann code `CLASS` [36], a modified version including DM-proton scattering [32], while $\mathcal{R}^{\rm nl}$ is computed by using $N$-body simulation code `GIZMO` [40]. In the Appendix A, we provide a detailed description of our $N$-body simulations. Note that $\mathcal{R}$ can be strongly dependent on a few parameters. Hence, we then use the widely adopted machine learning skill, principal component analysis (PCA), to investigate how many free parameters there are. For $n = 0, 2, 4$ with DM mass between 1 MeV to 1 TeV, we scrutinize the parameter space suggested by Ref. [41] to verify that the PCA machine can precisely interpolate the linear ratios $\mathcal{R}^{\rm lin}(k,z)$ by

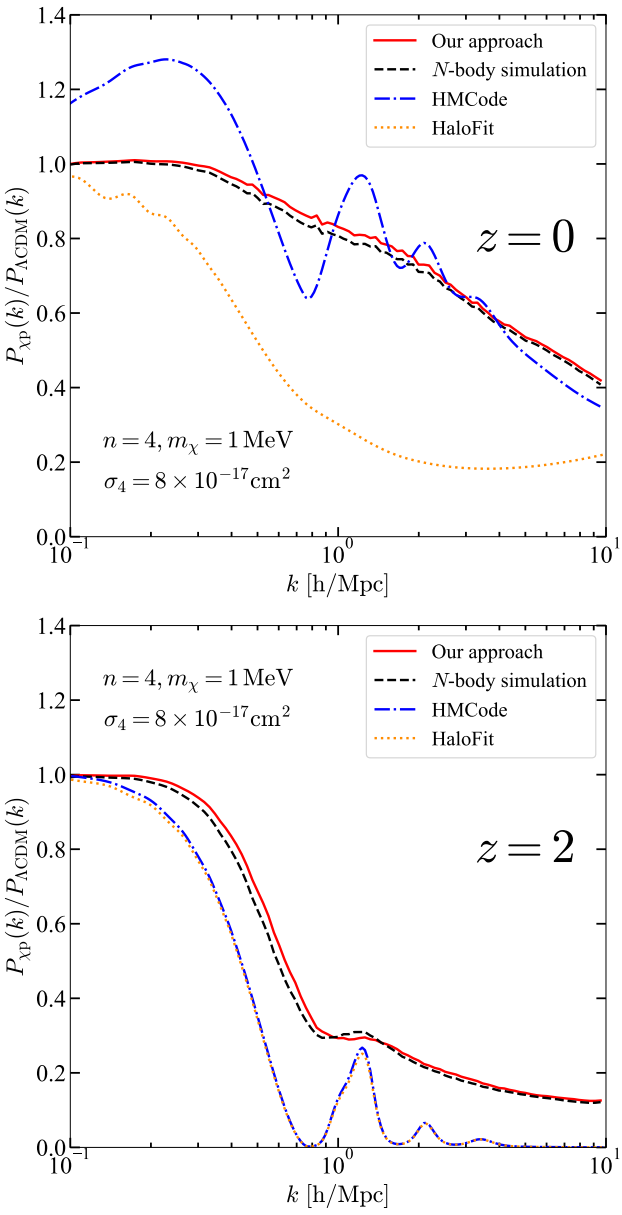



FIG. 2. The top (bottom) panel corresponding to the ratio of the non-linear matter power spectrum at $z = 0$ ($z = 2$) with example parameters $n = 4$, $m_\chi = 1$MeV, $\sigma_4 = 8 \times 10^{-17}\rm cm^2$. The red solid lines denote the outputs of our approach, the black dashed lines are the $N$-body simulation results, the blue dashed-dotted lines are the correction from `HMCode` and the orange dotted lines are the correction from `HaloFit`.

two parameters [1]. Such a collected set of $\mathcal{R}^{\rm lin}$ generated by parameter scan can be then used for training and validation. After training our PCA machine, we can simply use it to compress any $\mathcal{R}^{\rm lin}$ to only two PCA components.

Our next goal is to build reusable $\mathcal{R}^{\rm nl}$ tables as function of two PCA components. We uniformly select 205 points within this two-dimensional map from previously collected training and validating data. We then perform 205 $N$-body simulations accordingly to obtain corresponding $\mathcal{R}^{\rm nl}$. Consequently, for any linear power spectrum falling within our specified range of interest, a straightforward mapping from linear to non-linear matter power spectrum becomes achievable. Namely, we can interpolate the $\mathcal{R}^{\rm nl}$ tables with their corresponding two PCA components. Hence, the final non-linear matter power spectrum for DM-proton scattering can be obtained via

$$P^{\rm nl}_{\chi{\rm p}}(k,z) = \mathcal{R}^{\rm nl}(k,z) \times P^{\texttt{Halofit}}_{\Lambda{\rm CDM}}(k,z), \tag{6}$$

where $P^{\texttt{Halofit}}_{\Lambda{\rm CDM}}(k,z)$ is the non-linear matter power spectrum of ΛCDM cosmology computed by `Halofit`. We emphasize that the $N$-body simulations are only dependent on initial conditions making our approach applicable to various DAO models. This is the greatest advantage of our method, even though this paper only focuses on the DM-proton scattering scenario.

In Fig. 2, we present a comparison between our approach (red solid lines), $N$-body simulations (black dashed lines), `HMCode` (blue dashed-dotted lines), and `HaloFit` (orange dotted lines) for a benchmark point ($n = 4$, $m_\chi = 1$ MeV, $\sigma_4 = 8 \times 10^{-17}{\rm cm}^2$).

For both $z = 0$ (upper panel) and $z = 2$ (lower panel), we can see that our approach predicts the $N$-body simulation well. On the other hand, the `HMCode` and `HaloFit` cannot catch the features of the non-linear effects, because they exclusively rely on ΛCDM simulations. Quantitatively, our approach generates an acceptable error around 0.5%. For more information, the error estimation is given in Appendix B.

## IV. MARKOV CHAIN MONTE CARLO ANALYSIS

Based on PCA machine and the $\mathcal{R}^{\rm nl}$ tables developed in Sec. III, we can include non-linear corrections on the fly when performing a MCMC scan. The scan parameters are the six fundamental ΛCDM cosmological parameters, and $\sigma_{n=0,2,4}$ as given in Eq. 3. The six ΛCDM parameters consist of

$$\{\Omega_{\rm b}h^2,\ \Omega_{\rm cdm}h^2,\ \theta_{\rm s},\ \log(10^{10}A_{\rm s}),\ n_{\rm s},\ \tau_{\rm reio}\}, \tag{7}$$

corresponding to the baryon density, the cold DM density, the CMB peak position, the scalar amplitude, the scalar spectral index, and the optical depth to reionization, respectively. The scan range required for Eq. 7 can be found in previous work [43]. We assume all the DM particles can scatter with proton. Except for $\sigma_n$, we select seven benchmark masses $m_\chi \in \{10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}, 10^{3}\}$ GeV, and the three power indexes $n = 0, 2, 4$ for different scans. Namely, we have $7 \times 3$ scans in total for those benchmark masses and power indexes. We employ the cosmological MCMC package `MontePython` [44, 45] to undertake the task of the global fitting.

We include the following cosmological data into the likelihood: (i) The `3×2pt` likelihood based on the DES Y3 observations [43, 46], (ii) The CMB likelihoods are calculated based on Planck 2018 Legacy (P18) [47], including high-$\ell$ power spectra (`TT`, `TE`, and `EE`), low-$\ell$ power spectrum (`TT` and `EE`), and Planck lensing power spectrum (`lensing`), (iii) The BAO likelihood contains the BOSS DR12 dataset measurements at $z = 0.106$, $z = 0.15$ and $z = 0.2 - 0.75$ [48–50]. The details of DES Y3 `3×2pt` data and modeling are given in the Appendix C.

## V. RESULTS AND DISCUSSION

In this study, we determine the 95% upper limit of $\sigma_n$ using two methods: Bayesian Marginalized Posterior (MP) and Frequentist Profile Likelihood (PL). Both methods set upper limits by integrating $\sigma_n$ until 95% of the total probability is reached. The method MP integrates the probability density over nuisance parameters by marginalizing posterior densities, commonly applied in cosmology, particularly in the ΛCDM context. On the other hand, the PL method is preferred for null signal searches due to large volume effects and prior dependencies in unconstrained likelihoods, and it is used in DM direct detection [1–3] and indirect detection [51–53].

In our study on DM-proton scattering cross-section, we primarily present upper limits using the PL method, with results from the MP method included for comparative analysis [2].

In Fig. 3, we present the ratio spectra at $z = 0$ as a function of $k$, by varying $\sigma_{\chi p}$ in the left panels and $m_\chi$ in the right panels.

In the left panels, we take a constant DM mass at $m_\chi = 1$ GeV and investigate three different velocity scenarios: $n = 0$ (upper left panel), $n = 2$ (middle left panel), and $n = 4$ (lower left panel). Three colored lines correspond to the cross-sections $\sigma_{n=0,2,4}$ set at $0.1 \times \sigma_n^{95\%}$ (the red lines), $\sigma_n^{95\%}$ (the black lines), and $10 \times \sigma_n^{95\%}$ (the green lines). When comparing the non-linear and linear ratio spectra (represented by solid and dashed lines,

[1] A similar method has been utilized to investigate two-body decaying DM in [42].

[2] Recently, Ref. [54] has proposed that the Bayes factor surface may be an alternative good tool for dealing with null signal search.

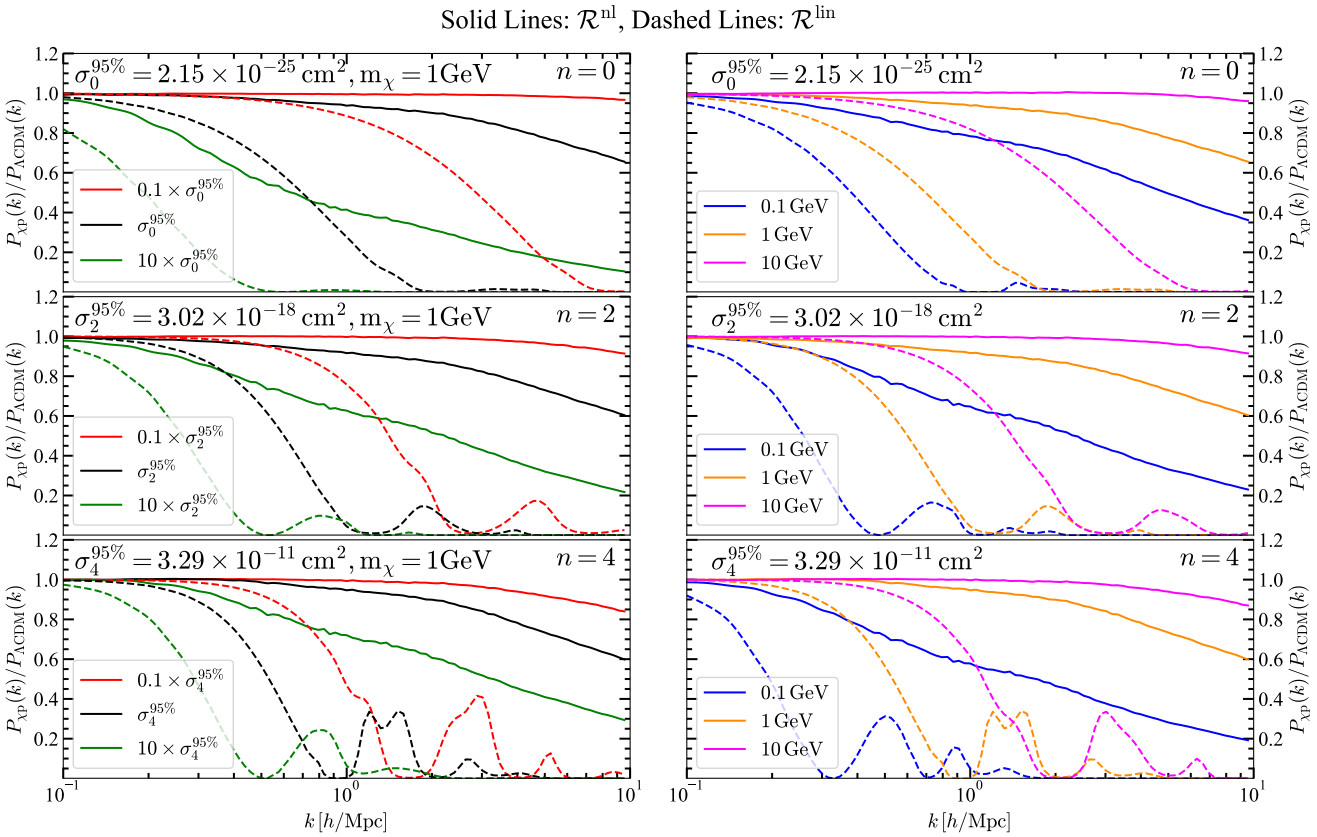



FIG. 3. The ratio power spectra caused by the DM-proton scatterings. The dashed lines are the linear results, and the solid lines are the non-linear results. Left Column: the ratio power spectra for DM mass $m_\chi = 1\,{\rm GeV}$ with the velocity scenarios $n = 0$ (left upper panel), $n = 2$ (left middle panel), and $n = 4$ (left lower panel). The reference cross-sections are $\sigma_0^{95\%} = 2.15 \times 10^{-25}\,{\rm cm}^2$, $\sigma_2^{95\%} = 3.02 \times 10^{-18}\,{\rm cm}^2$, and $\sigma_4^{95\%} = 3.29 \times 10^{-11}\,{\rm cm}^2$. These values represent the 95% upper limits based on the likelihood P18 + BAO + DES `3×2pt`, as discussed in Sec. V. Right column: the ratio spectra for DM mass $m_\chi = 100\,{\rm MeV}$ (blue lines), $m_\chi = 1\,{\rm GeV}$ (orange lines) and $m_\chi = 10\,{\rm GeV}$ (magenta lines) for $n = 0, 2, 4$, with the same cross-section given in the right panels.

respectively), the non-linear effects tend to washout the features of DAO and cause a shift in suppression towards smaller scales. The values of $\sigma_{n=0,2,4}^{95\%}$ displayed in the left corner of each figure are derived from the 95% upper limits of the likelihood (P18 + BAO + DES `3×2pt`). These values can be considered as characteristic cross-sections. For the sake of convenience, we subsequently define $k_{0.8}$ as the wavenumber where the ratio spectrum is measured at 0.8. Observing all non-linear ratio spectra with $\sigma_{n=0,2,4}^{95\%}$, we find that $k_{0.8}$ is approximately 5 $h$/Mpc. This value also indicates the sensitive scale of the DES `3×2pt` data.

An important trend emerges as reducing the cross-section value leads to the suppression of the matter power spectrum, shifting it toward smaller scales. This shift results from an earlier decoupling between DM and baryons due to reduced interaction between them. Therefore, for future experiments aiming to enhance detection sensitivity, efforts should be directed towards smaller-scale regions, such as $k > 1$ $h$/Mpc. However, theoretical challenges, including baryonic feedback and galaxy bias, introduce substantial and complex systematic uncertainties in WL surveys at small scales. Building on previous studies [55] and implementing masks on corresponding small angular scales in WL data, as described in further detail in Appendix C, aims to mitigate these uncertainties. This strategy also enables us to extend our analysis to future WL data.

In the right panels, we vary the parameter $m_\chi$ while keeping $\sigma_n$ fixed at $\sigma_n^{95\%}$. The ratio spectra for $m_\chi = 100\,{\rm MeV}$ (blue lines), $m_\chi = 1\,{\rm GeV}$ (orange lines), and $m_\chi = 10\,{\rm GeV}$ (magenta lines) are compared for scenarios of $n = 0, 2, 4$. Because the lighter DM results in a higher DM number density, the DM-proton interaction rate is also larger. This leads to the suppression of the matter power spectra for lighter DM shifting toward larger scales. When comparing with the left panels, we see that variations in the cross-section have a more significant impact on the location of $k_{0.8}$ than changes in the DM mass.

In Fig. 4, we show the 95% upper limits on the DM-proton velocity-dependent elastic scattering cross-section

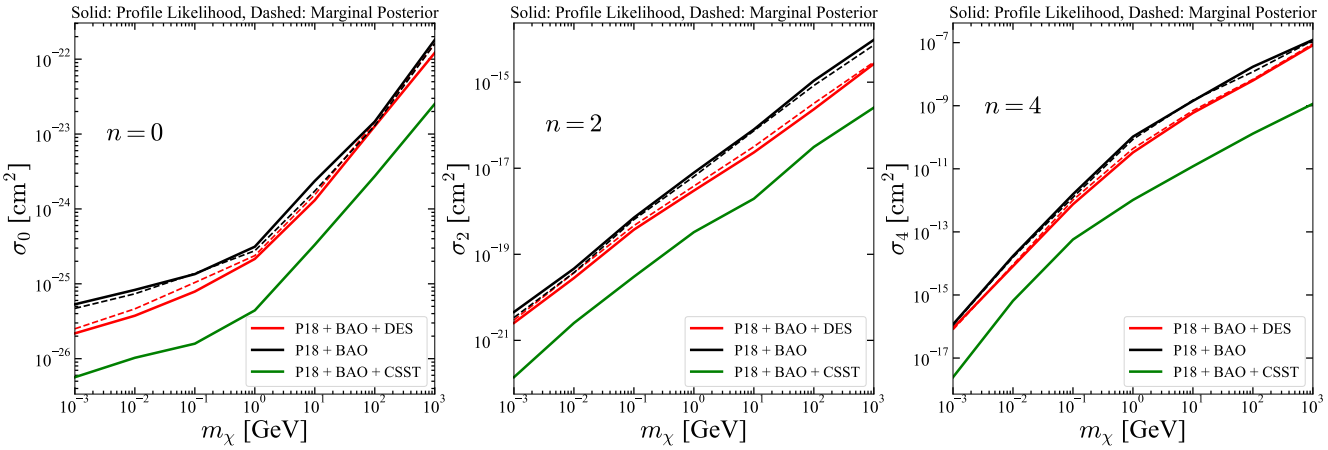



FIG. 4. The the 95% upper limits of cross sections $\sigma_n$ derived by the P18 + BAO + DES likelihoods (red lines) and by the P18 + BAO likelihoods (black lines). The solid lines are obtained by Profile Likelihood method and the dashed lines are obtained by Marginal Posterior method. The green lines correspond to the forecast bounds of CSST obtained by Profile Likelihood method.

$\sigma_n$. These limits are derived from the likelihood functions given in Sec. IV, making use of both the PL method (solid lines) and MP method (dashed lines). In all velocity-dependent scenarios, the PL method produces the most stringent upper limits, illustrated by red solid lines and based on the combined likelihoods of **P18 + BAO + DES** `3×2pt`. Black lines represent upper limits derived solely from the likelihoods of **P18 + BAO**. Additionally, we conducted a forecast to assess the constraining potential of upcoming CSST for DM-proton scattering, shown by green lines.

We see that incorporating DES `3×2pt` data improves constraints compared to early universe observations (P18 + BAO) by a factor ranging from one to five. Our results support conclusions from previous studies [10, 43] that observations of WL on small-scale matter distributions in the late universe provide a more precise exploration of DM physics. Finally, the projected 95% upper limits (the green lines in Fig.4) are derived from the CSST mock cosmic shear data with PL method. We generated the CSST mock cosmic shear data and the corresponding covariance matrix, employing CSST sensitivity settings based on P18 ΛCDM parameters: $\{\Omega_{\rm b}h^2 = 0.02238$, $\Omega_{\rm cdm}h^2 = 0.1201$, $\theta_{\rm s} = 0.01041$, $\log(10^{10}A_{\rm s}) = 3.045$, $n_{\rm s} = 0.9659$, $\tau_{\rm reio} = 0.0543\}$. The covariance matrix was computed using the `CosmoCov` code [56, 57], incorporating a conservative galaxy shape noise of $\sigma_e = 0.3$ [58] and the red-shift distributions of mock CSST galaxies from Ref. [59]. Again, we applied the same small angular scale mask as DES does in Appendix C. Remarkably, the future CSST cosmic shear sensitivity may substantially lower down the upper limit by around one to two orders compared to the DES `3×2pt` likelihood. Note that the primary theoretical uncertainties in our study arise from baryon feedback and the interpolation range of our PCA machine about percent level, see Appendix B for details. We expect to address and refine these aspects in future work to enhance the overall robustness of our results.

## VI. SUMMARY AND CONCLUSION

In order to study the role of DM-proton scattering in cosmology, we focused on the matter distribution at scales $k > 0.1, h/\text{Mpc}$, where non-linear effects become important. Accounting for the impact of DM-proton scattering in the non-linear evolution of structure formation, we computed 205 linear matter power spectra as initial conditions at $z = 127$ for subsequently conducted cosmological $N$-body simulations. To incorporate these effects into the MCMC global scan, we utilized the PCA method to establish a mapping between any linear ratio spectrum (the spectrum of the DM-proton scattering scenario to that of ΛCDM) and the corresponding ratio derived from $N$-body simulations.

Employing this approach, we utilized data from DES Y3 WL `3×2pt`, together with likelihoods derived from Planck 2018 CMB and BOSS DR12 BAO, to constrain the velocity-dependent elastic scattering cross-sections between DM particles and protons. We find null signals of DM-proton scattering, thus we set an upper limit on the cross section. Our comprehensive studies of the parameter space in the context of DM-proton scattering cosmology allowed us to estimate the 95% upper limits of cross-sections $\sigma_{n=0,2,4}$ (refer to Fig. 4). It is worthy to mention that the constraints derived from the DES `3×2pt` data showed a substantial improvement, achieving a factor of $\mathcal{O}(5)$ over previous results. Our results, introduced as the first Frequentist upper limits, are compared with those obtained using a Bayesian approach for a more comprehensive understanding.

In the near future, the precision of weak lensing measurements in galaxy surveys is set for further investigation into DM-proton scattering. To explore the sensi-

tivities of future WL data, we generated cosmic shear mock data and a covariance matrix for the CSST. The resulting sensitivities from CSST exhibit a remarkable improvement of one to two orders of magnitudes compared to the current DES `3×2pt` data, as illustrated in Fig. 4.

## VII. ACKNOWLEDGEMENTS

We are grateful to Rui An, Zi-qing Xia, Yuchao Gu, Lei Lei for useful discussions. This work is supported by the National Key Research and Development Program of China (No. 2022YFF0503304), and the Project for Young Scientists in Basic Research of the Chinese Academy of Sciences (No. YSBR-092).

### Appendix A: Settings of $N$-body simulations

We briefly introduce the settings of our $N$-body simulations. In this work, we conducted our $N$-body simulations using the TreePM code `GIZMO` [40], where the $N$-body part originates from the source code of `Gadget2` [60]. Our simulations start at $z = 127$, with the initial conditions generated by the code `2LPTic` [61]. The input power spectrum and corresponding $\sigma_8$ were obtained from `CLASS`. For all simulation points, we utilized a set of fiducial cosmological parameters $\{\Omega_m = 0.3118, \Omega_\Lambda = 0.6881, h = 0.6760, n_s = 0.9659\}$.

### Appendix B: Error estimation of our approach

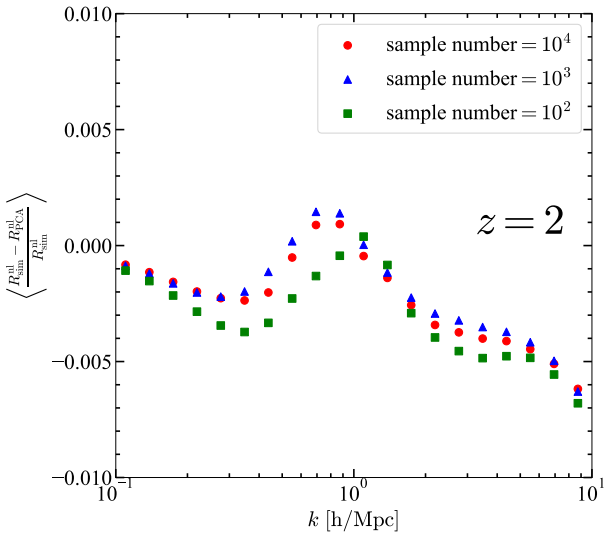



FIG. 5. The average non-linear ratio deviation between simulation results and our approach output at each wave number ($k$) at $z = 2$. The green cube, blue triangle and red dot marker correspond to sample number $N = 10^2, 10^3, 10^4$, respectively.

We test the robustness of our approach. Firstly, we randomly take 5 test points from the total 205 simulation points and using the left 200 points to build the interpolation map. Then we compute the average non-linear ratio deviation between the output of our approach (PCA) and simulation (sim) at each wave number ($k$) bins for the 5 test points, as once sampling. Finally, we compute the total average non-linear ratio deviation by summing over all the sample numbers,

$$\left\langle \frac{R^{\rm nl}_{\rm sim} - R^{\rm nl}_{\rm PCA}}{R^{\rm nl}_{\rm sim}} \right\rangle = \frac{1}{N} \sum_{N} \frac{1}{5} \sum_{i=1}^{i=5} \frac{R^{\rm nl}_{{\rm sim},i}(k,z) - R^{\rm nl}_{{\rm PCA},i}(k,z)}{R^{\rm nl}_{{\rm sim},i}(k,z)}, \tag{B1}$$

where $N$ is the sample number. We find the largest deviation appeared at $z = 2$ is shown in Fig. 5, which prove our approach could provide valid non-linear correction with an error less than 0.5% in DM-proton scattering scenario.

### Appendix C: WL observable and modeling

In this section, we outline the WL observable and model predictions according to reference [46, 55, 57, 62, 63]. Additionally, we illustrate the dependency of WL modeling on the non-linear matter power spectrum derived in Sec. III.

We start our analysis with the DES Y3 WL data employed in this study. This data-set encompasses measurements of galaxy shapes and distributions across 4143 deg$^2$ of the entire sky. The observed galaxies were sorted into two catalogs: the `METACALIBRATION` [64] source catalog and the `MagLim` [65] lens catalog. The source catalog contains $10^8$ galaxies and is categorized into four tomographic red-shift bins with bin edges at $z = [0.0, 0.36, 0.63, 0.87, 2.0]$, while the lens catalog encompasses $10^7$ galaxies, separated into six tomographic red-shift bins with bin edges at $z = [0.20, 0.40, 0.55, 0.70, 0.85, 0.95, 1.05]$. Each redshift bin $i$ or $j$ includes a red-shift density distribution of galaxies, denoted as $n^i_\kappa(z)$ for `METACALIBRATION` and $n^j_g(z)$ for `MagLim`. The measured shape distortions and angular positions of galaxies provide three sets of two-point correlations, referred to as '`3×2pt`'. These correlations include the angular separations of foreground lens galaxy pairs, measured through galaxy clustering $w(\theta)$ (position-position). The shape distortions of source galaxies can also be related to the distribution of foreground lens galaxies, measured via galaxy-galaxy lensing $\gamma_t(\theta)$ (position-shape). However, since galaxies act as a biased tracer of the matter field, cosmic shear $\xi_\pm(\theta)$ (shape-shape), as a cosmological WL method involves correlations between the shape distortions of source galaxy pairs, directly trace the foreground matter field and is extremely sensitive to the LSS, as well as the matter power spectrum in the late universe. The DES collab-

oration offers a comprehensive WL 3×2pt data [3], encompassing both auto and cross-correlations among different red-shift bins. Each observable is computed across 20 $\theta$-bins logarithmically spanning from 2.5 to 250 arcmin, along with its respective covariance matrix. However, the linear galaxy bias assumption breaks down on small angular scales, and complexities in baryonic physics introduce uncertainties. We adopt a mask method for the small angular scale data, consistent with [46]. Additionally, due to the poor model fit caused by high $z$ bins data in MagLim, we implement a high $z$ cut, resulting in the exclusion of bins 5 and 6 from the MagLim in our analysis [46]. It's important to note that the complete 3×2pt data vector consists of 1000 elements, only 462 elements remain after masking.

Now, we elucidate the model predictions for 3×2pt signals. Utilizing the Limber approximation [66, 67], we can compute the angular power spectrum for the 3×2pt observable between red-shift bins $i$ and $j$:

$$C^{ij}_{\kappa\kappa,\kappa\mathrm{g},\mathrm{gg}}(\ell) = \int \mathrm{d}\chi \frac{q^i_{\mathrm{g}/\kappa} q^j_{\mathrm{g}/\kappa}}{\chi^2} \times P^{\mathrm{nl}}_{\chi\mathrm{p}}\left(k = \frac{\ell + 0.5}{\chi}, z(\chi)\right), \tag{C1}$$

where $\kappa$ and g denote the source field and lens field, respectively. The symbol $\chi$ represents the comoving distance, while $\ell$ stands for the angular wave number. The 3D non-linear matter power spectrum, incorporating the characteristics of DM-proton scattering denoted as $P^{\mathrm{nl}}_{\chi\mathrm{p}}$, was derived in Eq. 6. The radial weight function for clustering in red-shift bin $i$ is

$$q^i_g(\chi) = b^i n^i_g(z(\chi)) \frac{\mathrm{d}z}{\mathrm{d}\chi} \tag{C2}$$

and the lensing efficiency kernel in red-shift bin $j$ is

$$q^j_\kappa(\chi) = \frac{3H_0^2\Omega_m}{2\mathrm{c}^2} \frac{\chi}{a(\chi)} \int_\chi^{\chi_{z_{\max}}} \mathrm{d}\chi' n^j_\kappa\left(z\left(\chi'\right)\right) \frac{\mathrm{d}z}{\mathrm{d}\chi'} \frac{\chi' - \chi}{\chi'}, \tag{C3}$$

where $b^i$ represents the linear galaxy bias within redshift bin $i$. The normalized distributions of lens galaxies in red-shift bin $i$ and source galaxies in bin $j$ correspond to $n^i_g(z)$ and $n^j_\kappa(z)$ respectively, which can be obtained from DES Y3 products [3]. The parameters $H_0$, $\Omega_m$, $c$, and $a(\chi)$ signify the Hubble constant, total matter density, speed of light, and scale factor. The angular correlation function for galaxy clustering can be computed from

$$w^i(\theta) = \int \frac{\mathrm{d}\ell\ell}{2\pi} J_0(\ell \cdot \theta) C^{ij}_{\mathrm{gg}}(\ell), \tag{C4}$$

the galaxy-galaxy lensing correlation function for lens galaxies in bin $i$ and source galaxies in bin $j$ can be writen as

$$\gamma^{ij}_{\mathrm{t}}(\theta) = \left(1 + m^j\right) \int \frac{\mathrm{d}\ell\ell}{2\pi} J_2(\ell \cdot \theta) C^{ij}_{\mathrm{g}\kappa}(\ell), \tag{C5}$$

and the cosmic shear correlation functions are

$$\xi^{ij}_{\pm}(\theta) = \left(1 + m^i\right)\left(1 + m^j\right) \int \frac{\mathrm{d}\ell\ell}{2\pi} J_{0/4}(\ell \cdot \theta) C^{ij}_{\kappa\kappa}(\ell), \tag{C6}$$

where $J_0$, $J_2$, and $J_4$ denote the zeroth-order, second-order, and fourth-order Bessel functions, respectively. The pre-factor $m^i$ represents the multiplicative shear bias in bin $i$. Additionally, several other critical systematic uncertainties must be taken into account. For instance, the red-shift distributions of galaxies in bin $i$ need calibration due to photo-$z$ uncertainties $\Delta z^i$, adjusted as $n^i(z - \Delta z^i)$. The intrinsic alignment (IA) of source galaxies significantly impacts $\gamma^{ij}_t(\theta)$ and $\xi^{ij}_\pm(\theta)$. We address this issue using a non-linear alignment (NLA) model [18, 63].


[1] PandaX collaboration, X. Ning et al., *Limits on the luminance of dark matter from xenon recoil data*, *Nature* **618** (2023) 47–50.

[2] XENON collaboration, E. Aprile et al., *First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment*, *Phys. Rev. Lett.* **131** (2023) 041003, [2303.14729].

[3] CDEX collaboration, W. H. Dai et al., *Exotic Dark Matter Search with the CDEX-10 Experiment at China's Jinping Underground Laboratory*, *Phys. Rev. Lett.* **129** (2022) 221802, [2209.00861].

[4] SENSEI collaboration, P. Adari et al., *SENSEI: First Direct-Detection Results on sub-GeV Dark Matter from SENSEI at SNOLAB*, 2312.13342.

[5] K. K. Boddy, V. Gluscevic, V. Poulin, E. D. Kovetz, M. Kamionkowski and R. Barkana, *Critical assessment of CMB limits on dark matter-baryon scattering: New treatment of the relative bulk velocity*, *Phys. Rev. D* **98** (2018) 123506, [1808.00001].

[6] T. R. Slatyer and C.-L. Wu, *Early-Universe constraints on dark matter-baryon scattering and their implications for a global 21 cm signal*, *Phys. Rev. D* **98** (2018) 023013, [1803.09734].

[7] K. K. Boddy and V. Gluscevic, *First Cosmological Constraint on the Effective Theory of Dark Matter-Proton Interactions*, *Phys. Rev. D* **98** (2018) 083510, [1801.08609].

[8] M. A. Buen-Abad, R. Essig, D. McKeen and Y.-M. Zhong, *Cosmological constraints on dark matter interactions with ordinary matter*, *Phys. Rept.* **961** (2022) 1–35, [2107.12377].

[9] E. O. Nadler, V. Gluscevic, K. K. Boddy and R. H.


[3] https://des.ncsa.illinois.edu/releases/y3a2/Y3key-products

Wechsler, *Constraints on Dark Matter Microphysics from the Milky Way Satellite Population*, *Astrophys. J. Lett.* **878** (2019) 32, [1904.10000].

[10] K. Maamari, V. Gluscevic, K. K. Boddy, E. O. Nadler and R. H. Wechsler, *Bounds on velocity-dependent dark matter-proton scattering from Milky Way satellite abundance*, *Astrophys. J. Lett.* **907** (2021) L46, [2010.02936].

[11] DES collaboration, E. O. Nadler et al., *Milky Way Satellite Census. III. Constraints on Dark Matter Properties from Observations of Milky Way Satellite Galaxies*, *Phys. Rev. Lett.* **126** (2021) 091101, [2008.00022].

[12] D. C. Hooper and M. Lucca, *Hints of dark matter-neutrino interactions in Lyman-$\alpha$ data*, *Phys. Rev. D* **105** (2022) 103504, [2110.04024].

[13] K. K. Rogers, C. Dvorkin and H. V. Peiris, *Limits on the Light Dark Matter–Proton Cross Section from Cosmic Large-Scale Structure*, *Phys. Rev. Lett.* **128** (2022) 171301, [2111.10386].

[14] J. F. Acevedo, H. An, Y. Boukhtouchen, J. Bramante, M. Richardson and L. Sansom, *Dark Matter-Induced Baryonic Feedback in Galaxies*, 2309.08661.

[15] DES collaboration, T. M. C. Abbott et al., *Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing*, *Phys. Rev. D* **98** (2018) 043526, [1708.01530].

[16] E. van Uitert et al., *KiDS+GAMA: cosmology constraints from a joint analysis of cosmic shear, galaxy–galaxy lensing, and angular clustering*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 4662–4689, [1706.05004].

[17] C. Heymans et al., *KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints*, *Astron. Astrophys.* **646** (2021) A140, [2007.15632].

[18] HSC collaboration, C. Hikage et al., *Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data*, *Publ. Astron. Soc. Jap.* **71** (2019) 43, [1809.09148].

[19] T. Hamana et al., *Cosmological constraints from cosmic shear two-point correlation functions with HSC survey first-year data*, *Publ. Astron. Soc. Jap.* **72** (2020) Publications of the Astronomical Society of Japan, Volume 72, Issue 1, February 2020, 16, https://doi.org/10.1093/pasj/psz138, [1906.06041].

[20] EUCLID collaboration, R. Laureijs et al., *Euclid Definition Study Report*, 1110.3193.

[21] LSST collaboration, v. Ivezić et al., *LSST: from Science Drivers to Reference Design and Anticipated Data Products*, *Astrophys. J.* **873** (2019) 111, [0805.2366].

[22] J. Green et al., *Wide-Field InfraRed Survey Telescope (WFIRST) Final Report*, 1208.4012.

[23] Z. Lou, M. Liang, D. Yao, X. Zheng, J. Cheng, H. Wang et al., *Optical design study of the Wide Field Survey Telescope (WFST)*, in *Advanced Optical Design and Manufacturing Technology and Astronomical Telescopes and Instrumentation* (M. Xu and J. Yang, eds.), vol. 10154, p. 101542A, International Society for Optics and Photonics, SPIE, 2016. DOI.

[24] L. Lei, Q.-F. Zhu, X. Kong, T.-G. Wang, X.-Z. Zheng, D.-D. Shi et al., *Limiting Magnitudes of the Wide Field Survey Telescope (WFST)*, *Res. Astron. Astrophys.* **23** (2023) 035013, [2301.03068].

[25] L. Lei, B.-Q. Chen, J.-D. Li, J.-T. Wu, S.-Y. Jiang and X.-W. Liu, *Identifications of RR Lyrae Stars and Quasars from the Simulated Data of Mephisto-W Survey*, *Res. Astron. Astrophys.* **22** (2022) 025004, [2111.08316].

[26] X. Liu, *Multi-channel Photometric Survey Telescope - Mephisto*, in *Galactic Archaeology in the Gaia Era*, p. 14, Jan., 2019.

[27] H. Zhan, *Consideration for a large-scale multi-color imaging and slitless spectroscopy survey on the Chinese space station and its application in dark energy research*, *Scientia Sinica Physica, Mechanica & Astronomica* **41** (Jan., 2011) 1441.

[28] H. Zhan, *The wide-field multiband imaging and slitless spectroscopy survey to be carried out by the survey space telescope of china manned space program*, *Chinese Science Bulletin* **66** (04, 2021) 1290–1298.

[29] Y. Gong, X. Liu, Y. Cao, X. Chen, Z. Fan, R. Li et al., *Cosmology from the Chinese Space Station Optical Survey (CSS-OS)*, *Astrophys. J.* **883** (2019) 203, [1901.04634].

[30] X.-l. Chen, S. Hannestad and R. J. Scherrer, *Cosmic microwave background and large scale structure limits on the interaction between dark matter and baryons*, *Phys. Rev. D* **65** (2002) 123515, [astro-ph/0202496].

[31] C. Dvorkin, K. Blum and M. Kamionkowski, *Constraining Dark Matter-Baryon Scattering with Linear Cosmology*, *Phys. Rev. D* **89** (2014) 023519, [1311.2937].

[32] D. V. Nguyen, D. Sarnaaik, K. K. Boddy, E. O. Nadler and V. Gluscevic, *Observational constraints on dark matter scattering with electrons*, *Phys. Rev. D* **104** (2021) 103521, [2107.12380].

[33] K. Sigurdson, M. Doran, A. Kurylov, R. R. Caldwell and M. Kamionkowski, *Dark-matter electric and magnetic dipole moments*, *Phys. Rev. D* **70** (2004) 083501, [astro-ph/0406355].

[34] A. Melchiorri, A. Polosa and A. Strumia, *New bounds on millicharged particles from cosmology*, *Physics Letters B* **650** (July, 2007) 416–420.

[35] Y. Ali-Haïmoud, S. S. Gandhi and T. L. Smith, *Exact treatment of weak dark matter-baryon scattering for linear-cosmology observables*, 2312.08497.

[36] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes*, *JCAP* **07** (2011) 034, [1104.2933].

[37] A. Mead, C. Heymans, L. Lombriser, J. Peacock, O. Steele and H. Winther, *Accurate halo-model matter power spectra with dark energy, massive neutrinos and modified gravitational forces*, *Mon. Not. Roy. Astron. Soc.* **459** (2016) 1468–1488, [1602.02154].

[38] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya and M. Oguri, *Revising the Halofit Model for the Nonlinear Matter Power Spectrum*, *Astrophys. J.* **761** (2012) 152, [1208.2701].

[39] D. J. E. Marsh, *WarmAndFuzzy: the halo model beyond CDM*, 1605.05973.

[40] P. F. Hopkins, *A new class of accurate, mesh-free hydrodynamic simulation methods*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 53–110, [1409.7395].

[41] Z. Li et al., *The Atacama Cosmology Telescope: limits on dark matter-baryon interactions from DR4 power spectra*, *JCAP* **02** (2023) 046, [2208.08985].

[42] J. Bucko, S. K. Giri, F. H. Peters and A. Schneider,

*Probing the two-body decaying dark matter scenario with weak lensing and the cosmic microwave background*, 2307.03222.

[43] L. Zu, C. Zhang, H.-Z. Chen, W. Wang, Y.-L. S. Tsai, Y. Tsai et al., *Exploring mirror twin Higgs cosmology with present and future weak lensing surveys*, *JCAP* **08** (2023) 023, [2304.06308].

[44] B. Audren, J. Lesgourgues, K. Benabed and S. Prunet, *Conservative Constraints on Early Cosmology: an illustration of the Monte Python cosmological parameter inference code*, *JCAP* **02** (2013) 001, [1210.7183].

[45] T. Brinckmann and J. Lesgourgues, *MontePython 3: boosted MCMC sampler and other features*, *Phys. Dark Univ.* **24** (2019) 100260, [1804.07261].

[46] DES collaboration, T. M. C. Abbott et al., *Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing*, *Phys. Rev. D* **105** (2022) 023520, [2105.13549].

[47] Planck collaboration, N. Aghanim et al., *Planck 2018 results. V. CMB power spectra and likelihoods*, *Astron. Astrophys.* **641** (2020) A5, [1907.12875].

[48] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell et al., *The 6dF Galaxy Survey: Baryon Acoustic Oscillations and the Local Hubble Constant*, *Mon. Not. Roy. Astron. Soc.* **416** (2011) 3017–3032, [1106.3366].

[49] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden and M. Manera, *The clustering of the SDSS DR7 main Galaxy sample – I. A 4 per cent distance measure at $z = 0.15$*, *Mon. Not. Roy. Astron. Soc.* **449** (2015) 835–847, [1409.3242].

[50] BOSS collaboration, S. Alam et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample*, *Mon. Not. Roy. Astron. Soc.* **470** (2017) 2617–2652, [1607.03155].

[51] Fermi-LAT, DES collaboration, A. Albert et al., *Searching for Dark Matter Annihilation in Recently Discovered Milky Way Satellites with Fermi-LAT*, *Astrophys. J.* **834** (2017) 110, [1611.03184].

[52] V. Gammaldi, J. Pérez-Romero, J. Coronado-Blázquez, M. Di Mauro, E. V. Karukes, M. A. Sánchez-Conde et al., *Dark matter search in dwarf irregular galaxies with the fermi large area telescope*, *Phys. Rev. D* **104** (Oct, 2021) 083026.

[53] Hess, HAWC, VERITAS, MAGIC, H.E.S.S., Fermi-LAT collaboration, H. Abdalla et al., *Combined dark matter searches towards dwarf spheroidal galaxies with Fermi-LAT, HAWC, H.E.S.S., MAGIC, and VERITAS*, *PoS* **ICRC2021** (2021) 528, [2108.13646].

[54] A. Fowlie, *The Bayes factor surface for searches for new physics*, 2401.11710.

[55] DES collaboration, A. Amon et al., *Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to data calibration*, *Phys. Rev. D* **105** (2022) 023514, [2105.13543].

[56] X. Fang, T. Eifler and E. Krause, *2D-FFTLog: Efficient computation of real space covariance matrices for galaxy clustering and weak lensing*, *Mon. Not. Roy. Astron. Soc.* **497** (2020) 2699–2714, [2004.04833].

[57] E. Krause and T. Eifler, *cosmolike – cosmological likelihood analyses for photometric galaxy surveys*, *Mon. Not. Roy. Astron. Soc.* **470** (2017) 2100–2112, [1601.05779].

[58] H. Miao, Y. Gong, X. Chen, Z. Huang, X.-D. Li and H. Zhan, *Cosmological Constraint Precision of the Photometric and Spectroscopic Multi-probe Surveys of China Space Station Telescope (CSST)*, 2206.09822.

[59] H. Lin, Y. Gong, X. Chen, K. C. Chan, Z. Fan and H. Zhan, *Forecast of neutrino cosmology from the CSST photometric galaxy clustering and cosmic shear surveys*, *Mon. Not. Roy. Astron. Soc.* **515** (2022) 5743–5757, [2203.11429].

[60] V. Springel, *The Cosmological simulation code GADGET-2*, *Mon. Not. Roy. Astron. Soc.* **364** (2005) 1105–1134, [astro-ph/0505010].

[61] M. Crocce, S. Pueblas and R. Scoccimarro, *Transients from Initial Conditions in Cosmological Simulations*, *Mon. Not. Roy. Astron. Soc.* **373** (2006) 369–381, [astro-ph/0606505].

[62] DES collaboration, L. F. Secco et al., *Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to modeling uncertainty*, *Phys. Rev. D* **105** (2022) 023515, [2105.13544].

[63] E. Krause, T. Eifler and J. Blazek, *The impact of intrinsic alignment on current and future cosmic shear surveys*, *Mon. Not. Roy. Astron. Soc.* **456** (2016) 207–222, [1506.08730].

[64] DES collaboration, M. Gatti et al., *Dark energy survey year 3 results: weak lensing shape catalogue*, *Mon. Not. Roy. Astron. Soc.* **504** (2021) 4312–4336, [2011.03408].

[65] DES collaboration, A. Porredon et al., *Dark Energy Survey Year 3 results: Optimizing the lens sample in a combined galaxy clustering and galaxy-galaxy lensing analysis*, *Phys. Rev. D* **103** (2021) 043503, [2011.03411].

[66] D. N. Limber, *The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II*, *Astrophys. J.* **119** (1954) 655.

[67] M. LoVerde and N. Afshordi, *Extended Limber Approximation*, *Phys. Rev. D* **78** (2008) 123506, [0809.5112].

# A direct probe of Λ potential in nuclear medium

Gao-Chan Yong[1,2]
[1] *Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730000, China*
[2] *School of Nuclear Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China*

Using the Liège intranuclear-cascade model together with the ablation model ABLA, an investigation is conducted into the effects of Λ potential in Λ-nucleus and Λ-hypernucleus-nucleus collisions across various beam energies. The findings show that the angle and transverse-momentum distributions of scattered Λ hyperon, the scattering cross section of the Λ hyperon in Λ-nucleus collisions as well as the rapidity distribution of Λ hyperon in Λ-hypernucleus-nucleus collisions are significantly influenced by the strength of the Λ potential in these scattering reactions across various beam energies. These demonstrations, unhindered by the uncertainties of Λ and hypernuclei productions in nuclear medium, allow for a direct investigation of the Λ potential, especially its momentum dependence. The criticality of probing the Λ potential is closely associated with the resolution of the "hyperon puzzle" in neutron stars.

arXiv:2402.18831v1 [nucl-th] 29 Feb 2024

Neutron stars emerge from the ashes of supernova explosions and were initially believed to predominantly consist of neutrons. Nonetheless, a sequence of theoretical studies has suggested that neutron stars might also harbor strange matter [1]. The concept of strangeness in neutron stars could profoundly impact our comprehension of these celestial bodies [2]. This aspect might influence their internal composition and cooling mechanisms, as highlighted in various studies [3–7]. Additionally, the presence of strangeness could alter the dynamics of neutron star mergers [8], potentially leading to distinct gravitational wave signatures [9]. The softening of the equation of state for dense matter, attributed to strangeness, could play a pivotal role in the shock-wave/neutrino-delayed-shock process, thereby affecting the explosive birth of core-collapse supernovae [10–13]. The presence of strangeness within neutron star matter has thus ignited the curiosity of the physics community [14–21], promising to shed light on the strong nuclear force and the characteristics of matter in ultra-extreme conditions, as further explored in references [19, 21–24]. Consequently, the study of hyperon-nucleon interactions and the significance of strangeness in neutron stars arises as a vital domain for astrophysicists, together with particle and nuclear physicists.

The resolution of the "hyperon puzzle", focusing on the inclusion of strangeness within the nuclear medium to ascertain the strong nuclear force, remains a fundamental query [22]. A variety of theoretical approaches have been adopted, such as utilizing nuclear many-body theories for calculating the properties of strange particles in dense environments [18, 24], applying effective field theories to elucidate the interactions between strange particles and nucleons [25], using perturbative quantum chromodynamics for the study of strange quark matter [26, 27]. Astrophysical observations of neutron stars also serve as a means to deduce the characteristics of strange particles within [28, 29].

Despite these methodologies, interactions between strangeness and non-strangeness in nuclear matter confront considerable theoretical ambiguities and are seldom directly examined through nuclear experiments in terrestrial labs. In Ref. [30], it is argued that combining studies of transport models with nuclear experimental data from facilities around the world could be a better way to address the "hyperon puzzle". Therefore, nuclear reactions may offer a more effective method for studying hyperon-nucleon interactions within nuclear medium: Hyperon-nucleus scattering experiments deal with hyperon-nucleon interactions in medium around saturation density, whereas hypernucleus-nucleus collisions examine these interactions at higher densities. By adjusting the beam energy, one can also explore the momentum dependence of hyperon-nucleon interactions in nuclear medium. In this study, the Liège intranuclear-cascade (INCL) model, combined with the ABLA deexcitation code, is employed to investigate the effects of the Λ potential around and above normal nuclear density. It's demonstrated that experiments involving Λ-nucleus and hypernucleus-nucleus collisions could directly unveil the Λ-nucleon interactions in nuclear medium.

The Liège intranuclear-cascade code, also known as the INCL model [31–33], is used to describe the collision between a projectile (such as nucleons, pions, hyperons and light ions) and a target nucleus. The INCL model incorporates classical physics principles, but also includes some quantum-mechanical features (such as Fermi motion, realistic space densities and Pauli blocking) to account for the initial conditions and dynamics of the collision. The INCL model treats an energy and isospin-dependent nucleon potential and an isospin-dependent constant hyperon potential with Woods-Saxon density distributions [32, 34, 35]. The model treats nuclear collisions as successive relativistic binary hadron-hadron collisions, where the positions and momenta of the hadrons are tracked over time. The extended INCL model includes the production of pion mesons and strange particles [36, 37]. The latest version of the model, INCL++6.32, includes the formation of hyperremnants [38, 39]. Based on different kinds of conservation laws (such as for baryon number, charge, energy, momentum, and angular momentum), the model predicts the formation of hot hyperremnants and characterizes them in terms of atomic and mass numbers, strangeness number,

excitation energy, and angular momenta.

On the other hand, the ABLA07 model [40] is a deexcitation model that describes the subsequent de-excitation processes after the collision. It describes the de-excitation of an excited remnant by emitting $\gamma$ rays, neutrons, light-charged particles, and intermediate-mass fragments (IMFs) or fission decays. The ABLA07 model was rewritten in C++ from its original FORTRAN version to be compatible with the GEANT4 simulation framework. The ABLA07 model is considered to be one of the best model for describing the deexcitation of hot remnants, according to benchmarks performed by the International Atomic Energy Agency [40]. Recently, the ABLA07 model has been extended to include the emission of hyperons and the formation of cold hypernuclei [39, 41]. Overall, the INCL model is used to describe the collision stage of the reactions, while the updated ABLA model is used to describe the subsequent deexcitation stage. The combination of the two models allows for the study of hyperon-related scatterings and provides insights into the properties of hyperon in nuclear matter.

Scattering experiments play a pivotal role in elucidating the complex interactions among particles. These experiments provide a profound understanding of particle characteristics and find wide applications in numerous scientific fields, including nuclear physics and quantum mechanics. In particular, scattering experiments are integral to deriving the optical potential, a critical element in the field of nuclear physics. The optical potential, usually connected to the potential energy of participant particles, foretells the particle movement within a quantum system. This potential can be extracted through a careful examination of the trajectories and speeds of both outgoing and incoming particles during scattering experiments. Indeed, the $\Lambda$-nucleon potential can be scrutinized through the production of hypernuclei, which helps in gauging the *momentum-independent* $\Lambda$ potential in medium around saturation density [35]. Likewise, studying the directed flows of hyperons or hypernuclei provides insights into the $\Lambda$ potential in medium at high densities [42]. Notwithstanding, such research poses prominent theoretical uncertainties, primarily linked to the production of $\Lambda$ and the assembly of hypernuclei.

Figure 1 illustrates the impact of the $\Lambda$ potential on the angular distributions of $\Lambda$ emissions in $\Lambda$+$^{40}$Ca scatterings across various $\Lambda$-beam energies. The $^{40}$Ca target (together with the latter used $^{16}$O) ensures the absence of isospin effects from the $\Lambda$ potential. Given that the $\Lambda$ potential can transition from a notably negative value at very low momenta to a positive one at high momenta around saturation density, I modify the $\Lambda$ potential value from the initially estimated -30 MeV at saturation density [35, 39] to a span of -50 to -20 MeV at normal density for illustrative purposes. At high densities, the $\Lambda$ potential might turn positive. To maintain simplicity, this range is still employed to demonstrate the effects of the $\Lambda$ potential. Observations indicate that at narrow polar angles, the influence of the $\Lambda$ potential on $\Lambda$ emissions

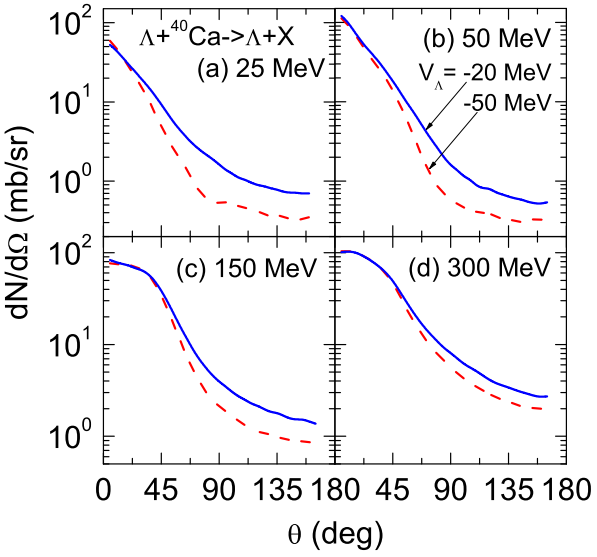



FIG. 1: Angle distributions of $\Lambda$ emissions in $\Lambda$+$^{40}$Ca scatterings under varying $\Lambda$ potentials and beam energies of 25 (a), 50 (b), 150 (c), 300 (d) MeV.

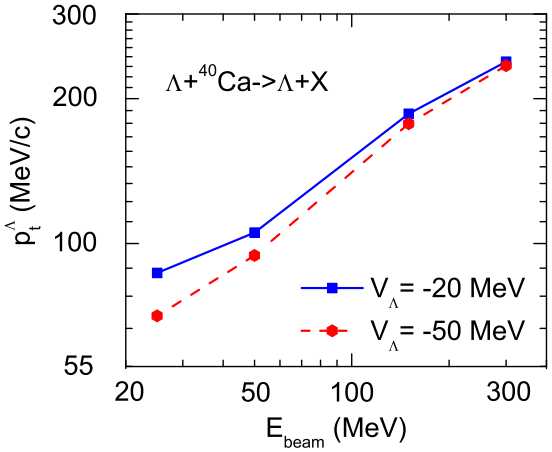



FIG. 2: Impact of the $\Lambda$ potential on the transverse momentum of $\Lambda$ in $\Lambda$+$^{40}$Ca scatterings at varying beam energies.

in $\Lambda$+$^{40}$Ca scatterings is subtle. However, at wider polar angles ($\theta > 60$), $\Lambda$ emissions exhibit pronounced effects of the $\Lambda$ potential. The less attractive the $\Lambda$ potential for $\Lambda$ hyperons (the larger the $\Lambda$ potential), the greater the number of $\Lambda$ emissions observed at these wider angles. Consequently, more $\Lambda$ emissions are observed with a $\Lambda$ potential of -20 MeV compared to -50 MeV. As the beam energy increases, the significance of the $\Lambda$ potential's effects diminishes progressively. Given that the $\Lambda$ potential may also be momentum-dependent, the beam-energy-dependent analysis presented here could serve as a means to investigate the momentum dependence of the $\Lambda$ potential in nuclear medium.

Considering the pronounced impact of the $\Lambda$ potential on $\Lambda$ emissions at large polar angles, it becomes signifi-

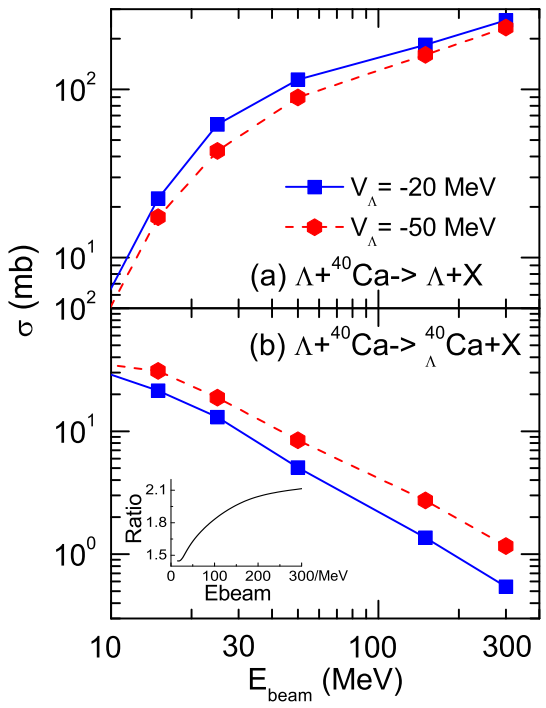



FIG. 3: Total cross sections for Λ (a) and hypernucleus $^{40}_{\Lambda}$Ca (b) productions in Λ+$^{40}$Ca scatterings across various beam energies and Λ potentials. The inset shows the ratio (Ratio = $\sigma^{\Lambda^{40}Ca}_{-50}/\sigma^{\Lambda^{40}Ca}_{-20}$) of $^{40}_{\Lambda}$Ca production cross sections with the two Λ potentials as a function of beam energy.

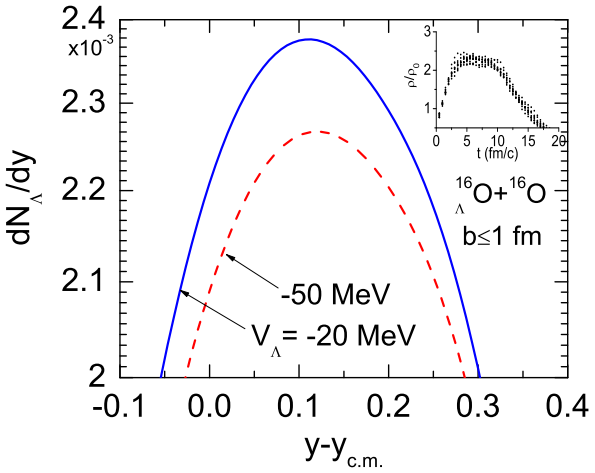



FIG. 4: Rapidity distributions for Λ production in central $^{16}_{\Lambda}$O+$^{16}$O collisions at 300 MeV/nucleon under various Λ potentials. The inset illustrates the central maximum compression density as a function of reaction time.

cant to explore its influence on Λ transverse momentum in Λ+$^{40}$Ca scatterings. Figure 2 illustrates the behavior of Λ transverse momentum under various beam energies. As anticipated, there is an observable increase in Λ transverse momentum corresponding to the rise in beam energies. Concurrently, the influence of the Λ potential on this transverse momentum diminishes as the beam energies escalate. This trend is logical, given that at higher energies, the Λ potential exerts a less noticeable effect on Λ+$^{40}$Ca scatterings. Additionally, Figure 2 highlights that an increase in the Λ potential (rendering it less attractive) leads to higher Λ transverse momenta. Hence, the Λ transverse-momentum spectra could serve as a viable tool for assessing the strength of the Λ potential in Λ+$^{40}$Ca scatterings.

To investigate the momentum dependence of the Λ potential around saturation density, examining the beam-energy dependent Λ scattering cross section in the Λ+$^{40}$Ca reaction is one approach. Panel (a) of Figure 3 presents the Λ scattering cross section across various beam energies and Λ potentials. It's observed that the total cross sections for Λ scatterings escalate as the beam energy increases. This increase is attributed to the diminished capturing capability of the Λ potential for higher energy Λ particles in the $^{40}$Ca target. The influence of the Λ potential on the total cross sections of Λ scatterings becomes less significant as the beam energy rises. A larger (less attractive) Λ potential leads to higher total Λ production cross sections, aligning with the findings depicted in Figure 1.

Panel (b) of Figure 3 illustrates the total cross sections of hypernucleus $^{40}_{\Lambda}$Ca production as a function of beam energy. Given that the total cross sections for Λ scatterings rise with increasing beam energy, an inverse trend in the total cross sections of hypernucleus $^{40}_{\Lambda}$Ca production with increasing beam energy is observed, which is logical. Intriguingly, the impact of the Λ potential on the total cross sections of hypernucleus $^{40}_{\Lambda}$Ca production intensifies with increasing beam energy, which can be clearly seen from the inset of Panel (b). This is because a higher energy Λ particle requires a greater number of scatterings – resulting in a longer duration within the target and a heightened influence from the Λ potential – to decelerate and eventually be captured by the Λ potential in the target $^{40}$Ca. Hence, the total cross sections of hypernucleus $^{40}_{\Lambda}$Ca production in Λ+$^{40}$Ca scatterings serve as a valuable means to explore the momentum-dependent Λ potential around saturation density.

The investigation into the Λ hyperon potential in dense nuclear environments is critically important for understanding neutron stars, especially when it comes to solving the "hyperon puzzle". This puzzle arises from the contradiction between theoretical models, which predict that hyperons would soften the equation of state (EoS), leading to the formation of lower mass neutron stars, and the recent observations of neutron stars with significantly higher masses. Research focused on the role of the Λ hyperon seeks to align the EoS with the existence of these more massive neutron stars. Therefore, exploring the high-density Λ potential sheds light on the behavior of matter under extreme conditions and could be key to resolving the hyperon puzzle.

To replicate high-density nuclear matter in terrestrial labs, heavy-ion collisions are frequently employed. Figure 4 illustrates the central $^{16}_{\Lambda}$O+$^{16}$O collisions at 300

MeV/nucleon under varying Λ potentials. The inset reveals that dense matter, approximately twice the saturation density, is achieved in these central collisions at 300 MeV/nucleon. The figure shows how the Λ hyperon's rapidity distributions in these collisions are significantly influenced by the Λ potential. A larger (less attractive) Λ potential results in increased Λ emissions at mid-rapidities. Given the Λ hyperon's presence exclusively in the projectile, it predominantly appears in the positive-rapidity region. Should the Λ potential be positive in high-density nuclear environments [43], the impact on the rapidity distributions of Λ production would be markedly greater.

The reaction involving Λ or Λ-hypernuclei can be realized through secondary-beam experiments in terrestrial labs. Given that current studies bypass the generations of Λ-hyperons or Λ-hypernuclei in the medium, theoretical uncertainties are substantially minimized.

To summarize, utilizing the Liège intranuclear-cascade model in tandem with the ABLA deexcitation model, this research delves into the dynamics of Λ-nucleus and hypernucleus-nucleus interactions. It aims to uncover how the Λ potential influences Λ particles in nuclear medium. Findings indicate that in the Λ+$^{40}$Ca collisions, Λ emissions are markedly influenced by the Λ potential, particularly at wide polar angles. Moreover, the impact of the Λ potential on the Λ particle's transverse momentum is distinctly observable. While the influence of the Λ potential on total cross sections of Λ scatterings diminishes as the beam energy escalates, its effect on the total cross sections for the production of the hypernucleus $^{40}_{\Lambda}$Ca intensifies with increasing beam energy. The rapidity distributions of Λ production in central $^{16}_{\Lambda}$O+$^{16}$O collisions at 300 MeV/nucleon significantly reflect the Λ potential's effects, offering a potential method to probe the Λ potential at high densities. This investigation aims at directly probing the Λ potential in the nuclear medium and addressing the "hyperon puzzle" in neutron stars.

The author thanks J. L. Rodríguez-Sánchez, J. Cugnon, J.-C. David and D. Mancusi for helpful communications. This work is supported by the National Natural Science Foundation of China under Grant Nos. 12275322, 12335008.

[1] F. Weber, Prog. Part. Nucl. Phys. **54**, 193 (2005).
[2] N. K. Glendenning, Phys. Rep. **342**, 393 (2001).
[3] F. Anzuini, A. Melatos, C. Dehman, D. Viganò, J. A. Pons, Mon. Not. Roy. Astron. Soc. **509**, 2609 (2021).
[4] J. A. Pons, S. Reddy, M. Prakash, J. M. Lattimer, and J. A. Miralles, Astrophys. J. **513**, 780 (1999).
[5] M. Prakash, M. Prakash, J. M. Lattimer, and C. J. Pethick, Astrophys. J. **390**, L77 (1992).
[6] Adriana R. Raduta, Jia Jie Li, Armen Sedrakian, Fridolin Weber, Mon. Not. Roy. Astron. Soc. **487**, 2639 (2019).
[7] Adriana R. Raduta, Armen Sedrakian, Fridolin Weber, Mon. Not. Roy. Astron. Soc. **475**, 4347 (2018).
[8] Mark G. Alford and Alexander Haber, Phys. Rev. C **103**, 045810 (2021).
[9] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **119**, 161101 (2017).
[10] Adriana R. Raduta, Eur. Phys. J. A **58**, 115 (2022).
[11] Armen Sedrakian, Arus Harutyunyan, Eur. Phys. J. A **58**, 137 (2022).
[12] Adriana R. Raduta, Micaela Oertel, Armen Sedrakian, Mon. Not. Roy. Astron. Soc. **499**, 914 (2020).
[13] Shmuel Balberg, Avraham Gal, Nucl. Phys. A **625**, 435 (1997).
[14] N. K. Glendenning, Astrophys. J. **293**, 470 (1985).
[15] J. Schaffner, I. N. Mishustin, Phys. Rev. C **53**, 1416 (1996).
[16] Norman K. Glendenning, Phys. Lett. B **114**, 392 (1982).
[17] Diego Lonardoni, Alessandro Lovato, Stefano Gandolfi, and Francesco Pederiva, Phys. Rev. Lett. **114**, 092301 (2015).
[18] D. Chatterjee, and I. Vidaña, Eur. Phys. J. A **52**, 1 (2016).
[19] Wei-Zhou Jiang, Bao-An Li, and Lie-Wen Chen, Astrophys. J. **756**, 56 (2012).
[20] D. Gerstung, N. Kaiser, W. Weise, Eur. Phys. J. A **56**, 175 (2020).
[21] A. Sedrakian, J. J. Li, F. Weber, Prog. Part. Nucl. Phys. **131**, 104041 (2023).
[22] I. Vidaña, Proc. R. Soc. Lond. A **474**, 0145 (2018).
[23] Xiangdong Sun, Zhiqiang Miao, Baoyuan Sun, Ang Li, Astrophys. J. **942**, 55 (2023).
[24] L. Tolos, L. Fabbietti, Prog. Part. Nucl. Phys. **112**, 103770 (2020).
[25] J. M. Alarcón, L. S. Geng, J. Martin Camalich, J. A. Oller, Phys. Lett. B **730**, 342 (2014).
[26] Fridolin Weber, Prog. Part. Nucl. Phys. **54**, 193 (2005).
[27] Cheng-Jun Xia, Shan-Gui Zhou, Nucl. Phys. B **916**, 669 (2017).
[28] Wen Hui Long, Bao Yuan Sun, Kouichi Hagino, and Hiroyuki Sagawa, Phys. Rev. C **85**, 025806 (2012).
[29] Thomas Hell and Wolfram Weise, Phys. Rev. C **90**, 045801 (2014).
[30] Gao-Chan Yong, Phys. Rev. D **108**, L091507 (2023).
[31] A. Boudard, J. Cugnon, S. Leray, and C. Volant, Phys. Rev. C **66**, 044615 (2002).
[32] A. Boudard, J. Cugnon, J.-C. David, S. Leray and D. Mancusi, Phys. Rev. C **87**, 014606 (2013).
[33] D. Mancusi, A. Boudard, J. Cugnon, J.-C. David, P. Kaitaniemi and S. Leray, Phys. Rev. C **90**, 054602 (2014).
[34] J. L. Rodríguez-Sánchez, Jean-Christophe David, Davide Mancusi, Alain Boudard, Joseph Cugnon, and Sylvie Leray, Phys. Rev. C **96**, 054602 (2017).
[35] J. L. Rodríguez-Sánchez, J.-C. David, J. Hirtz, J. Cugnon, and S. Leray, Phys. Rev. C **98**, 021602(R) (2018).
[36] S. Pedoux and J. Cugnon, Nucl. Phys. A **866**, 16 (2011).
[37] J. Hirtz, J.-C. David, A. Boudard, J. Cugnon, S. Leray, I. Leya, D. Mancusi, Eur. Phys. J. Plus **133**, 436 (2018).
[38] J. Hirtz, J.-C. David, A. Boudard, J. Cugnon, S. Leray,

I. Leya, J. L. Rodríguez-Sánchez, and G. Schnabel, Phys. Rev. C **101**, 014608 (2020).

[39] J. L. Rodríguez-Sánchez, J. Cugnon, J.-C. David, J. Hirtz, A. Kelić-Heil, S. Leray, Phys. Rev. C **105**, 014623 (2022).

[40] A. Kelić, M. V. Ricciardi, and K.-H. Schmidt, in *Proceedings of Joint ICTP-IAEA Advanced Workshop on Model Codes for Spallation Reactions, ICTP Trieste, Italy, 4-8 February 2008*, edited by D. Filges, S. Leray, Y. Yariv, A. Mengoni, A. Stanculescu, and G. Mank, (IAEA INDC, Vienna, 2008), pp. 181-221.

[41] J. L. Rodríguez-Sánchez, J. Cugnon, J.-C. David, J. Hirtz, A. Kelić-Heil, and I. Vidaña, Phys. Rev. Lett. **130**, 132501 (2023).

[42] B. E. Aboona *et al.* (STAR Collaboration), Phys. Rev. Lett. **130**, 212301 (2023).

[43] Asanosuke Jinno, Koichi Murase, Yasushi Nara, and Akira Ohnishi, Phys. Rev. C **108**, 065803 (2023).

arXiv:2402.18727v1 [hep-th] 28 Feb 2024

# Relativistic fluctuations in stochastic fluid dynamics

*Xin* An[1,*], *Gökçe* Başar[2,**], *Mikhail* Stephanov[3,***], and *Ho-Ung* Yee[3,****]

[1]National Centre for Nuclear Research, 02-093 Warsaw, Poland
[2]Department of Physics and Astronomy, University of North Carolina, Chapel Hill, North Carolina 27599, USA
[3]Department of Physics, University of Illinois, Chicago, Illinois 60607, USA

**Abstract.** The state-of-the-art theoretical formalism for a covariant description of non-Gaussian fluctuation dynamics in relativistic fluids is discussed.

## 1 Introduction

Fluctuating hydrodynamics is a powerful tool for exploring complex and critical phenomena in non-equilibrium systems that possess a small number of degrees of freedom. Such scenarios are realized in relativistic heavy-ion collisions where a few thousand particles are produced, especially when the hypothetical QCD critical point is approached. Therefore, fully establishing the framework for relativistic fluctuating hydrodynamics is necessary for interpreting observables in experiments that are sensitive to fluctuations and criticality. Developing a covariant description of non-Gaussian fluctuation dynamics in stochastic fluids is a crucial step toward achieving this ambitious goal.

In this proceeding, we will briefly formulate the essential theoretical development of deterministic fluctuating hydrodynamics from a general perspective. More specifically, we will begin by reviewing the results of Ref. [1] in Sec. 2.1 and 2.2, where we presented the covariant formalism for the non-equilibrium evolution of non-Gaussian fluctuations in relativistic fluids. It on one hand follows the approach used in earlier work [2] and [3], where the covariant dynamical description was developed but only for Gaussian fluctuations, and on the other hand extends the subsequent work [4], where a generic formalism for non-Gaussian fluctuation dynamics was established, yet it has not been implemented covariantly. In Sec. 2.3, we will illustrate how the number of independent $n$-point correlation functions can be significantly reduced, upon the use of certain approximation which facilitates easier numerical implementation. We shall keep the formulation as general as possible, allowing it to be applied to arbitrary $n$ and regimes not relying on the separation of relaxation time scales.

## 2 Theoretical framework

### 2.1 Fluctuation evolution equations

We start from the *covariant* Langevin equation for a set of stochastic fields $\psi_i$ (such as conserved quantities including charge density and energy-momentum density) where the

*e-mail: xin.an@ncbj.gov.pl
**e-mail: gbasar@unc.edu
***e-mail: misha@uic.edu
****e-mail: hyee@uic.edu

subscript $i$ labels different fields:

$$u \cdot \partial \psi_i = F_i + \xi_i \,, \tag{1}$$

where $F_i$ is the drift force, $\xi_i = H_{ij}\eta_j$ is the multiplicative noise whose amplitude $H_{ij}$ is related to the Onsager matrix via $Q_{ij} \equiv H_{ik}H_{kj}$, with a Gaussian form $\langle \eta_i(x_1)\eta_j(x_2)\rangle = 2\delta_{ij}\delta^{(4)}(x_1 - x_2)$. One should keep in mind that the velocity $u \equiv u(\psi_i)$ may or may not be chosen as one of the independent variables in $\psi_i$.

The importance of fluctuations can be studied by the multi-point correlation functions. The "raw" $n$-point correlation functions are defined by $G_n \equiv G_{i_1\ldots i_n} \equiv \langle \phi_{i_1}(x_1)\ldots\phi_{i_n}(x_n)\rangle$ where $\phi = \psi - \langle\psi\rangle$. The connected correlation functions $G^{\rm c}_n \equiv G^{\rm c}_{i_1\ldots i_n}$, which are directly related to the experimental observables (such as cumulants of particle multiplicity distribution), can be obtained via the relation

$$G^{\rm c}_{i_1\ldots i_n} = \sum_{k=1}^{n} \sum_{\{n_1,\ldots,n_k\}} (-1)^{k-1}(k-1)!\{n_1,\ldots,n_k\}!\, G_{\underbrace{i_1\ldots i_{n_1}}_{n_1}} G_{\underbrace{i_{n_1+1}\ldots i_{n_1+n_2}}_{n_2}} \ldots G_{\underbrace{i_{n-n_k+1}\ldots i_n}_{n_k}}\Big|_{\overline{1\ldots n}}, \tag{2}$$

where $\{n_1,\ldots,n_k\}! \equiv n!/n_1!\ldots n_k!\,k_1!\ldots k_n!$, the inner sum is over all ordered sets of integer numbers $\{n_1,\ldots,n_k\}$, such that $n_1 \leqslant n_2 \leqslant \cdots \leqslant n_k$ and $n_1 + \cdots + n_k = n$. Each set describes a partition of the $n$ indices $i_1,\ldots,i_n$ into $k$ groups in a way that each term in the sum in Eq. (2) is different.

Using Eqs. (1) and (2), it shall be straightforward to derive the evolution equation for $G^{\rm c}_n$. However, this system of equations is rather complicated and not in a closed form. Following Ref. [4] we expand and truncate these equations in two independent small parameters, $\varepsilon$ and $\varepsilon_q$, controlling the loop and gradient expansion respectively. The resulting power counting is specified as

$$G^{c}_{i_1\ldots i_n} \sim \varepsilon^{n-1}\,, \quad L_{i_1,i_2\ldots i_n} \sim \varepsilon_q + O(\varepsilon_q^2)\,, \quad Q_{i_1 i_2,\,j_1\ldots j_m} \sim \varepsilon_q^2\,\varepsilon\,, \quad u\cdot\partial \sim \varepsilon_q^2\,, \tag{3}$$

where indices following comma "," denote the derivative with respect to the fields $\psi$. With the help of Eq. (3), one finds the generic evolution equations for $n$-point connected correlation function at leading order $\varepsilon_q^2\,\varepsilon^{n-1}$:

$$\begin{aligned} u\cdot\partial^{(x)} G^{\rm c}_{i_1\ldots i_n}(x_1,\ldots,x_n) = n\Big[ &-(y_1\cdot\partial u)\cdot\frac{\partial}{\partial x_1} G^{\rm c}_{i_1\ldots i_n} \\ &+ \sum_{k=1}^{n-1} \sum_{\substack{\{n_1,\ldots,n_k\}\\ n_1+\cdots+n_k=n-1}} \{n_1,\ldots,n_k\}!\, L_{i_1,\,j_1\ldots j_k} G^{\rm c}_{j_1\,\underbrace{i_2\ldots}_{n_1}} \ldots G^{\rm c}_{j_k\,\underbrace{\ldots i_n}_{n_k}} \\ &+ (n-1)\sum_{k=0}^{n-2} \sum_{\substack{\{n_1,\ldots,n_k\}\\ n_1+\cdots+n_k=n-2}} \{n_1,\ldots,n_k\}!\, Q_{i_1 i_2,\,j_1\ldots j_k} G^{\rm c}_{j_1\,\underbrace{i_3\ldots}_{n_1}} \ldots G^{\rm c}_{j_k\,\underbrace{\ldots i_n}_{n_k}} \Big]_{\overline{1\ldots n}}, \end{aligned} \tag{4}$$

where the time derivative is taken in the rest frame of the fluid at midpoint $x \equiv \sum_{i=1}^n x_i/n$, $y_i \equiv x_i - x$ is the separation vector, $[\ldots]_{\overline{1\ldots n}}$ is the "averaged" permutation over indices labeled by $i$, and

$$L_{i,\,j_1 j_2\ldots j_n} = F_{i,\,j_1 j_2\ldots j_n} - u^{\mu}_{,\,j_1\ldots j_n}(\partial_\mu\psi_i) - \Big[n\delta_{ij_1} u^{\mu}_{,\,j_2\ldots j_n}\partial^{(j_1)}_{\mu}\Big]_{\overline{1\ldots n}} \tag{5}$$

is a multilinear operator where the last two terms are consequence of $u = u(\psi_i)$. The space-time arguments for $L$, $Q$ and $G^{\rm c}$, associated with the indices of those quantities, are suppressed on the right hand side. Eq. (4) can be more straightforwardly represented in diagrams in Fig. 1. These diagrams can be categorized into two groups, one from drift force and the other from the noise, corresponding to the second and third line in Eq. (4) respectively.

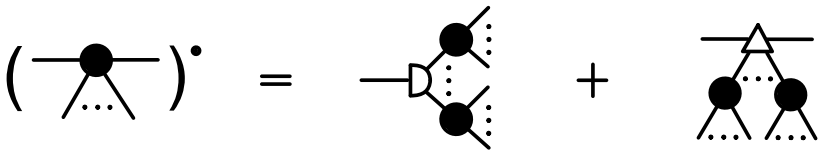


**Figure 1.** Diagrammatic representation of the evolution equations for generic multipoint connected correlation functions, with all possible combinatorial arrangements at tree level.

### 2.2 Confluent formalism

In Sec. 2.1 we have adopted an alternative strategy presented in Refs. [2] and [3] but generalize it to non-Gaussian fluctuations. This approach is more suitable to derive the Lorentz covariant fluctuation evolution equations. In addition to rewriting the equation in a covariant form with quantities measured in the local rest frame (like how it shall be done for a one-point function), we also need to implement it for multi-point equal-time functions where the confluent formalism is necessary to covariantly describe fluctuating fields at different points.

In the confluent formalism we first introduce the confluent $n$-point connected function

$$\bar{G}^{\rm c}_n \equiv \bar{G}^{\rm c}_{i_1\ldots i_n}(x_1,\ldots,x_n) \equiv \Lambda_{i_1}^{\ j_1}(x_1-x)\ldots\Lambda_{i_n}^{\ j_n}(x_n-x)G^{\rm c}_{j_1\ldots j_n}(x_1,\ldots,x_n)\,, \tag{6}$$

where $\Lambda(x_i-x)$ is a Lorentz boost defined by $\Lambda(\Delta x)u(x+\Delta x)=u(x)$. In Eq. (6) fluctuations at different points are boosted into the same frame (chosen as the local rest frame at the midpoint $x$). The separation four-vector $y_i(x) \equiv x_i - x$ can be expressed in terms of the local tetrad consisting of vector $u(x)$ and triad $e_a(x)$ with $a,b=1,2,3$: $y(x)=e_a(x)y^a$ where we have imposed the equal-time constraint $u\cdot y=0$. The generalized multi-point Wigner transform is then defined as

$$W_n(x;\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) = \int\left[\prod_{i=1}^n d^3y_i^a\, e^{-iq_{ia}y_i^a}\right]\delta^{(3)}\left(\frac{1}{n}\sum_{i=1}^n y_i^a\right)\bar{G}^{\rm c}_n(x+e_ay_1^a,\ldots,x+e_ay_n^a)\,, \tag{7}$$

where $\boldsymbol{q}=\{q^a\}$ is the wavenumber conjugate to $y^a$. This definition eliminates the dependence on $x$ of the transform kernel, making it convenient for a covariant theory. It is easy to check that the derivative $\mathring{\partial}$ commutes with the Wigner transform, and $\partial_a^{(y_i)} \to iq_{ia}$.

As the midpoint moves, only the changes respect the local rest frame of the midpoint are measured, not the changes resulted from the change of frame itself, we thus boost $G^{\rm c}_n$ to the frame before the movement to eliminate the kinematic change, at the same time keep the equal-time constraint to preserve the relative positions of the $n$ different points. Taking this two issues into account the confluent derivative is given by

$$\bar{\nabla}_\mu W_{i_1\ldots i_n} \equiv \mathring{\partial}_\mu W_{i_1\ldots i_n} - n\left(\bar{\omega}^{j_1}_{\mu i_1}W_{j_1 i_2\ldots i_n} - \mathring{\omega}^a_{\mu b}q_{1a}\partial_b^{(q_1)}W_{i_1\ldots i_n}\right)_{\overline{1\ldots n}}\,, \tag{8}$$

where $\bar{\omega}^j_{\mu i} = u_i\partial_\mu u^j - u^j\partial_\mu u_i$ and $\mathring{\omega}^a_{\mu b} = e^a_\nu\partial_\mu e^\nu_b$ are connections, $\mathring{\partial}$ is a derivative with respect to $x$ with $q_a$ fixed, while $\partial_a^{(q)}$ is a derivative with respect to $q_a$ with $x$ fixed.

The evolution equation of Wigner functions can be obtained by projecting Eq. (8) along $u(x)$ and using Eqs. (4) and (7):

$$\begin{aligned} u\cdot\bar{\nabla}W_{i_1\ldots i_n}(x;\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) = \int\left[\prod_{i=1}^n d^3y_i\, e^{-iq_{ia}y_i^a}\right]\delta^{(3)}\left(\frac{1}{n}\sum_{i=1}^n y_i\right)\Big\{u\cdot\partial G^{\rm c}_{i_1\ldots i_n} \\ - n\left[\left(u^\mu\bar{\omega}^\nu_{\mu\lambda}y_1^\lambda\partial_\nu^{(y_1)}\delta^{j_1}_{i_1} + y_1^\mu\bar{\omega}^{j_1}_{\mu i_1}u\cdot\partial + u^\mu\bar{\omega}^{j_1}_{\mu i_1}\right)G^{\rm c}_{j_1 i_2\ldots i_n}\right]_{\overline{1\ldots n}}\Big\} = \mathcal{P}[\{W_2,\ldots,W_n\}]\,, \end{aligned} \tag{9}$$

where in the last step one needs to perform the inverse Wigner transform of Eq. (7) in order to obtain a set of local evolution equations for $W_n$, whose explicit form depends on Eq. (4) that is to be substituted into Eq. (9).

### 2.3 Rotating-wave approximation

Although the evolution equations appear diagrammatically simple as in Fig. 1, the number of equations rapidly increases with the number of independent components of fluctuating fields, and the equations themselves also become more complicated. This complexity makes the numerical implementation challenging when considering additional hydrodynamic fields or higher-point functions. It is, therefore, instrumental to apply a method analogous to the rotating-wave approximation, under which the modes with rapid oscillation are averaged out.

In this approximation, we first change fluctuation variables from the set we are using, say $\phi_i$, to a new set $\Phi_i$ defined in such a way that the linear operator $L$ appearing in the linearized ideal hydrodynamic equations is diagonal under linear transform $\phi = U\Phi$:

$$u\cdot\partial\,\phi_i = L_{ij}\,\phi_j \quad \rightarrow \quad u\cdot\partial\,\Phi_i = \mathbb{L}_{ij}\Phi_j\,, \quad \mathbb{L}_{ij} = (U^{-1}LU)_{ij} = \lambda_i\delta_{ij}\,, \tag{10}$$

where $\lambda_i$ are the eigenvalues of $L$, the transformation matrix $U$ can be constructed from the eigenvectors of the operator $L$, since for each index $j$ the vector with $i's$ component given by $U_{ij}$ is the *left* eigenvector of $L$: $\sum_i L_{ki}U_{ij} = \lambda_j U_{kj}$. The Wigner function transforms as $W_n \equiv W_{i_1\ldots i_n} = U_{i_1 j_1}\ldots U_{i_n j_n}\mathbb{W}_{j_1\ldots j_n} \equiv (U)^n\mathbb{W}_n$ accordingly, and Eq. (9) becomes

$$u\cdot\bar{\nabla}\mathbb{W}_n = \left((U)^n u\cdot\bar{\nabla}(U^{-1})^n\right)\mathbb{W}_n + (U^{-1})^n\,\mathcal{P}[\{(U)^2\mathbb{W}_2,\ldots,(U)^n\mathbb{W}_n\}] = \left(\sum_{m=1}^{n}\lambda_{i_m}\right)\mathbb{W}_n + \ldots \tag{11}$$

If $\lambda_{i_1}+\ldots\lambda_{i_n}\neq 0$, $\mathbb{W}_n \equiv \mathbb{W}_{i_1\ldots i_n}$ is identified as a fast-oscillating mode that can be averaged out at a timescale much longer than the oscillation period. Consequently, we are left with only slow modes (with $\lambda_{i_1}+\ldots\lambda_{i_n}=0$) whose evolution is described by relaxation-type equations.

## 3 Summary

In this proceeding we highlighted the essential ideas in the recent theoretical development of fluctuating hydrodynamics from the bottom-up approach. It consists of three ingredients: a closed system of equations for arbitrary $n$-point correlation functions, a formalism to describe these equations covariantly in relativistic hydrodynamics, and a strategy to reduce the number of correlation functions and simplify the resulting equations. The explicit full equations for the correlation functions that are relevant for heavy-ion collisions especially for the critical point search, as well as more detailed discussions will be presented in Ref. [5].

## References


[1] X. An, G. Başar, M. Stephanov, and H.-U. Yee, Phys. Rev. C **108**, 034910 (2023), 2212.14029

[2] X. An, G. Başar, M. Stephanov, and H.-U. Yee, Phys. Rev. C **100**, 024910 (2019), 1902.09517

[3] X. An, G. Başar, M. Stephanov, and H.-U. Yee, Phys. Rev. C **102**, 034901 (2020), 1912.13456

[4] X. An, G. Başar, M. Stephanov, and H.-U. Yee, Phys. Rev. Lett. **127**, 072301 (2021), 2009.10742

[5] X. An, G. Başar, M. Stephanov, and H.-U. Yee, in progress

MDPI

*Article*

# Constraints on the mass of a bosonic dark matter candidate within the DD2Y-T model

**Mahboubeh Shahrbaf [1,2,*], Davood Rafiei Karkevandi [3,4] and Stefan Typel [5,6]**

[1] Incubator of Scientific Excellence–Centre for Simulations of Superdense Fluids, University of Wroclaw; m.shahrbaf46@gmail.com
[2] Frankfurt Institute for Advanced Studies, Giersch Science Center, D-60438 Frankfurt am Main, Germany; motlagh@fias.uni-frankfurt.de
[3] Department of Physics, Isfahan University of Technology, Isfahan 84156-83111, Iran; davood.rafiei64@gmail.com
[4] ICRANet-Isfahan, Isfahan University of Technology, 84156-83111, Iran; davood.rafiei64@gmail.com
[5] Technische Universität Darmstadt, Fachbereich Physik, Institut für Kernphysik, Schlossgartenstraße 9, D-64289 Darmstadt, Germany; stypel@ikp.tu-darmstadt.de
[6] GSI Helmholtzzentrum für Schwerionenforschung GmbH, Theorie, Planckstraße 1, D-64291 Darmstadt, Germany
* Correspondence: motlagh@fias.uni-frankfurt.de
† Present address: Frankfurt Institute for Advanced Studies, Giersch Science Center, D-60438 Frankfurt am Main, Germany

arXiv:2402.18686v1 [nucl-th] 28 Feb 2024

**Abstract:** Dark matter remains a topic of ongoing controversy. It has gained attention in the theoretical description of compact objects such as neutron stars with cores of very dense matter. Various candidates have been proposed for dark matter in the scientific literature. Among them, the sexaquark has been identified as a potential bosonic particle capable of being formed in neutron star matter based on its mass characteristics. In this study, we investigate the viability of the sexaquark as a candidate for dark matter, particularly under certain density conditions. Our goal is to address the challenges associated with the formation of a bosonic particle in a highly dense medium without compromising the stability of the neutron star. To achieve this, we introduce a straightforward linear mass shift for the sexaquark within the hadronic equation of state, utilizing a relativistic density functional approach. In our investigation, it is observed that the inclusion of Sexaquark as a candidate for dark matter within the hadronic matter equation of state, although featuring a repulsive interaction with baryonic matter, softens the equation of state. We suppose that the strength of the interaction of dark matter with baryonic matter increases linearly with the baryon density. We observe that raising the effective mass of the Sexaquark, as a result of increasing its vacuum mass, causes an increased stiffening of the equation of state as compared to the case of a constant mass. We determine the lower and upper mass boundaries for this bosonic dark matter based on observational constraints for neutron stars within the DD2Y-T model when a phase transition to quark matter phase is employed.

**Keywords:** Dark matter, Neutron star, Equation of State, Relativistic Mean-Field, Phase transition, Sexaquark

**Citation:** Shahrbaf, M.; Rafiei Karkevandi, D.; Typel, S. Constraints on the mass of bosonic dark matter candidate within DD2Y-T model. *Journal Not Specified* **2023**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

**Copyright:** © 2024 by the authors. Submitted to *Journal Not Specified* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

The accepted cosmological models that are presented and founded in General Relativity (GR), suggest that almost 27% of the matter-energy composition of the universe exists in the form of enigmatic Dark Matter (DM) [1].

Multi-messenger astrophysics is actively engaged in the search for and examination of new ideas concerning dark-matter particles. These candidates are believed to be non-relativistic particles that exist beyond the Standard Model of particle physics. They interact weakly with ordinary baryonic matter solely through gravitational forces [2]. In this context, the most widely accepted candidates for DM are weakly interacting massive

particles (WIMPs) [3,4]. However, hypothetical bosonic particles such as the axion [5,6] and the sexaquark [7,8] also exhibit the primary characteristics required to be considered as potential DM candidates. Their interaction with the baryonic matter must be very weak, which is why they have remained undetected thus far.

A broad spectrum for the mass of candidates for DM particles exists, spanning a wide range from approximately $10^{-22}$ eV to $10^{10}$ eV, contingent on the particular type of candidate. DM particles must be incorporated into an equation of state (EoS), which is a fundamental tool for characterizing matter, whereas the EoS for strongly interacting matter in neutron stars (NSs) is tightly bound by observational constraints. Therefore, efficient constraints on the features of DM candidates have been established using recent advancements in observational astronomy [9–15], mainly facilitated by gravitational wave (GW) interferometers and the Neutron Star Interior Composition ExploreR (NICER). One can refer to the works [16–18] and the references therein as illustrative examples of how the properties of DM can be constrained through multi-messenger observations of neutron stars. It is worth mentioning that Compact objects, such as white dwarfs, NSs, and strange stars, have the potential to include DM particles due to their high baryonic density. In particular, NSs are one of the well-known astrophysical laboratories for investigating DM because they are the densest environment known in the universe so far, except for black holes. Therefore, various studies have investigated the inclusion of either bosonic or fermionic DM candidates in NSs using different approaches [19–33].

The concept of a deeply bound Sexaquark (S), an alternative term for the hexaquark family of hypothetical particles, was first introduced as a bosonic candidate for DM in references [7,8,34]. The quark composition of S is identical to that of the H-dibaryon [35], with both consisting of uuddss quarks. However, S exhibits significantly stronger binding compared to the H-dibaryon. If S represents a molecular state like $\Lambda\Lambda$, it is important to note that $\Lambda$ is characterized as a color-neutral particle. In this context, the binding of two $\Lambda$ particles can exclusively occur through the exchange of color-neutral entities, such as mesons. When S comprises a complex system consisting of three colored diquarks, it becomes essential to acknowledge that these entities engage in interactions primarily governed by the color force. This force is notably more potent than the meson exchange force, particularly when considering short-range distances.

The proposed S is a boson at a spin-color-flavor singlet and even parity state, characterized by a zero electric charge, a baryon number of 2, and a strangeness value of -2. S may manifest itself as a deeply bound state, displaying a mass sufficiently low to render it either absolutely or effectively stable against decay. The potential of S to serve as DM hinges on its mass and its propensity to dissociate into two baryons, emphasizing the need for a small dissociation amplitude. An S particle lighter than deuterium cannot decay without violating baryon number conservation and, consequently, remains stable. This condition is met when the mass of S is less than the sum of the proton mass, electron mass, and Lambda baryon mass, which totals 2054.46 MeV — the lowest mass value consistent with a $\Delta \mathtt{S} = 1$ decay. Moreover, the ratio of densities between S dark matter and ordinary baryonic matter in the Universe that can be estimated through considerations of statistical equilibrium within the quark-gluon plasma, utilizing known parameters from Quantum Chromodynamics (QCD), aligns with observational data. Under such conditions, S emerges as a viable DM candidate [7]. A scheme in which such a state could be produced at rest through the formation of anti-protonic atoms has been proposed in [8].

In our prior work [36], referred to as 'paper I' hereafter, we established the lower-mass limit for S using a generalized relativistic density functional (GRDF) approach known as DD2Y-T, ensuring that the possible emergence of S as a DM candidate remains consistent with observational constraints derived from NS. Our findings disclosed the feasibility of S appearance in the hadronic phase of NS, demonstrating that a minimum mass of $m_{\mathtt{S}} = 1885$ MeV fulfills the necessary conditions that S does not appear at subsaturation densities. The stability of such stars is contingent on both a positive mass shift for S and a phase transition into deconfined quark matter. In this present study, we adopt the same approach

to ascertain the upper limit for the mass of S. We aim to investigate whether S remains a viable stable bosonic DM candidate across the entire accepted mass range within our model.

The paper is organized as follows: In Section 2, we present the EoS for hadronic matter including DM within the framework of the GDRF model with density-dependent couplings applied to NS matter. In Section 3, we introduce the EoS for quark matter. The hybrid EoS employing a cross-over transition is derived in section 4. Our findings and in-depth discussions concerning the properties of the resulting hybrid stars, as well as the mass of S, are presented in Section 5. Finally, we summarize our work and draw conclusions in Section 6.

## 2. Hadronic Matter Equation of State Including Dark Matter

The hadronic phase of matter is postulated to encompass all nucleons and hyperons, namely the neutron, proton, $\Lambda$, $\Sigma$, $\Xi$, and our DM candidate 'S'. These baryons are analyzed within the core of NS using a GRDF approach. In this approach, all particles are treated as quasi-particles with effective mass and effective chemical potential. The interactions in this model exclusively entail meson exchange interactions, and the couplings to these mesons depend on the density of the medium. The density-dependent couplings are meticulously adjusted to faithfully reproduce the characteristics of atomic nuclei, as established by earlier works [37–39].

For our investigation, we have employed an extended EoS, denoted as DD2Y-T, which accounts for all octet baryons. The interaction is described by the exchange of $\sigma$, $\omega$, $\rho$, and $\phi$ mesons. This EoS has been employed to assess the properties of NSs in both high-temperature, or "hot", and low-temperature, or "cold", hyperonic configurations [40]. Notably, it is essential to highlight that this model successfully addresses the 'hyperon puzzle', a conundrum that has been extensively deliberated in the literature and is considered one of the most challenging issues within the domain of hypernuclear physics [41–44].

In the DD2Y-T_S model, the S particle serves as an additional baryonic degree of freedom, featuring a mass shift that depends on the baryonic density, which captures the impact of its interaction with the surrounding medium. In principle, the interaction between the S particle and the surrounding medium can be theoretically elucidated by establishing couplings to mesons, akin to the approach employed for other baryonic degrees of freedom. However, this endeavor necessitates the introduction of several meson-sexaquark couplings, the precise values of which remain undetermined.

As established in paper I [36], we have demonstrated that a constant mass assumption for the S particle is incompatible with the structural stability of NS. Consequently, we introduced a linear density-dependent mass shift for the S particle, with a slope parameter harmonized with the corresponding mass shifts observed in other baryons. Detailed constraints regarding the lower threshold of the mass and the slope parameter have been thoughtfully presented and illustrated in Figure 8 of paper I. In our approach

$$S_S = -\Delta m_S, \tag{1}$$

is the scalar potential for the S particle and $\Delta m_S$ is the mass shift of S. The effective mass for S is thus defined as

$$m_S^* = m_S - S_S = m_S + \Delta m_S, \tag{2}$$

with the vacuum mass $m_S$ of S. The mass shift is given by

$$\Delta m_S = m_S x_S \frac{n_b}{n_0}. \tag{3}$$

where $n_b$ is the baryonic density, $x_S$ is a slope parameter and $n_0 = 0.15$ fm$^{-3}$ is the nuclear saturation density. A finite density of S exclusively emerges at zero temperature when the

condition $m_S^* = \mu_S^*$ is met, leading to the occurrence of Bose-Einstein condensation. The scalar density for S then is equal to the vector density and reads

$$n_S^{(s)} = n_S^{(v)} = n_b - n_n^{(v)} - n_p^{(v)} - n_\Lambda^{(v)} - n_{\Sigma^+}^{(v)} - n_{\Sigma^0}^{(v)} - n_{\Sigma^-}^{(v)} - n_{\Xi^0}^{(v)} - n_{\Xi^-}^{(v)}. \tag{4}$$

Here, $n_i^{(v)}$ signifies the vector density of the fermions. All relevant information pertaining to the thermodynamic properties of the system can be deduced from the grand canonical thermodynamic potential density, denoted as $\Omega(\mu_i)$, wherein $\mu_i$ signifies the chemical potential of each particle at zero temperature. The condensate contribution of the bosonic S in grand canonical thermodynamic potential is formally given by

$$\Omega_S = n_S^{(s)}(m_S^* - \mu_S^*), \tag{5}$$

where $\mu_S^* = \mu_S - V_S$ is the effective potential and $\mu_S = 2\mu_b$ is the chemical potential of S particle. $V_S$ is the vector potential for the S particle and reads

$$V_S = n_S^{(s)} \frac{\partial \Delta m_S}{\partial n_S^{(v)}}. \tag{6}$$

It appears as rearrangement contribution due to the density dependence of the mass shift and is required for thermodynamic consistency. For a more comprehensive exploration of the topic, we refer the reader to the extensive discussion presented in paper I.

The pressure ($P$) can be straightforwardly expressed as $P = -\Omega$. Moreover, the free energy density ($f$) coincides with the internal energy density ($\varepsilon$) and can be readily computed using the following relation:

$$f = \varepsilon = \Omega + \sum_i \mu_i n_i^{(v)}. \tag{7}$$

## 3. Quark-matter Equation of State

One of the best candidates for describing the color superconducting quark matter phase is a covariant nonlocal Nambu–Jona–Lasinio (nlNJL) model [45]. This effective model successfully accounts for the key features of low-energy QCD, such as dynamical chiral symmetry breaking, and also reasonably fulfills NS conditions.

In this effective model, it is worth noting that the precise values of the vector meson coupling and the diquark coupling remain undetermined. In one study, we have determined optimal parameters for those couplings in the nlNJL model [46]. This was achieved through Bayesian analysis, specifically in the context of a first-order phase transition from hadronic matter to deconfined quark matter within a systematic study of hybrid neutron star EoS.

In this study, we have utilized the selected parameters from reference [46] to establish a phase transition from the DD2Y-T_S model to the nlNJL model. The chosen parameters fall within the range of $0.10 < \eta_V < 0.12$ and $0.70 < \eta_D < 0.74$. As mentioned, these parameter selections were made through Bayesian analysis, considering the contemporary mass and radius constraints of NSs. For the Bayesian analysis in reference [46], two constraints derived from observations of neutron stars have been employed. The first constraint is based on the mass-radius measurements of the neutron star component within the binary system PSR J0740+6620, determined using the relativistic Shapiro time delay effect. The gravitational mass of this neutron star is estimated to be approximately $2.08^{+0.07}_{-0.07}$ M$\odot$ with a 68% credibility interval, as reported in references [14]. The neutron star's radius has been inferred through fitting rotating hot spot patterns to data from the Neutron Star Interior Composition Explorer (NICER) and X-ray Multi-Mirror (XMM-Newton) X-ray observations, resulting in an estimated radius of approximately $13.7^{+2.6}_{-1.5}$ km at the 68% confidence level, as reported in reference [11]. The second constraint utilized in our analysis pertains to the radius of a neutron star with a mass of 1.4 M$\odot$, estimated to be approximately $11.75^{+0.86}_{-0.81}$

km at a 90% confidence level. This constraint was derived through a combined analysis of the gravitational-wave event GW170817 with its electromagnetic counterparts AT2017gfo and GRB170817A, as well as the gravitational-wave event GW190425, both originating from neutron-star mergers, as described in reference [15].

## 4. Hybrid Equation of State

This study utilizes a crossover construction, often referred to as a mixed-phase approach, to model the phase transition between the hadronic and quark phases [44]. The hadronic equation of state (EoS) is considered appropriate at lower densities, particularly in the vicinity of nuclear saturation density, where the baryonic matter does not reveal its mixed structure. We establish the upper limit of applicability for our hadronic EoS as $n_H(\mu)$. In regions of relatively high densities, where quarks exist independently of specific baryons, the deconfined quark matter EoS is invoked. The lower threshold of validity for the color superconducting quark matter EoS is defined as $n_Q(\mu)$. In the intermediate density range, denoted by $n_H(\mu) < n(\mu) < n_Q(\mu)$, neither the hadronic EoS nor the quark matter EoS is suitable.

The determination of surface tension at the interface between the hadronic phase and quark phase in strongly interacting matter remains a topic of ongoing investigations. An infinite surface tension implies the applicability of a Maxwell construction, whereas a vanishing surface tension necessitates the use of a Glendenning construction. In scenarios where the surface tension is variable, a cross-over construction approach can be considered.

As the EoS for the hadronic and quark matter phases are expressed in terms of the relationship between pressure and chemical potential, denoted as $P_H(\mu)$ and $P_Q(\mu)$, respectively, specifically at $T = 0$ conditions relevant for NS, the effective cross-over EoS $P_M(\mu)$ can be conveniently represented as an interpolation function between these two phases. This interpolation requires that the pressure function seamlessly connects the values characteristic of hadronic and quark phases at their lower and upper limits. Moreover, it must satisfy fundamental thermodynamic constraints, including a positive slope of density versus chemical potential, i.e., $\frac{\partial n_M}{\partial \mu_M} = \frac{\partial^2 P_M}{\partial \mu_M^2} > 0$. Additionally, the causality condition, which ensures that the adiabatic speed of sound at zero frequency, $c_s^2 = \frac{\partial P}{\partial \epsilon}$, does not exceed the speed of light, must be adhered to. One effective and practical approach for achieving this interpolation is through the utilization of a polynomial function, which smoothly connects the pressure curves of the two phases.

The critical chemical potential, denoted as $\mu_c$, defines the point at which two phases achieve both mechanical and chemical equilibrium. It is determined through the Maxwell point, encompassing both the physical transition and a non-physical transition referred to as reconfinement, expressed as:

$$P_Q(\mu_c) = P_H(\mu_c) = P_c. \tag{8}$$

To model the pressure of the cross-over phase, we employ a polynomial ansatz:

$$P_M(\mu) = \sum_{q=1}^{N} \alpha_q (\mu - \mu_c)^q + (1 + \Delta_P) P_c. \tag{9}$$

Here, $\Delta_P$ represents a free parameter that characterizes the pressure of the cross-over phase at $\mu_c$:

$$P_M(\mu_c) = P_c + \Delta_P = P_M, \text{ where } \Delta_P = \frac{\Delta P}{P_c}. \tag{10}$$

The value of $\Delta_P$ is intricately related to the surface tension existing at the interface between the two phases. Typically, within the framework of equation (9), it is common to consider $N = 2$.

At the matching points $\mu_H$ and $\mu_Q$, the pressure of the cross-over phase aligns with the pressure of the hadronic and quark matter phases, respectively. This requires that both pressures and their first derivatives satisfy continuity conditions, as outlined below:

$$P_H(\mu_H) = P_M(\mu_H)\,, \tag{11}$$
$$P_Q(\mu_Q) = P_M(\mu_Q)\,, \tag{12}$$
$$\frac{\partial}{\partial\mu}P_H(\mu_H) = \frac{\partial}{\partial\mu}P_M(\mu_H)\,, \tag{13}$$
$$\frac{\partial}{\partial\mu}P_Q(\mu_Q) = \frac{\partial}{\partial\mu}P_M(\mu_Q)\,. \tag{14}$$

By considering the relationship $n(\mu) = \frac{\partial P(\mu)}{\partial\mu}$, we undertake a numerical solution of the continuity equations, i.e., (11-14) for baryon density. This approach enables us to deduce the unknown variables in the aforementioned equations.

It is notable to emphasize that for the employed EoSs, the first crossing between the hadronic phase and the deconfined quark matter phase takes place at extremely low densities. Consequently, we employ a cross-over construction for the second crossing corresponding to the reconfinement transition from the quark phase to the hadronic phase at higher densities, wherein $\Delta_P < 0$. This approach allows the hadronic equation of state to remain applicable at lower densities, while the quark matter equation of state remains valid at higher densities.

## 5. Results

In paper I, specific values for the slope parameter, denoted as $x_S$ which is in accordance with the coupling strength, and the vacuum mass of S $m_S$ in equation (3) were carefully chosen to satisfy the constraints arising from the interplay of these two parameters. The graphical representation of these constraints, depicting $x_S$ as a function of $m_S$ is illustrated in Figure 8 of Paper I. It is noteworthy that the DD2Y-T model, characterized by parameters fine-tuned to the properties of atomic nuclei around saturation density, offers an extrapolation of the EoS beyond the saturation point, rendering it less stringently constrained at high densities.

Consequently, the constraints derived from the saturation properties, as delineated in Figure 8 of Paper I, are notably more robust. Specifically, the lower boundaries for the mass and the coupling strength pertaining to S particles are established from the yellow and pink lines, yielding values of $m_S^{min} = 1885$ MeV and $x_S^{min} = 0.03$ respectively. Conversely, other constraints are based on the observations of NSs at high densities and may exhibit a degree of model dependency. Consequently, the forthcoming work will focus on investigating the model-dependent aspects of the current results.

This study aims to explore the upper limit for the mass of S as a candidate for DM within the framework of the DD2Y-T model. To ascertain the permissible range for the mass of S, it is important to venture into high-density regimes and ensure the alignment of outcomes with observational constraints associated with NSs. It is essential to note that a mass shift of S particles, $x_S$, serves the purpose of mimicking the interaction with the surrounding medium. Given the postulation that S is a candidate for DM, it is reasonable to anticipate a predominantly weak interaction with the medium. Consequently, even though elevating the mass of S would inherently enable us to increase the value of $x_S$ based on Figure 8 in paper I, it is a deliberate choice to maintain $x_S$ as the coupling strength of the DM interaction with baryonic matter at its minimal value, specifically $x_S = 0.03$ throughout the subsequent analyses.

It is imperative to note that in the lower panels of Figure 9 in Paper I, increasing the value of $x_S$ results in the outward displacement of the Mass-Radius (M-R) curve towards the outer region of the green band surrounding $1.4M_\odot$ which represents a crucial observational constraint relating to the mass and radius of NSs. This choice allows us to play with the

value of the vacuum (and consequently effective) mass of the DM candidate and investigate its role in the consistency of the EoS of the hybrid star with the observables.

We take the successful procedure of paper I and investigate it for the higher vacuum masses of S. Therefore, we employ the replacement interpolation construction (RIC) based on what was explained in Section 4 for a phase transition from hadronic matter to deconfined quark matter.

Figure 1 illustrates the pressure as a function of the baryonic chemical potential, showcasing both the hadronic phase and the quark phase. In the context of the hadronic phase, the coupling strength of S remains constant, as previously discussed, while different values of the vacuum mass of S are supposed but the effective mass exhibits an increasing trend. One should notice that in comparison to the hadronic EoS without S (as represented by DD2Y-T, which includes nucleons and hyperons), the inclusion of S as a DM candidate results in an obvious softening of the EoS, although a repulsive interaction between DM and baryonic matter is considered. In the quark phase, we depict various EoSs with specific parameters determined through Bayesian analysis, as outlined in reference [46]. These EoSs are characterized by three key parameters: the diquark coupling ($\eta_D$), the vector meson coupling ($\eta_V$), and the square of the speed of sound ($c_s^2 = \frac{dP}{d\epsilon}$), which serves as an indicator of EoS stiffness.

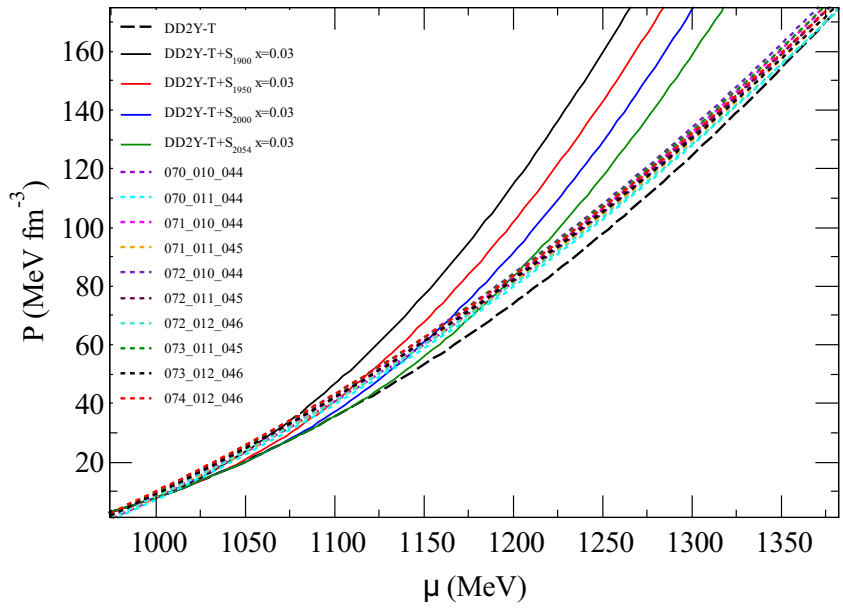



**Figure 1.** Pressure as a function of baryonic chemical potential for both hadronic phase and quark phase. The solid lines correspond to the hadronic phase with S for which the coupling strength of S remains constant while different values of the vacuum mass of S are supposed. The quark matter EoSs shown with dotted lines are characterized by three key parameters: the diquark coupling ($\eta_D$), the vector meson coupling ($\eta_V$), and the square of the speed of sound, respectively. The hadronic EoS without S is also shown with the black dashed line as DD2Y-T for comparison.

Figure 2 visually distinguishes and presents the low chemical potential region and the high chemical potential region, distinctly depicted in the left and right panels, respectively.

The purpose of presenting these two regions separately is twofold. Firstly, it serves to elucidate the physical transition from the hadronic phase to the quark phase in the left panel, which transpires at exceptionally low chemical potentials, near saturation density, where the emergence of deconfined quark matter or even S particles is not anticipated. Secondly, the high chemical potential region in the right panel reveals the reconfinement

point, where the RIC has been applied to correct the mischaracterization of the transition. As a result, at lower densities, the hadronic phase prevails, while at high densities, the quark phase remains a valid description. It is noteworthy that, for each hadronic EoS, we have chosen the initial intersection point with the quark phase lines in the right panel to serve as the basis for implementing the phase transition construction. But one aspect to highlight is that all the quark matter lines with favorable parameters as derived in [46] are applicable to the present investigation.

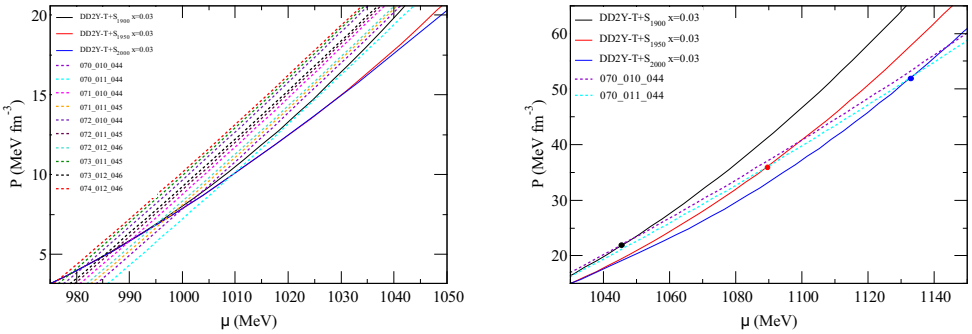



**Figure 2.** Same as Figure 1. The left panel shows the low chemical potential region while the right panel corresponds to the high chemical potential region. The dotts on the right panel show the critical chemical potential in Eq. (8) at which the RIC has been applied.

It is worth noticing in the right panel of Figure 2, when $x_S$ is held constant as a representative of the coupling strength of S interaction with the medium, increasing the vacuum mass of S particles (resulting in an increased effective mass of S) leads to a stiffening of the EoS, subsequently shifting the reconfinement onset to higher densities. As a result, a higher fraction of S particles as DM candidates will be produced in the stellar matter. A more detailed discussion regarding the fraction of DM within our model and our scenario is a topic for our future research.

Ultimately, Figure 3 delineates the distinct intersection points where the hadronic phase intersects with the quark matter phase, each accompanied by their respective hybrid EoS as obtained within the framework of RIC. The matching points defined earlier as $\mu_H$ and $\mu_Q$, represent the chemical potentials at which the pressure of the crossover phase is in concordance with the pressures of the hadronic and quark matter phases, clearly illustrated across the three panels.

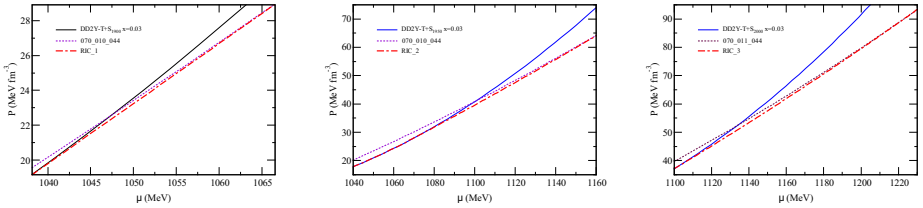



**Figure 3.** Pressure as a function of baryonic chemical potential for the hadronic phase, quark matter phase, and the hybrid EoS constructed within the framework of RIC for all three points shown in the right panel of Figure 2

It is clearly illustrated in Figure 3 how a negative value for $\Delta p$ in the context of RIC effectively addresses the issue of a nonphysical intersection point, often referred to as the reconfinement point. This adjustment restores the hadronic phase at lower densities and the quark matter phase at higher densities.

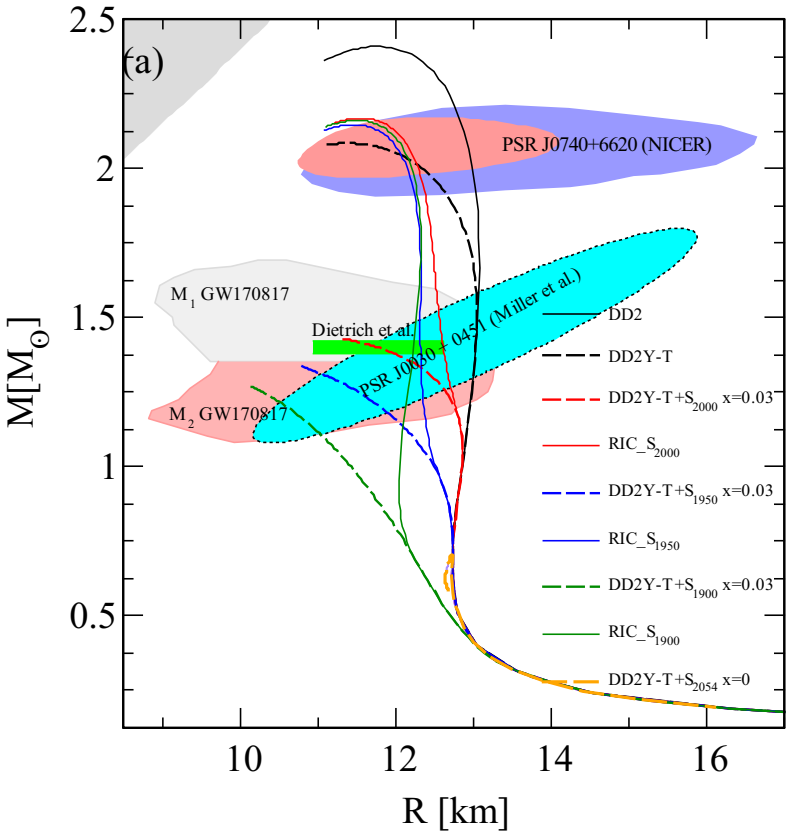



**Figure 4.** Mass-Radius curves for compact stars with a pure nucleonic core (black solid line), hyperonic core (black dashed line), hadronic core with S (dashed colorful lines), and hybrid stars with a quark matter core (solid colorful lines). For a comparison the new $1-\sigma$ mass-radius constraints from the NICER analysis of observations of the massive pulsar PSR J0740+6620 [14] are indicated in red [13] and blue [11] regions. Additionally, the green bar marks the radius of a 1.4 solar mass neutron star from a joint analysis of the gravitational-wave signal GW170817 with its electromagnetic counterparts at 90% confidence [15].

### *5.1. Discussion based on the observational constraints of neutron stars*

At the present time, the only way we can assess the accuracy of our hybrid solutions is by comparing them to observations of dense compact stars at high densities. To this end, both pure hadronic EoSs and hybrid ones are incorporated as inputs into the Tolman-Oppenheimer-Volkoff (TOV) equations, yielding the M-R diagram for each EoS. It is worth noting that the standard unified DD2 EoS, which also accounts for the crust of compact stars, has been utilized. The outcomes are presented graphically in Figure 4. The observational constraints are represented by color-coded regions in this figure.

In Figure 4, it is evident that none of the pure hadronic EoSs can simultaneously satisfy all imposed constraints. In contrast, all hybrid EoSs are successfully in favor of these constraints and this should be attributed to including DM in the hadronic phase. Notably, the figure illustrates the considerable trend in M-R curves as they pass the boundary defined by the green bar, which signifies the constraint on the radius of $1.4M_{\odot}$ stars. This behavior is attributed to the increased mass of S particles, which enhances the stiffness of the EoS and consequently results in larger stellar radii. Specifically, the red curve, corresponding to $m_S = 2000$ MeV, passes the border of the green bar. Therefore, we establish this mass value as the upper limit for the vacuum mass of S particles.

Furthermore, the recent detection of gravitational-wave (GW) signals arising from binary NS mergers, along with the subsequent determination of the tidal deformability parameter, have opened a new window for investigating the exotic internal structures of NSs [47,48]. Tidal deformability, which characterizes the susceptibility of a NS to deformation within a binary system, is exceptionally sensitive to the EoS governing NS matter and its compactness [49,50]. Consequently, we analyze the consistency of our model with respect to the tidal deformability constraint for a ($1.4M_{\odot}$) NS, $\Lambda_{1.4} = 190^{+390}_{-120}$.

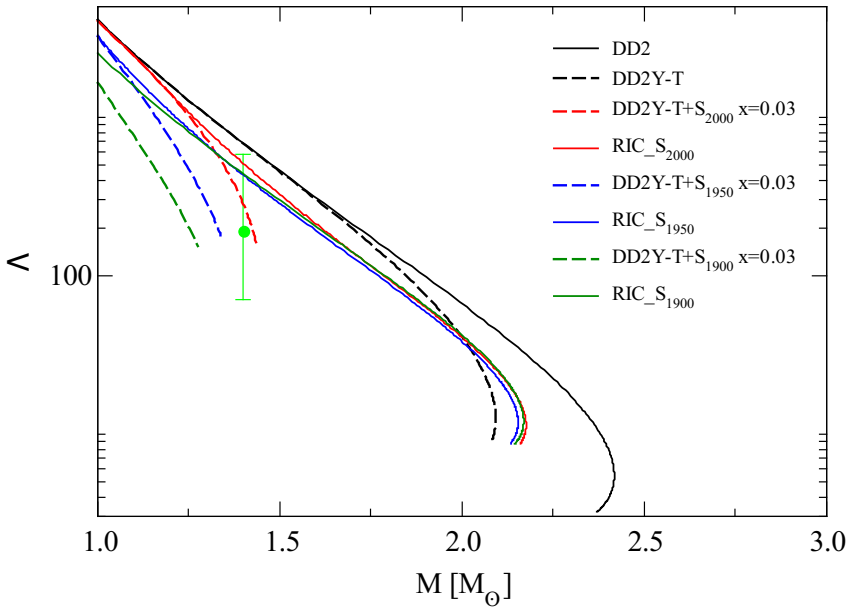



**Figure 5.** The dimensionless tidal deformability, denoted as $\Lambda$, is presented as a function of stellar mass for the ensemble of EoS illustrated in Figure 4. The green vertical line in the figure signifies the $\Lambda_{1.4}$ constraint derived from the low-spin prior analysis of GW170817, as reported in [9].

An intriguing outcome of DM production within the hadronic phase observed in Figure 5 is the fulfillment of the tidal deformability constraint. Without DM, both pure nucleonic EoS (DD2) and hyperonic EoS (DD2Y-T) fail to satisfy this constraint. However, the presence of S as a DM candidate plays a key role in softening the EoS around $1.4M_{\odot}$, ensuring that the tidal deformability constraint, as derived from the low-spin prior analysis of GW170817, is met for all hybrid EoSs constructed by incorporating S into the hadronic EoS.

## 6. Conclusions

Inspired by the presence of a stable sexaquark configuration, denoted as S=*uuddss*, with a mass below 2054 MeV, a threshold ensuring a lifetime exceeding the age of the Universe [7], which renders it a compelling bosonic DM, we have undertaken an examination to determine if the maximum mass, radius, and tidal deformability measurements of NSs are consistent with the potential existence of this intriguing particle in the core of NS. We employed the generalized relativistic density functional model, as introduced in reference [36], to describe the hadronic phase, incorporating the presence of S particles as a bosonic DM. Our observations indicate that the incorporation of S in the hadronic phase, featuring a repulsive interaction with baryonic matter within the parameter range employed in this study, leads to a softening of the EoS. Consequently, this softening effect contributes to the correction of the tidal deformability problem exhibited by pure hadronic EoS. The superconducting quark matter phase was modeled using the nlNJL model, as

outlined in reference [46], and we utilized a Replacement Interpolation Construction to develop a crossover phase transition between these two distinct phases. As a result, we modeled hybrid stars featuring a deconfined quark matter core surrounded by hypernuclear matter, including strange DM, all in accordance with the observational constraints of NSs. Our findings demonstrate that increasing the vacuum mass of the DM candidate is unfavorable to achieving the required softening of the EoS around $1.4M_{\odot}$ for satisfying the radius constraint of neutron stars. We have effectively constrained the upper limit of the sexaquark vacuum mass within our model to fall within the range of $1885 < m_{\mathrm{S}} < 2000$ MeV, drawing upon the consistency of our previous findings presented in reference [36] with the outcomes of the present study. However, an increase in the effective mass of S with medium density is required to find stable neutron stars above the density threshold for the appearance of S. Within the scope of our analysis, we refrain from making a definitive assessment regarding the necessity of a sexaquark in explaining NS properties. Eventually, the cooling behavior of NSs could potentially yield indications of the presence of S particles as a viable candidate for DM within the NS core. Furthermore, it is important to underscore that the approach taken in this paper is wholly phenomenological.

Outlook: In our upcoming work, we intend to conduct a more comprehensive parameter scan for our dark matter candidate and carry out a more detailed investigation of its precise role in shaping the equation of state of neutron stars. This will allow us to further refine our understanding of its effects within NSs. Moreover, investigating the model-dependency of our results is also a goal of our future works.

**Author Contributions:** M.Sh. The idea, investigation, methodology, formal analysis and calculations, writing and preparing the original draft of the paper, D.R. checking consistency with the observational constraints, completing the draft, reviewing and editing, S.T. Obtaining and programming the hadronic matter EoS, reviewing and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding

**Institutional Review Board Statement:** Not applicable

**Data Availability Statement:** Not applicable

**Acknowledgments:** M. Sh. has been supported by the program Excellence Initiative–Research University of the University of Wroclaw of the Ministry of Education and Science. M. Sh. acknowledges the hospitality of the Frankfurt Institute for Advanced Studies (FIAS).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

**DM** Dark Matter

**EoS** Equation of State

**NSs** Neutron Stars

**WIMPs** weakly interacting massive particles

**GW** gravitational-wave

**NICER** Neutron Star Interior Composition ExploreR

**S** Sexaquark

**QCD** Quantum Chromodynamics

**DD2** Parameterization of a GDRF for hadronic matter including only nucleons

**DD2Y-T** DD2 with hyperons by S. Typel

**DD2Y-T+S** DD2Y-T with sexaquark

**GRDF** Generalized relativistic density functional

**MC** Maxwell Construction

**RIC** Replacement Interpolation Construction

**M-R** Mass-Radius

## References


1. I. de Martino, S. S. Chakrabarty, V. Cesare, A. Gallo, L. Ostorero and A. Diaferio, "Dark matters on the scale of galaxies," Universe **2020** 6, 107.
2. P. Salucci, G. Esposito, G. Lambiase, E. Battista, M. Benetti, D. Bini, L. Boco, G. Sharma, V. Bozza and L. Buoninfante, *et al.* "Einstein, Planck and Vera Rubin: Relevant Encounters Between the Cosmological and the Quantum Worlds," Front. in Phys. **2021** 8, 603190.
3. C. Kouvaris, Phys. Rev. Lett. **2012** 108, 191301.
4. G. Bertone and T. Tait, M.P., "A new era in the search for dark matter," Nature **2018** 562, 51-56.
5. D. J. E. Marsh, "Axion Cosmology," Phys. Rept. **2016** 643, 1-79.
6. A. V. Balatsky, B. Fraser and H. S. Røising, "Dark sound: Collective modes of the axionic dark matter condensate," Phys. Rev. D **2022** 105, 023504.
7. G. R. Farrar, "A Stable Sexaquark: Overview and Discovery Strategies," [arXiv:2201.01334 [hep-ph]].
8. M. Doser, G. Farrar and G. Kornakov, "Searching for a dark matter particle with anti-protonic atoms," [arXiv:2302.00759 [hep-ph]].
9. B. P. Abbott *et al.* [LIGO Scientific and Virgo], "GW170817: Measurements of neutron star radii and equation of state," Phys. Rev. Lett. **2018** 121, 161101. [arXiv:1805.11581 [gr-qc]].
10. M. C. Miller, F. K. Lamb, A. J. Dittmann, S. Bogdanov, Z. Arzoumanian, K. C. Gendreau, S. Guillot, A. K. Harding, W. C. G. Ho and J. M. Lattimer, *et al.* "PSR J0030+0451 Mass and Radius from *NICER* Data and Implications for the Properties of Neutron Star Matter," Astrophys. J. Lett. **2019** 887, L24.
11. M. C. Miller, F. K. Lamb, A. J. Dittmann, S. Bogdanov, Z. Arzoumanian, K. C. Gendreau, S. Guillot, W. C. G. Ho, J. M. Lattimer and M. Loewenstein, *et al.* "The Radius of PSR J0740+6620 from NICER and XMM-Newton Data," Astrophys. J. Lett. **2021** 918, L28.
12. T. E. Riley, A. L. Watts, S. Bogdanov, P. S. Ray, R. M. Ludlam, S. Guillot, Z. Arzoumanian, C. L. Baker, A. V. Bilous and D. Chakrabarty, *et al.* "A *NICER* View of PSR J0030+0451: Millisecond Pulsar Parameter Estimation," Astrophys. J. Lett. **2019** 887, L21.
13. T. E. Riley, A. L. Watts, P. S. Ray, S. Bogdanov, S. Guillot, S. M. Morsink, A. V. Bilous, Z. Arzoumanian, D. Choudhury and J. S. Deneva, *et al.* "A NICER View of the Massive Pulsar PSR J0740+6620 Informed by Radio Timing and XMM-Newton Spectroscopy," Astrophys. J. Lett. **2021** 918, L27.
14. E. Fonseca, H. T. Cromartie, T. T. Pennucci, P. S. Ray, A. Y. Kirichenko, S. M. Ransom, P. B. Demorest, I. H. Stairs, Z. Arzoumanian and L. Guillemot, *et al.* "Refined Mass and Geometric Measurements of the High-mass PSR J0740+6620," Astrophys. J. Lett. **2021** 915, L12.
15. T. Dietrich, M. W. Coughlin, P. T. H. Pang, M. Bulla, J. Heinzel, L. Issa, I. Tews and S. Antier, "Multimessenger constraints on the neutron-star equation of state and the Hubble constant," Science **2020** 370, 1450-1453.
16. D. R. Karkevandi, S. Shakeri, V. Sagun and O. Ivanytskyi, "Bosonic dark matter in neutron stars and its effect on gravitational wave signal," Phys. Rev. D **2022** 105, 023001.
17. D. Rafiei Karkevandi, S. Shakeri, V. Sagun and O. Ivanytskyi, "Tidal deformability as a probe of dark matter in neutron stars," doi:10.1142/9789811269776_0307.
18. S. Shakeri and D. R. Karkevandi, "Bosonic Dark Matter in Light of the NICER Precise Mass-Radius Measurements," [arXiv:2210.17308 [astro-ph.HE]].
19. M. A. Perez-Garcia, J. Silk and J. R. Stone, "Dark matter, neutron stars and strange quark matter," Phys. Rev. Lett. **2010** 105, 141101.
20. S. C. Leung, M. C. Chu and L. M. Lin, Phys. Rev. D **2011**, 84, 107301 doi:10.1103/PhysRevD.84.107301 [arXiv:1111.1787 [astro-ph.CO]].
21. A. Li, F. Huang and R. X. Xu, "Too massive neutron stars: The role of dark matter?" Astropart. Phys. **2012** 37, 70-74.
22. G. Panotopoulos and I. Lopes, Phys. Rev. D **2017**, 96, no.8, 083004 doi:10.1103/PhysRevD.96.083004 [arXiv:1709.06312 [hep-ph]].
23. J. Ellis, G. Hütsi, K. Kannike, L. Marzola, M. Raidal and V. Vaskonen, "Dark Matter Effects On Neutron Star Properties," Phys. Rev. D **2018** 97, 123007.
24. A. Nelson, S. Reddy and D. Zhou, JCAP **2019**, 07, 012 doi:10.1088/1475-7516/2019/07/012 [arXiv:1803.03266 [hep-ph]].
25. O. Ivanytskyi, V. Sagun and I. Lopes, "Neutron stars: New constraints on asymmetric dark matter," Phys. Rev. D **2020** 102, 063028.
26. V. Sagun, E. Giangrandi, O. Ivanytskyi, I. Lopes and K. A. Bugaev, "Constraints on the fermionic dark matter from observations of neutron stars," PoS PANIC2021 **2022**, 313.
27. H. M. Liu, J. B. Wei, Z. H. Li, G. F. Burgio and H. J. Schulze, "Dark matter effects on the properties of neutron stars: Optical radii," Phys. Dark Univ. **2023** 42, 101338.
28. V. Parmar, H. C. Das, M. K. Sharma and S. K. Patra, "Influence of dark matter on magnetized neutron stars," Phys. Rev. D **2023** 108, 083003.

29. P. Thakur and T. K. Jha, "Neutron stars with fermionic dark matter: a two fluid approach," DAE Symp. Nucl. Phys. **2023** 66, 776-777.
30. M. Shahrbaf, "Appearance of sexaquark in the core of neutron stars as a candidate of dark matter," J. Phys. Conf. Ser. **2023** 2536, 012001.
31. N. Rutherford, G. Raaijmakers, C. Prescod-Weinstein and A. Watts, "Constraining bosonic asymmetric dark matter with neutron star mass-radius measurements," Phys. Rev. D **2023**, 107, 103051.
32. S. Shirke, S. Ghosh, D. Chatterjee, L. Sagunski and J. Schaffner-Bielich, "R-modes as a New Probe of Dark Matter in Neutron Stars," [arXiv:2305.05664 [astro-ph.HE]].
33. H. R. Rüter, V. Sagun, W. Tichy and T. Dietrich, "Quasi-equilibrium configurations of binary systems of dark matter admixed neutron stars," [arXiv:2301.03568 [gr-qc]].
34. G. R. Farrar, "A stable H dibaryon: Dark matter candidate within QCD?" Int. J. Theor. Phys. **2003** 42, 1211-1218.
35. R. L. Jaffe, "Perhaps a Stable Dihyperon," Phys. Rev. Lett. **1977** 38, 195-198 [erratum: Phys. Rev. Lett. **38** (1977), 617]
36. M. Shahrbaf, D. Blaschke, S. Typel, G. R. Farrar and D. E. Alvarez-Castillo, "Sexaquark dilemma in neutron stars and its solution by quark deconfinement," Phys. Rev. D **2022** 105, 103005.
37. S. Typel and H. H. Wolter, "Relativistic mean field calculations with density dependent meson nucleon coupling," Nucl. Phys. A **1999** 656, 331-364.
38. S. Typel, "Relativistic model for nuclear matter and atomic nuclei with momentum-dependent self-energies," Phys. Rev. C **2005** 71, 064301.
39. S. Typel, G. Ropke, T. Klahn, D. Blaschke and H. H. Wolter, "Composition and thermodynamics of nuclear matter with light clusters," Phys. Rev. C **2010** 81, 015803.
40. J. R. Stone, V. Dexheimer, P. A. M. Guichon, A. W. Thomas and S. Typel, "Equation of state of hot dense hyperonic matter in the Quark–Meson-Coupling (QMC-A) model," Mon. Not. Roy. Astron. Soc. **2021** 502 no.3, 3476-3490.
41. M. Shahrbaf and H. R. Moshfegh, "Appearance of hyperons in neutron stars within LOCV method," Annals Phys. **2019** 402, 66-77.
42. M. Shahrbaf, D. Blaschke, A. G. Grunfeld and H. R. Moshfegh, "First-order phase transition from hypernuclear matter to deconfined quark matter obeying new constraints from compact star observations," Phys. Rev. C **2020** 101, 025807.
43. M. Shahrbaf, H. R. Moshfegh and M. Modarres, "Equation of state and correlation functions of hypernuclear matter within the lowest order constrained variational method," Phys. Rev. C **2019** 100, 044314.
44. M. Shahrbaf, D. Blaschke and S. Khanmohamadi, "Mixed phase transition from hypernuclear matter to deconfined quark matter fulfilling mass-radius constraints of neutron stars," J. Phys. G **2020** 47, 115201.
45. A. E. Radzhabov, D. Blaschke, M. Buballa and M. K. Volkov, "Nonlocal PNJL model beyond mean field and the QCD phase transition," Phys. Rev. D **2011** 83, 116004.
46. M. Shahrbaf, S. Antić, A. Ayriyan, D. Blaschke and A. G. Grunfeld, "Constraining free parameters of a color superconducting nonlocal Nambu–Jona-Lasinio model using Bayesian analysis of neutron stars mass and radius measurements," Phys. Rev. D **2023** 107, 054011.
47. B. P. Abbott *et al.* [LIGO Scientific and Virgo], "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral," Phys. Rev. Lett. **2017**, 119, 161101.
48. B. P. Abbott *et al.* [LIGO Scientific and Virgo], "GW190425: Observation of a Compact Binary Coalescence with Total Mass $\sim 3.4 M_{\odot}$," Astrophys. J. Lett. **2020**, 892, L3.
49. T. Hinderer, "Tidal Love numbers of neutron stars," Astrophys. J. **2008**, 677, 1216-1220.
50. T. Hinderer, B. D. Lackey, R. N. Lang and J. S. Read, "Tidal deformability of neutron stars with realistic equations of state and their gravitational wave signatures in binary inspiral," Phys. Rev. D **2010**, 81, 123016.


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

# The Collective Coordinate Fix

**Arindam Bhattacharya,**[1] **Jordan Cotler,**[1,2] **Aurélien Dersy,**[1] **Matthew D. Schwartz**[1]

[1] *Department of Physics, Harvard University, Cambridge, MA 02138, USA*
[2] *Society of Fellows, Harvard University, Cambridge, MA 02138, USA*

arindamb@g.harvard.edu, jcotler@fas.harvard.edu,
adersy@g.harvard.edu, schwartz@g.harvard.edu

arXiv:2402.18633v1 [hep-th] 28 Feb 2024

**Abstract**

Collective coordinates are frequently employed in path integrals to manage divergences caused by fluctuations around saddle points that align with classical symmetries. These coordinates parameterize a manifold of zero modes and more broadly provide judicious coordinates on the space of fields. However, changing from local coordinates around a saddle point to more global collective coordinates is remarkably subtle. The main complication is that the mapping from local coordinates to collective coordinates is generically multi-valued. Consequently one is forced to either restrict the domain of path integral in a delicate way, or otherwise correct for the multi-valuedness by dividing the path integral by certain intersection numbers. We provide a careful treatment of how to fix collective coordinates while accounting for these intersection numbers, and then demonstrate the importance of the fix for free theories. We also provide a detailed study of the fix for interacting theories and show that the contributions of higher intersections to the path integral can be non-perturbatively suppressed. Using a variety of examples ranging from single-particle quantum mechanics to quantum field theory, we explain and resolve various pitfalls in the implementation of collective coordinates.

# Contents

**1 Introduction** **3**

**2 Collective coordinates: general results** **6**

**3 Free theory** **8**

**4 Double well potential** **11**

4.1 Normal modes at finite $T$ . . . . . 14

4.2 Collective coordinates . . . . . 16

4.3 Energy splitting . . . . . 17

4.4 Free twisted partition function . . . . . 20

**5 Quantum field theory** **23**

5.1 Basis and cross check . . . . . 23

5.2 Collective coordinates for dilatations . . . . . 25

5.3 Collective coordinates for translations . . . . . 29

**6 Conclusions** **30**

**A Metric approach to Jacobians in the path integral** **31**

**B Mapping QFT to the double well** **33**

# 1 Introduction

One of the great advantages of the path integral for a quantum theory is that it allows for expansions around saddle points other than the free theory, providing access to non-perturbative physics. For example, the lifetime of the Standard Model can be computed by expanding the path integral around a non-trivial configuration of the Higgs field [1–7]. In many field theories (including the Standard Model) there are families of equivalent non-trivial saddles related by symmetries, such as translations, dilatations, or internal symmetries. Expanding around any one of these saddles leads to a divergent 1-loop saddle point calculation because the zero eigenvalues from the flat directions make the functional determinant vanish. Such infinities are usually handled by introducing collective coordinates along the manifold of zero modes, effectively trading local coordinates around a single saddle for more global coordinates around the entire manifold of zero modes [8–10]. Then the integral over the zero mode manifold can be handled non-perturbatively, giving a factor of the volume of the symmetry group. The remaining path integral can then be evaluated using a saddle-point approximation, and can be regularized to be finite. While this procedure gives sensible answers in some situations where it can be checked (for example, the energy splitting in a symmetric double-well potential can be alternatively computed with the WKB approximation [11–16]), the procedure also happens to give very wrong answers in other situations, such as in free theories.

To see the problem clearly, consider in more detail the decay rate calculation in a quantum field theory with action $S = \int d^4x \left(\frac{1}{2}\,\partial_\mu\phi\,\partial^\mu\phi + \lambda\phi^4\right)$ with $\lambda < 0$. As we will review in Section 5, the decay rate involves an expansion around a non-trivial instanton background $\phi_b(x)$ normalized by an expansion around the trivial vacuum $\phi(x) = 0$. The decay rate $\Gamma$ can be written as [5–7]

$$\Gamma = \lim_{T\to\infty} \frac{1}{T} \mathrm{Im} \frac{\int_{C_b} \mathcal{D}\phi\, e^{-S[\phi]}}{\int_{C_{\mathrm{FV}}} \mathcal{D}\phi\, e^{-S_0[\phi]}}\,, \tag{1.1}$$

where $S_0[\phi] = \int d^4x \left(\frac{1}{2}\partial_\mu\phi\partial^\mu\phi\right)$ is the free action. Here $C_b$ is a path of steepest descent in the interacting theory passing through $\phi_b(x)$ and $C_{\mathrm{FV}}$ is the path of steepest descent in the free theory passing though the free vacuum $\phi(x) = 0$. To evaluate the numerator in the saddle point approximation, we expand

$$\phi(x) = \phi_b(x) + \sum_{m=0}^{\infty} a_m \phi_m(x)\,, \tag{1.2}$$

where $\{\phi_m(x)\}$ forms a basis of normal modes around the bounce $\phi_b(x)$. Five of these modes, corresponding to four translations and one dilatation mode, have zero eigenvalue with respect to the operator $L''[\phi_b]$, where $L$ is the Lagrangian, and so a naïve saddle point approach gives infinite contributions to the numerator of Eq. (1.1). A more careful treatment of steepest descent is required to resolve these infinities. To do so, we can change to a collective coordinate basis

$$\phi(x) = \phi_b(x+\xi, R) + \sum_{m\geq 5} \alpha_m \phi_m(x+\xi, R)\,, \tag{1.3}$$

where the five zero-mode coordinates $\alpha_m$ for $m = 0, 1, 2, 3, 4$ are traded for five collective coordinates $\xi^\mu$ and $R$. The integral over $\xi^\mu$ then gives a factor of the spacetime volume $T\mathcal{V}$ for the numerator of Eq. (1.1), as expected for a decay rate in spacetime. However, a complication arises if we use the same exact basis of Eq. (1.3), which happens to also diagonalize normal modes in the free theory, for the denominator of Eq. (1.1). In particular, the same $T\mathcal{V}$ factor apparently arises, which is unexpected. In this paper, we identify the cause of this unexpected factor, and show that the appropriate treatment of collective coordinates yields sensible results for both the numerator and denominator of Eq. (1.1). We more broadly study general coordinates on field space in which similar subtleties arise.

To understand the source of the subtleties, let us recall how $\delta$-functions can be used to change variables in integrals. In a finite-dimensional integral $\int_\Omega d^d x\, A(x)$ over $x \in \Omega \subseteq \mathbb{R}^d$, we can implement the change of variables $x = f(y)$ using the resolution of the identity

$$1 = \int_{f^{-1}(\Omega)} d^d y \left|\frac{\partial f(y)}{\partial y}\right| \delta^d(f(y) - x)\,, \tag{1.4}$$

so that

$$\int_\Omega d^d x\, A(x) = \int_\Omega d^d x \int_{f^{-1}(\Omega)} d^d y \left|\frac{\partial f(y)}{\partial y}\right| \delta^d(f(y) - x)\, A(x) \tag{1.5}$$

$$= \int_{f^{-1}(\Omega)} d^d y \left|\frac{\partial f(y)}{\partial y}\right| A(f(y))\,. \tag{1.6}$$

The above can be regarded as a trick to find the right Jacobian factor and domain of integration. Crucially, we notice that the above assumes that $f$ is an invertible function on the domain of integration. But what if $f$ multiply-covers the domain of integration? A simple example is the 1-dimensional integral $\int_0^\infty dx\, A(x)$ with the change of variables $x = y^2$. In this setting the salient resolution of the identity is

$$1 = \frac{1}{2}\int_{-\infty}^{\infty} dy\, |2y|\, \delta(y^2 - x)\,. \tag{1.7}$$

Here, we need the factor of $\frac{1}{2}$ out front because the map $f(y) = y^2$ taking $\mathbb{R} \to \mathbb{R}_{\geq 0}$ doubly covers the positive real line. Going back to Eq. (1.4), if $f(y) = x$ has $N_f[x]$ solutions for fixed $x$, then we need to correct Eq. (1.4) by

$$1 = \frac{1}{N_f[x]} \int_{f^{-1}(\Omega)} d^d y \left|\frac{\partial f(y)}{\partial y}\right| \delta^d(f(y) - x)\,, \tag{1.8}$$

where we call $N_f[x]$ the *intersection number*. A similar subtlety arises in gauge theories where there can be multiple solutions to a local gauge-fixing condition, and so we must introduce a multiplicative factor to correct for the so-called *Gribov copies* [17].

To see how the intersection number appears in a path integral, consider non-relativistic quantum mechanics where the path integral is over paths $x(t)$. For a time-independent Lagrangian, there

will always be a flat direction: $S[x(t)] = S[x(t+t_0)]$ for any path $x(t)$ and any $t_0$. To change from orthonormal coordinates on path space to collective coordinates we can employ the resolution of the identity[1]

$$1 = \frac{1}{N_X[x]} \int dt_0 \, \delta\Big[\langle X(t+t_0)|x(t)\rangle\Big] \Big|\langle X(t+t_0)|\dot{x}(t)\rangle\Big| \,. \tag{1.9}$$

Here $\langle x|y\rangle = \int dt\, x(t)y(t)$ denotes the $L^2$ inner product, $x(t)$ is a trajectory in the domain of the path integral, $\dot{x}(t) = \partial_t x(t)$ is the time derivative, $X(t+t_0)$ is the path whose coordinate we want to exchange for $t_0$, and $N_X[x]$ is the intersection number. That is, solving $\langle X(t+t_0)|x(t)\rangle = 0$ will replace the time-translation zero mode with the collective coordinate $t_0$ for any path $x(t)$. Then the intersection number $N_X[x]$ counts the number of values of $t_0$ for which $\langle X(t+t_0)|x(t)\rangle = 0$ for a given path $x(t)$. The intersection number is a complicated integer-valued functional of the path which must be carefully taken into account.

The textbook narrative is that a collective coordinate can be simply pulled out of the path integral giving a volume factor [18–22]. Unfortunately, this standard treatment does not properly account for the multi-valuedness of the coordinate change associated with going from local coordinates to collective coordinates. In this paper, we explore through examples how and when collective coordinates can be used. Our key observations are that (1) the transition to collective coordinates does not involve only an innocuous coordinate change; (2) in interacting theories, non-Gaussian terms in the action are important to suppress high-multiplicity intersections; and (3) the Jacobian factor incurred by changing to collective coordinates plays an important role in assuring the consistency of the implementation.

Our paper begins in Section 2 with an overview of how to change from local coordinates to collective coordinates. We then proceed through a set of examples of increasing complexity. In Section 3, we discuss a free particle with time compactified on a circle. For this system time-translation invariance is exact. We explicitly compute the intersection number associated with the collective coordinates and illustrate the proper treatment of the Jacobian. In Section 4, we consider the symmetric double-well potential in quantum mechanics, focusing on the computation of the energy splitting between the two lowest levels. This theory has a non-trivial instanton and a basis of excitations around the instanton whose analytic form is explicitly known. We further show that the intersection number apropos to the collective coordinates counts the number of times a given path $x(t)$ crosses $x = 0$. We then demonstrate that paths with $N_X[x] > 1$ give exponentially small contributions to the path integral in the interacting theory, but give $\mathcal{O}(1)$ contributions in the free theory. In Section 5 we consider $\lambda\phi^4$ quantum field theory and discuss collective coordinates for dilatation and translations. In Section 6 we conclude with a discussion.

[1] We have used $\Big|\langle \dot{X}(t+t_0)|x(t)\rangle\Big| = \Big|\langle X(t+t_0)|\dot{x}(t)\rangle\Big|$ obtained via integration by parts in $t$.

## 2 Collective coordinates: general results

We begin by discussing how collective coordinates are used in quantum mechanics. In evaluating the path integral, we often expand around some saddle point $x_I(t)$. This saddle can be the free solution $x = 0$, an instanton, or simply some classical path satisfying the boundary conditions of the path integral for the problem of interest. For concreteness, suppose we work in Euclidean signature with a time-independent Lagrangian of the form $L[x] = \frac{1}{2}\dot{x}^2 + V(x)$ with action $S[x] = \int dt\, L[x]$. Then we can expand a generic path in an orthonormal basis of eigenstates of $\frac{\delta^2 S}{\delta x^2}\big|_{x=x_I}$. That is, we write

$$x(t) = x_I(t) + a_0 x_0(t) + \sum_{n=1}^{\infty} a_n x_n(t) \tag{2.1}$$

where

$$[-\partial_t^2 + V''(x_I)]\, x_n = \lambda_n x_n\,. \tag{2.2}$$

In (2.1) we have separated out the zero mode $x_0(t) \propto \dot{x}_I(t)$. We can see that $\lambda_0 = 0$ using $\frac{\delta S}{\delta x}\big|_{x=x_I} = 0$ so that

$$0 = \partial_t \left.\frac{\delta S}{\delta x(t)}\right|_{x=x_I} = \int dt' \left.\frac{\delta^2 S}{\delta x(t)\delta x(t')}\right|_{x=x_I} \dot{x}_I(t')\,. \tag{2.3}$$

Then the path integral becomes, to 1-loop order in the saddle-point approximation,

$$Z = \mathcal{N}\int \mathcal{D}x\; e^{-S} = \mathcal{N}e^{-S[x_I]} \int \frac{da_0}{\sqrt{2\pi}} \prod_{n=1}^{\infty} \int \frac{da_n}{\sqrt{2\pi}}\; e^{-\lambda_n a_n^2} = \mathcal{N}e^{-S[x_I]} \sqrt{\frac{1}{\lambda_0}} \prod_{n=1}^{\infty} \sqrt{\frac{1}{\lambda_n}}\,. \tag{2.4}$$

Because of the zero eigenvalue, Eq. (2.4) is infinite. The existence of a zero mode is because the action $S[x(t+t_0)]$ is independent of $t_0$. We therefore would like to trade $a_0$ for $t_0$ and integrate over $t_0$ giving a factor of the volume of time $T$.

To trade $a_0$ for $t_0$ we transform from coordinates $\{a_0, a_n\}$ to collective coordinates $\{t_0, \alpha_n\}$ of the form

$$x(t) = x_I(t+t_0) + \sum_{n=1}^{\infty} \alpha_n x_n(t+t_0)\,. \tag{2.5}$$

Since $\{\alpha_0, \alpha_n\}$ are global coordinates, for which each path in the integration domain is represented exactly once, the zero mode can be projected out with

$$\big\langle x(t) - x_I(t+t_0)\big| x_0(t+t_0)\big\rangle = 0 \tag{2.6}$$

where $\langle x|y\rangle$ is the inner product under which the $x_n$ are orthonormal. Eq. (2.6) can be solved to find a coordinate $t_0$ for a given path $x(t)$. Importantly, however, there can be multiple solutions and we denote by $N_{x_0}[x]$ the number of solutions for $t_0$ given $x$. The Jacobian for the coordinate change is determined by the partial derivatives

$$\frac{\partial a_m}{\partial t_0} = \big\langle x_m(t)\big|\partial_{t_0} x_I(t+t_0) + \sum_{n=1}^{\infty} \alpha_n \partial_{t_0} x_n(t+t_0)\big\rangle, \quad \frac{\partial a_m}{\partial \alpha_n} = \big\langle x_m(t)\big| x_n(t+t_0)\big\rangle\,. \tag{2.7}$$

After some algebra (see Appendix B of [6]), this Jacobian can be simplified to

$$\mathcal{J} = \big\langle x_0(t+t_0)\big|\partial_{t_0} x_I(t+t_0) + \sum_{n=1}^{\infty} \alpha_n \partial_{t_0} x_n(t+t_0)\big\rangle \tag{2.8}$$

$$= \big\langle x_0(t)\big|\dot{x}_I(t) + \sum_{n=1}^{\infty} \alpha_n \dot{x}_n(t)\big\rangle\,. \tag{2.9}$$

Alternatively, the functional determinant can be computed as $\sqrt{\det G}$ where $G$ is the Riemannian metric on the space of fields in collective coordinates (see Appendix A). That the Jacobian is independent of $t_0$ is non-trivial. Thus we have

$$\boxed{Z = \mathcal{N} e^{-S[x_I]} \int \frac{dt_0}{\sqrt{2\pi}} \int \prod_{m=1}^{\infty} \frac{d\alpha_m}{\sqrt{2\pi}}\, e^{-\lambda_m \alpha_m^2} \frac{1}{N_{x_0}[x]} \Big|\big\langle x_0(t)\big|\dot{x}_I(t) + \sum_{n=1}^{\infty} \alpha_n \dot{x}_n(t)\big\rangle\Big|} \tag{2.10}$$

This is the formulation we should use if we want to implement the $\{t_0, \alpha_n\}$ coordinates as in Eq. (2.5).

From Eq. (2.10) we can integrate in $\alpha_0$ by including $\int d\alpha_0\, \delta(\alpha_0)$. If we then relabel $\alpha_0 \to a_0$ and $\alpha_n \to a_n$, the resulting expression would be exactly the same as if we had expanded

$$Z = \mathcal{N} \int dt_0 \left\{ \int \mathcal{D}x\; e^{-S[x]} \frac{1}{N_{x_0}[x]} \delta\Big[\big\langle x_0(t)\big|x(t)\big\rangle\Big] \Big|\big\langle x_0(t)\big|\dot{x}(t)\big\rangle\Big| \right\} \tag{2.11}$$

using Eq. (2.1). In the above equation we can no longer use $t_0$ as a coordinate: it is a constant from the point of view of the path integral (the part of the expression in the curly brackets). That is, we cannot evaluate the path integral expression using the coordinates in Eq. (2.5). If we want to use $t_0$ as a coordinate, we should instead use Eq. (2.10).

There was nothing essential about $x_I(t)$ or $x_0(t)$ in our derivation. Indeed, $x_I(t)$ does not even appear in Eq. (2.11). We can replace $x_0(t)$ by any path $X(t)$ leading to

$$\boxed{Z = \mathcal{N} \int dt_0 \left\{ \int \mathcal{D}x\; e^{-S[x]} \frac{1}{N_X[x]} \delta\Big[\big\langle X(t)\big|x(t)\big\rangle\Big] \Big|\big\langle X(t)\big|\dot{x}(t)\big\rangle\Big| \right\}} \tag{2.12}$$

where $N_X[x]$ counts the number of times $\langle X(t+t_0)|x(t)\rangle$ vanishes over the integration range of $t_0$ for fixed $x(t)$. Note that the above equation can also be derived inserting the identity of Eq. (1.9) inside the path integral and using $S[x(t+t_0)] = S[x(t)]$.

To summarize, the key new element in our analysis is the intersection number $N_X[x]$. If we write a generic path $x(t)$ in an orthonormal basis, then $N_X[x]$ is the number of solutions to

$$\langle X(t+t_0)|x(t)\rangle = 0 \tag{2.13}$$

over the range of $t_0$. Eq. (2.13) allows us to solve for $t_0$ in terms of the coordinates labeling an orthonormal basis, and there may be multiple solutions. When the collective coordinate $t_0$ is one

of the coordinates, Eq. (2.13) cannot be used. Instead, in collective coordinates, $N_X[x]$ counts the number of ways of constructing the same path $x(t)$ using different values of $t_0$. That is, the same function would be overcounted in the path integral in the collective coordinates if it can be described using multiple values of $t_0$, and so we must divide by a $N_X[x]$ factor to compensate. We will see examples of the use of $N_X[x]$ in the following.

# 3 Free theory

As a first example, we consider a free non-relativistic quantum theory for a particle with mass $m = 1$ and Lagrangian $L = \frac{1}{2}\dot{x}^2$. The propagator going from $a$ to $b$ in time $T$ is

$$D(a,b) = \langle b|e^{-HT}|a\rangle = \int_{-\infty}^{\infty} \frac{dp}{2\pi}\, e^{ip(b-a)} e^{-\frac{p^2}{2}T} = \sqrt{\frac{1}{2\pi T}}\, e^{-\frac{(b-a)^2}{2T}} \tag{3.1}$$

and the partition function is

$$Z_0 = \mathrm{Tr}(e^{-HT}) = \int_{-\mathcal{V}/2}^{\mathcal{V}/2} dx_0\, D(x_0,x_0) = \sqrt{\frac{1}{2\pi T}}\, \mathcal{V}\,, \tag{3.2}$$

where $\mathcal{V}$ is the volume of space used the regulate this integral. To compute the partition function with the path integral, we parameterize paths with periodic boundary conditions in time as

$$x(t) = x_0 + \sum_{n\,\text{even}} \left[a_n s_n(t) + b_n c_n(t)\right]\,, \tag{3.3}$$

where

$$s_n(t) = \sqrt{\frac{2}{T}} \sin\left(\frac{n\pi\left(t+\frac{T}{2}\right)}{T}\right), \quad c_n(t) = \sqrt{\frac{2}{T}} \cos\left(\frac{n\pi\left(t+\frac{T}{2}\right)}{T}\right). \tag{3.4}$$

Then the partition function is given by

$$Z_0 = \mathcal{N}\int \mathcal{D}x\; e^{-S[x]} = \mathcal{N}\int_{-\mathcal{V}/2}^{\mathcal{V}/2} dx_0 \prod_{n\ \text{even}} \left[\int_{-\infty}^{\infty} \frac{da_n}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda_n a_n^2} \int_{-\infty}^{\infty} \frac{db_n}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda_n b_n^2}\right] \tag{3.5}$$

$$= \mathcal{N}\mathcal{V}\left[\prod_{n\,\text{even}} \sqrt{\frac{1}{\lambda_n}}\right]^2 \tag{3.6}$$

where

$$\lambda_n = \left(\frac{n\pi}{T}\right)^2 \tag{3.7}$$

are the eigenvalues of the fluctuations under $\left.\frac{\delta^2 S}{\delta x(t)\,\delta x(t')}\right|_{x=0} = -\delta(t-t')\,\partial_t^2$. We then match on to Eq. (3.2) with

$$\mathcal{N} = \sqrt{\frac{1}{2\pi T}}\left[\prod_{n\,\text{even}} \frac{T}{n\pi}\right]^{-2} \tag{3.8}$$

which fixes the normalization of the path integral. Note that $\mathcal{V}$ is only needed to regulate the $x_0$ integral: we can take the integrals over $a_n$ and $b_n$ to go to from $-\infty$ to $\infty$ because the path integral is exponentially suppressed at large $a_n, b_n$.

Now we want to reproduce the partition function $Z_0$ with collective coordinates, using Eq. (2.12). We first need to choose which path $X(t)$ to project onto, i.e. which coordinate to swap for the collective coordinate $t_0$. The coordinate to swap should be one from a parameterization of paths where time is shifted by $t_0$, i.e. from

$$x(t) = x_0 + \sum_{n\,\text{even}} \left(\alpha_n s_n(t+t_0) + \beta_n c_n(t+t_0)\right). \tag{3.9}$$

In the simplest case, we would swap a single coordinate, say $\beta_m$, for $t_0$. That would mean choosing $X(t) = c_m(t)$. To change to coordinates involving $t_0$ we have to start from coordinates not involving $t_0$, such as the original coordinates in Eq. (3.3). The condition $\beta_m = 0$ is $\langle c_m(t+t_0)|x(t)\rangle = 0$ as in Eq. (2.13), which in the original coordinates amounts to

$$\tan\left(\frac{m\pi t_0}{T}\right) = -\frac{b_m}{a_m}. \tag{3.10}$$

This equation allows us to solve for $t_0$ and to transform from the $\{a_n, b_n\}$ coordinates to the $\{t_0, \alpha_n, \beta_n\}$ coordinates with no $\beta_m$ component. Note that over the range $0 < t_0 \leqslant T$, Eq. (3.10) has $m$ possible solutions so that[2] $N_X[x] = m$.

The Jacobian arising from Eq. (2.12) is

$$\mathcal{J} = \left\langle x(t)\middle|\dot{c}_m(t)\right\rangle = \sqrt{\lambda_m}\, a_m\,, \tag{3.11}$$

and thus in collective coordinates the partition function is

$$Z_{\text{coll}} = \mathcal{N}\frac{1}{m}\int_0^T dt_0 \int_0^{\mathcal{V}} dx_0 \prod_{\substack{n\,\text{even}\\ n\neq m}} \left(\int_{-\infty}^{\infty} \frac{da_n}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\lambda_n a_n^2} \int_{-\infty}^{\infty} \frac{db_n}{\sqrt{2\pi}} e^{-\frac{1}{2}\lambda_n b_n^2}\right) \times \int \frac{da_m}{\sqrt{2\pi}} \sqrt{\lambda_m}|a_m| e^{-\frac{1}{2}a_m^2\lambda_m}\,. \tag{3.12}$$

Using Eq. (3.8) the above agrees precisely with $Z_0$ in Eq. (3.2).

The next simplest case is to choose $X(t)$ to be a sum of two modes instead of one mode. Suppose we choose $X(t) = A\, c_m(t) + B\, c_k(t)$ which is a sum of cosine modes. Then the coordinate $t_0$ would be determined from the coordinates in Eq. (3.3) using the condition

$$\langle X(t+t_0)|x(t)\rangle = A\, b_m c_m(t_0) + A\, a_m s_m(t_0) + B\, b_k c_m(t_0) + B\, a_k s_m(t_0) = 0\,. \tag{3.13}$$

The number of solutions $N_X[x]$ for this equation is a complicated function of $a_m, b_m, a_k$, and $b_k$. It turns out to be somewhat easier to use polar coordinates. Let us write an arbitrary linear

[2] In this case, we can alternatively restrict the range $0 < t_0 \leqslant \frac{T}{m}$ and take $N_X[x] = 1$, but restricting the range is untenable in more complicated examples.

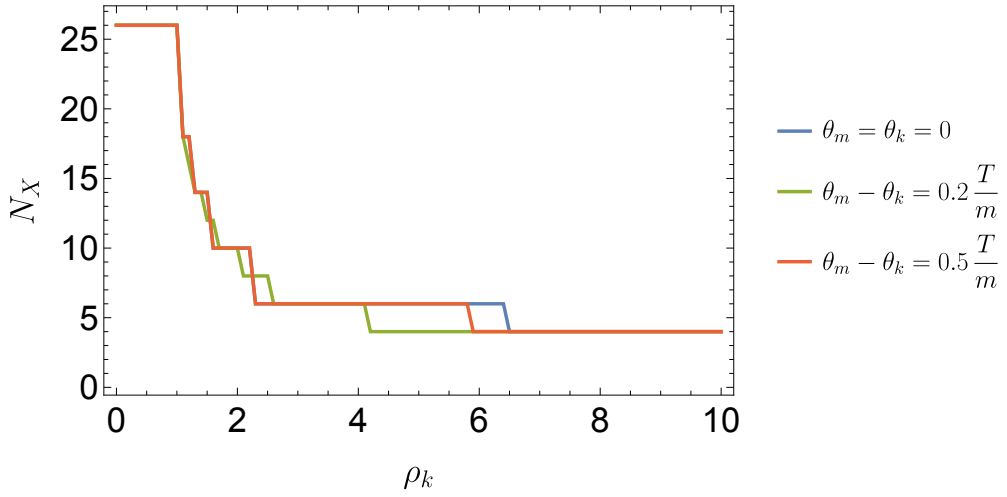



Figure 1: For the free quantum mechanical theory where the coefficient of a linear combination of two normal modes is traded for a collective coordinate, the intersection number $N_X$ cannot be neglected. $N_X$ is integer-valued. After integrating over the modulus $\rho_m$, we find that $N_X$ depends on the modulus $\rho_k$ and the relative angle $\theta_m - \theta_k$. Shown are the value of $N_X$ for $m = 26$ and $k = 4$ for a few choices of $\theta_m - \theta_k$.

combination of the $m$ and $k$ sines and cosine modes as

$$X(t) = R_m s_m(t + \phi_m) + R_k s_k(t + \phi_k)\,. \tag{3.14}$$

Changing to the polar coordinates $a_n = \rho_n s_n(\theta_n)$ and $b_n = \rho_n c_n(\theta_n)$ for convenience, we find the overlap

$$\langle X(t + t_0)|x(t)\rangle = R_m \rho_m s_m(t_0 + \theta_m + \phi_m) + R_k \rho_k s_k(t_0 + \theta_k + \phi_k) \tag{3.15}$$

which only involves two frequencies, as expected.

We now want to replace the $a_m, b_m, a_k, b_k$ integrals from the original partition function, which gave rise to a $\frac{1}{\lambda_m}\frac{1}{\lambda_k}$ term, with new integrals in our polar coordinates. As a cross check, we see that before introducing collective coordinates we have

$$Z_{\text{polar}} = \left(\frac{\lambda_m}{2\pi}\frac{\lambda_k}{2\pi} Z_0\right) \frac{m\pi}{T}\frac{k\pi}{T} \int_0^\infty d\rho_m \rho_m e^{-\frac{1}{2}\lambda_m \rho_m^2} \int_0^\infty d\rho_k \rho_k e^{-\frac{1}{2}\lambda_m \rho_m^2} \int_0^{\frac{2T}{m}} d\theta_m \int_0^{\frac{2T}{k}} d\theta_k = Z_0\,, \tag{3.16}$$

as expected. Here $\frac{\lambda_m}{2\pi}\frac{\lambda_k}{2\pi} Z_0$ represents the partition function where we have performed all integrals except for those over $a_m, b_m, a_k, b_k$. The $\frac{m\pi}{T}\frac{k\pi}{T}$ factors arise from the Jacobian induced by our polar coordinates.

Now we move to collective coordinates by inserting into the partially-integrated partition function our Eq. (2.12) which includes an integral over $t_0$, a $\delta$-function, and a Jacobian. Doing so and then shifting $\theta_m \to \theta_m - \phi_m$, $\theta_k \to \theta_k - \phi_k$, rescaling $\rho_m \to \frac{\rho_m}{R_m}, \rho_k \to \frac{\rho_k}{R_k}$, taking $\rho_k \to \rho_k \rho_m$, and performing the $\rho_m$ integral leads to

$$Z_{\rm coll} = Z_0 \times \left(\frac{\lambda_m}{2\pi}\frac{\lambda_k}{2\pi}\right)\frac{m\pi}{T}\frac{k\pi}{T}\int_0^\infty d\rho_k \frac{2\rho_k R_k^2 R_m^2}{(R_k^2\lambda_m + R_m^2\lambda_k\rho_k^2)^2}\int_0^{\frac{2T}{m}} d\theta_m \int_0^{\frac{2T}{k}} d\theta_k$$
$$\times \frac{1}{N_X(\rho_k,\theta_k,\theta_m)}\int_0^T dt_0\,\delta[s_m(\theta_m)+\rho_k s_k(\theta_k)]\left|\frac{m\pi}{T}c_m(\theta_m)+\rho_k\frac{k\pi}{T}c_k(\theta_k)\right|. \quad (3.17)$$

One can check that without the second line, we still have $Z_{\rm coll} = Z_0$. The intersection number $N_X$ counts the number of times the $\delta$-function fires for a given path. We can compute the intersection number by counting the roots of $\sin(\pi m(t_0+\theta_m)/T) + \rho_k \sin(\pi k(t_0+\theta_k)/T)$ on the interval $0 < t_0 < T$. Due to periodicity, the intersection number depends only on $\theta_k - \theta_m$. The function for $m = 26$ and $k = 4$ with three choices of $\theta_k - \theta_m$ is shown in Fig. 1.

To proceed, one can first observe that the integral is periodic in $\theta_m$ and $\theta_k$, so we can extend their limits of integration to $0 < \theta_m, \theta_k < T$ and pull out an overall factor of $\frac{4}{mk}$. Then we change variables from $(\theta_m, \theta_k)$ to $\theta_+ = \frac{1}{2}(\theta_m + \theta_k)$ and $\theta_- = \theta_m - \theta_k$ and integrate over $\theta_+$. In this integral, the $\delta$-function fires the number of times given by the number of solutions of $s_m(\theta_+ + \frac{1}{2}\theta_-) + \rho_k s_k(\theta_+ - \frac{1}{2}\theta_-)$ for $\theta_+$. But this number is identical to $N_X(\rho_k, \theta_-)$, so that again $Z_{\rm coll} = Z_0$ as expected. One can of course integrate over $\theta_m$ or $\theta_k$ first (rather than $\theta_+$) to get the same result, but the integrals are much more difficult.

In the general case where $X(t)$ has components of arbitrary frequency, the intersection number $N_X(\rho_n, \theta_n)$ would be a function of the entire configuration space. Thus the intersection number is generally a horrible mess, but accounting for the intersection number in the path integral is nonetheless required for collective coordinates to work. Luckily, for some choices of $X(t)$ the intersection number has a more physical interpretation and can be managed either analytically or approximately. A representative example is the double-well potential in quantum mechanics to which we now turn.

## 4 Double well potential

For the next example, we consider the symmetric double well potential in quantum mechanics whose Euclidean action is

$$S[x] = \int_{-T/2}^{T/2} dt\left[\frac{m}{2}\dot{x}^2 + V(x)\right], \quad V(x) = \frac{m\omega^2}{8a^2}(x^2 - a^2)^2\,. \quad (4.1)$$

There are two dimensionless scales in the problem: $ma^2\omega/\hbar$ which indicates the size of the barrier and $\omega T/\hbar$ which indicates time. We will set $m = \hbar = 1$ but keep $a, \omega$ and $T$ to assist with the physical interpretation of various limits.

The main quantity of interest in this theory to compute with path integral methods is the splitting of the two lowest energy eigenstates. This splitting, and the entire spectrum, have been studied in great detail both analytically and numerically [13, 23–29]. It is nevertheless difficult to find a rigorous analysis of the path integral computation of the energy splitting. Typically treatments are not only cavalier with regard to the introduction of collective coordinates, but also with regard to

the associated Jacobian and finite-time regulation. We will attempt to clarify all of these issues in this section.

Since the potential $V(x)$ is symmetric, the parity operator $P$ which takes $x \to -x$ commutes with the Hamiltonian. Thus energy eigenstates are parity even or parity odd. In the limit $a^2\omega \to \infty$, the barrier becomes enormous and the spectrum is that of two harmonic oscillators with frequency $\omega$. That is, $E_0^+ = E_0^- = \frac{\omega}{2}$ and the higher modes are spaced by $\omega$. Expanding the potential around a minimum and working perturbatively in the limit of large but finite $a^2\omega$, the energies receive corrections but the even and odd modes remain degenerate to all orders. The WKB approximation can be used to compute the energy splitting with the leading order result [13, 14, 29, 30]

$$\Delta_n^{\rm WKB} = E_n^- - E_n^+ = \omega \frac{2}{\sqrt{2\pi}} \left(8\omega a^2\right)^{n+1/2} e^{-\frac{2\omega a^2}{3}} . \tag{4.2}$$

In particular, the ground state splitting can be written as

$$\Delta_0^{\rm WKB} = 4\omega \sqrt{\frac{\omega a^2}{\pi}} e^{-\frac{2\omega a^2}{3}} . \tag{4.3}$$

In fact, the energies can be determined for this system by solving the Ricatti equation iteratively though the Exact WKB method relatively easily to arbitrary order [12, 28, 31–34]. However, our focus here is not on precision but on the role of collective coordinates in the path integral version of the calculation.

To compute the energy splitting with the path integral, it is helpful to set up the problem so that the leading result as $T \to \infty$ is directly proportional to $\Delta_0$. Defining $E_n^\pm = E_n \mp \frac{\Delta_n}{2}$ the partition function is

$$Z = {\rm Tr}[e^{-HT}] = \sum_n (e^{-E_n^+ T} + e^{-E_n^- T}) , \tag{4.4}$$

and the twisted partition function is

$$Z_P = {\rm Tr}[Pe^{-HT}] = \sum_n (e^{-E_n^+ T} - e^{-E_n^- T}) . \tag{4.5}$$

$Z_P$ is useful for extracting the energy since $E_n^+ = E_n^-$ to all orders in perturbation theory; that is, the splittings are non-perturbative. Since $Z - Z_P$ involves only odd modes and $Z + Z_P$ involves only even modes we see that

$$\lim_{T\to\infty} -\frac{1}{T} \ln\left(\frac{Z - Z_P}{Z + Z_P}\right) = E_0^- - E_0^+ = \Delta_0 . \tag{4.6}$$

The above formula makes no assumption about $\Delta_0$ being small, but is useful only if $Z$ and $Z_P$ are known non-perturbatively in $\Delta_0$. If we only know $Z_P$ in the small $\Delta_n$ limit, then we can use

$$\lim_{T\to\infty} \frac{2}{T}\Delta_n \lim_{\Delta_n\to 0} \frac{1}{\Delta_n}\frac{Z_P}{Z} = \lim_{T\to\infty} \frac{2}{T} \frac{\sum e^{-E_n^+ T}(\Delta_n T)}{\sum e^{-E_n^+ T}} = \Delta_0 . \tag{4.7}$$

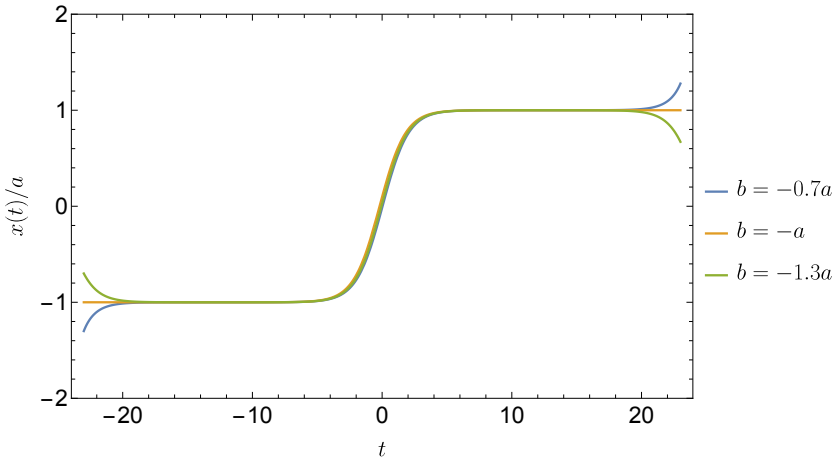



Figure 2: The twisted partition function sums over paths starting and ending at any $b$. At finite $T$, the double well admits exact instanton solutions for any $b$. For any $b$ and large values of $T$, these solutions spend most of the time in one of the two wells.

This formula will allow us to compute the splitting of the lowest energy states directly with the path integral: if we compute $Z_P$ and $Z$ in the saddle-point approximation and find at small splitting and large $T$ that $Z_P \sim ZAT$, then the energy splitting is $\Delta_0 = 2A$.

The twisted partition function can be computed with the path integral using twisted boundary conditions:

$$Z_P = \mathrm{Tr}(Pe^{-HT}) = \mathcal{N}\int_{-\infty}^{\infty} db\ \langle -b|e^{-HT}|b\rangle = \mathcal{N}\int_{-\infty}^{\infty} db \int_{x(-\frac{T}{2})=-b}^{x(\frac{T}{2})=b} \mathcal{D}x\ e^{-S[x]}\,. \tag{4.8}$$

All paths involved go from $-b$ to $b$ in time $T$ and we integrate over $b$ to take the trace. For the normalization $\mathcal{N}$, we can match to the partition function of the harmonic oscillator.

For any $b > 0$, the classical path from $-b$ to $b$ for a large time $T$ is exponentially close to the $T = \infty$ instanton

$$x_I(t) = a\tanh\frac{\omega t}{2}\,. \tag{4.9}$$

For $b < 0$, the classical path from $-b$ to $b$ for a large time $T$ is closer to the anti-instanton solution $-x_I(t)$. Both the instanton and anti-instanton have an action given by

$$S_I \equiv S[x_I] = a^2\omega\tanh\frac{\omega T}{4} - \frac{a^2\omega}{3}\tanh^3\frac{T\omega}{4} = \frac{2a^2\omega}{3} + \mathcal{O}(e^{-T\omega})\,. \tag{4.10}$$

In particular for $b \neq a$ these (anti-)instanton paths start at $-b$, shoot up to the top of hill of the inverted potential, stay there for almost the entire time, roll across, and then stay at the other hill, ending at $b$ after a total time $T$. Examples (found numerically with the shooting method) that have $b$ slightly lower or larger than $a$ are shown in Fig. 2. In particular, the part of the motion between $-b$ and $-a$ is given by a solution to the Euclidean equations of motion

$$\tilde{x}_b(t) = -a\coth\frac{\omega t}{2}\,, \tag{4.11}$$

and provides the corresponding action

$$S[\tilde{x}_b(t)] = \int_a^b dx\ \sqrt{2V(x)} = \frac{\omega}{6a}(a-b)^2(2a+b)\,. \tag{4.12}$$

Thus in the large $T$ limit, we find that the classical path from $-b$ to $b$ ($b>0$) has action

$$S[x_b(t)] = S[x_I] + 2S[\tilde{x}_b(t)] = S_I + \frac{\omega}{3a}(a-b)^2(2a+b)\,. \tag{4.13}$$

Note that when $\left|\frac{b}{a}\right| \gg 1$ then $S[x] \approx \frac{2a^2\omega}{3} + \frac{\omega|b|^3}{3a}$ and so the part of the action arising from the path approaching the hills is significant. The effect is that the contributions to the twisted partition function with $|b| \gg a$ will be small, however we still need to include the $b$ dependence to get the numerical prefactor correct.

For the fluctuations, the tail regions before and after $|x| \approx a$ give exponentially small contributions, so we can expand paths as $x(t) \to x_I(t) + x(t)$. Then the twisted partition function in the leading saddle point approximation becomes

$$Z_P \approx 2\mathcal{N}\int_0^\infty db\ e^{-S_I}e^{-2S[\tilde{x}_b(t)]}\int_{x(-\frac{T}{2})=0}^{x(\frac{T}{2})=0}\mathcal{D}x\ \exp\left[-\frac{1}{2}\int_{-\frac{T}{2}}^{\frac{T}{2}}dt\,x(t)[-\partial_t^2+V''(x_I)]x(t)\right] \tag{4.14}$$

$$\approx \mathcal{N}\sqrt{\frac{4\pi}{\omega}}e^{-S_I}\int_{x(-\frac{T}{2})=0}^{x(\frac{T}{2})=0}\mathcal{D}x\ \exp\left[-\frac{1}{2}\int_{-\frac{T}{2}}^{\frac{T}{2}}dt\,x(t)[-\partial_t^2+V''(x_I)]x(t)\right] \tag{4.15}$$

where we have performed the integral over $b$ in the large $S_I$ limit. The factor of 2 in front accounts for the anti-instanton solutions, which go from $-b$ to $b$ where $b<0$, and give the same contribution as the instanton solutions. Note that the instantons saturate the twisted boundary conditions and the fluctuations are Dirichlet. Next, we want to compute the path integral in the leading saddle-point approximation, carefully treating the collective coordinate and Gaussian integration over the fluctuations.

### 4.1 Normal modes at finite $T$

Fluctuations around the instanton are described by normal modes satisfying

$$\left[-\partial_t^2 + V''(x_I)\right]x_n = \lambda_n x_n \tag{4.16}$$

where $V''(x_I) = \omega^2\left(1-\frac{3}{2}\operatorname{sech}^2\frac{\omega t}{2}\right)$. Upon rescaling $t \to t/2$, we recognize that Eq. (4.16) is equivalent to the time-independent Schrödinger equation with a Pöschl-Teller potential [35] having $\lambda = 2$. The spectrum of the differential operator in Eq. (4.16) has bound states and a continuum. The continuum becomes discrete at finite $T$ and we can write a trajectory as

$$x(t) = x_I(t) + a_0x_0(t) + a_1x_1(t) + \sum_k(a_k^+x_k^+(t) + a_k^-x_k^-(t)) \tag{4.17}$$

with modes defined as follows.

The first bound state is[3]

$$x_0(t) = \sqrt{\frac{3\omega}{8}} \frac{1}{\cosh^2 \frac{\omega t}{2}} = \frac{\sqrt{6}}{2} \dot{x}_I(t), \quad \lambda_0 \approx 0\,. \tag{4.18}$$

Note that $x_0(t)$ is proportional to the velocity of the instanton.

The second bound state has odd parity and finite eigenvalue:

$$x_1(t) = \frac{\sqrt{3\omega}}{2} \frac{\sinh \frac{\omega t}{2}}{\cosh^2 \frac{\omega t}{2}}, \quad \lambda_1 = \frac{3}{4}\omega^2\,. \tag{4.19}$$

For $T \to \infty$ the remaining modes form a continuum. However, if we consistently maintain boundary conditions at finite $T$, these remaining modes are discrete. The eigenvalue equation is solved by

$$x_k(t) = \frac{1}{4}\left(1 + 4\frac{k^2}{\omega^2} + 6i\frac{k}{\omega}\tanh\frac{\omega t}{2} - 3\tanh^2\frac{\omega t}{2}\right) e^{ikt}, \quad \lambda_k = \omega^2 + k^2\,. \tag{4.20}$$

Paths $x_k(-t)$ are also eigenstates with the same eigenvalues. Since the Pöschl-Teller potential is symmetric, it is sensible to use parity eigenstates. Explicitly, these modes are

$$\begin{aligned} x_k^+(t) &= N_k[x_k(t) + x_k(-t)] \\ &= \frac{N_k}{2}\left[\left(1 + \frac{4k^2}{\omega^2}\right)\cos(kt) - \frac{6k}{\omega}\sin(kt)\tanh\left(\frac{\omega t}{2}\right) - 3\cos(kt)\tanh^2\left(\frac{\omega t}{2}\right)\right] \end{aligned} \tag{4.21}$$

with $N_k = \sqrt{\frac{2\omega^4}{T(4k^4+5k^2\omega^2+\omega^4)}}$, and

$$\begin{aligned} x_k^-(t) &= -iN_k[x_k(t) - x_k(-t)] \\ &= \frac{N_k}{2}\left[\left(1 + \frac{4k^2}{\omega^2}\right)\sin(kt) + \frac{6k}{\omega}\cos(kt)\tanh\left(\frac{\omega t}{2}\right) - 3\sin(kt)\tanh^2\left(\frac{\omega t}{2}\right)\right]. \end{aligned} \tag{4.22}$$

Note that the modes $x_k^\pm(t)$ are manifestly real. As written, the modes are normalized at large $T$ to $\langle x_k^\pm | x_k^\pm \rangle = 1 + \mathcal{O}\left(\frac{1}{T}\right)$. The eigenvalues are quantized by the boundary conditions $x_k^\pm\left(\frac{T}{2}\right) = 0$ which leads at large $T$ to

$$k^-(n) \approx \frac{2n\pi}{T}, \quad k^+(n) \approx \frac{(2n-1)\pi}{T}, \quad n = 1, 2, 3, \ldots\,. \tag{4.23}$$

As $T \to \infty$, the modes approach

$$x_k^\pm(t) \to s_n(t), \quad \lambda_n \approx \omega^2 + \left(\frac{n\pi}{T}\right)^2, \quad n = 1, 2, 3, \ldots \tag{4.24}$$

[3]This function as written has eigenvalue exactly 0, but does not satisfy the Dirichlet boundary conditions. If we impose the Dirichlet boundary conditions at $t = \pm\frac{T}{2}$, the eigenvalue gets lifted to $\lambda_0 \approx e^{-\omega T}$ (see [24, App. E] or [30, Eq. 23.55b]). While an exponentially small eigenvalue renders finite the functional determinant in the path integral, it does not produce the correct answer for the twisted partition function because higher order terms in fluctuations in the $x_0$ direction are important. Collective coordinates allow for inclusion of fluctuations due to time translations to all orders.

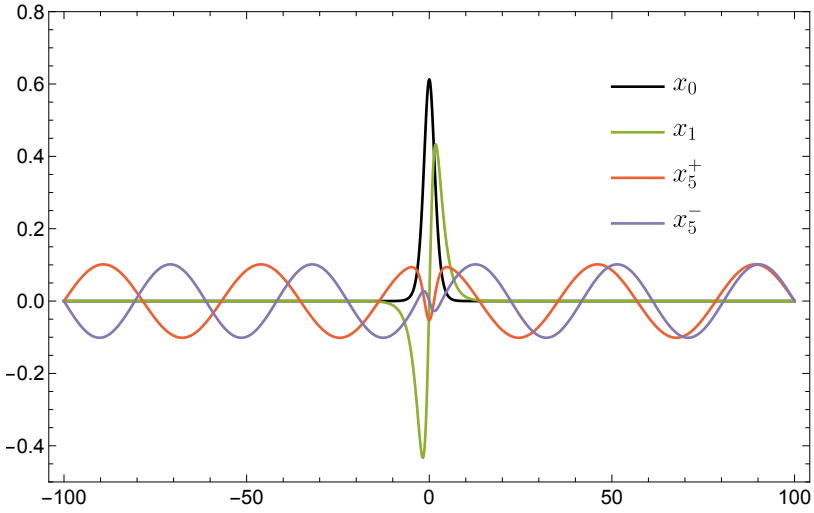



Figure 3: Normal modes of the double well at finite $T$ (plotted with $T = 200$), with Dirichlet boundary conditions at $t = \pm\frac{T}{2}$. Modes have even or odd parity. The two bound state modes $x_0$ and $x_1$ are shown, as are two of the continuum modes $x_k^\pm$ with $k = 5$. These continuum modes are similar to sine functions except near $t = 0$.

with $s_n$ as in Eq. (3.4). As $T \to \infty$ the $x_k^+$ become the even-$n$ modes and $x_k^-$ become the odd-$n$ modes. The modes $x_k^\pm$ and $s_n$ differ only close to $t = 0$, as shown in Fig. 3. For small $t$, it is enough to know the values of the modes and their derivatives at $t = 0$:

$$x_k^+(0) = \frac{1}{\sqrt{2T}}, \quad x_k^-(0) = 0, \quad \dot{x}_k^+(0) = 0, \quad \dot{x}_k^-(0) = \frac{4\sqrt{2}n\pi}{T^{3/2}}\,. \tag{4.25}$$

When $t$ is not small, the $s_n$ are excellent approximations to the $x_k^\pm$.

## 4.2 Collective coordinates

Now we would like to change to collective coordinates. To do so we need to consider the intersection number $N_X[x]$. This is the number of times a given path $x(t)$ has vanishing projection against $x_0(t+t_0)$ over the integration range of $t_0$. Parameterizing $x(t)$ as in Eq. (4.17), we need to find the number of zeros of

$$\begin{aligned}\big\langle x_0(t+t_0)\big|x(t)\big\rangle = \big\langle x_0(t+t_0)\big|x_I(t)\big\rangle + a_0\big\langle x_0(t+t_0)\big|x_0(t)\big\rangle + a_1\big\langle x_0(t+t_0)\big|x_1(t)\big\rangle \\ + \sum_k \Big[a_k^+\big\langle x_0(t+t_0)\big|x_k^+(t)\big\rangle + a_k^-\big\langle x_0(t+t_0)\big|x_k^-(t)\big\rangle\Big]\end{aligned} \tag{4.26}$$

for fixed values of the coordinates $\{a_0, a_1, a_k^\pm\}$. To evaluate the above expression we note that there are three relevant time scales: the overall time $T$, the collective coordinate time $t_0$, and the inverse frequency $\omega^{-1}$ associated with the curvature of the potential. The two bound states $x_0$ and $x_1$ have

support of order $\omega^{-1}$ while the continuum modes have support over all $t$. For $t_0 \ll \omega^{-1} \ll T$, we have

$$\langle x_0(t+t_0)|x(t)\rangle = -\sqrt{\frac{2a^2\omega}{3}}\, t_0 + a_0 - \frac{3\pi}{16\sqrt{2}}\, a_1 \omega t_0 + \cdots . \tag{4.27}$$

If we are permitted to drop the $a_1$ term and the additional terms thereafter, the above reduces to $t_0 \approx a_0\, S_I^{-1/2}$. This fiducial trade-off between $a_0$ and $t_0$ is the change to collective coordinates one normally anticipates: motion along the direction of the zero mode $x_0$ is exchanged for motion along $t_0$. However, the relationship between $a_0$ and $t_0$ only holds for $t_0 \ll \omega^{-1} \ll T$. For larger $t_0$ the other terms in Eq. (4.27) are important.

In the regime $\omega^{-1} \ll t_0 \ll T$, the detailed structure of $x_0(t+t_0)$ of order $\omega^{-1}$ cannot be resolved. In fact, as $\omega t_0 \to \infty$, the mode $x_0(t-t_0) \sim \frac{1}{\cosh^2 \frac{\omega(t-t_0)}{2}}$ approximates a $\delta$-function:

$$x_0(t-t_0) \approx \sqrt{\frac{6}{\omega}}\, \delta(t-t_0)\,, \quad \omega^{-1} \ll t_0 \ll T\,, \tag{4.28}$$

so that

$$\big\langle x_0(t-t_0)\big|x(t)\big\rangle \approx \sqrt{\frac{6}{\omega}} \int dt\, \delta(t-t_0) x(t) = \sqrt{\frac{6}{\omega}}\, x(t_0)\,. \tag{4.29}$$

The equation above leads to a nice interpretation of $N_{x_0}[x]$: it is the number of times $x(t_0) = 0$, i.e. the number of times a given path $x(t)$ crosses $x = 0$ on its way from $-b$ to $b$. Similarly, in the same limit the Jacobian in Eq. (2.12) $\mathcal{J} = \langle \dot{x}_0(t)|x(t)\rangle \approx -\dot{x}(0)$ becomes the velocity of the path at $t_0$, i.e. as the path crosses the origin.

### 4.3 Energy splitting

Now let us use the collective coordinates to compute the energy splitting. To do so, it is easiest to parameterize paths by

$$x(t) = x_I(t+t_0) + \alpha_1 x_1(t+t_0) + \sum_k \Big[\alpha_k^+ x_k^+(t+t_0) + \alpha_k^- x_k^-(t+t_0)\Big] \tag{4.30}$$

and to use Eq. (2.10). The Jacobian evaluates to

$$\mathcal{J} = \big\langle \dot{x}_0(t)\big|x_I(t) + \alpha_1 x_1(t) + \sum_k \Big[\alpha_k^+ x_k^+(t) + \alpha_k^- x_k^-(t)\Big]\big\rangle \tag{4.31}$$

$$\approx -\sqrt{\frac{2\omega a^2}{3}} - \frac{3\pi\omega}{16\sqrt{2}}\alpha_1 + \sum_k \alpha_k^+ \frac{4\sqrt{2}n\pi}{T^{3/2}} \tag{4.32}$$

where $n$ refers to the approximate wavenumber from Eq. (4.23).

The action of a path $x(t)$ parameterized in collective coordinates is

$$S[x] = \frac{1}{2}\alpha_1^2\lambda_1 + \frac{1}{2}\sum_k \Big[(\alpha_k^+)^2\lambda_k^+ + (\alpha_k^-)^2\lambda_k^-\Big]\,. \tag{4.33}$$

Since $\lambda_1 = \frac{3}{4}\omega^2$ and $\lambda_k \approx \omega^2 + \left(\frac{n\pi}{T}\right)^2$, the above action implies that the path integral is dominated by the regime $\alpha_1 \lesssim \omega^{-1}$ and $\alpha_k^{\pm} \lesssim \omega^{-1}$ (and even for $n > T\omega$ the scaling is $\alpha_k^{\pm} \lesssim \frac{T}{n} \lesssim \omega^{-1}$ as well). Thus, for purposes of computing the energy splitting where we are interested in the limit of large $a^2\omega$ and large $T$, the Jacobian factor in Eq. (4.32) reduces to

$$\mathcal{J} \approx -\sqrt{\frac{2\omega a^2}{3}} = -\sqrt{S_I}\,. \tag{4.34}$$

This is the result usually found in the literature.

In collective coordinate, the twisted partition function is then

$$Z_P = \mathcal{N}\sqrt{\frac{4\pi}{\omega}}\sqrt{S_I}\,e^{-S_I}\int_{-\frac{T}{2}}^{\frac{T}{2}}\frac{dt_0}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{d\alpha_1}{\sqrt{2\pi}}\prod_k\int_{-\infty}^{\infty}\frac{d\alpha_k^{\pm}}{\sqrt{2\pi}}\;e^{-\frac{1}{2}\lambda_1\alpha_1^2-\frac{1}{2}\lambda_k(\alpha_k^{\pm})^2}\frac{1}{N_{x_0}[x]} \tag{4.35}$$

where, as discussed, $N_{x_0}[x]$ counts the number of times a path $x(t)$ crosses $x = 0$. Although $N_{x_0}[x]$ is an incredibly complicated functional of paths, it can be simplified for the calculation of the twisted partition function in the large $T$ limit.

We reiterate that according to Eq. (4.29), $N_{x_0}[x]$ counts the number of times the path crosses the origin. Since all paths integrated over in the twisted partition function go from $-b$ to $b$ they must all cross the origin at least once – this is essential, otherwise the change to collective coordinate $t_0$ would not be valid. Because the potential $V(x)$ is largest at $x = 0$, larger values of $N_{x_0}[x]$ will give exponentially suppressed contributions to the path integral compared with the contributions at $N_{x_0}[x] = 1$. Consequently $N_{x_0}[x] \geqslant 1$ and paths with $N_{x_0}[x] > 1$ are negligible. Thus we can approximate $N_{x_0}[x] = 1$. So we have

$$Z_P \approx \mathcal{N}\sqrt{\frac{4\pi}{\omega}}T\sqrt{\frac{S_I}{2\pi}}\,e^{-S_I}\int_{-\infty}^{\infty}\frac{d\alpha_1}{\sqrt{2\pi}}\prod_k\int_{-\infty}^{\infty}\frac{d\alpha_k^{\pm}}{\sqrt{2\pi}}\;e^{-\frac{1}{2}\lambda_1\alpha_1^2-\frac{1}{2}\lambda_k(\alpha_k^{\pm})^2} \tag{4.36}$$

$$= \mathcal{N}\sqrt{\frac{4\pi}{\omega}}T\sqrt{\frac{S_I}{2\pi}}\,e^{-S_I}\sqrt{\frac{1}{\lambda_1}}\prod_k\sqrt{\frac{1}{\lambda_k^{\pm}}} \tag{4.37}$$

$$= \mathcal{N}\sqrt{\frac{4\pi}{\omega}}T\sqrt{\frac{S_I}{2\pi}}\,e^{-S_I}\left[\det{}'(-\partial_t^2 + V''(x_I))\right]^{-1/2}\,. \tag{4.38}$$

As such we have reduced our calculation to the textbook treatment.

We can also understand the $N_{x_0}[x] \approx 1$ replacement from a different perspective, which allows a more straightforward generalization to the QFT case in Section 5. If we solve the condition $\langle x_0(t+t_0)|x(t)\rangle = 0$ for $t_0$ using Eq. (4.27) and expand at large $S_I$, we find

$$t_0 = \frac{a_0}{\sqrt{S_I}} - \frac{3\pi\omega}{16\sqrt{2}S_I}\,a_0\,a_1 + \cdots\,. \tag{4.39}$$

Here $a_n \lesssim \omega^{-1}$ just as $\alpha_n \lesssim \omega^{-1}$. However, the mode with zero eigenvalue with coordinate $a_0$ is unsuppressed. Considering only the zero mode, the condition determining $t_0$ is

$$0 = \langle x_0(t+t_0)|x_I(t)\rangle + a_0\langle x_0(t+t_0)|x_0(t)\rangle \tag{4.40}$$

$$= \frac{3}{\omega}\sqrt{S_I}\,\frac{(\omega t_0 - \sinh(\omega t_0))}{\cosh \omega t_0 - 1} + a_0\,\frac{3(-2 + \omega t_0 \coth \frac{\omega t_0}{2})}{\cosh \omega t_0 - 1}\,. \tag{4.41}$$

The first function on the second line is a monotonically decreasing odd function of $t_0$ and the second function is a positive monotonically decreasing function of $|t_0|$ with a maximum at $t_0 = 0$; these conditions imply that there is exactly one solution to (4.40) for a given $a_0$. There are more solutions for $t_0$ if we turn on some of the other $a_n$'s, analogous to the non-linear terms in Eq. (4.39). Such additional solutions for $t_0$ require large values of $a_n \gtrsim \sqrt{S_I}$. However, when $a_n \gtrsim \omega^{-1}$, the path integral becomes exponentially suppressed. Thus any path with $N_{x_0}[x] > 1$ is exponentially suppressed by the contributions of large $a_n$ and so we can simply set $N_{x_0}[x] = 1$.

The above argument can be stated more physically. We expand around a particular instanton which crosses the origin at $t_0 = 0$. We can shift the path using the zero mode, changing the time the path crosses the origin. But to get an additional crossing, the quantum corrections must be enormous: as large as the classical contribution. Such large corrections are exponentially suppressed.

There is a close analogy between this calculation, where the energy splitting is given by paths that cross $x = 0$ exactly once, and the calculation of the tunneling rate in an asymmetric double well. In the latter, one way to derive the formula for the decay rate is to restrict to paths that cross the origin only once [5, 6]. This restriction can be understood physically: if the paths do not cross the origin at all, they do not mediate transitions from the false vacuum to the true one; to find a well-defined decay rate one must be in a regime where the back-scattering off the far end of the potential is irrelevant, which can be enforced by demanding at most one crossing. Thus the decay rate is determined by paths that cross the origin exactly once, akin to the energy splitting in the symmetric double well.

For completeness, we finish the energy-splitting calculation. The product over continuum eigenvalues is nearly the same as for a simple harmonic oscillator. For the harmonic oscillator, one expands around the classical solution $x_{\rm cl}(t) = b\cosh(\omega t)/\cosh(\omega T/2)$ where the fluctuations have eigenvalues $\lambda_n^{\rm SHO} = \omega^2 + (n\pi/T)^2$, and so the partition function is

$$Z_{\rm SHO} = \mathcal{N}\int_{-\infty}^{\infty} db\; e^{-\omega b^2}\prod_n \int_{-\infty}^{\infty}\frac{da_n}{\sqrt{2\pi}}\,e^{-\frac{1}{2}\lambda_n a_n^2} \tag{4.42}$$

$$= \mathcal{N}\sqrt{\frac{\pi}{\omega}}\prod_n \sqrt{\frac{1}{\lambda_n^{\rm SHO}}} = \mathcal{N}\sqrt{\frac{\pi}{\omega}}\,\det{}^{-1/2}(-\partial_t^2 + \omega^2) \tag{4.43}$$

which defines for us the normalization $\mathcal{N}$ in terms of $Z_{\rm SHO}$. For the double well, we can then plug $\mathcal{N}$ into Eq. (4.38), obtaining

$$Z_P = Z_{\rm SHO}\; 2T\sqrt{\frac{S_I}{2\pi}}\,e^{-S_I}\left(\frac{\det(-\partial_t^2 + \omega^2)}{\det'(-\partial_t^2 + V''(x_I))}\right)^{1/2}. \tag{4.44}$$

The ratio of determinants can be computed using the Gelfand-Yanglom method [36, 37] and yields

$12\omega^2$. Using Eq. (4.7), we find

$$\Delta_0 = 2\sqrt{\frac{6S_I}{\pi}}\, e^{-S_I}\omega \tag{4.45}$$

Recalling $S_I = \frac{2\omega a^2}{3}$, the above exactly matches the WKB result in Eq. (4.3). We have used that $Z \approx 2Z_{\rm SHO}$ in the small-energy-splitting limit.

## 4.4 Free twisted partition function

In the interacting theory, integration over the collective coordinate $t_0$ gave a factor of $T$ in the twisted partition function, as expected for the computation of the energy splitting. The free theory is also time-translation invariant, and so using collective coordinates would seem to give a factor of $T$. However, the free twisted partition function is just $Z_P = \frac{1}{2}$ (as we review shortly). So what is the fate of the putative factor of $T$? To find out, we need to properly compute the twisted partition function with action $S_0[x] = \int dt\, \frac{1}{2}\dot{x}^2$ using collective coordinates.

### Expected results

First we recall the expected result. The twisted partition function in the free theory can be computed as the zero frequency limit of the twisted partition function for the harmonic oscillator, namely

$$Z_P^0 = {\rm Tr}(Pe^{-HT}) = \lim_{\omega\to 0}\sum_{n=0}^{\infty} e^{-\omega\left(n+\frac{1}{2}\right)T}(-1)^n = \frac{1}{2}\,. \tag{4.46}$$

Alternatively, we can recall that the position-space propagator in quantum mechanics is

$$D(a,b) = \langle b|e^{-HT}|a\rangle = \int_{-\infty}^{\infty}\frac{dp}{2\pi}\langle b|e^{-\frac{1}{2}p^2T}|p\rangle\langle p|a\rangle = \int_{-\infty}^{\infty}\frac{dp}{2\pi}\, e^{ip(b-a)}e^{-\frac{p^2}{2}T} = \sqrt{\frac{1}{2\pi T}}\, e^{-\frac{(b-a)^2}{2T}}\,. \tag{4.47}$$

So the twisted partition function is

$$Z_P^0 = \int_{-\infty}^{\infty} db\; D(b,-b) = \int_{-\infty}^{\infty} db\,\sqrt{\frac{1}{2\pi T}}\, e^{-\frac{2b^2}{T}} = \frac{1}{2}\,. \tag{4.48}$$

In the path integral approach, we compute

$$Z_P^0 = \mathcal{N}\int_{-\infty}^{\infty} db \int_{x\left(-\frac{T}{2}\right)=-b}^{x\left(\frac{T}{2}\right)=b} \mathcal{D}x\; e^{-S[x]}\,. \tag{4.49}$$

Here the classical path going from $-b$ to $b$ in time $T$ is $x_c(t) = \frac{2t}{T}b$ with action $S_0[x_c] = \frac{2b^2}{T}$, as expected from the exponent of $D(b,-b)$. A general path is then parameterized by

$$x(t) = x_c(t) + c_n s_n(t), \quad n = 1, 2, 3, \ldots\,, \tag{4.50}$$

with $s_n(t)$ as in Eq. (3.4). These $s_n(t)$ modes satisfy $\partial_t^2 s_n = -\lambda_n s_n$ with $\lambda_n = \left(\frac{n\pi}{T}\right)^2$. Then

$$Z_P^0 = \mathcal{N}\int_{-\infty}^{\infty} db\; e^{-\frac{2b^2}{T}} \prod_{n=1}^{\infty}\int_{-\infty}^{\infty}\frac{dc_n}{\sqrt{2\pi}}\; e^{-\frac{1}{2}\lambda_n c_n^2} = \mathcal{N}\frac{\sqrt{2\pi T}}{2}\prod_n \frac{T}{n\pi} \tag{4.51}$$

so that to have $Z_P^0 = \frac{1}{2}$ we need

$$\mathcal{N} = \frac{1}{\sqrt{2\pi T}}\left[\prod_n \frac{T}{n\pi}\right]^{-1} . \tag{4.52}$$

**Collective coordinates**

Now we attempt to compute $Z_P^0$ using collective coordinates in the double-well basis. Noting that the formula in Eq. (2.12) is basis-independent, the only feature from the double well basis we carry over is that the function $X(t) = x_0(t)$ is the one used to project out the collective coordinate. Indeed, if we had tried to use the parametrization in Eq. (4.17) for the paths in the free theory case we would have run into problems because $x_I(t)$ is not a saddle point of the free action (in fact, it has much larger action than the classical path; for paths from $-a$ to $a$, $S_0[x_I] = \frac{a^2\omega}{3} \gg S_0[x_c] = \frac{2a^2}{T}$). Since $x_I(t)$ is not a saddle point, in attempting to use Eq. (4.17) in the path integral, we would find linear fluctuation terms which would require completing the square in the action, effectively reducing the computation to expanding around the classical path. The result would be the same as using the parameterization in Eq. (4.50) to begin with.

Marching forward, we want to evaluate

$$Z_P^{\text{coll},0} = \mathcal{N}\int_{-\infty}^{\infty} db\; e^{-\frac{2b^2}{T}} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt_0\; \prod_n \int_{-\infty}^{\infty}\frac{dc_n}{\sqrt{2\pi}}\; e^{-\frac{1}{2}\lambda_n c_n^2}\frac{1}{N_{x_0}[x]}\,\delta\Big[\big\langle x_0(t)\big|x(t)\big\rangle\Big]\,\Big|\big\langle x_0(t)\big|\dot{x}(t)\big\rangle\Big| . \tag{4.53}$$

As discussed in the double-well case, $\delta[\langle x_0(t)|x(t)\rangle]$ forces the path to cross $x = 0$ at $t = 0$ and $N_{x_0}[x]$ counts the number of times a given path crosses $x = 0$. To compute the argument of the $\delta$ function we need the projections

$$\langle x_0(t)|s_n(t)\rangle = \sqrt{\frac{2}{T}}\int_{-\infty}^{\infty} dt\; x_0(t)\sin\frac{n\pi\left(t+\frac{T}{2}\right)}{T} = \sqrt{\frac{12}{T\omega}}\sin\frac{n\pi}{2}\,Q_n \tag{4.54}$$

$$= (-1)^{\frac{n-1}{2}}\sqrt{\frac{12}{T\omega}}\,\delta_{n,\,\text{odd}}\,Q_n \tag{4.55}$$

which we calculate in the $T \to \infty$ limit, where

$$Q_n = \frac{\frac{n\pi^2}{T\omega}}{\sinh\frac{n\pi^2}{T\omega}} \approx 1 . \tag{4.56}$$

Thus the $\delta$ function involves only the odd modes, and sets $c_1 - c_3 + c_4 - c_5 + \cdots = 0$ up to corrections

in $\frac{1}{T}$. The Jacobian is

$$\langle x_0(t)|\dot{x}(t)\rangle = \langle x_0(t)|\dot{x_c}(t)\rangle + \sum_n c_n \langle x_0(t)|\dot{s}_n(t)\rangle = \frac{2b}{T}\sqrt{\frac{6}{\omega}} + \sum_n (-1)^{\frac{n}{2}} \sqrt{\frac{12}{\omega T}}\sqrt{\lambda_n}\, c_n\, Q_n\, \delta_{n,\,\text{even}} \tag{4.57}$$

which involves only the even modes. Since the $\delta$ function only involves the odd modes and the Jacobian only involves the even modes, the integration over the even and odd modes factorizes.

With the factor of $N_{x_0}[x]$, even though the path integral factorizes, the integrals are challenging. To show that $N_{x_0}[x]$ is necessary, it suffices to ignore it by setting it to 1 and show that the wrong answer results. Let us do so and behold the consequences of such ignorance.

The formulas

$$\int d^N c\; e^{-\pi \vec{c}\cdot A\cdot \vec{c}} |\vec{w}\cdot\vec{c}| = \frac{1}{\pi}\frac{\sqrt{\vec{w}\cdot A^{-1}\cdot\vec{w}}}{\sqrt{\det(A)}} \tag{4.58}$$

and

$$\int d^N c\; e^{-\pi \vec{c}^{\,2}} \delta(\vec{w}\cdot\vec{c}) = \frac{1}{\sqrt{\vec{w}^2}} \tag{4.59}$$

are helpful. For the odd modes we find

$$I_{\text{odd}} = \prod_{n\,\text{odd}} \int \frac{dc_n}{\sqrt{2\pi}}\, e^{-\frac{1}{2}c_n^2\lambda_n} \delta\left(c_n\sqrt{\frac{12}{\omega T}}\sin\frac{n\pi}{2}Q_n\right) = \left(\prod_{n\,\text{odd}}\sqrt{\frac{1}{\lambda_n}}\right)\sqrt{\frac{\omega}{3\pi T}} \tag{4.60}$$

and for the even modes we have

$$I_{\text{even}} = \int db\; e^{-\frac{2b^2}{T}} \prod_{n\,\text{even}} \int \frac{dc_n}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\lambda_n c_n^2} \left|\frac{2b}{T}\sqrt{\frac{6}{\omega}} + \sum_n (-1)^{\frac{n}{2}}\sqrt{\frac{12}{\omega T}}\sqrt{\lambda_n}\, c_n\, Q_n\, \delta_{n,\text{even}}\right| \tag{4.61}$$

$$= \left(\prod_{n\,\text{even}}\sqrt{\frac{1}{\lambda_n}}\right)\left(\sqrt{T} + \frac{3}{\omega\sqrt{T}} + \cdots\right). \tag{4.62}$$

So all together we obtain

$$Z_P^{\text{coll},0} = \frac{1}{\sqrt{2\pi T}}\left[\prod_n \sqrt{\frac{1}{\lambda_n}}\right]^{-1}\left(\prod_{n\,\text{odd}}\sqrt{\frac{1}{\lambda_n}}\right)\sqrt{\frac{\omega}{3\pi T}}\left(\prod_{n\,\text{even}}\sqrt{\frac{1}{\lambda_n}}\right)\sqrt{T}\int_{-T/2}^{T/2} dt_0 \tag{4.63}$$

$$= \sqrt{\frac{\omega T}{6\pi^2}}\,. \tag{4.64}$$

This does not agree with the correct answer $Z_P^0 = \frac{1}{2}$, as in Eq. (4.48). Indeed Eq. (4.64) is too large by a factor of $\sqrt{\omega T}$. Thus we confirm that using collective coordinates without the factor of $N_{x_0}[x]$ is incorrect.

To see that the right answer does emerge if $N_{x_0}[x]$ is included is challenging because of the extreme complexity of $N_{x_0}[x]$. However, incorporating $N_{x_0}[x]$ is essentially the same challenging problem discussed in Section 3. It does not seem that working in the double-well basis rather than a Fourier basis provides any additional insight, so we simply refer the reader back to Section 3 and move on to quantum field theory.

# 5 Quantum field theory

For our final example, we consider the $\lambda\phi^4$ quantum field theory in 4 Euclidean dimensions. The Lagrangian density is $\mathcal{L}[\phi] = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + V(\phi)$ with $V(\phi) = \frac{1}{4}\lambda\phi^4$. For $\lambda < 0$ this theory has a 5-parameter family of Fubini-Lipatov instanton saddle points [38, 39] called "bounces" given by

$$\phi_b^{R,\xi}(x) = \sqrt{\frac{8}{-\lambda}}\frac{R}{R^2 + (x^\mu + \xi^\mu)^2}\,, \tag{5.1}$$

with action

$$S_b = S[\phi_b] = -\frac{8\pi^2}{3\lambda}\,. \tag{5.2}$$

We will generally pick a fiducial member of this family to expand around and denote it by $\phi_b(r)$ with $R$ fixed and $\xi = 0$. Here $r = \sqrt{\sum_\mu x_\mu^2}$ is the Euclidean radial coordinate.

With $\lambda < 0$ the vacuum at $\phi = 0$ is unstable. The corresponding decay rate can be computed perturbatively as

$$\Gamma = \lim_{T\to\infty}\frac{1}{T}\mathrm{Im}\frac{\int_{C_b}\mathcal{D}\phi\, e^{-S[\phi_b]+\frac{1}{2}S''[\phi_b]\phi^2+\cdots}}{\int_{C_{\mathrm{FV}}}\mathcal{D}\phi\, e^{-\frac{1}{2}\int d^4x(\partial_\mu\phi)^2}} \tag{5.3}$$

where $C_b$ is a path of steepest descent passing though the bounce $\phi_b$ and $C_{\mathrm{FV}}$ is the expansion around the false vacuum $\phi = 0$. A derivation of this formula can be found in [7]. The numerator of this expression is expected to have a factor of $T\mathcal{V}$ from the translation invariance of the action (the $\xi^\mu$ moduli). Such a factor makes $\Gamma \propto \mathcal{V}$, i.e. the decay is a rate per unit volume, as expected in a translationally invariant theory. The denominator, which comprises an expansion around the free theory, is not expected to have a factor of $T\mathcal{V}$. The goal of this section is to understand why the numerator is proportional to $T\mathcal{V}$ but the denominator is not, despite the fact that both actions are translation and scale invariant.

## 5.1 Basis and cross check

Remarkably, the eigenstates for expanding around the bounce are known in closed form [7]. Even more remarkably, these same eigenstates also diagonalize fluctuations around the false vacuum. The modes are

$$\hat{\phi}_{ns\ell m}(r,\alpha,\theta,\phi) = \frac{\sqrt{2n+3}}{2\pi}\frac{1}{r}P_{n+1}^{-s-1}\left(\frac{R^2-r^2}{R^2+r^2}\right)\sqrt{\frac{\pi}{2}}\,Y^{s\ell m}(\alpha,\theta,\phi) \tag{5.4}$$

where $P_\ell^m(x)$ are associated Legendre polynomials and $Y^{s\ell m}(\alpha,\theta,\phi)$ are 3D spherical harmonics (which are also Legendre polynomials). We will use the "box" inner product

$$\langle\phi_i|\phi_j\rangle = -\int d^4x\,\phi_i(x)\Box\phi_j(x) \tag{5.5}$$

under which the modes are orthonormal, namely

$$\langle\hat{\phi}_{ns\ell m}|\hat{\phi}_{n's'\ell'm'}\rangle = -\int d^4x\,\hat{\phi}_{ns\ell m}\Box\hat{\phi}_{n's'\ell'm'} = \delta_{nn'}\delta_{ss'}\delta_{\ell\ell'}\delta_{mm'}\,. \tag{5.6}$$

The modes satisfy

$$\Box\,\hat{\phi}_{ns\ell m} = \lambda_n V''[\phi_b]\hat{\phi}_{ns\ell m}, \quad \lambda_n = \frac{(n+1)(n+2)}{6} = \frac{1}{3}, 1, 2, \frac{10}{3}, \ldots, \tag{5.7}$$

and there are $d_n = \frac{(n+1)(n+2)(2n+3)}{6} = 1, 5, 14, \ldots$ modes with eigenvalue $\lambda_n$.

The $n = 0$ mode with multiplicity $d_0 = 1$ is proportional to the bounce

$$\phi_{0000} = \frac{\sqrt{3}}{2\pi}\frac{R}{r^2+R^2} = \frac{1}{2\pi}\sqrt{\frac{-3\lambda}{8}}\,\phi_b = \frac{1}{2\sqrt{S_b}}\,\phi_b\,. \tag{5.8}$$

Normalizing with respect to our box inner product, the mode is independent of $\lambda$. Of the five $n = 1$ modes the one with $s = 0$ is the dilatation mode associated with changes in $R$, namely

$$\phi_d \equiv \phi_{1000} = \frac{\sqrt{5}R}{2\pi}\frac{R^2 - r^2}{(R^2+r^2)^2} = -\sqrt{\frac{5}{12S_b}}R\,\partial_R\phi_b\,. \tag{5.9}$$

The other four $n = 1$ modes, with $s = 1$, correspond to translations. They are linear combinations of the $\partial_\mu\phi_b$.

Field configurations around the bounce can be parameterized as

$$\phi(x) = \phi_b(x) + \sum_n c_n\hat{\phi}_n(x) \tag{5.10}$$

where we are using $n$ to abbreviate all the indices. The quadratic action around the bounce is

$$S[\phi] \approx S[\phi_b] + \sum_{n,\,m}\frac{1}{2}\int d^4x\, c_n\hat{\phi}_n[-\Box + V''(\phi_b)]c_m\hat{\phi}_m \tag{5.11}$$

$$= S_b + \sum_n \frac{1}{2}c_n^2\cdot\left(\frac{\lambda_n - 1}{\lambda_n}\right) \tag{5.12}$$

$$= S_b + \frac{1}{2}c_0^2\cdot(-2) + \frac{1}{2}c_1^2\cdot(0) + \frac{1}{2}c_2^2\cdot\left(\frac{1}{2}\right) + \cdots\,. \tag{5.13}$$

The unusual factor $(\lambda_n - 1)/\lambda_n$ is a result of the box inner product. If we had used the "$V$" inner product of Ref. [7], namely $\langle\phi_i|\phi_j\rangle_V = -\int d^4x\, V''[\phi_b]\phi_i\phi_j$, we would have found the action $S = S_b + \frac{1}{2}c_n^2(\lambda_n - 1) + \cdots$. For present purposes, the box inner product is preferable because it is well-defined even for $V = 0$ so we can use it in the free theory. Importantly, there is one mode ($n = 0$) with a negative eigenvalue and five modes ($n = 1$) with zero eigenvalue in either inner product.

Writing $\phi(x) = \sum_n b_n\hat{\phi}_n(x)$, the free action in these coordinates is

$$S[0] + \frac{1}{2}\int d^4x\, S''[0]\phi(x)^2 = \sum_{n,\,m}\frac{1}{2}\int d^4x(b_n\hat{\phi}_n)[-\Box](b_m\hat{\phi}_m) = \sum_n\frac{1}{2}b_n^2\,. \tag{5.14}$$

The reason that the same basis diagonalizes the expansion around the bounce and around the free theory is that the bounce is itself one of the the normal modes (the $n = 0$ one). Thus we can move from the bounce to the false vacuum by varying a single coordinate.

For a cross check, let us compute the propagator in the free theory. We can write the propagator in terms of the path integral as

$$D(x,0) = \frac{\int \mathcal{D}\phi\, e^{-S[\phi]}\phi(x)\phi(0)}{\int \mathcal{D}\phi e^{-S[\phi]}} = \frac{\int \prod_n db_n\, e^{-\frac{1}{2}b_n^2} \sum_{k,j}[b_k\phi_k(x)][b_j\phi_j(0)]}{\int \prod_n dc_n\, e^{-\frac{1}{2}c_n^2}} \tag{5.15}$$

$$= \sum_n \phi_n(x)\phi_n(0) \tag{5.16}$$

$$= \frac{1}{4\pi^2 Rr}\sum_n (2n+3)P_{n+1}^{-1}\left(\frac{R^2-r^2}{R^2+r^2}\right)\,. \tag{5.17}$$

Using the integral representation

$$P_{n+1}^{-1}(x) = \frac{1}{i\pi(n+1)}\int_0^\pi dt\cos(t)\left(\sqrt{x^2-1}\cos t + x\right)^{n+1} \tag{5.18}$$

we can perform the sum, leading to

$$D(x,0) = \frac{1}{4\pi^2 r^2} \tag{5.19}$$

which is the correct answer and as expected is independent of $R$.

## 5.2 Collective coordinates for dilatations

To go to collective coordinates, we consider first the $R$ coordinate associated with dilatations. The $\lambda\phi^4$ theory is classically scale invariant, including if we expand around the bounce or the false vacuum. Ultimately, the divergence associated with an integration over $R$ will be regulated by the scale anomaly, which shows up at higher orders in the saddle point approximation. We are not interested in this regulation here. Rather, we are interested in showing why a divergent integral over the collective coordinate $R$ should show up in the expansion around the bounce but not around the free theory.

Let us denote the dilatation mode as $\phi_d \equiv \phi_{1000}$. This is the mode we denoted by $X(t) = x_0(t)$ in the quantum mechanical case, and is the mode we would like to remove. While $R$ is already incorporated into our basis expansion, let us make the $R$-dependence explicit by fixing $R = \overline{R}$ and writing

$$\phi(x) = \phi_b(\overline{R},x) + c_d\phi_d(\overline{R},x) + \sum_{n\geq 1} c_n\phi_n(\overline{R},x)\,. \tag{5.20}$$

Then the collective coordinate parametrization where $R$ is a coordinate and $\phi_d$ is removed is

$$\phi(x) = \phi_b(R,x) + \sum_{n\geq 1}\alpha_n\phi_n(R,x)\,. \tag{5.21}$$

In these collective coordinates, the Jacobian factor from Eq. (2.10) is

$$\mathcal{J} = \langle\phi_d|\partial_R\phi\rangle = -\sqrt{\frac{3}{5}}\frac{1}{R}\left(2\sqrt{S_b}+\alpha_0\right) + \frac{8}{\sqrt{35}}\frac{1}{R}\alpha_2 \tag{5.22}$$

where $\alpha_2$ is the coefficient of $\phi_{2000}$. Only the radial modes with $\ell = 0$ (and only two of them, with $n = 0, 2$) contribute. Accordingly the expansion of the path integral around the bounce is

$$e^{-S_b} \int \frac{dR}{R} \int \prod_n d\alpha_n \, e^{-\frac{1}{2}\alpha_n^2 \frac{\lambda_n - 1}{\lambda_n}} \delta(\alpha_1) \left| \sqrt{\frac{3}{5}} \left( 2\sqrt{S_b} + \alpha_0 \right) - \frac{8}{\sqrt{35}} \alpha_2 \right| \frac{1}{N_d[\phi]} \tag{5.23}$$

where $N_d[\phi]$ counts the number of solutions to $\langle \phi_d(R, x) | \phi(\overline{R}, x) \rangle = 0$. For the free theory, the path integral in the collective coordinate basis is

$$\int \frac{dR}{R} \int \prod_n d\alpha_n \, e^{-\frac{1}{2}\alpha_n^2} \delta(\alpha_1) \left| \sqrt{\frac{3}{5}} \alpha_0 - \frac{8}{\sqrt{35}} \alpha_2 \right| \frac{1}{N_d[\phi]} \, . \tag{5.24}$$

The integrands of both Eqs. (5.23) and (5.24) have a dependence on $R$ which enters through $N_d[\phi]$.

In the interacting theory, $S_b \sim \frac{1}{-\lambda} \gg 1$ while $\lambda_n \sim 1$ so that the part of the path integral domain with $\alpha_n \lesssim 1$ dominates. Then we can approximate the Jacobian as $\sqrt{\frac{3}{5}} \left( 2\sqrt{S_b} + \alpha_0 \right) - \frac{8}{\sqrt{35}} \alpha_2 \approx \sqrt{\frac{12 S_b}{5}}$. If we can justify ignoring the $N_d[\phi]$ factor, the path integral reduces to the usual expression used to compute the decay rate in $\lambda\phi^4$ QFT.

In the free theory, we also have $\alpha_n \lesssim 1$. There $S_b$ does not appear and we can simply perform the integrals over $\alpha_0$ and $\alpha_2$ and compare to what we would have gotten without collective coordinates, where the $R$ integral is replaced by $\alpha_1$. The ratio of 'collective coordinate integral' and the 'non-collective coordinate integral' if we set $N_d[\phi] = 1$ is

$$\frac{\int \frac{dR}{R} \int d\alpha_0 \int d\alpha_2 e^{-\frac{1}{2}\alpha_0^2 - \frac{1}{2}\alpha_2^2} \left| \sqrt{\frac{3}{5}} \alpha_0 - \frac{8}{\sqrt{35}} \alpha_2 \right|}{\int d\alpha_0 d\alpha_1 d\alpha_2 e^{-\frac{1}{2}\alpha_0^2 - \frac{1}{2}\alpha_1^2 - \frac{1}{2}\alpha_2^2}} = \int \frac{dR}{R} \sqrt{\frac{17}{7\pi^2}} \neq 1 \, . \tag{5.25}$$

Since this ratio is not 1, we cannot expect $N_d[\phi] = 1$ to be a good approximation in the free theory. So if the standard computation for the interacting theory is correct, why can we drop $N_d[\phi]$ when expanding around the bounce but not around the false vacuum?

To proceed, we have to understand the intersection number $N_d[\phi]$. By analogy to the non-relativistic quantum mechanics case, $N_d[\phi]$ counts the number of solutions to $\langle \phi_d(x, R) | \phi \rangle = 0$ for a given $\phi$ over the range of $R$. As in the quantum mechanics case, we need to work in a basis other than the collective coordinate one to evaluate $N_\delta[\phi]$. In the basis with fixed $\overline{R}$ as in Eq. (5.20), we find

$$\left\langle \phi_d(x, R) \middle| \phi(x) \right\rangle = 2\sqrt{S_b} f_0(R, \overline{R}) + \sum_n c_n f_n(R, \overline{R}) \tag{5.26}$$

where

$$f_n(R, \overline{R}) = - \int d^4x \, \phi_d(x, R) \Box \phi_n(x, \overline{R}) \, . \tag{5.27}$$

For example, with $z = R/\overline{R}$ we have

$$f_0(z) = \frac{\sqrt{15} z \left( -1 - 9z^2 + 9z^4 + z^6 - 12 \left( z^4 + z^2 \right) \ln z \right)}{\left( z^2 - 1 \right)^4} \tag{5.28}$$

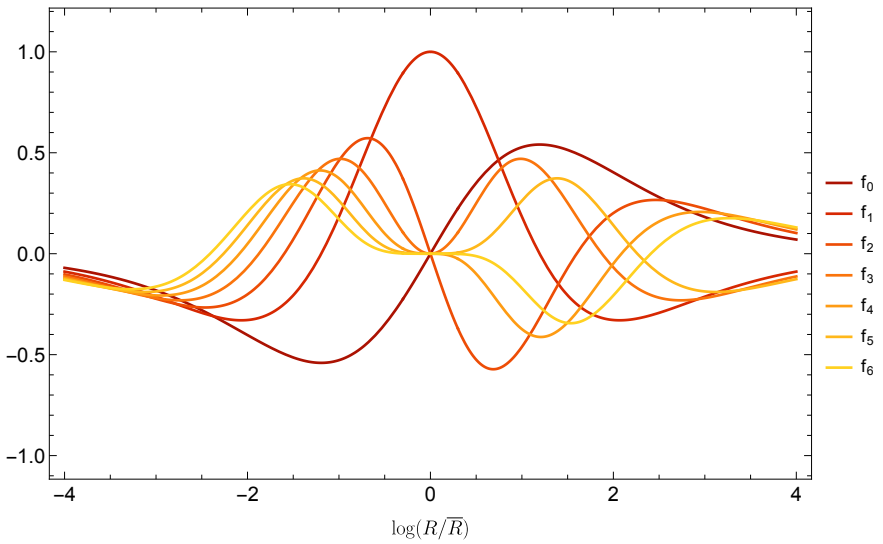



Figure 4: We depict the functions $f_n(R/\overline{R}) = \langle\phi_d(x,R)|\phi_n(x,\overline{R})\rangle$. The collective coordinate $R$ is determined by the condition $2\sqrt{S_b}f_0(R/\overline{R}) + c_n f_n(R/\overline{R}) = 0$ which may have multiple solutions.

$$f_1(z) = -\frac{5z\left(-1-46(z^2-z^6)+z^8-12z^2\left(3+z^2\right)\left(1+3z^2\right)\ln z\right)}{\left(z^2-1\right)^5}\,. \tag{5.29}$$

For the interacting theory, we can examine whether $N_d[\phi] = 1$ is a good approximation using the same approach as for the double-well potential in Section 4.3. Close to $R = \overline{R}$ we have

$$\begin{aligned}\left\langle\phi_d(x,R)\middle|\phi(x)\right\rangle = c_1 &+ \left(\sqrt{\frac{12S_b}{5}} + \sqrt{\frac{3}{5}}c_0 - \frac{8}{\sqrt{35}}c_2\right)\left(\frac{R-\overline{R}}{\overline{R}}\right) \\ &+ \frac{-7\sqrt{15}c_0 - 85c_1 + 8\sqrt{35}c_2 + 40\sqrt{5}c_3 - 14\sqrt{15S_b}}{70}\left(\frac{R-\overline{R}}{\overline{R}}\right)^2 + \cdots\,.\end{aligned} \tag{5.30}$$

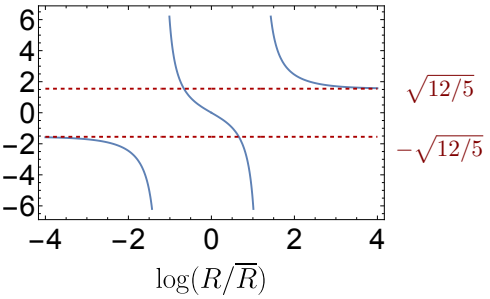
Since $\sqrt{S_b} \gg 1$ we can solve $\langle\phi_d|\phi\rangle = 0$ as a series in $\frac{1}{\sqrt{S_b}}$. The leading terms are

$$R = \overline{R}\left(1 - \sqrt{\frac{5}{12S_b}}c_1 + \frac{14\sqrt{15}c_0c_1 + 35c_1^2 - 16\sqrt{35}c_1c_2}{168S_b} + \cdots\right)\,. \tag{5.31}$$

Since $\lambda_n \approx 1$, the path integral becomes suppressed if any $c_n \gtrsim 1$ on account of the $e^{-c_n^2\lambda_n}$ terms. The only exception is the zero mode: excursions in the $c_1$ direction by arbitrary amounts only affect the integrand by an order 1 amount. From Eq. (5.31) we see that there is always at least one solution with nonzero $c_1$. To have multiple solutions the nonlinear terms need to become important. If we just consider the $c_n$ contributions, we require $c_n \gtrsim \sqrt{S_b} \gg 1$ and so the path integral is affected by an exponentially small amount (of order $e^{-c_1^2\lambda_R} \approx e^{-S_b}$). The situation is akin to our quantum mechanical examples: the quantum fluctuations are too weak to significantly alter the classical trajectory. It remains to check the possibility of additional solutions due to large values of $c_1$.

For large $c_1$, meaning $c_1 \gtrsim \sqrt{S_b}$, the expansion in Eq. (5.31) is no longer useful. If we consider only paths of the form $\phi(x) = \phi_b(x, \overline{R}) + c_1\phi_1(x, \overline{R})$, then a solution to $\langle \phi_d(x, R)|\phi(x)\rangle = 0$ requires

$$\frac{c_1}{\sqrt{S_b}} = -2\,\frac{f_0(R,\overline{R})}{f_1(R,\overline{R})} = \tag{5.32}$$




For a given $c_1$ there is always at least one solution to the above equation. More precisely, since $f_1(R, \overline{R}) = 0$ at $R/\overline{R} = x_c$ and $R/\overline{R} = 1/x_c$ where $x_c \approx 0.303$, there will be one solution with $x_c\overline{R} < R < \frac{1}{x_c}\overline{R}$. However, if $|c_1/\sqrt{S_b}| > \sqrt{12/5}$, there will be a second solution. So here $1 \leq N_d[\phi] \leq 2$, but we would like to determine whether the second solution is exponentially suppressed or not.

Although the action is exactly independent of the collective coordinate $R$, it is only independent of the coordinate $c_1$ to quadratic order. Expanding to higher orders there is a quartic term

$$S[\phi_b + c_1\phi_1] = S_b - \frac{5}{36S_b}c_1^4\,. \tag{5.33}$$

The sign is negative since $\lambda < 0$; when expanding around the bounce to quadratic order, there is only one direction in which the potential decreases, but to quartic order it decreases in every direction (except the flat ones)! So we do not want to include terms to quartic order in $c_1$. We do, however, want to include terms to quartic order in the collective coordinate $R$. If we did not include these terms, the action would not be independent of $R$. Large $R$ leads to a large factor of the volume of the dilatation group, but not an exponentially large factor from fluctuations in a direction of instability.

We are primarily interested in whether a field configuration has large quadratic action in our collective coordinates and not in the coordinates at fixed $\overline{R}$. The $\alpha_n$ coordinates are given by

$$\alpha_n = \big\langle \phi_b(x,\overline{R}) + c_1\phi_1(x,\overline{R}) - \phi_b(x,R)\big|\phi_n(x,R)\big\rangle \tag{5.34}$$

$$= g_n(R,\overline{R}) + c_1 f_n(R,\overline{R}) - 2\sqrt{S_b}\,\delta_{n0} \tag{5.35}$$

where $g_n(R,\overline{R}) = \langle\phi_b(x,\overline{R})|\phi_n(x,R)\rangle$ and we note that $g_1(R,\overline{R}) = f_0(R,\overline{R})$. When $c_1 \ll \sqrt{S_b}$, we have $c_1 \approx \sqrt{\frac{12S_b}{5}}(R-\overline{R})/\overline{R}$. In this regime $R$ is very close to $\overline{R}$ and all the $\alpha_n$ are small, namely $\alpha_n \ll \sqrt{S_b}$. When $c_1 \gtrsim \sqrt{S_b}$ then all $\alpha_n \gtrsim \sqrt{S_b}$ as well. The corresponding field configurations, which include all of the $N_d[\phi] = 2$ field configurations with only $c_1$ turned on, are all suppressed. Other field configurations with $N_d[\phi] > 1$ would have various $c_n \gtrsim \sqrt{S_b}$ leading to $\alpha_n \gtrsim \sqrt{S_b}$ as well. Thus we can approximate $N_d[\phi]$ as equal to 1 up to exponentially small corrections in the path integral.

The free field theory behaves very differently, akin to the free quantum mechanics example. In the free field theory, the $\sqrt{\frac{12S_b}{5}}(\frac{R-\overline{R}}{\overline{R}})$ term in Eq. (5.30) is not present. Then the leading solution in the regime $\frac{R-\overline{R}}{\overline{R}} \ll 1$ is

$$R = \overline{R}\left(1 - \frac{7\sqrt{5}c_1}{7\sqrt{3}c_0 - 8\sqrt{7}c_2} + \cdots\right). \tag{5.36}$$

There is no small parameter in the free theory and no zero mode. So even with $c_n \lesssim 1$ there are multiple solutions to $\langle\phi_d|\phi\rangle = 0$. That is, field configurations with $N_d[\phi] \gtrsim 1$ are not suppressed and in fact $N_d[\phi]$ can be arbitrarily large. Including these large values of $N_d[\phi]$ must compensate for the unregulated $\int \frac{dR}{R}$ integral, akin to the free quantum mechanics case. Showing the cancellation explicitly seems extraordinarily challenging: collective coordinates around the bounce are very cumbersome to use for the free theory.

An alternative approach to understand the exponential suppression when $N_d[\phi] > 1$ in the interacting theory, but not in the free theory, is to change coordinates to map directly to the double well. This approach is considered in Appendix B.

### 5.3 Collective coordinates for translations

In addition to the dilatation zero mode proportional to $\partial_R\phi_b$, the $\lambda\phi^4$ theory also has zero modes associated with translations which are linear combinations of the $\partial_\mu\phi_b$. When introducing collective coordinates for translations, one must also justify why the approximation $N_\delta = 1$ is valid in the interacting theory but not for the free theory. The argument is essentially the same as for the double-well potential and for the dilatation mode: in the interacting theory paths with $N_\delta > 1$ give an exponentially suppressed contribution to the action. Because $\partial_\mu\phi_b$ breaks Lorentz invariance, or rather $SO(4)$ invariance in Euclidean signature, the modes which contribute to the projections $\langle\partial_\mu\phi_b(x)|\phi(x^\mu + x_0^\mu)\rangle$ are not so easy to identify.

To prove that the analysis for the translation modes proceeds as with the dilatation mode, the easiest approach is to make a coordinate transformation to map any particular direction in spacetime to $r$. This can be done because the theory is scale invariant and the classical theory has an exact $SO(5)$ symmetry as part of a larger conformal symmetry. To make the map explicit, we employ the conformal mapping with coordinates

$$\eta^\mu = \frac{2x^\mu R}{R^2 + r^2}, \quad \eta^5 = \frac{R^2 - r^2}{R^2 + r^2}. \tag{5.37}$$

These new coordinates satisfy $\sum_{\mu=1}^4 \eta_\mu^2 + \eta_5^2 = 1$, i.e. they parameterize the 4-sphere. The inverse of this conformal mapping is the stereographic projection to Cartesian coordinates.

On the 4-sphere, the integration volume is

$$d^4x = \left(\frac{2R}{R^2 + r^2}\right)^{-4} d\Omega_4. \tag{5.38}$$

If we also rescale the fields as

$$\Phi(\eta) = \frac{r^2 + R^2}{2R}\phi(x) = \frac{R}{1+\eta_5}\phi(x)\,, \tag{5.39}$$

the action becomes

$$\int d^4x \left[-\frac{1}{2}\phi\Box\phi + \frac{\lambda}{4}\phi^4\right] = \int d\Omega_4 \left[\frac{1}{2}\Phi(-\vec{L}^2+2)\Phi + \frac{\lambda}{4}\Phi^4\right], \tag{5.40}$$

where $\vec{L}^2$ is the angular momentum operator. The equations of motion are now easily solved by a mode with $\vec{L}^2\Phi = 0$, namely $\Phi_b = \sqrt{\frac{2}{-\lambda}}$. Using Eq. (5.39) we find $\Phi_b = \frac{r^2+R^2}{2R}\phi_b$, as expected.

The normal modes from expanding around $\Phi_b$ are spherical harmonics in 5 dimensions, which are identical to the conformally mapped fluctuations in Eq. (5.4). Indeed, it is through this conformal mapping that the exact eigenfunctions were first derived [7,40,41]. In these coordinates there is an exact $SO(5)$ symmetry among the 5 zero modes. We can therefore rotate any of the translation modes to the dilatation mode. Mapping back to flat space, the analysis of the path integral and $N_\delta$ proceeds identically to the discussion in Section 5.2.

# 6 Conclusions

In this paper we have provided a resolution to a long-standing puzzle involving collective coordinates. Collective coordinates are a way to factor out an integral over the volume of a symmetry group from the path integral and still exploit a saddle point approximation in the non-symmetry directions. The puzzle is how collective coordinates can be consistently introduced through a change of coordinates, independent of the action. Doing so appears to produce correct results in interacting theories where the symmetry volume is expected, but not in free theories, where it is unexpected. The resolution to this puzzle is that the change of coordinates is allowed and valid, but it is multi-valued: a given path $x(t)$ in quantum mechanics may be described by multiple values of the collective coordinates $t_0$. We quantified this multi-valuedness with the intersection number $N_X[x]$: it counts the number of times two paths $X(t)$ and $x(t)$ intersect as one is shifted by the collective coordinate over the symmetry volume. The correct way to introduce a collective coordinate is therefore with a resolution of the identity of the form

$$1 = \int dt_0 \frac{1}{N_X[x]}\,\delta\Big[\langle X(t+t_0)|x(t)\rangle\Big]\,\Big|\langle X(t+t_0)|\dot{x}(t)\rangle\Big|\,. \tag{6.1}$$

Without the factor of $N_X[x]$, the path integral can give the wrong answer.

To call out and rectify the mistreatment of collective coordinates we have explored examples in quantum mechanics and quantum field theory. For the splitting between the two lowest energy levels in the symmetric double well potential in quantum mechanics, the $N_X[x]$ factor counts the number of times a path crosses the potential barrier. Because each crossing gives additional suppression by $e^{-S_I}$ where $S_I$ is the instanton action, we argued that one can approximate $N_X[x] = 1$ up

to exponentially small corrections. We made special effort to regulate the theory in finite time and carefully take the limit of a combination of the twisted and untwisted partition function to extract the energy splitting. While the final result is well-known through the WKB approximation, most path integral treatments involve unjustified approximations. We have attempted to provide a complete path integral treatment of the splitting where each step is under explicit control.

In the interacting theories we have explored here, namely the double well in quantum mechanics and interacting scalar quantum field theory, we found that the intersection number could be approximated as 1. In the free theories, the intersection number cannot be approximated as 1 and in fact it can be arbitrarily large, as needed to compensate for the volume factor which arises in collective coordinates.

It would be interesting to explore cases intermediate between these two extremes. An example is the instanton-anti-instanton pair in QCD as a function of separation. This configuration interpolates between the vacuum at zero separation and a saddle point of the action at infinite separation. The separation distance is a quasi-collective coordinate [42–44], becoming a flat direction at infinite separation and a unexceptional normal mode fluctuation at zero separation. In the dilute instanton gas approximation, the dependence of the action on the separation distance is neglected, so that the separation is treated as a collective coordinate. Thus the techniques developed here could provide insights into the validity of the dilute instanton gas approximation. Although QCD is a significant step up in complexity from the examples considered here, it is of direct phenomenological relevance and is certainly worth examining in more detail.

### Acknowledgements

We would like to thank G. Dunne, M. Mariño, M. Serone, and J. Stout for valuable discussions. M.D.S. and A.B. are supported in part by the U.S. Department of Energy under contract DE-SC0013607. J.C. is supported by a Junior Fellowship from the Harvard Society of Fellows.

## A Metric approach to Jacobians in the path integral

In this Appendix we explain an approach for computing Jacobians arising from changes of coordinates in path integrals. To this end, we review how to use Riemannian geometry to manipulate the path integral measure. For concreteness we consider a quantum mechanical path integral for a particle in $d$ spatial dimensions, although our arguments here readily generalize to field theory.

The path integral measure and its normalization depend on a metric on the space of paths. If $K$ is a positive-definite operator, we can begin by defining an inner product

$$\big\langle x(t)\big|y(t)\big\rangle_K = \int dt\, x(t)\cdot K\cdot y(t)\,. \tag{A.1}$$

It is most common to then choose a basis of functions $x_n(t)$ that are orthonormal with respect to this inner product. If we write $x(t)=\sum_n c_n x_n(t)$ where $\langle x_m|x_n\rangle=\delta_{mn}$ then the path integral

measure is

$$\mathcal{D}x = \prod_n dc_n \,. \tag{A.2}$$

Factors of $2\pi$ or an overall normalization $\mathcal{N}$ can be optionally absorbed into the measure; these will drop out of physical quantities calculated by ratios of path integrals.

Now we consider a set of coordinates which are not orthonormal. If we have any set of coordinates $\{a_n\}$ which can parametrize a general path $x(t)$ then we can define a distance element between paths as

$$ds^2 = \sum_{m,n} G_{mn}(a) da_m da_n = \sum_{m,n} \left\langle \frac{\partial x}{\partial a_m} \middle| \frac{\partial x}{\partial a_n} \right\rangle da_m \, da_n \,. \tag{A.3}$$

In this case the path integral measure becomes

$$\mathcal{D}x = \sqrt{\det G} \prod_n da_n \,, \tag{A.4}$$

which is the analog of the measure $\sqrt{\det g}\, d^d x$ in Riemannian geometry.

When we compute correlation functions, we divide by a partition function normalization computed using the same measure in Eq. (A.4). We observe that correlation functions are independent of the choice of positive-definite $K$ in Eq. (A.1), since changing $K$ to $K'$ affects the square root determinant multiplicatively by $\sqrt{\frac{\det K'}{\det K}}$ which cancels out the same factor in the partition function normalization.

Let us show how these more abstract ideas manifest in two examples.

**Example 1:** First, suppose we consider paths $x : [0, T] \to \mathbb{R}$ with $x(0) = x_i$ and $x(T) = x_f$. The space of such paths can be parameterized by

$$x(t) = x_i + \frac{t}{T}(x_f - x_i) + \sum_{n=1}^{\infty} c_n \sqrt{\frac{2}{T}} \sin\Big(\frac{\pi n}{T}\, t\Big) \tag{A.5}$$

with coordinates $\{c_n\}$. Taking $K = 1$, corresponding to the ordinary 'flat' $L^2$ inner product, the metric is

$$G_{mn} = \Big\langle \frac{\partial x}{\partial c_m} \Big| \frac{\partial x}{\partial c_n} \Big\rangle = \frac{2}{T} \int_0^T dt \, \sin\Big(\frac{\pi m}{T}\, t\Big) \, \sin\Big(\frac{\pi n}{T}\, t\Big) = \delta_{mn} \tag{A.6}$$

and so

$$ds^2 = \sum_n dc_n^2 \,. \tag{A.7}$$

For this metric $\sqrt{\det G} = 1$ and the path integral measure is simply

$$\mathcal{D}x = \prod_n dc_n \,. \tag{A.8}$$

**Example 2:** Now consider coordinates for paths around some instanton $x_I(t)$, namely $x(t) = x_I(t) + \sum_{n=0}^{\infty} a_n\, x_n(t)$. If $\langle x_m(t)|x_n(t)\rangle = \delta_{mn}$, then, by the same arguments as in the first example above, the path integral measure can be expressed in coordinates $\{a_n\}$ as $\prod_{n=0}^{\infty} da_n$. If $x_0(t)$ is a time translation zero mode of the action of our theory, then it is instead convenient to work with an expansion $x(t) = x_I(t+t_0) + \sum_{n=1}^{\infty} \alpha_n\, x_n(t+t_0)$ in coordinates $\{t_0, \alpha_{n>0}\}$. In these coordinates, the metric is

$$ds^2 = \langle \dot{x}(t)|\dot{x}(t)\rangle\, dt_0^2 + 2\sum_{n=1}^{\infty} \langle \dot{x}(t)|x_n(t+t_0)\rangle\, dt_0\, d\alpha_n + \sum_{n=1}^{\infty} d\alpha_n^2 \tag{A.9}$$

and so

$$\sqrt{\det G} = \left( \langle \dot{x}(t)|\dot{x}(t)\rangle - \sum_{n=1}^{\infty} \langle \dot{x}(t)|x_n(t+t_0)\rangle\langle x_n(t+t_0)|\dot{x}(t)\rangle \right)^{1/2} . \tag{A.10}$$

Since $\sum_{n=0}^{\infty} \langle a(t)|x_n(t+t_0)\rangle\langle x_n(t+t_0)|b(t)\rangle = \langle a(t)|b(t)\rangle$ by completeness of $\{x_{n\geq 0}(t+t_0)\}$ for fixed $t_0$, the above simplifies to

$$\sqrt{\det G} = |\langle x_0(t+t_0)|\dot{x}(t)\rangle| = \left| \langle x_0(t)|\dot{x}_I(t) + \sum_{n=1}^{\infty} \alpha_n \dot{x}_n(t)\rangle \right| . \tag{A.11}$$

Thus in the $\{t_0, \alpha_{n>0}\}$ coordinates the path integral measure is

$$\mathcal{D}x = dt_0 \prod_{m=1}^{\infty} d\alpha_m \left| \langle x_0(t)|\dot{x}_I(t)\rangle + \sum_{n=1}^{\infty} \alpha_n \langle x_0(t)|\dot{x}_n(t)\rangle \right| \tag{A.12}$$

which agrees with Eq. (2.9). Of course, we also have to be careful about the change in the domain of path integration coming from our change in coordinates. The modification of the domain and subtleties with multi-valued coordinate changes are more readily understood using the methods in the main text.

# B Mapping QFT to the double well

In the quantum mechanics case, we saw that $N_\delta$ counted the number of times a particular path crossed $x = 0$. This led to a simple physical argument for why paths with $N_\delta > 1$ were exponentially suppressed in the interacting theory where the potential is high at $x = 0$, but unsuppressed in the free theory where $x = 0$ is not special in any sense. In the scalar quantum field theory we justified why $N_\delta = 1$ is appropriate for the interacting theory, but did not provide a physical interpretation of $N_\delta$. We can do so by mapping the theory to quantum mechanics. This mapping provides an alternative way of thinking about the bounces in QFT.

To map QFT to QM we follow [44, 45] and use a conformal mapping to turn the dilatation direction into time. Let us take $r = Re^{\omega\tau}$ with $\omega$ an arbitrary frequency, and consider spherically-symmetric field configurations of the form

$$\phi(r) = e^{-\omega\tau} h(\tau) . \tag{B.1}$$

Then the action of $\lambda\phi^4$ theory evaluated on such field configurations is

$$S[\phi] = \frac{2\pi^2 R^2}{\omega} \int d\tau \left[ \frac{1}{2}(\partial_\tau h)^2 + \frac{1}{2}\omega^2 h^2 + R^2\omega^2 \frac{\lambda}{4} h^4 \right] . \tag{B.2}$$

The equations of motion are then $h''(\tau) - V'(h) = 0$ where $V(h) = \frac{1}{2}\omega^2 h^2 + R^2\omega^2 \frac{\lambda}{4} h^4$. Recalling that $\lambda < 0$, the above is an inverted double-well potential. The bounce is written as

$$h_b(\tau) = \sqrt{\frac{2}{-\lambda}} \frac{1}{R\cosh(\omega\tau)} . \tag{B.3}$$

The above satisfies the equation of motion of a particle in an inverted inverted double well, starting at the metastable local maximum at $h = 0$. The particle rolls slowly off of this hump starting at $\tau = -\infty$, bounces off a wall, and ends up back on top of the hump at $\tau = \infty$. The flat direction is associated with the time at the center of the bounce. The normal modes which had satisfied $\Box\phi_n = \lambda_n V''(\phi_b)\phi_n$ now satisfy

$$[\partial_\tau^2 - \omega^2] h_n = -\lambda_n \frac{6\omega^2}{\cosh^2(\omega\tau)} h_n . \tag{B.4}$$

However, there is a better basis we can use. Expanding around $h_b(\tau)$ the quadratic fluctuations satisfy

$$\left[ \partial_\tau^2 - \omega^2 + \frac{6\omega^2}{\cosh^2(\omega\tau)} \right] \eta_n = \hat{\lambda}_n \eta_n , \tag{B.5}$$

which reproduces the Pöschl-Teller potential with $\lambda = 2$. There are thus two bound states and a continuum. The $h_n$ and $\eta_h$ bases are different, except for the dilatation zero mode

$$h_d(\tau) = \eta_d(\tau) = \frac{\sqrt{5}}{4\pi R} \frac{\sinh(\omega\tau)}{\cosh^2(\omega\tau)} \tag{B.6}$$

which has the same form in both bases (it has $\hat{\lambda}_1 = 0$ and $\lambda_1 = 1$ so that $\frac{\lambda_1 - 1}{\lambda_1} = 0$ as in Eq. (5.13)).

The zero condition $\langle \phi_d(x, R) | \phi(x) \rangle = 0$ becomes $\langle \eta_1(\tau - \tau_0) | h(\tau) \rangle = 0$ after the mapping to the double well, and can be studied using $\eta$ modes quite similarly to the way we studied the double well in Section 4. We write

$$h(\tau) = b_0 \eta_0(\tau) + b_1 \eta_1(\tau) + \sum_{n \geq 2} b_n(\tau) \eta_n(\tau) \tag{B.7}$$

where $\eta_n$ are the continuum modes and $\eta_0, \eta_1$ are the bound states. As with the double well, the bound states have support for $\tau \sim \omega^{-1}$ while the continuum modes must be regulated with $-T < \tau < T$, or equivalently $Re^{-\omega T} < r < Re^{\omega T}$. The projection that needs to be studied now involves the collective coordinate $\tau_0$ and is

$$\langle \eta_1(\tau - \tau_0) | h(\tau) \rangle = 0 . \tag{B.8}$$

In the regime $\tau_0 \ll \omega^{-1}$ we reduce to the case where we can trade $\tau_0$ for $b_1$. In the regime $\omega^{-1} \ll \tau_0 \ll T$ the two bound states have support only near $\tau = 0$ and in particular

$$\eta_1(\tau - \tau_0) \approx \delta'(\tau - \tau_0), \quad \omega^{-1} \ll \tau_0 \ll T\,, \tag{B.9}$$

so that the zero overlap condition becomes $h'(\tau_0) = 0$. Thus $\tau_0$ is the time when $h(\tau)$ stops and $N_\delta[h]$ counts the number of times $h$ stops.

Plugging (B.1) into the action of the free theory, the potential is $V_0(h) = \frac{1}{2}\omega^2 h^2$ which is just a stable quadratic potential. Thus the $h$ actions for the free and interacting theories are similar, and the action of a given path in the interacting theory is always strictly lower than the action in the free theory. Nevertheless, since the saddle point we expand around has large action in the interacting theory, perturbations around this saddle must be large to slow the particle down so that it stops multiple times. Such large quantum fluctuations are exponentially suppressed. In the free theory, the saddle point is just $h = 0$ and the particle can stop multiple times without much penalty. Hence in the full quantum field theory setting, $N_d[\phi] = 1$ is a good approximation for the interacting theory but not for the free theory.

## References


[1] P.H. Frampton, *Vacuum Instability and Higgs Scalar Mass*, *Phys. Rev. Lett.* **37** (1976) 1378.

[2] S.R. Coleman, *The Fate of the False Vacuum. 1. Semiclassical Theory*, *Phys. Rev. D* **15** (1977) 2929.

[3] C.G. Callan, Jr. and S.R. Coleman, *The Fate of the False Vacuum. 2. First Quantum Corrections*, *Phys. Rev. D* **16** (1977) 1762.

[4] G. Isidori, G. Ridolfi and A. Strumia, *On the metastability of the standard model vacuum*, *Nucl. Phys. B* **609** (2001) 387 [hep-ph/0104016].

[5] A. Andreassen, D. Farhi, W. Frost and M.D. Schwartz, *Direct Approach to Quantum Tunneling*, *Phys. Rev. Lett.* **117** (2016) 231601 [1602.01102].

[6] A. Andreassen, D. Farhi, W. Frost and M.D. Schwartz, *Precision decay rate calculations in quantum field theory*, *Phys. Rev. D* **95** (2017) 085011 [1604.06090].

[7] A. Andreassen, W. Frost and M.D. Schwartz, *Scale Invariant Instantons and the Complete Lifetime of the Standard Model*, *Phys. Rev. D* **97** (2018) 056006 [1707.08124].

[8] J.-L. Gervais and B. Sakita, *Extended Particles in Quantum Field Theories*, *Phys. Rev. D* **11** (1975) 2943.

[9] C.W. Bernard, *Gauge Zero Modes, Instanton Determinants, and QCD Calculations*, *Phys. Rev. D* **19** (1979) 3013.

[10] G. 't Hooft, *Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle*, *Phys. Rev. D* **14** (1976) 3432.

[11] A. Voros, *The return of the quartic oscillator. The complex WKB method*, *Annales de l'I.H.P. Physique théorique* **39** (1983) 211.

[12] E. Delabaere and F. Pham, *Unfolding the quartic oscillator*, *Annals of Physics* **261** (1997) 180.

[13] G. Álvarez, *Langer–Cherry derivation of the multi-instanton expansion for the symmetric double well*, *Journal of Mathematical Physics* **45** (2004) 3095.

[14] G.V. Dunne and M. Unsal, *Uniform WKB, Multi-instantons, and Resurgent Trans-Series*, *Phys. Rev. D* **89** (2014) 105009 [1401.5202].

[15] G. Rastelli, *Semiclassical formula for quantum tunneling in asymmetric double-well potentials*, *Phys. Rev. A* **86** (2012) 012106.

[16] A. Garg, *Tunnel splittings for one-dimensional potential wells revisited*, *American Journal of Physics* **68** (2000) 430.

[17] V.N. Gribov, *Quantization of Nonabelian Gauge Theories*, *Nucl. Phys. B* **139** (1978) 1.

[18] S. Coleman, *Aspects of Symmetry: Selected Erice Lectures*, Cambridge University Press, Cambridge, U.K. (1985), 10.1017/CBO9780511565045.

[19] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*, WORLD SCIENTIFIC, 5th ed. (2009), 10.1142/7305.

[20] H.J.W. Müller-Kirsten, *Introduction to Quantum Mechanics: Schrödinger Equation and Path Integral*, World Scientific (2012), 10.1142/8428.

[21] M. Mariño, *Instantons and Large N: An Introduction to Non-Perturbative Methods in Quantum Field Theory*, Cambridge University Press (9, 2015), 10.1017/CBO9781107705968.

[22] M. Mariño, *Advanced Topics in Quantum Mechanics*, Cambridge University Press (12, 2021), 10.1017/9781108863384.

[23] E. Brezin, G. Parisi and J. Zinn-Justin, *Perturbation Theory at Large Orders for Potential with Degenerate Minima*, *Phys. Rev. D* **16** (1977) 408.

[24] E. Gildener and A. Patrascioiu, *Instanton Contributions to the Energy Spectrum of a One-Dimensional System*, *Phys. Rev. D* **16** (1977) 423.

[25] E.B. Bogomolny, *Calculation Of Instanton - Anti-Instanton Contributions In Quantum Mechanics*, *Phys. Lett. B* **91** (1980) 431.

[26] J. Zinn-Justin, *Multi - Instanton Contributions in Quantum Mechanics*, *Nucl. Phys. B* **192** (1981) 125.

[27] J. Zinn-Justin, *Multi - Instanton Contributions in Quantum Mechanics. 2.*, *Nucl. Phys. B* **218** (1983) 333.

[28] U.D. Jentschura and J. Zinn-Justin, *Higher order corrections to instantons*, *J. Phys. A* **34** (2001) L253 [math-ph/0103010].

[29] J. Zinn-Justin and U.D. Jentschura, *Multi-instantons and exact results I: Conjectures, WKB expansions, and instanton interactions*, *Annals Phys.* **313** (2004) 197 [quant-ph/0501136].

[30] H.J.W. Müller-Kirsten, *Introduction to Quantum Mechanics: Schrödinger Equation and Path Integral*, World Scientific (2012), 10.1142/8428.

[31] K. Iwaki and T. Nakanishi, *Exact WKB analysis and cluster algebras*, *J. Phys. A* **47** (2014) 474009 [1401.7094].

[32] E. Delabaere, H. Dillinger and F. Pham, *Exact semiclassical expansions for one-dimensional quantum oscillators*, *J. Math. Phys.* **38** (1997) 6126.

[33] B. Bucciotti, T. Reis and M. Serone, *An anharmonic alliance: exact WKB meets EPT*, *JHEP* **11** (2023) 124 [2309.02505].

[34] A. van Spaendonck and M. Vonk, *Exact instanton transseries for quantum mechanics*, 2309.05700.

[35] G. Poschl and E. Teller, *Bemerkungen zur Quantenmechanik des anharmonischen Oszillators*, *Z. Phys.* **83** (1933) 143.

[36] I.M. Gel'fand and A.M. Yaglom, *Integration in Functional Spaces and its Applications in Quantum Physics*, *Journal of Mathematical Physics* **1** (1960) 48.

[37] G.V. Dunne, *Functional determinants in quantum field theory*, *J. Phys. A* **41** (2008) 304006 [0711.1178].

[38] S. Fubini, *A New Approach to Conformal Invariant Field Theories*, *Nuovo Cim. A* **34** (1976) 521.

[39] L.N. Lipatov, *Divergence of the Perturbation Theory Series and the Quasiclassical Theory*, *Sov. Phys. JETP* **45** (1977) 216.

[40] I. Drummond and G. Shore, *Dimensional regularisation and instantons: A scalar field theory model*, *Annals of Physics* **121** (1979) 204.

[41] A.J. McKane and D.J. Wallace, *Instanton calculations using dimensional regularisation*, *Journal of Physics A: Mathematical and General* **11** (1978) 2285.

[42] I.I. Balitsky and A.V. Yung, *Instanton Molecular Vacuum in $\mathcal{N} = 1$ Supersymmetric Quantum Mechanics*, *Nucl. Phys. B* **274** (1986) 475.

[43] I.I. Balitsky and A.V. Yung, *Collective-Coordinate Method for Quasizero Modes*, *Phys. Lett. B* **168** (1986) 113.

[44] A.V. Yung, *Instanton Vacuum in Supersymmetric QCD*, *Nucl. Phys. B* **297** (1988) 47.

[45] V.V. Khoze and A. Ringwald, *Valley trajectories in gauge theories*, Tech. Rep. CERN, Geneva (1991).

Prepared for submission to JHEP

# Analytic solutions for the linearized first-order magnetohydrodynamics and implications for causality and stability

**Zhe Fang**[a] **Koichi Hattori**[a,b] **Jin Hu**[c]

[a] *Zhejiang Institute of Modern Physics, Department of Physics, Zhejiang University, Hangzhou, Zhejiang 310027, China*

[b] *Research Center for Nuclear Physics, Osaka University, 10-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan*

[c] *Department of Physics, Fuzhou University, Fujian 350116, China*

*E-mail:* koichi.hattori@zju.edu.cn, hu-j23@fzu.edu.cn

Abstract: We solve the first-order relativistic magnetohydrodynamics (MHD) within the linear-mode analysis performed near an equilibrium configuration in the fluid rest frame. We find two complete sets of analytic solutions for the four and two coupled modes with seven dissipative transport coefficients. The former set has been missing in the literature for a long time. Our method provides a simple and general algorithm for the solution search on an order-by-order basis in the derivative expansion, and can be applied to general sets of hydrodynamic equations. We also find that the small-momentum expansions of the solutions break down when the momentum direction is nearly perpendicular to an equilibrium magnetic field due to the presence of another small quantity, that is, a trigonometric function representing the anisotropy. We elaborate on the angle dependence of the solutions and provide alternative series representations that work near the right angle. Finally, we discuss the issues of causality and stability based on our analytic solutions and recent developments in the literature.

arXiv:2402.18601v1 [physics.plasm-ph] 28 Feb 2024

## Contents


**1 Introduction** **1**

**2 First-order MHD from the magnetic-flux conservataion** **3**

**3 Linear-mode analysis** **6**

**4 General solutions for the linearized equations** **8**
4.1 Alfven modes and issue of the small $k$ expansion in anisotropic systems 9
4.2 Magneto-sonic modes from analytic algorithm 11

**5 Causality and stability** **17**
5.1 The Alfven and magneto-sonic velocities in the ideal MHD 17
5.2 Dissipative nature of the first-order corrections in the fluid rest frame 17
5.3 Covariant stability in anisotropic systems 18

**6 Conclusion and outlook** **20**

**A inequalities** **21**
A.1 Inequalities for the Alfven and magneto-sonic velocities 21
A.2 Dissipative corrections in the fluid rest frame 23


## 1 Introduction

Relativistic magnetohydrodynamics (MHD) has been developing as a framework to describe various systems ranging from the femto-scale droplets realized in relativistic heavy-ion collisions at RHIC and LHC [1–5] to the cosmological/astronomical scales. The latter includes the accretion flows and jet formation near black holes [6–9], supernova explosions [10–12], and the binary mergers of neutron stars (NSs) [13–16] and of NS-BH [17, 18] (see also references therein). More recent developments include observations of the magnetic inverse cascade with coupled dynamics of the magnetic helicity (and the fermion chirality) [19–29] (see Refs. [30–32] for reviews).

In recent years, relativistic MHD was reformulated based on the conservation of the magnetic flux [33–36] (see Ref. [37] for a review). On the other hand, the conventional formulation is based on a coupled system of the Maxwell equation and the energy-momentum conservation law, where both equations have the source terms stemming from the electric current and the Joule heat and Lorentz force, respectively. This implies the presence of non-conserved charges, i.e., gapped modes, involved in the conventional formulation. In fact, electric fields are damped out or screened when systems approach equilibrium states. The new formulation canonically follows the spirit of hydrodynamics, that is, conservation

laws associated with symmetries. The magnetic flux conservation is identified as a consequence of generalized concept of global symmetries called the magnetic one-form symmetry [38].

In this paper, we focus on solving the set of MHD equations. We linearize the MHD equations with respect to perturbative disturbances applied to an equilibrium state and obtain the dispersion relations for the eigenmodes. This analysis, called the linear-mode analysis, has been conventionally used to discuss causality and stability of hydrodynamic theories [39–41] (see below for recent progress on causality and stability analyses). While the linear-mode analysis was applied to the relativistic MHD in the recent literature [33, 42–44], a complete set of solutions is still missing. The difficulty simply arises from the fact that one needs to diagonalize a large matrix in the absence of a spatial rotational symmetry broken by a magnetic field. In this paper, we develop an analytic algorithm for the solutions search on an order-by-order basis in the derivative expansion and obtain the complete set of analytic solutions. It is useful to obtain analytic solutions since the transport coefficients and the equations of state are often not (precisely) known in individual systems. Moreover, our method works as a general algorithm for the solution search, and can be applied to general sets of equations based on derivative expansions.

The complete set of solutions consists of six gappless modes, which are known as a pair of the Alfven waves and two pairs of the magneto-sonic waves. We fully include the first-order derivative corrections that consist of three bulk viscosities, two shear viscosities, and two electric resistivities. We focus on the Landau frame while a straightforward extension can be carried out for a general frame choice and general matching conditions of hydrodynamic variables (see, e.g., Ref. [37] for discussions about these choices in MHD and Ref. [44] for a recent linear-mode analysis without a specific choice). In the presence of the anisotropic corrections, the solutions for the magneto-sonic modes had been only known in the two particular limits where the momentum is oriented parallel or perpendicular to a background magnetic field [33, 42, 44].

We further discuss the convergence of the small-momentum expansion by inspecting the obtained solutions and find that the higher-order terms in the small-momentum expansion diverge when the momentum is taken nearly perpendicular to a magnetic field. This issue is caused by the presence of another small quantity, that is, a trigonometric function representing the angle dependence. We find this issue both in the Alfven and magneto-sonic modes. It can be a general issue in anisotropic systems. We identify the correct result obtained from the original equations before the small-momentum expansion is performed, giving a different result than that in the literature [33, 44]. We provide an alternative series representation that correctly captures the anisotropy near the right angle.

Finally, we discuss the causality and stability of relativistic MHD. We show that the phase velocities, that are called the Alfven velocities and the fast and slow magneto-sonic velocities, are always smaller than the speed of light. This means that the linear waves are causal *within* the ideal MHD. We also show that the first-order corrections, obtained within the fluid rest frame, always provide damping effects on the linear waves as long as the transport coefficients satisfy the inequalities required by the second law of thermodynamics.

This is expected since MHD only contains dissipative transport coefficients.[1] The absence of growing modes indicates stability of an equilibrium state *within the fluid rest frame* where the linear-mode analysis is performed. However, those conditions are not sufficient to guarantee the causality and stability of dissipative hydrodynamics in general Lorentz frames. It has been widely know that diffusive modes are acausal, leading to developments of the Israel-Steward theory [45–47] (see, e.g., Ref. [48] for a review), and observers in different Lorentz frames could see (unphysical) unstable modes (see, e.g., Ref. [49]).

Recent developments deepened our understanding of the issues of causality and stability. A stronger necessary condition for a stable dispersion relation was obtained from complex analysis of a retarded propagator in general causal theories [50]. Then, it was explicitly shown that, unless this necessary condition is satisfied, one finds a boost velocity that transforms a stable mode in the fluid rest frame to an unstable mode in the boosted frame [51, 52]. We briefly discuss a generalization of this crucial condition to anisotropic systems including MHD. The reader is referred to recent related works [53–56] [57, 58] [59] that may be classified in terms of employed criteria for causality and stability; with or without specific choices of a flow vector and hydrodynamic variables; and with or without the linearization. One is demanded to perform more analyses when including more conserved charges such as a vector charge [60], magnetic field [43, 44, 61], axial charge [62, 63], and/or spin [64–67].

This paper is organized as follows. We first recapitulate the recent formulation of relativistic MHD in Sec. 2 and show a set of linearized equations in Sec. 3. In Sec. 4, we introduce our method for the solution search. We elaborate on the convergence/breakdown of the small-momentum expansion. In Sec. 5, we discuss causality and stability of MHD, which is supported by Appendices. Finally, we conclude this paper in Sec. 6. Throughout the paper, we use the mostly plus metric convention $\eta^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ and the completely antisymmetric tensor with the convention $\epsilon^{0123} = +1$. Then, the fluid velocity $u^\mu$ is normalized as $u^\mu u_\mu = -1$. We define the projection operator $\Delta^{\mu\nu} = \eta^{\mu\nu} + u^\mu u^\nu$ such that $u_\mu \Delta^{\mu\nu} = 0$. To specify the direction of a magnetic field, we introduce a spatial unit vector $b^\mu$ such that $b^\mu b_\mu = 1$ and $u_\mu b^\mu = 0$ and accordingly another projection operator $\Xi^{\mu\nu} = \Delta^{\mu\nu} - b^\mu b^\nu$ such that $b_\mu \Xi^{\mu\nu} = 0 = u_\mu \Xi^{\mu\nu}$.

## 2 First-order MHD from the magnetic-flux conservataion

We recapitulate the formulation of relativistic MHD with the magnetic-flux conservation [33–36] (see Ref. [37] for a review). While the magnetic flux is conserved in the absence of a magnetic monopole, the electric flux can terminate at electric charges, implying that an electric field is not a conserved quantity. The conservation laws for the energy-momentum tensor $\Theta^{\mu\nu}$ and the (dual) electromagnetic field strength tensor $\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$ read

$$\partial_\mu \Theta^{\mu\nu} = 0, \quad \partial_\mu \tilde{F}^{\mu\nu} = 0. \tag{2.1}$$

[1] In the strict hydrodynamic limit, MHD does not contain the Hall terms due to the absence of net electric charge density.

Realizing the second equation as a conservation law for the magnetic flux led to a renewed formulation of magnetohydrodynamics along with the symmetry guideline, but with a generalized notion called the one-form symmetry.

The temporal components of the conserved currents provide the corresponding conserved charges

$$e = u_\mu u_\nu \Theta^{\mu\nu}, \quad B^\mu = -\tilde{F}^{\mu\nu} u_\nu. \tag{2.2}$$

We postulate that these quantities satisfy the first law of thermodynamics

$$Ts = e + p - H_\mu B^\mu, \tag{2.3a}$$
$$Tds = de - H_\mu dB^\mu, \tag{2.3b}$$
$$TDs = De - H_\mu DB^\mu, \tag{2.3c}$$

where we defined the temporal derivative $D = u^\mu \partial_\mu$. The translational symmetry of the system guarantees the conservation of the *total* energy density $e$ that contains not only matter contributions but also electromagnetic energy. The corresponding pressure $p$ should also be the total quantity.

To organize a closed system of equations, one needs to obtain the constitutive equations that express the spatial components of the conserved currents by the conserved charges. Based on the derivative expansion, the constitutive equations can be written down as

$$\Theta^{\mu\nu} = e u^\mu u^\nu + p_\parallel b^\mu b^\nu + p_\perp \Xi^{\mu\nu} + \Theta^{\mu\nu}_{(1)}, \tag{2.4a}$$
$$\tilde{F}^{\mu\nu} = B^\mu u^\nu - B^\nu u^\mu + \tilde{F}^{\mu\nu}_{(1)}, \tag{2.4b}$$

where we introduced a unit vector $b^\mu = B^\mu/\sqrt{B^\nu B_\nu}$ and the projection operator $\Xi^{\mu\nu} = \Delta^{\mu\nu} - b^\mu b^\nu$ such that $b_\mu \Xi^{\mu\nu} = 0 = u_\mu \Xi^{\mu\nu}$; Note that $u_\mu B^\mu = 0$ by definition (2.2). The explicitly written terms exhaust the zeroth-order terms that can be constructed with the available tensors in the absence of the vector and axial charges. The subscripts denote the first-order corrections that will be constrained by the entropy-current analysis below.

We are in position to compute the divergence of the entropy current $s^\mu = s u^\mu + s^\mu_{(1)}$, where $s^\mu_{(1)}$ is the first-order corrections to the entropy current. It can be expressed with the derivatives of the conserved quantities by the use of the first law of thermodynamics (2.3). Then, using the equations of motion (2.1) together with the constitutive equations (2.4), one finds that

$$\begin{aligned} \partial_\mu s^\mu &= s\partial_\mu u^\mu + Ds + \partial_\mu s^\mu_{(1)} \\ &= \beta(Ts - \epsilon - p_\perp + H_\mu B^\mu)\partial_\mu u^\mu - \beta\big[(p_\parallel - p_\perp) b^\mu b^\nu + B b^\mu H^\nu\big]\partial_\mu u_\nu \\ &\quad - \Theta^{\mu\nu}_{(1)} \partial_\mu \beta_\nu + \tilde{F}^{\mu\nu}_{(1)} \partial_\mu(\beta H_\nu) + \partial_\mu(s^\mu_{(1)} + \beta u_\nu \Theta^{\mu\nu}_{(1)} + \beta\omega_{\alpha\beta}\Sigma^{\mu\alpha\beta}_{(1)} - \beta H_\nu \tilde{F}^{\mu\nu}_{(1)}). \end{aligned} \tag{2.5}$$

The leading-order terms in derivative describe the *ideal MHD*. For the entropy production to vanish at the ideal order, one should have

$$p_\perp = p, \tag{2.6a}$$
$$(p_\parallel - p_\perp) b^\nu + H^\nu B = 0. \tag{2.6b}$$

The zeroth-order result can be summarized as

$$\Theta^{\mu\nu}_{(0)} = \epsilon u^\mu u^\nu - p\Xi^{\mu\nu} + (p - B_\mu H^\mu) b^\mu b^\nu, \tag{2.7}$$

where $H^\mu = \mu_m^{-1} B^\mu$ with $\mu_m$ being the magnetic permeability. The above result indicates the pressure anisotropy induced by the Maxwell stress.

The second law of thermodynamics requires that the first-order corrections be semi-positive definite. This condition is satisfied if individual contributions of thermodynamic forces take semi-positive values in Eq. (2.5), i.e.,

$$-\Theta^{\mu\nu}_{(1)} \partial_\mu \beta_\nu \geq 0, \tag{2.8a}$$

$$\tilde{F}^{\mu\nu}_{(1)} \partial_\mu (\beta H_\nu) \geq 0. \tag{2.8b}$$

These inequalities can be insured for general hydrodynamic configurations if the left-hand sides are organized in bilinear forms, constraining the possible forms of the constitutive equations as

$$\Theta^{\mu\nu}_{(1)} = -T\eta^{\mu\nu\rho\sigma} \partial_{(\rho}\beta_{\sigma)}, \tag{2.9a}$$

$$\tilde{F}^{\mu\nu}_{(1)} = -T\rho^{\mu\nu\rho\sigma} \partial_{[\rho}(\beta H_{\sigma]}). \tag{2.9b}$$

The fourth-rank tensors $\eta^{\mu\nu\rho\sigma}$ and $\rho^{\mu\nu\rho\sigma}$ can be constructed with the available tensors, $u^\mu$, $b^\mu$, $\Xi^{\mu\nu}$, and $\epsilon^{\mu\nu\alpha\beta}$, as [33, 36, 37]

$$\begin{aligned}\eta^{\mu\nu\rho\sigma} &= \begin{pmatrix} b^\mu b^\nu & \Xi^{\mu\nu} \end{pmatrix} \begin{pmatrix} \zeta_\parallel & \zeta_\times \\ \zeta'_\times & \zeta_\perp \end{pmatrix} \begin{pmatrix} b^\rho b^\sigma \\ \Xi^{\rho\sigma} \end{pmatrix} \\ &\quad + 2\eta_\parallel (b^\mu \Xi^{\nu(\rho} b^{\sigma)} + b^\nu \Xi^{\mu(\rho} b^{\sigma)}) + \eta_\perp (\Xi^{\mu\rho}\Xi^{\nu\sigma} + \Xi^{\mu\sigma}\Xi^{\nu\rho} + \Xi^{\mu\nu}\Xi^{\rho\sigma}),\end{aligned} \tag{2.10a}$$

$$\rho^{\mu\nu\rho\sigma} = -2\rho_\perp (b^\mu \Xi^{\nu[\rho} b^{\sigma]} - b^\nu \Xi^{\mu[\rho} b^{\sigma]}) + 2\rho_\parallel \Xi^{\mu[\rho}\Xi^{\sigma]\nu}. \tag{2.10b}$$

Note that we have chosen the Landau frame and the matching condition for the magnetic flux such that $e = u_\mu u_\nu \Theta^{\mu\nu}_{(0)}$ and $B^\mu = -\tilde{F}^{\mu\nu}_{(0)} u_\nu$ (see a review article [37] for more detailed discussions). Therefore, the tensors $\eta^{\mu\nu\rho\sigma}$ and $\rho^{\mu\nu\rho\sigma}$ are transverse to the flow vector $u^\mu$. It will be an interesting extension to discuss stability and causality in a general frame and matching conditions (see discussions in Sec. 5 and recent works [44, 53–55, 57, 58]).

Note also that we have $\zeta_\times = \zeta'_\times$ by virtue of Onsager's reciprocal relation and that there are no Hall terms in the charge-neutral systems. The two coefficients $\rho_{\perp,\parallel}$ are identified with the resistivities in the perpendicular and parallel directions with respect to the magnetic field [33, 36, 37]. The second law of thermodynamics, i.e., the inequalities (2.8), basically requires all the transport coefficients introduced above be semi-positive. An exception is the off-diagonal component $\zeta_\times$ that does not have to be semi-positive definite since the second law can be insured as long as the eigenvalues of the matrix are semi-positive definite. In summary, one finds the inequalities

$$\begin{aligned}&\zeta_\perp \geq 0, \quad \zeta_\parallel \geq 0, \quad \zeta_\parallel \zeta_\perp \geq \zeta_\times^2, \\ &\eta_\parallel \geq 0, \quad \eta_\perp \geq 0, \quad \rho_\perp \geq 0, \quad \rho_\parallel \geq 0.\end{aligned} \tag{2.11}$$

## 3 Linear-mode analysis

In this section, we solve the first-order hydrodynamic equations for the small perturbations near an equilibrium state, which is often called the linear-mode analysis. We apply perturbations on top of equilibrium values $u^\mu = (1,0,0,0)$ and $B^\mu = (0,0,0,B)$, where we took the direction of the magnetic field along the $z$ axis at the equilibrium without loss of generality. Namely, the conserved charges are displaced from their equilibrium values as

$$e \to e + \delta e(x), \quad u^\mu \to u^\mu + \delta u^\mu(x), \quad B^\mu \to B^\mu + \delta B^\mu(x). \tag{3.1}$$

We will linearize the hydrodynamic equations with respect to these perturbations. The perturbation $\delta B^\mu$ can have a perpendicular component to $B^\mu$. We assume a linear relation $H^\mu = B^\mu/\mu_m$ with $\mu_m$ being a constant in spacetime. For simplicity, we also assume that the contributions of the matter and magnetic components to the equilibrium energy density and pressure can be separated as

$$p = P + \frac{B^2}{2\mu_m}, \quad e = \epsilon + \frac{B^2}{2\mu_m}. \tag{3.2}$$

The conservation law of the energy-momentum tensor (2.1) can be projected as

$$u_\nu \partial_\mu \Theta^{\mu\nu} = 0, \quad \Xi^\rho_\nu \partial_\mu \Theta^{\mu\nu} = 0, \quad b_\nu \partial_\mu \Theta^{\mu\nu} = 0. \tag{3.3}$$

Plugging Eq. (2.4) into the above and focusing on the linear-order in the perturbations, one arrives at the linearized equations

$$0 = \partial_0 \delta\epsilon + \frac{B}{\mu_m}\partial_0 \delta B + h\partial_{\perp\mu}\delta u^\mu_\perp + (h - \frac{B^2}{\mu_m})\partial_z \delta u_z, \tag{3.4a}$$

$$\begin{aligned} 0 = {} & h\partial_0 \delta u_{x,y} + c_s^2 \partial_{x,y}\delta\epsilon + \frac{B}{\mu_m}(\partial_{x,y}\delta B_z - \partial_z \delta B_{x,y}) \\ & -[(\zeta_\perp + \eta_\perp)\partial^2_{x,y} + \eta_\perp \partial^2_{y,x} + \eta_\parallel \partial_z^2]\delta u_{x,y} - \zeta_\perp \partial_x\partial_y \delta u_{y,x} - (\zeta_\times + \eta_\parallel)\partial_z\partial_{x,y}\delta u_z, \end{aligned} \tag{3.4b}$$

$$\begin{aligned} 0 = {} & c_s^2 \partial_z \delta\epsilon + (h - \frac{B^2}{\mu_m})\partial_0 \delta u_z \\ & -(\zeta_\times + \eta_\parallel)\partial_z(\partial_x \delta u_x + \partial_y \delta u_y) - (\zeta_\parallel \partial_z^2 + \eta_\parallel(\partial_x^2 + \partial_y^2))\delta u_z, \end{aligned} \tag{3.4c}$$

where the subscripts $x, y, z$ denote the spatial components, but without minus signs from the metric, i.e., $(B_x, B_y, B_z) = (B^1, B^2, B^3)$, $(\partial_x, \partial_y, \partial_z) = (\partial^1, \partial^2, \partial^3)$, $\delta u^\mu_\perp = (0, \delta u_x, \delta u_y, 0)$. The second equation for the transverse components have the rotational symmetry around the magnetic-field direction. We also introduced the enthalpy $h = e + p = \epsilon + P + B^2/\mu_m$ with the equilibrium values and the (squared) sound velocity $c_s^2 = \delta P/\delta\epsilon$.

The equations for $\tilde{F}^{\mu\nu}$ can be projected and linearized in the same manner. The projected conservation law reads

$$u_\nu \partial_\mu \tilde{F}^{\mu\nu} = 0, \quad \Xi^\rho_\nu \partial_\mu \tilde{F}^{\mu\nu} = 0, \quad b_\nu \partial_\mu \tilde{F}^{\mu\nu} = 0. \tag{3.5}$$

The explicit forms of the linearized equations are obtained as

$$0 = \partial_i \delta B^i, \tag{3.6a}$$

$$0 = B\partial_z \delta u_{x,y} - \partial_0 \delta B_{x,y} \tag{3.6b}$$
$$-\rho'_\perp T\big[\,\partial_z \partial_{x,y}\delta(\beta B_z) - \partial_z^2 \delta(\beta B_{x,y})\,\big] + \rho'_\parallel T\big[\,\partial_\perp^2 \delta(\beta B_{x,y}) - \partial_{x,y}\partial_{\perp\mu}\delta(\beta B_\perp^\mu)\,\big],$$

$$0 = -B\partial_{\perp\mu}\delta u_\perp^\mu - \partial_0 \delta B_z + \rho'_\perp T\big[\,\partial_\perp^2 \delta(\beta B_z) - \partial_z \partial_{\perp\mu}\delta(\beta B_\perp^\mu)\,\big], \tag{3.6c}$$

where we used an identity $0 = (u_\mu + \delta u_\mu)(B^\mu + \delta B^\mu) = B\delta u_z - \delta B^0 + \mathcal{O}(\delta^2)$ and defined

$$\rho'_\parallel = \frac{\rho_\parallel}{\mu_m}, \quad \rho'_\perp = \frac{\rho_\perp}{\mu_m}. \tag{3.7}$$

It is useful to notice that the set of equations (3.6) contains only two independent dynamical equations. The first equation (3.6a) does not contain a time derivative and is nothing but the Gauss law constraint. Another redundancy can be identified with an identity

$$0 = \partial_\mu \partial_\nu \tilde{F}^{\mu\nu} = (\Xi_{\alpha\beta} - u_\alpha u_\beta + b_\alpha b_\beta)\partial^\alpha \partial_\mu \tilde{F}^{\mu\beta}. \tag{3.8}$$

This identify is satisfied by any antisymmetric tensor regardless of the actual components of $\tilde{F}^{\mu\nu}$, and serves as a sum-rule constraint on the set of equations (3.6). Then, we are left with two independent dynamical equations and, correspondingly, the two spatial components of $\delta B^\mu$.

The derivative of $\delta\beta$ in Eq. (3.6) can be expressed with that of $\delta\epsilon$ with the help of a relation obtained from the thermodynamic relation (2.3), that is,

$$\delta\beta = -\frac{c_s^2 \beta}{h - B^2/\mu_m}\delta\epsilon. \tag{3.9}$$

To summarize the above equations in the Fourier representation, we introduce a perturbation in a single mode

$$\delta\boldsymbol{u}(t, x, z) = \delta\tilde{\boldsymbol{u}}(\omega, k_\perp, k_\parallel) e^{-i\omega t + ik_\perp x + ik_\parallel z}, \tag{3.10}$$

and the same for $\delta e$ and $\delta\boldsymbol{B}$. Here, without loss of generality, we have set the transverse coordinate system in such a way that the dependence on the $y$ coordinate vanishes. Then, the equations (3.4) and (3.6) can be cast into two separate matrix equations

$$\Big(A_{(0)} + iA_{(1)}\Big)\begin{pmatrix}\delta B_y \\ \delta u_y\end{pmatrix} = 0, \tag{3.11a}$$

$$\Big(M_{(0)} + iM_{(1)}\Big)\begin{pmatrix}\delta\epsilon \\ \delta u_x \\ \delta u_z \\ \delta B_x \\ \delta B_z\end{pmatrix} = 0. \tag{3.11b}$$

The explicit forms of the first set of matrices are given as

$$A_{(0)} = \begin{pmatrix} \omega & Bk_\parallel \\ h\frac{v_A^2}{B}k_\parallel & h\omega \end{pmatrix}, \tag{3.12a}$$

$$A_{(1)} = \begin{pmatrix} \rho'_\perp k_\parallel^2 + \rho'_\parallel k_\perp^2 & 0 \\ 0 & \eta_\parallel k_\parallel^2 + \eta_\perp k_\perp^2 \end{pmatrix}, \tag{3.12b}$$

with the so-called Alfven-wave velocity

$$v_A = \frac{B}{\sqrt{\mu_m h}}. \tag{3.13}$$

The explicit forms of the second set of matrices are given as

$$M_{(0)} = \begin{pmatrix} 0 & -k_\perp B & 0 & 0 & \omega \\ 0 & k_\parallel B & 0 & \omega & 0 \\ \omega & -hk_\perp & h(v_A^2-1)k_\parallel & 0 & h\frac{v_A^2}{B}\omega \\ -c_s^2 k_\perp & h\omega & 0 & h\frac{v_A^2}{B}k_\parallel & -h\frac{v_A^2}{B}k_\perp \\ -c_s^2 k_\parallel & 0 & -h(v_A^2-1)\omega & 0 & 0 \end{pmatrix}, \tag{3.14a}$$

$$M_{(1)} = \begin{pmatrix} -\frac{Bc_s^2}{h(1-v_A^2)}\rho'_\perp k_\perp^2 & 0 & 0 & -\rho'_\perp k_\parallel k_\perp & \rho'_\perp k_\perp^2 \\ \frac{Bc_s^2}{h(1-v_A^2)}\rho'_\perp k_\perp k_\parallel & 0 & 0 & \rho'_\perp k_\parallel^2 & -\rho'_\perp k_\perp k_\parallel \\ 0 & 0 & 0 & 0 & 0 \\ 0 & (\zeta_\perp+\eta_\perp)k_\perp^2 + \eta_\parallel k_\parallel^2 & (\zeta_\times+\eta_\parallel)k_\perp k_\parallel & 0 & 0 \\ 0 & (\zeta_\times+\eta_\parallel)k_\perp k_\parallel & \zeta_\parallel k_\parallel^2 + \eta_\parallel k_\perp^2 & 0 & 0 \end{pmatrix}. \tag{3.14b}$$

We will solve these equations in the next section. For later use, we introduce an angle $\theta$ measured from the direction of the magnetic field, and the momenta can be expressed as

$$k_\parallel = k\cos\theta, \quad k_\perp = k\sin\theta. \tag{3.15}$$

We also normalize all the viscous coefficients by the enthalpy, i.e.,

$$\eta'_{\parallel,\perp} = \frac{1}{h}\eta_{\parallel,\perp}, \quad \zeta'_{\parallel,\perp,\times} = \frac{1}{h}\zeta_{\parallel,\perp,\times}. \tag{3.16}$$

## 4 General solutions for the linearized equations

In this section, we solve the matrix equations (3.11) to obtain the dispersion relations of the linear waves. The equations from the first-order hydrodynamics are accurate up to the order $k^2$, so that our goal is to obtain the dispersion relation up to this order. We will obtain a complete set of analytic solutions with all the transport coefficients being free parameters. This is useful since the magnitudes of the transport coefficients are often not (precisely) known in individual systems. However, we also find that the small $k$ expansion poses an issue of convergence in anisotropic systems. We investigate the solutions near the angle $\theta \sim \pi/2$ in detail and provide an alternative series representation that works well in this regime.

We first discuss the analytic solutions for Eq. (3.11a), which have been discussed in the literature [33, 42, 44]. We elaborate on this simpler equation to point out the convergence issue in the small $k$ expansion for anisotropic systems. We demonstrate the issue by comparing the limit of the angle $\theta \to \pi/2$ taken before and after the small $k$ expansion that does not agree with each other, and then identify the correct result, giving a different result than that in the literature [33, 44]. We then provide a series representation that correctly captures this limit as well as the corrections in $k^2$ near this angle.

Then, we proceed to tackle the larger matrix in Eq. (3.11b), of which the solutions have not been known in the literature. We will find the analytic solutions for the four modes fully including the dissipative effects. We introduce our simple algorithm for the solution search. We find that these modes also contain the convergence issue, and that the result at $\theta = \pi/2$ should be different than those in Refs. [33, 44]. An alternative series representation is provided accordingly.

### 4.1 Alfven modes and issue of the small $k$ expansion in anisotropic systems

The secular equation for Eq. (3.11a) is found to be

$$v_A^2 k_\parallel^2 - (\omega + i\tilde{\rho}k^2)(\omega + i\tilde{\eta}k^2) = 0, \tag{4.1}$$

where $\tilde{\rho} = \rho'_\perp \cos^2\theta + \rho'_\parallel \sin^2\theta$ and $\tilde{\eta} = \eta'_\parallel \cos^2\theta + \eta'_\perp \sin^2\theta$. The solutions are readily obtained as

$$\begin{aligned}\omega &= \pm\sqrt{v_A^2 k_\parallel^2 - \frac{1}{4}(\tilde{\rho}-\tilde{\eta})^2 k^4} - \frac{i}{2}(\tilde{\rho}+\tilde{\eta})k^2 \\ &= \pm v_A k_\parallel - \frac{i}{2}(\tilde{\rho}+\tilde{\eta})k^2 + \mathcal{O}(k^3).\end{aligned} \tag{4.2}$$

We performed the small $k$ expansion in the second line. These solutions are gapless in the limit $k \to 0$ and are known as the Alfven waves propagating along the equilibrium magnetic field. Since $\eta_{\parallel,\perp} \geq 0$, $\rho_{\parallel,\perp} \geq 0$, these modes are damped out in time by an exponential factor $e^{-\frac{1}{2}(\tilde{\rho}+\tilde{\eta})t}$. Without a parity-breaking effect, we have a pair of waves propagating in opposite directions with the same damping rate.

Below, we elaborate on an issue of the small $k$ expansion involved in the Alfven modes. It is important to clarify this issue here because one will find the same issue in the other matrix equation (3.11b) of which the analytic solutions have not been known. Anisotropic systems may potentially share the same issue. We investigate the limit of angle $\theta \to \pi/2$, i.e., the vanishing $k_\parallel$ limit in Eq. (4.2). Taking the limit *without* performing the small $k$ expansion, one finds that

$$\omega(\theta = \frac{\pi}{2}) = \pm\frac{i}{2}|\tilde{\rho}-\tilde{\eta}|k_\perp^2 - \frac{i}{2}(\tilde{\rho}+\tilde{\eta})k_\perp^2 = -i\eta'_\perp k_\perp^2, \quad -i\rho'_\parallel k_\perp^2, \tag{4.3}$$

irrespective of the sign of $\tilde{\rho} - \tilde{\eta}$. In this limit, these modes split into two distinct purely diffusive modes. These two modes are still invariant under the parity transformation, i.e., $k_\perp \to -k_\perp$, because the linear term vanishes in this limit. One can trace back the splitting

of the dispersion relations (4.3) to the original matrices (3.12). Taking the limit $\theta \to \pi/2$, one finds that the matrix equation (3.11a) reduces to a diagonal form

$$\begin{pmatrix} \omega + i\rho'_\parallel k_\perp^2 & 0 \\ 0 & \omega + i\eta'_\perp k_\perp^2 \end{pmatrix} \begin{pmatrix} \delta B_y \\ \delta u_y \end{pmatrix} = 0, \tag{4.4}$$

where the two perturbations $\delta B_y$ and $\delta u_y$ are decoupled from each other. One of the dispersion relations (4.3) is for the flow perturbation damped by the shear viscosity, while the other is for the magnetic flux diffusion by the resistivity.

Now, it should be noticed that the dispersion relations (4.3) are not reproduced by the limit taken *after* performing the small $k$ expansion in Eq. (4.2); The expanded result instead yields two degenerate purely diffusive modes [33, 44]. This disagreement occurs due to an invalid expansion of the terms containing $k_\parallel = k\cos\theta$ that is not a small quantity but is exactly zero when $\theta \to \pi/2$. Performing the small $k$ expansion first and then taking the limit $\theta \to \pi/2$, one finds that

$$\omega(\theta \to \frac{\pi}{2}) = -\frac{i}{2}(\rho'_\parallel + \eta'_\perp)k_\perp^2 \pm \lim_{\theta\to\pi/2} \sum_{n=1}^{\infty} \frac{c_n(\rho'_\parallel - \eta'_\perp)^{2n}}{(v_A\cos\theta)^{2n-1}} k_\perp^{2n+1}, \tag{4.5}$$

where $c_n$ is the numerical coefficients. In the above expansion, one encounters divergence of the higher-order terms as $\theta \to \pi/2$, which spoils the small $k$ expansion near $\theta = \pi/2$. Clearly, the small $k$ expansion and the limit of $\theta \to \pi/2$ do not commute with each other.

There is a transient angle $\theta$ (for a given $k$) where the propagating modes turn into the purely diffusive modes. We investigate this transition below. The disagreement about the limits originates from the ill-organized small $k$ expansion when there is another small quantity $\cos\theta$. In this case, we should specify which of $k$ or $\cos\theta$ is smaller before carrying out an expansion. Then, one can organize two pairs of series representations:

$$\omega = \pm v_A k_\parallel - \frac{i}{2}(\tilde{\rho} + \tilde{\eta})k^2 + \mathcal{O}\Big(\frac{k}{|\cos\theta|}k^2\Big), \quad \hat{k} < |\cos\theta|, \tag{4.6a}$$

$$\omega = -ik^2\rho'_\parallel - i\Big((\rho'_\perp - \rho'_\parallel)k^2 + \frac{v_A^2}{\eta'_\perp - \rho'_\parallel}\Big)\cos^2\theta + \mathcal{O}(\cos^4\theta), \tag{4.6b}$$

$$-ik^2\eta'_\perp - i\Big((\eta'_\parallel - \eta'_\perp)k^2 - \frac{v_A^2}{\eta'_\perp - \rho'_\parallel}\Big)\cos^2\theta + \mathcal{O}(\cos^4\theta), \quad |\cos\theta| < \hat{k}. \tag{4.6c}$$

While the first expression is the same as the expansion in the second line of Eq. (4.2), it should be emphasized that the correction terms are small only when $\hat{k}/|\cos\theta| < 1$. When $|\cos\theta| < \hat{k}$, we find two expansions with distinct pure imaginary coefficients shown in the second and third lines. They are smoothly connected to the two purely diffusive modes (4.3) at $\theta = \pi/2$. The inverse factor of $(\eta'_\perp - \rho'_\parallel)$ does not cause a divergence in general unless fine-tuned. In principle, one needs to make a dimensionless expansion parameter $\hat{k} \equiv k/k_c$ with an ultraviolet (UV) cutoff $k_c$ in order to compare the two small expansion parameters. While $k_c$ is implicit in the transport coefficients, it is useful to explicitly introduce $k_c$ so that one can maintain general values of the transport coefficients. The

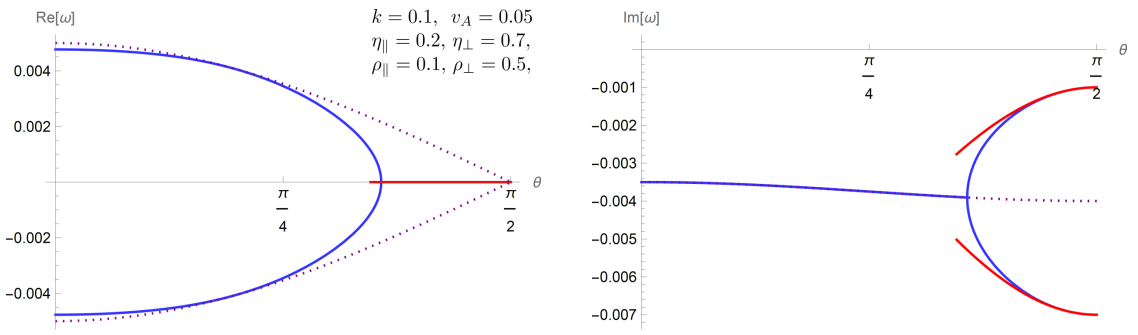



**Figure 1**. The dispersion relation of the Alfven modes. The original solution (blue-solid curves) in Eq. (4.2) is compared with the small-$k$ and small-cosine expansions in Eq. (4.6). The parameters are fixed as indicated in the legend for illustration.

cutoff in general depends on details of the microscopic dynamics, or more precisely, how we integrate out the UV degrees of freedom.

In Fig. 1, we plot the above series representations together with the original solution shown in Eq. (4.2). We confirm the agreement between the small-cosine expansion (red-solid curves) and the original solution (blue-solid curves) for the angle near $\theta = \pi/2$. As noted above, the small $k$ expansion (dotted curves) breaks down as the angle approaches $\theta = \pi/2$. There is a critical angle where the real part, and thus the velocity, vanishes. The critical angle is simply determined by the condition that the square root vanishes in Eq. (4.2). Above the critical angle, the Alfven modes turn into purely diffusive modes, and the degenerate imaginary parts split into two distinct values.

### 4.2 Magneto-sonic modes from analytic algorithm

In the previous subsection, we investigated the Alfven modes encoded in the matrix equation (3.11a). The analytic solutions for the other matrix equation (3.11b) have not been know to the best of our knowledge, except for the solutions at the particular momentum directions, i.e., $k_\parallel = 0$ or $k_\perp = 0$ [33, 42, 44]. It is challenging to obtain analytic solutions including the higher-order corrections in $k$. If one invokes brute-force efforts, one has to find general solutions for a quartic equation in $\omega$, which is possible but is not an efficient path to reach compact forms of solutions. Moreover, the solutions at $k_\parallel = 0$ suffer from the issue of the small $k$ expansion discussed in the previous subsection, giving rise to different solutions in this limit than those obtained in Refs. [33, 44]. As in the Alfven modes, we investigate the behavior near this limit carefully.

First, we provide a simple method for the solution search that works on an order-by-order basis in $k$. This method serves as a general algorithm that can be applied to general sets of hydrodynamic equations and any other equations based on derivative expansions (see Ref. [68] for further applications), while we here focus on the quartic secular equation in $\omega$ at $k^2$ order. We note again that the secular equations for hydrodynamic equations, and thus their solutions, are only accurate up to a given order in $k$. Thus, the order-by-order algorithm is a suitable method for the solution search.

We begin with the leading-order solutions by putting $M_{(1)} = 0$. The secular equation from $M_{(0)}$ is found to be

$$\omega\Big[\omega^4 - \mathcal{V}^2 k^2 \omega^2 + v_A^2 c_s^2 k^2 k_\parallel^2\Big] = 0, \tag{4.7}$$

with $\mathcal{V}^2 := (c_s^2 + v_A^2) - c_s^2 v_A^2 \sin^2\theta \geq 0$. A trivial solution, $\omega = 0$, originates from the redundancy mentioned below Eq. (3.8). We exclude this trivial solution in the following discussions. The leading-order dispersion relations are found to be

$$\omega = \pm v_1 k, \quad \pm v_2 k, \tag{4.8}$$

where the two distinct velocities are given as

$$v_{1,2} = \frac{1}{\sqrt{2}}\sqrt{\mathcal{V}^2 \pm \sqrt{\mathcal{V}^4 - 4 v_A^2 c_s^2 \cos^2\theta}}\ , \tag{4.9}$$

where the upper and lower signs are for $v_1$ and $v_2$, respectively. These modes are two pairs of counter-propagating waves called the fast and slow magneto-sonic waves. When $\theta = 0$, one of the pairs reduces to the Alfven waves (4.2) as required by the rotational symmetry and the other pair reduces to the sound modes without modification of the sound velocity because of the absence of a magnetic pressure according to the Gauss law $\delta B_z = -k_\perp \delta B_x / k_\parallel = 0$. In this limit, one finds that $\{v_1, v_2\} = \{c_s, v_A\}$ when $c_s \geq v_A$ and $\{v_1, v_2\} = \{v_A, c_s\}$ when $c_s < v_A$.

It is important to note that a pair of counter-propagating waves acquires the same dissipative corrections at $k^2$ order in the absence of parity-breaking effects. Therefore, the general solutions should be found in the forms

$$\omega = \pm v_1 k - i w_1 k^2, \quad \omega = \pm v_2 k - i w_2 k^2, \tag{4.10}$$

where $w_{1,2}$ are independent of $k$ and are determined below. Accordingly, one can make an ansatz for the factorized form of the quartic secular equation as

$$\begin{aligned} f_{\rm ans}(\omega) \ =\ & \{\omega - (v_1 k - i w_1 k^2)\}\{\omega - (-v_1 k - i w_1 k^2)\} \\ & \times \{\omega - (v_2 k - i w_2 k^2)\}\{\omega - (-v_2 k - i w_2 k^2)\} \\ & + \mathcal{O}(\omega^3 k^3) + \mathcal{O}(\omega^2 k^4) + \mathcal{O}(\omega^1 k^5) + \mathcal{O}(\omega^0 k^6). \end{aligned} \tag{4.11}$$

The uncertainties at $k^3$ order in each solution result in the uncertainties indicated in the last line. These uncertainties are not of our interest here, since they are not improved unless the constitutive equations are improved beyond the first-order derivative expansions. The following computation is greatly simplified by identifying these irrelevant higher-order terms and getting rid of them at this stage. Expanding $f_{\rm ans}(\omega)$, we should only retain the relevant terms as

$$f_{\rm ans}(\omega) = \omega^4 + g_2(k)\omega^3 + g_3(k)\omega^2 + g_4(k)\omega + g_5(k), \tag{4.12}$$

where $g_n$ denotes the $n$-th polynomial of $k$ stemming from the expansion of Eq. (4.11). As mentioned above, we should not, or do not have to, retain the terms higher than $n$ that could only be relevant beyond the first-order hydrodynamics.

The ansatz (4.12) is matched to the secular equation from Eq. (3.11b). Consistently to the above order counting, we only need to retain the terms at the same orders in $k$ in the secular equation. Then, the matching for the coefficients in the $\omega^3$ and $\omega^1$ terms lead to coupled linear equations

$$v_2^2 w_1 + v_1^2 w_2 = \frac{1}{2} W_1, \tag{4.13a}$$

$$w_1 + w_2 = \frac{1}{2} W_2, \tag{4.13b}$$

The explicit forms of $W_{1,2}$ are given as

$$\begin{aligned} W_1 =\ & \eta'_\parallel \Big( c_s^2 \cos^2(2\theta) + \frac{v_A^2 \sin^2\theta}{1-v_A^2} \Big) + \rho'_\perp c_s^2 \frac{1 - v_A^2 \cos^2\theta}{1 - v_A^2} \\ & + \zeta'_\parallel \frac{v_A^2}{1-v_A^2} \cos^2\theta + \big\{ \zeta'_\parallel - 2\zeta'_\times + (\zeta'_\perp + \eta'_\perp) \big\} c_s^2 \sin^2\theta \cos^2\theta, \end{aligned} \tag{4.14a}$$

$$\begin{aligned} W_2 =\ & \eta'_\parallel \Big( 1 + \frac{v_A^2}{1-v_A^2} \sin^2\theta \Big) + \rho'_\perp \Big( 1 + \frac{v_A^2 c_s^2}{1-v_A^2} \sin^2\theta \Big) \\ & + \frac{\zeta'_\parallel}{1-v_A^2} \cos^2\theta + \big( \zeta'_\perp + \eta'_\perp \big) \sin^2\theta, \end{aligned} \tag{4.14b}$$

where all the viscous coefficients are normalized by the enthalpy as in Eq. (3.16). Note that the coefficients in the $\omega^2$ and $\omega^0$ terms are automatically matched when one inserts $v_{1,2}$ in the leading-order solutions (4.9) because the unknowns $w_{1,2}$ are not involved. It is now a quite simple task to solve the above linear equations to find the solutions

$$w_1 = -\frac{W_1 - v_1^2 W_2}{2(v_1^2 - v_2^2)}, \quad w_2 = +\frac{W_1 - v_2^2 W_2}{2(v_1^2 - v_2^2)}. \tag{4.15}$$

As expected in the ansatz (4.11), $w_1$ and $w_2$ are interchanged when we interchange $v_1$ and $v_2$. The simple algorithm leading to these analytic solutions can be applied to general equations based on derivative expansions even with higher-order terms in $\omega$ and/or $k$.

Now, making the use of the lesson from the Alfven waves discussed in Sec. 4.1, we point out that the magneto-sonic modes also suffer from the breakdown of the small $k$ expansion near $\theta = \pi/2$. This is again caused by another small quantity, $\cos\theta$, that induces divergence in the higher-order terms in the small $k$ expansion. When the cosine factor becomes small near $\theta = \pi/2$, one should use the small cosine expansion to get a correct result. To see this issue, we first take the limit $\theta \to \pi/2$ in the velocities (4.9). In this limit, one finds that

$$v_1(\theta = \frac{\pi}{2}) = \sqrt{c_s^2 + v_A^2 - c_s^2 v_A^2}, \quad v_2(\theta = \frac{\pi}{2}) = 0. \tag{4.16}$$

The slow magneto-sonic waves do not propagate in the perpendicular direction and become purely diffusive modes. This implies the potential occurrence of the issue because, if there were a linear term, the counter-propagating modes should have a degenerate damping rate

because of the parity invariance. The damping rates $w_{1,2}$ in the same limit read

$$w_1(\theta=\frac{\pi}{2}) = \frac{1}{2}\Big(\eta'_\perp + \zeta'_\perp + \rho'_\perp(1-c_s^2)^2\frac{v_A^2}{v_1^2}\Big), \tag{4.17a}$$

$$w_2(\theta=\frac{\pi}{2}) \stackrel{?}{=} \frac{\eta'_\parallel}{2(1-v_A^2)} + \frac{\rho'_\perp c_s^2}{2v_1^2(1-v_A^2)}. \tag{4.17b}$$

Just below, we confirm that the above degenerate $w_2(\theta=\frac{\pi}{2})$ is *not* a correct result. To get the correct result, we take the limit $\theta\to\pi/2$ in Eq. (3.14). Then, the matrix equation reads

$$\begin{pmatrix} -i\frac{Bc_s^2}{h(1-v_A^2)}\rho'_\perp k_\perp^2 & -k_\perp B & \omega+i\rho'_\perp k_\perp^2 & 0 \\ \omega & -hk_\perp & h\frac{v_A^2}{B}\omega & 0 \\ -c_s^2 k_\perp & h\omega+i(\zeta_\perp+\eta_\perp)k_\perp^2 & -h\frac{v_A^2}{B}k_\perp & 0 \\ 0 & 0 & 0 & h(1-v_A^2)\omega+i\eta_\parallel k_\perp^2 \end{pmatrix}\begin{pmatrix}\delta\epsilon\\ \delta u_x\\ \delta B_z\\ \delta u_z\end{pmatrix} = 0, \tag{4.18}$$

where we used $\delta B_x = -k_\parallel \delta B_z/k_\perp = 0$ from the Gauss law (3.6a). Similarly to the case of the Alfven modes (4.4), one readily finds decoupling of a flow perturbation $\delta u_z$ of which the dispersion relation is solely governed by the shear viscosity $\eta_\parallel$. Diagonalizing the remaining three modes and retaining the terms in the $k^2$ order, the dispersion relations at $\theta=\pi/2$ are found to be

$$\omega_{1,2}(\theta=\frac{\pi}{2}) = \pm v_1 k_\perp - \frac{i}{2}\Big(\eta'_\perp+\zeta'_\perp+\rho'_\perp(1-c_s^2)^2\frac{v_A^2}{v_1^2}\Big)k_\perp^2 + \mathcal{O}(k_\perp^3), \tag{4.19a}$$

$$\omega_3(\theta=\frac{\pi}{2}) = -\frac{i\rho'_\perp c_s^2}{v_1^2(1-v_A^2)}k_\perp^2 + \mathcal{O}(k_\perp^4), \tag{4.19b}$$

$$\omega_4(\theta=\frac{\pi}{2}) = -\frac{i\eta'_\parallel}{1-v_A^2}k_\perp^2 + \mathcal{O}(k_\perp^4). \tag{4.19c}$$

Here, $v_1 = v_1(\theta=\pi/2)$ in Eq. (4.16) is understood. The fast magneto-sonic modes (4.19a) remain propagating modes, and still have the degenerate damping rate that agrees with $w_1(\theta=\pi/2)$ in Eq. (4.17a). In contrast, the slow magneto-sonic modes reduce to the two purely diffusive modes, and the damping rates split into two distinct ones that only depend on either $\eta'_\parallel$ or $\rho'_\perp$. They are different from the degenerate damping rate $w_2(\theta=\pi/2)$ in Eq. (4.17b) that was shown in Refs. [33, 44].[2]

Now, we investigate the behaviors near $\theta=\pi/2$. When $\cos\theta < \hat{k}$, one should organize a series representation with respect to $\cos\theta$, as we have discussed in Sec. 4.6. To find the solutions in the series representations, one can apply the same algorithm introduced above. Then, we find the solutions in the form

$$\omega_i(\cos\theta) = \tilde{\omega}_i - \frac{\tilde{\omega}_i^3 X_3 + \tilde{\omega}_i^2 X_2 + \tilde{\omega}_i X_1 + X_0}{\prod_{i\neq j}(\tilde{\omega}_i - \tilde{\omega}_j)}\cos^2\theta + \mathcal{O}(\cos^4\theta), \tag{4.20}$$

[2]The same results as in Eqs. (4.19b) and (4.19c) are shown in Ref. [42], where the limits are taken for $\theta\to\pi/2$ first and then $k\to 0$.

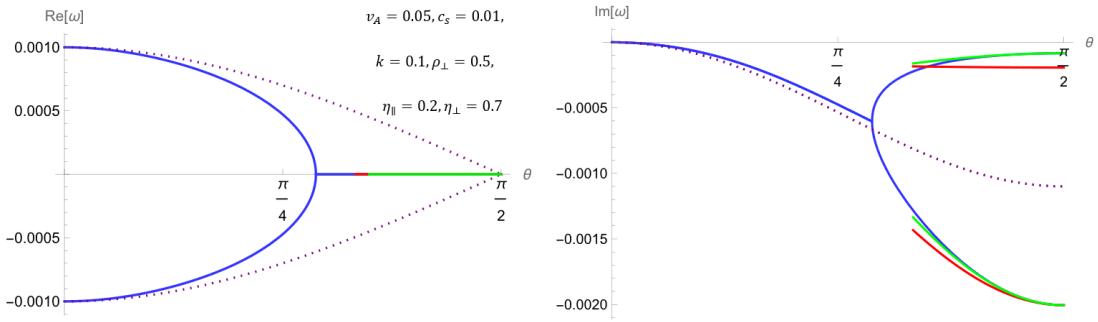



**Figure 2**. The dispersion relation for the slow magneto-sonic modes. Blue curves show the "exact solution" obtained by Mathematica. Dotted curves show the small-$k$ expansion in Eq. (4.15). Red and green curves show the small-cosine expansion (4.20), instead. In the red curve, we further expand the series coefficients up to $k^2$.

where $\tilde{\omega}_i$ on the right-hand side are the solutions for Eq. (4.18) at $\cos^2\theta = 0$, i.e., $\tilde{\omega} = \omega(\theta = \pi/2)$. It is interesting that the fast and slow sonic-modes, which were previously labeled as $\omega_{1,2}$ and $\omega_{3,4}$, are mixed among themselves in the correction terms of order $\cos^2\theta$. The explicit forms of $X_i$ are given as

$$X_0 = c_s^2 v_A^2 k^4 + k^6 c_s^2 \rho'_\perp \Big[ (\zeta'_\perp + \eta'_\perp) - \frac{2\zeta'_\times}{1-v_A^2} + \frac{\zeta'_\parallel - 2(2-v_A^2)\eta'_\parallel}{(1-v_A^2)^2} \Big], \tag{4.21a}$$

$$\begin{aligned} X_1 = {} & ik^4 \Big[ \frac{v_A^2(\eta'_\parallel - \zeta'_\parallel + c_s^2\rho'_\perp)}{1-v_A^2} - c_s^2(\zeta'_\parallel + \zeta'_\perp - 2\zeta'_\times - 4\eta'_\parallel + \eta'_\perp) \Big] \\ & + \frac{ik^6}{(1-v_A^2)^2} \rho'_\perp \Big[ c_s^2 v_A^2 \big( \zeta'^2_\times + 2\zeta'_\times \eta'_\parallel - \zeta'_\parallel(\zeta'_\perp + \eta'_\perp) + 3\eta'_\parallel(\zeta'_\perp + \eta'_\perp) \big) \\ & + (1-v_A^2)\big( \zeta'^2_\times - \zeta'_\parallel(\zeta'_\perp + \eta'_\perp) + 2\eta'_\parallel(\zeta'_\perp + \zeta'_\times + \eta'_\perp) \big) \Big], \end{aligned} \tag{4.21b}$$

$$\begin{aligned} X_2 = {} & -c_s^2 v_A^2 k^2 - \frac{k^4}{(1-v_A^2)^2} \Big[ c_s^2 v_A^2 \rho'_\perp \big( \zeta'_\parallel - 2(1-v_A^2)\zeta'_\perp - (1+v_A^2)\eta'_\parallel - 2(1-v_A^2)\eta'_\perp \big) \\ & - (1-v_A^2) \Big\{ \rho'_\perp \big( v_A^2 \eta'_\parallel + (1-v_A^2)(\zeta'_\perp + \eta'_\perp) \big) \\ & + \zeta'^2_\times + 2\eta'_\parallel(\zeta'_\perp + \zeta'_\times + \eta'_\perp) - \zeta'_\parallel(\zeta'_\perp + \eta'_\perp + \rho'_\perp) \Big\} \Big], \end{aligned} \tag{4.21c}$$

$$X_3 = \frac{ik^2}{1-v_A^2} \Big[ \zeta'_\parallel - (1-v_A^2)(\zeta'_\perp + \eta'_\perp) - v_A^2(\eta'_\parallel + c_s^2\rho'_\perp) \Big], \tag{4.21d}$$

These results should replace the naive small $k$ expansion when $\cos\theta < \hat{k}$.

In Fig. 2, we show the dispersion relations for the slow magneto-sonic modes. The "exact solution" for Eq. (3.11b) is shown by blue-solid curves, of which the analytic forms are obtained by an automated command in Mathematica; The complicated expressions are not useful to be shown here. Note that the "exact solution" contains the higher-order corrections beyond the $k^2$ order and does not mean that it is the true goal for the first-order hydrodynamics. As the angle approaches the right angle $\theta = \pi/2$, the small $k$ expansion,

shown by dotted curves, deviates from the "exact solution" even at a fixed small value of $k$. Instead of the small $k$ expansion, the small-cosine expansion (4.20) should be effective in this region as discussed above. The green curves show the small-cosine expansion (4.20) without an expansion for the momentum $k$, while the red curves show the same expansion but with a further expansion of the series coefficients up to $k^2$. Both curves well reproduce the branching in the imaginary part near the right angle. At $\theta = \pi/2$, i.e., $\cos\theta = 0$, the deviation between the green and red curves solely comes from the choice of $\tilde{\omega}_i$ with or without the higher-order terms in $k$ in Eq. (4.20). When $\cos\theta \neq 0$, the coefficients for the $\cos^2\theta$ corrections also contain the $k$ dependence in $X_i$ as well as $\tilde{\omega}_i$ both in the numerator and denominator.

Next, we comment on extracting magnitudes of the transport coefficients by comparing the linear-mode solutions with experiments/observations. Remarkably, the general solution obtained in Eq. (4.15) is necessary to determine the cross bulk viscosity $\zeta_\times$, which cannot be determined with the limiting solutions at $\theta = 0, \pi/2$ [33]. It is a cross quantity between the parallel and perpendicular directions with respect to the magnetic field, and cannot be induced along a single direction at $\theta \neq 0$ or $\pi/2$ (see Eqs. (4.17) and (4.22) below). Yet, even with the complete solution, one cannot determine a separation between $\eta'_\perp$ and $\zeta'_\perp$ that appears only in the sum in Eq. (3.14b) and the general solutions (4.14) accordingly. In the Alfven modes (4.2), $\eta'_\perp$ and $\rho'_\parallel$ also appear in the sum in the small $k$ expansion. However, the dependence on these two transport coefficients is split in Eq. (4.3) at $\theta = \pi/2$, and they can be determined with the linear-mode solutions. The other pair $\eta'_\perp$ and $\rho'_\parallel$ appear separately in the magneto-sonic modes in Eq. (4.19).

Before closing this section, it is also instructive to confirm the limit of $\theta \to 0, \pi$, i.e., $k_\perp \to 0$. As mentioned below Eq. (4.9), one of the pairs should become degenerate with the Alfven modes because of the rotational symmetry. When $\theta = 0$, we have

$$W_1(\theta = 0) = \eta'_\parallel c_s^2 + \rho'_\perp c_s^2 + \zeta'_\parallel \frac{v_A^2}{1 - v_A^2}, \tag{4.22a}$$

$$W_2(\theta = 0) = \eta'_\parallel + \rho'_\perp + \frac{\zeta'_\parallel}{1 - v_A^2}. \tag{4.22b}$$

Then, we find that, when $c_s \geq v_A$,

$$w_1(\theta = 0) = \frac{\zeta'_\parallel}{2(1 - v_A^2)}, \quad w_2(\theta = 0) = \frac{1}{2}(\eta'_\parallel + \rho'_\perp), \tag{4.23}$$

and that, when $c_s < v_A$,

$$w_1(\theta = 0) = \frac{1}{2}(\eta'_\parallel + \rho'_\perp), \quad w_2(\theta = 0) = \frac{\zeta'_\parallel}{2(1 - v_A^2)}. \tag{4.24}$$

In both cases, either of the pairs becomes degenerate with the Alfven modes (4.2). The other pair is the sound modes damped by the bulk viscosity. The shear viscosity does not contribute to the damping rate in this limit, differently from the usual sound modes in the absence of a magnetic field.

## 5 Causality and stability

In the last section, we have obtained the pairs of the Alfven waves (4.2) and the slow and fast magneto-sonic waves (4.10). In this section, we show that the phase velocities of the Alfven and magneto-sonic waves are always smaller than the speed of light and that the first-order derivative corrections in those solutions (4.2) and (4.10) always act as damping factors as long as the transport coefficients satisfy the inequalities (2.11) required by the second law of thermodynamics. The former implies causality in the ideal MHD. The latter implies a stability of equilibrium state in the *fluid rest frame* where the linear-mode analysis has been performed in the last section.

However, the above two properties in general do not guarantee causality beyond the ideal order or stability in an arbitrary Lorentz frame. We briefly discuss causality and stability of relativistic MHD along with the recent developments in the literature.

### 5.1 The Alfven and magneto-sonic velocities in the ideal MHD

First, we focus on the linear terms in $k$, putting the higher-order terms aside. This is the ideal MHD limit. We assume that the sound velocity satisfies an inequality

$$0 \leq c_s \leq 1. \tag{5.1}$$

We also assume that the Alfven velocity (3.13) satisfies an inequality

$$0 \leq v_A \leq 1. \tag{5.2}$$

The latter inequality is evident when the energy density and pressure are separated as in Eq. (3.2), where the Alfven velocity reads $v_A^2 = \frac{B^2/\mu_m}{\epsilon+P+B^2/\mu_m}$. This inequality should hold unless a strong coupling between the matter and magnetic components significantly reduces the total energy density and pressure.

Under the above inequalities, one can show that the velocities of the magneto-sonic waves (4.9) satisfy the inequalities

$$0 \leq v_2 \leq c_s \leq v_1 \leq 1 \quad \text{and} \quad 0 \leq v_2 \leq v_A \leq v_1 \leq 1. \tag{5.3}$$

Here, the relative magnitude between $c_s$ and $v_A$ is not assumed. A straightforward proof is given in Appendix A.1. The inequalities in Eqs. (5.2) and (5.3) indicate that the Alfven and magneto-sonic waves propagate with subluminal speeds in the ideal MHD.

### 5.2 Dissipative nature of the first-order corrections in the fluid rest frame

Next, we focus on the $k^2$ terms in the dispersion relations. Since all the present transport coefficients are dissipative in nature, the Alfven and magneto-sonic waves are expected to acquire damping effects. We show that the pure imaginary $k^2$ terms, which have been found in the previous section, take definite signs as long as the transport coefficients satisfy the inequalities (2.11) required by the second law of thermodynamics.

In the Alfven waves (4.2), it is clear that the pure imaginary coefficient in front of $k^2$ is always negative for $\eta_{\parallel,\perp} \geq 0$ and $\rho_{\parallel,\perp} \geq 0$ required by the inequalities (2.11).

The magneto-sonic waves also have the pure imaginary corrections at the $k^2$ order. The explicit forms of $w_{1,2}$ in Eq. (4.15) are given as

$$w_{1,2}(\theta) = \frac{\mp 1}{2(v_1^2 - v_2^2)}\Big[\eta_\parallel'\Big(c_s^2\cos^2(2\theta) - v_{1,2}^2 + \frac{(1-v_{1,2}^2)v_A^2}{1-v_A^2}\sin^2\theta\Big) \\ +\rho_\perp'\Big(c_s^2 - v_{1,2}^2 + \frac{(1-v_{1,2}^2)v_A^2 c_s^2}{1-v_A^2}\sin^2\theta\Big) + (\eta_\perp' + \zeta_\perp')(c_s^2\cos^2\theta - v_{1,2}^2)\sin^2\theta \\ +\zeta_\parallel'\frac{\mathcal{V}^2 - v_{1,2}^2 - c_s^2\cos^2\theta}{1-v_A^2}\cos^2\theta - 2\zeta_\times' c_s^2\cos^2\theta\sin^2\theta\Big], \tag{5.4}$$

where the upper and lower signs are for $w_1$ and $w_2$, respectively. As detailed in Appendix A.2, one can show that

$$w_{1,2}(\theta) \geq 0, \tag{5.5}$$

for any angle $\theta$. This means that the pure imaginary corrections always take negative signs [see the conventions in Eq. (4.10).] We note that the semi-positivity of $w_{1,2}$ can be shown *irrespective of* the sign of $\zeta_\times'$ as long as the transport coefficients satisfy the inequalities (2.11), none of which indeed specifies the sign of $\zeta_\times'$.

When the angle approaches $\theta = \pi/2$, one should refer to the small-cosine expansions in Eqs. (4.6) and (4.19) or (4.20). The leading-order terms of order $\cos^0\theta$, i.e., when $\theta = \pi/2$, indicate purely diffusive modes with definite signs. The cosine corrections are expected to be smaller than these leading-order terms within the regions of validity for the cosine expansions. Then, the signs should remain definite as seen in Figs. 1 and 2 with the corrections.

The above inequalities imply that both the Alfven and magneto-sonic waves are damped out by exponential factors for an observer in the fluid rest frame where the linear-mode analysis has been performed. However, those inequalities are not sufficient conditions for stability in an arbitrary Lorentz frame, but are necessary conditions. We discuss stability and causality conditions in a more general perspective below.

### 5.3 Covariant stability in anisotropic systems

It has been known that diffusive modes in the first-order hydrodynamics are acausal and that such diffusive modes, which are damped out in the fluid rest frame, can be transformed into growing modes in a general Lorentz frame [40]. Relativistic hydrodynamic theories containing such instability may not work in practice because the stability of local equilibria in a certain reference frame, e.g., a lab frame as often interested, is not guaranteed. Therefore, it is important to understand the origin of the instability and the necessary and/or sufficient conditions for the *covariant stability* where the local equilibria are stable in *any* Lorentz frame.

Here, we briefly discuss the covariant stability for MHD with a slight extension of the recent discussions by Gavassino for isotropic systems [51, 52]. We write the solutions for the linear perturbations (3.10) in a covariant form

$$\Psi = \tilde{\Psi} e^{-ik_\mu x^\mu}, \tag{5.6}$$

where $\Psi$ is an eigenvector and the corresponding momentum $k^\mu = (\omega, \boldsymbol{k})$ satisfies a dispersion relation obtained from the (linearized) hydrodynamic equations. However, as we have seen in the previous sections, the dispersion relations are not given in Lorentz-invariant forms like the free-particle on-shell conditions in quantum field theories, which is due to the derivative expansion. Therefore, we need to understand how the dispersion relations are transformed by a Lorentz boost.

First, notice that, in dissipative hydrodynamics, the spatial component of the momentum $\boldsymbol{k}$ develops an imaginary part in a general Lorentz frame due to the mixing between $\omega$ and $\boldsymbol{k}$ under Lorentz boosts. For the general solution (5.6) to be covariantly stable, one should have $\Im m\,[k_\mu x^\mu] \leq 0$. For this condition to be satisfied for observers in the forward light cone, the imaginary part of $k^\mu$ should be a vector lying outside the forward light cone, requiring that

$$\Im m\,\omega \leq -|\Im m\,\boldsymbol{k}| \quad \text{or} \quad (\Im m\,k^\mu)(\Im m\,k_\mu) = -(\Im m\,\omega)^2 + |\Im m\,\boldsymbol{k}|^2 \geq 0. \tag{5.7}$$

Combining the above inequalities, one finds a necessary condition for the covariant stability

$$\Im m\,\omega \leq |\Im m\,\boldsymbol{k}|\,. \tag{5.8}$$

This is an extension of the condition for an isotropic system obtained by Heller et al. from the analytic property of a general retarded propagator in causal quantum field theory theories [50] and interpreted by Gavassini as the covariantly stability condition for relativistic dissipative hydrodynamics [52]. The inequality (5.8) is a stronger condition than that for isotropic systems where there is essentially only one independent component of $\boldsymbol{k}$. For MHD, two of the three components should be treated independently as there remains a rotational symmetry around the magnetic-field direction.

When there are dissipative effects, i.e., $\Im m\,\omega \neq 0$, with $\Im m\,\boldsymbol{k} = 0$ in the fluid rest frame, one may consider successive multiple Lorentz boosts; The first boost generates an imaginary part of $\boldsymbol{k}$ due to the mixing with $\omega$, and the subsequent boosts require the inequality (5.8) for the causal stability. The result of these successive boosts is not equivalent to that of a single boost by a sum of the boost velocities, because of the non-Abelian nature of the Lorentz group. In analogy with the discussion about the Thomas precession [69], such a sequence of Lorentz boosts is required to move from one Lorentz frame to another, e.g., from the rest frame of a fluid cell to the lab frame, *when the fluid cell is accelerated*.

It is instructive to explicitly see the occurrence of instability when the inequality (5.8) is *not* satisfied [52]. In fact, the dispersion relations turn into unstable ones when boosted by a velocity

$$\boldsymbol{v} = \frac{\Im m\,\boldsymbol{k}}{\Im m\,\omega}\,, \tag{5.9}$$

that satisfies $|\boldsymbol{v}| \leq 1$ when the inequality (5.8) is not satisfied. Boosting the imaginary part of the momentum, one finds that

$$\begin{aligned}\Im m\,k^\mu \to \Im m\,k'^\mu &= \gamma(\Im m\,\omega - \boldsymbol{v}\cdot\Im m\,\boldsymbol{k}, -\gamma\boldsymbol{v}\Im m\,\omega + \Im m\,\boldsymbol{k} + \frac{(\gamma-1)\boldsymbol{v}\cdot\Im m\,\boldsymbol{k}}{v^2}\boldsymbol{v}) \\ &= \gamma(\frac{(\Im m\,\omega)^2 - |\Im m\,\boldsymbol{k}|^2}{\Im m\,\omega}, \boldsymbol{0})\,.\end{aligned} \tag{5.10}$$

This means that an observer in the new frame claims the existence of an unstable Fourier mode because $\Im m\, \omega' > 0$ and $\Im m\, \boldsymbol{k}' = 0$. Therefore, the inequality (5.8) is indeed necessary for the covariant stability.

If the inequality (5.8) is satisfied in one of the Lorentz frames, it is, by construction, satisfied in all the Lorentz frames connected by Lorentz boosts. Then, the covariant stability will be fulfilled *if* signals only reach observers inside the forward light cone, i.e., if theories respect causality (see a theorem in Sec. III.B in Ref. [70]). However, it has been known that, in dissipative hydrodynamics, signals reach observers outside the forward light cone. The acausal tails of dissipative modes are not only illegitimate in relativity in the first place but also observed as unstable modes outside the light cone (see a theorem in Sec. III.A in Ref. [70]). This is because, for a spacelike separation $(y-x)^2 > 0$, there is a Lorentz boost that inverts the chronicle ordering as $y^0 - x^0 \to y'^0 - x'^0 = -(y^0 - x^0)$, making the meaning of dissipation and growth observer-dependent concepts. Therefore, the covariant stability is fulfilled only in causal theories. The first-order MHD is stable in the fluid rest frame as shown above, but is not causal. The subluminal magneto-sonic velocities (5.3) only serve as necessary conditions for causality once the dissipative effects are included.

Causality is often a consequence of subtle cancellation among the acausal tails leaking across the light cone. A well-known example is the Klein-Gordon field in quantum field theory [71]. In Ref. [72], it is stated that any dispersion relation can leak across the light cone unless a medium is not dispersive or, in other words, dispersion relations are polynomials of the first order at most, i.e., $\omega(k) = a_0 + a_1 k$ (see also Ref. [50]). This implies that inspecting each dispersion relation alone does not guarantee causality.[3] Notions of velocities for a single dispersion relation, such as the front velocity, the group velocity, and the phase velocity, may be useful for screening apparently acausal theories, but do not serve as a sufficient causality test; Besides, it should be noticed that the font velocity is defined at the ultraviolet limit $k \to \infty$ outside the hydrodynamic regime. Instead, it will be useful to investigate causal structures of a set of partial differential equations with the method of characteristics (see, e.g., Refs. [53, 73] and references therein), though it will require more efforts in future works. It is worth adding that the covariantly stable condition (5.8) is derived independently of any notion of velocities (see also Theorem 2 in Ref. [52] where a criterion of causality is manifestly implemented without any notion of velocities).

## 6 Conclusion and outlook

In this paper, we investigated linear waves in relativistic magnetohydrodynamics in detail. Especially, in Sec. 4.2, we provided a simple and general analytic algorithm for the solution search. Based on this algorithm, we showed analytic solutions for the magneto-sonic waves

[3] The Israel-Stewart theory is one of such cases where the dissipative corrections, e.g., the viscous tensors, are promoted to independent variables, and the conservation laws are cast into a larger set of the first-order differential equations *both* in space and time (if the vorticity terms are neglected) [56]. Otherwise, one can convert a set of second-order differential equations to that of first-order differential equations by introducing auxiliary fields [55].

that have been missing in the literature for a long time. The algorithm can be applied to other hydrodynamic equations or any general set of equations based on a derivative expansion. We will provide an application elsewhere [68]. Also, while we focused on the Landau frame in the present work, it is interesting to investigate analytic solutions in a general choice of hydrodynamic variables (cf. Ref. [44]).

On the other hand, we also found that the small-momentum expansion for the solutions breaks down in MHD when the momentum direction is nearly or exactly perpendicular to an equilibrium magnetic field. This issue occurs both in the Alfven and magneto-sonic waves and stems from the competition between two small quantities involved in the solutions that are the momentum and the trigonometric functions representing the spatial anisotropy in MHD, i.e., a cosine function in the present convention. When the cosine becomes small near the right angle, we found that the higher-order terms in the small-momentum expansion diverge, spoiling the small-momentum expansion. The breakdown of the small-momentum expansion can be a general issue emerging in anisotropic systems. We provided alternative expressions of the solutions based on the small-cosine expansion in Eqs. (4.6) and (4.20) that work accurately near the right angle as shown in Figs. 1 and 2.

Lastly, we investigated the issues of causality and stability in the first-order relativistic MHD based on the analytic solutions. We showed that the Alfven and magneto-sonic velocities are less than the speed of light and that the first-order corrections always act as damping effects in the fluid rest frame. As mentioned in Sec. 5.3, these conditions are, however, not sufficient for the covariant stability, i.e., the stability in all the Lorentz frames. The main and general reason is that dissipative hydrodynamics exhibits acausal propagation across the forward light cone. Such acausal signals can be observed as unstable modes in the spacelike regions, indicating an imtimate connection between the issues of causality and stability. So far, the method of moment expansion, which leads to the Israel-Stewart theory, has been invoked in Refs. [74, 75] to formulate causal and stable MHD (see also Ref. [37] for a review). It is yet left as an open question to formulate covariantly stable MHD based on the magnetic-flux conservation (cf. Sec. 2). Other future works include computation of the transport coefficients (see [76–86] for recent studies). These developments will promote further numerical studies [1–5].

## Acknowledgments

We thank Yihui Tu, Shi Pu, and Dong-Lin Wang for useful discussions. This work is partially supported by the JSPS KAKENHI under grant Nos. 20K03948 and 22H01216, and the start-up Grant No. XRC-23112 of Fuzhou University.

## A inequalities

### A.1 Inequalities for the Alfven and magneto-sonic velocities

We assume that the equations of state satisfy the inequalities $0 \leq c_s \leq 1$ and $0 \leq v_A \leq 1$ as stated in Eqs. (5.1) and (5.2). Then, we show that the velocities of the magneto-sonic

waves (4.9) satisfy the inequalities (5.3), i.e.,

$$0 \le v_2 \le c_s \le v_1 \le 1 \quad \text{and} \quad 0 \le v_2 \le v_A \le v_1 \le 1. \tag{A.1}$$

Here, the relative magnitude between $c_s$ and $v_A$ is not assumed.

It is easy to show that $\mathcal{V}^2 \ge 0$ by comparing the magnitudes of the two terms in $\mathcal{V}^2$ as

$$(c_s^2 + v_A^2)^2 - (c_s^2 v_A^2 \sin^2\theta)^2 = c_s^4 + v_A^4 + c_s^2 v_A^2 (2 - c_s^2 v_A^2 \sin^4\theta) \ge 0. \tag{A.2}$$

Then, it is obvious that $v_{1,2} \ge 0$.

Next, we show that $v_{1,2} \le 1$. Comparing the two sides, one finds that

$$(2 - \mathcal{V}^2)^2 - \sqrt{\mathcal{V}^4 - 4v_A^2 c_s^2 \cos^2\theta}^{\,2} = -4(1 - v_A^2)(1 - c_s^2) \le 0. \tag{A.3}$$

Note also that $2 - \mathcal{V}^2 = (2 - c_s^2 - v_A^2) + c_s^2 v_A^2 \sin^2\theta \ge 0$. Then, one can conclude that $v_{1,2} \le 1$.

Lastly, we show the relative magnitudes of $v_{1,2}$ to $c_s$ and $v_A$. The difference between $v_{1,2}$ and $c_s$ reads

$$c_s^2 - v_{1,2}^2 = (c_s^2 - \frac{1}{2}\mathcal{V}^2) \mp \frac{1}{2}\sqrt{\mathcal{V}^4 - 4v_A^2 c_s^2 \cos^2\theta}. \tag{A.4}$$

The relative magnitudes of the two terms is examined as

$$\begin{aligned} \frac{1}{2^2}\left(\mathcal{V}^4 - 4v_A^2 c_s^2 \cos^2\theta\right) - (c_s^2 - \frac{1}{2}\mathcal{V}^2)^2 &= (-v_A^2 \cos^2\theta - c_s^2 + \mathcal{V}^2) c_s^2 \\ &= (1 - c_s^2) v_A^2 c_s^2 \sin^2\theta \\ &\ge 0. \end{aligned} \tag{A.5}$$

Therefore, the sign of the difference in Eq. (A.4) is determined by that of the square-root term regardless of the sign of the other term. Then, one can conclude that $v_2 \le c_s \le v_1$. By the same token, one can examine the difference

$$v_A^2 - v_{1,2}^2 = (v_A^2 - \frac{1}{2}\mathcal{V}^2) \mp \frac{1}{2}\sqrt{\mathcal{V}^4 - 4v_A^2 c_s^2 \cos^2\theta}. \tag{A.6}$$

The relative magnitudes of the two terms is examined as

$$\begin{aligned} \frac{1}{2^2}\left(\mathcal{V}^4 - 4v_A^2 c_s^2 \cos^2\theta\right) - (v_A^2 - \frac{1}{2}\mathcal{V}^2)^2 &= (-c_s^2 \cos^2\theta - v_A^2 + \mathcal{V}^2) v_A^2 \\ &= (1 - v_A^2) v_A^2 c_s^2 \sin^2\theta \\ &\ge 0. \end{aligned} \tag{A.7}$$

Then, one can conclude that $v_2 \le v_A \le v_1$.

Following the above proof, we conclude the inequalities (5.3).

### A.2 Dissipative corrections in the fluid rest frame

The next-to-leading order solutions in the magneto-sonic modes (4.10) are obtained with $w_{1,2}$ given in Eq. (5.4). Here, we show that $w_{1,2} \geq 0$ for the transport coefficients that satisfy the inequalities (2.11) required by the second law of thermodynamics. In the following discussion, one can forget about the positive overall factor like $1/(2(v_1^2 - v_2^2))$ that is irrelevant for examining the signs of $w_{1,2}$.

We begin with the terms associated with the bulk viscosities $\zeta_{\parallel,\perp,\times}$. Picking up these terms from $w_{1,2}$ in Eq. (5.4), we have

$$w_{1,2}^{\zeta} := s_{1,2}\zeta_{\perp} + t_{1,2}\zeta_{\parallel} + u\zeta_{\times}\,, \tag{A.8}$$

where

$$\begin{aligned} &s_{1,2} = \mp(c_s^2\cos^2\theta - v_{1,2}^2)\sin^2\theta, \quad t_{1,2} = \mp\frac{\mathcal{V}^2 - v_{1,2}^2 - c_s^2\cos^2\theta}{1 - v_A^2}\cos^2\theta, \\ &u = \pm 2c_s^2\cos^2\theta\sin^2\theta. \end{aligned} \tag{A.9}$$

One can show that $s_{1,2} \geq 0$ and $t_{1,2} \geq 0$ as we will see later. Assuming these positivities for the moment, one finds that

$$w_{1,2}^{\zeta} \geq 2\sqrt{s_{1,2}t_{1,2}\zeta_{\parallel}\zeta_{\perp}} + u\zeta_{\times}\,. \tag{A.10}$$

Further examining the relative magnitude of the two terms on the right-hand side, we have

$$4s_{1,2}t_{1,2}\zeta_{\parallel}\zeta_{\perp} - (u\zeta_{\times})^2 \geq (4s_{1,2}t_{1,2} - u^2)\zeta_{\times}^2 = 0\,, \tag{A.11}$$

where we used the inequality from the thermodynamic constraint (2.11) and the explicit forms of $s, t, u$ that, in both cases, lead to $(4s_{1,2}t_{1,2} - u^2) = 0$. Therefore, one can conclude that $w_{1,2}^{\zeta} \geq 0$ as long as $s_{1,2}, t_{1,2} \geq 0$ *irrespective of* the sign of $u$.

To show that $s_{1,2} \geq 0$, one can arrange it with the explicit forms of $v_{1,2}^2$ as

$$s_{1,2} = \mp\big(c_s^2\cos^2\theta - \frac{1}{2}\mathcal{V}^2\big) + \frac{1}{2}\sqrt{\mathcal{V}^4 - 4v_A^2c_s^2\cos^2\theta}. \tag{A.12}$$

As for $t_{1,2} \geq 0$, one can focus on the numerator

$$\begin{aligned} (1 - v_A^2)t_{1,2} &= \mp(\mathcal{V}^2 - v_{1,2}^2 - c_s^2\cos^2\theta) \\ &= \pm\big(c_s^2\cos^2\theta - \frac{1}{2}\mathcal{V}^2\big) + \frac{1}{2}\sqrt{\mathcal{V}^4 - 4v_A^2c_s^2\cos^2\theta}. \end{aligned} \tag{A.13}$$

In both cases, one can show that the square-root term is always larger than the absolute value of the first term, that is,

$$\begin{aligned} \frac{1}{2^2}(\mathcal{V}^4 - 4v_A^2c_s^2\cos^2\theta) - (c_s^2\cos^2\theta - \frac{1}{2}\mathcal{V}^2)^2 &= -c_s^2\cos^2\theta(v_A^2 + c_s^2\cos^2\theta - \mathcal{V}^2) \\ &= c_s^4\sin^2\theta\cos^2\theta(1 - v_A^2) \\ &\geq 0. \end{aligned} \tag{A.14}$$

Therefore, we have shown that $s_{1,2} \geq 0$ and $t_{1,2} \geq 0$, and accordingly that $w^{\zeta}_{1,2} \geq 0$.

We have three remaining terms associated with $\eta'_{\parallel}, \rho'_{\perp}, \eta'_{\perp}$. The last one $\eta'_{\perp}$ only appears with $\zeta'_{\perp}$ in Eq. (5.4), so that we have already shown the positivity of this term just above. We examine the remaining two terms below. The coefficients in front of $\rho'_{\perp}$ is arranged as

$$\frac{\mp 1}{1-v_A^2}\big[\,(1-v_A^2)(c_s^2-v_{1,2}^2)+(1-v_{1,2}^2)v_A^2c_s^2\sin^2\theta\,\big]. \tag{A.15}$$

According to the inequalities (A.1) for the velocities, the right-hand side is semi-positive definite for $v_2$. As for $v_1$, one can arrange the expression between the square brackets as

$$\begin{aligned}
&(1-v_A^2)(c_s^2-v_1^2)-(1-v_1^2)v_A^2c_s^2\sin^2\theta \\
&\quad = c_s^2(1-v_A^2+v_A^2\sin^2\theta)-v_1^2J \\
&\quad = \big[c_s^2(1-v_A^2+v_A^2\sin^2\theta)-\frac{1}{2}\mathcal{V}^2J\big]-\frac{J}{2}\sqrt{\mathcal{V}^4-4v_A^2c_s^2\cos^2\theta},
\end{aligned} \tag{A.16}$$

where $J = 1 - v_A^2 + v_A^2 c_s^2 \sin^2\theta$. Comparing the magnitudes of the two terms, one finds that

$$\begin{aligned}
\Big[\frac{J}{2}\sqrt{\mathcal{V}^4-4v_A^2c_s^2\cos^2\theta}\,\Big]^2-\big[c_s^2(1-v_A^2+v_A^2\sin^2\theta)-\frac{1}{2}\mathcal{V}^2J\big]^2 &= v_A^2c_s^2(1-v_A^2)(1-c_s^2)^2\sin^2\theta \\
&\geq 0\,.
\end{aligned} \tag{A.17}$$

This means that the sign of the left-hand side in Eq. (A.16) is determined by that of the square-root term, which is negative. Therefore, for both $w_{1,2}$, one can conclude that the coefficients in front of $\rho'_{\perp}$ are semi-positive definite in Eq. (A.15).

Lastly, the coefficients in front of $\eta'_{\parallel}$ in Eq. (5.4) can be arranged as

$$\frac{\pm 1}{1-v_A^2}\big[(1-v_A^2\cos^2\theta)v_{1,2}^2-\{(1-v_A^2)c_s^2\cos^2(2\theta)+v_A^2\sin^2\theta\}\big]. \tag{A.18}$$

Inserting the explicit forms of $v_{1,2}$, we have

$$(1-v_A^2\cos^2\theta)v_{1,2}^2-\{(1-v_A^2)c_s^2\cos^2(2\theta)+v_A^2\sin^2\theta\} = K_{1,2}+L. \tag{A.19}$$

where

$$K_{1,2} = \pm\frac{1}{2}(1-v_A^2\cos^2\theta)\sqrt{\mathcal{V}^4-4v_A^2c_s^2\cos^2\theta}, \tag{A.20a}$$

$$L = \frac{1}{2}(1-v_A^2\cos^2\theta)\mathcal{V}^2-(1-v_A^2)c_s^2\cos^2(2\theta)-v_A^2\sin^2\theta, \tag{A.20b}$$

where the upper and lower signs from $v_{1,2}$ are for $K_1$ and $K_2$, respectively. Examining the relative magnitude of the two terms, one finds that

$$K_{1,2}^2-L^2 = \frac{1}{4}(1-v_A^2)\sin^2(2\theta)\big\{v_A^2-c_s^2(2-v_A^2)\cos(2\theta)\big\}^2 \geq 0. \tag{A.21}$$

This inequality, together with $(1-v_A^2\cos^2\theta) \geq 0$, means that the overall signs in Eq. (A.19) are determined by that of $K_{1,2}$. Then, one can conclude that the coefficients in front of $\eta'_{\parallel}$ are semi-positive definite for both $w_{1,2}$.

From the above, we conclude that the magneto-sonic modes (4.10) are damped out by the semi-positive damping factor $w_{1,2}$ given in Eq. (5.4). We emphasize that the semi-positivity of $w_{1,2}$ has been shown irrespective of the sign of $\zeta'_\times$ as long as the transport coefficients satisfy the inequalities (2.11), none of which specifies the sign of $\zeta'_\times$.

## References


[1] G. Inghirami, L. Del Zanna, A. Beraudo, M. H. Moghaddam, F. Becattini and M. Bleicher, *Numerical magneto-hydrodynamics for relativistic nuclear collisions*, *Eur. Phys. J. C* **76** (2016) 659 [1609.03042].

[2] G. Inghirami, M. Mace, Y. Hirono, L. Del Zanna, D. E. Kharzeev and M. Bleicher, *Magnetic fields in heavy ion collisions: flow and charge transport*, *Eur. Phys. J. C* **80** (2020) 293 [1908.07605].

[3] K. Nakamura, T. Miyoshi, C. Nonaka and H. R. Takahashi, *Directed flow in relativistic resistive magneto-hydrodynamic expansion for symmetric and asymmetric collision systems*, *Phys. Rev. C* **107** (2023) 014901 [2209.00323].

[4] K. Nakamura, T. Miyoshi, C. Nonaka and H. R. Takahashi, *Relativistic resistive magneto-hydrodynamics code for high-energy heavy-ion collisions*, *Eur. Phys. J. C* **83** (2023) 229 [2211.02310].

[5] K. Nakamura, T. Miyoshi, C. Nonaka and H. R. Takahashi, *Charge-dependent anisotropic flow in high-energy heavy-ion collisions from a relativistic resistive magneto-hydrodynamic expansion*, *Phys. Rev. C* **107** (2023) 034912 [2212.02124].

[6] J. C. McKinney and R. D. Blandford, *Stability of Relativistic Jets from Rotating, Accreting Black Holes via Fully Three-Dimensional Magnetohydrodynamic Simulations*, *Mon. Not. Roy. Astron. Soc.* **394** (2009) 126 [0812.1060].

[7] J. C. McKinney, A. Tchekhovskoy and R. D. Blandford, *General Relativistic Magnetohydrodynamic Simulations of Magnetically Choked Accretion Flows around Black Holes*, *Mon. Not. Roy. Astron. Soc.* **423** (2012) 3083 [1201.4163].

[8] O. Bromberg and A. Tchekhovskoy, *Relativistic MHD simulations of core-collapse GRB jets: 3D instabilities and magnetic dissipation*, *Mon. Not. Roy. Astron. Soc.* **456** (2016) 1739 [1508.02721].

[9] S. W. Davis and A. Tchekhovskoy, *Magnetohydrodynamics Simulations of Active Galactic Nucleus Disks and Jets*, *Ann. Rev. Astron. Astrophys.* **58** (2020) 407 [2101.08839].

[10] M. Shibata, Y. T. Liu, S. L. Shapiro and B. C. Stephens, *Magnetorotational collapse of massive stellar cores to neutron stars: Simulations in full general relativity*, *Phys. Rev. D* **74** (2006) 104026 [astro-ph/0610840].

[11] J. Matsumoto, T. Takiwaki, K. Kotake, Y. Asahina and H. R. Takahashi, *2D numerical study for magnetic field dependence of neutrino-driven core-collapse supernova models*, *Mon. Not. Roy. Astron. Soc.* **499** (2020) 4174 [2008.08984].

[12] J. Matsumoto, Y. Asahina, T. Takiwaki, K. Kotake and H. R. Takahashi, *Magnetic support for neutrino-driven explosion of 3D non-rotating core-collapse supernova models*, *Mon. Not. Roy. Astron. Soc.* **516** (2022) 1752 [2202.07967].

[13] M. D. Duez, Y. T. Liu, S. L. Shapiro, M. Shibata and B. C. Stephens, *Collapse of magnetized hypermassive neutron stars in general relativity*, *Phys. Rev. Lett.* **96** (2006) 031101 [astro-ph/0510653].

[14] M. Shibata, S. Fujibayashi and Y. Sekiguchi, *Long-term evolution of a merger-remnant neutron star in general relativistic magnetohydrodynamics: Effect of magnetic winding*, *Phys. Rev. D* **103** (2021) 043022 [2102.01346].

[15] M. Shibata, S. Fujibayashi and Y. Sekiguchi, *Long-term evolution of neutron-star merger remnants in general relativistic resistive magnetohydrodynamics with a mean-field dynamo term*, *Phys. Rev. D* **104** (2021) 063026 [2109.08732].

[16] K. Kiuchi, A. Reboul-Salze, M. Shibata and Y. Sekiguchi, *A large-scale magnetic field produced by a solar-like dynamo in binary neutron star mergers*, *Nature Astronomy* (2023) [2306.15721].

[17] Z. B. Etienne, Y. T. Liu, V. Paschalidis and S. L. Shapiro, *General relativistic simulations of black hole-neutron star mergers: Effects of magnetic fields*, *Phys. Rev. D* **85** (2012) 064029 [1112.0568].

[18] K. Hayashi, S. Fujibayashi, K. Kiuchi, K. Kyutoku, Y. Sekiguchi and M. Shibata, *General-relativistic neutrino-radiation magnetohydrodynamic simulation of seconds-long black hole-neutron star mergers*, *Phys. Rev. D* **106** (2022) 023008 [2111.04621].

[19] A. Boyarsky, J. Frohlich and O. Ruchayskiy, *Self-consistent evolution of magnetic fields and chiral asymmetry in the early Universe*, *Phys. Rev. Lett.* **108** (2012) 031301 [1109.3350].

[20] I. Rogachevskii, O. Ruchayskiy, A. Boyarsky, J. Fröhlich, N. Kleeorin, A. Brandenburg et al., *Laminar and turbulent dynamos in chiral magnetohydrodynamics-I: Theory*, *Astrophys. J.* **846** (2017) 153 [1705.00378].

[21] A. Brandenburg, J. Schober, I. Rogachevskii, T. Kahniashvili, A. Boyarsky, J. Frohlich et al., *The turbulent chiral-magnetic cascade in the early universe*, *Astrophys. J. Lett.* **845** (2017) L21 [1707.03385].

[22] J. Schober, I. Rogachevskii, A. Brandenburg, A. Boyarsky, J. Fröhlich, O. Ruchayskiy et al., *Laminar and turbulent dynamos in chiral magnetohydrodynamics. II. Simulations*, *Astrophys. J.* **858** (2018) 124 [1711.09733].

[23] Y. Masada, K. Kotake, T. Takiwaki and N. Yamamoto, *Chiral magnetohydrodynamic turbulence in core-collapse supernovae*, *Phys. Rev. D* **98** (2018) 083018 [1805.10419].

[24] L. Del Zanna and N. Bucciatini, *Covariant and 3 + 1 equations for dynamo-chiral general relativistic magnetohydrodynamics*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 657 [1806.07114].

[25] J. Matsumoto, N. Yamamoto and D.-L. Yang, *Chiral plasma instability and inverse cascade from nonequilibrium left-handed neutrinos in core-collapse supernovae*, *Phys. Rev. D* **105** (2022) 123029 [2202.09205].

[26] J. Schober, I. Rogachevskii and A. Brandenburg, *Dynamo instabilities in plasmas with inhomogeneous chiral chemical potential*, *Phys. Rev. D* **105** (2022) 043507 [2107.13028].

[27] A. Brandenburg, K. Kamada, K. Mukaida, K. Schmitz and J. Schober, *Chiral magnetohydrodynamics with zero total chirality*, *Phys. Rev. D* **108** (2023) 063529 [2304.06612].

[28] A. Brandenburg, R. Sharma and T. Vachaspati, *Inverse cascading for initial magnetohydrodynamic turbulence spectra between Saffman and Batchelor*, *J. Plasma Phys.* **89** (2023) 905890606 [2307.04602].

[29] J. Schober, I. Rogachevskii and A. Brandenburg, *Chiral Anomaly and Dynamos from Inhomogeneous Chemical Potential Fluctuations*, *Phys. Rev. Lett.* **132** (2024) 065101 [2307.15118].

[30] A. Brandenburg, *Chirality in astrophysics*, in *CHIRAL MATTER: Proceedings of the Nobel Symposium 167*, pp. 15–35, World Scientific, 2023.

[31] K. Kamada, N. Yamamoto and D.-L. Yang, *Chiral effects in astrophysics and cosmology*, *Prog. Part. Nucl. Phys.* **129** (2023) 104016 [2207.09184].

[32] K. Hattori, K. Itakura and S. Ozaki, *Strong-field physics in QED and QCD: From fundamentals to applications*, *Prog. Part. Nucl. Phys.* **133** (2023) 104068 [2305.03865].

[33] S. Grozdanov, D. M. Hofman and N. Iqbal, *Generalized global symmetries and dissipative magnetohydrodynamics*, *Phys. Rev. D* **95** (2017) 096003 [1610.07392].

[34] K. Hattori, Y. Hirono, H.-U. Yee and Y. Yin, *MagnetoHydrodynamics with chiral anomaly: phases of collective excitations and instabilities*, *Phys. Rev.* **D100** (2019) 065023 [1711.08450].

[35] J. Armas and A. Jain, *Magnetohydrodynamics as superfluidity*, *Phys. Rev. Lett.* **122** (2019) 141603 [1808.01939].

[36] M. Hongo and K. Hattori, *Revisiting relativistic magnetohydrodynamics from quantum electrodynamics*, *JHEP* **02** (2021) 011 [2005.10239].

[37] K. Hattori, M. Hongo and X.-G. Huang, *New Developments in Relativistic Magnetohydrodynamics*, *Symmetry* **14** (2022) 1851 [2207.12794].

[38] D. Gaiotto, A. Kapustin, N. Seiberg and B. Willett, *Generalized Global Symmetries*, *JHEP* **02** (2015) 172 [1412.5148].

[39] W. A. Hiscock and L. Lindblom, *Stability and causality in dissipative relativistic fluids*, *Annals Phys.* **151** (1983) 466.

[40] W. A. Hiscock and L. Lindblom, *Generic instabilities in first-order dissipative relativistic fluid theories*, *Phys. Rev. D* **31** (1985) 725.

[41] W. A. Hiscock and L. Lindblom, *Linear plane waves in dissipative relativistic fluids*, *Phys. Rev. D* **35** (1987) 3723.

[42] J. Hernandez and P. Kovtun, *Relativistic magnetohydrodynamics*, *JHEP* **05** (2017) 001 [1703.08757].

[43] R. Biswas, A. Dash, N. Haque, S. Pu and V. Roy, *Causality and stability in relativistic viscous non-resistive magneto-fluid dynamics*, *JHEP* **10** (2020) 171 [2007.05431].

[44] J. Armas and F. Camilloni, *A stable and causal model of magnetohydrodynamics*, *JCAP* **10** (2022) 039 [2201.06847].

[45] W. Israel, *Nonstationary irreversible thermodynamics: A Causal relativistic theory*, *Annals Phys.* **100** (1976) 310.

[46] W. Israel, *The Dynamics of Polarization*, *Gen. Rel. Grav.* **9** (1978) 451.

[47] W. Israel and J. M. Stewart, *Transient relativistic thermodynamics and kinetic theory*, *Annals Phys.* **118** (1979) 341.

[48] G. S. Denicol, *Kinetic foundations of relativistic dissipative fluid dynamics*, *J. Phys. G* **41** (2014) 124004.

[49] G. S. Denicol, T. Kodama, T. Koide and P. Mota, *Stability and Causality in relativistic dissipative hydrodynamics*, *J. Phys. G* **35** (2008) 115102 [0807.3120].

[50] M. P. Heller, A. Serantes, M. Spaliński and B. Withers, *Rigorous Bounds on Transport from Causality*, *Phys. Rev. Lett.* **130** (2023) 261601 [2212.07434].

[51] L. Gavassino, *Can We Make Sense of Dissipation without Causality?*, *Phys. Rev. X* **12** (2022) 041001 [2111.05254].

[52] L. Gavassino, *Bounds on transport from hydrodynamic stability*, *Phys. Lett. B* **840** (2023) 137854 [2301.06651].

[53] F. S. Bemfica, M. M. Disconzi and J. Noronha, *Causality and existence of solutions of relativistic viscous fluid dynamics with gravity*, *Phys. Rev. D* **98** (2018) 104064 [1708.06255].

[54] F. S. Bemfica, M. M. Disconzi and J. Noronha, *Nonlinear Causality of General First-Order Relativistic Viscous Hydrodynamics*, *Phys. Rev. D* **100** (2019) 104020 [1907.12695].

[55] F. S. Bemfica, M. M. Disconzi and J. Noronha, *First-Order General-Relativistic Viscous Fluid Dynamics*, *Phys. Rev. X* **12** (2022) 021044 [2009.11388].

[56] F. S. Bemfica, M. M. Disconzi, V. Hoang, J. Noronha and M. Radosz, *Nonlinear Constraints on Relativistic Fluids Far from Equilibrium*, *Phys. Rev. Lett.* **126** (2021) 222301 [2005.11632].

[57] P. Kovtun, *First-order relativistic hydrodynamics is stable*, *JHEP* **10** (2019) 034 [1907.08191].

[58] R. E. Hoult and P. Kovtun, *Stable and causal relativistic Navier-Stokes equations*, *JHEP* **06** (2020) 067 [2004.04102].

[59] D.-L. Wang and S. Pu, *Stability and causality criteria in linear mode analysis: stability means causality*, 2309.11708.

[60] C. V. Brito and G. S. Denicol, *Linear stability of Israel-Stewart theory in the presence of net-charge diffusion*, *Phys. Rev. D* **102** (2020) 116009 [2007.16141].

[61] I. Cordeiro, E. Speranza, K. Ingles, F. S. Bemfica and J. Noronha, *Causality Bounds on Dissipative General-Relativistic Magnetohydrodynamics*, 2312.09970.

[62] E. Speranza, F. S. Bemfica, M. M. Disconzi and J. Noronha, *Challenges in solving chiral hydrodynamics*, *Phys. Rev. D* **107** (2023) 054029 [2104.02110].

[63] N. Abboud, E. Speranza and J. Noronha, *Causal and stable first-order chiral hydrodynamics*, 2308.02928.

[64] A. Daher, A. Das and R. Ryblewski, *Stability studies of first-order spin-hydrodynamic frameworks*, *Phys. Rev. D* **107** (2023) 054043 [2209.10460].

[65] G. Sarwar, M. Hasanujjaman, J. R. Bhatt, H. Mishra and J.-e. Alam, *Causality and stability of relativistic spin hydrodynamics*, *Phys. Rev. D* **107** (2023) 054031 [2209.08652].

[66] N. Weickgenannt, *Linearly stable and causal relativistic first-order spin hydrodynamics*, *Phys. Rev. D* **108** (2023) 076011 [2307.13561].

[67] X.-Q. Xie, D.-L. Wang, C. Yang and S. Pu, *Causality and stability analysis for the minimal causal spin hydrodynamics*, 2306.13880.

[68] Z. Fang, K. Hattori and J. Hu, *In preparation.*, .

[69] J. D. Jackson, *Classical Electrodynamics*. Wiley, 1999.

[70] L. Gavassino, M. Antonelli and B. Haskell, *Thermodynamic Stability Implies Causality*, *Phys. Rev. Lett.* **128** (2022) 010606 [2105.14621].

[71] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. 1995.

[72] L. Gavassino, M. M. Disconzi and J. Noronha, *Dispersion relations alone cannot guarantee causality*, 2307.05987.

[73] K. Izumi, *Causal Structures in Gauss-Bonnet gravity*, *Phys. Rev. D* **90** (2014) 044037 [1406.0677].

[74] G. S. Denicol, X.-G. Huang, E. Molnár, G. M. Monteiro, H. Niemi, J. Noronha et al., *Nonresistive dissipative magnetohydrodynamics from the Boltzmann equation in the 14-moment approximation*, *Phys. Rev. D* **98** (2018) 076009 [1804.05210].

[75] G. S. Denicol, E. Molnár, H. Niemi and D. H. Rischke, *Resistive dissipative magnetohydrodynamics from the Boltzmann-Vlasov equation*, *Phys. Rev. D* **99** (2019) 056017 [1902.01699].

[76] K. Hattori and D. Satow, *Electrical Conductivity of Quark-Gluon Plasma in Strong Magnetic Fields*, *Phys. Rev.* **D94** (2016) 114032 [1610.06818].

[77] K. Hattori, S. Li, D. Satow and H.-U. Yee, *Longitudinal Conductivity in Strong Magnetic Field in Perturbative QCD: Complete Leading Order*, *Phys. Rev.* **D95** (2017) 076008 [1610.06839].

[78] K. Hattori, X.-G. Huang, D. H. Rischke and D. Satow, *Bulk Viscosity of Quark-Gluon Plasma in Strong Magnetic Fields*, *Phys. Rev.* **D96** (2017) 094009 [1708.00515].

[79] S. Li and H.-U. Yee, *Shear viscosity of the quark-gluon plasma in a weak magnetic field in perturbative QCD: Leading log*, *Phys. Rev. D* **97** (2018) 056024 [1707.00795].

[80] K. Fukushima and Y. Hidaka, *Electric conductivity of hot and dense quark matter in a magnetic field with Landau level resummation via kinetic equations*, *Phys. Rev. Lett.* **120** (2018) 162301 [1711.01472].

[81] M. Kurian and V. Chandra, *Bulk viscosity of a hot QCD medium in a strong magnetic field within the relaxation-time approximation*, *Phys. Rev. D* **97** (2018) 116008 [1802.07904].

[82] W. Li, S. Lin and J. Mei, *Conductivities of magnetic quark-gluon plasma at strong coupling*, *Phys. Rev. D* **98** (2018) 114014 [1809.02178].

[83] K. Fukushima and Y. Hidaka, *Resummation for the Field-theoretical Derivation of the Negative Magnetoresistance*, *JHEP* **04** (2020) 162 [1906.02683].

[84] N. Astrakhantsev, V. V. Braguta, M. D'Elia, A. Y. Kotov, A. A. Nikolaev and F. Sanfilippo, *Lattice study of the electromagnetic conductivity of the quark-gluon plasma in an external magnetic field*, *Phys. Rev. D* **102** (2020) 054516 [1910.08516].

[85] K. Fukushima and A. Okutsu, *Electric conductivity with the magnetic field and the chiral anomaly in a holographic QCD model*, *Phys. Rev. D* **105** (2022) 054016 [2106.07968].

[86] H.-H. Peng, X.-L. Sheng, S. Pu and Q. Wang, *Electric and magnetic conductivities in magnetized fermion systems*, *Phys. Rev. D* **107** (2023) 116006 [2304.00519].

arXiv:2402.19440v1 [hep-ph] 29 Feb 2024

# Investigation of full-charm and full-bottom pentaquark states

K. Azizi*
*Department of Physics, University of Tehran, North Karegar Avenue, Tehran 14395-547, Iran*
*Department of Physics, Doğuş University, Dudullu-Ümraniye, 34775 Istanbul, Turkey and*
*School of Particles and Accelerators, Institute for Research in Fundamental Sciences (IPM) P.O. Box 19395-5531, Tehran, Iran*

Y. Sarac†
*Electrical and Electronics Engineering Department, Atilim University, 06836 Ankara, Turkey*

H. Sundu‡
*Department of Physics Engineering, Istanbul Medeniyet University, 34700 Istanbul, Turkey*
(Dated: March 1, 2024)

The continuous advancement of experimental techniques and investigations has led to observations of various exotic states in particle physics. Each addition to this family of states not only raises expectations for future discoveries but also focuses attention on such potential new states. Building upon this motivation and inspired by recent observations of various traditional and exotic particles containing an increased number of heavy quarks, our study provides a spectroscopic search for potential pentaquark states with spin-parity $\frac{3}{2}^-$ and composed entirely of charm or bottom quarks. We predict the masses for full-charm and full-bottom pentaquark states as $m = 7628 \pm 112$ MeV and $m = 21982 \pm 144$ MeV, respectively. We also compute the current couplings of these states to vacuum, which are main inputs in investigations of their various possible decays.

## I. INTRODUCTION

Since the proposal of the quark model, hadrons with non-conventional structures, which do not fit the conventional baryons composed of three quarks (antiquarks) and mesons composed of a quark and an antiquark, have been subjects of interests. The theory of strong interaction does not rule out the existence of such states, and this has attracted interest in these states. They were investigated extensively in both theory and experiments. Finally, the first evidence came out with the observation of $X(3872)$ state in 2003 [1]. And following this observation, many other such exotic state candidates were observed [2–10] and listed in Particle Data Group (PDG) [11]. These observations were also followed by many theoretical investigations trying to explain their internal structures, which still have ambiguity and need to be clearly identified with more scrutiny. It is evident that we will come across with other such possible exotic states in the future. This expectation necessitates their examinations in detail via different approaches to provide an understanding of their substructure and properties and provide feedback for future investigations. Besides, these states help to deepen our understanding of the dynamics of the strong interaction.

The pentaquark states are among these non-conventional states with their first observation reported in 2015 by the LHCb collaboration [3]. The investigation of the $\Lambda_b^0 \to J/\psi p K^-$ process resulted in two pentaquark states in the $J/\psi p$ invariant mass spectrum with the following resonance parameters [3]: $m_{P_c(4380)^+} = 4380 \pm 8 \pm 29$ MeV, $\Gamma_{P_c(4380)^+} = 205 \pm 18 \pm 86$ MeV and $m_{P_c(4450)^+} = 4449.8 \pm 1.7 \pm 2.5$ MeV, $\Gamma_{P_c(4450)^+} = 39 \pm 5 \pm 19$ MeV. This observation was supported by a full amplitude analysis for $\Lambda_b^0 \to J/\psi p \pi^-$ decays [4] in 2016. In 2019, using updated data, a new pentaquark state, $P_c(4312)^+$, was reported with $m_{P_c(4312)^+} = 4311.9 \pm 0.7^{+6.8}_{-0.6}$ MeV and $\Gamma_{P_c(4312)^+} = 9.8 \pm 2.7^{+3.7}_{-4.5}$ MeV by the LHCb collaboration and analyses revealed two narrow overlapping peaks for the previously observed peak of the $P_c(4450)^-$ state with masses and widths: $m_{P_c(4440)^+} = 4440.3 \pm 1.3^{+4.1}_{-4.7}$ MeV, $\Gamma_{P_c(4440)^+} = 20.6 \pm 4.9^{+8.7}_{-10.1}$ MeV and $m_{P_c(4457)^+} = 4457.3 \pm 0.6^{+4.1}_{-1.7}$ MeV, $\Gamma_{P_c(4457)^+} = 6.4 \pm 2.0^{+5.7}_{-1.9}$ MeV [5]. In the recent investigations, new states with the strange quark were also added to this family. The LHCb collaboration reported the $P_{cs}(4459)^0$ state through the investigation of $J\psi\Lambda$ invariant mass distribution in $\Xi_b^- \to J/\psi K^- \Lambda$ decays [9]. The mass and width for the $P_{cs}(4459)^0$ were given as $m = 4458.8 \pm 2.9^{+4.7}_{-1.1}$ MeV, and $\Gamma = 17.3 \pm 6.5^{+8.0}_{-5.7}$ MeV [9]. $P_{cs}(4338)$ state was reported with the mass $4338.2 \pm 0.7 \pm 0.4$ MeV and the width $7.0 \pm 1.2 \pm 1.3$ MeV from the amplitude analyses of $B^- \to J/\psi \Lambda \bar{p}$ [10].

Following the observations of the above pentaquark states, theoretical researches chasing the purpose of identifying their various properties have focused on these states. The sub-structures and quantum numbers of these observed

* kazem.azizi@ut.ac.ir; Corresponding author
† yasemin.sarac@atilim.edu.tr
‡ hayriyesundu.pamuk@medeniyet.edu.tr

pentaquark states still have uncertainty and this makes them attractive theoretically. Their investigations provide information not only about their nature and substructure but also for future experiments searching for new such states. Their scrutiny also supports and contributes to improving our understanding of the nonperturbative domain of quantum chromodynamics (QCD) with their distinct quark substructures compared to conventional hadrons with three quarks/antiquarks or a quark and an antiquark. With all these issues, these states were investigated thoroughly with the application of different approaches assigning them different substructures such as diquark-diquark-antiquark [12–26] and diquark-triquark [24, 27] models, and meson baryon molecular states [28–45]. Topological soliton model [46] and a variant of the D4-D8 model [47] were applied to search their properties, and in Refs [48–52] their being kinematical effects were taken into account. With the possibility for observation of new pentaquark states with quark content different from the observed ones, various candidate pentaquark states were also studied in the literature [53–67].

These observations and the advances in experimental facilities and techniques indicate the possibility of observing more exotic states in the future. With this expectation, these states with different possible structures and quark contents have been studied using various models [53–83]. The observed traditional baryons with double valence heavy quarks or exotic states with full heavy quarks in the recent years [84–90] enhance the expectation for future observation of such pentaquark states with two, three, or more heavy valence quarks in their quark substructures. This expectation motivates us to study such types of systems, and to this end, in this work, our focus is to study the pentaquark stets with full charm or full bottom quark contents. With the same motivation, the pentaquark states with full charm or bottom compositions were also studied in Refs. [64, 91–97]. However, to enhance our knowledge, better understand these states, and give support for possible future experiments, we need more analysis to explain their nature and possible structures. Furthermore, such analysis helps improve our understanding of nonperturbative regime of the quantum chromodynamics and refines our knowledge about the dynamics of heavy quarks. To investigate these states, we apply a method that has revealed its success so far with its predictions consistent with the experimental observations, namely the QCD sum rule method [98–100]. Using this method, we aim to calculate the masses and current coupling constants for considered states with spin-parity $\frac{3}{2}^-$ with a proper choice of interpolating current defining these states and composed of heavy quark fields. The information obtained about such states, which have the potential to be observed and investigated in future experiments, may provide support to conduct these experiments and help to explore and understand their properties.

The outline of the paper is as follows: In Section II, the QCD sum rules giving the masses and current coupling constants of the considered states were obtained. Section III gives the numerical analyses of the QCD sum rule results attained. The last section is devoted to the summary and conclusion.

## II. THE QCD SUM RULES FOR THE MASSES OF FULL-CHARM AND FULL-BOTTOM PENTAQUARKS

This section provides the QCD sum rule calculations that give the masses and corresponding current coupling constants of the considered full-charm and full-bottom pentaquarks. The mass calculations with the QCD sum rule method start with a two-point correlation function given as

$$\Pi_{\mu\nu}(p) = i \int d^4x e^{ip\cdot x} \langle 0|\mathcal{T}\{\eta_\mu^{P_Q}(x)\bar{\eta}_\nu^{P_Q}(0)\}|0\rangle, \tag{1}$$

where $\mathcal{T}$ stands for time ordering operator and $\eta_\mu^{P_Q}$ represents the full-heavy pentaquark currents either composed of completely charm quarks or completely bottom ones with the following form:

$$\eta_\mu^{P_Q} = [\epsilon^{ijk} Q_i^T C\gamma_\mu Q_j Q_k][\bar{Q}_l i\gamma_5 Q_l]. \tag{2}$$

The $Q$ in Eq. (2) is either charm quark or bottom quark field, $T$ represents transpose, $i,\ j,\ k,\ l$ are the color indices, and $C$ is the charge conjugation operator. We use this interpolating current to investigate both the fully-charmed pentaquark state using all its heavy quark fields as $c$ quark fields and the fully-bottom pentaquark state using them as all $b$ quark fields.

In the calculation of Eq. (1), two paths are followed; in the first path, it is calculated in terms of hadronic degrees of freedom. Since the result of the calculation contains the mass of the hadron and the current coupling constant, this path is called the hadronic side. On the other hand, the same equation is also computed in terms of QCD degrees of freedom, that is, in terms of the masses of the included quarks, quark-gluon condensates, and QCD coupling constant. Therefore, this path is called the QCD side of the calculation. After following these two paths to study the same correlation function, the obtained results are matched considering the coefficients of the same Lorentz structure on each part using a dispersion relation to obtain the QCD sum rules of the required quantities, masses, and current

coupling constants in this work. The obtained results also take contributions from higher states and continuum, which need to be suppressed to obtain the required physical parameters. Borel transformations and continuum subtractions are applied to both sides to provide this.

On the hadronic side, the interpolating currents are treated as annihilation or creation operators of the considered hadrons. A complete set of hadronic states carrying the same quantum numbers as the interpolating current is inserted inside the correlation function between the interpolating currents, and the integral over four-$x$ is performed. With these operations, the result becomes

$$\Pi^{\rm Had}_{\mu\nu}(p) = \frac{\langle 0|J_\mu|P_Q(p,s)\rangle\langle P_Q(p,s)|\bar{J}_\nu|0\rangle}{m^2_{P_Q}-p^2}+\cdots . \tag{3}$$

In Eq. (3) the $\cdots$ is used for the contribution of higher states and continuum, $|P_Q(p,s)\rangle$ is one particle satate with momentum $p$ and spin $s$, and the matrix elements are defined in terms of the current coupling constants, $\lambda_{P_Q}$, and Rarita-Schwinger spinor, $u_\mu(p,s)$, as

$$\langle 0|J_\mu|P_Q(p,s)\rangle \;=\; \lambda_{P_Q}u_\mu(p,s). \tag{4}$$

Though the states analyzed here are spin-$\frac{3}{2}$ states, the current used in the calculation also couples to the spin-$\frac{1}{2}$ state. Therefore, to isolate the contribution to the analyses from the spin-$\frac{3}{2}$ state and remove that from spin-$\frac{1}{2}$ state, we need to consider the matrix element entering the calculation relevant for the spin-$\frac{1}{2}$ states as

$$\langle 0|J_\mu|\frac{1}{2}(p,s)\rangle = A_{\frac{1}{2}^+}(\gamma_\mu+\frac{4p_\mu}{m_{\frac{1}{2}^+}})\gamma_5 u(q,s), \tag{5}$$

which indicates that the results getting contributions from these states are proportional to $\gamma_\mu$ or $p_\mu$. According to these, if we choose a proper Lorentz structure that does not contain $\gamma_\mu$ or $p_\mu$, we separate the part of the results giving contribution only to spin-$\frac{3}{2}$ state. To this end, we take into account the Lorentz structure $g_{\mu\nu}$ in our computation. After using the matrix element given in Eq. (4) inside the Eq. (3), we use the following summation over spin:

$$\sum_s u_\mu(p,s)\bar{u}_\nu(p,s) \;=\; -(\not{p}+m)\Big[g_{\mu\nu}-\frac{1}{3}\gamma_\mu\gamma_\nu-\frac{2p_\mu p_\nu}{3m^2}+\frac{p_\mu\gamma_\nu-p_\nu\gamma_\mu}{3m}\Big], \tag{6}$$

and obtain the hadronic side as

$$\Pi^{\rm Had}_{\mu\nu}(p) \;=\; \frac{{\lambda_{P_Q}}^2}{p^2-m^2_{P_Q}}(\not{p}+m_{P_Q})\Big[g_{\mu\nu}-\frac{1}{3}\gamma_\mu\gamma_\nu-\frac{2p_\mu p_\nu}{3m^2_{P_Q}}+\frac{p_\mu\gamma_\nu-p_\nu\gamma_\mu}{3m_{P_Q}}\Big]+\cdots . \tag{7}$$

As stated, to avoid the pollution of the spin-$\frac{1}{2}$ state, we work with the coefficient of the structure $g_{\mu\nu}$, which leads us to

$$\Pi^{\rm Had}_{\mu\nu}(p) \;=\; \frac{\lambda^2_{P_Q}}{p^2-m^2_{P_Q}}m_{P_Q}g_{\mu\nu}+\cdots , \tag{8}$$

where the contributions of other structures are represented by $\cdots$. After the Borel transformation with respect to $-p^2$ this result turns into

$$\Pi^{\rm Had}_{\mu\nu}(p) \;=\; \lambda^2_{P_Q}m_{P_Q}e^{-\frac{m^2_{P_Q}}{M^2}}g_{\mu\nu}+\cdots , \tag{9}$$

where $M^2$ is the Borel parameter arising from the Borel transformation.

After getting the hadronic side, we turn our attention to the QCD side of the calculation and use the same correlation function, Eq. (1), with the explicit form of the interpolating current given in terms of the quark fields. Using Wick's theorem, we obtain the possible contractions between the quark fields and attain the result in terms of the heavy

quark propagators as:

$$\begin{aligned}
\Pi_{\mu\nu}^{\rm QCD}(p) &= i\int d^4x e^{ip\cdot x} 2\epsilon_{abc}\epsilon_{a'b'c'}\Big\{{\rm Tr}[\gamma_\nu \tilde{S}_Q^{ba'}(x)\gamma_\mu S_Q^{ab'}(x)]{\rm Tr}[i\gamma_5 S_Q^{ee'}(x)i\gamma_5 S_Q^{e'e}(-x)]S_Q^{cc'}(x)\\
&- {\rm Tr}[\gamma_\nu \tilde{S}_Q^{ba'}(x)\gamma_\mu S_Q^{ab'}(x)]S_Q^{ce'}(x)i\gamma_5 S_Q^{e'e}(-x)i\gamma_5 S_Q^{ec'}(x)\\
&- 2{\rm Tr}[i\gamma_5 S_Q^{ee'}(x)i\gamma_5 S_Q^{e'e}(-x)]S_Q^{ca'}(x)\gamma_\nu \tilde{S}_Q^{ab'}(x)\gamma_\mu S_Q^{bc'}(x)\\
&+ 2S_Q^{ce'}(x)i\gamma_5 S_Q^{e'e}(-x)i\gamma_5 S_Q^{ea'}(x)\gamma_\nu \tilde{S}_Q^{ab'}(x)\gamma_\mu S_Q^{bc'}(x)\\
&+ 2S_Q^{ca'}(x)\gamma_\nu \tilde{S}_Q^{ab}(x)\gamma_\mu S_Q^{be'}(x)i\gamma_5 S_Q^{e'e}(-x)i\gamma_5 S_Q^{ec'}(x)\\
&- 2{\rm Tr}[S_Q^{ea'}(x)\gamma_\nu \tilde{S}_Q^{ab'}(x)\gamma_\mu S_Q^{be'}(x)i\gamma_5 S_Q^{e'e}(-x)i\gamma_5]S_Q^{cc'}(x)\\
&+ 2S_Q^{cb'}(x)\gamma_\nu \tilde{S}_Q^{ea'}(x)i\gamma_5 \tilde{S}_Q^{e'e}(-x)i\gamma_5 \tilde{S}_Q^{be'}(x)\gamma_\mu S_Q^{ac'}(x)\Big\},
\end{aligned} \tag{10}$$

where $\tilde{S}_Q^{ab}(x) = CS_Q^{abT}(x)C$ and $S_Q(x)$ is the heavy quark propagator given as

$$\begin{aligned}
S_{Q,ab}(x) &= \frac{i}{(2\pi)^4}\int d^4k e^{-ik\cdot x}\left\{\frac{\delta_{ab}}{\not{k}-m_Q} - \frac{g_s G_{ab}^{\alpha\beta}}{4}\frac{\sigma_{\alpha\beta}(\not{k}+m_Q)+(\not{k}+m_Q)\sigma_{\alpha\beta}}{(k^2-m_Q^2)^2}\right.\\
&\left. +\frac{\pi^2}{3}\langle\frac{\alpha_s GG}{\pi}\rangle\delta_{ij}m_Q\frac{k^2+m_Q\not{k}}{(k^2-m_Q^2)^4}+\cdots\right\}.
\end{aligned} \tag{11}$$

Using this propagator in Eq. (10), and taking the integral over $x$ and applying the Borel transformation, we obtain the final results, which are long and not presented here, given in the following form:

$$\mathcal{B}\Pi^{\rm QCD}(s_0,M^2) = \int_{25m_Q^2}^{s_0} ds e^{-\frac{s}{M^2}}\rho(s), \tag{12}$$

with $\rho(s)$ being the result corresponding to the coefficient of the Lorentz structure $g_{\mu\nu}$ obtained in the QCD side of the calculation.

The match of the results of the coefficients of the Lorentz structure $g_{\mu\nu}$ attained in hadronic and QCD sides gives the QCD sum rule for the masses and current coupling constants as

$$\lambda_{P_Q}^2 m_{P_Q} e^{-\frac{m_{P_Q}^2}{M^2}} = \mathcal{B}\Pi^{\rm QCD}(s_0,M^2), \tag{13}$$

from which we obtain the mass by taking the derivative of the Eq. (13) with respect to $-\frac{1}{M^2}$ and dividing this new result by Eq. (13) itself. This leads us to the mass as

$$m_{P_Q}^2 = \frac{\frac{d}{d(-\frac{1}{M^2})}\mathcal{B}\Pi^{\rm QCD}(s_0,M^2)}{\mathcal{B}\Pi^{\rm QCD}(s_0,M^2)}, \tag{14}$$

and current coupling constant is obtained from Eq. (13)using the mass result as

$$\lambda_{P_Q}^2 = \frac{e^{\frac{m_{P_Q}^2}{M^2}}}{m_{P_Q}}\mathcal{B}\Pi^{\rm QCD}(s_0,M^2). \tag{15}$$

The next stage is to analyze these results with the proper input parameters and fix the corresponding numerical values for these quantities, as explained in the next section.

## III. NUMERICAL ANALYSES

To attain the numerical results from the QCD sum rules obtained for masses and current coupling constants, some input parameters are required. These input parameters are provided in Table I. Besides, we need two auxiliary parameters in each case. These auxiliary parameters are Borel masses $M^2$ and threshold parameters $s_0$ that enter the calculations by the applications of Borel transformation and continuum subtraction, respectively. In the analyses, the

| Parameters | Values |
|---|---|
| $m_c$ | $1.27 \pm 0.02$ GeV [11] |
| $m_b$ | $4.18^{+0.03}_{-0.02}$ GeV [11] |
| $\langle \frac{\alpha_s}{\pi} G^2 \rangle$ | $(0.012 \pm 0.004)$ GeV$^4$[101] |

TABLE I. Necessary input parameters for the numerical analyses.

working windows for these parameters are fixed following the standard criteria of the QCD sum rule application. To determine the upper limits of the Borel parameters, dominance of the pole over the continuum is demanded, and for the lover limits, the convergence of the OPE calculation is taken into account. Regarding these, the proper intervals for the Borel parameters are fixed as follows:

$$8.0 \text{ GeV}^2 \leq M^2 \leq 10.0 \text{ GeV}^2, \tag{16}$$

for full-charm $P_c$ pentaquark state and

$$25.0 \text{ GeV}^2 \leq M^2 \leq 35.0 \text{ GeV}^2, \tag{17}$$

for the full-bottom $P_b$ pentaquark state. Regarding the intervals of the threshold parameters, these parameters are not completely arbitrary and have connections to the energies of the possible excited states. For this reason, their values are fixed by conjecturing the probable energies that may correspond to the first excited states of the considered pentaquarks. Therefore, the following intervals are applied in the analyses:

$$60.0 \text{ GeV}^2 \leq s_0 \leq 64.0 \text{ GeV}^2 \tag{18}$$

for $P_c$ state and

$$485.0 \text{ GeV}^2 \leq s_0 \leq 500.0 \text{ GeV}^2 \tag{19}$$

for $P_b$ states. Together with these above criteria, another criterion for these auxiliary parameters is the stability of the results with the variation of these parameters in their working intervals. Therefore, after fixing the intervals of these auxiliary parameters, we check the dependence of our results on them and depict their behaviors with the figures 1-4 for both the full-charm and full-bottom pentaquark states. As is seen from these figures, these expectations are satisfied enough at the chosen working intervals.

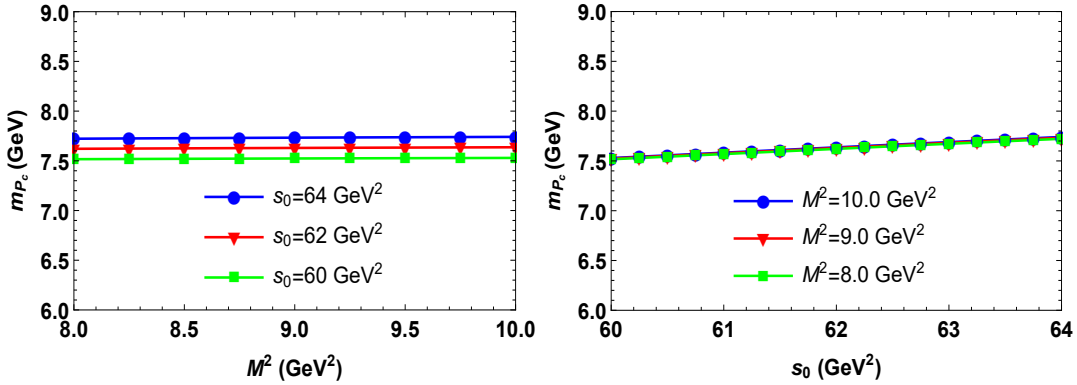



FIG. 1. **Left:** The variation of the mass of the $P_c$ state as a function of the Borel parameter $M^2$ at various values of the threshold parameter $s_0$. **Right:** The variation of the mass of the $P_c$ state as a function of the threshold parameter $s_0$ at various values of the Borel parameter $M^2$.

Having fixed the applicable intervals of the auxiliary parameters and by using all the inputs, the mass values and current coupling constants for these states can be reported. The results are presented in Table II. In the results, the errors arising from the errors of input parameters and uncertainties in the determination of the working intervals of the auxiliary parameters are also presented.

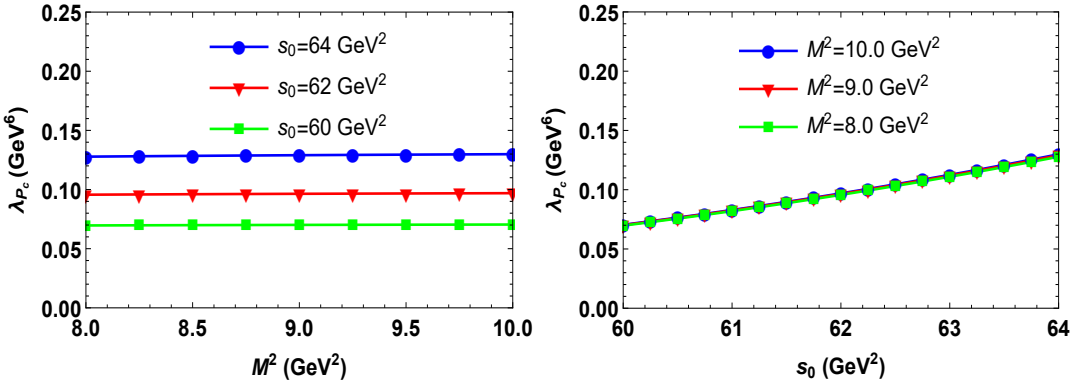



FIG. 2. **Left:** The variation of the current coupling constant $\lambda_{P_c}$ of the $P_c$ state as a function of the Borel parameter $M^2$ at various values of the threshold parameter $s_0$. **Right:** The variation of the current coupling constant $\lambda_{P_c}$ of the $P_c$ state as a function of the threshold parameter $s_0$ at various values of the Borel parameter $M^2$.

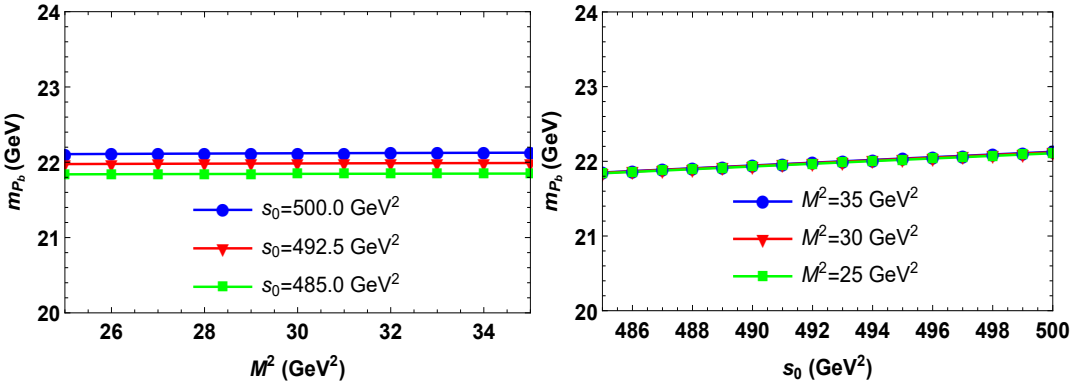



FIG. 3. **Left:** The variation of the mass of the $P_b$ state as a function of the Borel parameter $M^2$ at various values of the threshold parameter $s_0$. **Right:** The variation of the mass of the $P_b$ state as a function of the threshold parameter $s_0$ at various values of the Borel parameter $M^2$.

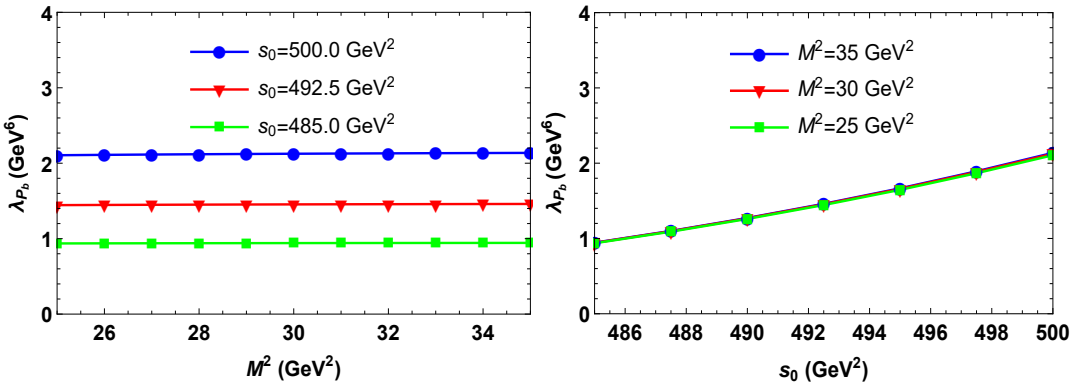



FIG. 4. **Left:** The variation of the current coupling constant $\lambda_{P_b}$ of the $P_b$ state as a function of the Borel parameter $M^2$ at various values of the threshold parameter $s_0$. **Right:** The variation of the current coupling constant $\lambda_{P_b}$ of the $P_b$ state as a function of the threshold parameter $s_0$ at various values of the Borel parameter $M^2$.

| State | $m$ (MeV) | $\lambda$ (GeV)$^6$ |
|---|---|---|
| $P_c$ | $7628 \pm 112$ | $(9.81 \pm 2.97) \times 10^{-2}$ |
| $P_b$ | $21982 \pm 144$ | $1.51 \pm 0.60$ |

TABLE II. : Masses and current coupling constants for $P_c$ and $P_b$ states.

## IV. SUMMARY AND CONCLUSION

The number of members of exotic states is increasing day after day as a result of improvements in experimental techniques and analyses. Each new member added to this family increases the expectations of future observations of such states and collects attention over these possible new states' investigations. With the motivations brought by the observed states, the ones with quark content different from the observed ones have been studied using various theoretical models to provide inputs or comparison grounds for future experimental findings. With the same motivation and the motivation brought by new observations of conventional or exotic states containing more numbers of heavy quarks, in the present work, we searched for possible pentaquark states composed of full charm or full bottom quarks and carrying spin-parity quantum numbers $J^P = \frac{3}{2}^-$. We predicted the masses for these states as $m = 7628 \pm 112$ MeV for full-charm pentaquark and $m = 21982 \pm 144$ MeV for full-bottom pentaquark. In the literature, there are a few works investigating the full-heavy pentaquark states. The similar mass predictions in Ref. [64] are $m = 7.41^{+0.27}_{-0.31}$ GeV and $m = 21.60^{+0.73}_{-0.22}$ GeV for full-charm and full-bottom states, respectively. In Ref. [92], the prediction for the mass of the full-bottom pentaquark state with spin-parity $\frac{3}{2}^-$ is $m = 23748.2 \sim 23752.3$ MeV, on the other hand, they obtained no bound state for the charm counterpart of this state. The predictions in Ref. [93] are $m = 7863.6$ MeV and $m = 23774.8$ MeV for full-charm and full-bottom pentaquarks with spin-parity $J^P = \frac{3}{2}^-$, respectively. Beside these, in Ref [94], the analysis gave no stable full-heavy pentaquark system. The results of the present work are consistent with those of Ref. [64] within the errors for both types of pentaquarks, on the other hand, they are smaller than the predictions presented in Refs. [92, 93]. As is seen, there is a need to conduct more investigations about these states to gain a deeper understanding of them.

Besides, since they are necessary for the calculation of the form factors that are used in the investigations of possible decay mechanisms of these states, we calculated the corresponding current coupling constants for these states.

It is evident that further analyses of these states are necessary to gain a better understanding of their structure and properties. These results may be supported by other works through comparisons of their results or used in further investigations delving into the interaction mechanisms of these states.

## ACKNOWLEDGEMENTS

K. Azizi is thankful to Iran Science Elites Federation (Saramadan) for the partial financial support provided under the grant number ISEF/M/99171.

[1] S. K. Choi *et al.* [Belle Collaboration], Phys. Rev. Lett. **91**, 262001 (2003) [hep-ex/0309032].
[2] P. A. Zyla *et al.* [Particle Data Group], PTEP **2020**, no.8, 083C01 (2020).
[3] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **115**, 072001 (2015) [arXiv:1507.03414 [hep-ex]].
[4] R. Aaij *et al.* [LHCb], Phys. Rev. Lett. **117**, no.8, 082003 (2016) [arXiv:1606.06999 [hep-ex]].
[5] R. Aaij *et al.* [LHCb], Phys. Rev. Lett. **122**, no.22, 222001 (2019) [arXiv:1904.03947 [hep-ex]].
[6] R. Aaij *et al.* [LHCb], Nature Commun. **13**, no.1, 3351 (2022) [arXiv:2109.01056 [hep-ex]].
[7] R. Aaij *et al.* [LHCb], Nature Phys. **18**, no.7, 751-754 (2022) [arXiv:2109.01038 [hep-ex]].
[8] R. Aaij *et al.* [LHCb], Phys. Rev. Lett. **128**, no.6, 062001 (2022) [arXiv:2108.04720 [hep-ex]].
[9] R. Aaij *et al.* [LHCb], Sci. Bull. **66**, 1278-1287 (2021) [arXiv:2012.10380 [hep-ex]].
[10] R. Aaij *et al.* [LHCb], Phys. Rev. Lett. **131**, no.3, 031901 (2023) [arXiv:2210.10346 [hep-ex]].
[11] R. L. Workman *et al.* [Particle Data Group], PTEP **2022**, 083C01 (2022)
[12] R. F. Lebed, Phys. Lett. B **749**, 454 (2015) [arXiv:1507.05867 [hep-ph]].
[13] L. Maiani, A. D. Polosa and V. Riquer, Phys. Lett. B **749**, 289 (2015) [arXiv:1507.04980 [hep-ph]].
[14] G. N. Li, X. G. He and M. He, JHEP **12**, 128 (2015) [arXiv:1507.08252 [hep-ph]].
[15] Z. G. Wang and T. Huang, Eur. Phys. J. C **76**, no. 1, 43 (2016) [arXiv:1508.04189 [hep-ph]].
[16] V. V. Anisovich, M. A. Matveev, J. Nyiri, A. V. Sarantsev and A. N. Semenova, arXiv:1507.07652 [hep-ph].

[17] Z. G. Wang, Nucl. Phys. B **913**, 163-208 (2016) [arXiv:1512.04763 [hep-ph]].
[18] Z. G. Wang, Eur. Phys. J. C **76**, no. 2, 70 (2016) [arXiv:1508.01468 [hep-ph]].
[19] R. Ghosh, A. Bhattacharya and B. Chakrabarti, Phys. Part. Nucl. Lett. **14**, no.4, 550-552 (2017) [arXiv:1508.00356 [hep-ph]].
[20] Z. G. Wang, Eur. Phys. J. C **76**, no.3, 142 (2016) [arXiv:1509.06436 [hep-ph]].
[21] J. X. Zhang, Z. G. Wang and Z. Y. Di, Acta Phys. Polon. B **48**, 2013 (2017) [arXiv:1711.10728 [hep-ph]].
[22] Z. G. Wang, H. J. Wang and Q. Xin, Chin. Phys. C **45**, 063104 (2021) [arXiv:2005.00535 [hep-ph]].
[23] A. Ali, I. Ahmed, M. J. Aslam, A. Parkhomenko and A. Rehman, PoS **ICHEP2020**, 527 (2021) [arXiv:2012.07760 [hep-ph]].
[24] G. J. Wang, R. Chen, L. Ma, X. Liu and S. L. Zhu, Phys. Rev. D **94**, no. 9, 094018 (2016) [arXiv:1605.01337 [hep-ph]].
[25] Z. G. Wang, Int. J. Mod. Phys. A **36**, no.10, 2150071 (2021) [arXiv:2011.05102 [hep-ph]].
[26] Z. G. Wang, Int. J. Mod. Phys. A **35**, no.01, 2050003 (2020) [arXiv:1905.02892 [hep-ph]].
[27] R. Zhu and C. F. Qiao, Phys. Lett. B **756**, 259 (2016) [arXiv:1510.08693 [hep-ph]].
[28] H. X. Chen, W. Chen, X. Liu, T. G. Steele and S. L. Zhu, Phys. Rev. Lett. **115**, no. 17, 172001 (2015) [arXiv:1507.03717 [hep-ph]].
[29] R. Chen, X. Liu, X. Q. Li and S. L. Zhu, Phys. Rev. Lett. **115**, no. 13, 132002 (2015) [arXiv:1507.03704 [hep-ph]].
[30] J. He, Phys. Lett. B **753**, 547 (2016) [arXiv:1507.05200 [hep-ph]].
[31] L. Roca, J. Nieves and E. Oset, Phys. Rev. D **92**, no. 9, 094003 (2015) [arXiv:1507.04249 [hep-ph]].
[32] K. Azizi, Y. Sarac and H. Sundu, Phys. Rev. D **95**, no. 9, 094016 (2017) [arXiv:1612.07479 [hep-ph]].
[33] U. G. Meißner and J. A. Oller, Phys. Lett. B **751**, 59 (2015) [arXiv:1507.07478 [hep-ph]].
[34] K. Azizi, Y. Sarac and H. Sundu, Phys. Lett. B **782**, 694-701 (2018) [arXiv:1802.01384 [hep-ph]].
[35] H. X. Chen, Chin. Phys. C **46**, no.9, 093105 (2022) [arXiv:2011.07187 [hep-ph]].
[36] H. S. Li, F. Guo, Y. D. Lei and F. Gao, [arXiv:2401.14767 [hep-ph]].
[37] H. S. Li, [arXiv:2401.14759 [hep-ph]].
[38] B. Wang, K. Chen, L. Meng and S. L. Zhu, [arXiv:2312.13591 [hep-ph]].
[39] H. X. Chen, W. Chen, X. Liu and X. H. Liu, Eur. Phys. J. C **81**, no.5, 409 (2021) [arXiv:2011.01079 [hep-ph]].
[40] F. Z. Peng, M. J. Yan, M. Sánchez Sánchez and M. P. Valderrama, Eur. Phys. J. C **81**, no.7, 666 (2021) [arXiv:2011.01915 [hep-ph]].
[41] R. Chen, Phys. Rev. D **103**, no.5, 054007 (2021) [arXiv:2011.07214 [hep-ph]].
[42] Z. G. Wang and X. Wang, Chin. Phys. C **44**, 103102 (2020) [arXiv:1907.04582 [hep-ph]].
[43] Z. G. Wang and Q. Xin, Chin. Phys. C **45**, no.12, 123105 (2021) [arXiv:2103.08239 [hep-ph]].
[44] Z. G. Wang, Int. J. Mod. Phys. A **34**, no.19, 1950097 (2019) [arXiv:1806.10384 [hep-ph]].
[45] K. Azizi, Y. Sarac and H. Sundu, Chin. Phys. C **45**, no.5, 053103 (2021) [arXiv:2011.05828 [hep-ph]].
[46] N. N. Scoccola, D. O. Riska and M. Rho, Phys. Rev. D **92**, no.5, 051501 (2015) [arXiv:1508.01172 [hep-ph]].
[47] Y. Liu and I. Zahed, Phys. Rev. D **95**, no.11, 116012 (2017) [arXiv:1704.03412 [hep-ph]].
[48] F. K. Guo, U. G. Meißner, W. Wang and Z. Yang, Phys. Rev. D **92**, no.7, 071502 (2015) [arXiv:1507.04950 [hep-ph]].
[49] M. Mikhasenko, arXiv:1507.06552 [hep-ph].
[50] F. K. Guo, U. G. Meißner, J. Nieves and Z. Yang, Eur. Phys. J. A **52**, no.10, 318 (2016) [arXiv:1605.05113 [hep-ph]].
[51] M. Bayar, F. Aceti, F. K. Guo and E. Oset, Phys. Rev. D **94**, no.7, 074039 (2016) [arXiv:1609.04133 [hep-ph]].
[52] X. H. Liu, Q. Wang and Q. Zhao, Phys. Lett. B **757**, 231-236 (2016) [arXiv:1507.05359 [hep-ph]].
[53] X. Liu, H. Huang and J. Ping, Phys. Rev. C **102** (2020) no.2, 025204 [arXiv:2005.09646 [hep-ph]].
[54] A. Feijoo, V. K. Magas, A. Ramos and E. Oset, Eur. Phys. J. C **76**, no.8, 446 (2016) [arXiv:1512.08152 [hep-ph]].
[55] H. X. Chen, L. S. Geng, W. H. Liang, E. Oset, E. Wang and J. J. Xie, Phys. Rev. C **93**, no.6, 065203 (2016) [arXiv:1510.01803 [hep-ph]].
[56] J. X. Lu, E. Wang, J. J. Xie, L. S. Geng and E. Oset, Phys. Rev. D **93**, 094009 (2016) [arXiv:1601.00075 [hep-ph]].
[57] Y. Irie, M. Oka and S. Yasui, Phys. Rev. D **97**, no.3, 034006 (2018) [arXiv:1707.04544 [hep-ph]].
[58] Q. Zhang, B. R. He and J. L. Ping, [arXiv:2006.01042 [hep-ph]].
[59] R. Chen, J. He and X. Liu, Chin. Phys. C **41**, no.10, 103105 (2017) [arXiv:1609.03235 [hep-ph]].
[60] T. Gutsche and V. E. Lyubovitskij, Phys. Rev. D **100**, no.9, 094031 (2019) [arXiv:1910.03984 [hep-ph]].
[61] K. Azizi, Y. Sarac and H. Sundu, Phys. Rev. D **98**, no.5, 054002 (2018) [arXiv:1805.06734 [hep-ph]].
[62] K. Azizi, Y. Sarac and H. Sundu, Phys. Rev. D **96**, no.9, 094030 (2017) [arXiv:1707.01248 [hep-ph]].
[63] X. Cao, F. K. Guo, Y. T. Liang, J. J. Wu, J. J. Xie, Y. P. Xie, Z. Yang and B. S. Zou, Phys. Rev. D **101**, no.7, 074010 (2020) [arXiv:1912.12054 [hep-ph]].
[64] J. R. Zhang, Phys. Rev. D **103**, no.7, 074016 (2021) [arXiv:2011.04594 [hep-ph]].
[65] Y. P. Xie and V. P. Goncalves, Phys. Lett. B **814**, 136121 (2021) [arXiv:2012.07501 [hep-ph]].
[66] E. Y. Paryev, [arXiv:2007.01172 [nucl-th]].
[67] F. L. Wang, R. Chen and X. Liu, Phys. Rev. D **103**, no.3, 034014 (2021) [arXiv:2011.14296 [hep-ph]].
[68] F. B. Duan, Q. N. Wang, Z. Y. Yang, X. L. Chen and W. Chen, [arXiv:2401.10078 [hep-ph]].
[69] S. Y. Kong, J. T. Zhu, S. Chen and J. He, [arXiv:2402.02703 [hep-ph]].
[70] A. Sharma and A. Upadhyay, [arXiv:2401.02146 [hep-ph]].
[71] W. X. Zhang, C. L. Liu and D. Jia, [arXiv:2312.12770 [hep-ph]].
[72] Y. Yan, H. Huang, X. Zhu and J. Ping, [arXiv:2312.04977 [hep-ph]].
[73] X. Liu, Y. Tan, X. Chen, D. Chen, H. Huang and J. Ping, [arXiv:2312.04390 [hep-ph]].
[74] Z. Y. Wang, C. W. Xiao, Z. F. Sun and X. Liu, [arXiv:2311.17736 [hep-ph]].

[75] F. L. Wang and X. Liu, Phys. Rev. D **109**, no.1, 1 (2024) [arXiv:2311.13968 [hep-ph]].
[76] G. Yang, J. Ping and J. Segovia, [arXiv:2311.01044 [hep-ph]].
[77] X. Liu, Y. Tan, X. Chen, D. Chen, H. Huang and J. Ping, Phys. Rev. D **108**, no.9, 094008 (2023).
[78] E. Y. Paryev, Nucl. Phys. A **1042**, 122792 (2024) [arXiv:2310.04123 [hep-ph]].
[79] J. X. Lin, H. X. Chen, W. H. Liang, W. Y. Liu and D. Zhou, Eur. Phys. J. A **60**, no.1, 15 (2024) [arXiv:2308.01007 [hep-ph]].
[80] Q. Xin, X. S. Yang and Z. G. Wang, Int. J. Mod. Phys. A **38**, no.22n23, 2350123 (2023) [arXiv:2307.08926 [hep-ph]].
[81] R. Chen and Q. Huang, Phys. Rev. D **108**, no.5, 054011 (2023) [arXiv:2307.04168 [hep-ph]].
[82] X. Zhu, Y. Wu, H. Huang, J. Ping and Y. Yang, Universe **9**, no.6, 265 (2023)
[83] Y. Yan, Y. Wu, H. Huang, J. Ping and X. Zhu, Eur. Phys. J. C **83**, no.7, 610 (2023) [arXiv:2305.16609 [hep-ph]].
[84] R. Aaij *et al.* [LHCb], Phys. Rev. Lett. **119**, no.11, 112001 (2017) [arXiv:1707.01621 [hep-ex]].
[85] R. Aaij *et al.* [LHCb], Phys. Rev. Lett. **121**, no.16, 162002 (2018) [arXiv:1807.01919 [hep-ex]].
[86] R. Aaij *et al.* [LHCb], Sci. Bull. **65**, no.23, 1983-1993 (2020) [arXiv:2006.16957 [hep-ex]].
[87] E. Bouhova-Thacker [ATLAS], PoS **ICHEP2022**, 806 (2022)
[88] J. Zhang *et al.* [CMS], PoS **ICHEP2022**, 775 [arXiv:2212.00504 [hep-ex]].
[89] A. Hayrapetyan *et al.* [CMS], [arXiv:2306.07164 [hep-ex]].
[90] G. Aad *et al.* [ATLAS], Phys. Rev. Lett. **131**, no.15, 151902 (2023) [arXiv:2304.08962 [hep-ex]].
[91] Z. G. Wang, Nucl. Phys. B **973**, 115579 (2021) [arXiv:2104.12090 [hep-ph]].
[92] Y. Yan, Y. Wu, X. Hu, H. Huang and J. Ping, Phys. Rev. D **105**, no.1, 014027 (2022) [arXiv:2110.10853 [hep-ph]].
[93] H. T. An, K. Chen, Z. W. Liu and X. Liu, Phys. Rev. D **103**, no.7, 074006 (2021) [arXiv:2012.12459 [hep-ph]].
[94] H. T. An, S. Q. Luo, Z. W. Liu and X. Liu, Phys. Rev. D **105**, no.7, 074032 (2022) [arXiv:2203.03448 [hep-ph]].
[95] W. X. Zhang, H. T. An and D. Jia, Eur. Phys. J. C **83**, no.8, 727 (2023) [arXiv:2304.14876 [hep-ph]].
[96] G. Yang, J. Ping and J. Segovia, Phys. Rev. D **106**, no.1, 014005 (2022) [arXiv:2205.11548 [hep-ph]].
[97] Z. B. Liang, F. X. Liu and X. H. Zhong, [arXiv:2402.17974 [hep-ph]].
[98] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **147**, 385 (1979).
[99] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Nucl. Phys. B **147**, 448 (1979).
[100] B. L. Ioffe, Nucl. Phys. B **188**, 317 (1981) Erratum: [Nucl. Phys. B **191**, 591 (1981)].
[101] V. M. Belyaev and B. L. Ioffe, Sov. Phys. JETP **57**, 716 (1983) [Zh. Eksp. Teor. Fiz. **84**, 1236 (1983)].

KEK-TH-2592, DESY-24-003, RESCEU-2/24, YGHP-24-01

# Neutrino zeromodes on electroweak strings in light of topological insulators

arXiv:2402.19417v1 [hep-ph] 29 Feb 2024

**Minoru Eto,** [a,b,c] **Yu Hamada,** [d,e,b] **Ryusuke Jinno,** [f] **Muneto Nitta,** [g,b,c] **and Masatoshi Yamada** [h]

[a] *Department of Physics, Yamagata University, Kojirakawa-machi 1-4-12, Yamagata, Yamagata 990-8560, Japan*

[b] *Research and Education Center for Natural Sciences, Keio University, 4-1-1 Hiyoshi, Yokohama, Kanagawa 223-8521, Japan*

[c] *International Institute for Sustainability with Knotted Chiral Meta Matter(SKCM$^2$), Hiroshima University, 1-3-2 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8511, Japan*

[d] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*

[e] *KEK Theory Center, Tsukuba 305-0801, Japan*

[f] *Research Center for the Early Universe (RESCEU), University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-003, Japan*

[g] *Department of Physics, Keio University, 4-1-1 Hiyoshi, Kanagawa 223-8521, Japan*

[h] *Center for Theoretical Physics and College of Physics, Jilin University, Changchun 130012, China*

*E-mail:* meto@sci.kj.yamagata-u.ac.jp, yu.hamada@desy.de, ryusuke.jinno@resceu.s.u-tokyo.ac.jp, nitta@phys-h.keio.ac.jp, yamada@jlu.edu.cn

Abstract: We examine neutrino zeromode solutions on the electroweak $Z$-string and their effect on the stability of the string in the standard model and its extensions. We propose using topological invariants constructed from the momentum (and real) space topology of Green's functions, often used for investigating edge modes in condensed matter physics. We analyze the standard model and then examine type-I and type-II extensions of the neutrino sector as well as their hybrid. Based on this analysis, we also comment on proposals in the literature to stabilize the $Z$-string.

## Contents


**1 Introduction** **1**

**2 Review on electroweak string in SM** **4**
2.1 Model setup and electroweak string solutions 4
2.2 Stability of $Z$ string 6

**3 Number of zeromodes on vortices and its relation with momentum space topology** **14**
3.1 Topological invariant for the number of zeromodes 14
3.2 Dirac fermion on Abrikosov-Nielsen-Olesen vortex 17
3.3 Application 1: Type-II seesaw model 19
3.4 Application 2: Type-I seesaw model 23
3.5 Application 3: Type-I seesaw model with one $\nu_L$ and two $\nu_R$'s 27
3.6 Application 4: Hybrid of type-I and type-II seesaw models 28
3.7 Comment on the stability of $Z$-string 29

**4 Discussion and Conclusions** **31**

**A Notation** **32**
A.1 Representation of Dirac matrices 32

**B Bulk momentum space topology in neutrino sector** **33**
B.1 Topological invariant for symmetry-protected topological phases 33
B.2 A simple example: Free Dirac fermion 34
B.3 Type II seesaw model 35
B.4 Type I seesaw model 35
B.5 Hybrid of type I and II models 36


## 1 Introduction

Spontaneous symmetry breaking is a widely observed phenomenon in statistical physics and quantum field theory. Its notion is crucial for understanding critical phenomena such as superconductivity [1] in condensed matter physics and dynamical mass generation [2–6] in high energy physics. A basic point in spontaneous symmetry breaking is that the trivial vacuum (or symmetric phase) associated to a symmetry in a system becomes an unstable (tachyonic) or metastable state which indicates existence of a lower energy state (broken phase) not obeying the symmetry. Such a simple phase-transition picture in spontaneous symmetry breaking can be described by the Ginzburg-Landau theory [7] with the free

energy as a function of an order parameter associated to a symmetry. In the broken phase with non-vanishing order parameter value, the equation of motion (EOM) could entail non-trivial solutions depending on spacial coordinates in addition to the vacuum solutions. One of well-known solutions is an Abrikosov-Nielsen–Olesen (ANO) vortex/string solution [8, 9] in a $U(1)$ gauge theory coupled with one complex scalar field, called the Abelian-Higgs model.[1] In addition, there are also other objects such as monopoles, domain walls, and Skyrmions. These solutions arise due to the existence of the non-trivial homotopy group characterizing maps between the order-parameter space (vacuum manifold) and the real space. Such topologically non-trivial solutions are called topological defects or solitons. In the context of particle cosmology, phase transitions result in creation of topological defects through the Kibble-Zurek mechanism [11, 12]. These objects may give great impacts on the thermal history of the universe.

If topological defects couple with fermionic fields, they may lead to non-trivial phenomena. In fact, it is well known that fermionic modes sometimes are trapped inside the defects. For example, fermionic zeromodes localized on vortices in topological superconductors [13–18], and Majorana fermion zeromodes on superfluid vortices [19] in neutron stars [20–22] and on non-Abelian vortices [23–28] in quantum chromodynamics (QCD) in high density region (quark matter) [26, 29–32]. Whether the zeromodes exist on the defects or not reflects different phases of matter that cannot be captured by the conventional Ginzburg-Landau theories. These phases are called topological phases and are classified by the homotopy groups of the *momentum space* of the fermions. The relation between the zeromodes and the momentum space topology has been widely studied utilizing Green's function [18, 28, 33–35] and (one-particle) Hamiltonian [36–40]. Thus, in such systems, it is important to consider the topology of both of real space and momentum space.

In contrast to these condensed matter, nuclear and quark matter systems, all of which are finite density matter and are therefore *non-relativistic*, vortex strings in *relativistic* field theories also contain fermion zeromodes [41–45]. Their qualitative properties are understood by the index theorem counting the number of zeromodes [44, 46, 47] and the anomaly inflow [45, 48]. In addition, Bogomol'nyi-Prasad-Sommerfield (BPS) vortices in supersymmetric field theories are always associated with fermion zeromodes as superpartners of translational (bosonic) zeromodes [49–51]. In the context of cosmology, fermion zeromodes on cosmic strings are also studied [52] in application for instance to superconducting cosmic strings [53].

The fermionic zeromodes might play important roles even in the electroweak theory of the Standard Model (SM) in particle physics. In fact, the electroweak theory admits a string-like solution of the EOMs called the electroweak $Z$-string [54–56], consisting of the $Z$ boson and Higgs fields. This is a non-topological soliton and its stability is not topologically ensured. Indeed, its stability is highly limited to the small parameter region [57–59].[2]

[1] The ANO vortex/string solution has been mainly studied for the scalar potential in polynomials of the scalar field $\#|\Phi|^2 + \#|\Phi|^4$. Recently, we have studied the ANO vortex/string solution in the Coleman-Weinberg type potential $|\Phi|^4 \log(|\Phi|^2/v_\Phi^2)$ in Ref. [10] and have shown non-trivial interactions between two vortices/strings depending on mass ratio between the scalar and gauge boson masses.

[2] Several mechanisms to stabilize the electroweak string or to enlarge its stable region have been sug-

See also Fig. 4 in Section 2.2 for the stability region of the electroweak string. The electroweak strings, however, have attracted attention because they are proposed to be useful for baryogenesis [66–68] and generation of cosmological magnetic fields [69, 70]. Because the SM fermions inevitably couple with the $Z$ boson and Higgs fields, the $Z$-string can trap fermion zeromodes inside them, and their effects on the stability of the $Z$-strings has been studied so far [71–78]. In particular, the existence of the zeromodes was originally proposed to improve the stability [71] while it was argued [75, 76, 79] that it rather destabilizes because the zeromodes appear as pairs of right-mover and left-mover modes along the $Z$-string and easily form massive modes to lower the energy of the configuration.

The SM introduces only the left-handed neutrinos to be massless. Turning the eye to the phenomenological aspects of high energy physics and cosmology, there is still a lot of possibilities for extensions of the neutrino sector. Such extensions could result in a finite number of zeromodes (or here equivalently topological number) to stabilize the electroweak string. The simplest extension would be the inclusion of the right-handed neutrinos with Majorana masses underlying the seesaw mechanism. It is found in Ref. [80] that in the simplest extension, the Dirac equations for the left- and right-handed neutrino fields have a (normalizable) zeromode solution. Together with a zeromode of the electron field, this indicates that the electroweak string is not stabilized. The zeromode counting is performed in another extension of the SM [81].

In this work, we revisit the neutrino zeromodes on the electroweak-string background in several extensions of the neutrino sector. In particular, we aim to understand the non-trivial topology of the electroweak string arising due to the existence of fermion zeromodes in terms of the topological invariant in momentum space of Green's functions [18, 28, 34, 35].[3] This way allows us to directly count the number of zeromodes from the topology of momentum space without looking for explicit zeromode solutions. Therefore, we do not have to worry about missing the number of zeromodes. Such a topological invariant has been mainly used in non-relativistic systems stated above. This method has several advantages compared to the argument based on the index theorem [44, 46, 47], which is more familiar to high energy physicists. Firstly, the arguments in those papers do not cover cases with Majorana masses while our method relies only on Green's functions and is general enough to cover Majorana cases. Secondly, our method directly calculates the topological invariant defined on real and momentum spaces, which allows us to deform the setup continuously as long as it does not change the topology. We intend to employ this also in particle physics and then demonstrate results obtained in several possible extensions of the neutrino sector including the early studies [80, 81] in order to show that this method successfully counts the number of zeromodes.

This paper is organized as follows: In Section 2, after summarizing model setups

gested: (i) effective potential at finite-temperature [60]; (ii) strong magnetic field [61]; (iii) thermal photon plasma [62]; (iv) axion in core of the $Z$-string [63]; (v) dark scalar condensation [64]; (vi) splitting the $Z$-string into two fractional $Z$-strings in the two-Higgs doublet model [65].

[3] More precisely, we utilize the topological invariant defined in momentum and real spaces since Green's function contains both the momenta and real space coordinates in presence of vortices. Nevertheless, to emphasize the dependence on the momenta, we use the word "momentum space topology" in this paper.

together with clarification of our notation, we briefly review the electroweak string (or $Z$-string) in the SM and its instability under small perturbations. In Section 3, we introduce the topological invariant that counts the number of fermion zeromodes appearing on the $Z$-string solutions. We apply this method to several models as extended neutrino sectors. By comparing these with explicit results obtained by solving the Dirac equations, we show that our method gives the correct values of the number of the zeromodes. Section 4 is devoted to discussion and conclusions. In Appendix B, we introduce the topological number in terms of the momentum-space topology [28, 34, 77] and evaluate it in several models for the neutrino sectors.

## 2 Review on electroweak string in SM

In this section, we briefly review the electroweak string solution (or the $Z$-string solution) in the SM sector.

### 2.1 Model setup and electroweak string solutions

We consider a string solution in the SM. It is known as the electroweak string or the $Z$-string, which has been first argued by Nambu in Ref. [54].

We start with the setup of the model. For our purpose, we focus on the first generation of the leptons (i.e. $e^-$ and $\nu_{eL}$), $SU(2)_L$ and $U(1)_Y$ gauge fields denoted by $W^a_\mu$ and $B_\mu$, respectively, and the Higgs boson $\Phi = (\phi_u \;\; \phi_d)^T$ in the electroweak sector. The first generation of the left-handed $SU(2)_L$ lepton doublet field is denoted by $\ell_L \equiv (\nu_L, e^-_L)^T$ and the right-handed electron singlet field by $e^-_R$. Besides, we introduce the right-handed neutrino $\nu_R$ for which the Majorana mass term is given in the model in addition to the Dirac mass term composed together with the left-handed neutrino and the Higgs field. The total Lagrangian is given by $\mathcal{L}_{\rm tot} = \mathcal{L}_{\rm SM} + \mathcal{L}_{\nu_R}$ where

$$\begin{aligned}\mathcal{L}_{\rm SM} = &-\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\Phi)^\dagger (D^\mu\Phi) - \lambda\left(\Phi^\dagger\Phi - \frac{v_\Phi^2}{2}\right)^2 \\ &+ \ell^\dagger_L i\bar\sigma^\mu D_\mu \ell_L + e^\dagger_R i\sigma^\mu D_\mu e_R - h'\left(e^\dagger_R \Phi^\dagger \ell_L + \ell^\dagger_L \Phi e_R\right), \end{aligned} \tag{2.1}$$

$$\mathcal{L}_{\nu_R} = \nu^\dagger_R i\sigma^\mu\partial_\mu\nu_R - \frac{1}{2}(\nu^c_R)^\dagger M^*_R \nu_R - \frac{1}{2}\nu^\dagger_R M_R \nu^c_R - h\left(\ell^\dagger_L i\tau_2\Phi^*\nu_R - \nu^\dagger_R \Phi^T i\tau_2^\dagger \ell_L\right). \tag{2.2}$$

Here, all spinors are given as two-component Weyl spinors and the superscript $c$ on the right- and left-handed spinors denote the charge conjugation such that $\nu^c_R = -i\sigma^2\nu^*_R$ and $\nu^c_L = i\sigma^2\nu^*_L$, respectively. The covariant derivatives $D_\mu$ act on the fields $\Psi$ with hypercharge $Y$ such that

$$D_\mu\Psi = \left(\partial_\mu - ig\frac{\sigma^a}{2}W^a_\mu - ig'\frac{Y}{2}B_\mu\right)\Psi, \tag{2.3}$$

where $g$ and $g'$ are the $SU(2)$ and $U(1)$ gauge couplings, respectively, and $\sigma^a$ are the Pauli matrices. The field strengths of $W^a_\mu$ and $B_\mu$ are given respectively by

$$W^a_{\mu\nu} \equiv \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\epsilon^{abc}W^b_\mu W^c_\nu, \tag{2.4}$$

$$F_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{2.5}$$

| field | $SU(2)_L$ | $Y$ |
|---|---|---|
| $\ell_L^T = (\nu_{eL} \;\; e_L)$ | **2** | $-1$ |
| $e_R$ | **1** | $-2$ |
| $\nu_R$ | **1** | 0 |
| $\Phi^T = (\phi_u \;\; \phi_d)$ | **2** | 1 |
| $W_\mu^a$ | **3** | 0 |
| $B_\mu$ | **1** | 1 |

**Table 1**. Charge assignment for the particle contents

The charge assignment for each particle is summarized in Table 1.

After the spontaneous symmetry breaking of $SU(2)_L \times U(1)_Y$ into $U(1)_{\rm em}$, i.e. the Higgs field obtains a non-vanishing expectation value $\langle\Phi\rangle = (0 \;\; v_\Phi/\sqrt{2})^T$, the gauge fields have mass terms characterized by $v_\Phi$. As is conventional, we define the $Z$ boson and photon fields respectively by

$$Z_\mu \equiv \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \,, \tag{2.6}$$

$$A_\mu \equiv \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \,, \tag{2.7}$$

where $\theta_W$ is the Weinberg angle which is related to the gauge couplings as

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}} \,, \qquad\qquad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}} \,. \tag{2.8}$$

Hereafter, we write $s_W \equiv \sin\theta_W$ and $c_W \equiv \cos\theta_W$ for simplicity. The charged $W$ boson fields are given by

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2) \,. \tag{2.9}$$

In the broken phase, the covariant derivative (2.3) is written in terms of the new basis as

$$D_\mu\Psi = \left(\partial_\mu - \frac{ig}{\sqrt{2}}\sigma^+ W_\mu^+ - \frac{ig}{\sqrt{2}}\sigma^- W_\mu^- - ig_Z\left(\frac{\sigma^3}{2} - s_W^2 Q\right) Z_\mu - ieQA_\mu\right)\Psi \,, \tag{2.10}$$

where $\sigma^\pm = (\sigma^1 \pm i\sigma^2)/2$ and $Q = \sigma^3/2 + Y/2$ is the electric charge defined through

$$g^2\frac{\sigma^3}{2} - g'^2\frac{Y}{2} = g_Z^2\frac{\sigma^3}{2} - g'^2 Q \,. \tag{2.11}$$

We have also defined

$$g_Z = \sqrt{g^2+g'^2} = \frac{g}{c_W} \,, \qquad\qquad e = \frac{gg'}{\sqrt{g^2+g'^2}} = g s_W \,. \tag{2.12}$$

We recast the Lagrangian for the gauge sector in the broken phase as

$$\mathcal{L}_{\rm gauge}^{\rm Kin} = -\frac{1}{4}A_{\mu\nu}A^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - \frac{M_Z^2}{2}Z_\mu Z^\mu - \frac{1}{2}W_{\mu\nu}^+ W^{-\mu\nu} - M_W^2 W_\mu^+ W^{-\mu} \,, \tag{2.13}$$

$$\begin{aligned}\mathcal{L}_{\rm gauge}^{\rm Int} = {} & iN_{\mu\nu}W^{+\mu}W^{-\nu} - iW_{\mu\nu}^+ N^\mu W^{-\nu} + iW_{\mu\nu}^- N^\mu W^{+\nu} - W_\mu^+ W^{-\mu} N_\nu N^\nu + W_\mu^+ W_\nu^- N^\mu N^\nu \\ & - \frac{g^2}{2}(W_\mu^- W^{-\mu} W_\nu^+ W^{+\nu} - W_\mu^+ W^{-\mu} W_\nu^- W^{+\nu}) \,,\end{aligned} \tag{2.14}$$

where the linear combination of $A_\mu$ and $Z_\mu$ has been defined as

$$N_\mu = eA_\mu - g_Z c_W^2 Z_\mu \,. \tag{2.15}$$

The field strengths of $A_\mu$, $Z_\mu$ and $W_\mu^\pm$ are given respectively by

$$A_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,, \qquad Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu \,, \qquad W_{\mu\nu}^\pm = \partial_\mu W_\nu^\pm - \partial_\nu W_\mu^\pm \,, \tag{2.16}$$

while the field strength of $N_\mu$ is $N_{\mu\nu} = eA_{\mu\nu} - g_Z c_W^2 Z_{\mu\nu}$. The gauge and the Higgs boson obtain the finite masses at the classical level as

$$M_W = \frac{1}{2} g v_\Phi \,, \qquad M_Z = \frac{1}{2} g_Z v_\Phi \,, \qquad M_H = \sqrt{2\lambda v_\Phi^2} \,. \tag{2.17}$$

Now, we focus on the $Z$-string configuration in the SM sector (2.1). To this end, we make the $Z$-string ansatz [54, 55]: In cylindrical coordinates $(r, \theta, z)$, the gauge and Higgs fields take the form

$$\begin{aligned} &qZ_\theta = -n\zeta(r) \,, && \phi_d = \frac{v_\Phi}{\sqrt{2}} f(r) e^{in\theta} \,, \\ &Z_r = Z_z = Z_t = W_\mu^\pm = A_\mu = 0 \,, && \phi_u = 0 \,. \end{aligned} \tag{2.18}$$

Here, we have defined $q \equiv g_Z/2$ and the winding number (or vorticity) $n$, and $\zeta(r)$ and $f(r)$ are the $Z$-string profile functions which are given as non-trivial solutions to the equations of motion for $Z_\mu$ and $\phi_d$:

$$f'' + \frac{1}{r} f' - \frac{n^2(1-\zeta)^2}{r^2} f - \beta f(f^2 - 1) = 0 \,, \tag{2.19}$$

$$\zeta'' - \frac{1}{r} \zeta' + 2(1-\zeta) f^2 = 0 \,, \tag{2.20}$$

with the boundary conditions $f(0) = \zeta(0) = 0$ and $f(\infty) = \zeta(\infty) = 1$. Here, we have rescaled the radius $\tilde{r} = M_Z r/\sqrt{2}$, and omitted tilde for simple notation. In Eq. (2.19), we have defined

$$\beta = \frac{8\lambda}{g_Z^2} = \frac{M_H^2}{M_Z^2} \,, \tag{2.21}$$

which is about 1.887 in the SM. Here and hereafter, we use the prime as the derivative with respect to $r$. The exact solution to Eqs. (2.19) and (2.20) is not known, so that we solve them numerically.

### 2.2 Stability of $Z$ string

In general, the $Z$-string is not a stable object. This is in contrast to the ANO string whose stability is guaranteed by topology. The stability analysis for the $Z$-string has been performed in Refs. [57–59]. In this section, we review its stability by following Ref. [59].

### 2.2.1 Setup

We first denote a set of fields by $\Phi = \{\phi_u, \phi_u^*, W_\mu^+, W_\mu^-, A_\mu, \phi_d, \phi_d^*, Z_\mu\}$ and that of the background fields by $\bar{\Phi} = \{\bar{\phi}_u, \bar{\phi}_u^*, \bar{W}_\mu^+, \bar{W}_\mu^-, \bar{A}_\mu, \bar{\phi}_d, \bar{\phi}_d^*, \bar{Z}_\mu\}$ which correspond to the configuration (2.18). We introduce an infinitesimal perturbation $\delta\Phi$ around the static background $\bar{\Phi}$ and substitute it into the full EOM for $\Phi = \bar{\Phi} + \delta\Phi$,

$$\frac{\delta S}{\delta \Phi(x)} = 0\,, \tag{2.22}$$

from which we get

$$\frac{\delta^2 S}{\delta\Phi\delta\Phi}\Big|_{\Phi=\bar{\Phi}} \delta\Phi = 0 \tag{2.23}$$

at the leading order of the fluctuation $\delta\Phi$. We have used that first derivative of the action vanishes since the background $\bar{\Phi}$ satisfies the static EOM. In particular, those for $\bar{\phi}_d$ and $\bar{Z}^\mu$ read Eqs. (2.19) and (2.20), respectively. This equation (2.23) is decomposed into the four block diagonal parts

$$\mathcal{D}_1 \begin{pmatrix} \delta\phi_u \\ \delta W_\mu^+ \end{pmatrix} = 0, \qquad \mathcal{D}_2 \begin{pmatrix} \delta\phi_u^* \\ \delta W_\mu^- \end{pmatrix} = 0, \qquad \mathcal{D}_3 \begin{pmatrix} \delta\phi_d \\ \delta\phi_d^* \\ \delta Z_\mu \end{pmatrix} = 0, \qquad \mathcal{D}_4 \delta A_\mu = 0\,. \tag{2.24}$$

The differential operators $\mathcal{D}_i$ correspond to the inverse propagators of fields around the background field $\bar{\Phi}$, so that it is convenient to use the gauge fixing action in order to remove zeromodes. To this end, we here employ

$$S_{\rm gf} = \int \mathrm{d}^4x\, \mathcal{L}_{\rm gf} = \frac{1}{2}\int \mathrm{d}^4x \sum_{i=1}^{4} |F_i|^2\,, \tag{2.25}$$

with the following $R_\xi$-gauge fixing functions

$$F_1(W_\mu^+) = \partial^\mu \delta W_\mu^+ - i c_W^2 \bar{Z}^\mu \delta W_\mu^+ - \frac{i c_W}{\sqrt{2}} \bar{\phi}_d^* \delta\phi_u\,, \tag{2.26a}$$

$$F_2(W_\mu^-) = \partial^\mu \delta W_\mu^- + i c_W^2 \bar{Z}^\mu \delta W_\mu^- - \frac{i c_W}{\sqrt{2}} \bar{\phi}_d \delta\phi_u^*\,, \tag{2.26b}$$

$$F_3(Z_\mu) = \partial^\mu \delta Z_\mu - \frac{i}{2}(\bar{\phi}_d \delta\phi_d^* - \bar{\phi}_d^* \delta\phi_d)\,, \tag{2.26c}$$

$$F_4(A_\mu) = \partial^\mu \delta A_\mu\,. \tag{2.26d}$$

These functions modify the differential operators $\mathcal{D}_i$ into $\tilde{\mathcal{D}}_i$. It turns out that the sector $(\delta\phi_d, \delta\phi_d^*, \delta Z_\mu, \delta A_\mu)$ always yields positive eigenvalues, whereas the sectors $(\delta\phi_u, \delta W_\mu^+)$ and its complex conjugate, namely

$$\tilde{\mathcal{D}}_1 \begin{pmatrix} \delta\phi_u \\ \delta W_\mu^+ \end{pmatrix} = 0\,, \qquad \tilde{\mathcal{D}}_2 \begin{pmatrix} \delta\phi_u^* \\ \delta W_\mu^- \end{pmatrix} = 0\,, \tag{2.27}$$

could contain unstable modes. Here, $\tilde{\mathcal{D}}_2$ can be obtained from $\tilde{\mathcal{D}}_1$, so that we concentrate on the $(\delta\phi_u, \delta W_\mu^+)$ sector. We expand the gauge field $\delta W_\mu^\pm$ in eigenstates of spin operator $S_z$ so as to be $S_z W_\uparrow^\pm = +W_\uparrow^\pm$ and $S_z W_\downarrow^\pm = -W_\downarrow^\pm$. More specifically, one has

$$\delta W_\uparrow^\pm = e^{-i\theta}\left(W_r^\pm - \frac{i}{r}W_\theta^\pm\right), \qquad \delta W_\downarrow^\pm = (\delta W_\uparrow^\pm)^\dagger . \tag{2.28}$$

To obtain the eigenequations, we perform the separation of spacetime variables such that

$$\delta\phi_u(x) = s_u(r)e^{i\ell\theta}e^{i\omega t}, \tag{2.29}$$

$$\delta W_\uparrow^+(x) = i w_{\uparrow,\ell}(r)e^{i(\ell-1-n)\theta}e^{i\omega t}, \tag{2.30}$$

$$\delta W_\downarrow^+(x) = i w_{\downarrow,\ell}(r)e^{i(\ell+1-n)\theta}e^{i\omega t}. \tag{2.31}$$

Inserting Eqs. (2.29)–(2.31) into the first equation in Eq. (2.27) leads to

$$\mathcal{M}_1 \begin{pmatrix} s_{u,\ell} \\ w_{\uparrow,\ell} \\ w_{\downarrow,\ell} \end{pmatrix} = \omega^2 \begin{pmatrix} s_{u,\ell} \\ w_{\uparrow,\ell} \\ w_{\downarrow,\ell} \end{pmatrix}, \tag{2.32}$$

where the stability matrix is given by

$$\mathcal{M}_1 = \begin{pmatrix} D_{1,\ell} & B_{1\uparrow,\ell} & B_{1\downarrow,\ell} \\ B_{1\uparrow,\ell} & D_{\uparrow,\ell} & 0 \\ B_{1\downarrow,\ell} & 0 & D_{\downarrow\ell} \end{pmatrix}, \tag{2.33}$$

with the differential operators

$$D_{1,\ell} = -\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{(\ell - n\cos 2\theta_W\zeta)^2}{r^2} + 2c_W^2 f^2 + \beta(f^2-1), \tag{2.34a}$$

$$D_{\uparrow,\ell} = -\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{(\ell - 1 - n - 2nc_W^2\zeta)^2}{r^2} + 2c_W^2 f^2 + 4nc_W^2\frac{\zeta'}{r}, \tag{2.34b}$$

$$D_{\downarrow,\ell} = -\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{(\ell + 1 - n - 2nc_W^2\zeta)^2}{r^2} + 2c_W^2 f^2 - 4nc_W^2\frac{\zeta'}{r}, \tag{2.34c}$$

$$B_{1\uparrow,\ell} = 2c_W\left(f' - \frac{nf(1-\zeta)}{r}\right), \tag{2.34d}$$

$$B_{1\downarrow,\ell} = -2c_W\left(f' + \frac{nf(1-\zeta)}{r}\right). \tag{2.34e}$$

Here, we have used the dimensionless quantities

$$A_\mu = \frac{v_\Phi}{\sqrt{2}}\tilde{A}_\mu, \qquad W_\mu^\pm = \frac{v_\Phi}{\sqrt{2}}\tilde{W}_\mu^\pm, \qquad Z_\mu = \frac{v_\Phi}{\sqrt{2}}\tilde{Z}_\mu, \qquad \Phi = \frac{v_\Phi}{\sqrt{2}}\tilde{\Phi}, \tag{2.35}$$

and then have omitted tildes on them. Note that in this energy unit, the eigenvalues $\omega$ are given in the energy unit $M_Z/\sqrt{2} \equiv 1$ for which the Higgs mass is given as $M_H^2 = 2\beta$. Evaluating the eigenvalues of the stability matrix (2.33), one can investigate the instability by exploring negative eigenvalues. In the operators $D_{\uparrow,\ell}$ and $D_{\downarrow,\ell}$, the term $2c_W^2 f^2$ corresponds to the $W$ boson mass due to the non-vanishing Higgs background field $|\phi_d| = f(r)$. The third terms in the operators $D_{i,\ell}$ are the angular momenta. We should especially note that the non-vanishing $Z$ boson field background $\bar{Z}_\theta = -\zeta(r)/q$ plays a role of an external "$SU(2)$-magnetic field", $\vec{B}_Z = \vec{\partial} \times \vec{\bar{Z}}$. Thus, due to the interaction between $\vec{B}_Z$ and the spin of the $W$ bosons ($S_z = \pm 1$), there exist the spin magnetic moments $\pm 4nc_W^2\zeta'/r$ in Eqs. (2.34b) and (2.34c). As we will see in Section 2.2.2, these effects become one of origins of instability.

### 2.2.2 Origin of instability

We are interested especially in the case of $n = 1$ and $\ell = 0$ for which a negative eigenvalue $\omega^2 < 0$ is observed. Below, we analyze the eigenequation (2.32) only in that case.

We first consider the case of $c_W = 0$ ($s_W = 1$) for which the $Z$ boson field is identified with $U(1)_Y$ gauge field as can be seen from Eq. (2.6): $Z_\mu = -B_\mu$. Therefore, the $Z$-string is regarded as a semilocal string [82]. The stability matrix (2.33) in this case becomes diagonal:

$$\begin{aligned}\mathcal{M}_1 &= \mathrm{diag}\,(D_{1,0},\ D_{\uparrow,0},\ D_{\downarrow,0}) \\ &= \begin{pmatrix} -\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\zeta^2}{r^2} + \beta(f^2-1) & 0 & 0 \\ 0 & -\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{4}{r^2} & 0 \\ 0 & 0 & -\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} \end{pmatrix}. \end{aligned} \tag{2.36}$$

Therefore, our task is to evaluate and to compare the eigenvalues of $D_{i,0}$ ($i = 1, \uparrow, \downarrow$),

$$D_{1,0}s_{u,0} = \omega_1^2 s_{u,0}\,, \qquad D_{\uparrow,0}w_{\uparrow,0} = \omega_\uparrow^2 w_{\uparrow,0}\,, \qquad D_{\downarrow,0}w_{\downarrow,0} = \omega_\downarrow^2 w_{\downarrow,0}\,. \tag{2.37}$$

Because the fluctuation of the $W$ bosons are decoupled from the background, the operators $D_{\uparrow,0}$ and $D_{\downarrow,0}$ are independent of both $\beta$ and the profile of the string solutions, $f(r)$ and $\zeta(r)$. Indeed, the second and third equations in Eq. (2.37),

$$\left(-\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{4}{r^2}\right) w_{\uparrow,0} = \omega_\uparrow^2 w_{\uparrow,0}\,, \qquad \left(-\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right) w_{\downarrow,0} = \omega_\downarrow^2 w_{\downarrow,0}\,. \tag{2.38}$$

are classes of the Helmholtz equation which yields positive eigenvalues, $\omega_\uparrow^2 > 0$ and $\omega_\downarrow^2 > 0$. Note that the solutions to these equations are given by

$$w_{\uparrow,0}(r) = c_1 J_2(\omega_\uparrow r) + c_2 Y_2(\omega_\uparrow r)\,, \qquad w_{\downarrow,0}(r) = c_3 J_0(\omega_\downarrow r) + c_4 Y_0(\omega_\downarrow r)\,, \tag{2.39}$$

where $J_n(x)$ and $Y_n(x)$ are the Bessel functions of the first and second kinds, respectively, and $c_i$ ($i = 1, \ldots, 4$) are constant. On the other hand, the eigenvalue of $D_{1,0}$ explicitly

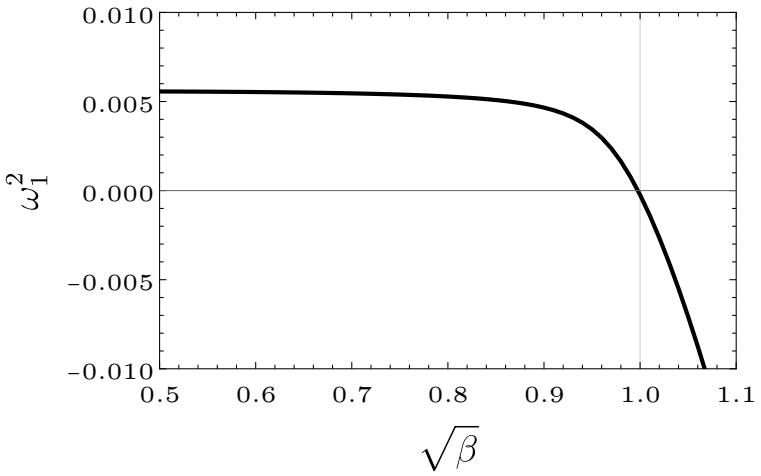



**Figure 1**. $\sqrt{\beta}$-dependence of the lowest eigenvalue of the stability matrix (2.33) with $s_W^2 = 1$. Note $\beta \equiv 8\lambda/g_Z^2 = M_H^2/M_Z^2$.

depends on $\beta$ and the profile functions. In Fig. 1, the $\sqrt{\beta}$-dependence of the lowest eigenvalues of the first equation in Eq. (2.37) is displayed. Here and hereafter, to perform numerical calculations, we set the boundary condition $s_{u,0}(r_{\max}) = 0$ with $r_{\max} = 50$ instead of $r_{\max} = \infty$. We see that $\omega_1^2$ is positive for $\sqrt{\beta} < 1$, while for $\sqrt{\beta} > 1$, it turns to negative. This is understood as that the Higgs potential gives a tachyonic contribution proportional to $\beta$, which is balanced at $\beta = 1$ with the positive contribution from the kinetic term. This agrees with the stability analysis in semilocal string [83, 84].

We next turn to the case $\beta$ to be small so that the tachyonic contribution from Higgs potential becomes insignificant and hence the instability comes only from the gauge sector. Especially when $s_W = 0$, the $Z$ boson is identified with $W_\mu^3$, while the electromagnetic field $A_\mu$ is just the $U(1)_Y$ field, $B_\mu$. In this case, the so-called "$W$-condensation" is clearly observed as an origin of instability. To explain $W$-condensation, let us here consider the "$SU(2)_L$-magnetic background field" defined by $B_{Z,i} = \epsilon_{ijk}\partial_j \bar{Z}_k$ which couples to the $W$ bosons with the coupling $g_Z c_W^2$. Assuming that $\vec{B}_Z = (0, 0, B_Z)$, the $x$- and $y$-direction momenta $p_x^2 + p_y^2$ of the charged $W$ bosons, $W_\mu^\pm$, are discretized as the Landau levels such that their energy dispersion is given by

$$E^2 = (2l+1)g_Z c_W^2 B_Z + p_z^2 + M_W^2 + 2g_Z c_W^2 B_Z S_z \,, \tag{2.40}$$

where $p_z$ is the momentum along with the $z$-direction and $l$ are the discrete positive integer numbers, $l = 0, 1, 2, \ldots$, denoting the Landau levels. Here, the last term in Eq. (2.40) corresponds to the magnetic moment between the spin $S_z = \pm 1$ of the $W$ bosons and $B_Z$. In the lowest energy of the spin $S_z = -1$ case, namely $p_z = 0$ and $l = 0$ for which one has $E^2 = M_W^2 - g_Z c_W^2 B_Z$, the energy could be negative when

$$B_Z > \frac{M_W^2}{g_Z c_W^2} \,. \tag{2.41}$$

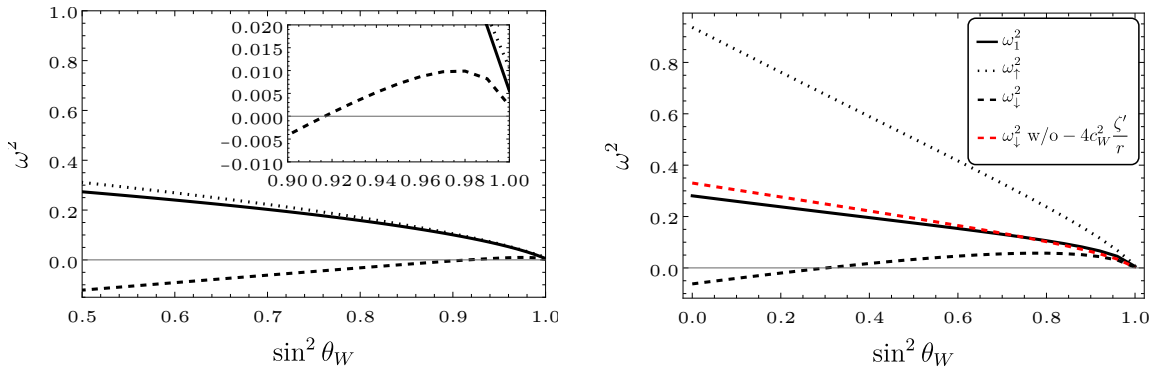



**Figure 2**. $\sin^2\theta_W$-dependence of the lowest three eigenvalues (black dashed, solid, and dotted lines) of stability matrices of Eq. (2.33) with $\beta = 0.01$ (left) and of the approximated version in Eq. (2.42) with $\beta = 0.01$ (right). The red dashed line in the right panel indicates the lowest eigenvalue for the stability matrix in which the term $4c_W^2\frac{\zeta'}{r}$ is eliminated by hand.

This instability is called "$W$ condensation" [85–88]. Thus, as the background field $B_Z$ is stronger, the $W$ bosons tend to form condensation at the Landau level $l = 0$. This is confirmed by analyzing the stability matrix (2.33) numerically.

In terms of the stability matrix (2.33), the last term in the operator (2.34c), namely $-4c_W^2\zeta'/r$, corresponds to the spin magnetic moment which causes an instability due to $W$-condensation. To see this, let us first consider the approximated stability matrix

$$\begin{aligned}\mathcal{M}_1 \approx \mathrm{diag}\Bigg( &-\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{(\cos 2\theta_W\zeta)^2}{r^2} + 2c_W^2 f^2 + \beta(f^2-1), \\ &-\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{4(1+c_W^2\zeta)^2}{r^2} + 2c_W^2 f^2 + 4c_W^2\frac{\zeta'}{r}, \\ &-\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{4c_W^4\zeta^2}{r^2} + 2c_W^2 f^2 - 4c_W^2\frac{\zeta'}{r}\Bigg),\end{aligned} \tag{2.42}$$

and then investigate eigenvalues of the each diagonal part. The right side panel of Fig. 2 depicts each eigenvalue of Eq. (2.42). Here, the eigenvalues are denoted as the same as of Eq. (2.37). We can see that $D_{1,0}$ and $D_{\uparrow,0}$ ($\omega_1^2$ and $\omega_\uparrow^2$) yield positive eigenvalues, while $D_{\downarrow,0}$ produces negatives $\omega_\downarrow^2$ for $\sin^2\theta_W < 0.3$. When we eliminate $-4c_W^2\zeta'/r$ from $D_{\downarrow,0}$ by hand, its eigenvalues become positive for all range of $\sin^2\theta_W$ as shown by the red dashed line in the right panel of Fig. 2. On the other hand, the left panel of Fig. 2 shows its three lowest eigenvalues of the stability matrix (2.33) without any approximation for $\beta = 0.01$. One of eigenvalues (the dashed line in Fig. 2) takes a negative value for $\sin^2\theta_W \lesssim 0.92$. One can see that qualitative tendency agrees with the approximated one (right panel).

We show the dependence of the lowest eigenvalue of stability matrix (2.33) on $s_W^2$ and $\sqrt{\beta}$ in Fig. 3. Note that the $\sqrt{\beta}$-dependence of $\omega^2$ at $s_W^2 = 1$ agrees with that in Fig. 1. The larger values of $\sqrt{\beta}$ are, the more negative the lowest eigenvalue tends to be. In particular,

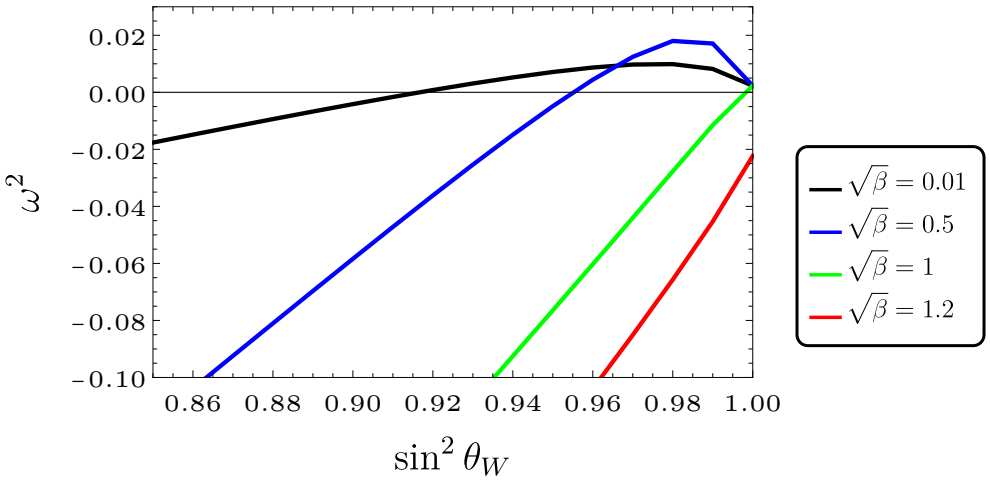



**Figure 3**. $\sin^2\theta_W$-dependence of the lowest eigenvalue of stability matrix (2.33) with different values of $\sqrt{\beta}$. Note $\beta \equiv 8\lambda/g_Z^2 = M_H^2/M_Z^2$.

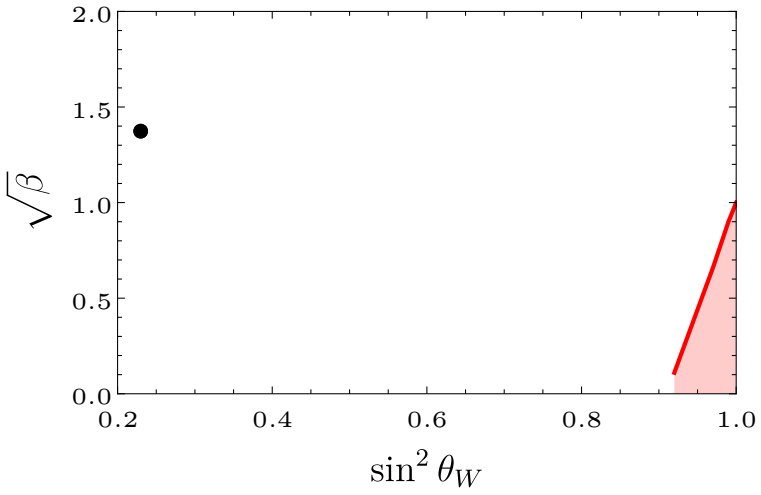



**Figure 4**. The $Z$-string is stabilized within the red region on $s_W^2$-$\sqrt{\beta}$ plane. The black point denotes the experimental values $s_W^2 = 0.23$ and $\sqrt{\beta} \simeq 1.37$.

for $\sqrt{\beta} > 1$, no value of $s_W^2$ yielding positive eigenvalues exists, which indicates that there is no stabilized region.

To summarize, the stable region on $s_W^2$-$\sqrt{\beta}$ plane is displayed in Fig. 4. Only in the triangle region, the $Z$-string can stably exist in sense that it is robust under small perturbations. Obviously, the experimental values, $s_W^2 = 0.23$ and $\sqrt{\beta} \simeq 1.37$ (which is shown by the black point in Fig. 4), are out of the stabilized region, so that the $Z$-string in the SM is an unstable object.

In the following, we discuss the impact of fermion zeromodes on the $Z$-string [80, 81] as a mechanics of its stabilization.

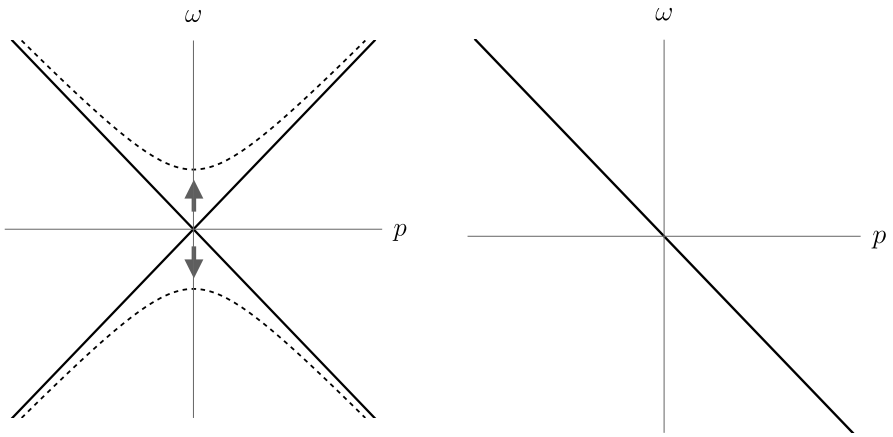



**Figure 5**. Illustration for two dimensional spectrum of fermion zeromodes in the string background. If the zeromodes appear as pairs of right and left-mover Weyl fermions (solid line in left panel), their degeneracy can be resolved by some perturbation, resulting in massive Dirac fermions in two dimensions (dashed line in left panel). In the sense of the Dirac index, this configuration is in the same topological sector as that of the vacuum (no zeromodes). On the other hand, if there exists a zeromode with single chirality, e.g., only left-mover, the zeromode cannot be made massive nor removed by any continuous perturbation (right panel). This string configuration might be regarded as in a different topological sector from that of the vacuum and is conjectured to be stable in Ref. [81].

### 2.2.3 Stabilization mechanism due to fermion zeromodes

Let us here briefly explain a fate of fermion zeromodes on the electroweak string. We consider the doublet fields $\Psi = (\nu, e)^T$, $(u, d)^T$, etc. on the electroweak string background. Note that zeromodes are degenerate energy eigenvalues. Solving the equations of motion (i.e., the Dirac equations) on the electroweak string background, one can see that the up and down components in $\Psi$ move in parallel and anti-parallel to the electroweak string flux, respectively [72, 74].

In this sense, when the electroweak string flux is in the $z$ direction, the up component is called a "right-mover" as proportional to $e^{-i\omega t+ipz}$ with a frequency $\omega$ and a momentum $p$, while the down one is a "left-mover" as proportional to $e^{-i\omega t-ipz}$. For the ANO string with a single fermion, only one zeromode solution appears as a Weyl fermion in two dimensional sense [43] and cannot be removed by any continuous deformation of the configuration. See the right panel in Fig.5. However, this does not happen in the case of the electroweak string which is not topological: The string necessarily couples to a fermion doublet $\Psi$, which provides both of the left-mover and right-mover Weyl fermions as a pair to form a two dimensional Dirac fermion. When the electroweak string is slightly perturbed with a parameter $\epsilon$, the degeneracy of the zeromode eigenvalues is resolved so that the Dirac fermion obtains finite masses of order of $\epsilon^2$. (See the solid and dashed lines in the left panel in Fig.5 for the unperturbed and perturbed spectrum.) This is shown by the standard perturbation theory as the correction to the zero energy eigenvalues at the first-order perturbation of order $\epsilon$ vanishes, i.e., $E^{(1)} = 0$ and the energy eigenvalues are corrected by

$E^{(2)}$ of the second order of $\epsilon$. It has been argued in Refs. [75, 76, 79] that fermions make the infinitely long electroweak string unstable due to the Dirac sea in which all negative energy eigenvalues should be summed up to give a negatively divergent contribution to $E^{(2)}$. Thus, it seems that the string cannot be stabilized under any perturbation.

Now, we turn to see the stabilization mechanism of the electroweak string in Ref. [81] due to the fermion. To this end, we assume that, in some extension of the neutrino sector, the up-component fields in $\Psi$ do not have zeromodes while the down-component fields still keep zeromodes. In such a situation, the zeromode-energy eigenmodes cannot be made massive in the same way as described above and thus survive under any continuous deformation (see the right panel in Fig.5). Then, the discrepancy of the number of zeromodes between the up and down components leads to the non-vanishing topological number in the sense of the index theorems. Hence, the electroweak string is a topologically distinct object from the vacuum. This implies that the string cannot decay by any continuous deformation. The main scope of this paper is to revisit this mechanism in several models of extensions in the neutrino sector.

## 3 Number of zeromodes on vortices and its relation with momentum space topology

Now, we are ready to discuss the fermion zeromodes that appear on the $Z$-string. To this end, we first present a convenient formalism based on the topological invariant defined in the momentum space (and real space), originally introduced in Refs. [28, 34]. In this section, we come back to the dimensionful unit for the coordinate $r$.

### 3.1 Topological invariant for the number of zeromodes

We consider a Green's function $\mathcal{G}(x^\mu, p_\mu)$ of a fermion field in a background of vortex strings, located on the $z$ axes. Due to the translational invariance with respect to the $z$ and $t$ directions, the derivatives with respect to $z$ and $t$ in the Green's function can be replaced by $p_3$ and $p_4 (\equiv ip_0)$, while in the $x$ and $y$ directions it contains $x$, $y$, $\partial_x$, and $\partial_y$ explicitly.

Here, we introduce the following topological quantity:

$$\mathcal{N}_1 \equiv \frac{1}{2\pi i}\int_C \mathrm{Tr}\left[\mathcal{G}(p_3,p_4)\mathrm{d}\mathcal{G}(p_3,p_4)^{-1}\right] , \tag{3.1}$$

where $\mathcal{G}(p_3, p_4)$ is the Green's function with $p_3$ and $p_4$ being $c$-numbers as stated above. For instance, one may give $\mathcal{G}^{-1} = \gamma^\mu \hat{p}_\mu + M(\hat{x}, \hat{y})|_{\hat{p}_3=p_3, \hat{p}_4=p_4}$. The integration contour $C$ is a circle with infinitesimal radius embracing $p_3 = p_4 = 0$ on $p_3$-$p_4$ plane. The trace is taken over the function space of $x$ and $y$ (on which $\partial_x$ and $\partial_y$ act) and the spinor indices. The fact that $\mathcal{N}_1$ is a topological invariant is understood immediately below.

To make the physical meaning of $\mathcal{N}_1$ clearer, we start with the fact that $\mathcal{G}^{-1}$ can be expressed as

$$\mathcal{G}^{-1} = (ip_4 + \hat{\mathcal{H}})\gamma^0 , \tag{3.2}$$

where $\hat{\mathcal{H}}$ corresponds to the one-particle Hamiltonian in the background of the strings. Since it is Hermitian and commutes with $\hat{p}_3$ and $\hat{p}_4$, we can diagonalize it independently of $p_3$ and $p_4$. Thus if one denotes the $n$-th eigenstates of $\hat{\mathcal{H}}$ by $|n\rangle$, the quantity (3.1) is given as follows:

$$\begin{aligned}\mathcal{N}_1 &= \frac{1}{2\pi i}\sum_n \langle n| \int_C \mathcal{G}(p_3,p_4)\mathrm{d}\mathcal{G}(p_3,p_4)^{-1}\,|n\rangle \\ &= \frac{1}{2\pi i}\sum_n \int_C G_n(p_3,p_4)\mathrm{d}G_n(p_3,p_4)^{-1},\end{aligned} \tag{3.3}$$

with

$$G_n^{-1} \equiv ip_4 + \mathcal{E}_n(p_3)\,, \tag{3.4}$$

where $\mathcal{E}_n$ is the eigenvalue of $\hat{\mathcal{H}}$ for the state $|n\rangle$. From this expression, one can find that $\mathcal{N}_1$ is a topological invariant. For each eigenstate $|n\rangle$, [4] the integration gives the winding number from the circle $C$ on $p_3$-$p_4$ plane to the complex phase of $G_n$, which is characterized by the elements of the first homotopy group $\pi_1(U(1)) = \mathbb{Z}$. Note here that we have assumed that there is no massless particle in the bulk, which corresponds to a continuous spectrum starting from zero in two dimensions. Such a massless particle makes the integrand in Eq. (3.1) singular with a branch cut starting from the origin on the original Lorentzian $p_0$-$p_3$ plane which leads to an obstacle to perform the calculation. Thus, the argument presented here is not applicable to the SM since the neutrinos are massless.

Defining a complex variable $z \equiv p_3 + ip_4$, we can rewrite Eq. (3.3) on the complex plane as

$$\mathcal{N}_1 = \frac{1}{2\pi i}\sum_n \int_C \mathrm{d}z \frac{\partial}{\partial z}\log G_n^{-1}(p_3,p_4)\,. \tag{3.5}$$

Note that the integration picks up residues of poles at $p_3 = p_4 = 0$, and thus vanishes for such $n$ that $\mathcal{E}_n \neq 0$ at $p_3 = p_4 = 0$. These are massive modes in two dimensions.[5] Therefore, only modes with vanishing $\mathcal{E}_n$ at $p_3 = p_4 = 0$ (just corresponding to massless modes) can contribute to the integration. For such zeromodes, denoting $\mathcal{E}_n \simeq \lambda_n p_3 + \mathcal{O}(p_3^2)$, we have

$$\mathcal{N}_1 = \frac{1}{2\pi i}\sum_n \int_C \mathrm{d}z \frac{\partial}{\partial z}\log(ip_4 + \lambda_n p_3)\,, \tag{3.6}$$

which means that $\mathcal{N}_1$ takes the sign of $\lambda_n$ for each massless mode. Because $\lambda_n$ determines the two-dimensional chirality of the massless modes, this is nothing but the Dirac index in two dimensions, namely,

$$\mathcal{N}_1 = n_+ - n_-\,, \tag{3.7}$$

where $n_+$ and $n_-$ are the numbers of the right- and left-mover modes on the vortex, respectively.

[4] For continuous spectrum, the summation over $n$ should be replaced with integration over continuous eigenvalues.

[5] Massive modes provide poles or cuts on $p_0$-$p_3$ plane, while they do not on $p_3$-$p_4$ plane thanks to the Wick rotation.

For later use, it is convenient to write the quantity (3.1) in terms of the Wigner transform of the Green's function. Here, the Wigner transformed Green's function is defined by

$$\tilde{G}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) \equiv \int \mathrm{d}^2 r_R \; e^{-i\boldsymbol{p}\cdot\boldsymbol{r}_R} \left\langle \boldsymbol{r}_1 \right| \mathcal{G}(p_3,p_4) \left| \boldsymbol{r}_2 \right\rangle \,, \tag{3.8}$$

where $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are vectors in $x$-$y$ position space with $\boldsymbol{R} \equiv (\boldsymbol{r}_1+\boldsymbol{r}_2)/2$ and $\boldsymbol{r}_R \equiv \boldsymbol{r}_1 - \boldsymbol{r}_2$ and $\boldsymbol{p}$ denotes a vector in the momentum $p_1$-$p_2$ space. The bra and ket are eigenfunctions in the position space. By definition, the following identity holds:

$$\int \mathrm{d}^2 R \int \frac{\mathrm{d}^2 p}{(2\pi)^2} \; \tilde{G}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) = \mathrm{Tr}\; \mathcal{G}(p_3,p_4)\,. \tag{3.9}$$

Thus, the Wigner transformation can be regarded as a map from an operator including the canonical operators $\hat{x}$, $\hat{p}$ into a *classical* distribution function defined on the phase space $(x,p)$.

By using the Wigner transformations for $\mathcal{G}$ and $\mathcal{G}^{-1}$, which are denoted by $\tilde{G}$ and $\mathcal{Q}$ hereafter, respectively, Eq. (3.1) reads

$$\mathcal{N}_1 = \frac{1}{2\pi i} \int \mathrm{d}^2 R \int \frac{\mathrm{d}^2 p}{(2\pi)^2} \int_C \mathrm{Tr}_D \left[ \tilde{G}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) \mathrm{d}\mathcal{Q}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) \right] \,, \tag{3.10}$$

where $\mathrm{Tr}_D$ denotes the trace acting on the spinor and other internal indices and $\mathcal{Q}$ is determined by the relation

$$\begin{aligned} 1 &= \mathcal{Q}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) * \tilde{G}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) \\ &\equiv \mathcal{Q}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) \exp\left[ \frac{i}{2} (\overleftarrow{\partial_R}\cdot\overrightarrow{\partial_p} - \overleftarrow{\partial_p}\cdot\overrightarrow{\partial_R}) \right] \tilde{G}(\boldsymbol{R},\boldsymbol{p},p_3,p_4)\,, \end{aligned} \tag{3.11}$$

which follows from the Wigner transformation of the identity $\mathcal{G}\mathcal{G}^{-1}=1$. Based on the derivative expansion of $\boldsymbol{R}$ and $\boldsymbol{p}$, Eq. (3.11) is solved iteratively as

$$\tilde{G}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) \equiv \tilde{G}^{(0)} + \tilde{G}^{(1)} + \tilde{G}^{(2)} + \cdots \tag{3.12}$$

with

$$\tilde{G}^{(0)}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) = \mathcal{Q}(\boldsymbol{R},\boldsymbol{p},p_3,p_4)^{-1}\,, \tag{3.13}$$

$$\tilde{G}^{(1)}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) = -\frac{i}{2}\tilde{G}^{(0)} \left[ \partial_{\boldsymbol{R}}\mathcal{Q}\cdot\partial_{\boldsymbol{p}}\tilde{G}^{(0)} - \partial_{\boldsymbol{p}}\mathcal{Q}\cdot\partial_{\boldsymbol{R}}\tilde{G}^{(0)} \right] \,, \tag{3.14}$$

$$\begin{aligned} \tilde{G}^{(2)}(\boldsymbol{R},\boldsymbol{p},p_3,p_4) = &-\frac{i}{2}\tilde{G}^{(0)} \left[ \partial_{\boldsymbol{R}}\mathcal{Q}\cdot\partial_{\boldsymbol{p}}\tilde{G}^{(1)} - \partial_{\boldsymbol{p}}\mathcal{Q}\cdot\partial_{\boldsymbol{R}}\tilde{G}^{(1)} \right] \\ &- \frac{1}{2}\left(\frac{i}{2}\right)^2 \Big[ \partial_{R_i}\partial_{R_j}\mathcal{Q}\,\partial_{p_i}\partial_{p_j}\tilde{G}^{(0)} + \partial_{p_i}\partial_{p_j}\mathcal{Q}\,\partial_{R_i}\partial_{R_j}\tilde{G}^{(0)} \\ &\qquad -2\partial_{p_i}\partial_{R_j}\mathcal{Q}\,\partial_{R_i}\partial_{p_j}\tilde{G}^{(0)} \Big] \,. \end{aligned} \tag{3.15}$$

This formula allows us to rewrite eventually the topological invariant $\mathcal{N}_1$ in the form of

$$\begin{aligned} \mathcal{N}_1 &= \frac{2}{5!(2\pi)^3 i} \int \mathrm{Tr}_D \left[ \tilde{G}^{(0)} \mathrm{d}(\tilde{G}^{(0)})^{-1} \wedge \mathrm{d}\tilde{G}^{(0)} \wedge \mathrm{d}(\tilde{G}^{(0)})^{-1} \wedge \mathrm{d}\tilde{G}^{(0)} \wedge \mathrm{d}(\tilde{G}^{(0)})^{-1} \right] \\ &= \frac{2}{5!(2\pi)^3 i} \int \mathrm{Tr}_D \left[ \tilde{G}^{(0)} \mathrm{d}(\tilde{G}^{(0)})^{-1} \right]^5 \,, \end{aligned} \tag{3.16}$$

where the integration is taken over the surface $C \otimes R^2 \otimes R^2$; $C$ is the circle embracing $p_3 = p_4 = 0$ in $p_3$-$p_4$ space, the first $R^2$ corresponds to $x$-$y$, and the second $R^2$ corresponds to $p_1$-$p_2$ space. See Refs. [28, 89]. Now, one can see that $\mathcal{N}_1$ is expressed by the winding number characterizing maps from the five-dimensional space to $GL(N, \mathbb{C})$ with $N$ being the size of the matrix $\tilde{G}^{(0)}$. Note that there are no singularities of $\tilde{G}$ in the entire six dimensional space other than the massless poles located on the origin. Due to the topological nature, we can rotate the integration surface into $C' \otimes R^4$ keeping the massless poles included inside the integral region, where $C'$ is a circle embracing $x = y = 0$ in $x$-$y$ space and $R^4$ corresponds to the momentum space $p^1, \cdots, p^4$. Furthermore, as long as the mass gap is not closed in the bulk, or equivalently, a new singularity appears inside the integration circle $C$ in Eq. (3.5), this quantity is topologically protected under any perturbation of the model. In the following, we calculate $\mathcal{N}_1$ (3.16) with the integration surface $C' \otimes R^4$ for some benchmark cases.

### 3.2 Dirac fermion on Abrikosov-Nielsen-Olesen vortex

As an exercise, we start with a demonstration for counting the number of zeromodes with Eq. (3.16) in case of a Dirac fermion in the presence of the (global) ANO vortex. In this system, there exists a zeromode [43], so that the topological invariant (3.16) has to entail nonzero value.

The Green's function for the fermion field in the ANO vortex background is given as

$$\mathcal{G}^{-1} = \gamma^i \hat{p}_i + \gamma^a p_a - m_D f(r) e^{in\gamma_5 \theta} \,, \tag{3.17}$$

where $m_D$ is the Dirac fermion mass, $n$ is the vorticity, $x + iy = re^{i\theta}$, $f(r)$ is the profile function satisfying $f(0) = 0$ and $f(\infty) = 1$, and the indices run as $i = 1, 2$, and $a = 3, 0$ (with $p_0 = -ip_4$). Note that $\gamma^i$ and $\gamma^0$ are anti-hermitian and hermitian, respectively. The Wigner transformation for the Green's function (3.17) is computed explicitly as

$$\begin{aligned}
\left(\tilde{G}^{(0)}\right)^{-1} &= \mathcal{Q}(\boldsymbol{R}, \boldsymbol{p}, p_3, p_4) \\
&= \int \mathrm{d}^2 r_R \; e^{-i\boldsymbol{p}\cdot\boldsymbol{r}_R} \left\langle \boldsymbol{R} + \frac{\boldsymbol{r}_R}{2} \right| \left( \gamma^i \hat{p}_i + \gamma^a p_a - m_D f(r) e^{in\gamma_5\theta} \right) \left| \boldsymbol{R} - \frac{\boldsymbol{r}_R}{2} \right\rangle \\
&= \int \mathrm{d}^2 r_R \; e^{-i\boldsymbol{p}\cdot\boldsymbol{r}_R} \left( \int \frac{\mathrm{d}^2 q}{(2\pi)^2} \left\langle \boldsymbol{R} + \frac{\boldsymbol{r}_R}{2} \right| \gamma^i \hat{p}_i + \gamma^a p_a \Big| q \Big\rangle \left\langle q \left| \boldsymbol{R} - \frac{\boldsymbol{r}_R}{2} \right\rangle - \delta^2(\boldsymbol{r}_R) m_D f(r) e^{in\gamma_5\theta} \right) \\
&= \int \mathrm{d}^2 r_R \int \frac{\mathrm{d}^2 q}{(2\pi)^2} \; e^{-i(\boldsymbol{p}-\boldsymbol{q})\cdot\boldsymbol{r}_R} \; (\gamma^i q_i + \gamma^a p_a) - \int \mathrm{d}^2 r_R \; e^{-i\boldsymbol{p}\cdot\boldsymbol{r}_R} \; \delta^2(\boldsymbol{r}_R) m_D f(r) e^{in\gamma_5\theta} \\
&= \gamma^\mu p_\mu - m_D f(r) e^{in\gamma_5\theta} \,.
\end{aligned} \tag{3.18}$$

For later use, we rewrite this as

$$\left(\tilde{G}^{(0)}\right)^{-1} = e^{in\gamma_5 \frac{\theta}{2}} \left(\gamma^\mu p_\mu - m_D f(r)\right) e^{in\gamma_5 \frac{\theta}{2}} \,. \tag{3.19}$$

Let us now evaluate $\mathcal{N}_1$ by substituting $\tilde{G}^{(0)}$ of Eq. (3.19) into Eq. (3.16). Noting here that the integration contour $C'$ reduces to an integration with respect to $\theta$ on a circle far

from the core of the vortex, we have

$$\begin{aligned}\mathcal{N}_1 = \frac{2\,(5\cdot 4)\;\epsilon_{ijk}}{5!(2\pi)^3}\int \mathrm{d}^3p \int \mathrm{d}p_4 \int_0^{2\pi}\mathrm{d}\theta \\ \times \mathrm{Tr}_D\Bigg[\left(\tilde{G}^{(0)}\frac{\partial}{\partial p_i}\left(\tilde{G}^{(0)}\right)^{-1}\right)\left(\tilde{G}^{(0)}\frac{\partial}{\partial p_j}\left(\tilde{G}^{(0)}\right)^{-1}\right) \\ \times\left(\tilde{G}^{(0)}\frac{\partial}{\partial p_k}\left(\tilde{G}^{(0)}\right)^{-1}\right)\left(\tilde{G}^{(0)}\frac{\partial}{\partial p_4}\left(\tilde{G}^{(0)}\right)^{-1}\right)\left(\tilde{G}^{(0)}\frac{\partial}{\partial \theta}\left(\tilde{G}^{(0)}\right)^{-1}\right)\Bigg]. \quad (3.20)\end{aligned}$$

Here, $\epsilon_{ijk}$ is the anti-symmetric tensor with $\epsilon_{123}=1$. From the explicit evaluations as

$$\tilde{G}^{(0)}\frac{\partial}{\partial\theta}\left(\tilde{G}^{(0)}\right)^{-1} = e^{-in\gamma_5\frac{\theta}{2}}\left(\gamma^\mu p_\mu - m_D\right)^{-1}(-inm_D\gamma^5)e^{in\gamma_5\frac{\theta}{2}}, \quad (3.21a)$$

$$\tilde{G}^{(0)}\frac{\partial}{\partial p_4}\left(\tilde{G}^{(0)}\right)^{-1} = e^{-in\gamma_5\frac{\theta}{2}}\left(\gamma^\mu p_\mu - m_D\right)^{-1}(-i)\gamma^0 e^{in\gamma_5\frac{\theta}{2}}, \quad (3.21b)$$

$$\tilde{G}^{(0)}\frac{\partial}{\partial p_i}\left(\tilde{G}^{(0)}\right)^{-1} = e^{-in\gamma_5\frac{\theta}{2}}\left(\gamma^\mu p_\mu - m_D\right)^{-1}\gamma^i e^{in\gamma_5\frac{\theta}{2}}, \quad (3.21c)$$

we obtain

$$\begin{aligned}\mathcal{N}_1 &= \frac{-2\epsilon_{ijk}}{6(2\pi)^2 i}\int \mathrm{d}^3p\int \mathrm{d}p_4\;\mathrm{Tr}_D\left[\frac{1}{\not p - m_D}\gamma^i\frac{1}{\not p - m_D}\gamma^j\frac{1}{\not p - m_D}\gamma^k\frac{1}{\not p - m_D}i\gamma^0\frac{1}{\not p - m_D}(-inm_D\gamma^5)\right] \\ &= \frac{-2n\epsilon_{ijk}m_D^2}{6(2\pi)^2}\mathrm{Tr}_D\left[i\gamma^5\gamma^i\gamma^j\gamma^k\gamma^0\right]\int \mathrm{d}^3p\int_{-\infty}^{\infty}\mathrm{d}p_4\frac{1}{(-p^2+m_D^2)^3}. \quad (3.22)\end{aligned}$$

Here, we can easily perform the $p_4$- and $\boldsymbol{p}$-integrations such that

$$\int \mathrm{d}^3p\int_{-\infty}^{\infty}\mathrm{d}p_4\frac{1}{(p_4^2+\boldsymbol{p}^2+m_D^2)^3} = \frac{3\pi}{8}\times(4\pi)\int_0^\infty \mathrm{d}|\boldsymbol{p}|\frac{\boldsymbol{p}^2}{(\boldsymbol{p}^2+m_D^2)^{5/2}} = \frac{(2\pi)^2}{8m_D^2}. \quad (3.23)$$

Together with $\mathrm{Tr}_D\left[i\gamma^5\gamma^i\gamma^j\gamma^k\gamma^0\right] = 4\epsilon^{ijk}$ and $\epsilon_{ijk}\epsilon^{ijk}=6$, the topological invariant ends up as

$$\mathcal{N}_1 = \frac{-2nm_D^2}{6(2\pi)^2}\times(4\epsilon_{ijk}\epsilon^{ijk})\times\frac{(2\pi)^2}{8m_D^2} = -n. \quad (3.24)$$

Therefore, it turns out that this calculation actually reproduces the well-known result given by Ref. [43] without solving the Dirac equation.[6]

From this argument, one may predict the number of the zeromodes arising from the electron sector in the presence of the $Z$-string background (2.18). Because it obtains the Dirac mass from the Yukawa coupling with the SM Higgs field in Eq. (2.1), the Green's function is the same as Eq. (3.17) with $m_D$ replaced by $h'v_\Phi/\sqrt{2}$. Thus, the number of the zeromodes is $\mathcal{N}_1 = -n$ for the $Z$-string of the vorticity $n$.

[6] We comment on the relation between the above result and the bulk topological invariant introduced in Appendix B. The former can be rewritten as $\mathcal{N}_1 = |\mathcal{N}_{CT}|$, from which it is clear that the number of the zeromodes $\mathcal{N}_1$ is related with the bulk topological invariant $\mathcal{N}_{CT}$ (Eq. (B.6)) protected by parity symmetry. Therefore, the vortex zeromode is understood via the bulk-edge correspondence, namely, the existence of the edge modes are ensured by the non-vanishing bulk topological invariant $\mathcal{N}_{CT}$. Note that, $\mathcal{N}_1$ itself does not depend on the sign of the Dirac mass $m_D$.

### 3.3 Application 1: Type-II seesaw model

In this subsection, we evaluate the topological invariant (3.16) for the type-II seesaw model. In Subsection 3.3.1, It turns out that the topological invariant for the neutrino sector entails a non-vanishing $\mathcal{N}_1$. This implies that the net topological invariant (3.7) vanishes due to the cancellation of the contributions from the electron and neutrino sectors. Therefore, the $Z$-string is not stabilized. To see this fact from another viewpoint, in Subsection 3.3.2, we explicitly construct non-trivial solutions to the equation of motion for the left-handed neutrino field and show that indeed there exist such zeromode solutions corresponding to two Majorana-Weyl fermions.

#### 3.3.1 Calculation of topological invariant

As a more non-trivial example, we consider the type-II seesaw model for neutrinos. Its Lagrangian is given by

$$\mathcal{L}_{\text{Type-II}} = \mathcal{L}_{\text{SM}}\big|_{\lambda\to 0} + \text{Tr}\left[(D^\mu\Delta)(D_\mu\Delta)\right] + V(\Phi,\Delta) - \left(\frac{1}{2}(y_M)_{ij}(\ell_i^c)^\dagger i\tau^2\Delta\ell_j + \text{h.c.}\right), \tag{3.25}$$

where $\mathcal{L}_{\text{SM}}\big|_{\lambda\to 0}$ is the SM Lagrangian given in Eq. (2.1) without the potential term proportional to $\lambda$ and Eq. (2.2), respectively, and $V(\Phi,\Delta)$ denotes the potential for the $SU(2)$-doublet Higgs field $\Phi$ and the $SU(2)$-triplet Higgs field $\Delta = \Delta^a\tau^a/2$ with $\tau^a$ being the Pauli matrices ($a = 1,2,3$). The triplet-Higgs field has $U(1)_Y$ hypercharge $Y = 2$, so that the covariant derivative acting on it is given by $D_\mu\Delta = \partial_\mu\Delta - ig[\frac{\tau^a}{2},\Delta]W_\mu^a - ig'B_\mu$ and the potential for the Higgs fields can contain $V(\Phi,\Delta) \ni \Phi^T i\tau^2\Delta^\dagger\Phi + \text{h.c.}$. In this model, the left-handed lepton doublet fields $\ell_i$ ($i = 1,2,3$) couple to the triplet Higgs field as Majorana Yukawa interactions $(y_M)_{ij}$, where $\ell^c$ is the charge conjugation of the lepton doublet fields, $\ell^c = i\sigma^2\ell^*$. Hereafter, for simplicity, we consider the one-flavor case ($i = 1$) in which Yukawa coupling matrices are reduced to single coupling constants.

Spontaneous symmetry breaking is realized when the $SU(2)$ double- and triplet-Higgs fields take the vacuum expectation values such that

$$\langle\Phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v_\Phi\end{pmatrix}, \qquad \langle\Delta\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 0\\ v_\Delta & 0\end{pmatrix}. \tag{3.26}$$

On this vacuum, the Majorana Yukawa terms in Eq. (3.25) turn to the Majorana-mass terms for the left-handed neutrino

$$\frac{M_L}{2}(\nu_L^c)^\dagger\nu_L + \frac{M_L^*}{2}\nu_L^\dagger\nu_L^c, \tag{3.27}$$

with $M_L = y_M v_\Delta/\sqrt{2}$.

Now, the $Z$-string ansatz is given by the Higgs double and $Z$ fields in (2.18) and the new triplet field $\Delta$,

$$\Delta = \frac{v_\Delta}{\sqrt{2}}\begin{pmatrix}0 & 0\\ f_\Delta(r)e^{2in\theta} & 0\end{pmatrix}, \tag{3.28}$$

where the functions $f(r)$ and $\zeta(r)$ in (2.18) and $f_\Delta(r)$ in (3.28) should be determined by solving the EOMs from the Lagrangian (3.25). Note that the vorticity of the triplet Higgs field is given by $2n$ which reflects its hypercharge $Y = 2$. This can be understood also from the term $\Phi^T i\tau^2 \Delta^\dagger \Phi + \text{h.c.}$ in $V(\Phi, \Delta)$. Therefore, in the $Z$-string background with vorticity $n$, the Majorana mass equips vorticity $2n$.

Thanks to the topological nature, we can turn off the gauge and Yukawa interactions. Thus, the non-interacting Lagrangian for the left-handed neutrino sector reads

$$\begin{aligned}\mathcal{L}_{\text{lh}\nu} &= \nu_L^\dagger \bar{\sigma}^\mu i\partial_\mu \nu_L - \frac{M_L}{2}\left(e^{i(2n)\theta}(\nu_L^c)^\dagger \nu_L + e^{-i(2n)\theta}\nu_L^\dagger \nu_L^c\right) \\ &= \frac{1}{2}\overline{N}_L \gamma^\mu i\partial_\mu N_L - \frac{M_L}{2}\overline{N}_L e^{-i(2n)\theta\gamma^5} N_L\,, \end{aligned} \tag{3.29}$$

where we have introduced the Nambu-Gorkov (four-component) spinors

$$N_L = \begin{pmatrix} \nu_L \\ \nu_L^c \end{pmatrix}\,, \quad \overline{N}_L \equiv -N_L^T i\gamma^2\gamma^0\,. \tag{3.30}$$

Now, one can see that the Lagrangian looks similar to that of the Dirac fermion coupled with the ANO vortex. By doing the same calculation given in the last subsection, we obtain

$$\mathcal{N}_1 = \frac{1}{2} \times (2n) = n\,. \tag{3.31}$$

Note that in general the net degrees of freedom of the Nambu-Gorkov spinor are a half of a Dirac spinor, so that $\mathcal{N}_1$ gets a factor $1/2$ compared to those of the Dirac fermion. This means that a minimal unit localized on the vortex is a half of a single Weyl fermion i.e., a Majorana-Weyl fermion in two dimensions. However, in the type-II seesaw model (3.25), because the left-handed Majorana mass is originated from the VEV of the triplet field $\Delta$ whose $U(1)_Y$ hypercharge is 2, the vorticity of the left-handed Majorana mass has to be $2n$ under a $Z$-string background of the vorticity $n$. Therefore, the topological invariant (3.31) yields $-1$ for $n = 1$. It implies that inside the $Z$-string, there are always two Majorana-Weyl fermions which form a single Weyl fermion.

On the other hand, the electron sector also gives a single Weyl fermion with the opposite chirality. Together with the neutrino sector, the net Dirac index (3.7) is found to be zero as $\mathcal{N}_1 = n_+ - n_- = 0$ with $n_+ = 1$ (electron) and $n_- = 1$ (neutrino) for the $Z$-string in the type-II seesaw model. The Dirac index in the quark sector also vanishes due to the same reason as the cancellation between zeromodes in up and down sectors. Consequently, the $Z$-string is not stabilized by fermion zeromodes in the type-II seesaw model.[7]

[7]Here, we comment on the "bulk-edge correspondence" in the current model setup. As discussed in Appendix B, one can define the bulk topological invariant $\mathcal{N}_K$ where $K$ is an appropriate operation acing on spinor space associated to a discrete symmetry the system has. See Eq. (B.1). The type-II seesaw model is invariant under time-reversal ($T$) symmetry (or $CP$ symmetry). Then, the topological quantity (3.31) is related with $N_T$ such that

$$\mathcal{N}_1 = |\mathcal{N}_T| \times 2n\,. \tag{3.32}$$

The bulk topological invariant $\mathcal{N}_T$ is explicitly calculated in Eq. (B.12). This relation ensures the existence of zeromodes on the $Z$-string as edge modes in terms of the bulk topology.

### 3.3.2 Explicit solution

This emergence of the Majorana-Weyl fermion on the $Z$-string can be verified explicitly by solving the Dirac equation. When we ignore the gauge coupling and the Yukawa coupling with the Higgs doublet, the Dirac equation for $\nu_L$ in the presence of the $Z$-string (2.18) and (3.28) reads

$$i\bar{\sigma}^\mu \partial_\mu \nu_L = M_L f_\Delta(r) e^{-2ni\theta} \nu_L^c \,. \tag{3.33}$$

In the cylindrical coordinate $(t, r, \theta, z)$, where the vortex core is placed at $r = \sqrt{x^2+y^2} = 0$, Eq. (3.33) is written as

$$\left[ i \begin{pmatrix} 0 & e^{-i\theta} \\ e^{i\theta} & 0 \end{pmatrix} \partial_r + i \begin{pmatrix} 0 & -ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix} \frac{1}{r} \partial_\theta + i(-\partial_0 + \sigma^3 \partial_z) \right] \nu_L = -M_L f_\Delta(r) e^{-2ni\theta} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \nu_L^* \,. \tag{3.34}$$

From this, it is obvious that the upper (lower) component corresponds to the left-mover (right-mover) mode traveling on the vortex since the $z$-$t$ direction in the Dirac equation is easily solved by multiplying a factor $e^{i\omega(z\pm t)}$ with an arbitrary constant $\omega$.

To solve Eq. (3.34), let us first take $n = 1/2$. This value is not allowed in the $Z$-string configuration since the Higgs doublet becomes a non-single-valued field $\propto e^{i\theta/2}$. However, it is noteworthy to consider that case to demonstrate that the Dirac equation provides a single Majorana-Weyl zeromode. Focusing on $\omega = 0$, the mode expansion for $\nu_L$,

$$\nu_L = \sum_{m=-\infty}^{\infty} e^{im\theta} \begin{pmatrix} \alpha_m(r) \\ \beta_m(r) \end{pmatrix} , \tag{3.35}$$

reduces the Dirac equation into two sets of equations

$$\alpha_m' - \frac{m}{r}\alpha_m = -iM_L f_\Delta(r) \alpha_{-2-m}^* \,, \tag{3.36a}$$

$$\alpha_{-2-m}' + \frac{2+m}{r}\alpha_{-2-m} = -iM_L f_\Delta(r) \alpha_m^* \,, \tag{3.36b}$$

$$\beta_m' + \frac{m}{r}\beta_m = iM_L f_\Delta(r) \beta_{-m}^* \,, \tag{3.36c}$$

$$\beta_{-m}' - \frac{m}{r}\beta_{-m} = iM_L f_\Delta(r) \beta_m^* \,. \tag{3.36d}$$

Let us here concentrate on the case of $m = 0$. Thus, we arrive at

$$\alpha_0' = -iM_L f_\Delta(r) \alpha_{-2}^* \,, \tag{3.37a}$$

$$\alpha_{-2}' + \frac{2}{r}\alpha_{-2} = -iM_L f_\Delta(r) \alpha_0^* \,, \tag{3.37b}$$

$$\beta_0' = iM_L f_\Delta(r) \beta_0^* \,. \tag{3.37c}$$

The first two equations (3.37a) and (3.37b) have no regular solution except for $\alpha_0(r) = \alpha_2(r) = 0$ (see Ref. [80]). On the other hand, the third one (3.37c) can entail a non-trivial solution. To see this, we make a redefinition $\beta_0 \equiv e^{-i\frac{\pi}{4}} \tilde{\beta}_0$ in order to remove the coefficient "$i$" in equation and decompose $\tilde{\beta}_0$ into the real and imaginary parts, $\tilde{\beta}_0 = \tilde{\beta}_{0,R} + i\,\tilde{\beta}_{0,I}$, for

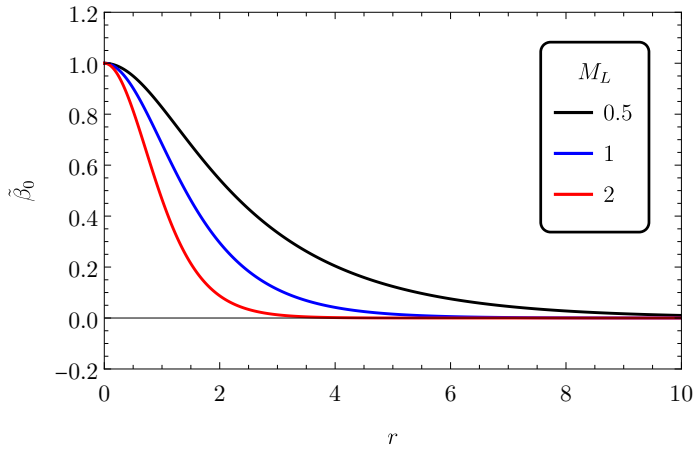



**Figure 6**. The solutions $\tilde{\beta}_0(r)$ to Eq. (3.38a) with $\tilde{\beta}_0(0) = 1$ in cases of $M_L = 0.5$ (black), 1 (blue) and 2 (red). For $f_\Delta(r)$, we use the numerical solution of $f(r)$ with $\beta\,(= M_H^2/M_Z^2) = 1$ and the winding number being 1.

which Eq. (3.37c) reads

$$\tilde{\beta}'_{0,R} = -M_L f_\Delta(r)\tilde{\beta}_{0,R}\,, \tag{3.38a}$$

$$\tilde{\beta}'_{0,I} = M_L f_\Delta(r)\tilde{\beta}_{0,I}\,. \tag{3.38b}$$

Since Eq. (3.38b) does not have any normalizable solution except for $\tilde{\beta}_{0,I} = 0$, these solutions are found to be

$$\tilde{\beta}_0(r) = \tilde{\beta}_{0,R} \propto \exp\left[-M_L \int_0^r \mathrm{d}r'\ f_\Delta(r')\right]. \tag{3.39}$$

This solution is displayed in Fig. 6. Therefore, we obtain a zeromode solution as

$$\nu_L \propto \begin{pmatrix} 0 \\ \tilde{\beta}_0(r) \end{pmatrix}, \quad \left(\tilde{\beta}_0^* = \tilde{\beta}_0\right) \tag{3.40}$$

up to an overall constant defined by the normalization condition. Note that this only has the lower component, which is a right-mover fermion traveling on the string. In addition, this solution has only one real degree of freedom, and hence behaves as a Majorana particle in two dimensions. Therefore, we have obtained one Majorana-Weyl zeromode.

We now come back to the case of the $Z$-string with integer $n$. For the single $Z$-string, $n = 1$, the equation of motion (3.34) yields

$$\alpha'_m - \frac{m}{r}\alpha_m = -iM_L f_\Delta \alpha^*_{-3-m}\,, \tag{3.41a}$$

$$\alpha'_{-3-m} + \frac{3+m}{r}\alpha_{3-m} = -iM_L f_\Delta \alpha^*_m\,, \tag{3.41b}$$

$$\beta'_m + \frac{m}{r}\beta_m = iM_L f_\Delta \beta^*_{-1-m}\,, \tag{3.41c}$$

$$\beta'_{-1-m} + \frac{-1-m}{r}\beta_{-1-m} = iM_L f_\Delta \beta^*_m\,. \tag{3.41d}$$

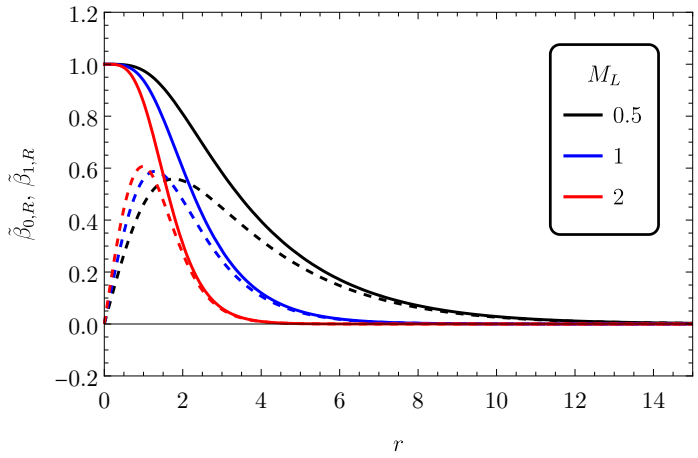



**Figure 7**. The solutions $\tilde{\beta}_{0,R}(r)$ (solid) and $\tilde{\beta}_{1,R}(r)$ (dashed) to Eqs. (3.42a) with $\tilde{\beta}_0(0) = 1$ in cases of $M_L = 0.5$ (black), 1 (blue) and 2 (red). For $f_\Delta(r)$, we use the numerical solution of $f(r)$ with $\beta\,(= M_H^2/M_Z^2) = 1$ and the winding number being 2.

Although there is no regular solution for $\alpha_m$ again, $\beta_m$ has a regular solution for $\beta_0$ and $\beta_1$. To see this, we again redefine as $\beta_0 \equiv e^{-i\frac{\pi}{4}}\tilde{\beta}_0$ and $\beta_1 \equiv e^{-i\frac{\pi}{4}}\tilde{\beta}_1$, and decompose them into real and imaginary parts. Then we obtain two pairs of coupled real equations,

$$\begin{cases} \tilde{\beta}'_{0,R} = -M_L f_\Delta \tilde{\beta}_{1,R}\,, \\ \tilde{\beta}'_{1,R} - \dfrac{1}{r}\tilde{\beta}_{1,R} = -M_L f_\Delta \tilde{\beta}_{0,R}\,, \end{cases} \tag{3.42a}$$

$$\begin{cases} \tilde{\beta}'_{0,I} = M_L f_\Delta \tilde{\beta}_{1,I}\,, \\ \tilde{\beta}'_{1,I} - \dfrac{1}{r}\tilde{\beta}_{1,I} = M_L f_\Delta \tilde{\beta}_{0,I}\,. \end{cases} \tag{3.42b}$$

The numerical solutions to Eqs. (3.42a) are shown in Fig. 7. Obviously, Eqs. (3.42b) have almost the same solutions as Eqs. (3.42a), but the relative sign between $\tilde{\beta}_{0,I}$ and $\tilde{\beta}_{1,I}$ is opposite to that of $\tilde{\beta}_{0,R}$ and $\tilde{\beta}_{1,R}$. Therefore, denoting particular *real* solutions of Eqs. (3.42a) by $\tilde{\beta}_{0,R} = \chi_0(r)$ and $\tilde{\beta}_{1,R} = \chi_1(r)$, we can express the general zeromode solutions as

$$\begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \end{pmatrix} = C_1 \begin{pmatrix} \chi_0(r) \\ \chi_1(r) \end{pmatrix} + i\, C_2 \begin{pmatrix} \chi_0(r) \\ -\chi_1(r) \end{pmatrix} \tag{3.43}$$

with $C_1, C_2$ being *real* arbitrary constants. This is a superposition of two independent real solutions, which are two Majorana-Weyl fermions with the same chirality (right-mover) in two dimensions and equivalent to a single Weyl fermion. Therefore, we have the index of the zeromodes $\mathcal{N}_1 = 1$, which is consistent with the result in Eq. (3.31).

### 3.4 Application 2: Type-I seesaw model

In the same manner as Subsection 3.3, we calculate the topological invariant for the neutrino sector in the case of the type-I seesaw model. The neutrino sector again produces a non-zero topological invariant. As is the case in the type-II seesaw model, the net topological

invariant (3.7) in the type-I seesaw model becomes zero. Hence, the present of fermion zeromodes cannot stabilize the $Z$-string. The fact that there are non-trivial zeromode solutions in the neutrino sector has been discussed in Ref. [80]. In Subsection 3.4.2, we review the zeromode solutions.

### 3.4.1 Calculation of topological invariant

The next example is the type-I seesaw model whose Lagrangian is given by

$$\mathcal{L}_{\text{Type-I}} = \mathcal{L}_{\text{SM}} + \mathcal{L}_{\nu_R} \,. \tag{3.44}$$

See Eqs. (2.1) and (2.2) for their explicit forms. Let us consider the case in which the Dirac mass is originated from the Dirac-Yukawa interaction $\sim (hv_\Phi/\sqrt{2})f(r)e^{in\theta}$ with the winding number $n$. In the bulk far from the center of the vortex, the Lagrangian for the neutrino sector becomes

$$\mathcal{L}_\nu = \nu_L^\dagger \bar{\sigma}^\mu i\partial_\mu \nu_L + \nu_R^\dagger \sigma^\mu i\partial_\mu \nu_R - \left(\frac{M_R}{2}(\nu_R^c)^\dagger \nu_R + \text{h.c.}\right) - m_D\left(\nu_R^\dagger e^{in\theta}\nu_L + \nu_L^\dagger e^{-in\theta}\nu_R\right) , \tag{3.45}$$

where $m_D = hv_\Phi/\sqrt{2}$. It is assumed here that the right-handed Majorana mass has no vorticity. When we write the mass term in the mass-matrix form as

$$\frac{1}{2}\begin{pmatrix}\nu_L^\dagger & (\nu_R^c)^\dagger\end{pmatrix}\begin{pmatrix}0 & m_D e^{-in\theta} \\ m_D e^{-in\theta} & M_R\end{pmatrix}\begin{pmatrix}\nu_L^c \\ \nu_R\end{pmatrix} + \text{h.c.}\,, \tag{3.46}$$

it is clear that we can vary the Majorana mass $M_R$ from finite values to zero without closing the massgap (in other words, keeping the determinant of the mass matrix as non-zero). Therefore, the topological invariant $\mathcal{N}_1$ in this case is proved to be equal (with the opposite sign) to that in the Dirac case in the presence of the ANO vortex, i.e.,

$$\mathcal{N}_1 = n \,. \tag{3.47}$$

This is consistent with the result in Ref. [80], in which a Weyl zeromode is obtained for $n = 1$.

### 3.4.2 Explicit solution (review)

Let us obtain the zeromode explicitly, which have been argued in Refs. [80]. The equations of motion for the neutrino fields read from Eqs. (2.1) and (2.2) as

$$i\bar{\sigma}^\mu D_\mu \nu_L = h\phi_d^* \nu_R \,, \tag{3.48}$$

$$i\sigma^\mu \partial_\mu \nu_R = h\phi_d \nu_L + M_R(\nu_R)^c \,, \tag{3.49}$$

where we have supposed the $Z$-string ansatz (2.18), i.e. the covariant derivative acting on the left-handed neutrino field is given by $D_\mu \nu_L = (\partial_\mu - iqZ_\mu)\nu_L$ with Eq. (2.18).

The Dirac equations are explicitly rewritten as

$$\left[i\sigma^r\partial_r + i\sigma^\theta \frac{1}{r}\left(\partial_\theta + in\zeta(r)\right) + i(-\partial_0 + \sigma^3\partial_z)\right]\nu_L = -m_D f(r)e^{-in\theta}\nu_R \,, \tag{3.50}$$

$$\left[i\sigma^r\partial_r + i\sigma^\theta \frac{1}{r}\partial_\theta + i(\partial_0 + \sigma^3\partial_z)\right]\nu_R = m_D f(r)e^{in\theta}\nu_L - M_R\, i\sigma^2 \nu_R^* \,. \tag{3.51}$$

From this, it is obvious that the lower (upper) component of $\nu_R$ corresponds to the left-mover (right-mover) mode traveling on the vortex, which are opposite to those of $\nu_L$. The $z$-$t$ direction is easily solved by multiplying a factor $e^{i\omega(z\pm t)}$ with an arbitrary constant $\omega$. Again, we focus on the zero-energy solution $\omega = 0$.

In order to solve Eqs. (3.50) and (3.51), we decompose them into the differential equations for each component of the spinor fields,

$$\nu_L = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \qquad \nu_R = \begin{pmatrix} \gamma \\ \delta \end{pmatrix}. \tag{3.52}$$

Plugging Eq. (3.52) into Eqs. (3.50) and (3.51), we obtain the equations of motion for each component of the neutrino fields. Those of the left-handed neutrino field read

$$ie^{i\theta}\left[\partial_r - \frac{1}{r}\Big(-i\partial_\theta + n\zeta(r)\Big)\right]\alpha = \Big(-m_D f(r) e^{-in\theta}\Big)\,\delta\,, \tag{3.53}$$

$$ie^{-i\theta}\left[\partial_r + \frac{1}{r}\Big(-i\partial_\theta + n\zeta(r)\Big)\right]\beta = \Big(-m_D f(r) e^{-in\theta}\Big)\,\gamma\,, \tag{3.54}$$

while for the right-handed neutrino field, we have

$$-ie^{i\theta}\left[\partial_r + \frac{i}{r}\partial_\theta\right]\gamma + M_R\gamma^* = \Big(-m_D f(r) e^{in\theta}\Big)\,\beta\,, \tag{3.55}$$

$$-ie^{-i\theta}\left[\partial_r - \frac{i}{r}\partial_\theta\right]\delta - M_R\delta^* = \Big(-m_D f(r) e^{in\theta}\Big)\,\alpha\,. \tag{3.56}$$

To solve these coupled differential equations, we perform the mode expansions for each component as

$$\alpha(t,r,\theta,z) = \sum_{m=-\infty}^{\infty} \alpha_m(r) e^{im\theta}\,, \tag{3.57a}$$

$$\beta(t,r,\theta,z) = -\,i\sum_{m=-\infty}^{\infty} \beta_m(r) e^{im\theta}\,, \tag{3.57b}$$

$$\gamma(t,r,\theta,z) = \sum_{m=-\infty}^{\infty} \gamma_m(r) e^{im\theta}\,, \tag{3.57c}$$

$$\delta(t,r,\theta,z) = -\,i\sum_{m=-\infty}^{\infty} \delta_m(r) e^{im\theta}\,, \tag{3.57d}$$

and then obtain the equations of motion for $\alpha_m(r)$, $\beta_m(r)$, $\gamma_m(r)$ and $\delta_m(r)$.

Hence, we look for solutions to the following recursive equations with the winding

number $n = 1$:

$$\alpha'_m - \frac{(m+\zeta)}{r}\alpha_m = m_D f \delta_{m+2}\,, \tag{3.58a}$$

$$\beta'_m + \frac{(m+\zeta)}{r}\beta_m = -\, m_D f \gamma_m\,, \tag{3.58b}$$

$$\gamma'_m - \frac{m}{r}\gamma_m + iM_R\gamma^*_{-1-m} = -\, m_D f \beta_m\,, \tag{3.58c}$$

$$\delta'_m + \frac{m}{r}\delta_m + iM_R\delta^*_{1-m} = m_D f \alpha_{m-2}\,. \tag{3.58d}$$

In the $M_R = 0$ limit, by replacing $m$ with $m+2$ in Eq. (3.58d), these differential equations become closed forms for $\alpha_m$, $\beta_m$, $\gamma_m$ and $\delta_{m+2}$ which are the standard equations for zeromodes of the Dirac particle.

We now focus on Eqs. (3.58b) and (3.58c). The replacement of $m$ with $-1-m$ gives

$$\beta'_{-1-m} + \frac{-1-m+\zeta}{r}\beta_{-1-m} = -\, m_D f \gamma_{-1-m}\,, \tag{3.59}$$

$$\gamma'_{-1-m} - \frac{-1-m}{r}\gamma_{-1-m} + iM_R\gamma^*_m = -\, m_D f \beta_{-1-m}\,. \tag{3.60}$$

Thus, the four equations, Eqs. (3.58b), (3.58c), (3.59), and (3.60), are closed. We eliminate $\gamma_m$ and $\gamma_{-1-m}$ from Eqs. (3.58c) and (3.60) by using Eqs. (3.58b) and (3.59) and then have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{1}{m_D f}\left(\tilde\beta'_m + \frac{m+\zeta}{r}\tilde\beta_m\right)\right] \\
&- \frac{m}{r}\frac{1}{m_D f}\left(\tilde\beta'_m + \frac{m+\zeta}{r}\tilde\beta_m\right) - M_R\left(\tilde\beta'^*_{-1-m} + \frac{-1-m+\zeta}{r}\tilde\beta^*_{-1-m}\right) = m_D f \tilde\beta_m\,,
\end{aligned} \tag{3.61}$$

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{1}{m_D f}\left(\tilde\beta'_{-1-m} + \frac{-1-m+\zeta}{r}\tilde\beta_{-1-m}\right)\right] \\
&+ \frac{1+m}{r}\frac{1}{m_D f}\left(\tilde\beta'_{-1-m} + \frac{-1-m+\zeta}{r}\tilde\beta_{-1-m}\right) - M_R\left(\tilde\beta'^*_m + \frac{m+\zeta}{r}\tilde\beta^*_m\right) = m_D f \tilde\beta_{-1-m}\,,
\end{aligned} \tag{3.62}$$

where we have again redefined as $\beta_m \equiv e^{-i\frac{\pi}{4}}\tilde\beta_m$ and $\gamma_m \equiv e^{-i\frac{\pi}{4}}\tilde\gamma_m$ for convenience. Note that all coefficients are real, so that we can separate the equations into real and imaginary parts.

Let us here solve the real parts $\tilde\beta_{0,R}$ and $\tilde\beta_{-1,R}$ of the above coupled two equations numerically. For this, it is actually sufficient to deal with the equations for $m = 0$. We fix the background bosonic fields ($Z_\mu$ and $\phi_d$) as the $Z$-string configuration. For various $M_R$, we find that the equations always have solutions for the real parts $\tilde\beta_{0,R}(r)$ and $\tilde\beta_{-1,R}(r)$. Fig. 8 displays these zeromode solutions for $M_R = 0.5$, 5, 50 (in the energy unit $M_Z/\sqrt{2} = 1$). One can see from Fig. 8 particularly that for large $M_R$, the solution becomes broader, namely slowly converges for $r \to \infty$. This is because each of the solutions asymptotically becomes a linear combination of two exponential tails with typical mass scales $M_R$ and $m_D^2/M_R$, as stated in Ref. [80]. For large $M_R$, the former tail ($\sim e^{-M_R r}$) suddenly decay,

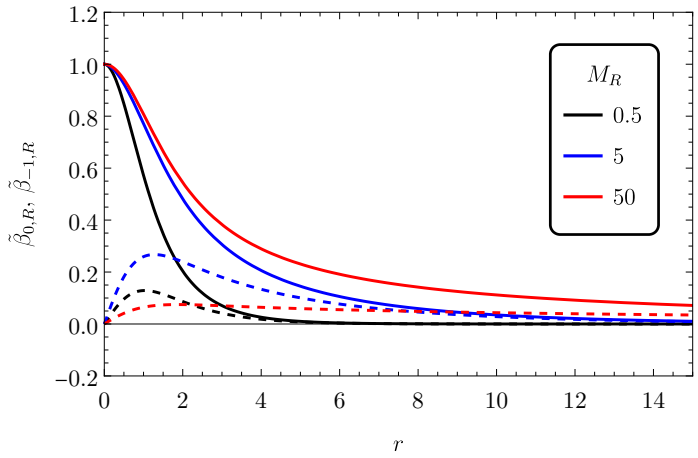



**Figure 8**. Plots for the zeromode solutions $\tilde{\beta}_{0,R}$ and $\tilde{\beta}_{-1,R}$ for $M_R = 0.5$ (black), 5 (blue), and 50 (red) as functions of $r$ in the energy unit $M_Z/\sqrt{2} = 1$. The Dirac mass is taken to be $m_D = 1$ in this unit. The solid and dashed lines stand for $\tilde{\beta}_0(r)$ and $\tilde{\beta}_{-1}(r)$, respectively. For $f(r)$, we take $\beta\,(= M_H^2/M_Z^2) = 1$.

while the latter tail ($\sim e^{-(m_D^2/M_R)r}$) survives even with large $r$, which provides the broad profile of the solutions. As $M_R \to \infty$, the solutions eventually become non-normalizable. This is consistent with the decoupling limit, that is, in the limit of $M_R \to \infty$ the model reduces to the SM without the right-handed neutrino, and hence the neutrino sector in the SM does not allow normalizable zeromodes on the electroweak string.

For the imaginary parts $\tilde{\beta}_{0,I}$ and $\tilde{\beta}_{-1,I}$, we get almost the same solutions as the real ones except for that their relative sign is opposite to that of $\tilde{\beta}_{0,R}$ and $\tilde{\beta}_{-1,R}$. Therefore, repeating the same argument presented in Sec. 3.3.2, it is found that we have obtained the two independent real solutions for $(\tilde{\beta}_0, \tilde{\beta}_{-1})$, corresponding to a right-mover Weyl fermion in two dimensions. Note that $(\tilde{\gamma}_0, \tilde{\gamma}_{-1})$ are not independent from $(\tilde{\beta}_0, \tilde{\beta}_{-1})$ since they are determined by Eqs. (3.58b) and (3.59). In addition, the solutions for $\alpha$ and $\delta$ are not found except for the trivial ones, $\alpha_m = \delta_m = 0$. Therefore, we conclude that the index $\mathcal{N}_1$ in the background of the single $Z$-string is given as $\mathcal{N}_1 = 1$, which is consistent with Eq. (3.47) and Ref. [80].

### 3.5 Application 3: Type-I seesaw model with one $\nu_L$ and two $\nu_R$'s

Now, we can consider a *modified* type-I seesaw model in which there are one left-handed and two right-handed neutrinos. The neutrino sector is extended to be the following Lagrangian

$$\begin{aligned}\mathcal{L}_{\text{neutrino}} = \nu_L^\dagger \bar{\sigma}^\mu i\partial_\mu \nu_L + \sum_{j=1,2} \nu_{R,j}^\dagger\, \sigma^\mu i\partial_\mu\, \nu_{R,j} \\ - \sum_{j,k}\left(\frac{M_{jk}}{2}(\nu_{R,j}^c)^\dagger \nu_{R,k} + \text{h.c.}\right) - \sum_j \left(m_j \nu_{R,j}^\dagger \nu_L + \text{h.c.}\right),\end{aligned} \tag{3.63}$$

where the two right-handed neutrinos are labeled by the index $j$ or $k$. The Majorana and Dirac masses are described by a $2 \times 2$ matrix $M_{jk}$ and a $1 \times 2$ matrix (vector) $m_j$. The

mass term is expressed by the following $3 \times 3$ mass-matrix form:

$$\left(\nu_L^\dagger \ (\nu_{R,1}^c)^\dagger \ (\nu_{R,2}^c)^\dagger\right) \begin{pmatrix} 0 & m_1 & m_2 \\ m_1 & M_{11} & M_{12} \\ m_2 & M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \nu_L^c \\ \nu_{R,1} \\ \nu_{R,2} \end{pmatrix} + \text{h.c.}\,. \tag{3.64}$$

Let us consider the case in which the two Dirac masses $m_1$ and $m_2$ have a vorticity $n$, i.e., $m_j\, e^{in\theta}$ for $j = 1, 2$. We calculate the topological invariant describing the number of the zeromodes, $\mathcal{N}_1$. Without loss of generality, we can diagonalize the Majorana mass matrix $M_{jk}$, leading to the mass matrix

$$\mathcal{M} \equiv \begin{pmatrix} 0 & m_1 e^{-in\theta} & m_2 e^{-in\theta} \\ m_1 e^{-in\theta} & M_1 & 0 \\ m_2 e^{-in\theta} & 0 & M_2 \end{pmatrix}, \tag{3.65}$$

and can take $M_1, M_2 \geq 0$. Because $\det \mathcal{M} \propto -M_1 m_2^2 - M_2 m_1^2$, we put $m_1 = M_2 = 0$ without closing the massgap. Note that we cannot take $M_1 = 0$ at the same time.

The Dirac equations are given as

$$i\bar{\sigma}^\mu \partial_\mu \nu_L = m_2 e^{-in\theta} \nu_{R,2}\,, \tag{3.66a}$$
$$i\sigma^\mu \partial_\mu \nu_{R,1} = M_1 \nu_{R,1}^c\,, \tag{3.66b}$$
$$i\sigma^\mu \partial_\mu \nu_{R,2} = m_2 e^{in\theta} \nu_L\,. \tag{3.66c}$$

From this, it is obvious that $\nu_{R,1}$ is decoupled from the system and $\nu_{R,2}$ and $\nu_L$ form one species of Dirac fermion. Therefore, the topological invariant $\mathcal{N}_1$ is equal to that in the Dirac case in the presence of the ANO vortex (with the opposite sign), i.e.,

$$\mathcal{N}_1 = n\,. \tag{3.67}$$

To obtain the zeromode explicitly, we can use the same argument for the Dirac fermion. See Subsection 3.2.

### 3.6 Application 4: Hybrid of type-I and type-II seesaw models

Let us consider the hybrid of the type-I and type-II seesaw. We concentrate on one generation of the SM fermions and one right-handed neutrino. The Lagrangian is given by a sum of $\mathcal{L}_{\text{Type-II}}$ (3.25) and $\mathcal{L}_{\nu_R}$ (2.2). The $Z$-string background is the same as that in the type-II seesaw model, (2.18) and (3.28). Thanks to the topological nature, we turn off the gauge and Yukawa interactions again, and thus the neutrino sector is described by the following Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{Type-I+Type-II}} = \nu_L^\dagger \bar{\sigma}^\mu i\partial_\mu \nu_L + \nu_R^\dagger \sigma^\mu i\partial_\mu \nu_R - \frac{M_L}{2}\left(e^{i(2n)\theta}(\nu_L^c)^\dagger \nu_L + e^{-i(2n)\theta}\nu_L^\dagger \nu_L^c\right) \\ - \left(\frac{M_R}{2}(\nu_R^c)^\dagger \nu_R + \text{h.c.}\right) - m_D\left(\nu_R^\dagger e^{in\theta}\nu_L + \nu_L^\dagger e^{-in\theta}\nu_R\right), \end{aligned} \tag{3.68}$$

which leads to the mass matrix of the neutrinos,

$$\frac{1}{2}\begin{pmatrix}\nu_L^\dagger & (\nu_R^c)^\dagger\end{pmatrix}\begin{pmatrix}M_L e^{-2ni\theta} & m_D e^{-in\theta}\\ m_D e^{-in\theta} & M_R\end{pmatrix}\begin{pmatrix}\nu_L^c\\ \nu_R\end{pmatrix} + \text{h.c.}\,. \tag{3.69}$$

Note that the left-handed Majorana and the Dirac masses has the winding phase, while the right-handed Majorana mass does not.

In contrast to the case given in Subsection 3.4, now we can turn off neither of $M_L$ nor $M_R$, and the system can be classified into two phases corresponding to

$$M_L M_R - m_D^2 \gtrless 0\,. \tag{3.70}$$

These are topologically distinct phases since in order to move from one phase into the other phase, one needs to get across the phase boundary on which the mass gap is closed.

Let us look at what value of $\mathcal{N}_1$ each topological phase has. For $M_L M_R - m_D^2 > 0$, one can take $m_D = 0$ and $M_L, M_R > 0$ as a representative point. In this case, the right-handed neutrino $\nu_R$ is completely decoupled from the SM sector, and the setup reduces to the case of Type-II seesaw model (discussed in Subsection 3.3), resulting in $\mathcal{N}_1 = n$. On the other hand, for $M_L M_R - m_D^2 < 0$, one can take $M_L = M_R = 0$ without loss of generality, and hence we have the same setup as the usual Dirac fermion in the background of the ANO string, resulting in $\mathcal{N}_1 = n$. Therefore, both cases have the same value $\mathcal{N}_1 = n$.

### 3.7 Comment on the stability of $Z$-string

So far, we have shown that in any seesaw model of the Type-I, Type-II, their hybrid, and the Type-I with different flavors, we have the same number of neutrino zeromodes on the $Z$-string background, $\mathcal{N}_1 = n$ with $n$ the vorticity. Accompanying with the electron sector, we have the total number of zeromodes in the lepton sector (for one generation) as

$$\mathcal{N}_1 = n\ (\text{neutrino}) - n\ (\text{electron}) = 0\,, \tag{3.71}$$

which means that the $1+1$-dimensional effective theory consisting of the zeromodes on the string is not a chiral theory but a vector-like one with the same degrees of freedom of $n$ (two-component) Dirac fermions. As stated above, the zero Dirac index on a string indicates the instability of the string itself because a pair of the fermion zeromodes with opposite chiralities forms a massive mode lowering the energy under the perturbative deformation of the string. Therefore, what we have found indicates that the $Z$-string cannot be stabilized by any extension of the neutrino sector in the SM.

On the other hand, in Ref. [81], it is shown that in the same setup as the last subsection (hybrid of type-I and -II), the number of the zeromodes from the neutrino sector can vanish, which means that the Dirac index from the neutrino and electron sectors becomes non-zero in total. It is further argued that this non-zero index can stabilize the $Z$-string. At a first look, our result is inconsistent with the result in Ref. [81]. We here give some comments on the apparent contradiction.

Firstly, one should note that this model admits another string solution when one imposes a global $U(1)_L$ ($L$ denotes the lepton number) symmetry associated with the

phase rotation of the SM leptons and the triplet $\Delta$. This symmetry prohibits the cubic term $\Phi^T i\tau^2\Delta^\dagger\Phi$ in the potential. The $U(1)_L$ symmetry is spontaneously broken by the VEV $v_\Delta$, which leads to a global string solution described by the following ansatz

$$\Phi = \frac{v_\Phi}{\sqrt{2}}\begin{pmatrix}0\\1\end{pmatrix}, \qquad \Delta = \frac{v_\Delta}{\sqrt{2}}\begin{pmatrix}0 & 0\\ f_\Delta(r)e^{i\ell\theta} & 0\end{pmatrix}, \qquad Z_\theta = 0\,, \tag{3.72}$$

with the vorticity $\ell \in \mathbb{Z}$ of the global string. Since the gauge coupling does not matter, the Dirac equation for $\nu_L$ in the background of this string is given as

$$i\bar\sigma^\mu\partial_\mu\nu_L = M_L f_\Delta(r)e^{-i\ell\theta}\nu_L^c\,, \tag{3.73}$$

which is equivalent to the Dirac equation for the $Z$-string (3.33) with $\ell$ replaced by $2n$. Note that, in this case, the minimal vorticity is $|\ell| = 1$ instead of $|\ell| = 2$ because the Higgs doublet is not charged under the $U(1)_L$ symmetry and is a single-value field for arbitrary $\ell$. Taking the minimal value $\ell = 1$, which is equivalent to substituting $n = 1/2$ into Eq. (3.33), then we again get the unique solution (3.40). Therefore the single global $U(1)_L$ string provides a single Majorana-Weyl fermion, or equivalently, a half of a Weyl fermion.

Next, let us point out that, in the study in Ref. [81], they consider a background described as follows

$$\Phi \sim \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v_\Phi e^{i\theta}\end{pmatrix}, \qquad \Delta \sim \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 0\\ v_\Delta & 0\end{pmatrix}, \qquad Z_\theta \sim \frac{1}{q}\,, \tag{3.74}$$

which is different from our $Z$-string ansatz (2.18) and (3.28) in the sense that the triplet Higgs $\Delta$ does not contain the winding phase. Instead, this is nothing but a superposition of the single $Z$-string and two global $U(1)_L$ strings with the phase of $\Delta$ anti-rotated, i.e., the vorticity for $Z$-string is 1, while that for the global $U(1)_L$ string is $\ell = -2$. In fact, in Ref. [81] it is mentioned that this configuration gives a logarithmic divergent tension from the covariant derivative $D_\mu\Delta$ to be $4\times 2\pi v_\Delta^2\log\Lambda$ with $\Lambda$ being the infrared cutoff, which is a feature of the global string.

In this case, thanks to the topological nature, the Dirac index on the string is given by the sum of contributions from the single $Z$-string and the two global $U(1)_L$ strings,[8]

$$\begin{aligned}\mathcal{N}_1 = &\underbrace{1}_{\nu\text{ zeromodes on }Z\text{-string}} - \underbrace{1}_{e\text{ zeromodes on }Z\text{-string}} - \underbrace{2\times\frac{1}{2}}_{\nu\text{ zeromodes from }U(1)_L\text{ string}}\\ &= -1\,.\end{aligned} \tag{3.75}$$

This is nothing but the value found in Ref. [81]. This states that this superposed configuration is in the same topological sector with the two global $U(1)_L$ strings without $Z$ flux, which implies that it evolves into the latter to reduce the energy. Therefore, it is not shown

[8] This is understood by considering a case in which these strings are separated far enough so that each of them is regarded as a single string.

in Ref. [81] that the $Z$-string itself is stabilized by the fermion zeromodes. In addition, because the global $U(1)_L$ symmetry cannot be an exact symmetry in the phenomenological viewpoint, the global $U(1)_L$ string is not stable in the phenomenologically viable setup. To conclude, we cannot expect any stable string object in the SM with the extensions in the neutrino sector.

## 4 Discussion and Conclusions

In this paper, we have applied the notion of topological invariants to various extensions of the SM, focusing on the emergence of zeromodes in $Z$-strings. In Sec. 2 we first have reviewed the electroweak string (or $Z$-string) solution in the SM. In Sec. 3 we have introduced a topological invariant defined in the momentum space and real space, that counts the number of fermion zeromodes along one-dimensional topological defects. While these solutions can be obtained explicitly by numerical methods, the language of topological invariant makes it easier to understand why and how these zeromodes appear. We have applied this formalism to several models with extensions in the neutrino sectors and have found that most seesaw models (type-I, type-II, their hybrid, and modified type-I), the number of the zeromode from the neutrino sector on the $Z$-string is always $n$ with $n$ the vorticity of the $Z$-string. We have also critically examined claims in the literature [81] from the viewpoint of these topological invariants. While it was claimed in the literature that the fermion zeromodes help to stabilize the $Z$-string, we pointed out that the stabilization is not the case: the topological object they consider is in fact the superposition of $Z$-string and a global string corresponding to $U(1)_L$, and the former is not stabilized by the zeromodes. To conclude, there remains a huge room to apply topological arguments to various models in high-energy physics and consider cosmological implications.

As the simplest extensions of the SM, the two Higgs doublet models (2HDMs) have been studied extensively. According to the fermion sectors, these models are further classified into several types. Similarly to the SM, the 2HDMs also admit $Z$-strings [90–92], which are non-topological and are unstable. The fermion zeromodes in 2HDMs were not studied yet, and they may or may not appear in various ways depending on the types of 2HDMs. A possibility of the stabilization by fermion zeromodes in this case is one of future directions. In addition, the 2HDMs admit topological $Z$-strings which are attached by domain walls [93–95] as axion string, and can support topological Nambu monopoles [96–98]. The stabilization of a non-topological $Z$-string by splitting into a pair of topological $Z$-strings was also proposed in Ref. [65]. The fermion zeromodes may enhance the stability region in such a mechanism.

In condensed matter physics, it is known that the number of zeromodes (edge modes) is strongly related with the topology of the bulk momentum space without defects, as the so-called bulk-edge correspondence. Thus, it can be expected that the topological invariant we have studied in this work would be understood in the context of the bulk-edge correspondence. This might help us to classify particle physics models in terms of the Altland-Zirnbauer classification [99] for symmetry-protected topological phases in condensed matter.

## Acknowledgements

The authors are grateful to Ken Shiozaki for helpful comments. This work is supported in part by the WPI program "Sustainability with Knotted Chiral Meta Matter (SKCM$^2$)" at Hiroshima University (M. E. and M. N.), JSPS KAKENHI Grant Numbers JP22H01221 (M. E. and M. N.), JP21J01117 (Y. H.), JP23K17687 (R. J.), JP23K19048 (R. J.), and the National Science Foundation of China (NSFC) under Grant No. 12205116 (M. Y.). This work is also supported by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy - EXC 2121 Quantum Universe - 390833306 and the Seeds Funding of Jilin University.

## A Notation

### A.1 Representation of Dirac matrices

We clarify the representation of the Dirac matrices. In the chiral representation, the Dirac matrices in Cartesian coordinates are given by

$$\gamma^\mu = \begin{pmatrix} \mathbf{0} & \sigma^\mu \\ \bar{\sigma}^\mu & \mathbf{0} \end{pmatrix}, \qquad \sigma^\mu = (\mathbf{1}, \sigma^i), \qquad \bar{\sigma}^\mu = (\mathbf{1}, -\sigma^i), \tag{A.1}$$

for which the chirality matrix becomes

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} \mathbf{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}, \tag{A.2}$$

and the charge conjugation matrix $C$ reads

$$C = -i\gamma^0\gamma^2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{A.3}$$

Since the $Z$-string configuration is obtained in cylindrical coordinates $(r, \theta, z)$, it is convenient to work in the same coordinates for the neutrino field configurations. More specifically, we intend to write

$$\gamma^0\partial_0 + \gamma^1\partial_1 + \gamma^2\partial_2 + \gamma^3\partial_3 = \gamma^0\partial_0 + \gamma^r\partial_r + \gamma^\theta\frac{1}{r}\partial_\theta + \gamma^3\partial_3\,. \tag{A.4}$$

Here, a representation of the Dirac matrices in cylindrical coordinates is found to be

$$\gamma^r = \begin{pmatrix} 0 & 0 & 0 & e^{-i\theta} \\ 0 & 0 & e^{i\theta} & 0 \\ 0 & -e^{-i\theta} & 0 & 0 \\ -e^{i\theta} & 0 & 0 & 0 \end{pmatrix}, \qquad \gamma^\theta = \begin{pmatrix} 0 & 0 & 0 & -ie^{-i\theta} \\ 0 & 0 & ie^{i\theta} & 0 \\ 0 & ie^{-i\theta} & 0 & 0 \\ -ie^{i\theta} & 0 & 0 & 0 \end{pmatrix}, \tag{A.5}$$

while $\gamma^0$ and $\gamma^3$ are the same as in Eq. (A.1).

For the two-by-two matrices $\sigma^\mu$, we have

$$\sigma^\mu \partial_\mu = \sigma^0 \partial_0 + \sigma^1 \partial_1 + \sigma^2 \partial_2 + \sigma^3 \partial_3 = \sigma^0 \partial_0 + \sigma^r \partial_r + \sigma^\theta \frac{1}{r} \partial_\theta + \sigma^3 \partial_3 \tag{A.6}$$

with

$$\sigma^r = \begin{pmatrix} 0 & e^{-i\theta} \\ e^{i\theta} & 0 \end{pmatrix}, \quad \sigma^\theta = \begin{pmatrix} 0 & -ie^{-i\theta} \\ ie^{i\theta} & 0 \end{pmatrix}. \tag{A.7}$$

# B Bulk momentum space topology in neutrino sector

In this appendix, we consider the neutrino models in light of topology. We first introduce another topological invariant defined in the momentum space of the neutrino fields by closely following Refs. [28, 34, 77].

## B.1 Topological invariant for symmetry-protected topological phases

Let us start by reviewing general prescriptions to define topological invariants characterizing topological phases protected by symmetry of gapped systems. Since topological properties are robust under continuous deformations, we switch off interaction in the system, so that we can specify ourselves into one-particle states. Let $\mathcal{G}$ be a free Green's function with a non-zero mass gap defined on the four-dimensional momentum space $p_\mu$ and $K$ be an operator associated with a discrete symmetry that satisfies $K^2 = 1$ and acts on the Green's function and (one-particle) Hamiltonian. Then, one can define the following quantity [18, 28, 34, 100, 101]:

$$\mathcal{N}_K \equiv \frac{\epsilon_{ijk}}{24\pi^2} \mathrm{Tr} \left[ \int \mathrm{d}^3 p \; K (\mathcal{G} \partial_{p_i} \mathcal{G}^{-1}) (\mathcal{G} \partial_{p_j} \mathcal{G}^{-1}) (\mathcal{G} \partial_{p_k} \mathcal{G}^{-1}) \right], \tag{B.1}$$

where the trace is performed over indices of the Dirac spinors and other internal degrees of freedom (if exist). Let $n$ be the dimension over which the trace is taken (i.e., $\mathcal{G}$ is a $n \times n$ matrix). In order for this integration to be well-defined, the integral domain should not contain on-shell pole nor branch cut, and hence we are led to take $p_0 = 0$. In this case, $\mathcal{G}$ is regular everywhere and an element of $GL(n, \mathbb{C})$ (general linear group of degree $n$ over complex numbers).

Let us next see why this quantity (B.1) plays a role of the topological invariant. When $\mathcal{G}$ does not commute nor anti-commute with $K$, this quantity does not have any topological meaning in general. On the other hand, when $\mathcal{G}$ satisfies

$$\mathcal{G} K = K \mathcal{G} \quad \text{or} \quad \mathcal{G} K = -K \mathcal{G}, \tag{B.2}$$

one can evaluate the integral (B.1) within each $K = \pm 1$ subspace,

$$\mathcal{N}_K|_{K=\pm 1} = \pm \frac{\epsilon_{ijk}}{24\pi^2} \underset{K=\pm 1}{\mathrm{Tr}} \left[ \int \mathrm{d}^3 p \; (\mathcal{G} \partial_{p_i} \mathcal{G}^{-1}) (\mathcal{G} \partial_{p_j} \mathcal{G}^{-1}) (\mathcal{G} \partial_{p_k} \mathcal{G}^{-1}) \right], \tag{B.3}$$

where $\underset{K=\pm 1}{\mathrm{Tr}}$ indicates the trace over the subspaces with $K = \pm 1$. Noting that $GL(n, \mathbb{C})$ homotopes to the unitary group $U(n)$, this integration resembles the Pontryagin index and

gives a topological invariant of a map from the three-dimensional torus $T^3$ into $U(n)$. Since this quantity is topologically quantized,[9] it is useful to characterize (symmetry-protected) topological phases.

### B.2 A simple example: Free Dirac fermion

As a toy model, let us consider a free Dirac fermion, whose Lagrangian is given by

$$\mathcal{L} = \bar{\psi} i\gamma^\mu \partial_\mu \psi + m_D \bar{\psi}\psi \,. \tag{B.4}$$

This Lagrangian is invariant under the $CT$ transformation (which is combination of the charge and time-reversal transformations):

$$\psi \xrightarrow{CT} -\gamma^1\gamma^3(-i)(\bar{\psi}\gamma^0\gamma^2)^T = -i\gamma^1\gamma^2\gamma^3\psi^* = i\gamma^5\gamma^0\psi^* \,. \tag{B.5}$$

Following Subsection B.1, let us evaluate the quantity (B.1) with $K = i\gamma^5\gamma^0$,

$$\mathcal{N}_{CT} \equiv \frac{\epsilon_{ijk}}{24\pi^2}\mathrm{Tr}\left[\int \mathrm{d}^3p \; i\gamma^5\gamma^0(\mathcal{G}\partial_{p_i}\mathcal{G}^{-1})(\mathcal{G}\partial_{p_j}\mathcal{G}^{-1})(\mathcal{G}\partial_{p_k}\mathcal{G}^{-1})\right]_{p_0=0} , \tag{B.6}$$

with $\mathcal{G}^{-1} \equiv \gamma^\mu p_\mu + m_D$. If the Lagrangian were not invariant under the $CT$ transformation, this quantity is not topological. Thus, it is meaningful only when the system respects the symmetry. One can explicitly check now that the quantity (B.6) is indeed a topological invariant as follows:

$$\mathcal{N}_{CT} = \frac{\epsilon_{ijk}}{24\pi^2}\mathrm{Tr}\left[\int \mathrm{d}^3p \; \frac{i\gamma^5\gamma^0(\not{p} - m_D)\gamma^i(\not{p} - m_D)\gamma^j(\not{p} - m_D)\gamma^k}{(p^2 - m_D^2)^3}\right]_{p_0=0} , \tag{B.7}$$

where the numerator is simplified to be $i\gamma^5\gamma^0\gamma^i\gamma^j\gamma^k(-m_D)(\boldsymbol{p}^2 + m_D^2)$ with $\boldsymbol{p}$ the three-dimensional momentum ($p^2|_{p_0=0} = -\boldsymbol{p}^2$). Using a formula $\mathrm{Tr}[i\gamma^5\gamma^0\gamma^i\gamma^j\gamma^k] = 4\epsilon^{ijk}$ and $\epsilon_{ijk}\epsilon^{ijk} = 6$, we have

$$\mathcal{N}_{CT} = \frac{1}{\pi^2}\int \mathrm{d}^3p \; \frac{m_D}{(\boldsymbol{p}^2 + m_D^2)^2} = \frac{m_D}{|m_D|} = \mathrm{sign}(m_D) \,. \tag{B.8}$$

Thus, this is a topological invariant, which only depends on the sign of the Dirac mass. Note that, in this free Dirac fermion system, the sign of the Dirac mass can be easily changed by redefinition of the field $\psi$, which means that $\mathcal{N}_{CT}$ itself cannot have any physical meaning.[10]

[9] Because the momentum space considered here is not three-sphere $S^3$ but $T^3$, this quantity is different from the Pontryagin index to measure $\pi_3(U(n)) = \mathbb{Z}$. As a result, it is not topological in a strict sense but can be continuously changed when one introduces interactions violating spatial translational symmetry [102, 103]. Nevertheless, we can practically regard it topological because the translational symmetry is usually respected in particle physics setups.

[10] Once higher order terms with respect to momenta are taken into account in the dispersion relation, the sign of the Dirac mass becomes physically meaningful in relation to the sign of the coefficient of such terms, and one may discuss whether the system is in a topologically (non)trivial phase, see e.g., Ref. [104]. In relativistic field theories in continuum space, such higher order terms may naturally come from regularizations [105].

### B.3 Type II seesaw model

We here give a similar quantity in the type-II seesaw model with the left-handed neutrino having a Majorana mass. Let $\nu_L$ be a two-component spinor field describing the left-handed neutrino in the SM. We consider free neutrinos and ignore all interactions to the SM gauge and Higgs fields. The non-interacting Lagrangian for the type II seesaw model is given by

$$\mathcal{L} = \nu_L^\dagger i\bar{\sigma}^\mu \partial_\mu \nu_L + \left( \frac{M_L}{2} \nu_L^\dagger \nu_L^c + (\text{h.c.}) \right) = \frac{1}{2} \overline{N} \left( i\gamma^\mu \partial_\mu + M_L \right) N \,, \tag{B.9}$$

where the four-component Nambu-Gorkov spinor (Majorana spinor) $N$ defined in Eq. (3.30) has been used. Since the system (B.9) is not invariant under the $CT$ (nor $P$) transformation, the topological invariant (B.6) in this model trivially vanishes. However, one can define a similar quantity associated with the $CP$ (or $T$) symmetry.

As discussed in Section 3.3, the Majorana mass $M_L$ originates from a Majorana-Yukawa interaction in which a $SU(2)$-triplet Higgs field $\Delta$ takes a vacuum expectation value (3.26), although we do not take into account its origin since we are interested in the broken phase and ignore all interactions. In this basis, the time-reversal symmetry acts on $N$ as

$$N = \begin{pmatrix} \nu_L \\ \nu_L^c \end{pmatrix} \xrightarrow{T} \begin{pmatrix} \sigma^1\sigma^3 \nu_L \\ \sigma^1\sigma^3 \nu_L^c \end{pmatrix} = \begin{pmatrix} -i\sigma^2 \nu_L \\ -i\sigma^2 i\sigma^2 (\nu_L^\dagger)^T \end{pmatrix} = \begin{pmatrix} -(\nu_L^{c\dagger})^T \\ (\nu_L^\dagger)^T \end{pmatrix} = \gamma^5\gamma^0 (N^\dagger)^T \,. \tag{B.10}$$

Therefore, we can define a topological invariant with $K = i\gamma^5\gamma^0$,

$$\mathcal{N}_T \equiv \frac{1}{2} \frac{\epsilon_{ijk}}{24\pi^2} \text{Tr} \left[ \int \text{d}^3p \; i\gamma^5\gamma^0 (\mathcal{G}\partial_{p_i}\mathcal{G}^{-1})(\mathcal{G}\partial_{p_j}\mathcal{G}^{-1})(\mathcal{G}\partial_{p_k}\mathcal{G}^{-1}) \right]_{p_0=0} , \tag{B.11}$$

with $\mathcal{G}^{-1} \equiv \gamma^\mu p_\mu + M_L$. This is the same form as $\mathcal{N}_{CT}$ defined by Eq. (B.6) except for the overall factor $1/2$, so that we can calculate this in the same way, which leads to

$$\mathcal{N}_T = \frac{1}{2} \frac{M_L}{|M_L|} = \frac{1}{2} \text{sign}(M_L) \,. \tag{B.12}$$

The topological invariant depends on the sign of the Majorana mass.

### B.4 Type I seesaw model

We here consider the type-I seesaw model, in which $\nu_L$ and $\nu_R$ are two-component spinors describing the left-handed and left-handed neutrinos, respectively. The non-interacting Lagrangian is given by

$$\mathcal{L} = \nu_L^\dagger i\bar{\sigma}^\mu \partial_\mu \nu_L + \nu_R^\dagger i\sigma^\mu \partial_\mu \nu_R + \left( m_D \nu_L^\dagger \nu_R - \frac{M_R}{2} \nu_R^\dagger \nu_R^c + (\text{h.c.}) \right) , \tag{B.13}$$

where

$$\nu_R^c \equiv -i\sigma^2 (\nu_R^\dagger)^T \,, \qquad \nu_L^c \equiv i\sigma^2 (\nu_L^\dagger)^T \,. \tag{B.14}$$

Here, we perform a field rotation with angle $\theta = m_D/M_R$ assuming $|m_D| \ll |M_R|$ such that the neutrino fields mix as

$$\nu_L^{c\prime} \simeq \nu_L^c + \theta \nu_R \,, \qquad \nu_R' \simeq \nu_R - \theta \nu_L^c \,, \tag{B.15}$$

for which the Lagrangian is rewritten as

$$\mathcal{L} = \nu_L'^{\dagger} i\bar{\sigma}^\mu \partial_\mu \nu_L' + \nu_R'^{\dagger} i\sigma^\mu \partial_\mu \nu_R' + \left( \frac{m_D^2}{2M_R} \nu_L'^{\dagger} \nu_L'^{c} - \frac{M_R}{2} (\nu_R'^{c})^\dagger \nu_R' + \text{h.c.} \right) . \tag{B.16}$$

Note that we have used $\nu_L'^T (\nu_R'^{\dagger})^T = -\nu_R'^{\dagger} \nu_L'$ since they are Grassmann variables.

Now, the signs of the mass terms for $\nu_L'$ and $\nu_R'$ are opposite, so that the total topological invariant vanishes:

$$\mathcal{N}_T = \frac{1}{2} \,\text{sign}\left( -\frac{m_D^2}{M_R} \right) + \frac{1}{2} \,\text{sign}(M_R) = 0 \,. \tag{B.17}$$

### B.5 Hybrid of type I and II models

It is straight forward to extend the above result into hybrid models of type I and II, in which the Lagrangian is given by

$$\mathcal{L} = \nu_L^\dagger i\bar{\sigma}^\mu \partial_\mu \nu_L + \nu_R^\dagger i\sigma^\mu \partial_\mu \nu_R + \left( \frac{M_L}{2} \nu_L^\dagger \nu_L^c + \frac{M_R}{2} \nu_R^\dagger \nu_R^c + m_D \nu_L^\dagger \nu_R + (\text{h.c.}) \right) . \tag{B.18}$$

By diagonalizing the mass matrix on the field basis $(\nu_L, \nu_R)$,

$$\boldsymbol{M}_\nu = \begin{pmatrix} M_L & m_D \\ m_D & M_R \end{pmatrix} , \tag{B.19}$$

one obtains mass eigenvalues,

$$m_1 = \frac{1}{2} \left[ M_L + M_R - \sqrt{(M_L - M_R)^2 + 4m_D^2} \right] , \tag{B.20}$$

$$m_2 = \frac{1}{2} \left[ M_L + M_R + \sqrt{(M_L - M_R)^2 + 4m_D^2} \right] . \tag{B.21}$$

Since the total topological invariant is again given as

$$\mathcal{N}_T = \frac{1}{2} \text{sign}(m_1) + \frac{1}{2} \text{sign}(m_2) \,, \tag{B.22}$$

one can easily see that there are two phases of $\mathcal{N}_T$:

$$\begin{cases} m_1, m_2 > 0 & \Leftrightarrow \mathcal{N}_T = 1 \,, \\ m_1, m_2 < 0 & \Leftrightarrow \mathcal{N}_T = -1 \,, \\ m_1 < 0 \wedge m_2 > 0 & \Leftrightarrow \mathcal{N}_T = 0 \,. \end{cases} \tag{B.23}$$

Note here that $m_1 < m_2$ always holds.

It follows that the transition of the sign of $\det \boldsymbol{M}_\nu = m_1 m_2$ implies the transition of the phases, which means that the phases cannot be continuously connected unless at least one mass eigenstate becomes massless (i.e., the mass gap is closed).

## References


[1] J. Bardeen, L. N. Cooper and J. R. Schrieffer, *Microscopic theory of superconductivity*, *Phys. Rev.* **106** (1957) 162.

[2] Y. Nambu and G. Jona-Lasinio, *Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. 1.*, *Phys. Rev.* **122** (1961) 345.

[3] Y. Nambu and G. Jona-Lasinio, *Dynamical model of elementary particles based on an analogy with superconductivity. II.*, *Phys. Rev.* **124** (1961) 246.

[4] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.

[5] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.

[6] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.

[7] V. L. Ginzburg and L. D. Landau, *On the Theory of superconductivity*, *Zh. Eksp. Teor. Fiz.* **20** (1950) 1064.

[8] A. A. Abrikosov, *On the Magnetic properties of superconductors of the second group*, *Sov. Phys. JETP* **5** (1957) 1174.

[9] H. B. Nielsen and P. Olesen, *Vortex Line Models for Dual Strings*, *Nucl. Phys.* **B61** (1973) 45.

[10] M. Eto, Y. Hamada, R. Jinno, M. Nitta and M. Yamada, *Abrikosov-Nielsen-Olesen strings from the Coleman-Weinberg potential*, *Phys. Rev. D* **106** (2022) 116002 [2205.04394].

[11] T. W. B. Kibble, *Topology of Cosmic Domains and Strings*, *J. Phys.* **A9** (1976) 1387.

[12] W. H. Zurek, *Cosmological Experiments in Superfluid Helium?*, *Nature* **317** (1985) 505.

[13] N. B. Kopnin and M. M. Salomaa, *Mutual friction in superfluid* $^3$He*: Effects of bound states in the vortex core*, *Phys. Rev. B* **44** (1991) 9667.

[14] G. E. Volovik, *Fermion zero modes on vortices in chiral superconductors*, *JETP Lett.* **70** (1999) 609 [cond-mat/9909426].

[15] N. Read and D. Green, *Paired states of fermions in two-dimensions with breaking of parity and time reversal symmetries, and the fractional quantum Hall effect*, *Phys. Rev. B* **61** (2000) 10267 [cond-mat/9906453].

[16] D. A. Ivanov, *Non-abelian statistics of half-quantum vortices in p-wave superconductors*, *Phys. Rev. Lett.* **86** (2001) 268 [cond-mat/0005069].

[17] T. Fukui and T. Fujiwara, *Topological stability of Majorana zero-modes in superconductor-topological insulator systems*, *J. Phys. Soc. Jap.* **79** (2010) 033701 [0911.2558].

[18] G. E. Volovik, *The Universe in a Helium Droplet*. Oxford University Press, 02, 2009, 10.1093/acprof:oso/9780199564842.001.0001.

[19] A. Sedrakian and J. W. Clark, *Superfluidity in nuclear systems and neutron stars*, *Eur. Phys. J.* **A55** (2019) 167 [1802.00017].

[20] Y. Masaki, T. Mizushima and M. Nitta, *Microscopic description of axisymmetric vortices in* $^3P_2$ *superfluids*, *Phys. Rev. Res.* **2** (2020) 013193 [1908.06215].

[21] Y. Masaki, T. Mizushima and M. Nitta, *Non-Abelian half-quantum vortices in 3P2 topological superfluids*, *Phys. Rev. B* **105** (2022) L220503 [2107.02448].

[22] Y. Masaki, T. Mizushima and M. Nitta, *Non-Abelian Anyons and Non-Abelian Vortices in Topological Superconductors*, in *Encyclopedia of Condensed Matter Physics (Second Edition)*, pp. 755–794, 2024, DOI [2301.11614].

[23] Y. Nishida, *Is a color superconductor topological?*, *Phys. Rev. D* **81** (2010) 074004 [1001.2555].

[24] S. Yasui, K. Itakura and M. Nitta, *Fermion structure of non-Abelian vortices in high density QCD*, *Phys. Rev. D* **81** (2010) 105003 [1001.3730].

[25] T. Fujiwara, T. Fukui, M. Nitta and S. Yasui, *Index theorem and Majorana zero modes along a non-Abelian vortex in a color superconductor*, *Phys. Rev. D* **84** (2011) 076002 [1105.2115].

[26] M. Eto, Y. Hirono, M. Nitta and S. Yasui, *Vortices and Other Topological Solitons in Dense Quark Matter*, *PTEP* **2014** (2014) 012D01 [1308.1535].

[27] C. Chatterjee and M. Nitta, *Low-energy Effective Worldsheet theory of a non-Abelian vortex in high-density QCD Revisited: A regular gauge construction*, *Phys. Rev. D* **95** (2017) 085013 [1612.09419].

[28] M. A. Zubkov, *Momentum space topology of QCD*, *Annals Phys.* **393** (2018) 264 [1610.08041].

[29] A. P. Balachandran, S. Digal and T. Matsuura, *Semi-superfluid strings in high density QCD*, *Phys. Rev.* **D73** (2006) 074009 [hep-ph/0509276].

[30] E. Nakano, M. Nitta and T. Matsuura, *Non-Abelian strings in high density QCD: Zero modes and interactions*, *Phys. Rev.* **D78** (2008) 045002 [0708.4096].

[31] M. Eto, E. Nakano and M. Nitta, *Effective world-sheet theory of color magnetic flux tubes in dense QCD*, *Phys. Rev.* **D80** (2009) 125011 [0908.4470].

[32] M. Eto, M. Nitta and N. Yamamoto, *Instabilities of Non-Abelian Vortices in Dense QCD*, *Phys. Rev. Lett.* **104** (2010) 161601 [0912.1352].

[33] V. Gurarie, *Single-particle green's functions and interacting topological insulators*, *Physical Review B* **83** (2011) 085426.

[34] G. E. Volovik and M. A. Zubkov, *Standard Model as the topological material*, *New J. Phys.* **19** (2017) 015009 [1608.07777].

[35] A. M. Essin and V. Gurarie, *Bulk-boundary correspondence of topological insulators from their respective green's functions*, *Phys. Rev. B* **84** (2011) 125132.

[36] A. Schnyder, S. Ryu, A. Furusaki and A. Ludwig, *Classification of topological insulators and superconductors in three spatial dimensions*, *Phys. Rev. B* **78** (2008) 195125 [0803.2786].

[37] A. Kitaev, *Periodic table for topological insulators and superconductors*, *AIP Conf. Proc.* **1134** (2009) 22 [0901.2686].

[38] A. P. Schnyder, S. Ryu, A. Furusaki and A. W. Ludwig, *Classification of topological insulators and superconductors*, in *AIP Conference Proceedings*, vol. 1134, pp. 10–21, American Institute of Physics, 2009.

[39] J. C. Y. Teo and C. L. Kane, *Topological Defects and Gapless Modes in Insulators and Superconductors*, *Phys. Rev. B* **82** (2010) 115120 [1006.0690].

[40] S. Ryu, A. P. Schnyder, A. Furusaki and A. W. W. Ludwig, *Topological insulators and superconductors: Tenfold way and dimensional hierarchy*, *New J. Phys.* **12** (2010) 065010 [0912.2157].

[41] C. R. Nohl, *Bound State Solutions of the Dirac Equation in Extended Hadron Models*, *Phys. Rev. D* **12** (1975) 1840.

[42] H. J. de Vega, *Fermions and Vortex Solutions in Abelian and Nonabelian Gauge Theories*, *Phys. Rev. D* **18** (1978) 2932.

[43] R. Jackiw and P. Rossi, *Zero Modes of the Vortex - Fermion System*, *Nucl. Phys. B* **190** (1981) 681.

[44] E. J. Weinberg, *Index Calculations for the Fermion-Vortex System*, *Phys. Rev. D* **24** (1981) 2669.

[45] C. G. Callan, Jr. and J. A. Harvey, *Anomalies and Fermion Zero Modes on Strings and Domain Walls*, *Nucl. Phys. B* **250** (1985) 427.

[46] G. W. Semenoff, *Index Theorems and Superconducting Cosmic Strings*, *Phys. Rev. D* **37** (1988) 2838.

[47] N. Ganoulis and G. Lazarides, *A Generalized Index Theorem for String Superconductivity in Realistic Models*, *Phys. Rev. D* **38** (1988) 547.

[48] H. Bagherian, K. Fraser, S. Homiller and J. Stout, *Zero Modes of Massive Fermions Delocalize from Axion Strings*, 2310.01476.

[49] M. Eto, Y. Isozumi, M. Nitta, K. Ohashi and N. Sakai, *Solitons in the Higgs phase: The Moduli matrix approach*, *J. Phys. A* **39** (2006) R315 [hep-th/0602170].

[50] M. Shifman and A. Yung, *Supersymmetric Solitons and How They Help Us Understand Non-Abelian Gauge Theories*, *Rev. Mod. Phys.* **79** (2007) 1139 [hep-th/0703267].

[51] M. Shifman and A. Yung, *Supersymmetric solitons*, Cambridge Monographs on Mathematical Physics. Cambridge University Press, 5, 2009, 10.1017/CBO9780511575693.

[52] M. Hindmarsh, *Massless modes on cosmic strings*, *Physica B* **178** (1992) 47.

[53] E. Witten, *Superconducting Strings*, *Nucl. Phys. B* **249** (1985) 557.

[54] Y. Nambu, *String-Like Configurations in the Weinberg-Salam Theory*, *Nucl. Phys.* **B130** (1977) 505.

[55] T. Vachaspati, *Vortex solutions in the Weinberg-Salam model*, *Phys. Rev. Lett.* **68** (1992) 1977.

[56] A. Achucarro and T. Vachaspati, *Semilocal and electroweak strings*, *Phys. Rept.* **327** (2000) 347 [hep-ph/9904229].

[57] M. James, L. Perivolaropoulos and T. Vachaspati, *Stability of electroweak strings*, *Phys. Rev.* **D46** (1992) R5232.

[58] M. James, L. Perivolaropoulos and T. Vachaspati, *Detailed stability analysis of electroweak strings*, *Nucl. Phys.* **B395** (1993) 534 [hep-ph/9212301].

[59] M. Goodband and M. Hindmarsh, *Instabilities of electroweak strings*, *Phys. Lett. B* **363** (1995) 58 [hep-ph/9505357].

[60] R. Holman, S. Hsu, T. Vachaspati and R. Watkins, *Metastable cosmic strings in realistic models*, *Phys. Rev. D* **46** (1992) 5352 [hep-ph/9208245].

[61] J. Garriga and X. Montes, *Stability of Z strings in strong magnetic fields*, *Phys. Rev. Lett.* **75** (1995) 2268 [hep-ph/9505424].

[62] M. Nagasawa and R. Brandenberger, *Stabilization of the electroweak Z string in the early universe*, *Phys. Rev. D* **67** (2003) 043504 [hep-ph/0207246].

[63] L. Masperi, A. Megevand and S. Savaglio, *Stability of modified electroweak strings*, *Z. Phys. C* **72** (1996) 171 [hep-ph/9410211].

[64] P. Forgács and A. Lukács, *Electroweak strings with dark scalar condensates and their stability*, *Phys. Rev. D* **102** (2020) 023009 [1909.07447].

[65] M. Eto, Y. Hamada and M. Nitta, *Stable Z-strings with topological polarization in two Higgs doublet model*, *JHEP* **02** (2022) 099 [2111.13345].

[66] R. H. Brandenberger and A.-C. Davis, *Electroweak baryogenesis with electroweak strings*, *Phys. Lett.* **B308** (1993) 79 [astro-ph/9206001].

[67] M. Barriola, *Electroweak strings that produce baryons*, *Phys. Rev.* **D51** (1995) 300 [hep-ph/9403323].

[68] T. Vachaspati and G. B. Field, *Electroweak string configurations with baryon number*, *Phys. Rev. Lett.* **73** (1994) 373 [hep-ph/9401220].

[69] T. Vachaspati, *Estimate of the primordial magnetic field helicity*, *Phys. Rev. Lett.* **87** (2001) 251302 [astro-ph/0101261].

[70] R. Poltis and D. Stojkovic, *Can primordial magnetic fields seeded by electroweak strings cause an alignment of quasar axes on cosmological scales?*, *Phys. Rev. Lett.* **105** (2010) 161301 [1004.2704].

[71] T. Vachaspati and R. Watkins, *Bound states can stabilize electroweak strings*, *Phys. Lett. B* **318** (1993) 163 [hep-ph/9211284].

[72] M. A. Earnshaw and W. B. Perkins, *Stability of an electroweak string with a fermion condensate*, *Phys. Lett. B* **328** (1994) 337 [hep-ph/9402218].

[73] J. M. Moreno, D. H. Oaknin and M. Quiros, *Fermions on the electroweak string*, *Phys. Lett. B* **347** (1995) 332 [hep-ph/9411411].

[74] J. Garriga and T. Vachaspati, *Zero modes on linked strings*, *Nucl. Phys. B* **438** (1995) 161 [hep-ph/9411375].

[75] S. G. Naculich, *Fermions destabilize electroweak strings*, *Phys. Rev. Lett.* **75** (1995) 998 [hep-ph/9501388].

[76] S. Kono and S. G. Naculich, *Electroweak strings and fermions*, in *PASCOS / HOPKINS 1995 (Joint Meeting of the International Symposium on Particles, Strings and Cosmology and the 19th Johns Hopkins Workshop on Current Problems in Particle Theory)*, pp. 297–306, 7, 1995, hep-ph/9507350.

[77] M. Groves and W. B. Perkins, *The Dirac sea contribution to the energy of an electroweak string*, *Nucl. Phys. B* **573** (2000) 449 [hep-ph/9908416].

[78] G. E. Volovik, *Condensation of fermion zero modes in the vortex*, *JETP Lett.* **104** (2016) 201 [1509.05006].

[79] H. Liu and T. Vachaspati, *Perturbed electroweak strings and fermion zero modes*, *Nucl. Phys. B* **470** (1996) 176 [hep-ph/9511216].

[80] G. D. Starkman, D. Stojkovic and T. Vachaspati, *Neutrino zero modes on electroweak strings*, *Phys. Rev. D* **63** (2001) 085011 [hep-ph/0007071].

[81] G. Starkman, D. Stojkovic and T. Vachaspati, *Zero modes of fermions with a general mass matrix*, *Phys. Rev. D* **65** (2002) 065003 [hep-th/0103039].

[82] T. Vachaspati and A. Achucarro, *Semilocal cosmic strings*, *Phys. Rev. D* **44** (1991) 3067.

[83] M. Hindmarsh, *Existence and stability of semilocal strings*, *Phys. Rev. Lett.* **68** (1992) 1263.

[84] A. Achucarro, K. Kuijken, L. Perivolaropoulos and T. Vachaspati, *Dynamical simulations of semilocal strings*, *Nucl. Phys. B* **388** (1992) 435.

[85] J. Ambjorn and P. Olesen, *Electroweak Magnetism: Theory and Application*, *Int. J. Mod. Phys. A* **5** (1990) 4525.

[86] T. Vachaspati, *Electroweak strings*, *Nucl. Phys.* **B397** (1993) 648.

[87] A. Achucarro, R. Gregory, J. A. Harvey and K. Kuijken, *Cinderella strings*, *Phys. Rev. Lett.* **72** (1994) 3646 [hep-th/9312034].

[88] W. B. Perkins, *W condensation in electroweak strings*, *Phys. Rev. D* **47** (1993) R5224.

[89] M. A. Silaev and G. E. Volovik, *Topological superfluid $^3$He-B: Fermion zero modes on interfaces and in the vortex core*, *J. Low Temp. Phys.* **161** (2010) 460 [1005.4672].

[90] H. La, *Vortex solutions in two Higgs systems and tan Beta*, hep-ph/9302220.

[91] M. A. Earnshaw and M. James, *Stability of two doublet electroweak strings*, *Phys. Rev. D* **48** (1993) 5818 [hep-ph/9308223].

[92] G. Bimonte and G. Lozano, *Vortex solutions in two Higgs doublet systems*, *Phys. Lett. B* **326** (1994) 270 [hep-ph/9401313].

[93] G. R. Dvali and G. Senjanovic, *Topologically stable electroweak flux tubes*, *Phys. Rev. Lett.* **71** (1993) 2376 [hep-ph/9305278].

[94] M. Eto, M. Kurachi and M. Nitta, *Constraints on two Higgs doublet models from domain walls*, *Phys. Lett. B* **785** (2018) 447 [1803.04662].

[95] M. Eto, M. Kurachi and M. Nitta, *Non-Abelian strings and domain walls in two Higgs doublet models*, *JHEP* **08** (2018) 195 [1805.07015].

[96] M. Eto, Y. Hamada, M. Kurachi and M. Nitta, *Topological Nambu monopole in two Higgs doublet models*, *Phys. Lett. B* **802** (2020) 135220 [1904.09269].

[97] M. Eto, Y. Hamada, M. Kurachi and M. Nitta, *Dynamics of Nambu monopole in two Higgs doublet models. Cosmological Monopole Collider*, *JHEP* **07** (2020) 004 [2003.08772].

[98] M. Eto, Y. Hamada and M. Nitta, *Topological structure of a Nambu monopole in two-Higgs-doublet models: Fiber bundle, Dirac's quantization, and a dyon*, *Phys. Rev. D* **102** (2020) 105018 [2007.15587].

[99] A. Altland and M. R. Zirnbauer, *Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures*, *Physical Review B* **55** (1997) 1142.

[100] G. E. Volovik, *Topological invariant for superfluid He-3 - B and quantum phase transitions*, *JETP Lett.* **90** (2009) 587 [0909.3084].

[101] G. E. Volovik, *Topological superfluid 3He-B in magnetic field and Ising variable*, *Pisma Zh. Eksp. Teor. Fiz.* **91** (2010) 61 [1001.1514].

[102] R. Kennedy and C. Guggenheim, *Homotopy theory of strong and weak topological insulators*, *Phys. Rev. B* **91** (2015) 245148.

[103] S. Ryu, A. P. Schnyder, A. Furusaki and A. W. W. Ludwig, *Topological insulators and superconductors: tenfold way and dimensional hierarchy*, *New Journal of Physics* **12** (2010) 065010.

[104] S.-Q. SHEN, W.-Y. SHAN and H.-Z. LU, *Topological insulator and the dirac equation*, *SPIN* **01** (2011) 33 [1009.5502].

[105] S. Aoki, H. Fukaya, N. Kan, M. Koshino and Y. Matsuki, *Magnetic monopole becomes dyon in topological insulators*, *Phys. Rev. B* **108** (2023) 155104 [2304.13954].

JLAB-THY-24-3998

# Hard-scattering approach to strongly hindered electric dipole transitions between heavy quarkonia

Cai-Ping Jia*,[1,2] Yu Jia†,[2,3] Junliang Lu‡,[4,2] Zhewen Mo§,[5,2] and Jia-Yue Zhang¶[6,2]

[1]*School of Nuclear Science and Technology, Lanzhou University, Lanzhou 730000, China*

[2]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*

[3]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

[4]*Institute of Particle and Nuclear Physics, Henan Normal University, Xinxiang 453007, China*

[5]*CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*

[6]*Theory Center, Jefferson Lab, Newport News, Virginia 23606, USA*

(Dated: March 1, 2024)

arXiv:2402.19413v1 [hep-ph] 29 Feb 2024

## Abstract

The conventional wisdom in dealing with electromagnetic transition between heavy quarkonia is the multipole expansion, when the emitted photon has a typical energy of order quarkonium binding energy. Nevertheless, in the case when the energy carried by the photon is of order typical heavy quark momentum, the multipole expansion doctrine is expected to break down. In this work, we apply the "hard-scattering" approach originally developed to tackle the strongly hindered magnetic dipole ($M1$) transition [Y. Jia *et al.*, Phys. Rev. D. 82, 014008 (2010)] to the strongly hindered electric dipole ($E1$) transition between heavy quarkonia. We derive the factorization formula for the strongly hindered $E1$ transition rates at the lowest order in velocity and $\alpha_s$ in the context of the non-relativistic QCD (NRQCD), and conduct a detailed numerical comparison with the standard predictions for various bottomonia and charmonia $E1$ transition processes.

* jiacp20@lzu.edu.cn
† jiay@ihep.ac.cn
‡ lujl@ihep.ac.cn
§ mozw@itp.ac.cn
¶ jzhang@jlab.org

## I. INTRODUCTION

Electromagnetic (EM) transitions provide important means to unravel the internal structure of quarkonium and search for the new quarkonium states [1, 2]. From the experimental perspective, many EM transition channels of charmonia and bottomonia have been precisely measured in BES, $B$ factories, and CLEO experiments [3]. From the theoretical standpoint, the standard approach to tackle the EM transition between quarkonium is based on the multipole expansion, essentially identical to the well-known methodology in dealing with the atomic EM transition. The dominant transition mode is the electric dipole ($E1$) transition, where the orbital angular momentum of heavy quarks is changed by one unit while the spin is conserved. The subleading transition mode is the magnetic dipole ($M1$) transition, where the orbital angular momentum does not change but the spin is flipped.

In the standard picture of multipole expansion, the radiated photon is assumed to carry a very small (*ultrasoft*) momentum, $k \sim mv^2$, where $m$ denotes the heavy quark mass, and $v$ signifies the typical velocity of the quark within the quarkonium. Consequently, the wavelength of the emitted photon is much longer than the size of the quarkonium, so that the transition amplitude can be expressed as an overlap integral between the wave functions of the initial-state and final-state quarkonia. The leading $E1$ transition formula has long been known in the spontaneous radiation of atoms about one century ago, as one of the triumphs of the quantum theory of electromagnetic fields [4]. The corresponding $E1$ transition formula for quarkonium is essentially the same. The modern treatment of EM transitions of quarkonia is expedited by invoking the machinery of nonrelativistic effective field theory (EFT) approach, the so-called potential NRQCD (pNRQCD), where the relativistic corrections can be systematically accounted [5, 6].

Since the binding force for an atom in QED is Coulombic, it is appropriate to assume the energy difference between any two atomic levels is of order $mv^2$. However, the story is quite different for heavy quarkonium. The interquark potential is Coulombic only at short distance and becomes linearly rising confinement potential at long distance. As a consequence, the mass difference between the highly-excited quarkonium and the lowest-lying state could be rather large, actually could be arbitrarily large in the $N_c \to \infty$ limit or if neglecting the string breaking effect. For example, the mass difference between $\Upsilon(3S)$ and $\eta_b$ is as large as 1 GeV. In such a case, it is obviously inappropriate to regard the photon as ultrasoft, and the standard multipole expansion doctrine simply breaks down.

In fact, for the $M1$ transition process $\Upsilon(3S) \to \eta_b + \gamma$, it is more appropriate to regard the photon to be *semi-hard* ($k \sim mv$) rather than ultrasoft. Assuming that perturbative QCD is still applicable at the scale $mv \sim 1$ GeV, the "hard-scattering" approach has been developed to tackle a class of strongly hindered $M1$ transitions exemplified by $\Upsilon(3S) \to \eta_b + \gamma$ in 2009 [7]. The key idea is that a semi-hard gluon has to be exchanged between heavy quark and antiquark in such a process, so that the transition amplitude can be expressed as the convolution between the radial wave functions of the parent and daughter quarkonia and the perturbatively calculable hard-scattering kernel. Without adjustable parameters, the predictions made in [7] are found in decent agreement with the measurements.

The goal of this paper is to extend the theoretical framework in [7] to the case of the strongly hindered $E1$ transitions between heavy quarkonia. For concreteness, we concentrate on the $E1$ transitions between $P$-wave and $S$-wave quarkonium states. We derive the corresponding factorization formula in the context of nonrelativistic QCD (NRQCD), and infer the "hard-scattering" kernel at the leading order in $v$ and $\alpha_s$. We also numerically compare

the predictions from this new formalism with those from the conventional multipole formula.

The rest of the paper is organized as follows. In Section II, we introduce our theoretical framework to deal with $E1$ transition, the nonrelativistic EFT with $SU(3)_c \times U(1)_{\rm em}$ gauge group. In Section III, we recap the standard $E1$ transition formula from the conventional multipole expansion context. Section IV is the main body of the paper. Starting from the NREFT, we apply the hard-scattering mechanism to derive the formula for strongly hindered $E1$ transition, taking the $P$- to $S$-wave quarkonium transition as an explicit example. In Section V, we present the numerical predictions for various $E1$ transition processes in bottomonium and charmonium family, comparing the predictions from the new mechanism with the conventional one. We finally summarize in Section VI. In Appendix we illustrate that identical formula for strongly hindered $E1$ transition can also be derived in the underlying relativistic field theory.

## II. THEORETICAL SETUP: NONRELATIVISTIC EFFECTIVE FIELD THEORY

A heavy quarkonium is a nonrelativistic bound state that contains towers of distinct scales: the heavy quark mass $m$, its typical momentum $mv$, and its typical kinetic energy $mv^2$, obeying the hierarchy $m \gg mv \gg mv^2$. Needlessly to say, there also exists the characteristic QCD scale $\Lambda_{\rm QCD}$. From phenomenological potential models, one empirically accepts that $v^2 \approx 0.3$ for charomonium, and $v^2 \approx 0.1$ for bottomonium. The momentum scale $mv \sim 0.9$ GeV for charmonium and $mv \sim 1.2$ GeV for bottomonium. The typical binding energy $mv^2 \sim \Lambda_{\rm QCD} \approx 0.5$ GeV for both charmonium and bottomonia.

The field-theoretical tool to tackle heavy quarkonium is an effective field theory dubbed nonrelativistic QCD (NRQCD) [8], which describes slowly-moving heavy quark and heavy antiquark to interact with the soft gluons. The predictions for a physical observable can be organized in velocity expansion. In order to describe the electromagnetic transition between heavy quarkonium, it is convenient to enlarge NRQCD by including the low-energy photon as the explicit degree of freedom. For this purpose, we promote the NRQCD to a nonrelativistic EFT (NREFT) with gauge group $SU(3)_c \times U(1)_{\rm em}$. Since the $E1$ transition does not flip the heavy quark/antiquark spin, suffices it to retain only the lowest-order NREFT Lagrangian:

$$\mathcal{L}_{\rm NREFT} = \psi^\dagger \left( iD_0 + \frac{\boldsymbol{D}^2}{2m} \right) \psi + \chi^\dagger \left( iD_0 + \frac{\boldsymbol{D}^2}{2m} \right) \chi + \mathcal{L}_{\rm light} + \cdots , \tag{1}$$

where $\psi$ ($\chi^\dagger$) denotes the Pauli spinor fields that annihilate a heavy quark (antiquark), $m$ signifies the heavy quark mass, $\mathcal{L}_{\rm light}+$ refers to the standard QCD Lagrangian for gluons and light quarks. The covariant derivatives $iD_0 = i\partial_0 - gt^a A_0^a - ee_q A_0^{\rm em}$, $i\boldsymbol{D} = i\nabla + gT^a \boldsymbol{A}^a + ee_Q \boldsymbol{A}^{\rm em}$, with $t^a$ ($a = 1, \cdots, 8$) denoting the $SU(3)_c$ generators in fundamental representation, and $ee_Q$ denoting the electric charge of the heavy quark. For simplicity, we have neglected all the higher-order terms in heavy quark velocity expansion. Note the NREFT at lowest order in $v$ emerges a rigorous heavy quark spin symmetry, which is manifest in the fact that (1) does not explicitly entail Pauli matrices.

One obtains the NRQCD Lagrangian by integrating out the relativistic quantum fluctuation (greater or equal to $m$) from QCD. As elucidated by Beneke and Smirnov [9], three distinct dynamical low-energy modes are present in NRQCD, with the dispersive scaling

behaviors specified by [1]

$$\text{semi-hard mode}: \quad k^0 \sim |\boldsymbol{k}| \sim mv; \tag{2a}$$

$$\text{potential mode}: \quad k^0 \sim mv^2, \qquad |\boldsymbol{k}| \sim mv; \tag{2b}$$

$$\text{ultrasoft mode}: \quad k^0 \sim |\boldsymbol{k}| \sim mv^2. \tag{2c}$$

## III. REVIEW OF CONVENTIONAL $E1$ TRANSITION FORMULA

It is the $ee_Q \boldsymbol{A}^{\rm em} \cdot \boldsymbol{\nabla}$ term in the NREFT Lagrangian (1) that mediates the $E1$ transition. In the standard textbook treatment, the emitted photon is assumed to be ultrasoft ($|\boldsymbol{k}| \sim mv^2$), whose wavelength is much longer than the typical quarkonium size ($\sim 1/mv$). Thus it is legitimate to carry out the multipole expansion and the leading EM transition is dominated by the $E1$ transition. In such case, the semi-hard quanta are irrelevant and can be integrated out, so that one descends from NRQCD to potential NRQCD (pNRQCD), which turns out to be the appropriate EFT to describe the electromagnetic transition between quarkonia.

The spin-averaged quarkonium $E1$ transition rate in the traditional multipole expansion framework reads [1]

$$\Gamma(^{2S+1}L_J \to {}^{2S'+1}L'_{J'} + \gamma) = \frac{4\alpha e_Q^2}{3} (2J'+1) S_{if}^E k^3 |\varepsilon_{if}|^2, \tag{3}$$

where a quarkonium state is characterized by the principle quantum number $n$, the total angular momentum $J$, the orbital angular momentum $L$ and total spin $S$. The statistical factor $S_{if}^E = S_{fi}^E$ is

$$S_{if}^E = \max(L, L') \begin{Bmatrix} J & 1 & J' \\ L' & S & L \end{Bmatrix}^2, \tag{4}$$

where the last term in the brackets denotes the 6-$j$ symbol. The QCD dynamics is encapsulated in the overlap integral $\varepsilon_{if}$, which at the lowest order in $v$ is given by

$$\varepsilon_{if} = \int_0^\infty dr\, r^3\, R_{nL}(r) R_{n'L'}(r), \tag{5}$$

where $R_{nL}(r)$ denotes the radial wave function with principle quantum number $n$ and orbital angular momentum $L$. For the $\mathcal{O}(v^2)$ corrections to the quarkonium $E1$ transition processes in the framework of pNRQCD, we refer the interested readers to [6].

## IV. STRONGLY HINDERED $E1$ TRANSITION BETWEEN $P$-WAVE AND $S$-WAVE QUARKONIA

In contrast to the atomic systems, the mass difference between highly excited quarkonium and the lowest-lying quarkonium state could be as large as 1 GeV (imagine the mass difference between $\chi_{bJ}(3P)$ and $\Upsilon(1S)$). In such a situation, it is much more appropriate to regard the emitted photon as *semi-hard* rather than ultrasoft, and one definitely ought to

[1] Note that the semi-hard mode is often dubbed the soft mode in the literature.

give up the traditional multipole expansion approach. Upon emitting a semi-hard photon, an almost on-shell (potential) heavy quark will inevitably become semi-hard, which bears a virtuality of order-$(mv)^2$ and can only survive in a relatively short time ( $1/mv$) according to the uncertainty principle. In order to bring this virtual heavy quark back to almost on-shell (potential), one semi-hard gluon must be exchanged between it and the spectator heavy anti-quark. This is the essential physical consideration that leads to the "hard-scattering" mechanism [7]. This new mechanism has been first applied to predict the strongly hindered $M1$ transition process $\Upsilon(3S) \to \eta_b\gamma$, and found notable phenomenological success. It is the major goal of this work to apply this "hard-scattering" mechanism to tackle the strongly hindered $E1$ transition between heavy quarkonia.

### A. Spin-conserving quark amplitude from "hard-scattering" mechanism

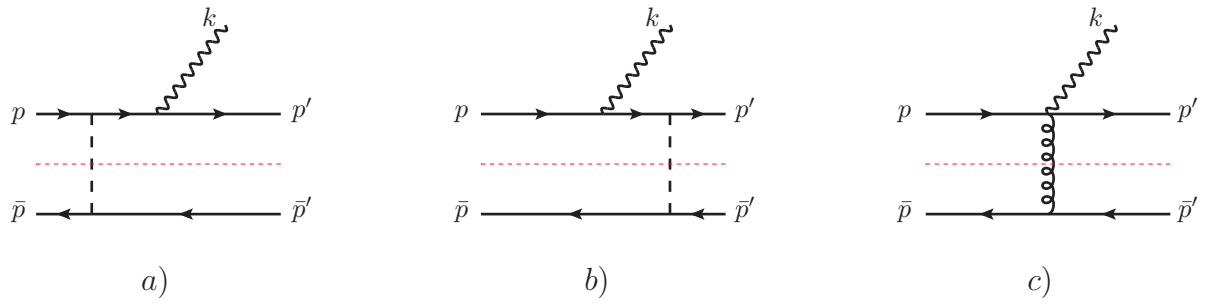



FIG. 1: Three lowest-order NREFT diagrams that contribute to the strongly hindered $E1$ transition between heavy quarkonia arising from the hard-scattering mechanism. The vertical dashed and curly lines signify the temporal and spatial gluon propagators. The horizontal thin dashed cut emphasizes that the upper half of the diagrams are reminiscent of the Compton scattering process, while the lower half of the diagrams are simply a quark-gluon vertex. For simplicity, we have omitted three mirror diagrams where the photon is emitted from the heavy anti-quark line.

Let us first consider the on-shell amplitude for $Q(p,s_1)\overline{Q}(\bar{p},s_2) \to Q(p',s_1')\overline{Q}(\bar{p}',s_2')+\gamma(k)$, where all the incoming and outgoing quarks and antiquarks move nonrelativistically. For simplicity, we choose to work in the rest frame of the parent quarkonium, and assume the photon moves along the $\hat{\mathbf{z}}$ axis. $s_1$, $s_2$ refer to the magnetic quantum number of the incoming $Q$ and $\overline{Q}$, and $s_1'$, $s_2'$ refer to the magnetic quantum number of the outgoing $Q$ and $\overline{Q}$. For our interest, we are only concerned in the case where the final-state $Q\overline{Q}$ pair inherits the same total spin from the initial-state $Q\overline{Q}$ pair. Denoting the total and relative momenta of the incoming (outgoing) heavy quark-antiquark pair by $P$ and $q$ ($P'$ and $q'$), respectively, we then have

$$p = \frac{P}{2} + q, \qquad \bar{p} = \frac{P}{2} - q; \tag{6a}$$

$$p' = \frac{P'}{2} + q', \qquad \bar{p}' = \frac{P'}{2} - q'. \tag{6b}$$

The on-shell condition demands that $P \cdot q = 0$ and $P' \cdot q' = 0$. For the incoming quark-antiquark pair, we have $P^\mu = (M_i, \mathbf{0})$ and $q^\mu = (0, \mathbf{q})$, with $M_i \approx 2m$ signifying the mass

of the parent quarkonium. It is natural to assume $\mathbf{q} \sim \mathcal{O}(mv)$. We will regard the emitted photon as semi-hard, *i.e.*, $\mathbf{k} \sim \mathcal{O}(mv)$. Since the daughter quarkonium recoils against the emitted photon, the daughter quarkonium acquires a nonzero yet small spatial momentum $\mathbf{P}' = -\mathbf{k} \sim \mathcal{O}(mv)$. As a consequence of $P' \cdot q' = 0$, the relative momentum for the outgoing quark-antiquark pair has a typical spatial component $\mathbf{q}' \sim \mathcal{O}(mv)$, yet acquires a strongly suppressed temporal component $q'^0 \sim \mathcal{O}(mv^2)$.

The punch line of the "hard-scattering" mechanism is that a semihard gluon must be exchanged between the heavy quark and heavy antiquark. Half of the lowest-order Feynman diagrams have been depicted in Fig. 1. The Feynman rules can be straightforwardly read off from the NREFT Lagrangian (1). Although it is a popular practice to use Coulomb gauge in computing the NRQCD diagrams, it turns out that Feynman gauge is a more convenient choice to handle the process entailing a semi-hard gluon exchange. The corresponding NREFT amplitudes affiliated with three diagrams in Fig. 1 are

$$i\mathcal{A}_a = \xi^{s_1'\dagger}\left(iee_Q\frac{2\mathbf{p}'\cdot\boldsymbol{\varepsilon}_\gamma^*}{2m}\right)\frac{i}{l_{Q1}^0-\frac{\mathbf{l}_{Q1}^2}{2m}}(ig_st^a)\xi^{s_1}\left(\frac{-i}{l_g^2+i\epsilon}\right)\eta^{s_2^\dagger}(ig_st^a)\eta^{s_2'}, \tag{7a}$$

$$i\mathcal{A}_b = \xi^{s_1'\dagger}(ig_st^a)\frac{i}{l_{Q2}^0-\frac{\mathbf{l}_{Q2}^2}{2m}}\left(iee_Q\frac{2\mathbf{p}\cdot\boldsymbol{\varepsilon}_\gamma^*}{2m}\right)\xi^{s_1}\left(\frac{-i}{l_g^2+i\epsilon}\right)\eta^{s_2^\dagger}(ig_st^a)\eta^{s_2'}, \tag{7b}$$

$$i\mathcal{A}_c = \xi^{s_1'\dagger}\left(\frac{ig_see_Q}{2m}t^a\right)\xi^{s_1}\left(\frac{i}{l_g^2+i\epsilon}\right)\eta^{s_2^\dagger}\left(ig_s\frac{(\bar{\mathbf{p}}+\bar{\mathbf{p}}')\cdot\boldsymbol{\varepsilon}_\gamma^*}{2m}t^a\right)\eta^{s_2'}, \tag{7c}$$

where the transversity condition $\mathbf{k}\cdot\boldsymbol{\varepsilon}_\gamma^* = 0$ has been utilized. For simplicity, we have suppressed the color indices associated with quark Pauli spinors.

The momenta carried by the quark propagators and gluon propagators in Fig. 1 are

$$l_{Q1}^0 \equiv p'^0 + |\mathbf{k}| - m \sim \mathcal{O}(mv), \qquad \mathbf{l}_{Q1} = \mathbf{p}' + \mathbf{k} \sim \mathcal{O}(mv), \tag{8a}$$

$$l_{Q2}^0 \equiv p^0 - |\mathbf{k}| - m \sim \mathcal{O}(mv), \qquad \mathbf{l}_{Q2} = \mathbf{p} - \mathbf{k} \sim \mathcal{O}(mv), \tag{8b}$$

$$l_g \equiv p - p' - k = q - q' - \frac{k}{2} \sim \mathcal{O}(mv). \tag{8c}$$

Note we have subtracted the quark mass in the zeroth-component of the quark 4-momentum, in accordance with the nonrelativistic dispersion relation. From (8) one immediately observes that the momenta affiliated with both quark and gluon propagators belong to the semi-hard family. The gluon propagator can be simplified as

$$\frac{1}{l_g^2+i\epsilon} = \frac{1}{(q-q'-\frac{k}{2})^2+i\epsilon} \approx \frac{-1}{(\mathbf{q}'-\mathbf{q})^2+\mathbf{k}\cdot(\mathbf{q}'-\mathbf{q})-i\boldsymbol{\epsilon}}, \tag{9}$$

which scales homogenously as $1/(mv)^2$.

The nonrelativistic quark propagators become inhomogeneous in the soft region. The quark kinetic energy scales as $mv^2$, which is subleading in $v$ compared with $l_{Q1}^0$. At the lowest order in $v$, it is legitimate to retain only the $l_{Q1}^0$ piece and resulting the following eikonal propagators:

$$\frac{1}{l_{Q1}^0-\frac{\mathbf{l}_{Q1}^2}{2m}} = \frac{1}{l_{Q1}^0}\left(1+\frac{\mathbf{l}_{Q1}^2}{2ml_{Q1}^0}+\cdots\right) \approx \frac{1}{l_{Q1}^0} \approx \frac{1}{|\mathbf{k}|}, \tag{10a}$$

$$\frac{1}{l_{Q2}^0-\frac{\mathbf{l}_{Q2}^2}{2m}} = \frac{1}{l_{Q2}^0}\left(1+\frac{\mathbf{l}_{Q2}^2}{2ml_{Q2}^0}+\cdots\right) \approx \frac{1}{l_{Q2}^0} \approx -\frac{1}{|\mathbf{k}|}, \tag{10b}$$

which simply scales as $1/mv$. This recipe for the soft quark propagator has been expounded in the context of strategy of region [10, 11]. To put in other word, we can obtain this result by re-interpreting (1) as the heavy quark effective theory (HQET) Lagrangian [12, 13] which contains the spin-independent $\mathcal{O}(1/m)$ operator. The eikonal propagators in (10) are nothing but the quark propagators given in HQET. It is interesting to note that in Fig. 1, the upper half diagrams above the horizontal cut are reminiscent of the soft Compton scattering process, in which the soft massless gauge bosons scatter off a static heavy quark. In a sense, it is more appropriate to account for such a process from the angle of HQET rather than NRQCD [2].

Inspecting (7), we notice that the amplitude of Fig. 1$c$) (with a seagull vertex) is suppressed by a factor $\propto |\mathbf{k}|/m \sim v$ relative to the amplitudes of Fig. 1$a$) and $b$), hence can be safely discarded.

One can also employ the equality $\mathbf{p}' \cdot \boldsymbol{\varepsilon}_\gamma^* = \mathbf{q}' \cdot \boldsymbol{\varepsilon}_\gamma^*$ and $\mathbf{p} \cdot \boldsymbol{\varepsilon}_\gamma^* = \mathbf{q} \cdot \boldsymbol{\varepsilon}_\gamma^*$ in (7), since $\mathbf{P}' \cdot \boldsymbol{\varepsilon}_\gamma^* = \mathbf{P} \cdot \boldsymbol{\varepsilon}_\gamma^* = 0$. Further implementing the homogenized soft gluon and quark propagators given in (9) and (10), we obtain the full quark-level amplitude at the lowest order in $v$:

$$\mathcal{A}[Q\overline{Q} \to Q\overline{Q} + \gamma(k)] = \frac{ee_Q g_s^2}{2m|\mathbf{k}|}\, \xi^{s_1'\dagger} t^a \xi^{s_1} \eta^{s_2\dagger} t^a \eta^{s_2'}\, \boldsymbol{T}(\mathbf{q}' - \mathbf{q}) \cdot \boldsymbol{\varepsilon}_{\gamma^*}, \tag{11}$$

where

$$\boldsymbol{T}(\mathbf{w}) = \frac{2\mathbf{w}}{\mathbf{w}^2 + \mathbf{w} \cdot \mathbf{k} - i\epsilon} + (\mathbf{k} \to -\mathbf{k}). \tag{12}$$

The second term originates from those mirror diagrams where the photon is emitted from the heavy antiquark.

It is worth emphasizing that, the above derivation amounts to integrating out the semi-hard quantum fluctuation in perturbative NRQCD. A tacit assumption is that the perturbation theory is still applicable, *i.e.*, the characteristic strong coupling constant $\alpha_s(mv) \ll 1$. This assumption is expected to be solid for bottomonia, and might be reasonable even for charmonia.

### B. Folding hard-scattering kernel with bound-state wave functions

Having obtained the quark amplitude for $Q\overline{Q} \to Q\overline{Q} + \gamma(k)$, we proceed to derive the corresponding $E1$ transition rate between two physical quarkonia states. To form a bound

[2] A cautious reader might worry about the incompatibility between the kinematic setup in (6) and the assumption of the semi-hard gluon propagator. Some inconsistency may arise from the lower half of Fig. 1, since it is impossible for a potential quark to remain potential after being stuck by a semi-hard gluon. The root of this inconsistency is that in order for $k \sim mv$, the parent quarkonium has to be a highly-excited state, so that at the quark level, the on-shell condition $\bar{p}^2 = m^2$ cannot be simultaneously satisfied in addition to the condition $p^2 = m^2$. Nevertheless, one may hypothesize that the spectator antiquark in the initial state is off-shell by an amount of $(mv)^2$, or it still remains on-shell albeit with a larger effective mass, $\bar{p}^2 \equiv m_{\rm eff}^2 \approx m^2 + \mathcal{O}(mv)^2$. In any event, the role of the lowest half diagram of Fig. 1 is just to generate a simple mass-independent NRQCD amplitude $\eta^{s_2\dagger}(ig_s t^a)\eta^{s_2'}$. Therefore, we conclude that the conflict between the kinematic assignment in (6) and the semi-hard gluon propagator will not spoil the validity of the NREFT amplitude (7).

state, the heavy quark and antiquark are assumed to be nearly on-shell (potential), exchanging infinite ladders of potential gluons. The corresponding binding dynamics are governed by nonperturbative QCD. To make a physical prediction for $E1$ transition, we have to fold the "hard-scattering" kernel in (11) with the nonperturbative wave functions of the initial- and final-state quarkonia.

At the lowest order in $v$, the state vector of a quarkonium $H$ bearing the quantum number $n^{2S+1}L_J$, can be explicitly constructed in terms of the nonrelativistic heavy quark-antiquark basis [14]

$$\left|H(n^{2S+1}L_J)(\mathbf{P},J_z)\right\rangle = \sum_{i,j,s_1,s_2,S_z,L_z} \frac{\delta^{ij}}{\sqrt{N_c}} \left\langle \frac{1}{2}s_1;\frac{1}{2}s_2 \middle| S\,S_z \right\rangle \langle S\,S_z;L\,L_z|J\,J_z\rangle$$
$$\times \int \frac{d^3\mathbf{q}}{(2\pi)^3}\widetilde{R}_{nL}(|\mathbf{q}|)Y_{LL_z}(\hat{\mathbf{q}}) \left|Q^i\left(\frac{\mathbf{P}}{2}+\mathbf{q},s_1\right)\overline{Q}^j\left(\frac{\mathbf{P}}{2}-\mathbf{q},s_2\right)\right\rangle, \tag{13}$$

where $i,j=1,2,3$ denote the color indices, $s_1$ and $s_2$ characterize the magnetic quantum number of the quark and antiquark. $\widetilde{R}_{nL}(|\mathbf{q}|)$ denotes the radial wave function for quarkonium in momentum space, $Y_{LL_z}(\hat{\boldsymbol{q}})$ signifies the spherical harmonics. $\langle\frac{1}{2}s_1;\frac{1}{2}s_2|S\,S_z\,\rangle$ and $\langle\,S\,S_z\,;L\,L_z|J\,J_z\,\rangle$ correspond to the Clebsch-Gordan coefficients affiliated with spin-spin coupling and spin-orbital coupling, where $S_z$ and $L_z$ signify the magnetic quantum numbers corresponding to the spin and orbital angular momentum.

One readily verifies that the quarkonium state $H$ defined in (13) obeys the standard non-relativistic normalization:

$$\langle H(n^{2S+1}L_J)\,(\mathbf{P}_1)\,|H(n^{2S+1}L_J)\,(\mathbf{P}_2)\rangle = (2\pi)^3\,\delta^{(3)}(\mathbf{P}_1-\mathbf{P}_2), \tag{14}$$

provided that the momentum-space quarkonium wave function is subject to the following orthogonality condition:

$$\int\frac{d^3\mathbf{q}}{(2\pi)^3}\widetilde{R}_{nL}(|\mathbf{q}|)\widetilde{R}^*_{n'L'}(|\mathbf{q}|)Y_{LL_z}(\hat{\mathbf{q}})Y^*_{L'L'_z}(\hat{\mathbf{q}}) = \delta_{nn'}\delta_{LL'}\delta_{L_zL'_z}. \tag{15}$$

The selection rule of the $E1$ transition is that the orbital angular momentum changes by one unit while keeping spin unchanged. For definiteness, in this work we will concentrate on the situation where the $P$-wave quarkonium transitions into a $S$-wave quarkonium.

The $E1$ transition amplitude becomes

$$\mathcal{A}[n^{2S+1}P_J \to n'^{2S+1}S_S+\gamma] = \int\frac{d^3\mathbf{q}}{(2\pi)^3}\int\frac{d^3\mathbf{q}'}{(2\pi)^3}\frac{1}{N_c}\sum_{i,s_1,s_2,S_z,L_z}\left\langle\frac{1}{2}s_1;\frac{1}{2}s_2\middle|S\,S_z\right\rangle\langle S\,S_z;1\,L_z|J\,J_z\rangle$$
$$\times\sum_{i',s'_1,s'_2}\left\langle S'\,S'_z\middle|\frac{1}{2}s'_1;\frac{1}{2}s'_2\right\rangle\widetilde{R}_{nP}(|\mathbf{q}|)Y_{1L_z}(\hat{\mathbf{q}})\widetilde{R}^*_{n'S}(|\mathbf{q}'|)Y^*_{00}(\hat{\mathbf{q}}') \tag{16}$$
$$\times\mathcal{A}[Q^i(s_1)\overline{Q}^i(s_2)\to Q^{i'}(s'_1)\overline{Q}^{i'}(s'_2)+\gamma(k)],$$

The quark and anti-quark can form either spin singlet or spin triplet. It is convenient to

adopt the following spin-spin coupling formula:

$$\sum_{s_1,s_2}\langle\frac{1}{2}s_1;\frac{1}{2}s_2|0\,0\rangle\xi^{s_1}\eta^{s_2\dagger}=\frac{1}{\sqrt{2}}\,\mathbf{1}_{2\times2},\qquad\sum_{s_1',s_2'}\langle 0\,0|\frac{1}{2}s_1';\frac{1}{2}s_2'\rangle\eta^{s_2'}\xi^{s_1'\dagger}=\frac{1}{\sqrt{2}}\,\mathbf{1}_{2\times2},\tag{17a}$$

$$\sum_{s_1,s_2}\langle\frac{1}{2}s_1;\frac{1}{2}s_2|1\,S_z\rangle\xi^{s_1}\eta^{s_2\dagger}=\frac{1}{\sqrt{2}}\boldsymbol{\sigma}\cdot\boldsymbol{\varepsilon}(S_z),\qquad\sum_{s_1',s_2'}\langle 1\,S_z'|\frac{1}{2}s_1';\frac{1}{2}s_2'\rangle\eta^{s_2'}\xi^{s_1'\dagger}=\frac{1}{\sqrt{2}}\boldsymbol{\sigma}\cdot\boldsymbol{\varepsilon}^*(S_z'),\tag{17b}$$

where $\boldsymbol{\varepsilon}(S_z)$ refers to the polarization vector affiliated with the unit spin.

Implementing (17) into (16), we find

$$\mathcal{A}[n^1P_1(J_z)\to n'^1S_0+\gamma]=C_F\frac{ee_Qg_s^2}{2m|\mathbf{k}|}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\int\frac{d^3\mathbf{q}'}{(2\pi)^3}\left(\frac{1}{\sqrt{2}}\right)^2\mathrm{Tr}[\mathbf{1}_{2\times2}^2]\tag{18a}$$

$$\times\ \widetilde{R}_{nP}(|\mathbf{q}|)Y_{1J_z}(\hat{\mathbf{q}})\widetilde{R}^*_{n'S}(|\mathbf{q}'|)Y^*_{00}(\hat{\mathbf{q}'})\,\boldsymbol{T}(\mathbf{q}'-\mathbf{q})\cdot\boldsymbol{\varepsilon}^*_\gamma,$$

$$\mathcal{A}[n^3P_J(J_z)\to n'^3S_1(S_z')+\gamma]=C_F\frac{ee_Qg_s^2}{2m|\mathbf{k}|}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\int\frac{d^3\mathbf{q}'}{(2\pi)^3}\sum_{S_z,L_z}\langle 1\,S_z;1\,L_z|J\,J_z\rangle\tag{18b}$$

$$\times\ \widetilde{R}_{nP}(|\mathbf{q}|)Y_{1L_z}(\hat{\mathbf{q}})\widetilde{R}^*_{n'S}(|\mathbf{q}'|)Y^*_{00}(\hat{\mathbf{q}'})\left(\frac{1}{\sqrt{2}}\right)^2\mathrm{Tr}[\boldsymbol{\sigma}\cdot\boldsymbol{\varepsilon}^*(S_z')\boldsymbol{\sigma}\cdot\boldsymbol{\varepsilon}(S_z)]\boldsymbol{T}(\mathbf{q}'-\mathbf{q})\cdot\boldsymbol{\varepsilon}^*_\gamma,$$

where the hard-scattering kernel $\boldsymbol{T}$ is given in (12). Note we have conducted the color sum using $\mathrm{Tr}[t^at^a]=C_FN_c$, with $C_F=\frac{N_c^2-1}{2N_c}$.

Since the bound-state wave functions are usually solved in the coordinate space, we proceed to Fourier transforming (18) using

$$\widetilde{R}_{nL}(|\mathbf{q}|)Y_{LL_z}(\hat{\mathbf{q}})=\int d^3\mathbf{r}\,e^{-i\mathbf{q}\cdot\mathbf{r}}R_{nL}(r)Y_{LL_z}(\hat{\mathbf{r}}).\tag{19}$$

Since the hard-scattering kernel $\boldsymbol{T}$ depends on the quark relative momenta $\mathbf{q}$ and $\mathbf{q}'$ only through the linear combination $\mathbf{q}'-\mathbf{q}$, one readily conducts three-fold integration and only retains the integration over $\mathbf{r}$:

$$\begin{aligned}&\int\frac{d^3\mathbf{q}}{(2\pi)^3}\int\frac{d^3\mathbf{q}'}{(2\pi)^3}\int d^3\mathbf{r}\int d^3\mathbf{r}'\,e^{-i\mathbf{q}\cdot\mathbf{r}}\,e^{i\mathbf{q}'\cdot\mathbf{r}'}\,f(\mathbf{r})\,\boldsymbol{T}(\mathbf{q}'-\mathbf{q})g(\mathbf{r}')\\
&=\int d^3\mathbf{r}\,f(\mathbf{r})\int\frac{d^3\mathbf{w}}{(2\pi)^3}\,\boldsymbol{T}(\mathbf{w})\int d^3\mathbf{r}'\,g(\mathbf{r}')e^{i\mathbf{w}\cdot\mathbf{r}'}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\,e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\\
&=\int d^3\mathbf{r}\,f(\mathbf{r})\,g(\mathbf{r})\int\frac{d^3\mathbf{w}}{(2\pi)^3}\,e^{i\mathbf{w}\cdot\mathbf{r}}\boldsymbol{T}(\mathbf{w})\\
&=\int d^3\mathbf{r}\,f(\mathbf{r})\,g(\mathbf{r})\,\frac{1}{2\pi}e^{\frac{i}{2}kr}\left[\hat{\mathbf{r}}\,\frac{2i+kr}{r^2}\cos\left(\frac{1}{2}kr\cos\theta\right)+i\frac{\hat{\mathbf{k}}}{r}\sin\left(\frac{1}{2}kr\cos\theta\right)\right]\end{aligned}\tag{20}$$

where we have changed the integration variable from $\mathbf{q}'$ to $\mathbf{w}=\mathbf{q}'-\mathbf{q}$ in the second line.

It is convenient to employ the following representation of the spherical harmonics:

$$Y_{00}(\hat{\mathbf{r}})=\sqrt{\frac{1}{4\pi}},\qquad Y_{1L_z}(\hat{\mathbf{r}})=\sqrt{\frac{3}{4\pi}}\hat{\mathbf{r}}\cdot\boldsymbol{\varepsilon}(L_z),\tag{21}$$

where $\boldsymbol{\varepsilon}(L_z)$ refers to the polarization vector associated with the unit orbital angular momentum.

In light of (20) and (21), the $P$-to-$S$-wave one-photon transition amplitudes in (18) can be reduced to

$$\mathcal{A}[n^1P_1(J_z)\to n'^1S_0+\gamma]=\sqrt{\frac{1}{4\pi}}\sqrt{\frac{3}{4\pi}}C_F\frac{ee_Qg_s^2}{2m|\mathbf{k}|}\varepsilon^i(J_z)\varepsilon_\gamma^{j*}$$
$$\times\int dr\,r^2\,\frac{2i+kr}{r^2}e^{\frac{i}{2}kr}R_{nP}(r)R^*_{n'S}(r)\,F^{ij}(kr), \tag{22a}$$

$$\mathcal{A}[n^3P_J(J_z)\to n'^3S_1(S'_z)+\gamma]=\frac{\sqrt{3}}{4\pi}C_F\frac{ee_Qg_s^2}{2m|\mathbf{k}|}\sum_{S_z,L_z}\langle 1\,S_z;1\,L_z|J\,J_z\rangle(\boldsymbol{\varepsilon}^*(S'_z)\cdot\boldsymbol{\varepsilon}(S_z))\,\varepsilon^i(L_z)\varepsilon_\gamma^{j*}$$
$$\times\int dr\,r^2\frac{2i+kr}{r^2}e^{\frac{i}{2}kr}R_{nP}(r)R^*_{n'S}(r)\,F^{ij}(kr), \tag{22b}$$

where the rank-2 symmetric tensor can be readily worked out upon performing the angular integration:

$$F^{ij}(kr)\equiv\int\frac{d\Omega}{4\pi}\cos\left(\frac{1}{2}kr\cos\theta\right)\hat{\mathbf{r}}^i\hat{\mathbf{r}}^j=\frac{2}{kr}j_1\left(\frac{kr}{2}\right)\delta^{ij}+\left[j_0\left(\frac{kr}{2}\right)-\frac{6}{kr}j_1\left(\frac{kr}{2}\right)\right]\hat{\mathbf{k}}^i\hat{\mathbf{k}}^j. \tag{23}$$

Due to the transversity condition $\mathbf{k}\cdot\boldsymbol{\varepsilon}_\gamma^*=0$, the second term in the bracket in (20) as well as the $\hat{\mathbf{k}}^i\hat{\mathbf{k}}^j$ piece in $F^{ij}(r)$ yields a vanishing contribution in (22), thus can be discarded.

For the spin-triplet $P$-wave states, there exists the following spin-orbital coupling formula [15]:

$$\sum_{S_z,L_z}\langle 1\,S_z;1\,L_z|0\,0\rangle\varepsilon^i(S_z)\varepsilon^j(L_z)=-\frac{1}{\sqrt{3}}\delta^{ij}, \tag{24a}$$

$$\sum_{S_z,L_z}\langle 1\,S_z;1\,L_z|1\,J_z\rangle\varepsilon^i(S_z)\varepsilon^j(L_z)=-\frac{i}{\sqrt{2}}\epsilon_{ijk}\varepsilon^k_{^3P_1}(J_z), \tag{24b}$$

$$\sum_{S_z,L_z}\langle 1\,S_z;1\,L_z|2\,J_z\rangle\varepsilon^i(S_z)\varepsilon^j(L_z)=\varepsilon^{ij}_{^3P_2}(J_z), \tag{24c}$$

where $\boldsymbol{\varepsilon}^k_{^3P_1}(J_z)$ denotes the polarization vector of the $^3P_1$ state, and $\varepsilon^{ij}_{^3P_2}(J_z)$ signifies the symmetric and traceless polarization tensor of the $^3P_2$ state.

With the aid of the spin-orbital coupling relation (24), the $P$-to-$S$-wave strongly hindered $E1$ transition amplitudes in (22) finally become

$$\mathcal{A}(n^1P_1\to n'^1S_0+\gamma)=\frac{-\sqrt{3}ee_Qg_s^2C_F}{2m\pi}\varepsilon_{nn'}(r)\,\boldsymbol{\varepsilon}_{^1P_1}\cdot\boldsymbol{\varepsilon}_\gamma^*, \tag{25a}$$

$$\mathcal{A}(n^3P_0\to n'^3S_1+\gamma)=\frac{ee_Qg_s^2C_F}{2m\pi}\varepsilon_{nn'}(r)\,\boldsymbol{\varepsilon}^*_{^3S_1}\cdot\boldsymbol{\varepsilon}_\gamma^*, \tag{25b}$$

$$\mathcal{A}(n^3P_1\to n'^3S_1+\gamma)=\frac{\sqrt{3}iee_Qg_s^2C_F}{2\sqrt{2}m\pi}\varepsilon_{nn'}(r)\,\boldsymbol{\varepsilon}_{^3P_1}\cdot\boldsymbol{\varepsilon}^*_{^3S_1}\times\boldsymbol{\varepsilon}_\gamma^*, \tag{25c}$$

$$\mathcal{A}(n^3P_2\to n'^3S_1+\gamma)=\frac{-\sqrt{3}ee_Qg_s^2C_F}{2m\pi}\varepsilon_{nn'}(r)\,\varepsilon^{ij}_{^3P_2}\varepsilon^{i*}_{^3S_1}\varepsilon_\gamma^{j*},. \tag{25d}$$

where the overlap integral $\varepsilon_{nn'}$ is defined by

$$\varepsilon_{nn'}(r) = \int dr\, r^2 R_{nP}(r) R_{n'S}(r) e^{\frac{i}{2}kr} \frac{2-ikr}{k^2\, r^3}\, j_1(\frac{kr}{2}). \tag{26}$$

Curiously, the overlap integral for the strongly hindered $E1$ transition looks similar to, but somewhat simpler than its counterpart in the strongly hindered $M1$ transition formula [7]:

$$\varepsilon^{M1}_{nn'}(r) = \int dr\, r^2 R_{nP}(r) R_{n'S}(r) e^{\frac{i}{2}kr} \frac{1}{r}\left[ j_0\left(\frac{kr}{2}\right) - \frac{2-ikr}{kr}\, j_1\left(\frac{kr}{2}\right)\right]. \tag{27}$$

Both of the overlap integrals in the strongly hindered $E1$ and $M1$ transitions are complex-valued. It looks a little bit unusual since we have only considered tree-level diagrams. From the calculational angle, the occurrence of the imaginary part can be attributed to the fact that the exchanged semi-hard gluon can become on-shell as $\mathbf{q} - \mathbf{q}' = \mathbf{k}$. Nevertheless, numerical study reveals that the real part dominates over the imaginary part in the overlap integral for both $E1$ and $M1$ transition processes. In most quarkonium $E1$ transition processes investigated in Sec. V, we find the real part of the overlap integral is several orders of magnitude greater than the imaginary part. Therefore, one may safely discard the contribution of the imaginary part in the overlap integrals in phenomenological analysis.

It is also interesting to examine the asymptotic expression of the overlap integral in the limit $kr \ll 1$. For the strongly hindered $M1$ transition, one finds asymptotically $e^{\frac{i}{2}kr}\frac{1}{r}\left[j_0\left(\frac{kr}{2}\right) - \frac{2-ikr}{kr}\, j_1\left(\frac{kr}{2}\right)\right] \to \frac{2}{3r} + \frac{ik}{2} + \cdots$ [7]. For the strongly hindered $E1$ transition, we find asymptotically $e^{\frac{i}{2}kr}\frac{2-ikr}{k^2\, r^3}\, j_1(\frac{kr}{2}) \to \frac{1}{3kr^2} + \frac{ik^2 r}{72} + \cdots$. At the limit $kr \ll 1$, it appears that the suppression of the imaginary part relative to the real part in the strongly hindered $E1$ transition is much more severe than the strongly hindered $M1$ transition. Finally, substituting this asymptotic expression into (26), the numerical values are very close to the rigorous results of the overlap integrals for various quarkonium $E1$ transition processes.

### C. Transition rates for strongly hindered $E1$ processes

Squaring the amplitudes, summing over final-state spins and averaging upon the initial-state polarizations, and using the formula

$$\Gamma(^{2S+1}L_J \xrightarrow{\mathrm{E}_1} {}^{2S+1}L'_{J'} + \gamma) = \frac{k}{2\pi}\frac{1}{2J+1}\sum_{\mathrm{Pol.}} |\mathcal{A}|^2\,, \tag{28}$$

we then obtain

$$\Gamma(n^3P_J \to n'^3S_1 + \gamma) = \frac{64 e_Q^2 \alpha \alpha_s^2 C_F^2 k}{M_i^2} |\varepsilon_{nn'}|^2, \tag{29a}$$

$$\Gamma(n^3S_1 \to n'^3P_J + \gamma) = \frac{64(2J+1) e_Q^2 \alpha \alpha_s^2 C_F^2 k}{3M_i^2} |\varepsilon_{nn'}|^2, \tag{29b}$$

$$\Gamma(n^1P_1 \to n'^1S_0 + \gamma) = \frac{64 e_Q^2 \alpha \alpha_s^2 C_F^2 k}{M_i^2} |\varepsilon_{nn'}|^2, \tag{29c}$$

$$\Gamma(n^1S_0 \to n'^1P_1 + \gamma) = \frac{192 e_Q^2 \alpha \alpha_s^2 C_F^2 k}{M_i^2} |\varepsilon_{nn'}|^2, \tag{29d}$$

for $J = 0, 1, 2$. For simplicity, we have approximated the heavy quark mass $m$ by $M_i/2$, half of mass of the initial-state quarkonium.

Equations (29) and (26) constitute the major new results of this work.

If we neglect the fine splitting among different $P$-wave spin triplet members, the transition rates satisfy the relation $\Gamma_0 : \Gamma_1 : \Gamma_2 = 1 : 1 : 1$ for $n^3P_{0,1,2} \to n'^3S_1 + \gamma$, and $\Gamma_0 : \Gamma_1 : \Gamma_2 = 1 : 3 : 5$ for the inverse processes $n^3S_1 \to n'^3P_{0,1,2} + \gamma$. These patterns are identical to what are derived from the conventional $E1$ transition formula (3), which is a consequence of heavy quark spin symmetry [16]. This is not surprising, since our new $E1$ transition mechanism also respects the heavy quark spin symmetry, which is a rigorous consequence of the lowest-order NRQCD Lagrangian.

Comparing (3) and (29), we immediately notice one important difference between the "hard-scattering" mechanism and conventional formula, that the sensitivity of the $E1$ transition rate to the photon momentum is much weaker in the former case.

## V. NUMERICAL RESULTS

In this section, we make a detailed numerical comparison between our new predictions arising from the hard-scattering mechanism and the conventional ones for $E1$ transition processes between $P$ and $S$-wave bottomonia and charmonia.

The key nonperturbative inputs are the Schrödinger wave functions for the parent and daughter quarkonium states. Assuming that color Coulomb force and linear confinement force are equally important for quarkonium dynamics, we choose to work with the Cornell potential model [17]:

$$V_{\rm Cornell}(r) = \frac{-\kappa}{r} + \frac{r}{a^2}, \tag{30}$$

where $\kappa = 0.52$ and $a = 2.34\,{\rm GeV}^{-1}$. In solving for the charmonia and bottomonia energy levels and the corresponding wave functions, we also take $m_c = 1.84$ GeV and $m_b = 5.18$ GeV [18]. The QCD coupling constant entering (29) should be evaluated at the characteristic semi-hard scale, and we take $\alpha_s(m_c v) = 0.59$ and $\alpha_s(m_b v) = 0.43$ [19].

TABLE I: The $E1$ transition rates for ${}^3S_1 \to {}^3P_J + \gamma$ in bottomonia family, predicted from both the "hard-scattering" mechanism and traditional multipole expansion method. We also include experimental measurements whenever available.

| process | $k$(MeV) | $\Gamma_{\rm exp}$(keV) | $\Gamma_{\rm hard\ scatt.}$ (keV) | $\Gamma_{\rm multipole}$(keV) |
|---|---|---|---|---|
| $\Upsilon(3S) \to \chi_{b0}(1P)\gamma$ | 484 | 0.055 | 0.20 | 0.12 |
| $\Upsilon(3S) \to \chi_{b1}(1P)\gamma$ | 452 | 0.018 | 0.64 | 0.29 |
| $\Upsilon(3S) \to \chi_{b2}(1P)\gamma$ | 434 | 0.203 | 1.11 | 0.43 |
| $\Upsilon(3S) \to \chi_{b0}(2P)\gamma$ | 121 | 1.199 | 1.18 | 1.53 |
| $\Upsilon(3S) \to \chi_{b1}(2P)\gamma$ | 100 | 2.560 | 4.34 | 2.54 |
| $\Upsilon(3S) \to \chi_{b2}(2P)\gamma$ | 86 | 2.662 | 8.42 | 2.69 |
| $\Upsilon(2S) \to \chi_{b0}(1P)\gamma$ | 163 | 1.215 | 1.00 | 1.46 |
| $\Upsilon(2S) \to \chi_{b1}(1P)\gamma$ | 129 | 2.207 | 3.82 | 2.19 |
| $\Upsilon(2S) \to \chi_{b2}(1P)\gamma$ | 110 | 2.287 | 7.48 | 2.28 |

TABLE II: Same as Table I, but the $E1$ processes are replaced with ${}^3P_J \to {}^3S_1+\gamma$ in bottomonium family.

| process | $k$(MeV) | $\Gamma_{\rm exp}$(keV) | $\Gamma_{\rm hard\ scatt.}$ (keV) | $\Gamma_{\rm multipole}$(keV) |
|---|---|---|---|---|
| $\chi_{b0}(2P) \to \Upsilon(1S)\gamma$ | 744 | - | 18.05 | 10.74 |
| $\chi_{b1}(2P) \to \Upsilon(1S)\gamma$ | 764 | - | 17.54 | 11.64 |
| $\chi_{b2}(2P) \to \Upsilon(1S)\gamma$ | 777 | - | 17.23 | 12.25 |
| $\chi_{b0}(2P) \to \Upsilon(2S)\gamma$ | 208 | - | 12.44 | 9.22 |
| $\chi_{b1}(2P) \to \Upsilon(2S)\gamma$ | 229 | - | 11.25 | 12.40 |
| $\chi_{b2}(2P) \to \Upsilon(2S)\gamma$ | 243 | - | 10.61 | 14.75 |
| $\chi_{b0}(1P) \to \Upsilon(1S)\gamma$ | 391 | - | 44.82 | 20.00 |
| $\chi_{b1}(1P) \to \Upsilon(1S)\gamma$ | 424 | - | 41.24 | 25.43 |
| $\chi_{b2}(1P) \to \Upsilon(1S)\gamma$ | 442 | - | 39.48 | 28.85 |

TABLE III: Same as Table II, but the $E1$ processes are replaced with ${}^1P_1 \to {}^1S_0+\gamma$ in bottomonium family.

| process | $k$(MeV) | $\Gamma_{\rm exp}$(keV) | $\Gamma_{\rm hard\ scatt.}$(keV) | $\Gamma_{\rm multipole}$(keV) |
|---|---|---|---|---|
| $h_b(2P) \to \eta_b(1S)\gamma$ | 825 | - | 16.37 | 14.64 |
| $h_b(2P) \to \eta_b(2S)\gamma$ | 258 | - | 10.04 | 17.58 |
| $h_b(1P) \to \eta_b(1S)\gamma$ | 487 | - | 36.08 | 38.75 |
| $\eta_b(2S) \to h_b(1P)\gamma$ | 99 | - | 15.03 | 3.01 |

After numerically obtaining quarkonia radial wave functions, we then use (29) and (26) to calculate the partial width from "hard-scattering" mechanism for various $E1$ transition processes. In Table I-IV we enumerate our predictions for various $E1$ transition rates for bottomonia and charmonia family. For the sake of comparison, we also juxtapose the predictions from the conventional multipole expansion method, (3), together with the measured values whenever available [3].

For $E1$ transitions in both bottomonium and charmonium, the numerical predictions from the hard scattering mechanism are different from those obtained in the conventional $E1$ formula [3] In the case of the small photon momentum, the multipole expansion picture is expected to hold, and we do not anticipate that the predictions from our new mechanism are trustworthy anymore. As a matter of fact, in most observed channels with small photon momentum, the predictions from the conventional formula are much closer to the measured values than those from the "hard-scattering" mechanism.

We certainly anticipate our predictions become more reliable than the conventional predictions when the photon momentum exceeds 800 MeV. The transition processes $\chi_{b0,1,2} \to \Upsilon(1S)+\gamma$ are of special interest, since the photon momenta in these processes become quite close to 800 MeV. From Table II we observe that the predictions from the "hard-scattering" mechanism are considerably larger than those from the conventional mechanism. Similarly,

[3] It is curious to note that, the predictions from two very different mechanisms appear to differ only by a few times. This is nontrivial since there are no free adjustable parameters in both (3) and (29).

TABLE IV: The $E1$ transition rates in charmonium family, predicted from both the "hard-scattering" mechanism and traditional multipole expansion method. We also include experimental measurements whenever available.

| process | $k$(MeV) | $\Gamma_{\rm exp}$(keV) | $\Gamma_{\rm hard\ scatt.}$ (keV) | $\Gamma_{\rm multipole}$(keV) |
|---|---|---|---|---|
| $\chi_{c0}(1P) \to J/\psi\gamma$ | 303 | 151.2 | 436.29 | 141.07 |
| $\chi_{c1}(1P) \to J/\psi\gamma$ | 390 | 288.12 | 329.41 | 299.29 |
| $\chi_{c2}(1P) \to J/\psi\gamma$ | 429 | 374.30 | 295.12 | 400.66 |
| $\psi(2S) \to \chi_{c0}(1P)\gamma$ | 261 | 28.783 | 8.67 | 50.52 |
| $\psi(2S) \to \chi_{c1}(1P)\gamma$ | 171 | 28.665 | 41.54 | 42.49 |
| $\psi(2S) \to \chi_{c2}(1P)\gamma$ | 128 | 27.989 | 94.11 | 29.58 |
| $h_c(1P) \to \eta_c(1S)\gamma$ | 499 | 350 | 264.39 | 630.70 |
| $\eta_c(2S) \to h_c(1P)\gamma$ | 111 | - | 200.63 | 35.19 |

the momentum of the radiated photon in the process $h_b \to \eta_b(1S) + \gamma$ is greater than 800 MeV. From Table III, one observes that the prediction from the new mechanism is slightly bigger than that from the conventional one. Hopefully, the future precise measurements of these $E1$ transition channels will unambiguously test whether the "hard-scattering" mechanism is more successful relative to the conventional one or not.

From Table I-IV, we also numerically confirm that the predictions from the "hard-scattering" mechanism indeed have a weaker sensitivity to the photon momentum than those from the conventional mechanism. This is supported by the fact that the predictions from the new mechanism respects the pattern dictated by the heavy quark spin symmetry to a better extent.

Last but not least, it is worth mentioning a decade-long puzzle, *i.e.*, the measured $E1$ transition rates for $\Upsilon(3S) \to \chi_{b0,1,2}\gamma$ processes exhibit an anomalous pattern, which badly violates the cherished heavy quark spin symmetry. Clearly this anomaly cannot be accounted by both the conventional and the new hard-scattering mechanism. There have been some efforts attempting to resolve this puzzle by including a variety of relativistic corrections in quark potential models [20–25]. It is difficult to conceive that the heavy quark spin symmetry can be badly violated for the bottom quark sector, especially concerning the fact $v^2 \approx 0.1$ for bottomonium. Obviously, future independent experiments will be essential to finally settle down this puzzle.

## VI. SUMMARY

The electromagnetic transition between heavy quarkonia is a widely studied topic. For a very small photon momentum, the dominant $E1$ transition is adequately described by the standard multipole expansion approach.

In sharp contrast to the atomic system, the linear confinement potential in QCD allows the mass difference between the highly excited quarkonium and the ground state to be as large as 1 GeV. In such situation, the photon can no longer be viewed as ultrasoft, and the multipole expansion doctrine ceases to work. It is more appropriate to regard the photon momentum as semihard, and adopt the hard-scattering mechanism to account for these strongly hindered EM transitions, where a semihard gluon must be exchanged between

heavy quark and antiquark during the single-photon emission.

In this work, we apply the "hard-scattering" approach originally developed to tackle the strongly hindered $M1$ transition a decade ago to deal with the strongly hindered $E1$ transitions between quarkonia. We derive the compact formulas for the corresponding transition rates in the context of the non-relativistic EFT, at the lowest order in the strong coupling and velocity expansion. Although it is possible to deduce the identical expression from the relativistic QCD, the physics becomes much more transparent from the angle of NREFT. We have also conducted a detailed numerical study for various $E1$ transition processes between $P$- and $S$-wave bottomonia and charmonia, and confront our new predictions with the conventional ones. Unfortunately, unlike the $M1$ case, the emitted photon in all the observed $E1$ transition channels in bottomonia and charmonia is not energetic enough to justify the predictions from the hard-scattering mechanism. Nevertheless, the photon energy already reaches 800 MeV in the decay channels $\chi_{b0,1,2} \to \Upsilon(1S) + \gamma$ and $h_b \to \eta_b(1S) + \gamma$, and the predictions from the "hard-scattering" mechanism are considerably larger those from the conventional approach. Therefore, we hope that the future precise measurements of these $E1$ transition channels can examine the validity of the "hard-scattering" mechanism. In the future, we also plan to further apply this hard-scattering formalism to predict the strongly hindered $D$-to-$P$-wave quarkonia $E1$ transition, and make some testable predictions in close contact with phenomenology.

**Acknowledgments**

We are grateful to Jichen Pan for discussions. This work is supported in part by the National Natural Science Foundation of China under Grants No. 11925506. The work of Z. W. M. is also supported in part by the National Science Foundation of China No. 12347145, No. 12022514, No. 12375099 and No. 12047503, and National Key Research and Development Program of China No. 2020YFC2201501 and No. 2021YFA0718304. The work of J.-Y. Z. is also supported in part by the US Department of Energy (DOE) Contract No. DE-AC05-06OR23177, under which Jefferson Science Associates, LLC operates Jefferson Lab.

**Appendix A: strongly hindered $E1$ transition from QCD**

In this appendix, we show that our key results, (26) and (29), can also be directly reproduced from the underlying relativistic field theory, *e.g.*, QCD plus QED. As a matter of fact, the strongly hindered $M1$ transition in the hard-scattering scattering has been derived from this angle [7]. At the lowest order in $\alpha_s$, there are four Feynman diagrams for $Q\overline{Q} \to Q\overline{Q} + \gamma$ in the hard-scattering mechanism, two of which are depicted in Fig. 2. Note the seagull vertex in Fig. 1 arising from the NREFT is absent in the underlying relativistic field theory. The kinematical setting is identical with (6).

Let us consider the single-photon transition process $n^{2S+1}P_J \to n'^{2S+1}S_S + \gamma$. The starting point is (16), except the quark amplitude is now computed from underlying relativistic theory

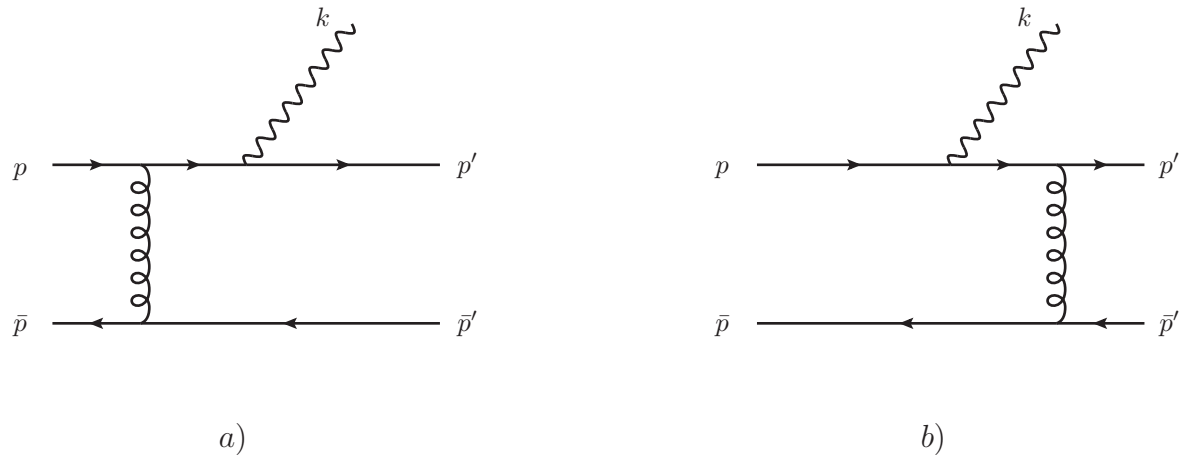



FIG. 2: The lowest-order Feynman diagrams in QCD+QED that contribute to the strongly hindered $E1$ transition between heavy quarkonia arising from the hard-scattering mechanism. For simplicity, we have suppressed two mirror diagrams where the photon is emitted from the heavy anti-quark line.

rather than from NREFT:

$$\mathcal{A}[n^{2S+1}P_J \to n'^{2S+1}S_S + \gamma] = \int \frac{d^3\mathbf{q}}{(2\pi)^3} \int \frac{d^3\mathbf{q}'}{(2\pi)^3} \frac{1}{N_c} \sum_{i,s_1,s_2,S_z,L_z} \left\langle \frac{1}{2}s_1; \frac{1}{2}s_2 \middle| S\, S_z \right\rangle \langle S\, S_z; 1\, L_z | J\, J_z \rangle$$
$$\times \sum_{i',s_1',s_2'} \left\langle S'\, S_z' \middle| \frac{1}{2}s_1'; \frac{1}{2}s_2' \right\rangle \widetilde{R}_{nP}(|\mathbf{q}|) Y_{1L_z}(\hat{\mathbf{q}}) \widetilde{R}^*_{n'S}(|\mathbf{q}'|) Y^*_{00}(\hat{\mathbf{q}}') \tag{A1}$$
$$\times \mathcal{A}[Q^i(s_1)\overline{Q}^i(s_2) \to Q^{i'}(s_1')\overline{Q}^{i'}(s_2') + \gamma(k)].$$

The quark amplitude $\mathcal{A} = \mathcal{A}_a + \mathcal{A}_b +$ two mirror diagrams, which in Feynman gauge reads

$$\mathcal{A}_a = iee_Q g_s^2 \frac{[\bar{u}(p')\not{\epsilon}^*_\gamma(\not{p}' + \not{k} + m)t^a\gamma^\alpha u(p)]\,[\bar{v}(\bar{p})t^a\gamma_\alpha v(\bar{p}')]}{((p'+k)^2 - m^2)((q - q' - \frac{k}{2})^2 + i\epsilon)}, \tag{A2a}$$

$$\mathcal{A}_b = iee_Q g_s^2 \frac{[\bar{u}(p')t^a\gamma^\alpha(\not{p} - \not{k} + m)\not{\epsilon}^*_\gamma u(p)]\,[\bar{v}(\bar{p})t^a\gamma_\alpha v(\bar{p}')]}{((p-k)^2 - m^2)((q - q' - \frac{k}{2})^2 + i\epsilon)}. \tag{A2b}$$

It is convenient to employ the covariant projector technique [26] to enforce the $Q\overline{Q}$ pair to bear the spin-singlet or spin-triplet quantum numbers. For the $Q\overline{Q}$ pair in the initial state, one can make the following substitution in (A2):

$$\sum_{s_1,s_2} u(p,s_1)\bar{v}(\bar{p},s_2)\langle \frac{1}{2}s_1\frac{1}{2}s_2|00\rangle = \frac{-1}{2\sqrt{2}E(E+m)}(\not{p} + m)\frac{\not{P} + 2E}{4E}\gamma_5(\not{\bar{p}} - m) \otimes \frac{\mathbf{1}_c}{\sqrt{N_c}}, \tag{A3a}$$

$$\sum_{s_1,s_2} u(p,s_1)\bar{v}(\bar{p},s_2)\langle \frac{1}{2}s_1\frac{1}{2}s_2|1S_z\rangle = \frac{1}{2\sqrt{2}E(E+m)}(\not{p} + m)\frac{\not{P} + 2E}{4E}\not{\epsilon}(S_z)(\not{\bar{p}} - m) \otimes \frac{\mathbf{1}_c}{\sqrt{N_c}}, \tag{A3b}$$

with $P^2 \equiv 4E^2 \approx 4m^2$. $\mathbf{1}_c$ signifies the $3 \times 3$ unit matrix to enforce the projection of the $Q\overline{Q}$ pair onto the color-neutral state. $\varepsilon(S_z)$ denotes the polarization vector for the unit spin. For the $Q\overline{Q}$ pair in the final state, one can make the analogous substitution in (A2), by adapting (A3) accordingly with the aid of the identity $v\bar{u} = \gamma^0(u\bar{v})^\dagger\gamma^0$. Note that the Dirac spinors in (A3) are normalized in nonrelativistic convention.

In the hard-scattering picture, both quark and gluon propagators in Fig. 2 should be viewed as semi-hard, with power counting of various momenta expounded in Sec. IV A. Similar to the line of reasoning leading to (9) and (10), the quark and gluons propagators in Fig. 2 can be simplified as

$$\frac{1}{(p'+k)^2-m^2} = \frac{1}{k\cdot P' + 2k\cdot q'} \approx \frac{1}{k\cdot P'} \approx \frac{1}{k\cdot P}, \tag{A4a}$$

$$\frac{1}{(p-k)^2-m^2} = -\frac{1}{k\cdot P + 2k\cdot q} \approx -\frac{1}{k\cdot P}, \tag{A4b}$$

$$\frac{1}{(q-q'-\frac{k}{2})^2+i\epsilon} \approx \frac{-1}{(\mathbf{q}'-\mathbf{q})^2 + \mathbf{k}\cdot(\mathbf{q}'-\mathbf{q}) - i\epsilon}. \tag{A4c}$$

Implementing the expanded propagators in (A4), performing the spin and color projection using (A3), retaining only the linear terms in $\boldsymbol{q}$ and $\boldsymbol{q'}$ in the numerators, the transition amplitude for $n^{2S+1}P_J \to n'^{2S+1}S_S + \gamma$ in (A1) can be reduced to

$$\mathcal{A}[n^1P_1(J_z) \to n'^1S_0 + \gamma] = C_F\frac{ee_Qg_s^2}{k\cdot P}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\int\frac{d^3\mathbf{q}'}{(2\pi)^3} \tag{A5a}$$
$$\times\widetilde{R}_{nP}(|\mathbf{q}|)Y_{1J_z}(\hat{\mathbf{q}})\widetilde{R}^*_{n'S}(|\mathbf{q}'|)Y^*_{00}(\hat{\mathbf{q}}')\,\boldsymbol{T}(\mathbf{q}'-\mathbf{q})\cdot\boldsymbol{\varepsilon}^*_\gamma,$$
$$\mathcal{A}[n^3P_J(J_z) \to n'^3S_1(S'_z) + \gamma] = C_F\frac{ee_Qg_s^2}{k\cdot P}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\int\frac{d^3\mathbf{q}'}{(2\pi)^3}\sum_{S_z,L_z}\langle 1\,S_z; 1\,L_z|J\,J_z\rangle \tag{A5b}$$
$$\times\widetilde{R}_{nP}(|\mathbf{q}|)Y_{1L_z}(\hat{\mathbf{q}})\widetilde{R}^*_{n'S}(|\mathbf{q}'|)Y^*_{00}(\hat{\mathbf{q}}')\,(\boldsymbol{\varepsilon}^*(S'_z)\cdot\boldsymbol{\varepsilon}(S_z))\,\boldsymbol{T}(\mathbf{q}'-\mathbf{q})\cdot\boldsymbol{\varepsilon}^*_\gamma,$$

where the hard-scattering function $\boldsymbol{T}(\mathbf{q}'-\mathbf{q})$ has been given in (12). Hearteningly, these equations are exactly identical to (18) derived in NREFT. Following the remaining steps as elaborated in Sec. IV, we then reproduce the key formula of this work, (26) and (29).

---


[1] N. Brambilla *et al.* [Quarkonium Working Group], doi:10.5170/CERN-2005-005 [arXiv:hep-ph/0412158 [hep-ph]].

[2] E. Eichten, S. Godfrey, H. Mahlke and J. L. Rosner, Rev. Mod. Phys. **80**, 1161-1193 (2008) doi:10.1103/RevModPhys.80.1161 [arXiv:hep-ph/0701208 [hep-ph]].

[3] R. L. Workman *et al.* [Particle Data Group], PTEP **2022**, 083C01 (2022) doi:10.1093/ptep/ptac097

[4] P. A. M. Dirac, Proc. Roy. Soc. Lond. A **114**, 243 (1927) doi:10.1098/rspa.1927.0039

[5] N. Brambilla, Y. Jia and A. Vairo, Phys. Rev. D **73**, 054005 (2006) doi:10.1103/PhysRevD.73.054005 [arXiv:hep-ph/0512369 [hep-ph]].

[6] N. Brambilla, P. Pietrulewicz and A. Vairo, Phys. Rev. D **85**, 094005 (2012) doi:10.1103/PhysRevD.85.094005 [arXiv:1203.3020 [hep-ph]].

[7] Y. Jia, J. Xu and J. Zhang, Phys. Rev. D **82**, 014008 (2010) doi:10.1103/PhysRevD.82.014008 [arXiv:0901.4021 [hep-ph]].
[8] W. E. Caswell and G. P. Lepage, Phys. Lett. B **167**, 437-442 (1986) doi:10.1016/0370-2693(86)91297-9
[9] M. Beneke and V. A. Smirnov, Nucl. Phys. B **522**, 321-344 (1998) doi:10.1016/S0550-3213(98)00138-2 [arXiv:hep-ph/9711391 [hep-ph]].
[10] M. Beneke, A. Signer and V. A. Smirnov, Phys. Rev. Lett. **80**, 2535-2538 (1998) doi:10.1103/PhysRevLett.80.2535 [arXiv:hep-ph/9712302 [hep-ph]].
[11] H. W. Griesshammer, Phys. Rev. D **58**, 094027 (1998) doi:10.1103/PhysRevD.58.094027 [arXiv:hep-ph/9712467 [hep-ph]].
[12] E. Eichten and B. R. Hill, Phys. Lett. B **234**, 511-516 (1990) doi:10.1016/0370-2693(90)92049-O
[13] H. Georgi, Phys. Lett. B **240**, 447-450 (1990) doi:10.1016/0370-2693(90)91128-X
[14] S. Weinberg, *Lectures on Quantum Mechanics*, Cambridge University Press, 2015, ISBN 978-1-107-11166-0 doi:10.1017/CBO9781316276105
[15] J. H. Kuhn, J. Kaplan and E. G. O. Safiani, Nucl. Phys. B **157**, 125-144 (1979) doi:10.1016/0550-3213(79)90055-5
[16] P. L. Cho and M. B. Wise, Phys. Lett. B **346**, 129-136 (1995) doi:10.1016/0370-2693(94)01658-Y [arXiv:hep-ph/9411303 [hep-ph]].
[17] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane and T. M. Yan, Phys. Rev. D **17**, 3090 (1978) [erratum: Phys. Rev. D **21**, 313 (1980)] doi:10.1103/PhysRevD.17.3090
[18] E. J. Eichten and C. Quigg, Phys. Rev. D **52**, 1726-1728 (1995) doi:10.1103/PhysRevD.52.1726 [arXiv:hep-ph/9503356 [hep-ph]].
[19] E. Braaten, Nucl. Phys. B Proc. Suppl. **71**, 441-450 (1999) doi:10.1016/S0920-5632(98)00376-4 [arXiv:hep-ph/9810390 [hep-ph]].
[20] P. Moxhay and J. L. Rosner, Phys. Rev. D **28**, 1132 (1983) doi:10.1103/PhysRevD.28.1132
[21] S. N. Gupta, S. F. Radford and W. W. Repko, Phys. Rev. D **30**, 2424 (1984) doi:10.1103/PhysRevD.30.2424
[22] H. Grotch, D. A. Owen and K. J. Sebastian, Phys. Rev. D **30**, 1924 (1984) doi:10.1103/PhysRevD.30.1924
[23] T. A. Lahde, Nucl. Phys. A **714**, 183-212 (2003) doi:10.1016/S0375-9474(02)01362-3 [arXiv:hep-ph/0208110 [hep-ph]].
[24] D. Ebert, R. N. Faustov and V. O. Galkin, Phys. Rev. D **67**, 014027 (2003) doi:10.1103/PhysRevD.67.014027 [arXiv:hep-ph/0210381 [hep-ph]].
[25] A. M. Badalian and B. L. G. Bakker, Phys. Rev. D **86**, 074001 (2012) doi:10.1103/PhysRevD.86.074001 [arXiv:1203.0936 [hep-ph]].
[26] G. T. Bodwin and A. Petrelli, Phys. Rev. D **66**, 094011 (2002) [erratum: Phys. Rev. D **87**, no.3, 039902 (2013)] doi:10.1103/PhysRevD.66.094011 [arXiv:hep-ph/0205210 [hep-ph]].

arXiv:2402.19393v1 [hep-ph] 29 Feb 2024

# New Physics with PeV Astrophysical Neutrino Beams

By

Ibrahim Safa

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

University of Wisconsin – Madison

2022

Date of final oral examination: August 17, 2022

The dissertation is approved by the following members of the Final Oral Committee:

Francis Halzen, Professor, Physics

Carlos Argüelles, Assistant Professor, Physics

Ke Fang, Assistant Professor, Physics

Amy Barger, Professor, Astronomy

Benjamin Jones, Professor, Physics

*Yeah, this album is dedicated*

*To all the teachers that told me I'd never amount to nothin'*

*To all the people that lived above the buildings that I was hustlin' in front of*

*Called the police on me when I was just tryin' to make some money to feed my daughter (it's all good).*

The Notorious B.I.G (1972 - 1997)

## Abstract

Astrophysical neutrinos allow us to access energies and baselines that cannot be reached by human-made accelerators, offering unique probes of new physics phenomena. This thesis aims to address the challenges currently facing searches for Beyond Standard Model (BSM) physics in the high-energy universe using astrophysical neutrinos, particularly in the contexts of flavor measurements and connections with dark matter.

The search for new physics with astrophysical neutrinos requires as a prerequisite understanding standard neutrino sources, which remain ambiguous. We begin by performing a multi-wavelength search for astrophysical neutrino sources using nine years of IceCube data. We find hints of neutrino emission from radio-bright Active Galactic Nuclei (AGN), further supporting recent claims that neutrino emission occurs near the core of AGNs.

Next we turn our attention to BSM searches. Accurate flavor measurements of the astrophysical flux provide a smoking gun signature to BSM physics. This requires a precise measurement of the tau neutrino fraction. However, tau identification proved a major hurdle in the current generation of observatories. We confront the problem of astrophysical neutrino flavor measurements by first introducing `TauRunner`, a simulation tool that accurately models the propagation of tau neutrinos including previously neglected effects such as tau lepton energy losses and depolarization in matter. We show that better modeling of tau neutrino propagation improves IceCube transient point-source sensitivities by more than an order of magnitude at EeV energies, and diffuse flux sensitivities by a factor of two. Second, we use this software to model IceCube counterparts to anomalous events reported by the ANITA experiment. After performing an analysis using IceCube data, we show that all Standard Model explanations are ruled out. Looking ahead to the future of flavor measurements, we also present a study that predicts the production of tau neutrinos via the propagation of electron and muon neutrinos in Earth, finding an irreducible but quantifiable background to next-generation tau neutrino observatories.

Finally, we attempt to address the field's shared ignorance of the origin of neutrino and dark matter masses by exploring potential connections between the two. Specifically, we present an analysis of dark matter annihilation and decay to neutrinos. We obtain limits from MeV to ZeV masses using more than a dozen neutrino experiments. Notably, using recent data from the SuperKamiokande experiment, we place the first-ever limit on dark matter annihilation that reaches

the thermal relic abundance in the neutrino sector, challenging notions that studies with neutrinos cannot be sensitive enough to make strong claims about the nature of dark matter.

## Acknowledgements

I am not just thankful but honestly quite confounded by the chain of circumstances that led to this thesis because only ten years ago I had dropped out of high-school. Since then, undoubtedly a lot of things had to work out in my favor and I for one think it's really neat that they did.

First and foremost I want to address Francis. Thank you for always being in my corner, for never questioning why I wasn't making progress on the projects I was supposed to be working on, and for always being open to new ideas no matter how crazy. Your compassionate guidance and trust gave me security in myself, my abilities, and the science I am pursuing. That is a privilege that not many students have, and it sure made things a whole lot easier. I only hope to carry forward a fraction of the passion and vigor with which you approach research, and I promise to continue to "do things" rather than reading books.

Although Francis (pre-covid) was traveling more often than not, he always answered my messages within minutes. He also made sure I was surrounded by people who could teach me, hands-on, the grueling details. Carlos Argüelles, of course, immediately comes to mind. You are a fearless scientist. You told me that everyone wants glory, but few are willing to mop the floors first. You taught me that a scientist needs to know how to "actually do things"; to write code, to develop new tools, solve pesky problems, think critically especially about concepts that are eerily widely accepted. Only then can we find the anomalies, and that's where the good stuff is. I will always be grateful for these lessons, which you taught me as a mentor, friend, and eventually boss. Another mentor I had the privilege of working with is Ali Kheirandish. Your humane approach to science, idealism, and attention to detail imprinted on me the desire to do not just good science, but honest science.

I thank Ben Jones for introducing me to neutrinos, for believing in me when I had a lot of enthusiasm but very little knowledge to back it up, even though I proclaimed to be a PMT expert. You are a selfless advisor who is always looking out for the little guy, and you showed me that I can actually make a career out of this thing.

I also thank those whose love for the subject sparked in me the interest in Physics: Mohammad Cherri my high-school physics teacher, Greg Sherman at Collin College, and Collin Thomas who invited students to "come do crazy things" and then delivered on his promise when he gave me a microwave, some steel wool, and a resilient bacteria he'd found in his backyard to test theories of panspermia.

To my grad school companion and lifelong friend Jeff Lazar, thanks dawg it was good to hang. I also thank Leslie for stepping up when I needed it most. My first two years of graduate school were mostly held together by Lisa Everett and Jeff Schmidt. Their commitment to every student's success at UW-Madison was inspiring, I will never forget that. I thank my good friends Neil, Alex, Zach, Brent, Praful, and Milena for commiserating about graduate school and helping me get through it. And to the many friends who kept me sane and helped me along the way: Youssef, Jenny, Khalid, Chloe, Ahmad, Joe and Sam, Lindsey. Additionally, to those who taught me how to overcome hurdles by putting them smack dab in my way, thank you.

Most of all, I want to thank my mother, Ghada, my sister Sara, and my brothers Mo and Karim. Through all our struggles, you always had my back, and never once questioned my decisions. It means the world.

And finally, to Daphne. I'm so very glad I found you.

# Contents

Abstract . . . . . ii
Acknowledgements . . . . . iv

**1 Introduction** **1**

**2 Multiwavelength Search for Neutrino Emitters** **6**
2.1 Gamma rays are not the whole story . . . . . 7
2.2 Source Selection . . . . . 8
2.3 Analysis Method . . . . . 11
2.4 Hypotheses and Sensitivity . . . . . 13
2.5 Results and Interpretation . . . . . 14

**3 Tau Neutrinos Require Some Care** **17**
3.1 Tau Neutrinos through Matter . . . . . 18
3.2 Tau Energy Losses and Decays . . . . . 20
3.3 Tau depolarization in matter . . . . . 21
3.4 Production of Tau Neutrinos in Earth . . . . . 26

**4 `TauRunner`: A Comprehensive Solution to Lepton Propagation** **34**
4.1 Algorithm Overview . . . . . 34
4.2 Structure of the code . . . . . 35
4.3 Performance . . . . . 40
4.4 Results and Comparisons . . . . . 43
4.5 Implementation of tau depolarization . . . . . 46

**5 Implications of Accurate Tau Neutrino Modeling** **51**
5.1 GZK Neutrinos . . . . . 51
5.2 Point Source Transient Limits . . . . . 57
5.3 ANITA Followup deux façons . . . . . 59

**6 Neutrinos and Dark Matter** **68**
6.1 Motivation . . . . . 69
6.2 Dark Matter Density Profile . . . . . 71
6.3 Annihilation to neutrinos . . . . . 72
6.4 Velocity-dependent annihilation . . . . . 76
6.5 Decay to neutrinos . . . . . 79
6.6 Experimental methods . . . . . 82
6.7 Results . . . . . 90

**7 Conclusion** **106**

**List of Figures** **108**

**List of Tables** **111**

**Bibliography** **112**

# Chapter 1

# Introduction

> "Abstruse thought and profound researches I [Nature] prohibit, and will severely punish, by the pensive melancholy which they introduce, by the endless uncertainty in which they involve you, and by the cold reception which your pretended discoveries shall meet with, when communicated."
>
> David Hume

In 1930, Pauli hypothesized a neutral particle to explain anomalies in beta decay experiments and save energy conservation [1]. The neutrino, a particle so elusive that Pauli himself thought would never be detected, was indeed observed near nuclear reactors merely two decades later by Cowan and Reines [2]. Not long after their discovery, neutrinos started producing anomalies of their own.

In 1964, John Bahcall predicted a flux of electron neutrinos from the Sun produced by yet unconfirmed nuclear fusion at its core [3]. Davis and Bahcall then proposed a chlorine experiment which was sensitive to the capture of electron neutrinos on chlorine atoms. A few years later Davis, Harmer, and Hoffman would report the first detection of solar neutrinos [4, 5], confirming that nuclear fusion is indeed the fuel of stars; however, the measured flux from the Sun was only a third of what Bahcall had predicted. A host of experiments then confirmed this measurement. This became known as the solar neutrino problem. It would not have been a problem, of course, if Bahcall hadn't predicted (accurately) the solar neutrino flux, and then stood by his calculation following the first measurements.

Four decades later Super-Kamiokande would measure a similar effect, namely a deficit of the predicted flux, but this time using muon neutrinos produced in the atmosphere [6]. When the Sudbury Neutrino Observatory measured the solar flux with all neutrino flavors via neutral current

neutrino interactions [7], it became clear that both Bahcall and the measurements were correct. The truth was that neutrinos were changing flavor.

What was required to solve this problem was to measure the neutrino flux at several baselines and energies. These results would then be cemented by measurements of neutral current interactions, which are flavor-independent. The result meant that neutrinos must have mass and the Standard Model (SM) is incomplete.

We were able to accommodate neutrino oscillations with a parametric extension of the SM. Neutrinos would change flavor in-flight, and this flavor change was dependent on the neutrino energy (E) and the distance it traveled (L); however, the underlying mechanism which provides neutrinos their mass remains unknown. In fact, this ad-hoc extension of the SM has already started to break down. At the turn of the century, the Liquid Scintillator Neutrino Detector [8] measured a low-energy electron neutrino excess, followed by a confirmation by the MiniBoone detector, again at different energies and baselines [9]. This flavor anomaly, aptly named the MiniBoone anomaly, remains unresolved and suggests that the story is not yet over. What had started as a bold idea put forward by Pauli, followed by an accurate calculation by Bahcall, is now forcing us to rethink our most successful theory of interactions in nature to date.

The recipe is now clear. In order to understand the underlying neutrino mass mechanism, we must rely on a symbiosis of accurate theoretical predictions and a variety of neutrino measurements at all possible baselines and energies, until new anomalies reveal themselves exposing the incompleteness of our theories.

In 2013, the IceCube Neutrino Observatory reported the first detection of a high-energy astrophysical neutrino flux extending to $\sim$ 10 PeV energies [10]. Interpretations of this discovery gave footholds to address two long-standing problems in physics. First, the energy content in the Universe is the same in cosmic-rays and neutrinos, suggesting that finding the sources of astrophysical neutrinos will lead to the discovery of cosmic-ray accelerators; this would solve the century-old cosmic-ray mystery. Second, the astrophysical neutrino flux is predominantly extragalactic [11], allowing us to access neutrino baselines extending to gigaparsecs, and energies millions of times higher than what can be reached with earth-based neutrino factories. Fig. 1.1 shows expected and measured flavor composition of astrophysical neutrinos at Earth, after oscillations at cosmic scales. Different neutrino production scenarios are represented by different colors in the triangle. These are the three most likely neutrino production scenarios. However, one can start from any point on this triangle

and after oscillation, end up in the gray "bowtie" near the center. The uncertainties on the final flavor composition are driven by the uncertainties on measured neutrino oscillation parameters. In the next two decades, experiments such as Juno [12], Hyper-Kamiokande [13], DUNE [14], and the IceCube upgrade [15] will significantly improve our knowledge of oscillation parameters. The impact is reflected in the flavor triangle as a significant shrinking of the predicted standard oscillation region, shown in the lighter hue colors. Any measurement that deviates from the expected astrophysical flavor composition will have discovered new physics. It is therefore important to measure the astrophysical neutrino flavor composition accurately. This, today, is hindered by the tau neutrino flavor identification.

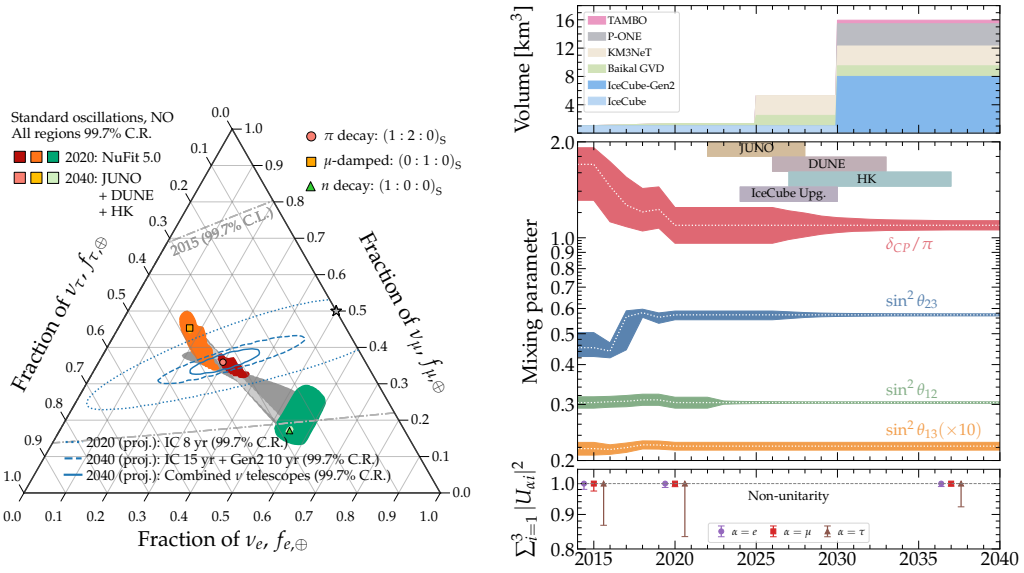



Figure 1.1: Left: Flavor composition of astrophysical neutrinos at Earth. The triangle includes current measurements and future projections. Different colors correspond to different production scenarios. Dark hues use the current uncertainties on oscillation parameter measurements, while light hues use the expected uncertainties in 2040. Current constraints on the flavor composition are shown as gray lines, while projections are shown in blue. Right: Projections for the next two decades that will affect flavor measurements. This includes the evolution of high-energy neutrino telescopes, neutrino mixing angles, and unitarity constraints. Top: Stacked effective volume of all high-energy neutrino telescopes. Middle: Uncertainties on the neutrino oscillation parameters as a function of time. Experiments relevant for neutrino oscillation parameter measurements are also shown with their proposed timelines. Bottom: Constraints on the unitarity of the neutrino mixing matrix for different neutrino flavors. Figures are from Ref. [16].

Beyond flavor measurements, neutrinos offer another unique probe into one of the most confounding cosmological mysteries; The nature of almost 85% of mass density in the Universe is unknown. Overwhelming astrophysical and cosmological evidence shows the bulk of the mass density of the

universe cannot be accounted for in the SM. Observations of the bullet cluster and dwarf galaxies hint at a possible corpuscular nature of this so called dark matter. If true, this particle must interact at or below the weak scale, and could require an extension to the SM. Although no signal has yet been detected to confirm a connection between the standard and dark sectors, another chink in the SM armor, neutrinos, may provide a path forward. Specifically, there is a distinct possibility that the dark sector is only tied to the SM via neutrinos, a scenario motivated by scotogenic neutrino-mass-generation models[17, 18]. In these models, neutrinos acquire some or all of their mass effectively through interactions with a dark field. [19, 20, 21, 22, 23, 24, 25, 26]. These models introduce heavy neutrino states, sometimes called dark neutrinos, which could also provide a possible explanation of anomalies observed at short-baseline neutrino experiments such as MiniBoone [27, 28, 29, 30].

Astrophysical neutrinos, therefore, offer exciting potential in the search for new physics which could manifest at production or propagation. Neutrinos could be produced by dark matter, implying interactions with dark fields that may also explain their mass dynamically. On their way to Earth, new physics could also manifest in non-standard oscillations, altering their flavor composition. But above all, understanding their standard sources is a crucial component to answering all such questions. These three avenues will be explored here.

But before we begin, a kind warning to the reader. This document is not self-contained, and serves rather as a listing of contributions I made as a graduate student under the supervision of Francis Halzen and Carlos Argüelles. Familiarity with particle physics at the level of Halzen and Martin [31] is assumed, as well as knowledge of standard statistical and experimental techniques as described in the Particle Data Group [32]. Further, I make no attempt to introduce or describe the IceCube experiment as I had no part in designing, building, or operating it; I leave that honor to those who contributed, and point you instead to the excellent detector paper here [33].

In this thesis we search for new physics using high-energy astrophysical neutrinos. As a necessary pre-requisite to new physics searches, however, we begin with an attempt to understand standard neutrino sources. In chapter 2, we describe a multi-wavelength search for astrophysical neutrino sources using nine years of IceCube data and find hints of neutrino emission from radio-bright Active Galactic Nuclei (AGN), further supporting recent claims that neutrino emission occurs at the core of AGNs.

We then turn our attention to astrophysical flavor measurements, and focus on improving predictions for one of the more pesky flavors, tau neutrinos, in chapters 3-5. Chapter 3 describes several

improvements in high energy tau neutrino propagation modeling. We present the modeling of tau energy losses in matter, and depolarization of taus as a result of multiple interactions. Further, we present a mechanism by which tau neutrinos are produced in Earth as a result of electron and muon neutrino propagation. This has a direct impact on future astrophysical neutrino flavor composition measurements. The work presented in this chapter resulted in two publications. The first [34] was co-authored with Alfonso Garcia Soto, Pavel Zhelnin, and Carlos Argüelles and published in *Physical Review Letters*. The second [35] was co-authored with Diksha Garg, Mary Hall Reno, Sameer Patel, and Carlos Argüelles and published in *Physical Review D*.

Chapter 4 presents `TauRunner`, a simulation package that incorporates these effects in a single publicly available software. The work presented in this chapter was published [36] in *Computer Physics Communications* and co-authored with Jeffrey Lazar, Alex Pizzuto, Oswaldo Vasquez, and Carlos Argüelles.

Chapter 5 discusses several advantages that result from accurate modeling of tau neutrinos. First, we present an improvement in IceCube's sensitivity to GZK neutrinos by a factor of two. Second, we show an improvement in IceCube's transient source sensitivity by more than an order of magnitude. Third, we perform an analysis using IceCube data and use tau neutrinos to rule out standard model explanations of anomalous events reported by the ANITA experiment. Work presented in this chapter resulted in three publications. The first [37] was co-authored with Francis Halzen, Ali Kheirandish, and Carlos Argüelles and published in *ApJ Letters*. The second [38] was co-authored with Alex Pizzuto, Carlos Argüelles, Francis Halzen, Ali Kheirandish and Justin Vandenbroucke and published in the *Journal of Cosmology and Astroparticle Physics*. The third [39] was co-authored with Alex Pizzuto and Anastasia Barbano and published together with the IceCube collaboration in *The Astrophysical Journal*.

Then, motivated by scotogenic neutrino mass generation models in which interactions with dark matter effectively provide some or all of the neutrino mass, we present a study of possible connections between dark matter and neutrinos in chapter 6. Specifically we derive constraints on dark matter annihilation and decay to neutrinos in our galaxy and from all dark matter halos in the observable universe using more than a dozen neutrino experiments. Our limits cross the thermal relic abundance in the neutrino sector for the first time. This work [40] was published in *Reviews of Modern Physics* and co-authored with Carlos Arguelles, Alejandro Diaz, Ali Kheirandish, Andres Olivares Del-Campo, and Aaron Vincent. Finally, in chapter 7, we conclude.

# Chapter 2

# Multiwavelength Search for Neutrino Emitters

> "See first, think later, then test. But always see first. Otherwise you will only see what you were expecting. Most scientists forget that."
>
> Douglas Adams

Before our quest for new physics may begin, it is imperative that we understand the standard sources of astrophysical neutrinos for two reasons. First, subtracting the known standard sources from the total measured astrophysical flux improves our sensitivity to neutrinos produced by exotic sources like dark matter. Second, knowing the production location of detected neutrinos opens up a multitude of new physics tests like, for example, neutrino decay, non-standard neutrino interactions, a fourth neutrino state, right-handed neutrinos, lorentz-invariance violation, CPT violation etc.

In this chapter we present an analysis that uses nine years of IceCube data to search for neutrino sources. We first summarize IceCube's observations in the last decade in Sec. 2.1 and motivate a move away from an exclusive reliance on gamma rays; Instead, we advocate for a multi-wavelength approach relying on radio, x-ray, infrared, and optical data to correlate neutrino and electromagnetic sources. In Sec. 2.2 we describe our chosen catalog and the cuts made based on radio variability. We then present the statistical method and event selection in Sec 2.3. The tested hypotheses and analysis sensitivity are then shown in Sec. 2.4. Finally, in Sec 2.5 we present the results of this search and what it means in the context of other recent results.

## 2.1 Gamma rays are not the whole story

IceCube has detected an astrophysical neutrino flux extending from energies as low as a few TeV to as high as $\sim$ 10 PeV [10]. These measurements showed that the energy content in the Universe is the same in neutrinos as it is in cosmic-rays, pointing to a common origin. It gave us a foothold into answering the century-old mystery: Where do cosmic-rays come from, and how are they accelerated to such immense energies?

The connection is rather simple: Active galaxies or neutron stars transform some of their gravitational energy via relativistic outflows into particle acceleration. Accelerated protons are then stopped by surrounding ambient radiation or dust. Astrophysical neutrinos in the Standard Model are then created by the decay of mesons or neutrons, which in turn could be created by proton-proton or proton-photon interactions. A necessary byproduct of all such interactions is also high-energy gamma-rays. Therefore neutrino astronomers' first instinct was to search for correlation of high-energy gamma-rays with astrophysical neutrinos. This led to a multitude of correlation studies between IceCube and GeV-TeV gamma ray sources in the last decade [41, 42, 43]. These studies, however, have indicated that Fermi-detected blazars cannot be the dominant sources of IceCube's astrophysical neutrinos [44, 45].

In 2017, IceCube launched a realtime alert program that informs the astronomy community of any interesting neutrino candidates that have more than 50% chance of being astrophysical in origin, enabling followups by any telescope that has an overlapping field-of-view [46]. On September 22nd, 2017, one such alert with estimated 290 TeV neutrino energy was found to be in spatial coincidence with TXS 0506+056, an active galaxy that is bright in multiple wavelengths of light [47]. This IceCube event was coincident with a gamma-ray flare detected by the Fermi telescope. A followup search in IceCube's archival data in the direction of TXS 0506+056 found a flare of 13 neutrinos in 100 days, a period spanning mid 2014 to 2015 [48]. This 2014-2015 neutrino flare of TXS-0506+056 was correlated with a rise in its radio flux at 15GHz, and a significant drop in the optical flux; however, no enhancement in the gamma-ray activity was detected [49]. Recent studies of this source in fact pointed to hints of gamma-ray suppression at the time of neutrino emission not just in TXS 0506+056, but also in two other gamma-ray sources spatially coincident with high-energy neutrino events [50]. These observations provided the first clues that sources may not simultaneously be efficient neutrino and gamma-ray emitters.

In hindsight, this may seem quite obvious. In order to efficiently produce neutrinos, a source must have in their vicinity sufficient target matter or radiation to stop a fraction of the cosmic-rays they accelerate. A simple comparison of the proton and photon interaction cross sections reveals that material or radiation dense enough to stop protons is more likely to stop photons with equivalent energy. Of course, the photons are not absorbed, but cascade down in energy until their energy is low enough to escape the source environment. It is therefore prudent in the selection of sources for our analysis to relax the requirement that neutrino and gamma-ray sources are directly correlated. It's important to note we do not claim that neutrinos and gamma-rays are unrelated. Rather we claim that this relationship is more complicated than a simple direct correlation, and we proceed in our source selection without making any assumptions about the gamma-ray connection.

## 2.2 Source Selection

We select a special class of AGN for this analysis using radio, infrared, and x-ray data to pick out sources more likely to be active and variable, pointing to an active core. These scenarios could include but are not limited to galaxy or supermassive black-hole mergers, major matter accretion activity near the core of AGN, and tidal disruption events [51, 52, 53]. All such events can cause outbursts in most wavelengths of photons, and manifest as heightened variability at various timescales. Such signatures are indicators of necessary accelerator and target material to produce neutrinos. Our main discriminant will be the variability of the 15GHz radio flux. Previously, two alerts issued by IceCube pointed to possible radio enhancement accompanying high-energy neutrino emission (TXS 0506+056 and PKS 1502+106). At the time of the neutrino alerts, these two sources were either close to or at their peak radio emission at 15GHz over a period of more than two decades of continuous observation [50].

The sources are selected from the Candidate Gamma Ray Blazar Survey (CGRaBS) catalog [54]. CGRaBS is a statistically complete sub-selection of the CRATES catalog (Combined Radio All-Sky Targeted Eight GHz Survey). It consists of 1625 radio sources with flat/hard radio spectra. This catalog was initially created in 2007 to guide Fermi observations, which was set to launch in 2008. Therefore the sources were selected based on their radio and x-ray spectra to resemble properties of gamma-ray blazars. Interestingly, only 39% of these sources were seen by Fermi in 10 years.

To pick out more likely neutrino emitters from this catalog, we first characterize the radio variability of the sources using two measures. First, the intrinsic modulation index (V)

$$V = \frac{s}{\bar{I}} = \frac{1}{\bar{I}}\sqrt{\frac{N}{N-1}\left(\overline{I^2} - \bar{I}^2\right)} \tag{2.1}$$

where $s$ is the standard deviation of flux measurements, and $\bar{I}$ is the mean measured flux density. This is a simple estimate of the deviation of a set of measurements from the mean measured flux density. This does not, however, account for statistical uncertainties in the individual measurements and could therefore be inaccurate, especially for fainter sources. The second measure we use is $\eta$, a reduced $\chi^2$ test-statistic that quantifies the deviation of individual flux measurements from the weighted mean measured flux density

$$\eta = \frac{1}{N-1}\sum_{i=1}^{N}\frac{(I_i - \xi_I)^2}{\sigma_{I,i}^2} \tag{2.2}$$

where N is the number of observations, $\sigma_{I,i}$ is the uncertainty on a single measurement, $I_i$ is a single flux measurement and

$$\xi_I = \frac{\sum_{i=1}^{N} I_i/\sigma_{I,i}^2}{\sum_{i=1}^{N} 1/\sigma_{I,i}^2} \tag{2.3}$$

is the weighted mean of flux measurements.

We use radio data obtained from a fast-cadence 15 GHz radio monitoring program with the 40m telescope at the Owens Valley Radio Observatory (OVRO) [55]. The telescope is located in the Owens Valley, California at an elevation of 1236m. Construction of the telescope was completed in 1966 and has been used for many missions since then. The blazar monitoring program ran between 2007 and 2020 and collected data at a cadence of 3 days from more than 1500 sources. Although OVRO has so far only published lightcurves from 2007 till 2016, this provides thousands of measurements per source and is sufficient to characterize their variability using the measures explained above. A 2-dimensional scatter plot of the modulation index as a function of the mean flux density is shown in Fig. 2.1. The color scale is the $\chi^2$ test-statistic defined above. The black line in the figure corresponds to a standard deviation of 65mJy, which is the mean standard deviation of all sources. We are interested in the most variable sources, and place a cut to keep only those sources whose variability exceeds the mean standard deviation of the population. This corresponds to the region above the black line, and results in a selection for sources whose test-statistic ($\eta$) is highest. The final selection, after variability cuts, includes 390 sources.

Finally, we cross-correlate these sources with surveys at other wavelengths, starting with infrared photons. Infrared photons coming from the dust torus of an AGN are a good candidate for target

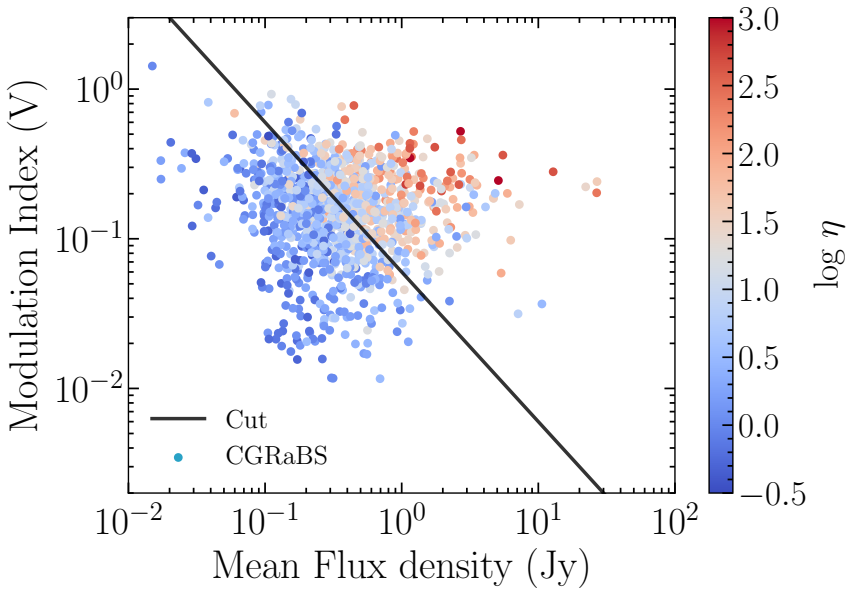



Figure 2.1: Variability measures for all sources in the CGRaBS catalog above $-20°$ declination. Shown is the modulation index (V) as a function of mean radio flux density at 15GHz as measured by the OVRO 40m telescope. The color scale shows the log of the variability test-statistic used to characterize these sources. The black line is where we place the cut to keep more variable sources.

radiation. Therefore we are interested in loud radio AGN that also have associated infrared emission. However, there are other physical scenarios, such as star-forming regions, where there is significant radio and infrared emission, which doesn't fit the hypothesis we're trying to test. Since this catalog is composed of blazar candidates, however, we expect the IR emission to be pointing mostly to target radiation associated with the central engine, as described by the blazar sequence [56]. To make our selection, we use a value called $q_{22}$, which is simply the ratio of IR to radio intensities. Because active central engines are expected to cause an IR flux, this value $q_{22}$ is then supposed to minimize the contamination from sources whose radio emission is not associated with AGN activity. In other words, IR emission associated with the activity causing radio emission will show a correlation, and this correlation can be captured in the ratio of the two intensities. We use radio-loud WISE (Widefield Infrared Survey Explorer) blazars detected in all four WISE filters, whose mid-IR colors are similar to the typical colors of confirmed gamma-ray emitting blazars [57]. The cross-correlation with CGRaBS yields 238 sources.

We then turn to X-ray data. The X-ray flux can act as a good proxy to column depth, which in turn is a possible indicator of gamma-ray opaque sources. We cross-correlate our selection with

XMM-Newton [58], an X-ray telescope sensitive to the $0.2-12$ keV flux. The cross-correlation yields 87 sources.

A skymap of all 1158 sources in the CGRaBS catalog above a declination of -20 degrees is shown in Fig. 2.2. This declination band represents the field of view of the OVRO 40m telescope. We also show skymaps of all cross-correlated selections discussed above, which includes sources after the variability cut, the X-ray cross-correlation, and Infrared cross-correlation.

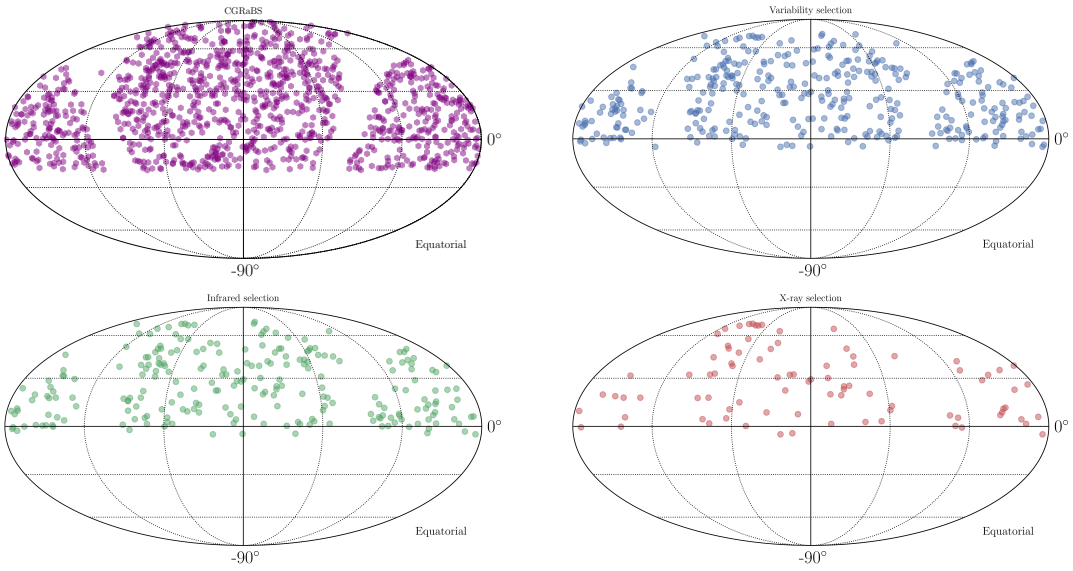



Figure 2.2: Skymap of CGRaBS sources. Top-left shows all of the sources in the field of view of OVRO. Top-right shows the 390 sources remaining after applying variability cuts, and a declination cut of $-5°$ due to the IceCube event selection used. Bottom-left shows the remaining 283 sources after cross-correlation with the radio-loud WISE catalog, and bottom-right shows the remaining 87 sources after cross-correlation with XMM-Newton.

## 2.3 Analysis Method

Our analysis searches for neutrino spatial and energy correlation from a collection of sources. We test the hypothesis that all of the sources in a catalog are emitting neutrinos. This method is called a stacking analysis [59]. Stacking sources together has been shown to improve the sensitivity of source searches.

We use an unbinned maximum-likelihood to search for neutrino emission above the atmospheric and diffuse astrophysical components. The likelihood is defined as

$$\mathcal{L}\left(n_s, \gamma\right) = \prod_{i=1}^{N} \left[\frac{n_s}{N}\mathcal{S}_{\text{stack}}\left(\mathbf{x}_i, \left\{\mathbf{x}_{S_k}\right\}, \sigma_i, E_i \mid \gamma\right) + \left(1 - \frac{n_s}{N}\right)\mathcal{B}\left(\delta_i, E_i\right)\right] \tag{2.4}$$

where $n_s$ is the number of signal events and $\gamma$ the spectral index of the sources assuming a simple power-law. $N$ is the total number of events in the sample, $\mathbf{x}_i$ is the sky coordinate of event $i$, $\mathbf{x}_{S_k}$ is the sky coordinate of source $S_k$, $\sigma$ is the angular uncertainty, $E$ is the energy proxy. The background pdf, $\mathcal{B}$ is

$$\mathcal{B}\left(\delta_i, E_i\right) = \underbrace{\frac{\mathcal{P}_{\mathcal{B}}\left(\delta_i\right)}{2\pi}}_{\text{spatial}} \times \underbrace{\mathcal{E}_{\mathcal{B}}\left(E_i, \delta_i\right)}_{\text{energy}}, \tag{2.5}$$

where $\mathcal{P}_{\mathcal{B}}$ is the background probability of an event and depends only its declination $\delta$ since IceCube is located at the geographic South Pole. The stacking signal PDF, $\mathcal{S}_{\text{stack}}$ can be written as

$$\mathcal{S}_{\text{stack}}\left(\mathbf{x}_i, \left\{\mathbf{x}_{S_k}\right\}, \sigma_i, E_i \mid \gamma\right) = \frac{\sum_{k=1}^{M}\left[W_k R_k\left(\delta_{S_k} \mid \gamma\right) \times \mathcal{S}\left(\mathbf{x}_i, \mathbf{x}_{S_k}, \sigma_i, E_i \mid \gamma\right)\right]}{\sum_{k=1}^{M} W_k R_k\left(\delta_{S_k} \mid \gamma\right)} \tag{2.6}$$

where the sum runs over all $M$ sources, and each individual event signal PDF is weighted by a combination of the detector acceptance at the source location $R_k$ and the theoretical source weight $W_k$, which will be discussed later on. Finally, the signal PDF $\mathcal{S}$ is built from simulation and using kernel density estimators to extract the energy dependence of the angular estimators, including intrinsic uncertainties such as the kinematic angle between the neutrino and the muon.

The dataset used for this analysis consists of nine years (2011-2019) of IceCube muon events in the Northern Sky for a total livetime of 3186 days. We use deep neural networks to reconstruct the muon energies which significantly improves the energy resolution over previous treatments. A detailed description of the dataset and reconstruction techniques was published in ()

The best estimates for the parameters assuming a signal hypothesis are those which maximize the likelihood function, $\mathcal{L}$. This maximized likelihood is then compared against the likelihood of the background only hypothesis such that

$$\mathcal{TS} = 2\ln\left(\frac{\mathcal{L}\left(\hat{n}_s, \hat{\gamma}\right)}{\mathcal{L}\left(n_s = 0\right)}\right). \tag{2.7}$$

where variables with ($\hat{}$) are best-fit parameters. This is the test-statistic used for the analysis.

## 2.4 Hypotheses and Sensitivity

Starting with our source selection after radio variability cuts, we test three hypotheses in this analysis. These hypotheses use the same methods described above but vary in the final source selection and source weights $W_k$. The first hypothesis tests for a correlation directly between neutrino emission and radio variability using, as weights, the variability test-statistic $\eta$ described in 2.2. The second hypothesis uses the correlation between our selected sources and the radio-loud WISE catalog. We weight this selection by the ratio of radio intensity to infrared intensity at 22 mm. Finally, the third hypothesis tests for a correlation with soft X-rays weighting by the X-ray flux measured with XMM Newton. Skymaps for these selections can be seen in Fig. 2.2, and the distributions of weights is shown in Fig. 2.3. We build background distributions for these locations and weights

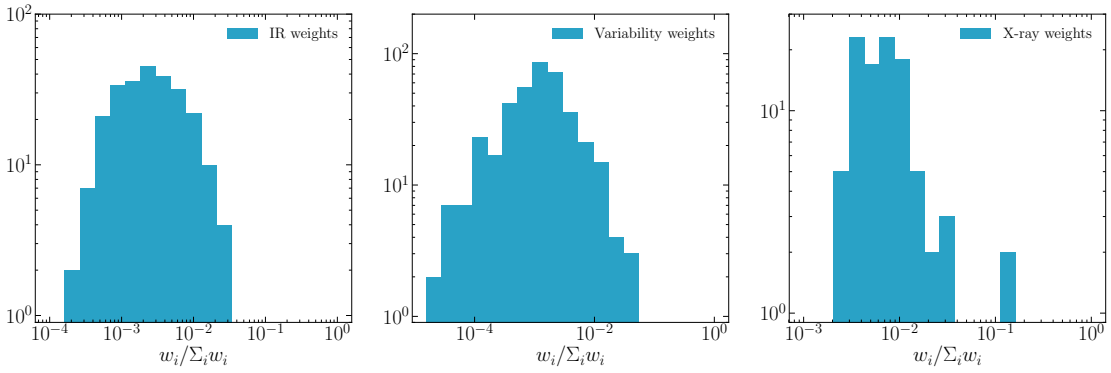



Figure 2.3: Sum-normalized weight distributions shown for all three hypotheses tested. Notably, the X-ray flux weights show two sources that each carry at least 10% of the weights, whereas the other two hypotheses are more evenly distributed.

by repeating many trials with randomized data realizations; that is, we assume $n_s = 0$ and obtain the corresponding $\mathcal{TS}$. The background distributions obtained for all three hypotheses are shown in Fig. 2.4. The sensitivity for this analysis is then estimated via injection tests where a signal is inserted in a fake data realization, and the $\mathcal{TS}$ is maximized. We repeat this test many times for various injected spectral indices and fluxes. The sensitivity is set at 90% C.L when 90% of the trials return a $\mathcal{TS}$ value greater than the mean of background-only trials. The $5\sigma$ discovery potential is set at the point where 50% of trials exceed the $5\sigma$ threshold background test-statistic. Sensitivities and discovery potentials as a function of spectral index are shown in Fig. 2.5 for all three hypotheses.

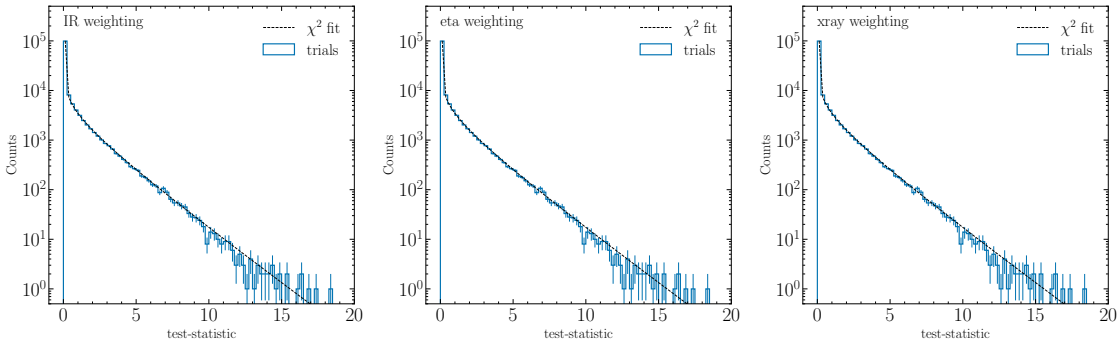



Figure 2.4: Distribution of test-statistic values assuming a background-only hypothesis, shown for all three weighting schemes. Overlayed on top of the histogram is a functional fit assuming a $\chi^2$ distribution with a delta function at zero.

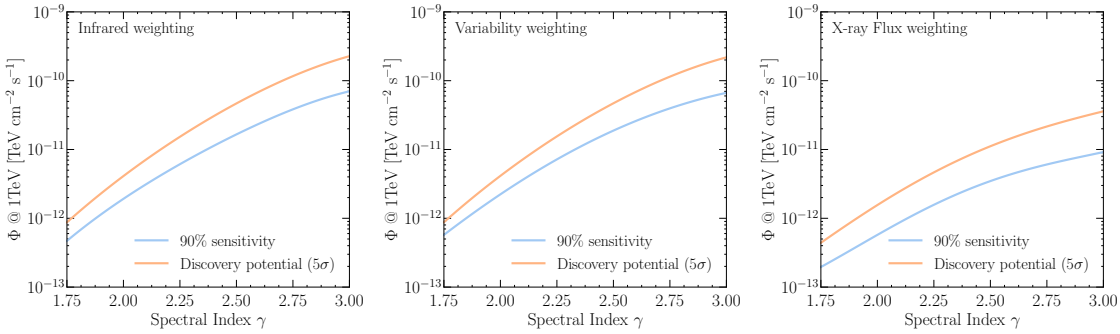



Figure 2.5: Sensitivity and discovery potential as a function of spectral index, for all three weighting schemes. Sensitivity is defined as the flux where 90% of signal-injected trials exceed the mean background $\mathcal{TS}$. Discovery potential is defined as the flux needed for 50% of signal injection trials to exceed the $\mathcal{TS}$ value corresponding to the quoted significance.

## 2.5 Results and Interpretation

Our most significant result is the infrared sub-selection. The unblinded test-statistic overlayed on the background distribution is shown in Fig. 2.6. We find hints for neutrino emission in the case of our infrared selection at the level of $2.15\sigma$ one-sided significance above background expectation. We fit 16.33 events with a spectral index of 2.08. The unblinded log-likelihood scan is shown in Fig 2.7 where we plot the fitted flux as a function of spectral index. The best-fit flux and spectral index as well as the 1 and $2\sigma$ contours are shown. This observation further supports a recent result from the IceCube experiment testing the correlation of neutrinos with AGN cores [60] which showed an excess of $2.6\sigma$ after stacking $32,249$ sources. We obtain a similar significance but with a factor of $\sim 100$

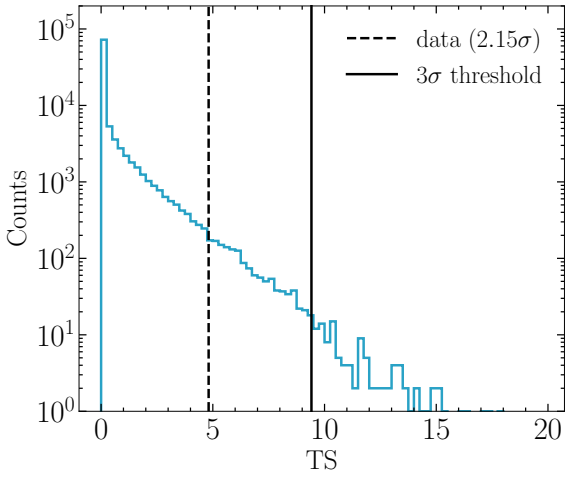



Figure 2.6: Unblinded $\mathcal{TS}$ value overlaid on the background distribution for the infrared selection. The p-value corresponds to a one-sided significance of $2.15\sigma$. Also shown is the $3\sigma$ threshold $\mathcal{TS}$ value.

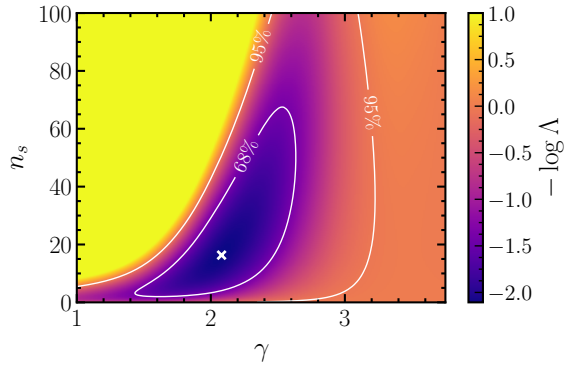



Figure 2.7: Unblinded log-likelihood scan showing the best-fit point (white marker) as well as the $1\sigma$ and $2\sigma$ contours. The best-fit point corresponds to 16.33 events with a spectral index of 2.08. $\Lambda$ is the ratio of the signal and background likelihood values.

less sources, and more importantly with a negligible overlap between the two source selections. This gives us increased confidence in the underlying physics scenario being tested. Namely, cosmic-rays are likely accelerated near the cores of AGN, and are stopped by the surrounding ambient radiation, producing neutrinos and obscuring the high-energy photons. Independent groups using IceCube's public alerts with estimated energies $> 200$ TeV have also seen a correlation of IceCube neutrinos with radio AGN, and concluded that neutrinos are likely produced by proton interactions with ambient radiation near the core of AGN[61]. Further, an analysis of public icecube muon events in the TeV-PeV energy range found evidence for neutrino emission from compact radio sources [62].

These observations, along with the analysis presented here, all point to a coherent picture. Astrophysical neutrinos point to cosmic-ray accelerators with significant radio emission coming from an active core. Relativistic outflows from accreting matter onto that core accelerate cosmic-rays. These cosmic-rays subsequently interact with ambient radiation near the source where the radiation density is high, making efficient neutrino sources that become gamma-ray opaque. The sources are not necessarily gamma-ray quiet, but the gamma-rays likely escape at lower energies, and the observed high-energy gamma rays are produced further downstream from the site of neutrino production. Undoubtedly, further years of observation will make this conclusion clear and robust.

# Chapter 3

# Tau Neutrinos Require Some Care

> "And then there's Ibrahim.. I have no idea what he's working on anymore. I think taus. There's always things to do with the taus"
>
> Francis Halzen

As briefly highlighted in the introduction, measuring the flavor composition of astrophysical neutrinos is a direct test of all physics scenarios that would alter neutrino oscillations. Such a test is especially strong with astrophysical neutrinos given the immense baselines these neutrinos traverse, and the energies they possess. Flavor measurements in this generation are limited by the tau fraction, mainly due to low statistics. That is why a multitude of experiments are now proposed which focus primarily on detecting tau neutrinos. These experiments will require accurate modeling of tau neutrino propagation in matter, which is not as straight-forward as other flavors. In fact we show here that accurate modeling can have significant impact even on the current generation of experiments, namely IceCube. In this Chapter, we present improvements made to modeling of tau neutrino propagation at high-energies. As shown in Fig. 3.1, tau neutrinos above 10 PeV energies can appear as tracks in detectors. They may also decay to muons before reaching the detector, thus contaminating muon event selections, if not properly accounted for. We introduce neutrino interactions in matter is Sec. 3.1. We then describe in Sec. 3.2 charged tau energy losses in matter, which become relevant above $\mathcal{O}(10)$ PeV energies. In Sec. 3.3 we derive a treatment for the depolarization of taus in matter. Finally, in Sec. 3.4 we quantify the effect of tau neutrino production in Earth due to the propagation of muon and electron neutrinos. We find that the production of tau neutrinos at ultra-high energies is significant and will be the largest irreducible background to

future tau neutrino experiments. These effects are then all discussed in the context of a Monte Carlo software package that is presented in Chapter 4.

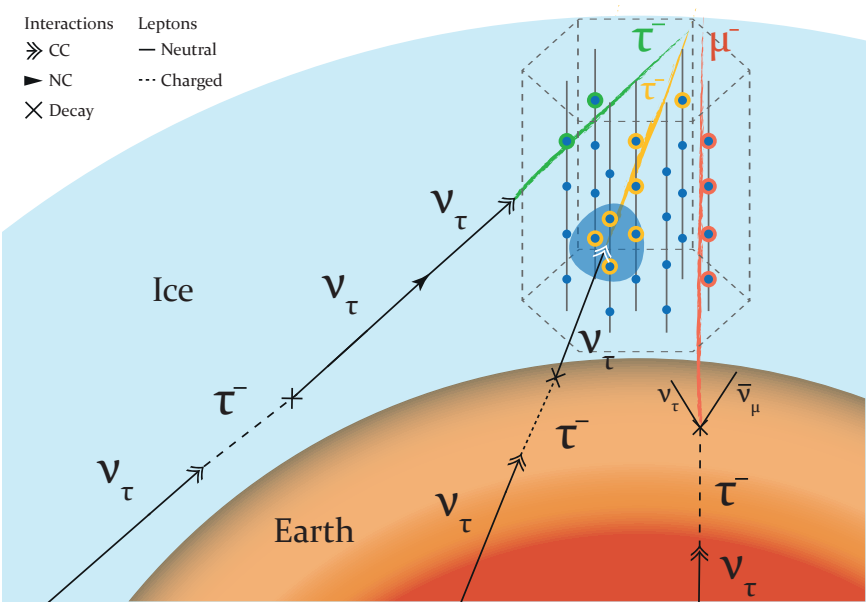



Figure 3.1: Schematic of lepton propagation through Earth followed by a measurement with the IceCube detector. There are three possible signatures from EeV tau-neutrino secondaries, described here from left to right. Left: A throughgoing tau track, which is possible for taus at or above 10 PeV. Center: The interaction vertex is contained in the fiducial volume of the detector in this case, producing a cascade from the charged-current interaction, along with an outgoing tau track. Right: The tau decays before reaching the detector, producing a muon in ∼18% of the cases, which subsequently enters the detector. For clarity, not all particles involved in the interaction are shown. An additional contribution included in the results but not shown here is an NC interaction inside the detection volume.

## 3.1 Tau Neutrinos through Matter

The propagation of a flux of neutrinos through a medium can be described by the following cascade equation [63]

$$\frac{d\varphi(E,x)}{dx} = -\sigma(E)\varphi(E,x) + \int_E^\infty d\tilde{E}\ f(\tilde{E},E)\varphi(\tilde{E},x), \tag{3.1}$$

where $E$ is the neutrino energy, $x$ is the target column density, $\sigma(E)$ the total neutrino cross section per target nucleon, $f(\tilde{E},E)$ is a function that encodes the migration from larger to smaller neutrino energies, and $\varphi(E,x)$ is the neutrino spectrum. The first term on the right hand side accounts for the loss of flux at energy $E$ due to charged-current (CC) and neutral-current (NC) interactions, whereas the second term is the added contribution from neutrinos at higher energy, $\tilde{E}$, to $E$ through NC

interactions of $\nu_{e,\mu,\tau}$ and CC interactions in the $\nu_\tau$ channel. In Earth, where the column density can be on the order of $10^{33}$ nucleons/cm$^2$, neutrino attenuation is important for 100 TeV energies and higher [64]. In the Sun, where the column density is, on average, $10^{37}$ nucleons/cm$^2$, attenuation is relevant for energies greater than 100 GeV [65, 66]. For dense astrophysical environments capable of producing high-energy neutrinos—one of the denser ones being radio galaxies—the column density is around $10^{24}$ nucleons/cm$^2$ [67]. Given that the column density required for a neutrino to undergo a single interaction is, on average, $10^{29}$ nucleons/cm$^2$, astrophysical sources are thus transparent even to neutrinos of EeV energies and higher. In this work, the secondaries produced in CC interactions of other flavors are neglected due to the fact that the electrons and muons lose energy rapidly. On the other hand, taus produced in CC tau neutrino interactions have a much higher probability of decaying yielding high-energy neutrinos. This is due to the fact that weak decays scale as $m^5$ and that the tau mass is significantly larger than that of the muon, allowing for more decay modes, which results in a ratio of lifetimes between muons and taus of approximately $10^7$. While the lifetimes are drastically different, the energy losses above $\sim$1 PeV, where stochastic losses are dominant, are only a factor of 10 smaller for taus than for muons. These two facts set the critical energy in ice—the energy at which the decay and interaction lengths are equal—to be approximately $\sim 10^9$ GeV for taus, while for muons it is $\sim$10 GeV [68]. This implies that tau energy losses can be safely neglected below 10 PeV and the decay-on-the-spot approximation is a good one, see e.g., [69]. However, in this work we consider neutrino propagation at EeV energies and higher, where this approximation no longer holds and careful treatment of tau energy losses is required; see [70, 71] for recent implementations and discussions.

Measurements of neutrino cross sections have been performed at energies below GeV up to a few PeV [72]. This includes a multitude of results using human-made neutrinos in accelerator [73, 74] and reactor [75, 76] experiments as well as natural sources such as solar [77], atmospheric [78], and astrophysical neutrinos [79, 80]; for recent reviews, see [81, 82]. In the future, measurements of high-energy neutrinos from collider experiments will be available in the TeV range [83, 84].

Unfortunately, these measurements stop short of the region of interest for this work where the fractional momenta, $x_{\rm Bjorken}$, of the quarks probed by the neutrino can reach $x_{\rm Bjorken} \ll 10^{-8}$. The nucleon structure function is not measured at such low $x_{\rm Bjorken}$ and is extrapolated in cross section calculations [85, 86]. Such extrapolations neglect gluon color screening making perturbative QCD calculations of the neutrino cross section grow faster than allowed by unitarity at extremely high

energies [87]. Phenomenological approaches to include gluon screening parameterize the extremely small $x_{\text{Bjorken}}$ behavior using a dipole model [88] of the nucleon so as to result in a $\ln^2(s)$ dependence of the cross section at extremely high energies [89]. Calculations using this approach were shown to be in good agreement with the total proton-proton cross section measurements from Auger [90] and TOTEM at LHC [91] data. This ultimately results in a difference of a factor $\sim 2$ at $10^{12}$ GeV; see Fig. 3.2 for a comparison of two PDF models: a perturbative QCD calculation [85], and a dipole model [88].

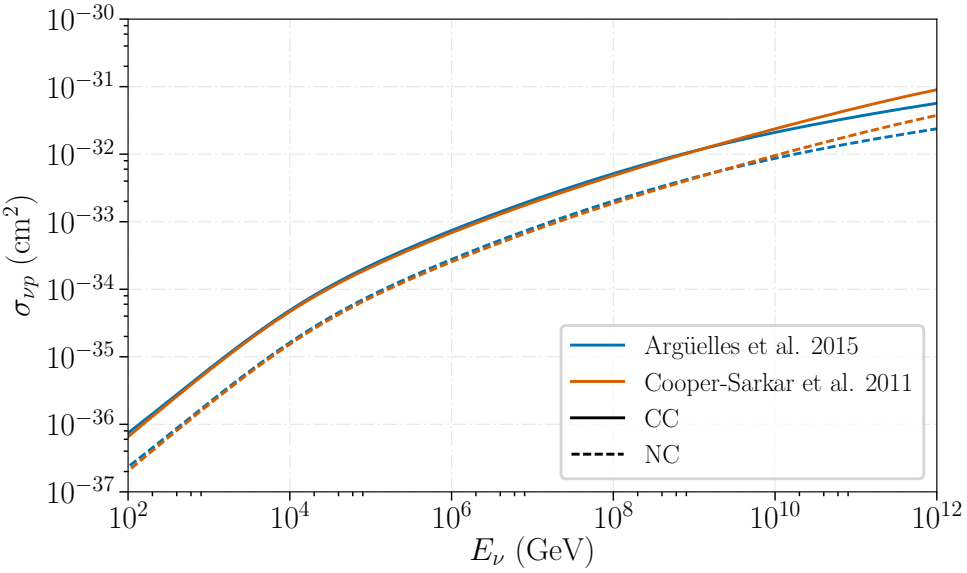



Figure 3.2: The neutrino-proton cross section as a function of the neutrino energy. Solid (dashed) lines correspond to charged-current (neutral-current) cross sections. Blue lines [88] correspond to the model used for this work. Orange lines [85] are also implemented in the software and can be chosen by the user.

## 3.2 Tau Energy Losses and Decays

When neutrinos undergo CC interactions, they convert to their charged partners through the exchange of a $W$ boson. Charged particles lose energy in dense media through many processes, and the relative importance of each process depends on the lepton's mass and its energy [92]. At lower energies, a charged lepton can ionize atoms as it traverses the medium. This process is described by the Bethe-Bloche equation, and at higher energies scales logarithmically and becomes sub-dominant for all flavors. A charged lepton can also interact with the electric field of a nucleus, losing energy in the process through the emission of a photon. This process, called bremsstraahlung, scales

like the inverse-sqaured mass of the lepton, and is therefore the dominant energy loss mechanism for electrons. Another possible interaction with the field of a nucleus leads to the production of electron-positron pairs. This process scales like the inverse of the lepton mass, and is one of the leading energy-loss mechanisms for $\mu$ and $\tau$. Finally, the leptons can also lose energy by exchanging a photon with a nucleon, in what is referred to as a photonuclear interaction. This process dominates tau energy losses at the highest energies ($\geq 10^9$ GeV). This cross section depends on the nucleon structure function, and thus it has the same source of uncertainty as the neutrino-nucleon cross section. Fig. 3.3 shows distributions of final tau energies and total distance traveled before decay for several initial tau energies, obtained by incorporating the discussed models by into a lepton propagation tool, PROPOSAL [93]. We observe that the tau range and minimum energy is bounded. For the range, it grows logarithmically above 10 kilometers, mainly due to the growing interaction cross section. For the energy at decay, it's bounded around 100 PeV below which the decay length is much shorter than the interaction length.

Apart from interacting, $\mu$ and taus can also undergo weak decays. This process scales like the mass of the lepton to the fifth power, and is therefore the most likely outcome for taus propagating in Earth up to $10^9$ GeV. Above this energy, the total interaction length for other processes becomes shorter than the decay length. $\mu$, on the other hand, are much more likely to lose all of their energy before decaying at rest, or getting absorbed by a nucleus. The decay distribution of taus is well described by parametrizations in [94], which we use in this work.

## 3.3 Tau depolarization in matter

Tau depolarization in electromagnetic interactions is dominated by scatterings in which the incoming tau loses a substantial fraction of its initial energy. As discussed in Sec. 3.2 bremsstrahlung is suppressed relative to $e^+e^-$ pair production and photonuclear interactions for taus [95]. Pair production that yields a change in tau energy of more than 10% are rare. For example, for $10^9$ GeV tau neutrinos propagating through Earth's mantle, only 0.02% of the pair production tau scatterings have significant energy loss. On the other hand, photonuclear interactions have more frequent hard scatterings, in which the final tau energies are less than 90% of their initial energies. Thus, here we focus on tau photonuclear energy loss. We begin by describing the formalism of tau spin polarization, and then move to connect that to the photonuclear interaction cross section.

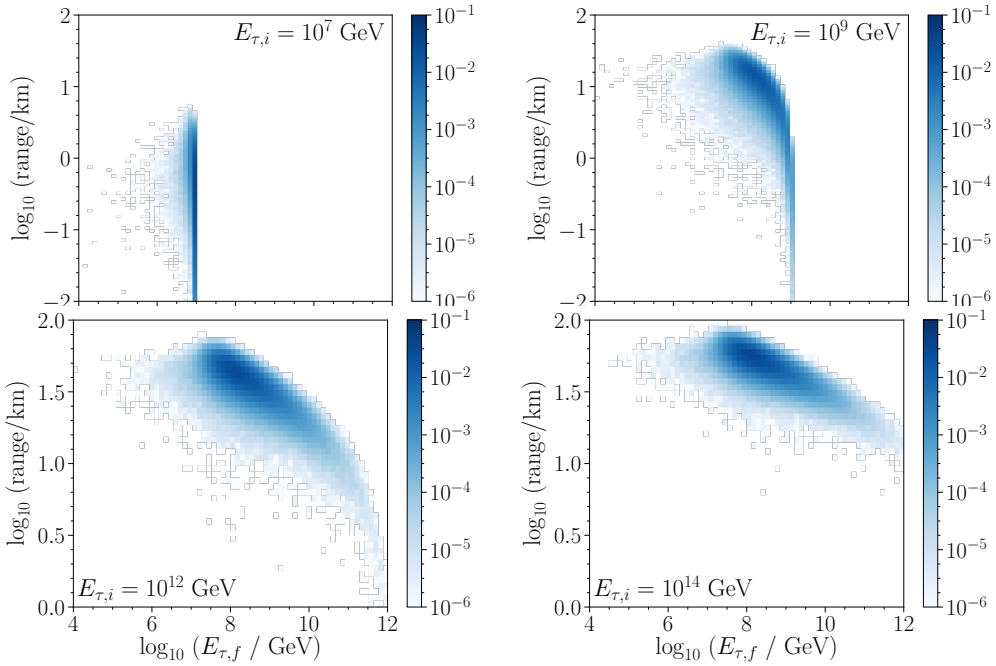



Figure 3.3: Distribution of final tau energies and total distance traveled before decay for several initial tau energies. At 10 PeV (upper left) and below, taus lose little energy before decay, while at 1 EeV (upper right) taus reach the critical energy and losses become appreciable. In this regime, the median range increases linearly as the tau becomes more boosted. At 1 ZeV (bottom left) and above (bottom right), these distributions show asymptotic behavior, with taus decaying around 100 PeV and traveling, on average, tens of kilometers.

The spin polarization three-vector of a tau produced in a CC neutrino interaction, in its own rest frame, can be written as [96]

$$\vec{s} = (s_x, s_y, s_z) = \frac{P}{2}(\sin\theta_P \cos\phi_P, \sin\theta_P \sin\phi_P, \cos\theta_P) = \frac{P}{2}(\Lambda \sin\theta_P, 0, \cos\theta_P) \tag{3.2}$$

where the spin direction is relative to the final state tau momentum direction in the lab frame, taken to be the $z$-axis. In Eq. (3.2), $\theta_P$ and $\phi_P$ are the polar and azimuthal angle of the spin vector in the $\tau$ rest frame, and $P$ is the degree of polarization. The polarization vector lies in the scattering plane [96], thus $\phi_P = 0$ or $\pi$. Therefore in Eq. (3.2), $\Lambda$ takes value $+1$ or $-1$ according to the azimuthal angle. In what follows, we define $\mathcal{P}_z \equiv 2s_z$. For a single scattering, we denote the polarization as $\mathcal{P}_{\tau,z}$. We are interested in the final, or net, polarization of the tau just before it decays, which we denote as $\mathcal{P}_z$. This quantity $\mathcal{P}_z$ enters into the energy distribution of the $\nu_\tau$ from tau decay, which can be described as a function of $z_\nu \equiv E_\nu/E_\tau$ in the form

$$\frac{1}{\Gamma}\frac{d\Gamma(\tau \to \nu_\tau)}{dz_\nu} = \sum_i B_i \Big( g_0^i(z_\nu) + \mathcal{P}_z g_1^i(z_\nu) \Big) , \tag{3.3}$$

where $g_0^i(z_\nu)$ and $g_1^i(z_\nu)$ depend on $i$ decay channels with branching fraction $B_i$. The full decay width can be modeled as the sum of $\nu_\tau$ plus $e\bar{\nu}_e$, $\mu\bar{\nu}_\mu$, $\pi$, $\rho$, $a_1$, and $4\pi$ final states [97, 98]. We note that for $\bar{\tau} \equiv \tau^+$ decays, the differential decay distribution is

$$\frac{1}{\Gamma}\frac{d\Gamma(\bar{\tau} \to \bar{\nu}_\tau)}{dz_{\bar{\nu}}} = \sum_i B_i \Big( g_0^i(z_{\bar{\nu}}) - \mathcal{P}_z g_1^i(z_{\bar{\nu}}) \Big) , \tag{3.4}$$

for $z_{\bar{\nu}} \equiv E_{\bar{\nu}}/E_{\bar{\tau}}$. Therefore, CC production of a right-handed (RH) $\bar{\tau}$ will yield the same decay distribution in $z_{\bar{\nu}}$ as the $z_\nu$ decay distribution of LH $\tau$. ***Scattering kinematics***– We first define the kinematic variables for the EM interactions. For an incoming $\tau$ momentum for EM interaction ($k$), target nucleon momentum ($p$) and outgoing tau momentum ($k'$) in the laboratory frame, we write

$$k^\mu = (E_i, 0, 0, p_i), \quad p^\mu = (M, 0, 0, 0), \quad k'^\mu = (E_\tau, p_\tau \sin\theta, 0, p_\tau \cos\theta). \tag{3.5}$$

Here $E_i$ and $E_\tau$ are the incoming tau and outgoing tau energies in the laboratory frame, respectively, and $k^2 = m_\tau^2$. The Lorentz invariant variables, in terms of energy and angles in the lab frame where the target is at rest, are given by

$$\begin{aligned} \nu &= p \cdot q/M = E_i y = (E_i - E_\tau), \\ Q^2 &= -q^2 = -(k - k')^2 = 2(E_i E_\tau - p_i p_\tau \cos\theta) - k^2 - k'^2, \\ x &= Q^2/(2p \cdot q)\,. \end{aligned} \tag{3.6}$$

***Tau photonuclear scattering***– Fig. 3.4 shows the average values of the $z$ component of the initial spin polarization vector, $-\langle {\mathcal{P}_{\nu,z}}^{CC} \rangle$, as a function of $E_\tau/E_\nu = (1-y)$. It can be observed that a tau produced by neutrino CC interaction is almost fully polarized for high incident neutrino energies. The shaded band shows where 10% or less of the $\tau$ emerge with this energy fraction. Thus, it is a good approximation to assume $\nu_\tau \to \tau$ produces left-handed taus in agreement with results found in ref. [99]. Therefore in what follows, we assume that taus produced via CC neutrino interactions begin their journey as fully polarized, and proceed to derive the spin polarization vector for $\tau \to \tau$ electromagnetic scattering in terms of the structure functions. For tau electromagnetic interactions of initially left-polarized tau, the EM leptonic current is

$$j_\lambda^\mu(\tau\,, EM) = \bar{u}_\tau(k', \lambda)\gamma^\mu u_\tau(k, -) \tag{3.7}$$

The hadronic tensor for EM case will only have structure functions $W_1$, $W_2$ and reduces to

$$W_{\mu\nu} = -g_{\mu\nu} W_1 + \frac{p_\mu p_\nu}{M^2} W_2, , \tag{3.8}$$

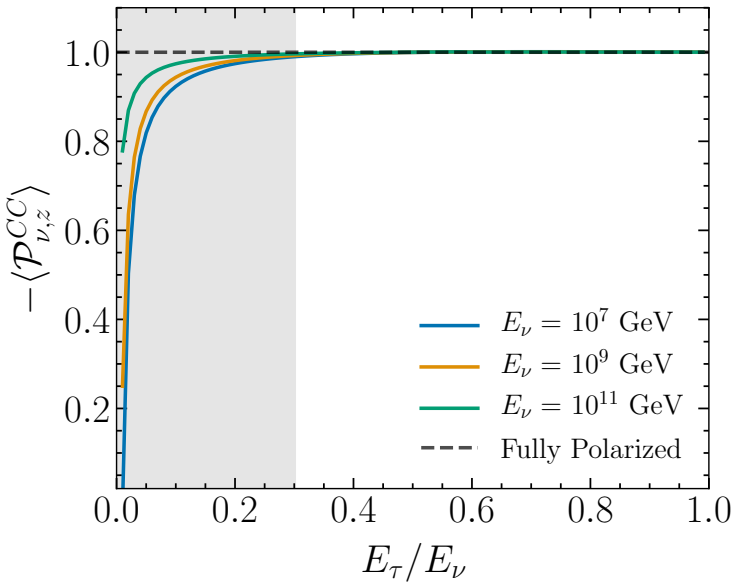



Figure 3.4: Initial polarization of taus produced from charged-current scattering. The average value of $\langle \mathcal{P}^{CC}_{\nu,z}(y) \rangle$ as a function of tau energy fraction for $\nu_\tau$ CC scattering for different incident neutrino energies. The shaded region shows where 10% or less of the taus emerge with this energy fraction. In the massless lepton limit, $-\langle \mathcal{P}^{CC}_{\nu,z}(y) \rangle = 1$ for all $y$, corresponding to left-handed out-going taus.

by taking into account gauge invariance. The differential cross section for tau photonuclear scattering is

$$\frac{d^2\sigma_{\rm EM}}{dy\,dQ^2} = \frac{2\pi\alpha^2}{Q^4}\frac{1}{E_i}F_\tau\,, \tag{3.9}$$

where

$$F_\tau = 2W_1(E_iE_\tau - p_ip_\tau\cos\theta - 2m_\tau^2) + W_2(E_iE_\tau + p_ip_\tau\cos\theta + m_\tau^2)$$

is obtained by contracting the hadronic and leptonic tensors. The structure functions $W_1$ and $W_2$ are more commonly written in terms of $F_1$ and $F_2$ as

$$W_1 = F_1/M, \quad W_2 = F_2/\nu \tag{3.10}$$

with

$$F_1 = \frac{1}{2x(1+R)}\left(1 + \frac{4M^2x^2}{Q^2}\right)F_2\,. \tag{3.11}$$

The differential cross section translates to

$$\frac{d^2\sigma_{\rm EM}}{dy\,dQ^2} = \frac{4\pi\alpha^2}{Q^4}\frac{F_2(x,Q^2)}{y}\left[1 - y - \frac{Q^2}{4E_i^2} + \left(1 - \frac{2m_\tau^2}{Q^2}\right)\frac{y^2(1+4M^2x^2/Q^2)}{2[1+R(x,Q^2)]}\right]$$

The quantity $R(x,Q^2)$, implicitly defined in eq. (3.11), can be written in terms of the longitudinal structure function $F_L \simeq F_2(x,Q^2) - 2xF_1(s,Q^2)$ (in the small $x$ limit) and $F_1(x,Q^2)$. This gives

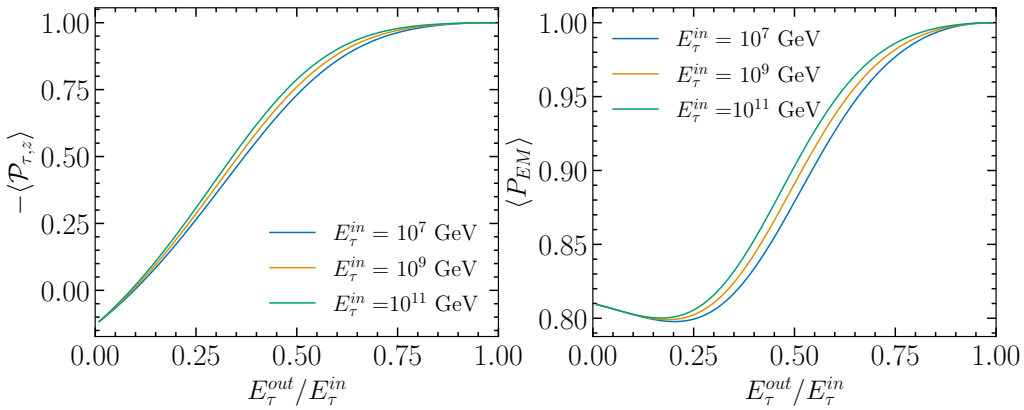



Figure 3.5: Average polarization of outgoing tau after a single scattering. The average polarization about $z$-axis $\langle \mathcal{P}_{\tau,z} \rangle$ (left) and the average total polarization $\langle P_{EM} \rangle$ (right) as a function of outgoing tau energy fraction for single photonuclear electromagnetic scattering of $\tau$'s with rock for different incident tau energies. Here, $\langle P^{EM} \rangle = (\langle \mathcal{P}^{EM}_{\tau,z} \rangle^2 + \langle \mathcal{P}^{EM}_{\tau,x} \rangle^2)^{\frac{1}{2}}$

$R(x, Q^2) \equiv F_L(x, Q^2)/(2xF_1(x, Q^2))$. At high energies, we can take $R(x, Q^2) = 0$. The tau spin polarization vector can be related to the tau EM scattering cross section as follows: The spin density matrix gives the relation [96]

$$dR_{\lambda\lambda'} \sim L^{\mu\nu}_{\lambda\lambda'} W_{\mu\nu} \,. \tag{3.12}$$

We can then relate the elements of the spin polarization vector to the lepton-hadron contractions. Up to an overall normalization factor $N$,

$$\begin{aligned} dR_{++} + dR_{--} &= d\sigma_{\rm sum} = N(L^{\mu\nu}_{++} + L^{\mu\nu}_{--})W_{\mu\nu}, && (3.13)\\ dR_{+-} &= s_x d\sigma_{\rm sum} = \frac{L^{\mu\nu}_{+-} W_{\mu\nu}}{(L^{\mu\nu}_{++} + L^{\mu\nu}_{--})W_{\mu\nu}} d\sigma_{\rm sum}, && (3.14)\\ \frac{dR_{++} - dR_{--}}{2} &= s_z d\sigma_{\rm sum} = \frac{(L^{\mu\nu}_{++} - L^{\mu\nu}_{--})W_{\mu\nu}/2}{(L^{\mu\nu}_{++} + L^{\mu\nu}_{--})W_{\mu\nu}} d\sigma_{\rm sum}. && (3.15) \end{aligned}$$

Finally, using Eqs. (3.2), (3.13), (3.14), and (3.15) we can evaluate the spin polarization vector components for tau EM scattering:

$$\begin{aligned} s_x &= -\frac{m_\tau}{2} \sin\theta \Big[ 2E_i\, W_1 - W_2(E_i + E_\tau) \Big] / F_\tau \\ s_y &= 0 && (3.16)\\ s_z &= -\frac{1}{2} \Big[ 2W_1(p_i p_\tau - E_i E_\tau \cos\theta) + W_2(p_i p_\tau + E_i E_\tau \cos\theta) + W_2 m_\tau^2 \cos\theta \Big] / F_\tau. && (3.17) \end{aligned}$$

The energy distribution of $\nu_\tau$ from $\tau$ decays depend on $\mathcal{P}_{\tau,z}$ (Eq. 3.3), so only the $s_z$ component of the spin polarization vector is relevant. The average value of $\mathcal{P}^{EM}_{\tau,z}$ as a function of $y$, integrated

over $Q^2$ is:

$$\langle \mathcal{P}^{EM}_{\tau,z}(y)\rangle \equiv \frac{d\langle P\cos\theta_P\rangle_{\rm EM}}{dy} = \int dQ^2\, 2s_z \frac{d^2\sigma_{\rm EM}}{dy\, dQ^2}\left(\frac{d\sigma_{\rm EM}}{dy}\right)^{-1}. \tag{3.18}$$

In Fig. 3.5 (left), the average value of $-\langle \mathcal{P}^{EM}_{\tau,z}(y)\rangle$ after a single scattering, as a function of $E^{out}_\tau/E^{in}_\tau$, is shown for three incident tau energies. A high-energy tau passing through rock can get partially depolarized. The depolarization effect becomes significant for $y \gtrsim 0.2$, namely, for $E^{out}_\tau/E^{in}_\tau \lesssim 0.8$. The value of $\langle \mathcal{P}^{EM}_{\tau,z}\rangle$ depends on the quantities $P$ and $\theta_P$. We calculate the average of the polarization, $\langle P^{EM}\rangle = (\langle \mathcal{P}^{EM}_{\tau,z}\rangle^2 + \langle \mathcal{P}^{EM}_{\tau,x}\rangle^2)^{\frac{1}{2}}$, where $\langle \mathcal{P}^{EM}_{\tau,x}\rangle$ is evaluated according to eq. (3.18) with $s_z \to s_x$. In Fig. 3.5 (right), we show the average total polarization for different incident tau energies. We see that $\langle P^{EM}\rangle$ lies between 1 and 0.8 throughout the range of $E^{out}_\tau/E^{in}_\tau$. The main contribution to depolarization observed in Fig. 3.5 (left) is from the polar angle $\theta_P$.

## 3.4 Production of Tau Neutrinos in Earth

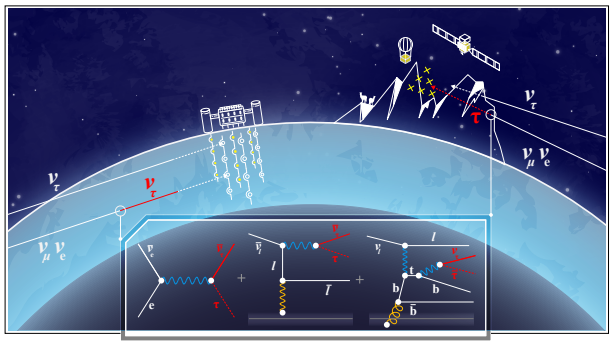



Figure 3.6: Artistic rendition of relevant interactions for tau appearance. The Feynman diagrams summarize the muon- and electron-neutrino interactions that would produce a pair $\tau + \nu_\tau$: Glashow resonance (left), $W$-boson production (middle) and top-quark production (right). Ice- and water-Cherenkov detectors are represented on the center left. Earth-skimming and space-based observatories are shown on the center right.

We show in this section that high-energy electron or muon neutrinos can also produce tau neutrinos and prove that this effective tau appearance phenomenon has a significant and direct impact on the inferred astrophysical neutrino spectra and flavor measurements. For the first time, we identify channels that would yield a significant contribution of secondary tau neutrinos as shown in Fig. 3.6. The secondary contribution arises from channels that produce on-shell $W$ bosons, which decay 10% of the time to $\nu_\tau + \tau$ [32]. The dominant interactions are neutrino-nucleus [100, 101, 102,

103, 104] or neutrino-electron [105, 106, 107] $W$-boson production, which become relevant above PeV energies. In the neutrino-nucleus interaction, when the energy transferred from a neutrino to a nucleon is above the top-quark mass, a $W$ boson can also be produced from the decay of this heavy quark to a $b + W$ pair [102]. The latter process is intrinsically related to the parton distribution functions (PDF) of the sea bottom in the nucleons. Depending on the PDFs and mass scheme models, top-quark production accounts for $5 - 15\%$ of the total deep-inelastic-scattering (DIS) cross sections above 100 TeV [108, 109, 110, 111, 112, 113, 88, 114]. The structure functions for neutrino-nucleon deep inelastic scattering have been implemented in GENIE. In this study, the HEDIS-CSMS model (`GHE19_00b`) is used. This calculation uses the NLO HERA1.5 PDF set [115] and uses expressions for the massless coefficient functions at NLO. This is the preferred model in IceCube's analyses. Additionally, the $W$-boson production and Glashow resonance are computed using the differential cross sections as described in [86]. Fig. 3.7 shows the total cross section for channels in which a tau neutrino is produced from 10 TeV to 100 EeV. For comparison, the deep inelastic scattering contribution is also shown. The flux at the detector of produced tau neutrinos

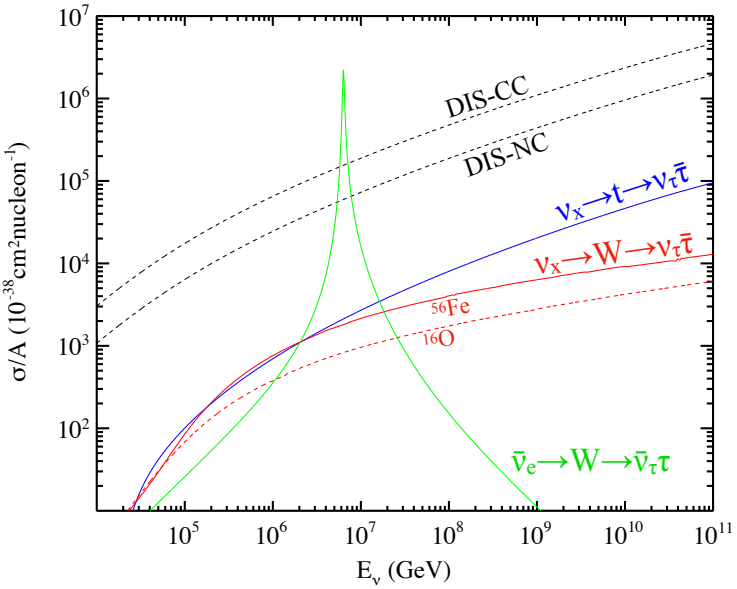



Figure 3.7: Neutrino-nucleus cross sections per nucleon as a function of the neutrino energy assuming an isoscalar target. Continuous lines represent the interaction channels in which a $\nu_\tau + \tau$ pair is produced: Glashow resonance (green), $W$-boson production (red), and top-quark production (blue). The deep inelastic scattering contribution is shown with a dashed line (summing CC and NC). The $W$-boson production is shown for two different nuclei (O and Fe) in order to illustrate the $Z^2/A$ scaling effect.

in different angular regions is shown in Fig. 3.8. In addition, the relative fractions of secondary tau neutrinos to the astrophysical component at 1 PeV and 100 PeV are shown.

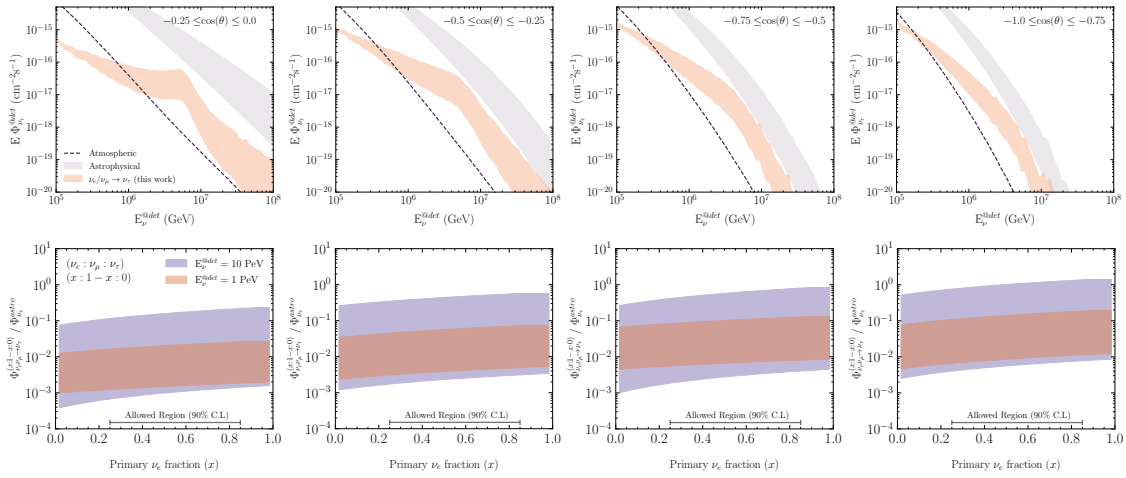



Figure 3.8: Top: Components of tau-neutrino flux at the detector integrated over different angular regions. Grey shaded band shows the expected flux between the best-fit spectra from the HESE and throughgoing muons analyses [116, 117]. Dashed line shows the atmospheric flux using `MCeq` (H3a-SIBYLL23C) [118]. The red shaded regions show the secondary tau neutrinos contribution produced by the propagation of muon and electron neutrinos assuming the best fits from IceCube's analysis [116, 117] and $(2:1:0)$ flavor composition. Bottom: Ratio of the secondary flux assuming different muon- and electron-neutrino fractions to the HESE's best-fit flux assuming $(1:1:1)$ at 1 PeV (red) and 10 PeV (blue). The bands represent the uncertainty in the primary spectrum. The black, horizontal line indicate the combined 90% C.L. from different flavor composition measurements [119, 120].

***Diffuse cosmic neutrino fluxes*** — The astrophysical neutrino flux has been characterized by the IceCube Collaboration using multiple complementary channels [121, 122, 123, 124, 125, 119, 126, 116, 127]. As a benchmark scenario, motivated by the Fermi acceleration mechanism [128], these analyses model the astrophysical neutrino flux as an unbroken power law in energy. The results of these analyses agree within errors, and the best-fit spectral index ranges from $2.37 \to 2.89$. The softest spectral index is measured using high-energy starting events (HESE) with deposited energy above 60 TeV [116], while the hardest spectrum is measured when selecting for throughgoing muons in the northern sky [117, 127]. Both analyses set constraints on the fraction of muon neutrinos, whereas similar signatures of electrons and taus in the detector give rise to a degeneracy between the other two flavors. The smoking-gun signature of tau neutrinos is two separated energy depositions, known as a *double bang*. These have recently been detected [129]; still, the sample size does not yield significant evidence of a tau-neutrino component [130].

Here we estimate the expected flux of tau neutrinos at the detector under two different hypotheses. First, we assume the best-fit results from HESE and throughgoing muons under the canonical $(1:1:1)$ flavor composition. Second, we use the best-fit results assuming a zero tau-neutrino component in the primary flux. Fig. 3.9 shows the expected tau-neutrino flux at the detector for events below the horizon ($-0.5 < \cos\theta < -0.25$). Tau neutrinos produced from muon and electron neutrinos alone can contribute as much as 50% of the primary astrophysical flux for energies above 10 PeV. This fraction mainly depends on the primary spectrum but not on the relative composition of muon and electron neutrinos as shown in Fig. 3.9. At these energies the yield of secondary tau neutrinos arising from top-quark and W-boson productions is similar, while Glashow resonance dominates at $\sim$ 6 PeV. Additionally, the secondary tau neutrino flux has an angular dependence since it varies with the column depths that neutrinos travel. Hence, the relative fraction will be higher for steeper angles. Consequently, even in a scenario where the astrophysical neutrino flux is only made of muon and electron neutrinos, there is a guaranteed flux of tau neutrinos in the detector, which is larger than the atmospheric component (from the prompt decay of heavy mesons) above $\mathcal{O}(100\text{TeV})$.

***Impact of tau appearance on discovery of cosmic tau neutrinos*** — We now examine how this additional component manifests in the HESE analysis. To do so, we combine the expected fluxes at the detector with the publicly available effective areas, assuming the best-fit HESE flux and a democratic primary flavor composition. Fig. 3.10 shows the expected energy and angular distribution of up going tau neutrinos in HESE after ten years of data taking. Above 300 TeV, we expect 1.6 upgoing tau neutrinos, but only 0.02 events of atmospheric origin. The expected rate from secondary tau neutrinos is twice that of the prompt atmospheric component. Although the expected number of events is small, this secondary component is the dominant irreducible background for any astrophysical tau-neutrino search. Tab. 3.1 summarizes the capabilities for rejecting the non-tau cosmic component, assuming that all upgoing tau-neutrino events in the HESE sample above an energy threshold are identified. We conclude that this additional tau-neutrino contribution reduces the discovery potential for all energy threshold criteria considered. A thorough breakdown using realistic tau-neutrino identification criteria as well as a likelihood approach that includes the energy and angular information must be performed by neutrino observatories to fully quantify the impact of this intrinsic background.

Currently, flavor composition measurements from IceCube are statistically limited due to the size of the detector. However, the IceCube-Gen2 optical array will be able to detect a sizable amount

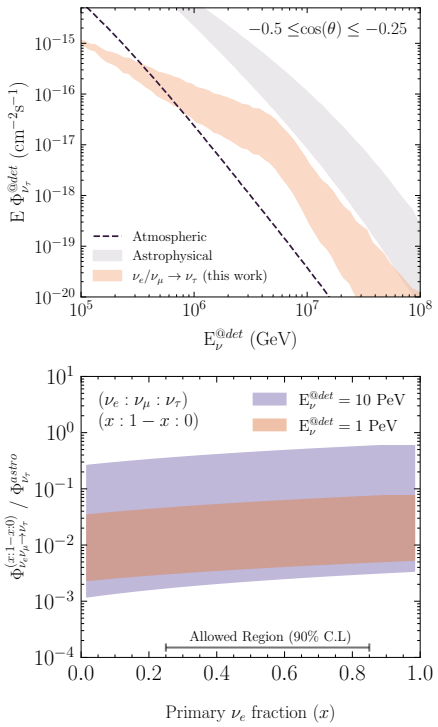



Figure 3.9: Primary neutrino fluxes and appearing tau-neutrino component. Top: Different components of tau-neutrino flux at the detector integrated over $-0.5 < \cos\theta < -0.25$. Grey shaded band shows the expected flux between the best-fit spectra from the HESE and throughgoing muons analyses [116, 117]. Dashed line shows the atmospheric flux using `MCeq` (H3a-SIBYLL23C) [118]. The red shaded region represents tau neutrinos appearing from the propagation of muon and electron neutrinos, assuming $(2 : 1 : 0)$ flavor composition. Bottom: Ratio of the secondary flux assuming different muon- and electron-neutrino fractions to the astrophysical flux assuming $(1 : 1 : 1)$ at 1 PeV (salmon) and 10 PeV (lavender). The bands represent the uncertainty in the primary spectrum. The black, horizontal line indicates the 90% C.L. allowed flavor composition [119, 120].

| $E_{th}$ | $P_{\tau>15m}$ | HESE | Atmos. | $\nu_\mu/\nu_e \to \nu_\tau$ |
|---|---|---|---|---|
| 100 TeV | 1% | 6.63 | 0.13 (6.3) | 0.05-0.11 (6.0-5.7) |
| 200 TeV | 9% | 3.00 | 0.05 (4.3) | 0.03-0.09 (4.1-3.7) |
| 300 TeV | 17% | 1.57 | 0.02 (3.2) | 0.02-0.07 (2.9-2.5) |
| 400 TeV | 23% | 1.12 | 0.01 (2.7) | 0.01-0.06 (2.4-2.1) |

Table 3.1: Impact of appearing tau-neutrino component on cosmic tau-neutrino discovery. Each row shows a different assumed threshold, as indicated in the first column, to identify tau neutrinos in the HESE analysis. The second column shows the percentage of tau-neutrino interactions with $E_\nu = E_{th}$ producing a tau lepton that travels more than 15m [130]. The third column shows the number of upgoing tau-neutrino events in HESE in ten years. The significance ($\sigma$) to reject the non-tau cosmic component for different background hypothesis is shown in parenthesis in the fourth and fifth columns when considering atmospheric and appearing-tau backgrounds.

of PeV neutrinos [131]. The rate of secondary neutrinos due to $W$-boson production, Glashow resonance, and top-quark production is $5 - 20\%$ above 10 PeV, as shown in Fig. 3.10. In fact, the

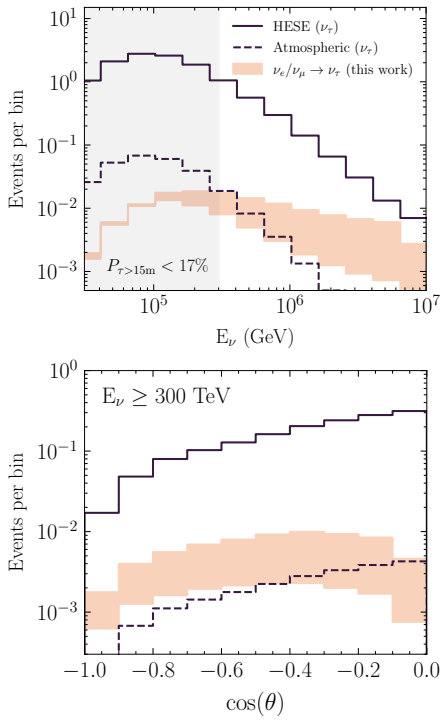



Figure 3.10: Rates of tau-neutrino events in ten years of HESE. Energy (top) and angular (bottom) distribution of tau-neutrino events that pass the HESE selection. The angular distribution is shown for events with energies above 300 TeV. Lines and shaded regions correspond to the fluxes described in Fig. 3.9. The grey area represents the region where tau-neutrino identification with current reconstruction methods is challenging.

fraction of any flavor is affected by these new channels since $W$ bosons can also decay to muon and electron neutrinos. Therefore, any analysis studying the composition of the astrophysical flux using upgoing events must account for these secondary neutrinos. Thus, our conclusions are generic and extend beyond the HESE analysis.

***Ultra-high-energy neutrinos*** — Currently, IceCube measurements extend to 10 PeV in neutrino energy [116, 127, 132]. At higher energies, many experiments have placed upper limits on the neutrino flux [133, 134, 135, 136]; however, their effective volumes and Earth's opacity for UHE neutrinos limit their capabilities. One of the most promising ideas to detect EeV neutrinos is to look for Earth-skimming tau leptons [137, 138, 139, 140, 141, 142]. Almost horizontal tau neutrinos at energies above 10 PeV will interact in Earth's crust, producing an energetic tau-lepton, which can emerge in the atmosphere and decay. Subsequently, electromagnetic radiation from upgoing EASs can be detected at optical and radio wavelengths. Using this technique, experiments like POEMMA [143], PUEO [144], GRAND [138], Trinity [142] and Beacon [139] claim to have significantly

better sensitivities than current experiments. In fact, the projected performance of the aforementioned experiments shows that they would be able to discover the cosmogenic neutrino flux for most UHECRs scenarios.

Fig. 3.11 shows the predicted energy spectrum of tau leptons emerging almost horizontally from Earth's surface. In this calculation, we have assumed a cosmogenic model that includes more than a single UHECR source population [145]. The expected secondary flux coming from muon- and electron-neutrino interactions — as illustrated in Fig. 3.6 — is also shown together with the $E^{-2}$ flux for which GRAND, Trinity, and Beacon will observe more than 2.44 events (90% C.L. upper limit) after ten years of data taking. At energies above 10 PeV, the dominant mechanism for the secondary flux is via top-quark production. In the case of ultra-high-energy neutrinos, unlike the lower energy astrophysical component, the angular distribution of the primary and secondary components are very similar. Two main conclusions can be derived from the predictions shown in Fig. 3.11. First, we expect a non-negligible contribution of emerging taus from a primary flux of muon and electron neutrinos. This holds true even if the flavor composition of the cosmogenic flux has no primary tau component. Thus for optimistic models, GRAND, Trinity, and Beacon will detect taus regardless of primary flavor composition. Second, when they observe emerging taus, they cannot infer the type of neutrino that produced them. Therefore, these experiments will not set strong constraints on the normalization of the cosmogenic flux without making assumptions about its flavor composition; *i.e.*, the yield of taus is degenerated between small $(1:1:1)$ and large $(1:1:0)$ primary fluxes.

There are some nascent reconstruction techniques to detect the optical Cherenkov emission of EASs induced by muons with detectors like POEMMA or Trinity [146]. These novel methods would make Earth-skimming experiments sensitive to multiple flavors, allowing them to break the degeneracy previously described. Similarly, experiments that can detect geo-synchrotron and Askaryan radiation [147] — *e.g.*, IceCube-Gen2 [131], RNO-G [148] or PUEO [144] — would not be affected by this degeneracy because they are equally sensitive to all neutrino flavors. Therefore, a combined analysis between the mentioned experiments will be fundamental to constrain the normalization and flavor composition of the cosmogenic neutrino flux.

We have shown that the flux of secondary tau-neutrinos is larger than the prompt atmospheric component above a few hundred TeV. It is the dominant irreducible background for astrophysical tau-neutrino searches in current and future neutrino telescopes. Hence, future flavor composition measurements of the cosmic-neutrino flux must account for it. The relative impact of this back-

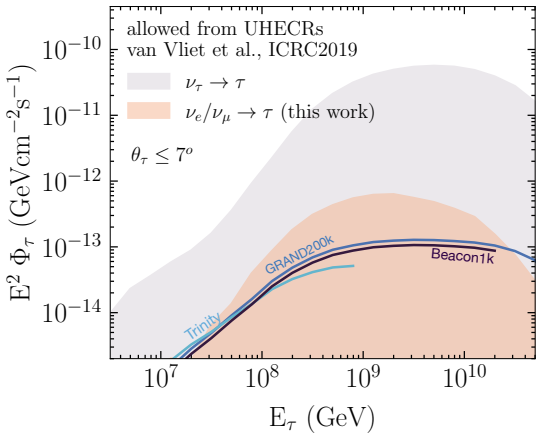



Figure 3.11: Tau-lepton fluxes arriving at Earth-skimming experiments. Expected flux of emerging tau for Earth-skimming neutrinos ($\theta_\tau < 7°$). Shaded curves represent the allowed regions derived from a model fit to Auger data [145] with (gray) and without (salmon) tau neutrinos in the primary flux. Solid lines indicate the flux for which Earth-skimming experiments would collect more than 2.44 events in ten years assuming a $E^{-2}$ primary neutrino spectrum.

ground highly depends on the primary neutrino spectrum, so it is fundamental to constrain its shape and normalization further. We have proved that Earth-skimming experiments are sensitive to cosmogenic fluxes with multiple source populations of UHECRs, independent of the primary flavor composition. Even in extreme scenarios without primary tau-neutrinos, there will be a distinguishable flux of emerging tau leptons coming from muon- and electron-neutrino interactions. Nevertheless, experiments sensitive to a single flavor cannot set constraints on the normalization of the flux since there is a degeneracy between primary and secondary contributions. We therefore conclude that a symbiotic ecosystem of neutrino telescopes with different flavor identification capabilities is needed to understand the origin of UHE neutrinos.

# Chapter 4

# `TauRunner`: A Comprehensive Solution to Lepton Propagation

Now that we have described all the nuances of tau neutrino propagation, we aim to incorporate all these effects into a single, comprehensive, customizable and user-friendly package. Enter `TauRunner`, a public `Python`-based package designed to have minimal dependencies, to allow the user to construct arbitrary neutrino trajectories and propagation media, and to provide interfaces to modify physics inputs such as neutrino cross sections easily. This chapter describes the software and provides benchmarks and comparisons to other packages that have similar aims. We expect this software to be useful for next-generation neutrino detectors operating in liquid water (P-ONE [149]), solid water (IceCube-Gen2 [131]), mountains (Ashra NTA [137], TAMBO [140]), and outer space (POEMMA [141]).

## 4.1 Algorithm Overview

`TauRunner` solves the neutrino transport equation Eq. (3.1) using a Monte-Carlo approach. A flowchart of the `TauRunner` Monte-Carlo algorithm is shown in Fig. 4.1. Given an initial neutrino type, energy, and incident angle, it begins by calculating the mean interaction column depth, $\lambda_{\text{int}}$, which depends on the medium properties and neutrino cross section. A column depth is then randomly sampled from an exponential distribution with parameter $\lambda_{\text{int}}$, and the neutrino advances the corresponding free-streaming distance. If the neutrino does not escape the medium, either an NC or CC interaction is chosen via the accept/reject method. In the case of an NC interaction,

the neutrino energy loss is sampled from the differential cross section, and the process repeats. In the case of a CC interaction, a charged lepton is created with energy sampled from the neutrino differential cross section.

The treatment of the charged lepton then varies according to the initial neutrino flavor. Electrons are assumed to be absorbed and the propagation stops there. $\mu$ and $\tau$, however, are recorded and passed to `PROPOSAL` [68] to be propagated through the same medium. $\mu$ that do not escape will either decay at rest resulting in neutrinos that are below the energies supported by `TauRunner`, or get absorbed. Therefore a $\mu$ that does not escape is not tracked further. Finally, $\tau$s can either escape or decay. In the latter case, a secondary $\nu_\tau$ is created whose energy is sampled from tau decay distributions provided in [94]. Additionally, if the $\tau$ decays leptonically, $\nu_e$ or $\nu_\mu$ will be created. When this happens, the properties of the resulting secondaries are recorded and added to a basket which stores all secondary particles to be propagated together after the primary particle propagation is complete.

## 4.2 Structure of the code

`TauRunner` may be run either from the command line by running `main.py` or may be imported to run within another script or `Jupyter` notebook. To run from the command line, the user must minimally specify the initial energy, the incident nadir angle, and the number of events to be simulated. These can be specified with the `-e`, `-t`, and `-n` command line flags respectively. This will run the `TauRunner` algorithm in Earth with a chord geometry. The `TauRunner` output will be printed in the terminal unless an output file is specified with the `-save` flag. If this option is specified, `TauRunner` will save both a `numpy` array and a `json` file with the configuration parameters at the specified location. In order to ensure reproducibility, the user may specify a seed for the random number generator with the `-s` flag. By default, `main.py` propagates an initial $\nu_\tau$ flux, but a user may specify other initial particle types by using the `-flavor` flag. Additional options that may be specified by the user can be found in the `initialize_args` function of `main.py` or by running `main.py` with the `-h` flag.

To run within another script or `Jupyter` notebook the user must import the `run_MC` function from `main.py`. In this latter case one must also create a `TauRunner` `Particle`, `Track`, `Body`, `CrossSection` objects and a `PROPOSAL` propagator. The `Particle` class contains the particle properties as well as methods for particle propagation. The `Track` class parametrizes the geometry of the particle trajectories. The `Body` class defines the medium in which the propagation is to occur.

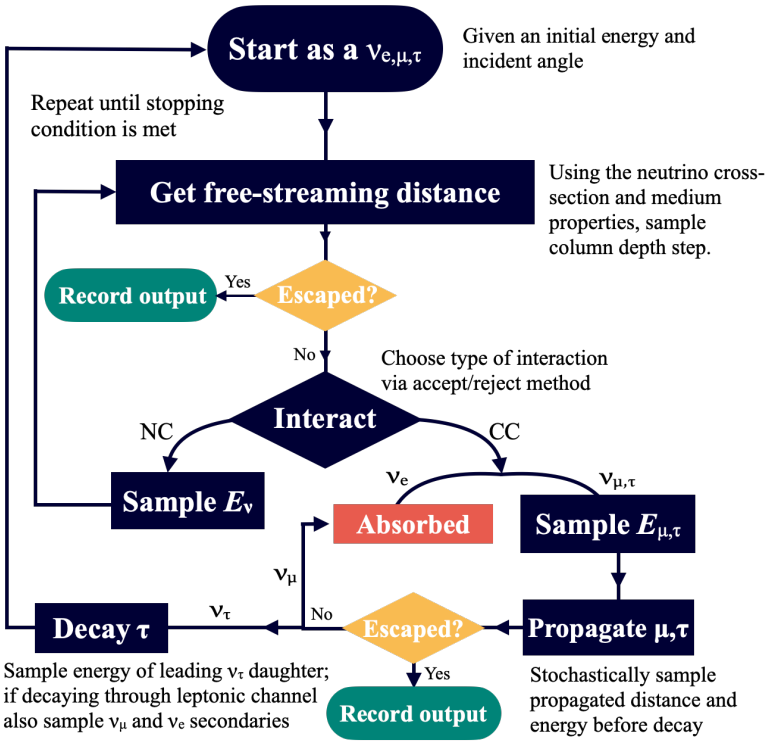



Figure 4.1: Flowchart of the `TauRunner` propagation algorithm. Square boxes indicate actions performed by the software. Diamond boxes indicate decision-making stopping points. Rounded-corner squared boxes indicate beginning and end of the algorithm. Note that users can select also charged leptons as the initial state, in which case the algorithm skips straight to the charged particle propagation step.

The `CrossSection` class defines neutrino cross section model. Additionally, `TauRunner` provides a convenience function for constructing `PROPOSAL` propagators, `make_propagator`, which can be imported from the `utils` module. `Casino.py` combines these classes according to the logic outlined in Fig. 4.1. After discussing the package broadly, we will discuss conventions and describe `TauRunner`'s output.

In what follows we briefly describe each module:

***`Particle`***— A `Particle` instance contains the structure of a `TauRunner` event. This includes, among other quantities, the particle's initial and current energies, particle type, and position. Additionally, it has a number of methods for particle decay and interaction as well as charged lepton propagation. Finally, the $\tau$ decay parametrization is contained in `particle/utils.py`.

The user may propagate $\nu_e$, $\nu_\mu$, $\nu_\tau$, $\mu^-$, $\tau^-$, or any of the corresponding anti-particles in `TauRunner`. To do this, the user should initialize the the `Particle` object with the correspond-

ing Particle Data Group Monte Carlo number [92]. It should be noted that the user may create an $e^{\pm}$, but the internal logic of `TauRunner` assumes all $e^{\pm}$ are immediately absorbed and thus no propagation occurs; see Fig. 4.1.

***`Track`***— The `Track` class contains the geometrical information about the particle's trajectory. A track is parametrized by an affine parameter which defines the position along the trajectory: 0 is the beginning of the trajectory, and 1 is the end. Almost all of the methods of the `Track` class are mappings between the affine parameter and physically relevant quantities, *e.g.* radius, distance traveled, and column depth. The only argument which is generic to the `Track` class is `depth` which specifies the distance below the surface of the body at which to stop propagation. This may intuitively be thought of as the depth of the detector to which the particles are propagated. An illustration of the `TauRunner` geometry and a diagram of the functional relation of physical quantities to the affine parameter is shown in Fig. 4.2

The `Track` class allows the user to make custom trajectories. The user need only specify mappings between the affine parameter and these variables. Different trajectories may require additional arguments from the user, depending on the nature of the trajectory. To illustrate this point, we can look at the two tracks which are implemented by default, the `Chord` and `Radial` trajectories. The former is used for paths which originate outside the `Body` and cross a section of `Body`. The latter is used for paths which originate at the center of the `Body`. The former `Track` describes neutrinos coming from space and passing through Earth on the way to a detector, as in the case of Earth-skimming $\tau$ searches, while the latter gives the trajectory of a neutrino originating in the center of the planet, relevant for searches for neutrinos from gravitationally trapped dark matter. Clearly, an incoming angle needs to be specified for the `Chord` trajectory. Thus, we can see that the necessary arguments for specifying a `Track` may vary from one geometry to another.

***`Body`***— The `Body` class specifies the medium in which the `Particle` is to be propagated. In `TauRunner`, we require that all bodies be spherically symmetric, and so a `Body` may be minimally specified by a physical radius, and a density profile. The density profile may be a positive scalar, a unary function which returns a positive scalar, or a potentially-mixed list of positive scalars and such functions. The sole argument of the functions used to specify the density should be the radius at which the density is to be given, in units of the radius of the body, *i.e.* the domains should be $[0, 1]$. In this system $r = 0$ is the center of the body and $r = 1$ the surface. If the user wishes to make a layered body, *i.e.* one where a list specifies the density profile, they must pass a list of tuple

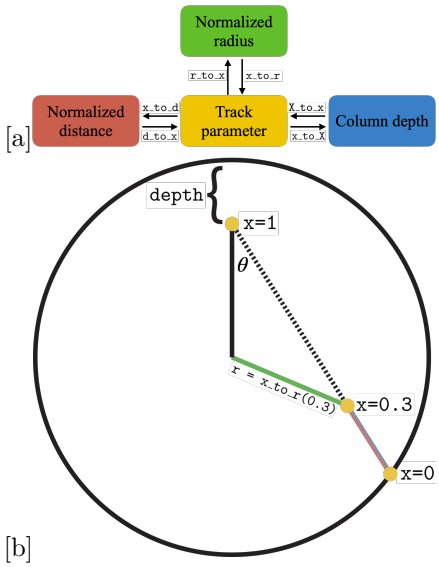



Figure 4.2: Schematic of `TauRunner` geometry as contained within the `Track` class. (a) shows the relation between the physical quantities relevant to propagation and the affine parameter that parametrizes the `Track`. The arrows connecting these quantities are labeled with the functions used to convert between them in `TauRunner`. Specifically, these are the functions a user must define in order to specify a custom `Track` geometry. All distances are normalized with respect to the radius of the body in which the track sits. (b) shows a diagram of these parameters within a spherical `TauRunner` body. Colors correspond to the boxes in (a). Additionally, it illustrates the `depth` parameter which intuitively gives the depth of the detector.

with the length of this list equal to the number of layers. The first element of each tuple should be the scalar or function which gives the density, and the second element should be the right hand boundary of the layer in units of the radius. The last right hand boundary should always be 1 since $r = 1$ is the outer edge of the body. Lastly, all densities should be specified in g/cm$^3$.

In addition to a radius and a density profile, the user may also provide the `proton_fraction` argument to specify the fraction of protons to total nucleons in the body. By default, we assume that the propagation medium is isoscalar, *i.e.* we set the proton fraction to 0.5 throughout the entire body. As in the case of the density profile, this argument may be a scalar, a function, or a list of function-boundary tuples. The domains of any functions provided must be [0, 1], and the ranges must be in this same interval.

While the user can construct bodies themselves, there are five bodies implemented by default in `TauRunner`: the Earth, a high-metallicity Sun, and low-metallicity Sun, the moon, a constant density slab. We use the PREM parametrization to model the densities of Earth [150]. For the Sun, we use fits provided by [151]. To instantiate the `Earth` object, one calls the `construct_earth`

function, which returns an `Earth` object. Additionally, this function allows one to pass in a list of additional layers which will be placed radially outward from the edge of the PREM Earth. This functionality may be useful for *e.g.* adding a layer of water or ice or adding the atmosphere for simulating atmospheric air showers. To initialize the Sun, one can use the `construct_sun` function. With this function, the user may specify 'HZ_Sun' or 'LZ_Sun' to use the high- and low-metallicity `TauRunner` suns respectively, or a path to a user defined solar model.

***`CrossSection`***— The `TauRunner` cross sections module defines the neutrino interactions. Internally, `TauRunner` assumes that cross sections are equal for all neutrino flavors. Additionally, `TauRunner` uses the isoscalar approximation by default, *i.e.* it assumes a medium is made of equal parts $p^+$ and $n$; however, this assumption may be changed by altering the `proton_fraction` of the `Body` object. The software includes both CSMS [85] and dipole [152] cross sections implemented by default; however, it is straightforward for the user to implement other cross section models by providing `scipy` splines in the appropriate format. For the total neutrino cross section these splines are `scipy.interpolate.UnivariateSpline` objects whose $x$-axis is the $\log_{10}$ of the neutrino energy in eV and whose $y$-axis is the $\log_{10}$ of cross section in $\text{cm}^2$. The differential cross section splines are `scipy.interpolate.RectBivariateSpline` objects whose $x$-axis is the $\log_{10}$ of the neutrino energy in eV, whose $y$-axis is a convenience variable which combines the incoming and outgoing neutrino energies, $E_{\text{in}}$ and $E_{\text{out}}$, given by

$$\eta = \frac{E_{\text{out}} - 10^9 \text{ eV}}{E_{\text{in}} - 10^9 \text{ eV}},$$

and whose $z$-axis is the $\log_{10}$ of incoming energy times the differential cross section in $\text{cm}^2$.

***`PROPOSAL`***— To propagate charged leptons, `TauRunner` relies on `PROPOSAL`, an open source C++ program with `Python` bindings. A utility module to interface with `PROPOSAL`, `utils/make_propagator.py`, is provided with `TauRunner`. This function instantiates `PROPOSAL` particle and geometry objects, which are then used to create a propagator instance. Since `PROPOSAL` does not support variable density geometries, the `segment_body` function is used to segment the `TauRunner` body into a number of constant density layers. The number of layers is determined by solving for points in the body where fractional change in the density is equal to a constant factor, called `granularity`. This argument may be specified by the user, and by default is set to 0.5. A single propagator object is created for all $\tau^\pm$ and, if needed, for all $\mu^\pm$. Since `TauRunner` assumes $e^\pm$ are always absorbed, a propagator will never be made for these. Whenever a new geometry is used, `PROPOSAL` creates energy loss tables which are saved in `resources/proposal_tables`. The tables require a few minutes to generate,

resulting in an overhead for new configurations, but subsequent simulations with the same geometry will not suffer any slow down.

### Conventions

`TauRunner` uses a natural unit system in which $\hbar = c = \text{eV} = 1$. As a consequence of this system, any energy passed to `TauRunner` must be in eV. `TauRunner` includes a `units` package to easily convert common units to the units `TauRunner` expects. This may be imported from the `utils` module, and its usage is demonstrated in several examples. Additionally, since `TauRunner` assumes that propagation occurs in a spherical body, the radius of this body establishes a natural length scale. Thus all distances are expressed as a fraction of this radius.

### Output

The `run_MC` function, which carries out the logic of `TauRunner`, returns a `numpy.recarray`. This array may be set to a variable if running `TauRunner` from a script of notebook, or printed or saved if running `TauRunner` from the command line.

In this paragraph, we will describe the fields of this output. The `"Eini"` field reports the initial energy of the lepton in eV. The `"Eout"` field reports the energy of the particle when propagation has stopped in eV. In the case that the particle was absorbed, this field will always read `0.0`. The `"theta"` field reports the incident angle of the lepton in degrees. The `"nCC"` and `"nNC"` fields report the number of charged and neutral current interactions the particle underwent in its propagation. The `"PDG_Encoding"` field reports the particle type, using the Particle Data Group MC numbering scheme. The `"event_ID"` is a number carried byfield reports which initial lepton the particle comes from. The `"final_position"` field reports the track parameter when the propagation was ended. This may be used to physical quantities of a particle when it was absorbed, or when a user-defined stopping condition was met

## 4.3 Performance

For a given primary spectrum and medium through which to propagate, there are a variety of related factors that determine the runtime of the program, including, but not limited to: (1) the

initial energy of the neutrinos, (2) the total column depth of the path, (3) the settings for computing energy losses, and (4) which particles are being tracked.

We show example runtimes for a few different use cases in Fig. 4.3. For a fixed `Track` propagating through Earth, neutrinos with higher initial energy take longer to propagate as they undergo more interactions and as a result experience more stochastic energy losses. Additionally, those particles that are only being propagated through Earth-skimming trajectories ($\cos(\theta) \approx 0$) can be simulated much quicker than those with large column depths. This is especially advantageous for proposed Earth-skimming next generation neutrino observatories, *e.g.* [141, 153, 138, 154, 140].

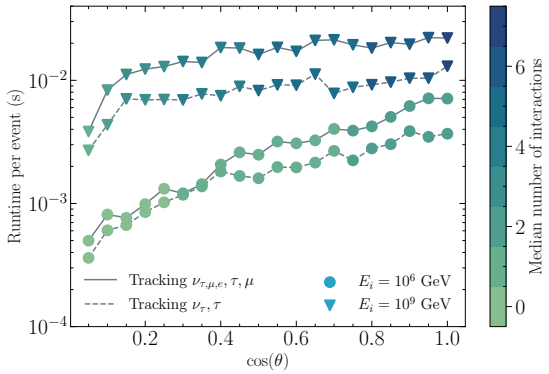



Figure 4.3: Runtime per $\nu_\tau$ event. Average runtime per event for various monochromatic fluxes of neutrinos through the Earth, as a function of nadir angle, $\theta$ for incident $\nu_\tau$ with energies of 1 PeV (circles) and 1 EeV (triangles). In general, runtime scales with the average number of interactions, which is a function of the energy of the particles and the column depth through which they propagate. The colorbar indicates the median number of NC+CC interactions that the initial beam of $\nu_\tau$ undergo. Tracking secondary particles (solid lines) created in $\nu_\tau$ CC interactions increases the runtime as the number of particles to propagate increases. Each point represents the average runtime from a simulation including $10^6$ events on a single CPU.

By default, all secondary particles that are created as a result of interactions are recorded, meaning that every $\nu_\tau$ CC interaction has a chance to increase the number of particles that need to be simulated. If the user is only interested in outgoing $\nu_\tau$ and $\tau$ lepton distributions, this option can be disabled with by setting `no_secondaries=True`, which can improve the overall runtime by as much as a factor of two.

Runtime can further be reduced depending on the treatment of energy losses of charged leptons. By default, energy losses are handled by `PROPOSAL` [68], which treats them stochastically. The user has the choice to ignore energy losses completely, with the setting `no_losses=True`, which can improve the runtime by as much as 40%, although this approximation can only be used in

certain scenarios, such as when the initial tau lepton energy is small enough that the interaction length becomes much smaller than the decay length. This has potential applications for recently proposed indirect searches of ultra-high-energy neutrinos by looking for PeV neutrinos through the Earth [38] using large current and next-generation ice or water Cherenkov detectors, such as IceCube-Gen2 [131]. Within `PROPOSAL`, there is also an option to treat energy losses that are below a certain threshold continuously. We find that setting this parameter to `vcut=1e-3`, meaning all energy losses that represent less than that fraction of the initial particle energy are treated without stochasticity, achieves an optimal runtime while not neglecting any of the important features that are a result of treating energy losses stochastically.

The first time that a user runs the code, there may be additional overhead while `PROPOSAL` calculates energy loss distributions for charged leptons. However, these tables are stored so that future iterations can run more efficiently. Once the user has run the code at least once and the `PROPOSAL` energy loss tables are stored, then current runtimes allow users to propagate approximately one million initial EeV $\nu_\tau$ through Earth's diameter in approximately eight hours with one CPU. For an initial energy of one PeV, one million $\nu_\tau$ take approximately one hour, depending on the incident angle. We also found that this runtime varied marginally from machine to machine, and the runtimes in Figure 4.3 and the numbers quoted thus far were all found using a heterogeneous distributed cluster of Linux machines. The code was also tested on a machine running MacOS with the Apple M1 chip, where the runtimes were found to extremely comparable to those presented above. For example, $10^4$ $\nu_\tau$ with initial energy of one EeV and $\theta = 0^\circ$ with no secondaries took 0.0127 s per event, on average, and those in the figure above took 0.0124 s per event, on average.

In terms of memory, `TauRunner` can be run on most modern machines, requiring only a few GB of RAM to run. For example, propagating $10^4$ $\nu_\tau$ through the Earth with initial energies of an EeV requires only approximately 1 GB of memory when tracking only $\nu_\tau$ and $\tau$, and approximately 3 GB when tracking all particles. The vast majority of this memory is allocated for calculating energy losses with PROPOSAL, *e.g.* for various trajectories through the Earth and for various initial energies, we found that $\sim 50 - 90\%$ of the memory usage was due to `PROPOSAL`. Because most of the memory is due to overhead from the energy losses, there is only a marginal increase in memory usage from propagating many more particles, *e.g.* two sample iterations of the code both took between 2.5 GB and 3.0 GB when propagating $10^4$ or $10^6$ $\nu_\tau$ through the Earth with the same initial energies and angles.

## 4.4 Results and Comparisons

The results of several tau neutrino simulation sets are illustrated in this section. Fig. 4.4 shows column-normalized distributions of outgoing neutrino energy fraction as a function of initial neutrino energy. Interestingly, the dashed line showing the median outgoing tau neutrino energy fraction varies with a constant slope, corresponding to the energy at which Earth becomes transparent. That energy is roughly 10 PeV at the horizon (top left), $\mathcal{O}(1)$ PeV in the mantle (top right and bottom left), and $\mathcal{O}(10)$ TeV through the core (bottom right). This means that for a large fraction of the Northern Sky, tau neutrinos pile-up and escape at energies where the atmospheric neutrino background is relatively low. This idea is also made clear when illustrated for a monochromatic flux. In Fig. 4.5, EeV tau neutrinos are propagated and the outgoing energies are plotted as a function of nadir angle. A similar feature can be seen, where a majority of neutrinos in this simulation escape with energy above 100 TeV.

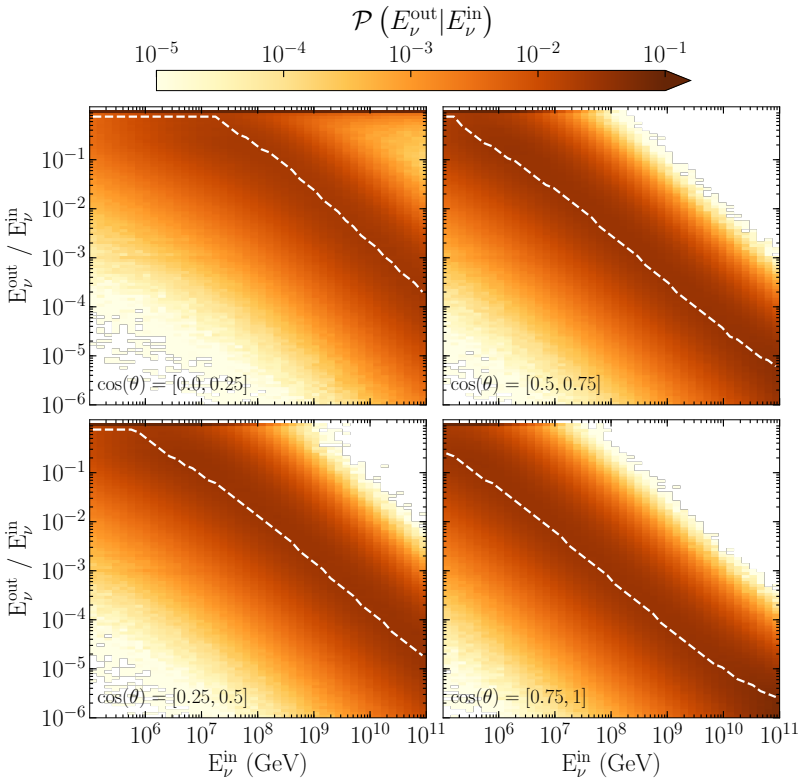



Figure 4.4: Outgoing $\nu_\tau$ distributions for an incident $E^{-1}$ power-law flux. Shown are the outgoing tau neutrino energy fraction as a function of the primary tau-neutrino neutrino flux injected as an $\mathrm{E}^{-1}$ power-law from 100 TeV to 10 EeV, shown in slices of equal solid angle in the Northern Sky. The dashed line indicates the median outgoing energy

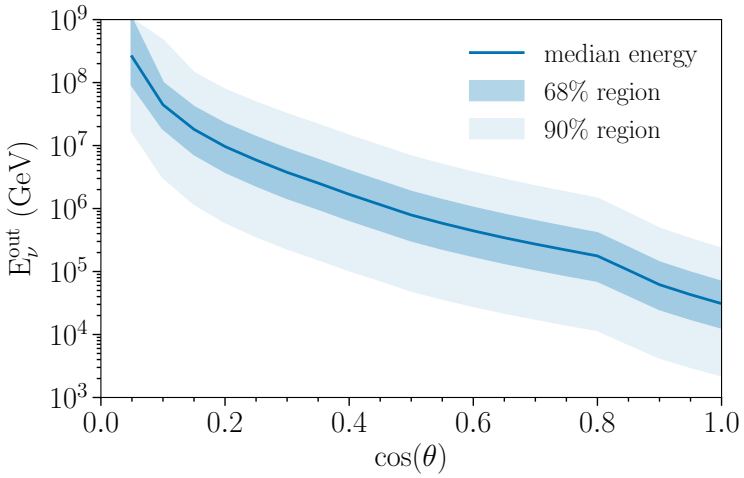



Figure 4.5: EeV tau neutrinos in Earth. Median outgoing energies of secondary tau neutrinos shown as a function of nadir angle. Also, 68% and 90% probability contours for outgoing energies are included. The feature at approximately $\cos\theta$ of 0.8 is caused by the core.

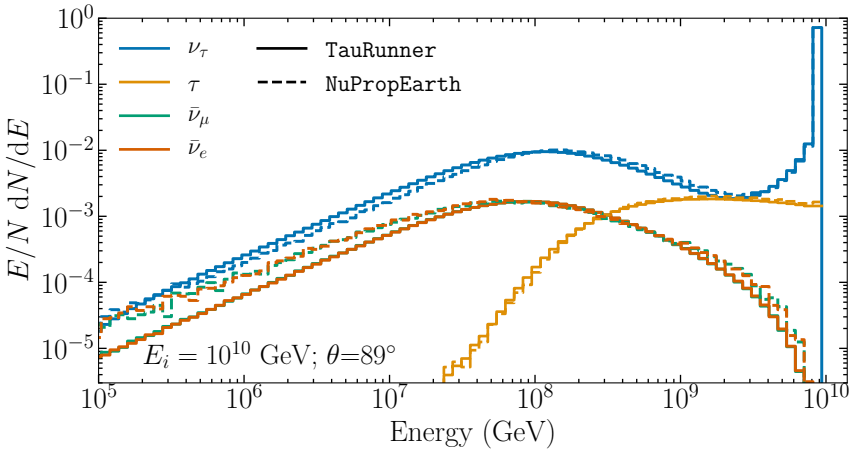



Figure 4.6: A monochromatic flux of tau neutrinos. Outgoing particle energy distributions for a fixed angle and energy. We include secondary anti-electron and -muon neutrinos, as well as charged taus. `TauRunner` shows good agreement with `NuPropEarth`. This set assumes Earth as a body with a 4km layer of water.

`TauRunner` has also been compared to several publicly available packages that perform similar tasks. A summary of the various tested packages and their features is shown in Tab. 4.1. Besides `TauRunner`, only `NuPropEarth` offers a full solution in the case of tau neutrinos. To illustrate this, we show in Fig. 4.6 the output of both packages for an injected monochromatic flux of tau neutrinos at $10^{10}$ GeV and one degree below the horizon. For secondary taus and tau neutrinos, the two packages show excellent agreement. We note that comparisons with `NuPropEarth` use the trunk version of

the code, which has a new treatment for charged particle propagation using `PROPOSAL` instead of `TAUSIC`. Secondary anti-muon and -electron neutrino distributions show slight disagreement in the tails, likely due to different tau polarization treatments. These differences are still being investigated, and will be addressed in an upcoming work.

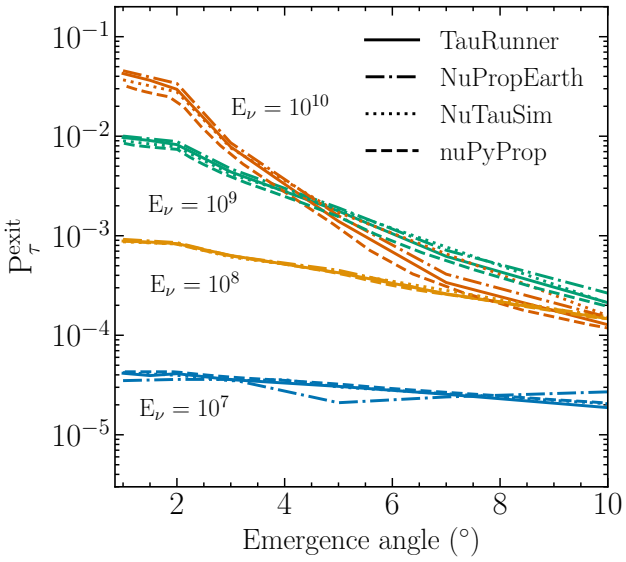



Figure 4.7: Charged tau lepton exit probability. Different colors correspond to four different monochromatic neutrino energies. The emergence angle is measured with respect to horizon. The `TauRunner` prediction (solid line) is compared to NuTauSim, NuPropEarth, and nuPyProp, which are shown in different linestyles.

Fig. 4.7 shows a comparison of the charged tau exit probability in Earth as a function of nadir angle. $P^{\tau}_{\text{exit}}$ is the probability that an incoming neutrino will exit Earth as a charged tau. This quantity is especially relevant for future neutrino observatories hoping to detect Earth-skimming tau neutrinos. In that scenario, exiting taus make up the bulk of the expected signal. `TauRunner` again shows great agreement overall with other packages.

| **Software** | **Language** | **Input** | **Output** | **Medium** | |
|---|---|---|---|---|---|
| `TauRunner` | `Python` | $\nu_{\tau,\mu,e}, \tau, \mu$ | $\nu_{\tau,\mu,e}, \tau, \mu$ | Earth/Sun/Moon/Custom | |
| `NuPropEarth`[86] | `C++` | $\nu_{\tau,\mu,e}$ | $\nu_{\tau,\mu,e}, \tau$ | Earth/Custom | |
| `nuPyProp`[155] | `Python/FORTRAN` | $\nu_\tau$ | $\tau$ | Earth | |
| `NuTauSim`[156] | `C++` | $\nu_\tau$ | $\tau$ | Earth | |

Table 4.1: Each row of this table represents a given package. Input and output particles include their not explicitly mentioned antiparticles. Custom medium refers to a user-defined `Body` in `TauRunner`.

## 4.5 Implementation of tau depolarization

The theory and derivation of tau depolarization were discussed in Sec. 3.3. Here, we present the implementation of the ideas discussed in Sec 3.3 and focus on tau propagation in rock, first to determine the distribution of the tau polarization just before it decays, then to illustrate impact of depolarization on the energy distribution of the tau neutrinos that come from tau decays.

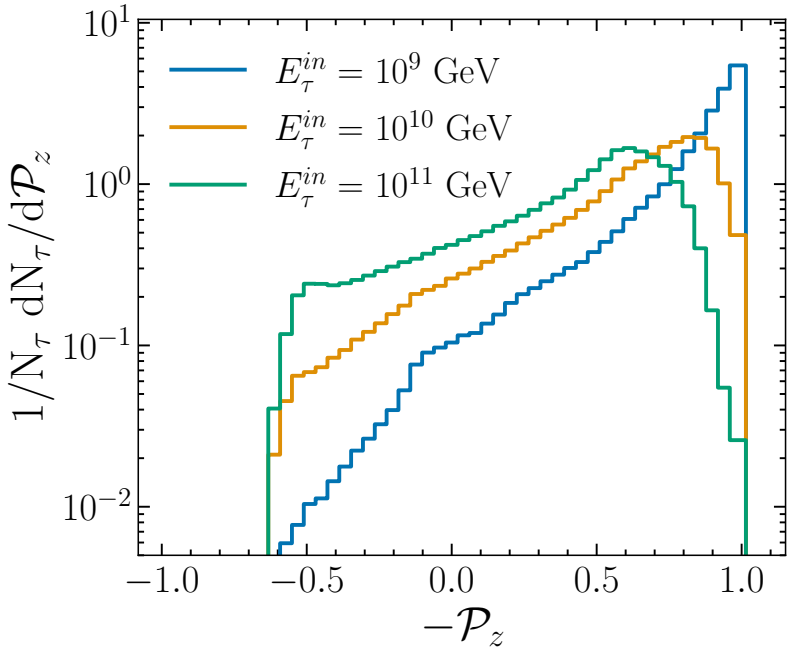



Figure 4.8: Distribution of final tau polarization before decay, for monoenergetic taus in rock. The differential number of taus as a function of final polarization from multiple EM interactions for $10^7$ $\tau$'s propagated through a slab of rock for each initial tau energy.

We use `TauRunner` to propagate tau leptons at multiple initial energies through a slab of 200 km.w.e. of standard rock ($A = 22$, $Z = 11$, and $\rho = 2.65$ g/$cm^3$), schematically illustrated in the left panel of Fig. 4.8. With this depth of rock, all of the taus decay in the slab. The simulation records $y$ (inelasticity) for each EM interaction of the tau. For photonuclear interactions, the corresponding $\langle \mathcal{P}_{\tau,z}^{EM} \rangle$ and $\langle P^{EM} \rangle$ (see Fig. 3.5) are used to determine $\theta_P$ and $P$ for all photonuclear interactions, combined according to:

$$\cos\theta_P = \frac{\langle \mathcal{P}_{\tau,z}^{EM} \rangle}{\langle P^{EM} \rangle} \tag{4.1}$$

$$\theta_{P,f} = \theta_{P,1} \pm \theta_{P,2} \pm \theta_{P,3} \pm \ldots \tag{4.2}$$

$$P_f = P_1 \cdot P_2 \cdot P_3 \cdot \ldots, \tag{4.3}$$

Then, the final polarization about the $z$-axis for each tau is given by

$$\mathcal{P}_z = P_f \cos\theta_{P,f}\,, \tag{4.4}$$

(For a single interaction of the tau, we have defined its polarization about $z$-axis with the notation $\mathcal{P}_{\tau,z}$.) Using Eq. 4.4 with our simulated data we show our results in the right panel of Fig. 4.8 for three incident tau energies, $E_\tau^{in} = 10^9$ GeV, $10^{10}$ GeV, and $10^{11}$ GeV. It is observed that for $E_\tau^{in} = 10^9$ GeV, the taus are more polarized and therefore the distribution of $-P_z$ peaks closer to one. On the other hand, for $E_\tau^{in} = 10^{11}$ GeV, there is more depolarization and we see a shift in the peak away from one. This is due to the increase in number of interactions for higher initial tau energy which causes more depolarization.

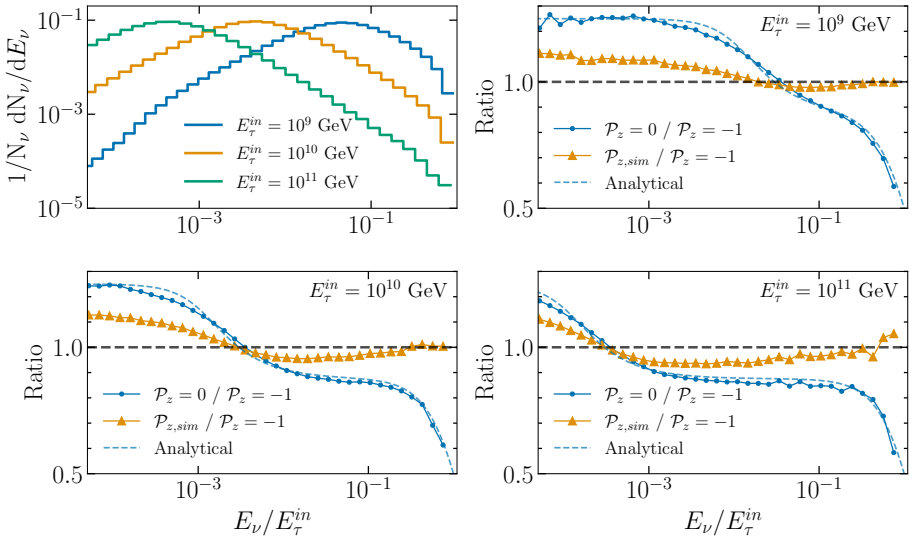



Figure 4.9: Effect of tau depolarization on outgoing tau neutrinos for multiple starting energies. Upper left: Differential number of tau neutrinos as a function of $\nu_\tau$ energy fraction for simulated depolarization taus, shown for three different energies. Upper right: The orange (blue) markers show the ratio of the tau neutrino energy distribution with simulated ($\mathcal{P}_z = 0$) polarization to the tau neutrino energy distribution for left-handed tau decays as a function of $E_\nu/E_\tau^{in}$ for taus of energy $E_\tau^{in} = 10^9$ GeV in rock. The dashed blue curve shows the approximate analytic evaluation of the ratio of decay neutrino distributions with $\mathcal{P}_z = 0$ to $\mathcal{P}_z = -1$. The lower plots show these ratios for $E_\tau^{in} = 10^{10}$ GeV (left) and $E_\tau^{in} = 10^{11}$ GeV (right).

The bump in the right panel of Fig. 4.8 for $E_\tau^{in} = 10^{11}$ GeV at $\mathcal{P}_z \simeq -0.6$ arises because $\cos\theta_{P,f}$ can have negative values, i.e., $\theta_{P,f} > \pi/2$. The distribution for $P_f$ peaks at $\sim 0.6$ for $E_\tau^{in} = 10^{11}$ GeV, which combined with negative values of $\cos\theta_{P,f}$ gives $\mathcal{P}_z \simeq -0.6$ for some fraction of taus.

We now turn to the neutrino energy distribution from tau decays after propagating in rock. For each incident tau at a fixed energy, we calculate the energy of the tau neutrino from tau decay

including the final tau energy and polarization. This is done by creating a cumulative distribution function from the neutrino energy distribution equation given in Eq. 3.3.

We show the effect of depolarization of taus on $\nu_\tau$ energy distribution in Fig. 4.9. The upper left plot shows the differential number of tau-neutrinos as a function of tau-neutrino energy fraction ($E_\nu/E_\tau^{in}$). Higher energy taus lose more of their initial energy as they propagate farther before their decays.

For three initial tau energies, the remaining plots in Fig. 4.9 show ratio of $\nu_\tau$'s produced from unpolarized ($\mathcal{P}_z = 0$) to fully polarized ($\mathcal{P}_z = -1$) taus (blue markers and curve), and simulated ($\mathcal{P}_{z,sim}$) to fully polarized ($\mathcal{P}_z = -1$) taus (orange markers and curve). The cross-over points of the ratio plots approximately correspond to the peaks in the upper left plot of Fig. 4.9.

To cross-check our results for the neutrino energy distribution from tau decays given taus incident on rock, we used an approximate analytical equation to get the $\nu_\tau$ spectrum. We show dashed blue curves with this semianalytic approximation in Fig. 4.9. The ratio of the analytic evaluation of the neutrino energy distribution of unpolarized to fully polarized tau agrees very well with the ratio of distributions for $\mathcal{P}_z = 0$ to $\mathcal{P}_z = -1$ from the Monte Carlo.

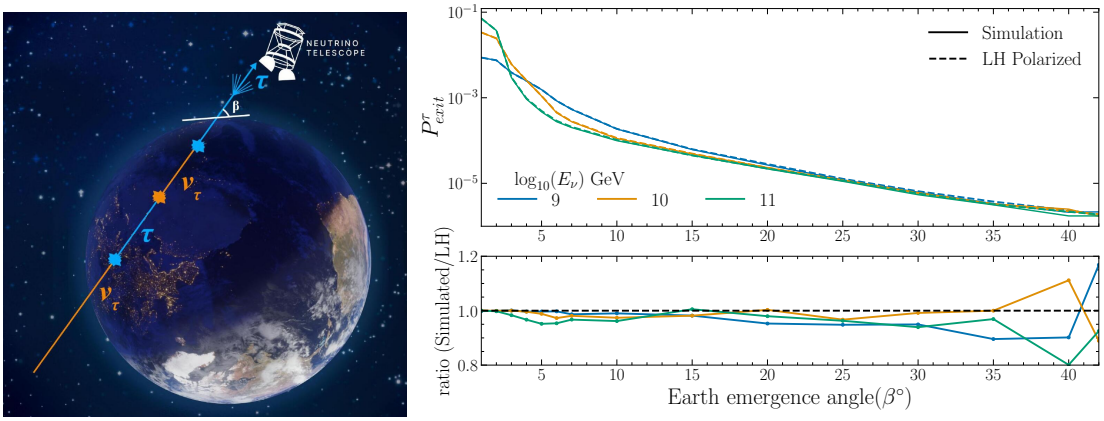



Figure 4.10: Impact of tau depolarization on the tau exit probability. Left: Schematic diagram of a $\nu_\tau$ incident on the Earth that results in an emerging tau after series of CC interactions and tau decays (regeneration). The Earth emergence angle is $\beta$. Right: Exit probability of taus as a function of different Earth emergence angles, for three different initial tau-neutrino energies. It shows a comparison when we consider LH polarization and simulated depolarization for EM interactions of the taus.

***Results for Earth-skimming tau neutrinos***— Now we implement this depolarization technique into our full neutrino simulation, allowing us to study its effects on neutrino telescopes. We simulate $\nu_\tau$'s skimming Earth, at different Earth emergence angles ($\beta$), taking into account tau

depolarization. A schematic of Earth-skimming tau neutrino trajectories to produce a tau that emerges from Earth is shown in the left panel of Fig. 4.10.

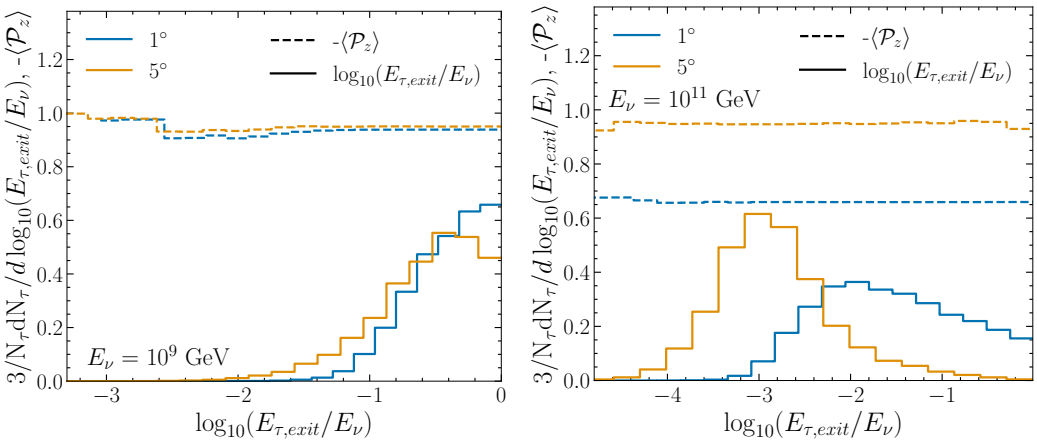



Figure 4.11: Final energy of Earth-emerging taus and their polarization starting from a neutrino beam. Differential number of exiting taus and average polarization ($-\langle\mathcal{P}_z\rangle$) of exiting taus as a function of exiting tau's energy fraction for initial neutrino energies $E_\nu = 10^9$ GeV (left) and $10^{11}$ GeV (right) and for $\beta = 1°, 5°$. The energy distribution normalization was chosen so the energy distributions and polarizations can appear in the same figures. Note that the $x$-axes have different ranges in the two panels.

For Earth-based, suborbital and satellite instruments that detect signals of tau decay-induced extensive air showers, modeling requires the probability that a neutrino produces an exiting tau, the energy of the emerging tau, and its final polarization upon exit. In the right panel of Fig. 4.10, the exit probability of the taus is plotted for different Earth emergence angles, for three different initial tau neutrino energies. It shows a comparison between fully polarized and simulated depolarized taus. We observe that the exit probability is changed by 5% for smaller angles, and 10% for higher angles, when we consider depolarization in the EM interactions. This shows us that depolarization has a small impact on the exit probability of the taus.

Fig. 4.11 shows us the final energy of the exiting taus with corresponding average polarization. The energy distributions are normalized so that they appear on the same scale as the polarization curves included in the figures. The plot on left is for $E_\nu = 10^9$ GeV for $\beta = 1°$ and $5°$. For this energy, the regeneration rate is negligible for both angles. The taus that exit the Earth are the ones created from the initial tau neutrinos that interact close to the surface of the Earth. Thus there is no significant depolarization and the final energy of the exiting taus is close to the initial tau-neutrino energy. The right-hand plot in Fig. 4.11 for $E_\nu = 10^{11} GeV$, for $\beta = 1°$, the high energy

taus that are produced are partially depolarized as they propagate and lose energy on their way to exit Earth. For $\beta = 5^\circ$, the taus able to exit Earth are created from the regenerated tau neutrinos which interact closer to the surface of Earth. Because of regeneration, the energy distribution of the emerging taus is lower than for $\beta = 1^\circ$. Since the taus that emerge are produced close to the Earth surface, and with each tau production, its polarization is reset to $-1$, the average polarization for exiting taus is close to $-1$ for $\beta = 5^\circ$ for incident neutrinos with $E_\nu = 10^{11}$ GeV. In principle, tau depolarization effects can affect the flux normalization and the energy distributions of tau neutrinos and taus that arrive at underground detectors or taus that emerge to produce upgoing air showers. We have performed an analysis of the dominant contribution to the depolarization of CC interaction produced left-handed taus as they transit materials.

The depolarization of taus is not complete for tau energies up to $10^{11}$ GeV. With our simulations of tau energy loss in rock, Fig. 4.9 shows that the neutrino energy distributions from tau decays are shifted from LH tau decays by $\sim \pm 10\%$. We have shown that the tau exit probability is modified at most by $\sim 10\%$ for large Earth emergence angles, where the exit probability is already low. The energy distribution of the emerging taus is essentially the same with and without accounting for EM depolarization effects. Our results for the polarization of Earth-emerging taus show that the average polarization depends on the incident neutrino energy and angle, but it is largely independent of the final tau energy. Improved modeling of the initial energy of the extensive air shower from tau decays in the atmosphere by including $\mathcal{P}_z$ is therefore straightforward to implement in Monte Carlo simulations with stochastic energy loss like `TauRunner`.

# Chapter 5

# Implications of Accurate Tau Neutrino Modeling

This chapter builds on chapters 3 and 4 to showcase the physics potential of accurate tau neutrino modeling. We discuss the impact on GZK neutrino searches for Cherenkov telescopes in Sec. 5.1 and perform a sensitivity analysis for IceCube, showing that current limits can be improved by up to a factor of two. We then highlight the promise of detecting ultra-high-energy transients with tau neutrinos in Sec. 5.2, showing that current IceCube differential sensitivities are in fact orders of magnitude better in the Northern Sky when using muons from tau neutrinos. Finally, We discuss two followups to anomalous ANITA events using tau neutrinos in Sec. 5.3 and show that all standard model explanations are ruled out.

## 5.1 GZK Neutrinos

At energies beyond the so-called Greisen-Zatsepin-Kuzmin (GZK) cutoff ($E \geq 40$ EeV), proton interactions with the CMB restrict the mean free path of cosmic-ray nuclei primaries to less than a few hundred Mpc from sites of cosmic acceleration. A suppression compatible with the GZK cutoff has indeed been observed in cosmic-ray experiments [157, 158, 159, 160]. The subsequent decay of the mesons from these interactions leads to an observable, yet currently undetected, flux of GZK neutrinos.

Although the GZK flux should be isotropic at Earth, searches for this flux have been limited to either half of the sky (downgoing) or often smaller solid angles, specifically looking for Earth-

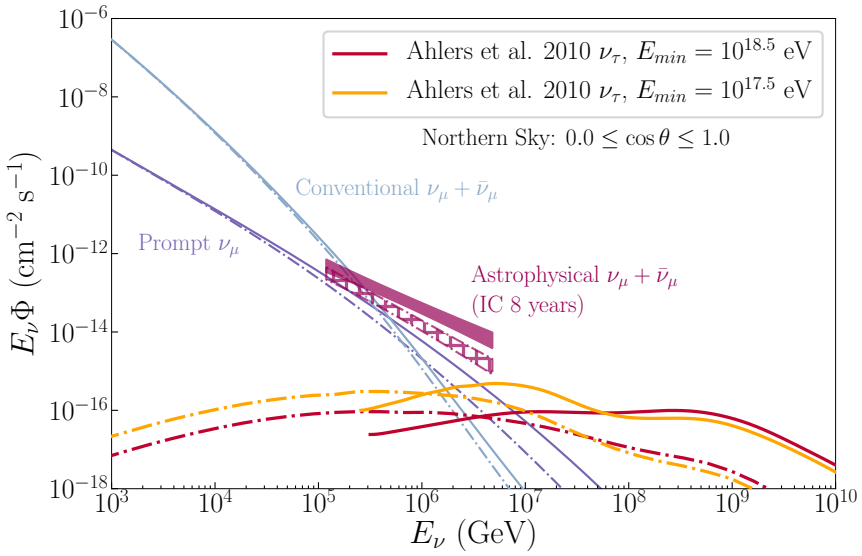



Figure 5.1: Per flavor neutrino fluxes from 1 TeV to 10 EeV, integrated over the northern sky. Primary fluxes are shown as solid lines, and fluxes present at the IceCube detector are shown in dashed-dotted lines and hatches. The $\nu_\tau$ components for various models of the cosmogenic flux are shown in red and orange [170]. These spectra are compared to models of both the conventional [171] and prompt [172] components of the atmospheric flux as well as measurements of the diffuse astrophysical flux [173]. Secondary $\nu_\tau$ spectra peak at PeV energies, a region of parameter space optimal for neutrino telescopes such as IceCube.

skimming neutrinos, where the probability of detecting a tau in the atmosphere after a single neutrino interaction in Earth is optimized [161, 162, 163, 164, 165, 166, 167, 168, 169]. Here, we show how using the secondary tau neutrino flux will extend this search to the entire sky. Specifically, we look for neutrinos after several interactions in Earth, which emerge at $\mathcal{O}$(PeV) energies. We also show that the rate from Earth-traversing neutrinos is not negligible, and quantify the gain in sensitivity when including them.

Fig. 5.1 displays the secondary tau neutrino flux of GZK neutrinos compared to atmospheric and diffuse astrophysical per-flavor neutrino fluxes in the northern sky. For a representative GZK flux, we choose a model produced from a fit to HiRes data [170]. We have also considered other fluxes in the literature [174, 169, 175] and find that the shape of the secondary flux at the detector is roughly consistent and differences manifest mostly in the overall normalization. Predictions of number of events varied from 1 to 18 for the proton-dominated fluxes considered. In the figures, we show our benchmark flux [170] as an example, as it provides a picture of the average expectation for proton-dominated fluxes. The conventional component in Fig. 5.1 shows the $\nu_\mu$ flux produced in cosmic-ray showers in the atmosphere, using the model in [171]. The prompt component is the

expected muon neutrino flux arising from atmospheric charm production in cosmic-ray showers; we use the model in [172]. Although there is a predicted tau neutrino component of this prompt flux, predominantly from D-meson decays, the level of this flux is much smaller compared to the prompt electon- and muon-neutrino components. The astrophysical muon neutrino flux we use is based on eight years of northern sky muon track data from IceCube [173]. All of these primary fluxes are propagated to the detector. The fluxes arriving at the detector are then compared to the secondary flux from GZK neutrinos. The spectrum of the secondary GZK flux is much harder and strongly dependent on declination, providing additional handles to distinguish GZK secondaries from other astrophysical or atmospheric events. The tau neutrino component of the astrophysical neutrino flux (produced directly at cosmic accelerators and not from proton interactions with the CMB) could also contribute to the secondary neutrino flux in the northern sky. This contribution, and its relation to the GZK secondaries, depends on the spectral index and maximum energy of the primary astrophysical flux. If the primary astrophysical flux is assumed to be a hard power law (such as that of the through going muon neutrino sample [117]) and unbroken out to energies exceeding $10^9$ GeV, then the secondary rate from such a flux would be comparable to or higher than that of the GZK flux up to $\mathcal{O}$(10 PeV). However, if the primary flux is softer (such as the measurement from the high energy starting events sample [176]) or if the energy spectrum cuts off below $10^9$ GeV, then the secondary rate would be lower than that from the GZK flux at energies above $\sim 3$ PeV.

To further highlight the expected signal shape, we show the resulting expected signal distribution of this benchmark model in Fig. 5.2. The number of expected signal events at IceCube is calculated by propagating a $\nu_\tau$ flux isotropically over the Northern Hemisphere from incidence on the Earth to a few kilometers away from IceCube. The number of events expected at IceCube is then given by

$$\mathcal{N}_\nu^{\mathrm{GZK}} = \int dE' d\Omega\ \Phi_\nu(E'_\nu)\Delta T \left[\sigma_{\nu N}^{CC}(E'_\nu)\cdot \frac{\Gamma_{\tau\to\mu}}{\Gamma_{total}}\cdot N_N^{CC}(E'_\nu) + \sigma_{\nu N}^{NC}(E'_\nu)\cdot N_N^{NC}\right], \tag{5.1}$$

where $\Phi_\nu(E'_\nu)$ is the emerging flux near the detector, $\sigma_{\nu N}^{CC}$ and $\sigma_{\nu N}^{NC}$ is the neutrino-nucleon isoscalar cross section for CC and NC, respectively, $\Gamma_{\tau\to\mu}/\Gamma_{total}$ is the tau to muon branching fraction, and $N_N$ is the effective number of isoscalar targets. This number is fixed to be $N$ targets in 1 km$^3$ of ice for the NC channel, but has an energy dependence for the CC channel due to the extended muon range, and is given by,

$$N_N^{CC}\left(E_\nu\right) = \int d\tilde{E}_\mu d\tilde{E}_\tau \frac{dN_\tau}{d\tilde{E}_\tau}\left(\tilde{E}_\tau; E'_\nu\right) \frac{dN_\mu}{dE_\mu}\left(E_\mu; \tilde{E}_\tau\right) R_\mu\left(E_\mu\right) A^{geo} \frac{\rho^{ice}}{M_{iso}}. \tag{5.2}$$

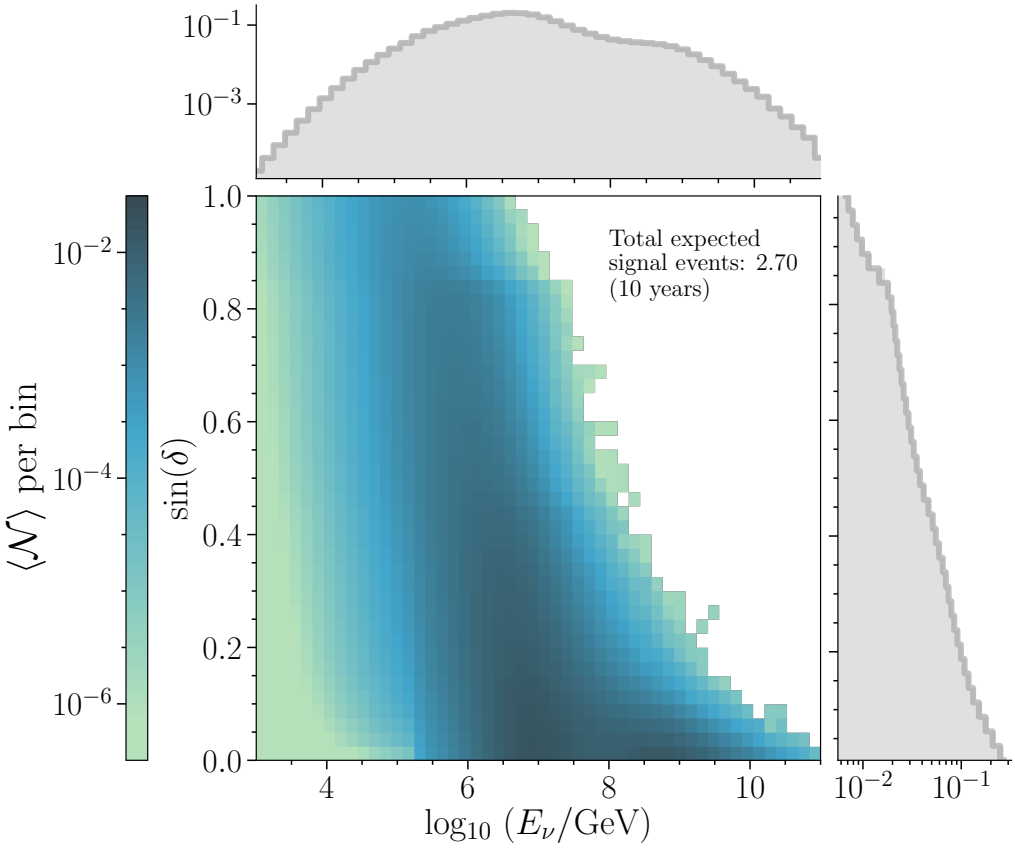



Figure 5.2: Expected signal of GZK neutrinos at IceCube, assuming the model from [170] and assuming a cosmic-ray composition that is dominated by protons above energies of $10^{17.5}$ eV, after 10 years of data collection. The Earth-skimming contribution represents only about one third of the total expectation, and the majority of events are expected to peak around 10 PeV in true neutrino energy.

In Eq. 5.2, the first and second terms are the tau and muon energy distributions, respectively, $R_\mu\left(E_\mu\right)$ is the average muon range calculated with MMC, $A^{geo}$ is the geometrical transverse area (1 km$^2$ in this case), $\rho^{ice}$ is the density of ice, and $M_{iso}$ is the isoscalar nucleon mass. Fig. 5.2 shows the expected number of events at IceCube binned in true neutrino energy and declination. We find that, assuming a proton-dominated UHECR flux with a minimum crossover energy of $10^{17.5}$ eV ($10^{18.5}$ eV), IceCube should see 2.70 (1.25) upgoing neutrinos with a hard energy spectrum, peaking at 10 PeV, in ten years of data taking. These events are dominated by the CC channel, with only around 10% of the signal coming from NC interactions in the fiducial volume of the detector. Of

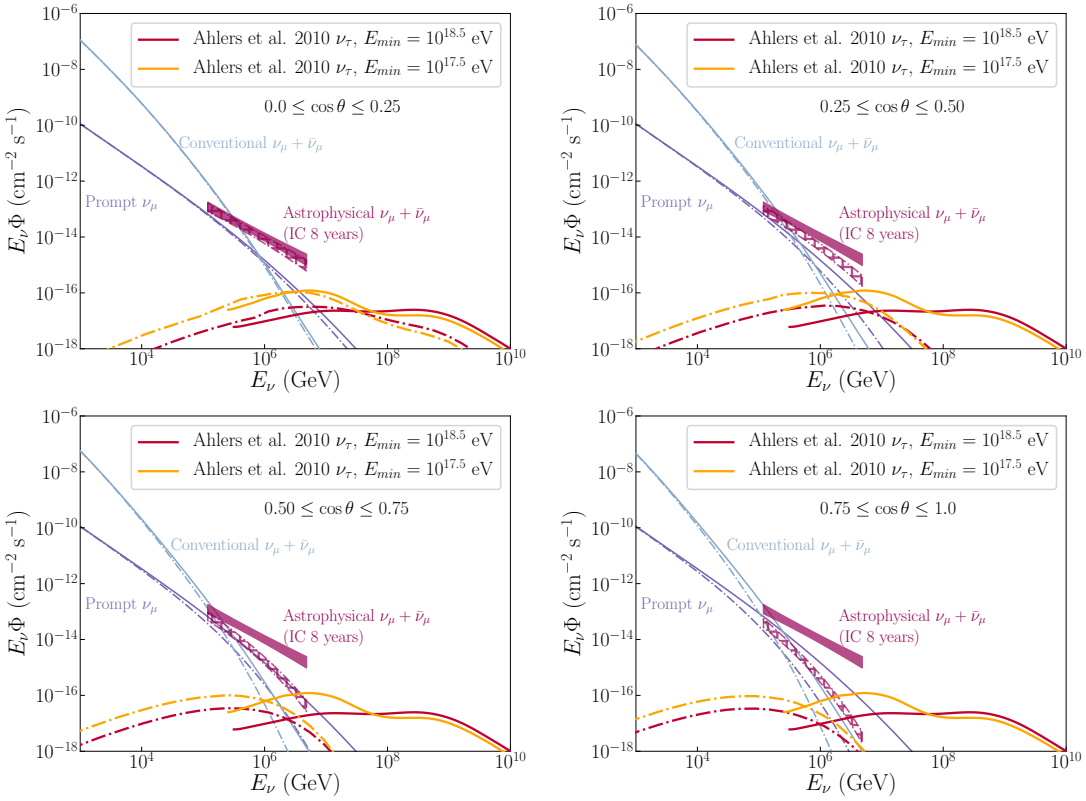



Figure 5.3: Per flavor neutrino fluxes from 1 TeV to 10 EeV, integrated over various zenith bands in the Northern Sky. Solid lines are primary fluxes, while secondary fluxes are represented by dashed-dotted lines and hatches. The secondary GZK tau neutrino spectrum is strongly dependent on the incoming zenith angle. For arrival directions towards Earth's core, it contributes equally to the astrophysical flux at IceCube above 2 PeV.

all of the events, we find that only ∼0.8 (0.5) would be Earth-skimming, where we have defined Earth-skimming to be up to 5 degrees below the horizon. Therefore, in total, we expect the rate from Earth-traversing neutrinos to be at least twice that from Earth-skimming neutrinos.

Fig. 5.3 further demonstrates the declination dependence of this flux through comparison to atmospheric backgrounds, and shows that in certain zenith angle bands with large enough column depth through Earth, the flux arriving at IceCube is higher than the atmospheric background at and above 2 PeV. Given that the event expectation is dominated by the composition of the primary cosmic rays, searching for neutrino secondaries can provide constraints on the composition of primary nuclei, independent of measurements performed by direct cosmic-ray experiments such as the Pierre Auger Observatory and Telescope Array[177, 160].

| | Model Rejection | Factors |
|---|---|---|
| **Time** | **Horizon** | **Horizon+Upgoing** |
| years | $\delta \leq 8^\circ$ | $\delta \leq 90^\circ$ |
| 1 | 1.65 | 0.77 |
| 5 | 0.64 | 0.27 |
| 10 | 0.42 | 0.17 |

Table 5.1: Sensitivity improvement for GZK neutrino searches. Each row corresponds to the number of years of data taking for IceCube-Gen2. The middle column shows the model rejection factor when measuring neutrinos near the horizon, while the rightmost column lists the corresponding number when measuring both horizon and upgoing events. The model rejection factor is defined as the ratio of the baseline GZK flux normalization with respect to the limit obtained in this analysis at 90% CL. As a baseline GZK fit we use the flux given by Ahlers *et. al.* 2010 [170].

Next, we illustrate the importance of tau regeneration in searching for GZK neutrinos with a statistical analysis. We evaluate the response of IceCube to neutrinos with energy above 100 TeV, where the atmospheric neutrino component is subdominant [178]. The muon-neutrino effective area is taken from [179]. We consider two isotropic components: an astrophysical component modeled by a power-law whose parameters have been determined using an independent cascade set of data and a GZK spectrum that we model according to [170]. The two components can be differentiated by their energy dependence as well as their angular distribution. This difference is caused by a contrasting amount of absorption as a function of energy and traversed column depth.

To quantify the significance of the Earth-traversing contribution, we use a binned Poisson likelihood to compare our prediction to a baseline expectation. We compute the number of muons produced by the interactions of muon and tau neutrinos in the vicinity of the detector. We bin the resulting muons linearly using twenty bins in zenith and logarithmically using two bins per decade in reconstructed energy. We model the detector energy resolution by introducing an 80% smearing of the initial energy assuming a normal distribution [180]. We compute the 90% CL. upper limit on the Ahlers-Halzen GZK model normalization, $\phi_0$, using an Asimov data set, i.e., we set the observed number of events to be the expected mean number of events in the absence of a GZK component. Our results are shown for one, five, and ten years of data taking in Tbl. 5.1. The numbers shown are nominally referred to as the model rejection factor (MRF), defined as the ratio of the nominal prediction of the model to the normalization that can be constrained by data at 90% C.L. In other words, numbers less than one in Table. 5.1 imply that the baseline model can be detected at greater

than 90% C.L. When considering only Earth-skimming neutrinos, namely declinations such that $\delta < 10°$, we get sensitivities that are comparable to current IceCube constraints [180]. However, when we include the Earth-traversing component, the sensitivity improves by approximately a factor of two, with most of the improvement coming from the range between 10° to 30° below the horizon.

## 5.2 Point Source Transient Limits

Astrophysical beam dumps in the vicinity of the sources of UHE cosmic rays provide the opportunity for the production of charged and neutral pions, and therefore, of neutrinos with energies exceeding tens of PeV. Potential sources of UHE neutrinos include binary neutron star (BNS) mergers [181], magnetars [182, 183, 184], young and fast spinning pulsars [185], supermassive black hole mergers [186], gamma-ray burst (GRB) afterglows [187, 188], and blazars with a luminous dust-torus [56, 189, 190]. The predicted neutrino energy spectrum peaks in the range of 10–100 PeV, energies that are typically below the sensitivity of the current generation of neutrino telescopes. However, the presence of a flux extending beyond 10 PeV presents an opportunity to probe the sources via the regenerated flux from tau neutrinos.

In order to demonstrate the power of the regeneration-based technique for identifying sources of UHE neutrinos, we compute the differential sensitivity of IceCube for a neutrino flux in the range of 1 PeV to 1 ZeV. Neutrinos with this energy traversing Earth will generate a cascade of muon and tau neutrino fluxes at the detector. In contrast to the primary all-flavor flux that is attenuated, the regenerated flux results in a substantial intensity of neutrinos of PeV energy, where the sensitivity is optimal for detection by optical Cherenkov detectors. Here, we utilize the IceCube differential sensitivity for transients computed in [191]. For each energy decade, we inject a flux of tau neutrinos, assuming the spectrum follows a $E^{-2}$ distribution. To obtain the normalization for each bin, we determine the corresponding flux of secondary neutrinos that reaches the detector. The sensitivity of IceCube to a flux of muons arriving at the horizon constrains the flux that arrives at the detector in any direction, when excluding Earth effects. Thus, we use the sensitivity at $\delta = 0$ to constrain the initial flux at higher energies; that is, we constrain the normalization of the injected spectrum at different declinations by folding the sensitivity at the horizon with the expected number of muons from tau decay produced by the propagated secondary neutrinos. In Fig. 5.4, we show the corresponding differential sensitivity of IceCube to the time-integrated flux in the energy range of $10^4$ GeV – $10^{11}$ GeV. For instance, for a declination of 30°, the $\nu_\tau$ sensitivity to

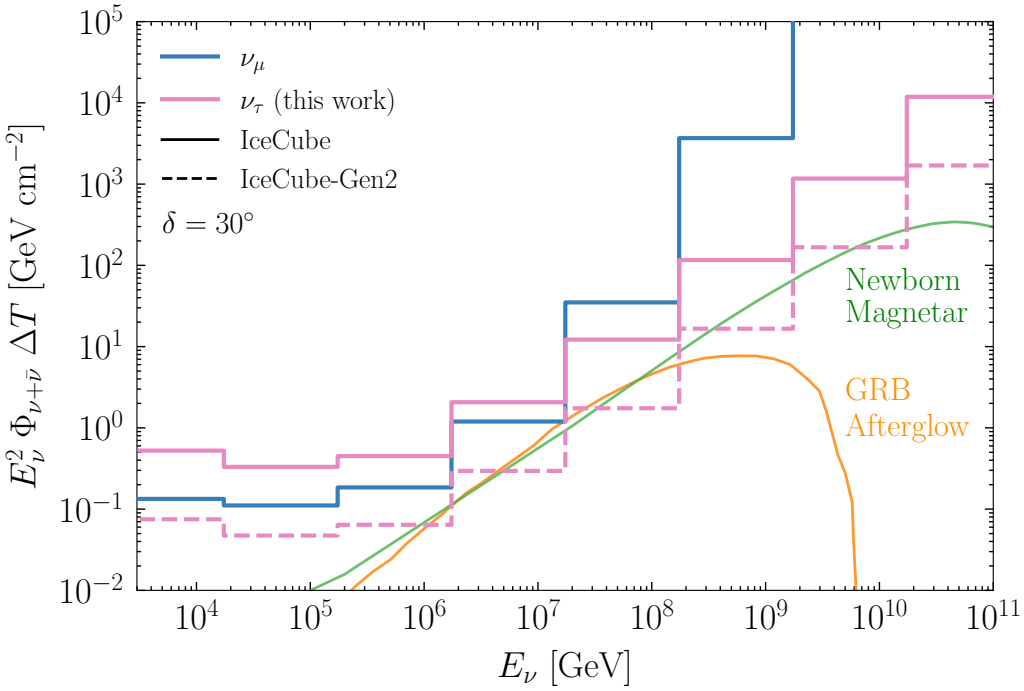



Figure 5.4: Improved point source sensitivity with $\nu_\tau$. Differential sensitivity to an $E^{-2}$ time-integrated flux for transient UHE sources at a fixed declination ($\delta = 30°$). The current sensitivity of IceCube is shown in blue, while the sensitivity to muons from tau neutrinos is shown in pastel pink. The projected sensitivity resulting from our method for IceCube-Gen2 (optical-only) is shown as a dashed line. To illustrate the power of our method, we compare it to the predicted time-integrated neutrino flux from newborn magnetars (green) [184] for a distance of 0.3 Mpc and GRB afterglows (orange) [187] at 20 Mpc.

a transient outperforms the conventional time-dependent search with $\nu_\mu$ for energies greater than $10^7$ GeV. The improvement grows with energy, exceeding an order of magnitude at $10^9$ GeV. We also show the corresponding sensitivity for IceCube-Gen2 demonstrating the enhanced reach of time-dependent searches with water-Cherenkov telescopes to sources of UHE neutrinos. The significance of this improvement is underscored by the power of this methodology to probe neutrinos from GRB afterglows and newborn magnetars, which is out of reach for conventional time-dependent follow-up. We note that the relative improvement in the sensitivity increases with declination, and reaches more than three orders of magnitude at $10^9$ GeV and 60 degrees below the horizon. The method creates new opportunities for the discovery of cosmic accelerators that produce the most energetic particles and were previously inaccessible to Cherenkov detectors.

## 5.3 ANITA Followup deux façons

The Antarctic Impulsive Transient Antenna (ANITA) project is a balloon experiment, designed with the primary purpose of detecting the UHE cosmogenic neutrino flux [192, 193, 166]. Although this is the project's primary scientific goal, the experiment is sensitive to a wide array of impulsive radio signals, and ANITA's first three flights have resulted in a few interesting detections. For a summary of the event properties, see Tab. 5.2. Two of these events had signatures consistent with upgoing air showers produced by a tau [194, 195]. This interpretation requires the decay of a tau (from a tau neutrino CC interaction) to occur in the atmosphere, producing an extensive air shower (EAS). This is distinguishable from a reflected EAS initiated by a cosmic ray, in which the radio signal acquires a phase reversal from reflection off of the Antarctic ice, while an upgoing EAS does not display such a phase reversal. However, this interpretation is problematic as tau neutrinos with energies to which ANITA is sensitive are not likely to travel through the large Earth column depths required for these events. While it has been noted that these events are unlikely to be caused by an isotropic neutrino flux [196, 197, 198], discrete-source emission could evade these constraints. Beyond the Standard Model (BSM) explanations have also been proposed. This includes axion-photon conversion [199], sterile neutrinos [198, 200, 201, 202], and heavy SUSY partners or dark matter particle decays [203, 204, 205, 206, 207, 208, 209, 210, 211]. Here, we examine the discrete-source emission hypothesis and show that any detection of EeV neutrinos from steep incident angles at ANITA implies a large and detectable flux of TeV-PeV neutrinos with IceCube.

The number of events detected by ANITA due to tau showers in the atmosphere from a primary neutrino flux, $\Phi(E_\nu)$, is given by

$$\mathcal{N}_\nu = \int dE_\nu dE'_\nu \; \Phi(E_\nu) \frac{dN_\nu}{dE'_\nu}(E'_\nu; E_\nu) \, \xi_{acc}(E'_\nu) \, \Delta T \; , \tag{5.3}$$

where $E_\nu$ and $E'_\nu$ are the primary and secondary neutrino energies, respectively, $dN_\nu(E'_\nu; E_\nu)/dE'_\nu$ is the energy distribution of secondary tau neutrinos near the ice surface, $\Delta T$ is the duration of observation, and $\xi_{acc}(E'_\nu)$ is the ANITA acceptance [196] in units of cm$^2$sr. The acceptance incorporates the probability of neutrinos interacting in the ice as well as the probability of a tau decay shower occurring in the atmosphere. Given that the reported acceptance in [196] includes neutrino propagation through Earth, we set the acceptance at all angles to be that near the horizon to remove the Earth absorption effects, which we account for separately with `TauRunner`. For

Table 5.2: Properties of the neutrino candidate events from the first three flights of ANITA, from [166, 194, 195]. The two anomalous ANITA events (AAE) are those consistent with a steeply upgoing $\nu_\tau$ interpretation. Sky coordinates are projections from event arrival angles at ANITA. Localization uncertainty is expressed as major and minor axis standard deviations, position angle. This angle describes the rotation of the major axis relative to the north celestial pole turning positive into right ascension.

| | **AAE-061228** | **AAE-141220** | **AAC-150108** |
|---|---|---|---|
| Event, Flight | 3985267, ANITA-I | 15717147, ANITA-III | 83139414, ANITA-III |
| Detection Channel | Geomagnetic | Geomagnetic | Askaryan |
| Date, Time (UTC) | 2006-12-28, 00:33:20 | 2014-12-20, 08:33:22.5 | 2015-01-08, 19:04:24.237 |
| RA, Dec (J2000) | 282°.14, +20°.33 | 50°.78, +38°.65 | 171°.45, +16°.30 |
| Localization | 1°.5 × 1°.5, 0°.0 | 1°.5 × 1°.5, 0°.0 | 5°.0 × 1°.0, +73°.7 |
| Reco. Energy (EeV) | 0.6 ± 0.4 | $0.56^{+0.30}_{-0.20}$ | ≥ 10 |
| Chord Length (km) | 5740 ± 60 | 7210 ± 55 | - |

the incoming flux, we take the minimalistic assumption of a delta function in energy, $\Phi(E_\nu) = \frac{dN}{dAdtdE_\nu} = \Phi_0 \delta(E_\nu - E_0)$, where $\Phi_0$ is the normalization with units $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Probabilities of tau neutrinos exiting Earth with energies greater than 0.1 EeV are shown in Fig. 5.5 for the chord lengths corresponding to AAE141220. For both taus and tau neutrinos, the probability of exiting Earth with an energy larger than 0.1 EeV seems to be fairly independent of energy, for initial tau neutrino energies above 1 EeV. Therefore, in what follows, we choose $E_0 = 1$ EeV as the primary energy.

As was discussed above, this primary flux of EeV neutrinos is guaranteed to be associated with a secondary flux of TeV to PeV neutrinos. Such a large rate of TeV muons simultaneously crossing the IceCube volume would deposit a large amount of charge. Large-charge events are promptly reported by IceCube via the EHE and HESE streams[1]. For example, the EHE stream requires three thousand photoelectrons and thirty channels to trigger an alert. A ∼1 PeV muon typically deposits ∼200 TeV, which on average results in over four thousand photoelectrons in more than 40 channels, when crossing the full detector [212]. In fact, the largest energy deposit reported in these streams[2] corresponds to 5 PeV, but it's for a downgoing event; horizontal and upgoing events have not had multi-PeV announcements in these streams. Thus, we conclude that IceCube has not observed catastrophic events that would be produced by bundles of TeV-neutrino-induced muons.

[1] https://gcn.gsfc.nasa.gov/amon.html
[2] https://gcn.gsfc.nasa.gov/gcn3/24028.gcn3

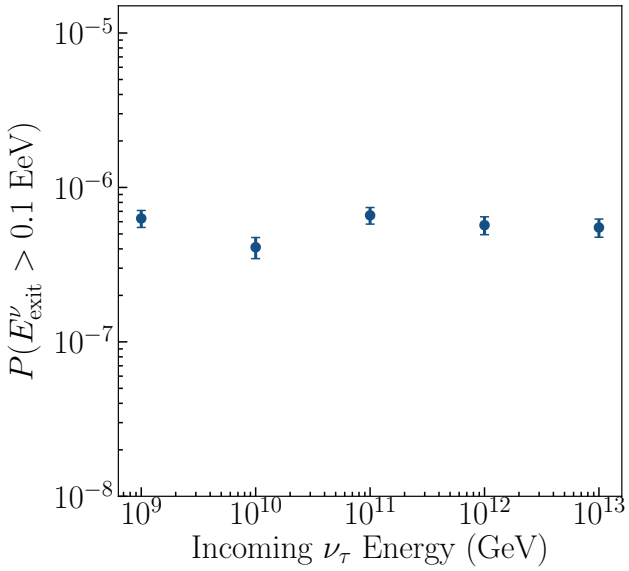



Figure 5.5: Probability for a tau neutrino to exit the Earth with a minimum energy of 0.1 EeV (approximate ANITA threshold), after Earth propagation (based on inferred chord length for AAE141220), assuming $\nu_\tau$ incidence at a particular initial energy. Errors are statistical only.

In what follows, we then take the conservative assumption that a single muon makes it through. Such events have been observed, and we can compare this expected yield to IceCube's measurement of the high-energy events. We find the maximum allowed normalization of the incident flux by comparing the secondary neutrino distribution with the measured IceCube astrophysical flux from the high-energy starting event selection (HESE) [213]. Results for AAE141220 are shown in Fig. 5.6. The unfolded HESE spectrum is propagated to the detector using `TauRunner`.

The 90% C.L. upper limit on the EeV primary flux normalization is set by comparing both secondary distributions and requiring that the secondaries produced by the primary EeV flux do not exceed those of HESE at 90% C.L. Given that the time profile of the intrinsic flux is unknown, we place limits on the time-integrated flux. We take the duration to be 22 days ($\Delta T$ in Eq. (5.3)), corresponding to the entire ANITA-III flight. We find the maximum allowed time-integrated flux to be $E^2\Phi\Delta T \simeq 10^2$ GeV cm$^{-2}$. Using the maximum allowed time-integrated flux, we calculate the expected number of events from ANITA. This yields a maximum expected number of neutrinos of fewer than $\mathcal{O}\left(10^{-7}\right)$ in 22 days. This is illustrated in Fig. 5.6 where we show the flux required to produce one event from ANITA as a reference. It is therefore extremely unlikely for the reported event to be caused by a high-energy tau neutrino.

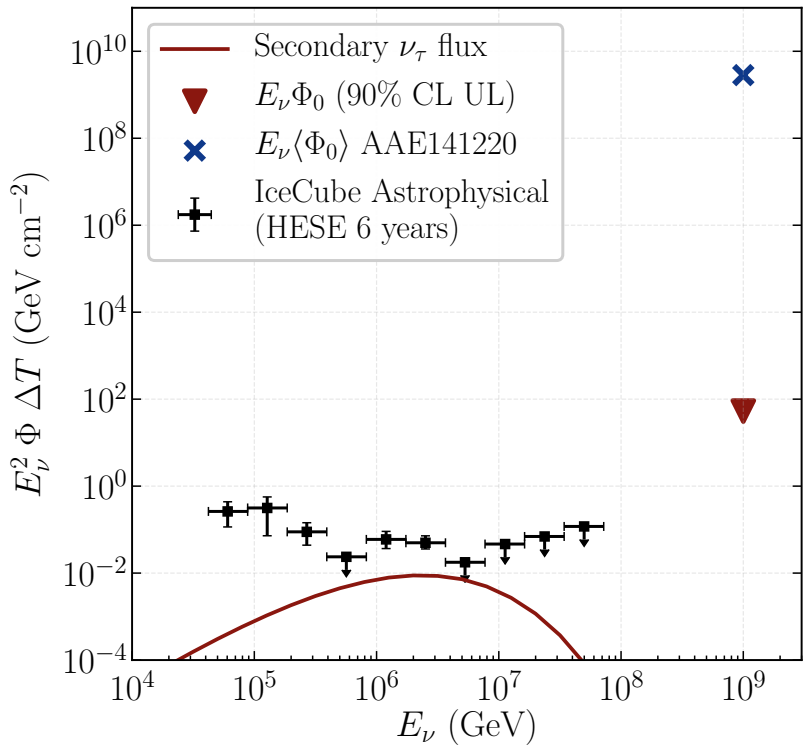



Figure 5.6: Maximum allowed flux of EeV neutrinos (maroon arrow), given an injected monoenergetic neutrino flux at or above the detected ANITA event (AAE141220) energy. The normalization of the secondary flux is set to the maximum that does not exceed IceCube's diffuse astrophysical flux (black bins). The flux needed to produce one event in the third flight of ANITA (blue marker) exceeds the upper limit by many orders of magnitude. We use the published spectrum based on six years of high-energy starting events.

It is worth noting, however, that short-timescale transients are allowed to overproduce the measured astrophysical flux as long as they do not overproduce the astrophysical flux integrated over the duration of the measurement. Therefore in what follows we present a dedicated transient source search using seven years of IceCube data. For this analysis, we focus on data collected from 2011 to 2018, containing approximately 900,000 events from 2532 days of operation.

We adopt the procedure described in [214] to search for counterparts to ANITA events. Namely, we perform an analysis to test for temporal structure in IceCube data in the location of the ANITA events. This analysis incorporates the information from the localization of the ANITA events through a joint likelihood. The sky is divided into grid positions, $\mathbf{x}_s$, and at each point we maximize the likelihood, $\mathcal{L}$, with respect to the expected number of signal events, $n_s$, and other signal parameters

contained in the variable $\alpha$ depending on the different signal hypotheses tested. This likelihood is given by

$$\mathcal{L} = \lambda \prod_{i=1}^{N} \left( \frac{n_s}{n_s + n_b} S(\mathbf{x}_i, \mathbf{x}_s, \alpha) + \frac{n_b}{n_s + n_b} B(\mathbf{x}_i, \mathbf{x}_s) \right) P_A(\mathbf{x}_s), \tag{5.4}$$

where $n_b$ is the expected number of observed background events and $N$ is the total number of observed events in the time window. The vector $\mathbf{x}_i$ contains the event observables such as its reconstructed energy, direction, and reconstruction uncertainty. $P_A$ is the spatial probability distribution function (PDF) of ANITA events, which are included in Table 5.2. $B$ describes the energy and declination PDF of our background, which is parameterized from data and is the same among all analyses. Temporal terms in $B$ vary according to the assumed emission time profile. While the signal PDF $S$ describes the signal hypothesis, the parameter $\lambda$ modifies the likelihood formalism in order to take into account low-statistics problems in some of the analyses. In general, the signal PDF, $S$, is defined as

$$S = S^{space}(\mathbf{x}_i, \mathbf{x}_s, \sigma_\mathbf{i}) \cdot S^{energy}(E_i, \delta_i, \gamma) \cdot S^{time} \ . \tag{5.5}$$

These three terms reflect the spatial, energy, and time PDFs, respectively, of our signal hypothesis. The spatial term, $S^{space}$, expresses the probability for an event with best-fit reconstructed direction $\mathbf{x}_i$ to originate from a source at the direction $\mathbf{x}_s$, according to a two-dimensional Gaussian function with angular resolution $\sigma_i$. The energy PDF, $S^{energy}$, describes the probability of obtaining an event with reconstructed energy $E_i$ given a declination $\delta_i$ under the hypothesis of an $E^{-\gamma}$ power-law energy spectrum, which helps differentiate signal from the known atmospheric backgrounds in our event selection. The time term, $S^{time}$, describes the time PDF of events observed from the source.

We search for IceCube events in spatial coincidence with the ANITA events in short time windows, $\Delta t$, centered on each ANITA event. We call this period the *on-time window*. This is equivalent to setting $S^{time}$ equal to a uniform PDF in this on-time window and to zero for all times outside this window. To help distinguish potential signals for time windows in which the expected number of background events is small, we set

$$\lambda = \frac{(n_s + n_b)^N}{N!} \cdot \mathrm{e}^{-(n_s + n_b)} \tag{5.6}$$

as in [215, 216]. Due to the small statistics for short time windows, the likelihood is only maximized with respect to $n_s$, and the energy dependence in $S^{energy}$ is fixed to an $E^{-2}$ spectrum. To account

Table 5.3: Analysis results and upper limits. Upper limits (90% C.L) are on the time-integrated $\nu_\mu + \bar{\nu}_\mu$ power law flux ($E^{-2}$ ) from a point source following the spatial probability distribution provided by ANITA. Limits are set assuming constant emission over a fixed time window.

| Event | Time Window | $p$-value | Upper limit (GeV $\cdot$ cm$^{-2}$) |
|---|---|---|---|
| AAE-141220 | 10 s | - | - |
| | $10^3$ s | 1.0 | 0.053 |
| | $10^5$ s | 1.0 | 0.051 |
| AAC-150108 | 10 s | 1.0 | 0.040 |
| | $10^3$ s | 1.0 | 0.041 |
| | $10^5$ s | 1.0 | 0.032 |

for the temperature dependence of atmospheric muon rates [217], we determine $n_b$ by calculating the rate of events from the surrounding five days of data on either side of our on-time window. Taking the logarithmic likelihood ratio between the maximum likelihood and that of the null hypothesis results in our test statistic (TS), defined as

$$\mathrm{TS} = -2\hat{n_s} + \sum_{i=1}^{N} 2\log\left[1 + \frac{\hat{n_s} S(\mathbf{x}_i, \mathbf{x}_s)}{n_b B(\mathbf{x}_i)}\right] + 2\log\left[\frac{P_A(\mathbf{x}_s)}{P_A(\mathbf{x}_0)}\right], \tag{5.7}$$

where $\mathbf{x}_0$ is the reported best-fit location of the ANITA event and $\hat{n}_s$ is the value of $n_s$ that maximizes the likelihood. TS is calculated for all $\mathbf{x}_s$, and the maximum value is reported. For this analysis, $P_A$ is a two-dimensional Gaussian assuming the localization uncertainties reported in Table 5.2. As we are not motivated by a specific astrophysical class of objects with characteristic timescales of emission, we consider constant emission over various time windows for each of the ANITA events. This technique is similar to previous IceCube searches for gamma-ray bursts and fast radio bursts [215, 216]. AAE-061228 is excluded from this analysis because it occurred before IceCube had attained a full detector configuration. For AAC-150109 we consider three separate time windows: 10 s, $10^3$ s, and $10^5$ s. During the event time of AAE-141220, IceCube was temporarily not collecting data, due to a run transition that had begun approximately 0.5 seconds before the event and lasting for about one minute. Because of this, we only investigate hypotheses of constant emission over two time windows ($10^3$ s and $10^5$ s), where the period of time from the run transition is not an appreciable portion of our on-time window.

The limits we set on muon neutrinos in the TeV–PeV energy range can constrain generic fluxes of incident tau neutrinos with EeV energies. As discussd earlier, any incident flux with an EeV $\nu_\tau$ component that traverses large Earth chord lengths will result in a secondary flux of lower energy

neutrinos, to which IceCube would be sensitive. We use the same prescription here to constrain a generic point source flux that includes EeV neutrinos.

For any incident flux of neutrinos from the northern sky, $\Phi(E_\nu, t)$, the number of expected detected tau-neutrino–induced muon events at IceCube is given by

$$\begin{aligned}\langle N^{\mu}_{\text{IceCube}} \rangle = &\int dE_\mu \int dE_\tau \int dE_\nu \Phi\left(E_\nu, t\right) P_\tau^{surv}\left(E_\nu\right) \frac{dN_\tau\left(E_\tau\right)}{dE_\tau} \frac{\Gamma_{\tau\to\mu}}{\Gamma_{\text{total}}} \frac{dN_\mu}{dE_\mu}\left(E_\tau, E_\mu\right) A^{\mu}_{eff}\left(E_\mu\right) \Delta T \\ &+ \int dE_\mu \int dE_\tau \int dE'_\nu \int dE_\nu \Phi\left(E_\nu, t\right) P_\nu\left(E_\nu, E'_\nu\right) \frac{dN_\nu}{dE'_\nu}\left(E'_\nu\right) N^p\left(E'_\nu\right) \\ &\times \frac{dN_\tau}{dE_\tau}\left(E'_\nu; E_\tau\right) \frac{\Gamma_{\tau\to\mu}}{\Gamma_{\text{total}}} \frac{dN_\mu}{dE_\mu}\left(E_\tau; E_\mu\right) A^{\mu}_{eff}\left(E_\mu\right) \Delta T, \quad (5.8)\end{aligned}$$

where the first contribution is from emerging tau-leptons that would decay to muons and then pass an IceCube event selection. The second contribution is from the remaining $\nu_\tau$ flux, the majority of which has cascaded down in energy. $N^p(E_\nu)$ is the number of targets effectively seen by an incident neutrino with energy $E_\nu$. The effective area of this event selection to muons incident on the detector is displayed in Figure 5.7. $P_\tau^{surv}\left(E_\nu\right)$ and $P_\nu(E_\nu)$ represent the survival probability of a $\tau$-lepton and $\nu_\tau$, given an incident neutrino energy, respectively. $\Gamma_{\tau\to\mu}/\Gamma_{\text{total}}$ represents the branching ratio for the tau-decay to muon channel, which is approximately 18%.

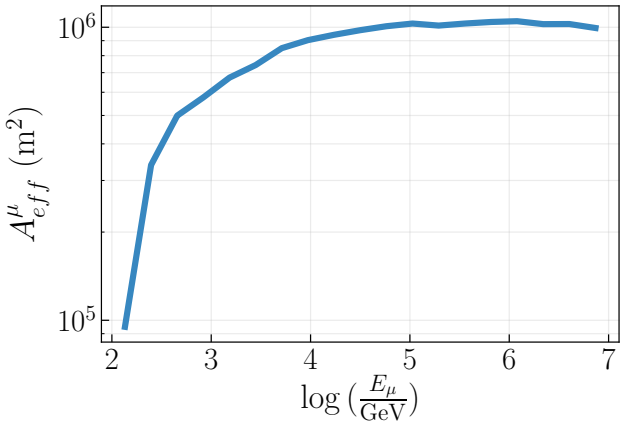



Figure 5.7: Effective area of the IceCube event selection to muons from the northern sky, incident on a volume 1.5 km away from the edge of the detector. $E_\mu$ is the muon energy incident on this volume.

For ANITA, the number of expected events from upgoing taus can be computed using Eq. 5.3. We focus on the non-observation of coincident events in IceCube at $\Delta T = 10^3$ s. A similar procedure can be applied to longer time windows. Qualitatively, it would result in similar limits up to the lifetime of the ANITA flight. For longer emission timescales, limits from IceCube become even more

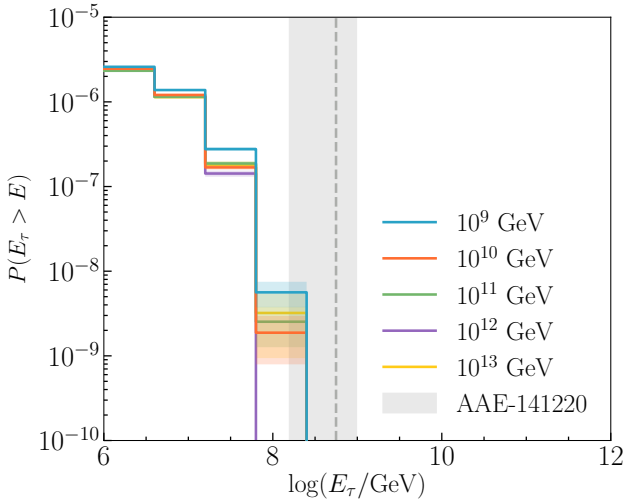



Figure 5.8: Normalized cumulative distributions for Earth-emerging tau-leptons. Colors correspond to the incoming tau-neutrino energy, and the gray band is the 95% containment on the error of the reconstructed shower energy of AAE-141220.

constraining as the implied normalization on the ANITA flux would have to increase to compensate for the fraction of time during which ANITA was not taking data.

We inject fluxes described by delta functions in energy, $\Phi(E_\nu, t) = \Phi_0 \delta(E_\nu - E_0)$, where now the normalization carries units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. As explained previously, the secondary tau neutrinos at or above an EeV are degenerate and therefore these limits apply to any assumed initial flux shape. After propagating these mono-energetic fluxes, we record what fraction of the incident flux results in a detectable signal at ANITA. We repeat this procedure for a variety of injected initial neutrino energies so that we can find the energy that yields the maximum probability of a $\tau$-lepton arriving at ANITA with an energy within the quoted reconstructed energy bounds. We find that the optimal flux for ANITA corresponds to an injected $\nu_\tau$ flux with $E_0 = 1$ EeV. Normalized cumulative distributions from secondary $\tau$-leptons are shown in Figure 5.8 for injected neutrinos at angles corresponding to the best-fit reconstructed direction of AAE-141220.

We next inject a flux of EeV tau neutrinos and find the spectral shape of the secondary $\nu_\tau$ flux that would be incident on IceCube. As we observed zero coincident events in the time window of $10^3$ s around AAE-141220, we calculate the maximum allowed flux normalization at 90% C.L. on the primary flux that would evade this non-observation. The results are displayed in Figure 5.3.

Although IceCube's sensitivity peaks many orders of magnitude below the reconstructed energies of the ANITA events, the limits set on any potential neutrino source that created AAE-141220 are

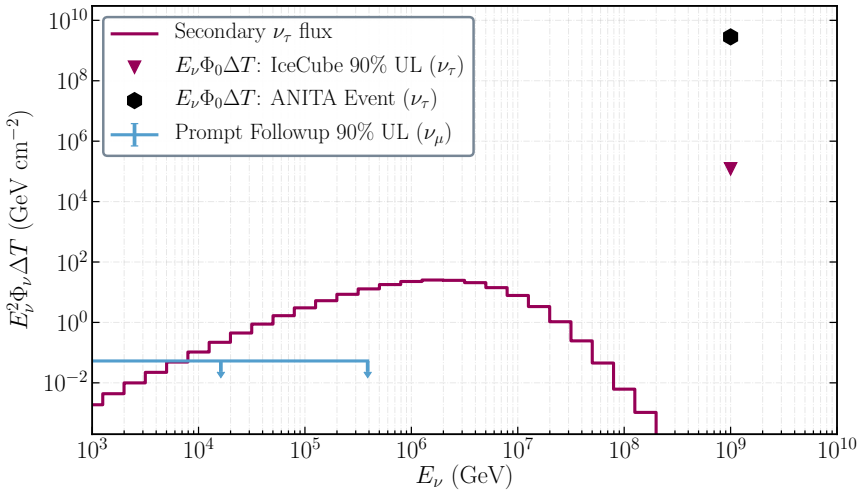



Figure 5.9: Upper limits (90% C.L.) placed by calculating the secondary neutrino flux (purple histogram) from an incident flux of EeV neutrinos assuming constant emission over $10^3$ s and comparing to the non-observation of IceCube events for AAE-141220. The flux implied by the ANITA observations (black) overshoots our upper limit (purple arrow) by many orders of magnitude. For comparison, upper limits on the time-integrated muon-neutrino flux from the prompt analysis are shown in blue. All fluxes are per flavor $\nu + \bar{\nu}$.

more constraining by several orders of magnitude than the implied flux by the ANITA observations. These limits are conservative, and severely constrain any incident spectrum which could produce an observable event at ANITA consistent with the AAE.

# Chapter 6

# Neutrinos and Dark Matter

> "We can, in our conception, join the head of a man to the body of a horse; but it is not in our power to believe that such an animal has ever really existed."
>
> David Hume

The goal of this chapter is to compute constraints on the $\chi\chi \to \nu\bar{\nu}$ annihilation channel and $\chi \to \nu\bar{\nu}$ decay channel from all available extraterrestrial neutrino data. We focus on the two most promising sources of indirect dark matter signal: 1) the dark matter halo of the Milky Way, in which we are deeply embedded, and 2) the full cosmic flux from the sum of all DM halos within our cosmological horizon. We cover a mass range from 1 MeV to $10^{15}$ MeV.

The structure of this chapter is as follows: we begin in Sec. 6.1 with a theoretical overview motivating possible links between the neutrino and dark sectors. In Sec. 6.3 we review the annihilation signal we are constraining, from the Milky Way halo and from the isotropic background of extragalactic halos. In Sec. 6.4, we detail the calculations needed to extend our analysis to velocity-dependent annihilations, namely $p$-wave and $d$-wave processes. Sec. 6.5 introduces the flux of neutrinos from dark matter decay in both the Milky Way halo and extragalactic halos. Sec. 6.6 briefly summarizes the experimental techniques used for neutrino detection in a wide energy range, and describes the statistical methods employed in this work to constrain the neutrino flux from dark matter. Our results are presented in Sec. 6.7, including results from previous analyses that we recast to be consistent with our halo assumptions, wherever possible. We then show the results of varying these assumptions in the range allowed by stellar dynamic observations for the Galactic component and simulation results for the extragalactic one.

## 6.1 Motivation

The Standard Model (SM) of particle physics is the framework that describes matter and its interactions at the most fundamental level. Despite overwhelming success as a predictive theory, observations indicate that the SM is incomplete. Neutrinos have nonzero masses, yet the Higgs mechanism that provides masses for the other SM fermions cannot account for the chiral nature of neutrinos and their interactions unless additional particle content is added to the model. Additionally, overwhelming astrophysical and cosmological evidence points to the existence of a new species of weakly-interacting particles – dark matter (DM) – which accounts for $\sim 85\%$ of the mass budget of the Universe. Local stellar dynamics, galactic rotation curves [218, 219], cluster dynamics [220, 221], and gravitational lensing *e.g.*[222, 223] all point to mass-to-light ratios in astrophysical objects that are much higher than could be accounted for by stellar objects and gas; for a historical overview see [224]. Measured primordial abundances of light elements tell us that Big Bang Nucleosynthesis requires a total baryon density[1] of only $\Omega_b \sim 0.05$, while the Cosmic Microwave Background (CMB) and other probes of large scale structure require the *total* density of nonrelativistic matter to be $\Omega_m \sim 0.3$. [2] A leading hypothesis for the nature of this new nonbaryonic component is the Weakly Interacting Massive Particle (WIMP). The relic abundance of WIMPs today was set as they fell out of equilibrium with the high-temperature plasma of the early Universe. When the temperature, $T$, fell below the DM mass, $m_\chi$[3], the equilibrium distribution became Botlzmann-suppressed, namely $\sim \exp(-m_\chi/T)$. At some point, the expansion rate $H(t)$, became larger than the thermally-averaged self-annihilation rate, preventing further annihilation into SM particles, *freezing-out* the relative density of DM particles. The WIMP scenario predicts the observed relic abundance of DM for values of the thermally-averaged self-annihilation rate $\langle \sigma v \rangle \simeq 3 \times 10^{-26}$ cm$^3$s$^{-1}$ regardless of the final annihilation channel.

Thermal production of weakly-interacting DM in the Early Universe implies possible ongoing self-annihilation to SM particles wherever DM exists today. Significant effort has gone into searches for indirect signatures of DM annihilation. Annihilation to most SM states yields an abundance of photons with energies on the order of 10% of the DM mass, such that some of the strongest

[1] By *baryonic* we refer here to stable nonrelativistic matter made of SM particles including neutrons, protons and electrons.

[2] More precisely, the baryon density is inferred to be $\Omega_b h^2 = 0.0224 \pm 0.0001$ and the (cold) DM density is $\Omega_c h^2 = 0.120 \pm 0.001$ [225], where $\Omega_i$ is the ratio of the density of component $i$ to the critical density, and the Hubble constant is $H_0 \equiv h100$ km s$^{-1}$Mpc$^{-1}$.

[3] We work in natural units where $c = \hbar = k_B = 1$.

constraints on particle DM models are from the (non) observation of X- and gamma-ray signals from the Milky Way and its satellite galaxies; see *e.g.* [226, 227]. Cosmic-ray signatures provide similarly constraining limits, reports of excesses notwithstanding; see [228] and references therein.

As X- and gamma-ray experiments rely, by design, on electromagnetic signals, they are optimal for probing links between the dark sector and quarks or charged leptons, although neutrino detectors can still play a role in these searches [229].

There is a distinct possibility that the principal portal through which the DM interacts with the SM is via the neutrino sector [18]. This naturally arises in "scotogenic" models, in which neutrino mass generation occurs through interactions with the dark sector [19, 20, 21, 22, 23, 24, 25, 26]. These models introduce heavy neutrino states, sometimes called dark neutrinos, which could also provide a possible explanation of the MiniBooNE anomaly [27, 28, 29, 30]. "Secret" neutrino interactions with dark matter have recently become a very active field of investigation, where constraints have been obtained using high-energy astrophysical neutrinos [230, 231, 232, 233, 234, 235, 236, 237, 238, 239], solar neutrinos [240], cosmology [241, 242], accelerator neutrino experiments [243, 244, 245] and colliders [246]. The past two decades have seen extraordinary progress in the field of extraterrestrial neutrino detection. Observations span a wide energy range, from the MeV *pp* solar neutrino flux [77] to the PeV ($10^6$ GeV) high-energy astrophysical neutrinos [121, 122, 176]. Furthermore, neutrino flux limits exist all the way up to $\sim$ ZeV ($10^{12}$ GeV) [163, 167]. With these observations, a multitude of experimental constraints have been derived on the dark matter annihilation cross section to neutrino pairs and decay lifetime, either by experimental collaborations themselves or by independent authors recasting results of previous searches.

Neutrinos are the most weakly interacting stable particles in the SM and, consequently, the hardest to detect. In the context of indirect detection, this implies that models where DM annihilates predominantly to neutrinos are difficult to rule out. This makes the study of neutrinos as a final state particle particularly interesting as, so far, all direct and indirect searches for the footprints of DM–SM interactions have come up empty [247, 248].

## 6.2 Dark Matter Density Profile

To parametrize the galactic DM halo, we use a generalized Navarro-Frenk-White (NFW) profile, which is given by

$$\rho_\chi(r) = \rho_s \frac{2^{3-\gamma}}{\left(\frac{r}{r_s}\right)^\gamma \left(1+\frac{r}{r_s}\right)^{3-\gamma}}. \tag{6.1}$$

The galactocentric distance is

$$r = \sqrt{R_0^2 - 2\,x\,R_0\,\cos\psi + x^2}, \tag{6.2}$$

where $\psi$ is the angle between the Galactic center (GC) and the line of sight, and $R_0$ is the distance from the Sun to the GC. We take the Sun to be located $R_0 = 8.127$ kpc from the GC, as determined by recent measurements of the four-telescope interferometric beam-combiner instrument GRAVITY [249]. We use DM halo parameters compatible with the best-fit values of [250], *i.e.*: a local density[4] of $\rho_0 = 0.4$ GeV cm$^{-3}$, a slope parameter $\gamma = 1.2$, and a density $\rho_s$ at scale radius $r_s = 20$ kpc.

Annihilation specifically is sensitive to the clustering of DM, rather than the overall number of particles as it requires the collision of two DM particles. Therefore extragalactic neutrino fluxes from DM annihilation are enhanced by clustering. Accounting for this enhancement requires integrating the described DM density profile over redshift, accounting for the the mass distribution of halos. This is often called the "boost" factor $G(z)$, and is given by

$$G(z) = \frac{1}{\Omega_{DM,0}^2 \rho_c^2} \frac{1}{(1+z)^6} \int dM \frac{dn(M,z)}{dM} \int dr 4\pi r^2 \rho_\chi^2(r). \tag{6.3}$$

The first integral is over halo masses $M$ whose distribution is specified by the halo mass function (HMF), $dn/dM$, while the second integral is over the halo overdensities themselves. We model the latter as self-similar NFW profiles whose densities and radii are specified by a concentration parameter uniquely determined by their mass and redshift. The parametrization that we employ is based on fits to the MultiDark/BigBolshoi [251] simulations and can be found in Appendix B of [252].

Two uncertainties arise from the integral over $M$. First is the choice of integration limits, specifically the lower limit, $M_{min}$. This is because smaller halos are more concentrated, thus contributing more to the injected neutrino energy. This means that choosing arbitrarily low-minimum halo masses results in unrealistic limits. It is common in the literature to use $M_{min} = 10^{-6} M_\odot$ as a benchmark,

[4] It is customary to specify $\rho_0 \equiv \rho_\chi(R_0)$ rather than $\rho_s$, as the former can be more directly measured. The two are related by inverting Eq. (6.1).

although there is no data-driven motivation for this choice. $M_{min}$ is not well-constrained, and will ultimately depend on model details [253, 254]. Therefore, in this work we pick $M_{min} = 10^{-3} M_\odot$ as a conservative limit choice. In section 6.7, we show the effect of varying $M_{min}$ down to $10^{-9} M_\odot$. The other uncertainty arises from the choice of HMF, $dn/dM$, parametrization. We use the results of the N-body simulation by [255], as parametrized in [252, 256]. Several other HMF parametrizations are tested, and the uncertainties due to choice of HMF are quantified in Sec. 6.7.

## 6.3 Annihilation to neutrinos

The limits derived on the DM annihilation to neutrinos can be interpreted as an upper bound on the total DM annihilation cross section to SM particles [257, 258], since the latter is larger by definition. From a particle physics point of view, the direct annihilation of DM to neutrinos at tree level requires the addition of a neutrino-DM term to the SM Lagrangian that couples them. Since neutrinos belong to an $SU(2)$ doublet, naïve SM gauge invariance implies that coupling neutrinos with DM would also induce an interaction between the DM and the charged leptons, mediated, *e.g.*, by a new $Z$-like particle. Such interactions are highly constrained, as they lead to production of dijet or dilepton signatures observable at colliders (see *e.g.* [259, 260]), fixed target experiments [261], and direct detection experiments (see *e.g.* [262] and references therein).

Nevertheless, there exist viable models in which the DM phenomenology is dominated by its interactions with neutrinos [18]. Coupling only to the heavier lepton generations can strongly mitigate bounds from electron interactions, *e.g.* by introducing a $U(1)_{L_\mu - L_\tau}$ symmetry [263, 264]. A more elegant option allows the DM to interact with a sterile neutrino that then mixes with the active neutrinos, leading to direct annihilations of DM to neutrinos if the mass of the sterile neutrino is larger than the DM mass [265, 30]. If the sterile-light mixing is sizable, DM–neutrino interactions will provide the best window to understand such DM models. A comprehensive review of these scenarios can be found in [18].

Finally, we are considering direct annihilation to neutrinos without including electroweak (EW) corrections, which severely complicate the spectral shape computations. These are important at energies above the electroweak scale, and will have two main consequences: 1) the peak of the spectrum will be slightly broadened, and 2) A lower-energy continuum will be produced. Given the typical energy resolution $\gtrsim 10\%$ [180] for high-energy neutrino detectors, the former effect is not likely to be important. The second effect could potentially lead to stronger bounds from the

additional flux at lower energies. A detailed computation of this effect up to ultra-high-energies has only recently been performed [266]; as these were not available at the time of this analysis we do not include these here. At sub-TeV energies, these corrections are accurately implemented in numerical codes such as PYTHIA [267, 268]; a comparison between our limits and ones derived using these additional corrections show very little difference [see e.g. [269]].

A more important consequence is the presence of gamma radiation from the decay of EW products, which can potentially provide complementary constraints to dedicated neutrino-line searches [270]. Using these secondary products, current constraints on the thermally averaged annihilation cross section to neutrinos from Fermi-LAT and HESS hover around $10^{-23}$ cm$^3$s$^{-1}$ in the 300 GeV to 3 TeV mass range [271]. These gamma-ray based constraints are at the same level as current bounds from ANTARES [272], but are expected to be improved by the next generation gamma-ray experiments such as the Cherenkov Telescope Array (CTA) [271]. We will provide an example using these projections for CTA in Sec. 6.7, noting that this only includes prompt gamma rays. Inverse-Compton scattering of primary electrons and positrons with interstellar photons will strengthen the sensitivity of gamma-ray searches. This effect has been studied for DM decay searches, but not for annihilation $\chi\chi \to \nu\bar{\nu}$ [273, 274, 275].

### Galactic contribution

We begin by setting limits on DM annihilation to neutrino pairs in the Milky Way (MW) dark matter halo. The expected flux per flavor of neutrinos plus antineutrinos at Earth, assuming equal flavor composition[5], is given by

$$\frac{d\Phi_{\nu+\bar{\nu}}}{dE_\nu} = \frac{1}{4\pi}\frac{\langle\sigma v\rangle}{\kappa m_\chi^2}\frac{1}{3}\frac{dN_\nu}{dE_\nu}J(\Omega), \tag{6.4}$$

where $\kappa$ is 2 for Majorana DM and 4 for Dirac DM, $m_\chi$ is the DM mass, and $\langle\sigma v\rangle$ is the thermally averaged self-annihilation cross section into all neutrino flavors. Going forward we set $\kappa = 2$ (Majorana DM). The spectrum in the case of annihilation to two neutrinos is simply $dN_\nu/dE_\nu = 2\delta(1 - E/m_\chi)m_\chi/E^2$. $J(\Omega)$ is a three-dimensional integral over the target solid angle in the sky, $d\Omega$, and the distance $dx$ along the line of sight (l.o.s.) of the DM density $\rho_\chi$, namely

$$J \equiv \int d\Omega \int_{\text{l.o.s.}} \rho_\chi^2(x)dx. \tag{6.5}$$

[5]If the flavor composition at the source is not democratic, neutrino oscillation will yield a flavor composition at Earth that is close, but not equal to $(\nu_e : \nu_\mu : \nu_\tau) = (1 : 1 : 1)$. Annihilation to $\nu_e$ only will give $\sim (0.55 : 0.25 : 0.2)$; to $\nu_\mu$: $\sim (0.25 : 0.36 : 0.38)$ and $\nu_\tau$ yields $\sim (0.19 : 0.38 : 0.43)$.

It is referred to as the $J$-factor and has units of $\mathrm{GeV}^2\,\mathrm{cm}^{-5}\,\mathrm{sr}$.[6] In practice, the upper limit of integration can be set at

$$x_{\max} = \sqrt{R_{\mathrm{halo}}^2 - \sin^2\psi R_0^2} + R_0\cos\psi, \tag{6.6}$$

for some maximum halo radius $R_{\mathrm{halo}}$. Details of the dark matter profile and the values of the variables used in the integration are explained in Sec. 6.2.

The resulting $J$-factors for $s$, $p$, and $d$-wave annihilation are shown in Tbl. 6.1; the latter cases will be discussed in Sec. 6.4. Some experiments, such as ANITA, AUGER, and GRAND, are only sensitive to a certain region of the sky. In these cases, the corresponding $J$-factors must be re-computed by converting their respective sensitivity from elevation/azimuth to galactic coordinates, and integrating over the resulting region. A value of the $J$-factor is not given for some experiments, where the flux cannot be factored out as in Eq. (6.4). This could be due *e.g.* to an energy-dependent acceptance. These are also shown in Tbl. 6.1. When the exposure is not a simple declination window, we provide the reference to where it can be obtained. Recent works [276, 277, 278, 250] have constrained the halo shape and density parameters, using observations of stellar dynamics in the MW. In Sec. 6.7, we illustrate the effect on the dark matter limits obtained in this work when varying these parameters within those constraints.

### Extragalactic contribution

In addition to DM annihilation in the MW, annihilation of extragalactic dark matter integrated over all redshifts should provide a diffuse isotropic neutrino signal [257]. As in the search for extragalactic background light, there are two contributions to this isotropic flux: 1) a "background" flux from the diffuse (non-collapsed) distribution of DM, whose rate grows with redshift as $\Omega_{DM}^2 \sim (1+z)^6$, and 2) a late-time contribution from the large overdensities in galactic halos.

In this case, the expected flux of neutrinos plus antineutrinos per flavor at Earth from DM annihilation is given by

$$\frac{d\Phi_{\nu+\bar{\nu}}}{dE_\nu} = \frac{1}{4\pi}\frac{\Omega_{DM}^2\rho_c^2\langle\sigma v\rangle}{\kappa m_x^2}\frac{1}{3}\int_0^{z_{up}} dz \frac{(1+G(z))\,(1+z)^3}{H(z)}\frac{dN_{\nu+\bar{\nu}}\,(E')}{dE'}, \tag{6.7}$$

where $H(z) = H_0\left[(1+z)^3\Omega_m + (1+z)^4\Omega_r + \Omega_\Lambda\right]^{1/2}$ is the time-dependent Hubble parameter, $\rho_c$ is the critical density of the Universe, $\Omega_m$, $\Omega_r$, and $\Omega_\Lambda$ are respectively the fraction of $\rho_c$ made up of

[6] Another equivalent convention used in the literature is to report the dimensionless quantity $\mathcal{J} = J/\Delta\Omega R_0\rho_0^2$ [258].

| Experiment | Exposure | $J_s/10^{23}$ | $J_p/10^{17}$ | $J_d/10^{11}$ |
|---|---|---|---|---|
| ♡ All-sky | All-sky | 2.3 | 2.2 | 3.6 |
| ♡ GRAND | Fig. 24 of [279] | 0.28 | 0.28 | 0.46 |
| ♡ ANITA | dec $= [1.5°, 4°]$ | 0.018 | 0.018 | 0.028 |
| CTA [271] | Galactic Center [271] | 0.074 | 0.12 | 0.16 |
| ♡ TAMBO | Fig. 3 & 4 of [140] | 0.0009 | – | – |
| ♡ Auger | zenith $= [90°, 95°]$ | 0.10 | – | – |
| | zenith $= [75°, 90°]$ | 0.28 | | |
| | zenith $= [60°, 75°]$ | 0.27 | | |
| ♡ P-ONE | $\cos(\text{zenith}) = [-1, -0.5]$ | 0.87 | 0.85 | 1.4 |
| | $\cos(\text{zenith}) = [-0.5, 0.5]$ | 1.2 | 1.2 | 2.0 |
| | $\cos(\text{zenith}) = [0.5, 1]$ | 0.13 | 0.12 | 0.18 |

Table 6.1: J-factors used to constrain DM annihilation for different experiments and their associated halo parameters. $J$-factors, given in units of $\text{GeV}^2\ \text{cm}^{-5}\ \text{sr}$, are computed according to Eq. (6.5). We use these to find the expected neutrino flux as described in Eq. (6.4). Each row corresponds to a different experimental setup given its angular exposure. The first column names the experiment; the second column summarizes their angular acceptance; and the last three columns give the $s$-wave, $p$-wave, and $d$-wave $J$-factors, respectively. The hearts, ♡, indicate new results given in this work.

matter, radiation and dark energy. While the upper limit on redshift, $z_{up}$, can in principle be as high as the neutrino decoupling time at $T \sim$ MeV, neutrinos produced at that epoch are redshifted to the point of being invisible to existing detectors. $dN_\nu\left(E'\right)/dE$ is the neutrino spectrum at the detector, where $E'$ ($E$) is the energy at the source (detector). The spectrum is related to the source production spectrum via a Jacobian transformation to take cosmological redshift into account, namely

$$\frac{dN_{\nu+\bar{\nu}}\left(E'\right)}{dE'} = 2\frac{m_\chi}{E'^2}\delta\left(\frac{m_\chi}{E'} - 1\right) = \frac{2}{E}\delta\left[z - \left(\frac{m_\chi}{E} - 1\right)\right]. \tag{6.8}$$

In Eq. (6.7), $\langle\sigma v\rangle$ is the thermally averaged cross section. The first part of the factor $1+G(z)$ in the integrand of Eq. (6.7) represents the isotropic background DM contribution, while $G(z)$ is the halo boost factor at redshift $z$. It accounts for the enhancement to the annihilation rate in DM clusters and their evolution with redshift, and is defined in Sec. 6.2.

The expected spectrum of DM annihilation to two neutrinos from cosmological sources is shown in Fig. 6.1, for different DM masses. These are overlaid on the Super-Kamiokande (SK) [280] and IceCube [281, 179] unfolded atmospheric $\nu_e$ and $\nu_\mu$ fluxes as well as the isotropic astrophysical flux [178].

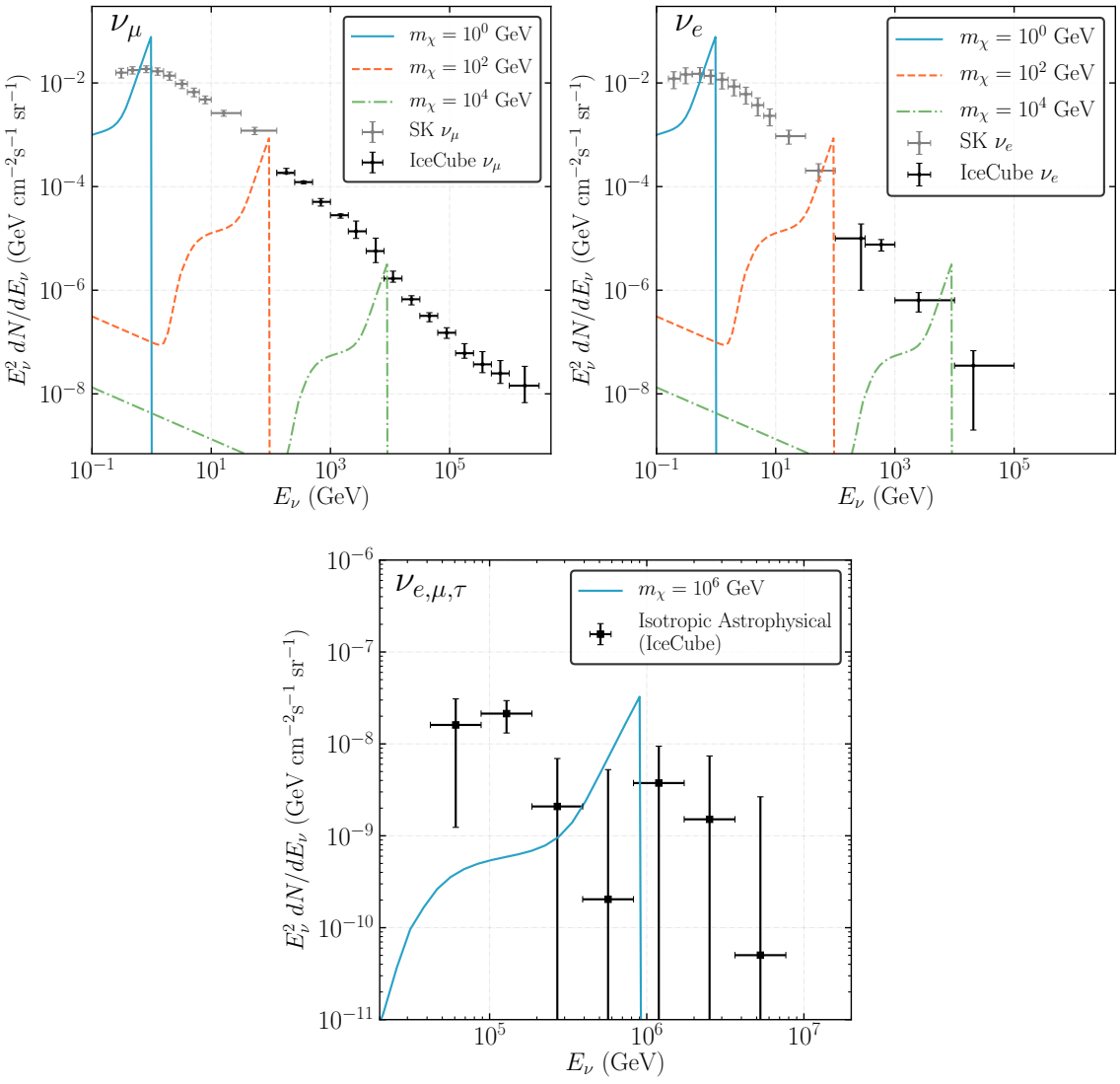



Figure 6.1: Examples of neutrino fluxes produced by dark matter annihilation overlayed on the observed neutrino distributions. Expected flux of neutrinos from extragalactic dark matter annihilation as a function of energy, shown for several dark matter masses. Fluxes are computed using the value of the cross section corresponding to the 90% C.L. limit derived in this work. Here, the extragalactic dark matter annihilation fluxes are compared to the unfolded atmospheric fluxes from both Super-Kamiokande [280] and IceCube [281, 179]. Top left is the $\nu_\mu$ channel; top right is the $\nu_e$ channel; the bottom shows a comparison to IceCube's measured per-flavor isotropic Astrophysical flux using 7.5 years of Starting Events [178].

## 6.4 Velocity-dependent annihilation

Certain matrix element vertex structures lead to a suppression of the constant ($s$-wave) part of the self-annihilation cross section. Expanding in powers of $v/c$, the dominant term may be $p$-wave

($\propto v^2$) or $d$-wave ($\propto v^4$) in the nonrelativistic limit. The DM velocity distribution depends on the kinematic details of the structure in which it is bound, as well as its distance from the center of that distribution. Assuming a normalized Maxwellian distribution, $f(v,r)$, with dispersion $v_0(r)$, the annihilation rate will be proportional to

$$\langle v^n \rangle = \int d^3v f(v,r) v^n. \tag{6.9}$$

For $p$- and $d$-wave, this respectively yields

$$\langle v^2 \rangle = 3v_0^2(r), \tag{6.10}$$

$$\langle v^4 \rangle = 15v_0^4(r). \tag{6.11}$$

We obtain the dispersion velocity, $v_0$, by solving the spherical Jeans equation, assuming isotropy. This is given by

$$\frac{d(\rho(r)v_0^2(r))}{dr} = -\rho(r)\frac{d\phi(r)}{dr}, \tag{6.12}$$

where $\phi(r)$ is the total gravitational potential at radius $r$. For Galactic constraints, we include not only the contribution of the DM halo to $\phi(r)$, but also follow [282] and include a parametrization of the MW bulge and disk potentials to account for their masses. These are given by

$$\phi(r)_{\rm bulge} = -\frac{G_N M_b}{r+c_b}, \tag{6.13}$$

$$\phi(r)_{\rm disk} = -\frac{G_N M_d}{r}\left(1-e^{-r/c_d}\right), \tag{6.14}$$

where $G_N$ is Newton's gravitational constant, $M_b = 1.5\times10^{10}M_\odot$, and $c_b = 0.6$ kpc are the bulge mass and scale radius, while $M_d = 7\times10^{10}M_\odot$ and $c_d = 4$ kpc are the disk mass and scale radius [282]. Galactic $J$-factors can then be reevaluated via

$$J_{v^n} = \int d\Omega \int_{\rm l.o.s.} \frac{\langle v^n(r)\rangle}{c^n}\rho_\chi^2(r)dx. \tag{6.15}$$

In the case of our extragalactic analysis, we only include the potential from the DM halos themselves. This is conservative, in that the addition of the uncertain baryonic contributions would only strengthen our constraints. In a similar manner to the Galactic case, Eqs. (6.7) and (6.3) must be modified to include the dependence on $\langle v^n\rangle(r)$. As long as the annihilation remains a two-to-two process (unlike scenarios in e.g.[283]), Eq. (6.7) becomes:

$$\frac{d\Phi_\nu}{dE_\nu} = \frac{c}{4\pi}\frac{\Omega_{DM}^2\rho_c^2\langle\sigma v\rangle}{2m_x^2}\int_0^{z_{up}} dz \frac{\left(\left[\frac{1+z}{1+z_{\rm KD}}\right]^n + G_n(z)\right)(1+z)^3}{H(z)}\frac{dN_\nu(E')}{dE}, \tag{6.16}$$

where the redshift $z_{\rm KD}$ is related to the temperature at kinetic decoupling $T_{\rm KD}$ and the temperature of the CMB today $T_{\rm CMB,0}$ via $1+z_{\rm KD} = T_{\rm KD}/T_{\rm CMB,0} \simeq 4.2\times10^9\,(T_{\rm KD}/{\rm MeV})$ [256]. [253] obtained a temperature of kinetic decoupling:

$$T_{\rm KD} \simeq 2.02\,{\rm MeV}\,\left(\frac{m_\chi}{\rm GeV}\right)^{3/4}. \tag{6.17}$$

In general, kinetic decoupling occurs later than chemical freeze-out and depends on the number of relativistic degrees of freedom $g_\star(T_{\rm KD})$. At redshifts where the annihilation products are still measurable by earth-based detectors, the factor of $((1+z)/(1+z_{\rm KD}))^n$ still leads to a strong enough suppression that it will always be subdominant to the halo contribution proportional to $G_n(z)$. The exact value of $T_{\rm KD}$ in Eq. (6.17) is thus inconsequential. Eq. (6.3) including velocity dependence is rewritten as follows:

$$G_n(z) = \frac{1}{\Omega_{DM,0}^2\rho_c^2}\frac{1}{(1+z)^6}\int dM\frac{dn(M,z)}{dM}\int dr 4\pi r^2\frac{\langle v^n(r)\rangle}{c^n}\rho_\chi^2(r), \tag{6.18}$$

where we have used the same HMF as in the velocity-independent case, with the addition of the velocity dispersion, $\langle v^n(r)\rangle$, in the rightmost integral. [256] provides the detailed method of solving the Jeans equation to compute $\langle v^n(r)\rangle$ as a function of the DM halo concentration. For convenience, we provide the following function for the $p-$ and $d-$wave cases:

$$\ln(G_n) \simeq \sum_i c_i\,\alpha^i, \tag{6.19}$$

where $c_i$ are the coefficients provided in Tbl. 6.2, and $\alpha \equiv \ln(z)$. This parametrization is valid down to redshifts $\gtrsim 10^{-3}$.

| | $p-$wave | $d-$wave |
|---|---|---|
| $c_0$ | $-7.004$ | $-19.88$ |
| $c_1$ | $-1.821$ | $-2.493$ |
| $c_2$ | $-0.5793$ | $-0.804$ |
| $c_3$ | $-0.09559$ | $-0.1636$ |
| $c_4$ | $-0.006148$ | $-0.02101$ |
| $c_5$ | 0 | $-0.001181$ |

Table 6.2: Coefficients of the polynomial fit to velocity dependent halo boost factors. The coefficients corresponding to Eq. (6.19), which is a parametrization to the numerical solution of Eq. (6.18).

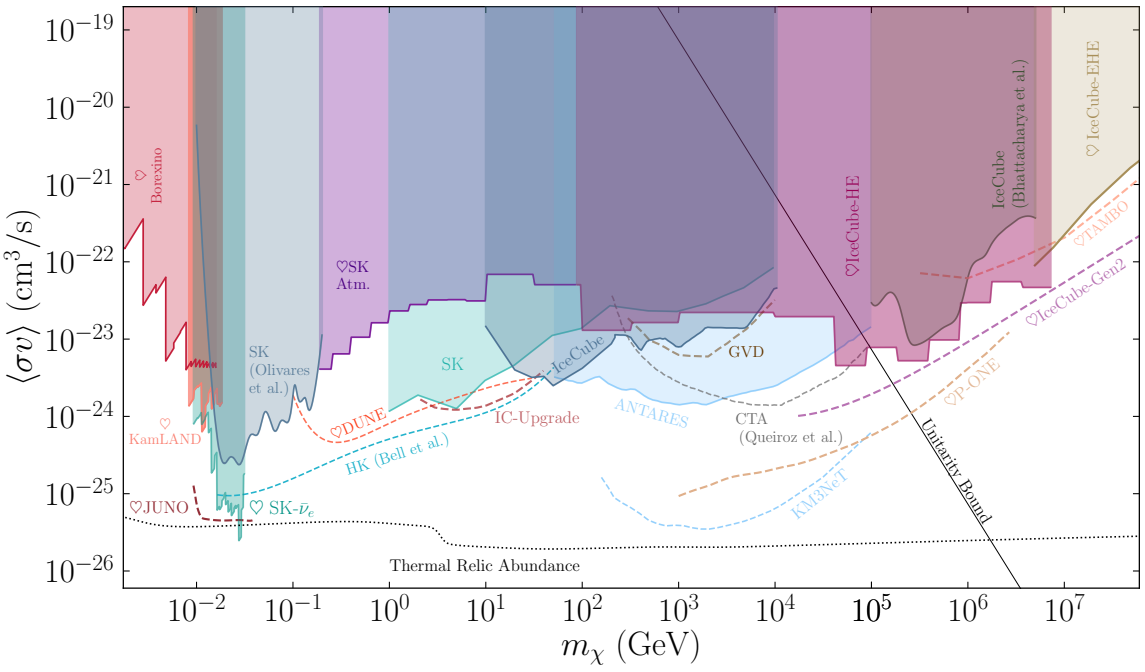



Figure 6.2: The landscape of dark matter annihilation into neutrinos up to $10^8$ GeV. We show results from this work, as well as previously published limits. Data and corresponding references are detailed in Sec. 6.7. Solid and dashed lines represent 90% C.L. limits and sensitivities, respectively. Projected sensitivities assume five years of data taking for neutrino experiments and 100 hours of observation for CTA. The dotted line corresponds to the value required to explain the observed abundance via thermal freeze-out. The straight diagonal line, labeled as "Unitarity Bound," gives the maximum allowed cross section for a non-composite DM particle. These results assume 100% of the dark matter is composed of a given Majorana particle. If instead only a fraction, $f$, is considered these results should be multiplied by $1/f^2$. In the case of Dirac DM, limits would be scaled up by a factor of two. The heart symbols (♡) indicate new results obtained in this work. See Fig. 6.5 for constraints and projections up to $10^{11}$ GeV.

## 6.5 Decay to neutrinos

Here we focus on direct decay to neutrino pairs $\chi \to \nu\bar{\nu}$. At masses greater than $\sim$ TeV, electroweak corrections can "decloak" the dark matter, producing high-energy photons detectable at current and next-generation gamma-ray observatories. We will find, however, that current and future neutrino telescopes retain superior sensitivity across nearly the entire range of masses that we consider. This is due to three factors: the loop suppression of the gamma-ray production rate, the growth of the electroweak cross section with energy, and the loss of high-energy gamma rays to interactions in the interstellar and intergalactic medium.

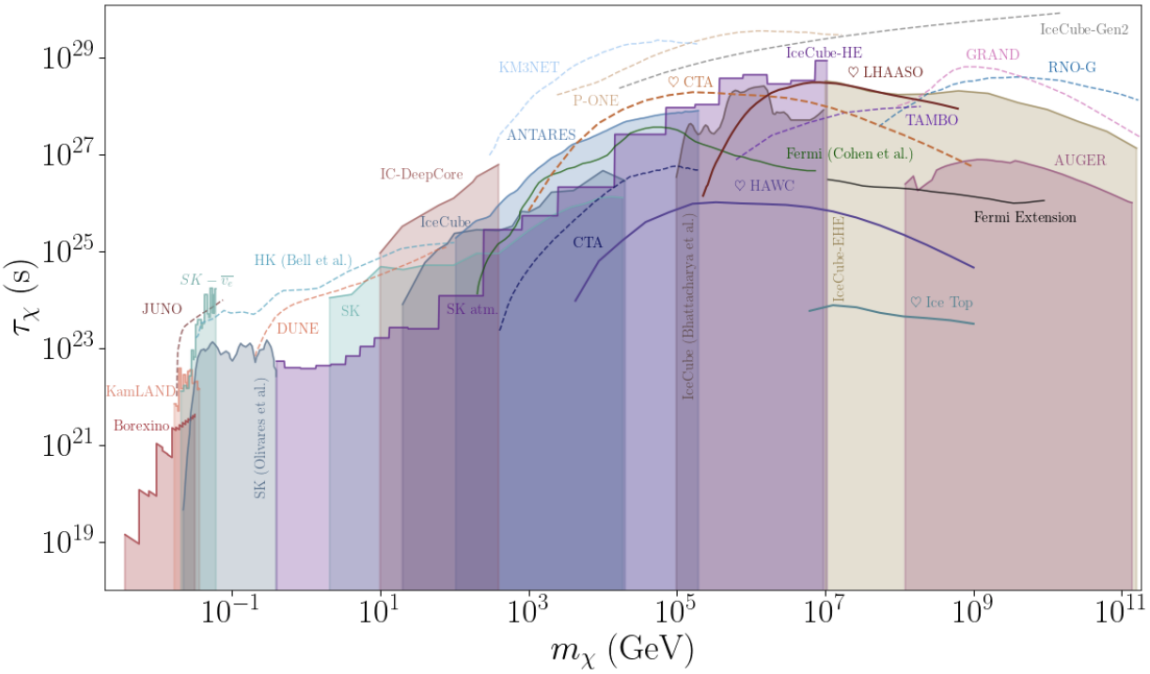



Figure 6.3: Limits on the lifetime of dark matter decay to neutrino pairs $\chi \to \bar{\nu}\nu$. DM masses span a few MeV to $10^{14}$ MeV. Limits are especially strong for heavy DM where neutrino observatories such as IceCube and ANTARES dominate. Future projections are shown as dashed lines.

### Galactic contribution

The flux of neutrinos and antineutrinos at the detector due to dark matter decay is given by

$$\frac{d\Phi_{\nu+\bar{\nu}}}{dE_\nu} = \frac{1}{4\pi}\frac{\Gamma_\chi}{m_\chi}\frac{1}{3}\frac{dN_\nu}{dE_\nu}D(\Omega), \tag{6.20}$$

where $\Gamma_\chi = \tau_\chi^{-1}$ is the decay rate of dark matter and $m_\chi$ is the dark matter mass. A single decay produces two neutrinos each with half the DM mass such that the produced spectrum is

$$\frac{dN_\nu}{dE_\nu} = 2\delta(E - m_\chi/2). \tag{6.21}$$

The amount of dark matter in the field of view of a given experiment is captured by the factor $D(\Omega)$. Sometimes referred to as the $D$-factor, it is given by the line of sight (l.o.s.) integral of dark matter density over the target angle in the sky such that

$$D \equiv \int d\Omega \int_{\text{l.o.s.}} \rho_\chi(x)dx. \tag{6.22}$$

where $x$ is the l.o.s. distance from Earth and $\rho_\chi$ is the dark matter density, the details of which are described in Sec. 6.2. The $D$-factor used for all experiments are shown in Table 6.3. A lot of neutrino

experiments are sensitive to the entire sky, those are grouped together. For experiments that are only sensitive to a some portion of the sky, the angular bounds of integration used in determining the $D$-factor are calculated by converting their respective sensitivity from elevation/azimuth to galactic coordinates.

| Experiment | Exposure | $D/10^{23}$ |
|---|---|---|
| ♡ All-sky | All-sky | 2.65 |
| ♡ GRAND | Fig. 24 of [279] | 0.298 |
| ♡ ANITA | dec $= [1.5°, 4°]$ | 0.052 |
| ♡ CTA | Galactic Center [271] | 0.003 |
| ♡ TAMBO | Fig. 3 & 4 of [140] | 0.001 |
| ♡ Auger | zenith $= [90°, 95°]$ | 0.11 |
| | zenith $= [75°, 90°]$ | 0.35 |
| | zenith $= [60°, 75°]$ | 0.33 |
| ♡ P-ONE | $\cos(\text{zenith}) = [-1, -0.5]$ | 0.83 |
| | $\cos(\text{zenith}) = [-0.5, 0.5]$ | 1.35 |
| | $\cos(\text{zenith}) = [0.5, 1]$ | 0.47 |

Table 6.3: D-factors for different experiments used to constrain dark matter decay and their associated halo parameters. $J$-factors, given in units of GeV cm$^{-2}$ sr, are computed according to Eq. (6.5). We use these to find the expected neutrino flux as described in Eq. (6.4). Each row corresponds to a different experimental setup given its angular exposure. The first column names the experiment; the second column summarizes their angular acceptance; and the last three columns give the $s$-wave, $p$-wave, and $d$-wave $J$-factors, respectively. The hearts, ♡, indicate new results given in this work.

### Extragalactic contribution

DM decay outside the Milky Way also contributes to the observed neutrino flux at Earth. Extragalactic decays produce a diffuse isotropic signal originating from all redshifts [257]. The total expected flux is given by

$$\frac{d\Phi_{\nu+\bar{\nu}}}{dE_\nu} = \frac{1}{4\pi}\frac{\Omega_{DM}\rho_c}{m_x\tau_\chi}\frac{1}{3}\int_0^{z_{up}}\frac{dz}{H(z)}\frac{dN_{\nu+\bar{\nu}}(E')}{dE'}, \tag{6.23}$$

where $Omega_{DM}$ is the dark matter fractional energy density, $rho_c$ is the critical density today, $H(z)$ is the Hubble parameter, and $E' = (1+z)E$ is the starting energy of a neutrino accounting for the expansion of the Universe.

## 6.6 Experimental methods

In this section we will briefly review the different methodologies and technologies used for neutrino detection relevant for the discussion of the experimental results discussed in this chapter. The results presented in Sec. 6.7 rely on our understanding of the backgrounds in the region of interest. Depending on whether the background flux is known, limits can be either background-agnostic or background informed. Moreover, the limits highly depend on the systematics that govern neutrino detection, for instance the energy resolution and flavor identification capability. Below, we first outline the statistical framework for limit-setting, before describing detector physics used over energy ranges considered here, from a few MeV up to $10^{12}$ GeV and beyond.

### Statistical Methods

To contextualize the variety of experimental capabilities, we will first outline the principal statistical treatments used to infer the properties of the flux of neutrinos from dark matter annihilation or decay. We will explain them in increasing order of complexity and strength.

#### Background-agnostic methods

In this method we use the observed data and the detector signal efficiency to constrain the flux of neutrinos from DM. This method can inform us of the maximum allowed flux, but, by construction, it cannot be used to claim the observation of dark matter. This technique is predicated on comparing the observed and expected number of events in a given bin, by means of the following likelihood function:

$$\mathcal{L}(\mu) = \begin{cases} \mathcal{P}(d|\mu) & (d < \mu), \\ 1 & (d \geq \mu), \end{cases} \tag{6.24}$$

for which the likelihood is less than one only if the predicted number of events $\mu$ is larger than the recorded data, $d$. The probability distribution $\mathcal{P}$ could be a Poisson or Gaussian distribution depending on the sample size. Using this likelihood one can construct one-sided confidence upper limits on $\mu$ and, in turn, on the dark matter model parameters. The strength of this method is determined by experiment exposure, signal efficiency, and the amplitude of unmodeled backgrounds; these determine the statistical uncertainty and the phase-space over which the bins are defined. In the case of dark matter, one would ideally bin the events in: energy, direction, and morphology; but

often this is either not done due to decreasing statistical power, insufficient Monte Carlo certainty, or increasing difficulty in modeling the systematic uncertainties.

Here we take advantage of this approach in a number of experimental settings. As examples, we compare the Super-Kamiokande unfolded neutrino energy distribution [280] to the dark matter expectation using this technique and perform a similar comparison to the IceCube PeV astrophysical neutrino segmented fit. We also use this technique when experiments have not seen neutrino events and upper limits are reported, such as the Pierre Auger Observatory's limit on neutrino flux at very high energies.

**Background-informed methods**

A higher statistical power can be achieved by simultaneously modeling the signal – the event rate due to dark matter – and background – any other contribution to the observed rate. This requires signal and background efficiencies, as well as a model for the background distribution over each observable. A prototypical likelihood function is:

$$\mathcal{L}(\theta, \eta) = \mathcal{P}(d|\mu_s(\theta, \eta) + \mu_b(\eta))\Pi(\eta), \tag{6.25}$$

where $\mu_s(\theta, \eta)$ and $\mu_b(\eta)$ are the expected signal and background counts respectively, $d$ represents the observed counts, and $\theta$ and $\eta$ are the dark matter parameters and nuisance parameters, respectively. The latter parameters incorporate the effect of the systematic uncertainties in the signal and background distributions and are often constrained by previous knowledge or *in situ* measurements represented in the function $\Pi(\eta)$. When the signal and background predictions are well defined, the probability function, $\mathcal{P}$, is taken to be a Poisson function in the small-count regime or a Gaussian function in the large-count regime.

If the signal or background predictions carry large uncertainties, which is often the case for rare backgrounds or signals that cover very specific parts of phase space such as dark matter lines [284], stochastic likelihood models can be used [285, 286, 287]. For other treatments proposed to tackle this problem see also [288, 289, 290].

In either case, the treatment of systematic uncertainties is often done by using the profile likelihood method, in which the likelihood function is maximized over the nuisance parameter at each physics parameter point [291]. Alternatively, in Bayesian treatments (see *e.g.* [292]) or hybrid frequentist-Bayesian treatments [293] the nuisance parameters are marginalized over by integrating the likelihood function. In the case that the bin content is large, such that a Gaussian likelihood

function is a good approximation, the expectations can be computed accurately. Often, a multidimensional Gaussian is used where the covariance between bins incorporates both the systematic and statistical uncertainties. The latter approach does not require additional parameters to incorporate systematic uncertainties into the likelihood, making it computationally advantageous.

With this formalism, background-informed analyses have additional power compared to the background agnostic scenario, provided that experiments are capable of constraining the background size, and separating it from signal. The ability to constrain background is encapsulated in systematic uncertainties, whereas the separation of background from signal depends on the features of both. The features in the case of neutrinos from dark matter are a democratic flavor composition, spatial clustering predominantly around the Galactic center, and an energy distribution which is maximal close to the dark matter mass. Separating dark matter from background using these three features then depends on the experimental direction and energy resolutions, as well as its flavor identification capabilities dictated by the event morphological classification. The latter is important since natural and anthropogenic sources often have a non-democratic flavor composition. This is a characteristic of the stronger constraints. For example, we use the fact that for MeV dark matter one of the main backgrounds are solar neutrinos, which can be efficiently removed by selecting only for antineutrinos in Super-Kamiokande or JUNO; we also rely on this in our predictions of the sensitivities for DUNE and Hyper-Kamiokande in the 100 MeV to 30 GeV energy range, where we use the fact that one can do morphological event analysis to remove muon neutrinos which are the dominant component of the atmospheric flux.

### Neutrino Detection Methods

Because neutrinos only interact via the weak nuclear force, neutrino detection must proceed in at least two steps: first, interaction between a neutrino and a detector electron or nucleus, and second, the detection of the resulting electromagnetic signal. Typically, energy from a gamma-ray or electron cascades down via scintillation, additional ionization or Cherenkov radiation and is subsequently measured by optical sensors or charge readout.

The small neutrino detection cross section poses a great challenge in the search for the expected fluxes from dark matter annihilation to neutrinos. As the dark matter mass increases, larger detectors are necessary to compensate for the smaller flux, which scales as $m_\chi^{-2}$. Such a scaling can come at the cost of energy and angular resolution, as well as flavor identification, all of which allow differ-

entiation between the dark matter induced neutrinos from other natural or anthropogenic neutrino sources as discussed in the previous section. In this section, we review the techniques used to detect neutrinos in different energy ranges; see also [82, 294] for a discussion in the context of neutrino oscillation experiments. Note that the energy ranges detailed here are approximate, and there is naturally some overlap between techniques and physics discussed in each respective subsection.

**Neutrino energies below 10 MeV**

Coherent elastic neutrino-nucleus scattering, namely $\nu A_N^Z \to \nu A_N^{*Z}$, dominates the cross section at the lowest energies [295]. This process, sometimes abbreviated as CE$\nu$NS, has no kinematic threshold and scales quadratically with the atomic number. However, the maximum recoil energies are very small making its detection difficult; in fact it has only recently been observed using anthropogenic neutrinos in detectors of $\mathcal{O}(10)$ kg of mass [296]. Future ton-scale dark matter direct detection experiments such as DARWIN [297] expect to see solar and atmospheric neutrinos via CE$\nu$NS. Because of the trade-off between detector size and nuclear recoil threshold, they would only be sensitive to DM above $m_\chi \sim 10$ MeV, and provide only marginal improvement over existing dedicated neutrino experiments that use different detection channels.

Neutrino-electron scattering also has no kinematic threshold at detectable energies, and the cross section is predicted without ambiguities that arise from form factors in hadron-neutrino interactions. This interaction's well-understood kinematics, together with the fact that a single outgoing charged particle is produced, makes it a good channel to use for DM annihilation searches. This is because precise energy and directional information can be inferred. The angle between the neutrino and the electron is tightly constrained by the kinematics, $E_e\theta_e < 2m_e$, allowing for an accurate reconstruction of the neutrino direction (it was through this process that in 1998 the Super-Kamiokande experiment made the first image of the Sun in neutrinos [298]; see also [299, 300, 301] for subsequent measurements by SNO and Borexino). Angular information is used to mitigate the $\sim$ 1-10 MeV solar neutrino backgrounds and to search for correlations with the expected angular distribution of DM via $J(\Omega)$. Unfortunately, the neutrino-electron cross section is approximately $10^{-43}$ cm$^2$ at 5 MeV, which is about a factor of 10 smaller than the dominant neutrino-nucleon process.

The other commonly-used technique to detect sub-10 MeV neutrinos is inverse beta decay (IBD), $\bar{\nu}_e p \to n e^+$. This is due to three reasons: first, the large and well-measured IBD cross section, approximately $10^{-42}$ cm$^2$ at 5 MeV [75, 302], with an uncertainty of $\sim 0.2\%$ [75, 76]; second, the

low-threshold: $E_\nu > 1.806$ MeV; and finally, the ability to reduce background by searching for the prompt positron signature followed by the neutron capture. This detection method is often used with hydrocarbon-based scintillator since it contains a large number of free protons and emits large number of photons, typically $10^4$ per MeV of deposited energy [303]. The energy deposited by the prompt signal is the kinetic energy of the positron plus two 511 keV gamma-rays from electron-positron annihilation, and a 2.2 MeV gamma ray from the delayed capture of the neutron on free protons. In hydrogen-based detectors the neutron capture time is typically 300 $\mu s$. If the detector is doped with 1% Gadolinium, this time is reduced to about 20 $\mu s$ and the prompt gamma-ray energy is 8 MeV allowing for an improved background suppression [304]; *e.g.* in the case of Super-Kamiokande a hundredfold background suppression efficiency can be achieved [305]. In the search for dark matter this process has the advantage that it is only triggered by $\bar{\nu}_e$ allowing for very efficient suppression of the solar neutrino flux that dominates the natural backgrounds at sub-10 MeV energies. In fact, our strongest limit across all dark matter masses comes from an IBD search by Super-Kamiokande; see Fig. 6.2.

**Neutrino energies between 10 MeV and 1 GeV**

Between $\sim$ 10 MeV and $\sim$ 1 GeV, in Cherenkov detectors the proton is invisible since it is Cherenkov threshold – approximately 1.3 GeV in mineral oil, 1.4 GeV in water, and can be as low as 1.2 GeV in the Antarctic ice [306]. This has advantages and disadvantages compared to scintillator detectors, on the one hand it simplifies identification and classification of events since the observed Cherenkov light must be associated with the outgoing charged lepton. On the other hand, the lack of proton kinematics means that the energy and angular resolution can be greatly degraded. The dominant neutrino-nucleon process in this energy range is that of charged-current quasi-elastic (CCQE) scattering, namely $\nu_\alpha N \to \alpha \tilde{N}$ where $\alpha$ is a charged lepton and $N$ ($\tilde{N}$) is a proton or neutron. At high enough energies, muon neutrinos can have CCQE interactions, producing muons which can be identified by the morphology of the Cherenkov ring. Due to the larger mass, muons tend to preserve their direction as they travel through the detector producing sharper rings than electrons. Cherenkov detectors can be constructed out of mineral oil, water, or ice. Although oil-based detectors boast a larger Cherenkov angle and the ability to run without a purification system, they are only utilized in smaller detectors [294] due to the higher filling cost. For this reason, multi-kiloton detectors available as of 2020 are all water or ice based. As early as 2022, JUNO will become the first

multi-kiloton liquid scintillator detector. Since the DM-induced flux is expected to be very small, the larger water or ice Cherenkov detectors currently dominate the constraints over oil-Cherenkov detectors and we will not discuss them further.

**Neutrino energies from 1 GeV to $10^7$ GeV**

Resonant light-meson production is important between approximately 1 and 10 GeV. Due to the difficulty in cross section modeling, neutrino detection in this range is subject to large uncertainties. Above 10 GeV the contribution of deep inelastic scattering (DIS), where the neutrino exchanges a $W$ or $Z$ boson with one of the partons inside the nucleon becomes the dominant process. The production of taus in tau-neutrino charged-current interactions becomes possible above the threshold $m_\tau = 1.777$ GeV, though the cross section is only around 15% of the charged-current muon-neutrino cross section at 10 GeV, rising to 75 % at 100 GeV [307].

Though unsegmented Cherenkov detectors are still used in this energy range, the use of tracking calorimeters, often constructed as segmented scintillators, are popular as they allow for improved reconstruction of outgoing muon tracks, as well as electromagnetic and hadronic showers produced in the interaction vertex. Notable examples of these types of detectors in contemporary neutrino physics are the NO$\nu$A experiment and the T2K near-detector. Sampling calorimeters have also been used to increase the target density, though this comes at the expense of a degraded energy resolution. In this case a dense material like iron is interleaved with scintillator panels. This design was used by the MINER$\nu$A experiment [308] to perform precision measurements of the neutrino cross section and has been used in the past to measure neutrino oscillations by MINOS [309]. In these detectors the morphological features observed in the trackers have been used to identify the different neutrino interaction processes by comparing them to generated event libraries [310, 311] or convolutional neural networks [312, 313]. Given the size of these detectors they are not expected to play a role in the detection of dark matter and are not included in this work.

The newest neutrino detectors in this energy range are the so-called liquid argon time projection chambers (LArTPC) [314]. These detectors consist of an electric field cage filled with liquid argon. When a charged particle is produced in the argon, it travels through the medium and ionizes the argon atoms, liberating electrons. An electric field then drifts the electrons to wire planes on one side of the detector, recording a projected footprint of the interaction. Three dimensional reconstruction is also possible by using the timing of the charge deposition on the wires. To localize

the event in the third dimension, the drift time of electrons in argon and the initial interaction time need to be known. The initial interaction time can be known in the case of generic neutrino interactions via the scintillation light produced by the charged particles in argon or, in the case of neutrinos produced in bunches in a beam, by the beam timing. In the case of dark matter searches, relevant for this work, only the former technique is relevant. Even though the neutrino-argon cross section is currently poorly understood compared to other materials conventionally used in neutrino physics, these detectors have the potential for unprecedented particle identification: see *e.g.* [315, 316, 317]. Examples of currently operating LArTPC neutrino detectors are MicroBooNE [318] and ICARUS [319] at Fermilab. The next generation experiment in this category is DUNE [320].

At the higher end of this energy range, neutrino telescopes such as ANTARES and IceCube have the largest neutrino collection volumes. These detectors operate at energies above 10 GeV where DIS is the dominant cross section process [321]. These detectors use natural media, such as the Mediterranean water or the Antarctic ice, as targets for the neutrino interaction. Cherenkov light produced by charged particles by products of these interactions are then observed by photomultiplier tubes (PMTs) arranged on sparse arrays. In these detectors the different neutrino interactions map onto different morphologies of the time and spatial distribution of charge in the array. Neutral-current interactions, charged-current electron-neutrino interactions, and most of the charged-current tau-neutrino interactions produce a morphology known as a cascade. Because cascades can be contained in the detector, this morphology has the best energy resolution. Charged-current muon neutrino interactions produce a morphology known as tracks, due to the long travel-time of the muon. This morphology provides the best directional information. In water, photons tend to scatter less than in ice, providing more direct light. This means that the muon angular resolution in water-based detectors is better than those in ice. On the other hand, given the longer absorption length of photons in ice compared to water, the effective detector volume is larger for detectors deployed deep in the ice. Finally, charged-current tau neutrino interactions can produce a variety of morphologies depending on the boost factor of the tau and its decay channel. For example, around 1 PeV, a tau can travel on average 50 m before decaying producing separated energy depositions known as *double bangs* [322, 323]; in 2018 IceCube announced the first candidate astrophysical tau events [129, 324]. Finally, in these detectors one can also observe the electron-neutrino scattering, since at approximately 6.3 PeV an electron antineutrino can resonantly scatter with an atomic electron producing a $W$ on shell [105, 325]; $W$-production of coherent photon scattering can also

be important at these energies see [100, 101, 326, 103, 86]. The observation of this process provides a unique handle on the ratio of neutrinos to antineutrinos, as well as providing exquisite energy resolution; and in fact, a candidate event has recently been detected [327].

**Neutrino energies above $10^7$ GeV**

At extremely-high energies, the neutrino flux expected from dark matter and other astrophysical sources such as cosmogenic neutrinos is very small, necessitating the construction of detectors with effective volumes much larger than a cubic kilometer. Neutrino interactions in this energy range occur overwhelmingly via deep inelastic scattering [321]. Two main techniques are used to search for neutrinos in this energy range, both of which rely on identifying horizontal or upgoing particles to mitigate the larger cosmic-ray backgrounds. The first method involves looking for air showers induced by neutrino-nucleus interactions in the atmosphere or just below the surface of the Earth, while the second uses the radio signature produced in very-high-energy neutrino interaction [328, 329], known as Askaryan radiation [330, 331]. This former technique can be detected in a number of ways: sparse surface arrays of water Cherenkov tanks are used to identify charged particles from showers as they develop over an area that may span many square-km. Air fluorescence telescopes and optical air Cherenkov telescopes can also be used alone or in combination with water tanks as is the case for Auger [332]. The timing, morphology, and amount of light deposition is used to infer the energy of the incoming particle, its direction, and its nature. In particular, a neutrino will typically travel much deeper into the atmosphere than a cosmic ray or gamma ray before interacting. Tau neutrinos are particularly promising, as $\tau$ leptons can be produced in a nearby mountain or below the horizon [333]. If the tau survives the journey out of the mountain, its decay yields an upgoing air shower [71, 334]; an EeV $\tau$ typical interaction length is a few kilometers in rock and is shorter than its decay length. The expected event rate for such processes at cosmic ray observatories like Auger turns out to be higher than from neutrino-induced atmospheric showers, thanks to the high density of rock. Radio arrays such as GRAND [279] have been proposed to cover as large an effective area as possible (up to two-hundred thousand square-km) to search for such a signal.

The second method, Askaryan radiation detection, aims to observe neutrinos via the radio emission generated by charge displacement caused by the developing electromagnetic or hadronic shower after DIS scattering. This emission is distinct from down-going cosmic-ray showers in that the polarization of the radio signal is expected to be different. This technique has been implemented

by using radio antennae either suspended from balloons [335] or buried in the ice [336, 337] in the Antartic continent. The ability to cover a large area with a single antenna cluster makes this a very scalable and relatively low-cost technique.

## 6.7 Results

Our main results are shown in Figs. 6.2-6.7. Fig. 6.2 shows the results derived according to the procedures described in Secs. 6.3, in addition to previous results available in the literature. Fig. 6.4 shows a more detailed view of the low-mass (sub-GeV) range; Fig. 6.5 shows results for the high-mass ($10^3$-$10^{11}$ GeV) region. Finally, Figs. 6.6-6.7 provide the constraints and projections in the case of velocity-dependent p-wave and d-wave annihilation, respectively. Further, Fig. 6.3 shows results for decay of dark matter according to procedures described in Sec. 6.5. We label the results derived specifically for this work with a heart (♡).

In the rest of this section, we describe the data that we used to produce or recast limits according to the procedures outlined in Secs. 6.3 and 6.5. We split the data into three lists: 1) data used to construct constraints in Fig. 6.2 and Fig. 6.3; 2) previous limits that we have recast; and 3) data used to place limits in the high mass ($m_\chi > 10^3$ GeV) region.

When reporting literature results, where possible, we have rescaled them to use the same halo parameters, *i.e.* consistent $J$- and $D$-factors, as computed in Sec. 6.3 and Sec. 6.5. In this way, we ensure that the constraints we present can be properly compared one with another. The rescaling could not be done in the case of ANTARES [272], SK [338], and IceCube [339], since these were event-by-event analyses for which data is not publicly available. This is unfortunate since the halo parameters used in these studies are no longer preferred (see discussion in Sec. 6.7). Shaded regions correspond to experimental limits, whereas dashed lines are projections based on future experimental sensitivity. Finally, in the case of DM annihilation we include two lines for reference. First, the dotted black line corresponds to the cross section required to produce the observed relic abundance from thermal freeze-out computed as in [340], and second, the solid black line labeled "unitarity bound" corresponds to the perturbative unitarity limit on non-composite WIMP dark matter [341]; see [342] for a recent discussion.

The limits shown in Figs. 6.2 and 6.3, employing the approach of Secs. 6.3 and 6.5 use the following data, which we also summarize in Tbl. 6.4.

1. ***Borexino:*** Borexino is a large-volume unsegmented liquid scintillator detector located underground at the *Laboratori Nazionali del Gran Sasso* in Italy [343]. The collaboration has released two event selections: one which has a livetime of 736 days selecting electron-antineutrino candidate events over the entire fiducial volume and another one with 482 days of livetime designed to search for geo-neutrinos [344]. These event selections are combined into a single set designed to obtain a pure sample of electron-antineutrinos by means of searching for signatures of inverse beta decay. Using this selection, they derive upper limits on the all-sky monochromatic electron-antineutrino flux ranging from $\sim 10^5$ to $\sim 10^2$ $\bar{\nu}_e \mathrm{cm}^{-2}\mathrm{s}^{-1}$, for energies ranging from $\sim 2$ to 17 MeV, respectively. We use the flux upper limits produced by [345] and recently updated by [346] and compare it with one-sixth of the all-flavor expected flux from dark matter to set our constraints.

2. ***KamLAND:*** KamLAND is an unsegmented liquid scintillator detector located in the Kamioka observatory near Toyama, Japan. The approximately one kiloton of mineral oil fiducial volume is contained in a 13 meter balloon. Beyond its well-known work on reactor neutrinos, KamLAND has measured the $^8$B solar spectrum [347], searched for geoneutrinos [348], and placed limits on the flux of extraterrestrial neutrinos above $\sim 8.3$ MeV [349] which constrains the supernovae relic neutrino flux. In the latter work, an upper limit on the extraterrestrial flux of $\bar{\nu}_e$ is derived, which is at the $\mathcal{O}(10)$ $\bar{\nu}_e$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{MeV}^{-1}$ level and is given from 8.3 MeV to 18.3 MeV. Using this result, we derive a constraint on the dark matter parameters, shown in salmon. Note that in [349], the KamLAND collaboration also derives a similar constraint, but with outdated $J$-factors; their result and ours are comparable. These are the leading constraints in the $\sim$10 MeV mass range, but we expect that they will be improved by the next-generation liquid scintillator detector in China, JUNO [350].

3. ***SK:*** Super-Kamiokande (SK) is a 50kt ultrapure water Cherenkov detector located in Kamioka, Japan [351]. SK can use the morphology of the Cherenkov ring produced by charged particles to perform particle identification, energy measurement, and obtain directional information of the events. The unfolded electron- and muon-neutrino fluxes in the sub-GeV to several TeV energy range has been published by SK [280]. This unfolding uses data from the four stages, SK-I, SK-II, SK-III, and SK-IV, resulting in a total livetime of 4799 days for the fully contained and partially contained event selection and 5103 for the upward-going muon sample. The un-

folded fluxes are expected to be dominated by the atmospheric neutrino flux; in fact they are in agreement with model predictions, *e.g.* the HKKM model [171], within systematic uncertainties. The dominant source of uncertainties on the unfolded fluxes is the neutrino interaction cross section, which introduces an uncertainty of approximately 20% in the unfolded flux. In the case of electron-neutrinos, the second largest uncertainty is due to the small statistics at high energies; which can be up to 10% in the highest energy bins. For all flavors, all other sources of uncertainty are less than 5% across all energy bins. We compare the unfolded flux with the expected flux from dark matter to produce limits on Galactic and extragalactic dark matter annihilation and decay. These results are shown in purple in Figs. 6.2, 6.3, 6.6, and 6.9, and labeled as ♡***SK-Atm.*** In order to obtain these limits we used a background-agnostic approach as described in Sec. 6.6, and a binned truncated Gaussian likelihood in energy with two degrees of freedom. Additionally, we perform an analysis using 2853 days of low energy data from SK I/II/III, as well as 2778 days of data from SK phase IV, which led to an upper limit on the relic supernova electron antineutrino ($\bar{\nu}_e$) flux [352]; labeled ♡***SK-***$\bar{\nu}_{\mathbf{e}}$. The resulting limits on $\langle \sigma v \rangle$ turn out to be the strongest over the entire mass range that we consider, crossing the relic abundance line for masses between 27 and 30 MeV.

4. ***IceCube***: The IceCube Neutrino Observatory is a gigaton ice Cherenkov neutrino detector located at the geographic South Pole [353]. IceCube has measured the atmospheric neutrino spectrum in the 100 GeV to 100 TeV energy range. By separating the events into their observed morphologies ("cascades" and "tracks"), the collaboration recently published the unfolded electron- and muon-neutrino flux in this energy range [281, 179]. At energies greater than 60 TeV, using events whose interaction vertex starts in the inner part of the detector [121, 176], they have also reported the result of a piece-wise power-law fit to the astrophysical neutrino component using more than six years of data [354]. We use these to produce background-agnostic limits comparing the produced neutrino flux from DM with the reported unfolding or spectral fits. The obtained limits are shown for dark matter masses from 200 GeV to 10 PeV, labeled ♡***IceCube-HE*** and colored in dark magenta. Limits use the same likelihood construction as in the case of the SK limits described above. Note that the muon neutrino atmospheric unfolding reported by IceCube uses northern tracks, which are unfortunately in the wrong hemisphere for the Galactic center. Therefore, for that sample, we only constrain extragalactic emission. Dedicated neutrino line searches have not been yet performed by the

IceCube collaboration, although sensitivities have been estimated in [355, 356] to be stronger than current IceCube constraints in that region. We describe the region labeled ***IceCube-EHE*** below, in the description of the high-mass region.

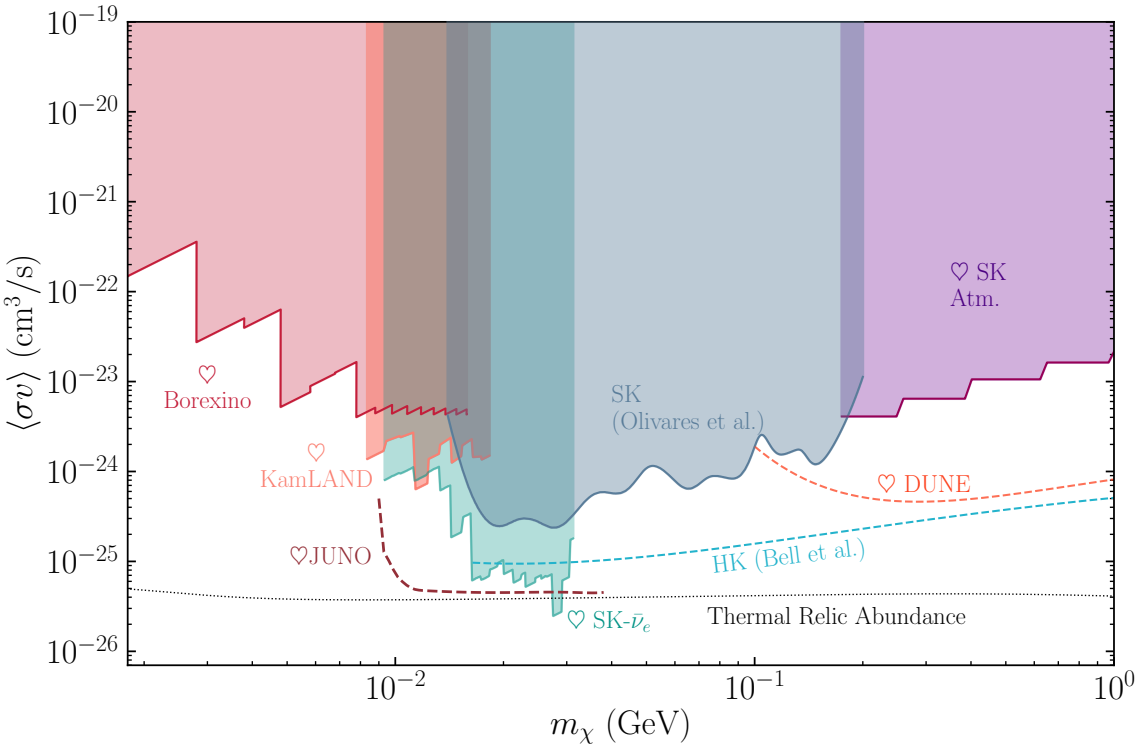



Figure 6.4: The landscape of sub-GeV dark matter annihilation into neutrinos. Same as Fig. 6.2, but restricted to dark matter masses below one GeV.

Additionally, we use the following previously-published limits on dark matter annihilation and decay obtained by constraining the Galactic flux, rescaled to account for the galactic halo parameters used here unless indicated otherwise:

1. ***Super-Kamiokande diffuse supernovae flux search***: The gray region labeled ***SK Olivares et al.*** is an independent analysis of SK all-sky low-energy data which uses SK phases I through III to derive an upper bound on the supernova relic neutrinos [357, 358, 359]. This analysis covers neutrino energies from 10 MeV to 200 MeV; see [360] for a recent discussion of backgrounds in the low-energy range. The upper limit on supernova relic neutrinos was then converted into dark matter annihilation constraints, and was originally presented in [241, 361, 362]. Recently, SK phase-IV data has placed new constraints on the $\bar{\nu}_e$ flux

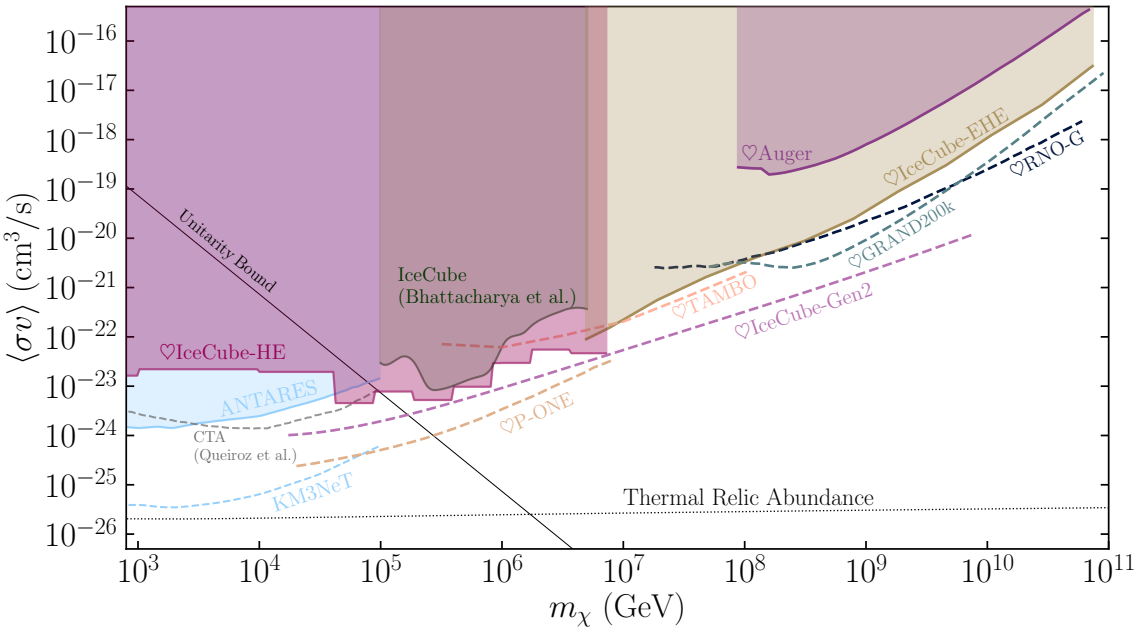



Figure 6.5: The landscape of supra-TeV dark matter annihilation into neutrinos. Same as Fig. 6.2, but for the high-mass region. All the experimental constraints in this plot are calculated by converting either the detected flux or the reported upper limit into a conservative upper bound on the DM annihilation cross section.

in the 10 to 30 MeV energy range [352]. These observations improve over KamLAND constraints [349] by a factor between 3 and 10 in their overlapping energy range. Thus these observations dominate the constraints for dark matter masses below $\sim 20$ MeV. Where they overlap, the *Olivares et al.* limits are not quite as strong as the *SK-*$\bar{\nu}_e$ limits that we have presented, because their background modelling could not use angular information which is not publicly available.

2. ***Super-Kamiokande Galactic dark matter search***: The teal region, labeled ***SK***, is from [363]. This analysis uses muon-neutrino data in the energy range between 1 GeV and 10 TeV collected by SK over 5325.8 days. Since this analysis relies on angular information that is not public, it has not been rescaled to account for our choice of galactic halo parameters.

3. ***IceCube/DeepCore Galactic dark matter search***: The IceCube limits are from [339] and use 329 days of IceCube data. These place constraints for masses in between 25 GeV and 10 TeV. At the lowest masses, these limits include data from DeepCore, an array of more closely spaced inner strings in IceCube. In addition, we include a limit derived from 3 years of

data using primarily tracks to constrain Galactic center emission [364]. For display purposes, we join these two lines, choosing the best limit at each point, and show it in navy blue, simply labeled as ***IceCube***.

4. ***IceCube-Bhattacharya et al.*** is taken from [365]'s channel-by-channel unbinned likelihood analysis of the High-Energy Starting Event (HESE) data, including energy, angular, and topology information. They include both Galactic and extragalactic constraints. Constraints that we derive (IceCube-HE) using only spectral information follow these limits quite closely at higher energies since the small sample size prevent angular information from contributing significantly.

5. ***ANTARES dedicated Galactic dark matter search***: The light blue region, labeled ANTARES, is from a Galactic center analysis of nine years of ANTARES muon neutrino and antineutrino data [366, 272]. This covers the dark matter mass range from 53 GeV to 100 TeV.

6. ***Baikal dedicated Galactic dark matter search* (not shown)**: The Baikal underwater neutrino telescope [367, 368], NT-200, is a water Cherenkov detector deployed in Lake Baikal, Russia. It has an instrumented volume of approximately 100 kt and is comprised of 192 optical modules arranged on eight strings, with a typical distance between strings of 21 m. The collaboration performed an analysis looking for dark matter annihilation in the Galactic center into neutrinos using data recorded between April of 1998 to February of 2003 [369]. This analysis claimed to place limits on the cross section at the $10^{-22}$ cm$^3$s$^{-1}$ level for a 1 TeV dark matter mass. We do not add this result to our constraint summary because there are stronger results in this mass range, but we do show the projections of the next generation detector at Lake Baikal, GVD.

Finally, Fig. 6.2, includes next-generation sensitivities that can be reached by future experiments. These are shown as dashed lines:

1. ***DUNE***: The Deep Underground Neutrino Experiment (DUNE) far detector will be a 46.4 kiloton liquid argon Time Projection Chamber (TPC) [370, 371] constructed at the Sanford Underground Research Facility (SURF) in South Dakota, USA. Its main advantage in detecting neutrinos from DM annihilation and decay is its improved particle identification, using

morphological reconstruction, with respect to Cherenkov detectors like Super-Kamiokande, ANTARES, or IceCube, which *e.g.* can be exploited to make improved measurements of solar neutrinos [372]. Thus, a dedicated DUNE analysis utilizing the expected improved directional capability can prove effective in a search for Galactic dark matter annihilation to neutrinos. We derive projected sensitivities for dark matter masses in the range from 100 MeV to 30 GeV and show them in Figs. 6.2 and 6.3 as dashed orange lines. The dominant background in this energy range is from atmospheric neutrinos. We use the predictions provided by [373] at the Homestake gold mine at SURF, taking into account oscillations through the Earth using the nuSQuIDS package [374, 375, 376]. In our analysis, we consider $e$- and $\tau$-flavored charged-current interactions and compare the expected energy distribution; *i.e.* we do not take into account event-by-event directional information. We use a fractional charged lepton energy resolution of $2\% + 15\%/\sqrt{E/\mathrm{GeV}}$ [370] and assume the idealized condition of 100% efficiency. In our analysis, charged-current electron-neutrino interactions are assumed to deposit all their energy in the detector, while tau-neutrino charged-current interactions will deposit less visible energy due to the invisible neutrinos produced in the prompt $\tau$ decay. Since we expect that DUNE morphological identification will be able to single out muon-neutrino charged-current processes, we choose to remove them from the analysis as they are the primary contributor to the atmospheric neutrino background. Limits are derived using a binned Poisson likelihood and a background-informed method as described in Sec. 6.6. We note that, due to liquid argon TPC's morphological reconstruction capabilities, a proper Galactic center analysis including directionality would benefit from the inclusion of muon-neutrino charged-current interactions, and thus our projections are conservative.

2. ***Hyper-Kamiokande***: Building on SK's technology, a new water Cherenkov detector with a fiducial mass of 187 kton called Hyper-Kamiokande (HK) will be built in Kamioka, Japan [377]. Due to its larger size, this detector will be able to place stronger limits on the DM annihilation and decay to neutrinos than its predecessor [361]. In fact, Hyper-Kamiokande is estimated to reach $\sim 10^{-25}$ cm$^3$s$^{-1}$ for 1 GeV dark matter and $\sim 10^{-22}$ cm$^3$s$^{-1}$ at $10^4$ GeV with ten years of data taking [378]. Furthermore, the possibility of doping both the SK and the HK detectors with gadolinium (Gd) will reduce the dominant background for low-energy analyses by a factor of five and, consequently, improve the constraints on DM parameters [379, 380, 381].

Similar to our DUNE analysis, we assume that the dominant background in this energy range is due to atmospheric neutrinos, where we use the predictions provided by [373] at the Kamioka mines, and allow these neutrinos to oscillate through the Earth using the nuSQuIDS package [374, 376]. We only consider $e$- and $\tau$-flavored charged-current interactions, without taking into account directionality. We make the same assumptions as our DUNE analysis regarding energy deposition, while using an energy resolution of $1.5\% + 2\%/\sqrt{E/\mathrm{GeV}}$ [382]. We use total energy rather than lepton (visible) energy, which leads to a sensitivity overestimate of $\sim 40\%$ but simplifies the analysis. In principle, it is be possible to record lepton and proton energy above the proton Cherenkov threshold, [see e.g [383]. We follow the same statistical procedure as in DUNE and, like DUNE, the sensitivity strength derives primarily from the expected electron- and tau-neutrinos signal. Taking advantage of this channel explains why our estimates are better than ones presented by [378]; see [384] for a discussion on "shower power." These projected sensitivities, especially at low energies, are subject to a $\sim 30\%$ uncertainty due to a combination of atmospheric background uncertainties and neutrino cross sections.

3. ***JUNO***: The Jiangmen Underground Neutrino Observatory [350] is a 20 kt unsegmented liquid scintillator detector under deployment in the Guangdong province of China. The detector has a muon tracker on top of it and is also surrounded by water. Both of these systems can be used to veto cosmic-ray muons by either tagging them in the muon tracker or by detecting their Cherenkov light in water. Due to its large volume and good energy resolution (estimated to be $3\%/\sqrt{E/\mathrm{MeV}}$) we expect that this experiment will have good sensitivity for neutrino line searches. We estimate the sensitivity of JUNO to dark matter annihilation to neutrinos in the electron antineutrino channel following the proposal given in [385]. We use background estimates derived for diffuse supernova background searches, as presented in [350]. Below 11 MeV, reactor antineutrinos dominate the background. Between 11 and 40 MeV, the backgrounds are primarily neutral current interactions from atmospheric neutrinos, with subdominant charge current contributions. According to our projection, JUNO is expected to constrain the velocity-averaged annihilation cross section better than $10^{-25}$ cm$^3$s$^{-1}$ in the 10 to 40 MeV mass range. The estimate is shown in dark red in Fig. 6.2, and its equivalent dark matter lifetime is shown in Fig. 6.3.

4. ***IceCube Upgrade***: The IceCube Upgrade is an extension of the current IceCube/DeepCore array with seven closely-packed strings. These new strings will be separated by approximately 20 meters and each contain 100 photomultiplier tubes spaced vertically by 3 meters [15]. Additionally, a number of calibration devices and sensors will be deployed to improve the modelling of the ice [386, 387]. In [388] a preliminary estimation of the IceCube Upgrade sensitivity was performed. It is expected to be better than $10^{-24}$ cm$^3$s$^{-1}$ for a 10 GeV dark matter mass.

5. ***IceCube Gen-2***: The next-generation ice Cherenkov neutrino observatory in Antarctica is a substantial expansion to the current IceCube observatory, aiming at enhancing the detector volume by a factor of ten [389]. This increased effective area is expected to provide a better sensitivity to resolve sources of high-energy cosmic neutrinos and identify components of cosmic neutrino flux. Dark matter limits from IceCube presented here should therefore scale by at least the increased sample size due to the larger effective area. We have recast the estimates of diffuse flux sensitivity given in [390] to estimate the sensitivity to dark matter.

6. ***Baikal-GVD***: The Baikal Gigaton Volume Detector (GVD) is a planned expansion to the existing NT-200 detector, and is currently being deployed in Lake Baikal, Russia. The detector has recently reached an effective volume of $\sim 0.35$ km$^3$ and has already seen first $\nu$-light [391]. The full array will contain 10,386 optical modules divided among 27 clusters of strings, and is expected to have a final instrumented volume of around 1.5 km$^3$. The sensitivity of GVD to Galactic dark matter has been estimated in [392] and is shown as a dashed brown line labeled **GVD**.

7. ***KM3Net***: The km$^3$-scale water Cherenkov detector currently under construction in the Mediterranean sea is designed to provide high-purity increased effective areas in the Southern Hemisphere. The larger effective area and improved angular resolution, compared to ANTARES, are expected to provide better constraints on dark matter. Two separate sites are under construction for low- and high-energy regimes [393]. The high-energy site, called KM3NeT/ARCA, will consist of two detector array blocks located approximately 100 km offshore from *Porto Palo di Capo Passero*, Sicily, Italy [394]. Each block is expected to have 115 strings with an average spacing of 90 m. The low-energy site, called KM3NeT/ORCA, consists of one array block and is under deployment approximately 40 km south of Toulon, France;

close to the ANTARES site. The array is made out of 115 strings with an average horizontal spacing of 20 m. Each string contains 18 optical modules; in KM3NeT/ARCA they are spaced vertically by 36 m, while in KM3NeT/ORCA they are spaced 9 m. The horizontal spacing and number of strings are proportional to the effective volume of the experiment, while the vertical spacing is related to the energy threshold [395]. KM3NeT/ARCA's science program is mainly oriented towards higher-energy (astrophysical) neutrino searches, while KM3NeT/ORCA will measure neutrino oscillations using atmospheric neutrinos. Assuming an $E^{-2}$ democratic-flavor astrophysical neutrino flux with a normalization of $\sim 1.8 \times 10^{-8}\text{GeV}^{-1}\text{s}^{-1}\text{cm}^{-1}\text{sr}^{-1}$ and an exponential cut-off at 3 PeV they expect to see 11 $\nu_\mu$'s, 41 $\nu_e$'s, and 26 $\nu_\tau$'s in five years of KM3NeT/ARCA operation [393]. In Figs. 6.2 and 6.3 we show the KM3NeT/ARCA expected sensitivity to dark matter annihilation to neutrinos in five years of data taking [396]. Their sensitivity is within a factor of a few from the expected relic abundance cross section for dark matter masses around a TeV.

8. ***P-ONE***: The Pacific-Ocean Neutrino Experiment (P-ONE) is a newly proposed multi-cubic kilometer neutrino detector utilizing sea water as Cherenkov medium [397]. P-ONE would be deployed in the Cascadia Basin, off the coast of Vancouver island in the Pacific Ocean, taking full advantage of the Ocean Network Canada infrastructure and expertise already in place. The main goal of the experiment is to explore the origin of the extraterrestrial neutrino flux. A pair of test strings, named STRAW [398], has already been successfully deployed and has collected water absorption data. The first phase of the detector, known as the Pacific Ocean Neutrino *Explorer*, involving ten strings is planned to begin deployment in 2023. Each string is planned to be equipped with twenty photomultiplier tubes. The full detector is expected to be complete by 2030 with 70 strings. Projected limits include backgrounds from atmospheric and diffuse astrophysical neutrinos, and use the exposures shown in [397].

9. ***TAMBO***: The Tau Air-Shower Mountain-Based Observatory is a proposed array of small water-Cherenkov tanks to be deployed on either the Colca Valley or Cotahuasi Canyon in Peru [399, 140]. These are two of the world's four deepest valleys and their unique geometry allows for efficient detection of Earth-skimming PeV $\nu_\tau$. Most of the Colca Valley runs along a North-South corridor, though a smaller section of it has an East-West corridor. If deployed in the East-West corridor of the Colca valley, the declination band covered is $-15.5 \pm 10$ de-

grees, while in the North-South corridor it would be $-15.5 \pm 50$ degrees. These two provide two extreme configurations in terms of its GC exposure, while a deployment in the Cotahuasi canyon, which has an approximately diagonal corridor, would provide an intermediate exposure. TAMBO's effective area is expected to be 10 times larger than IceCube $\nu_\tau$ [121] at a PeV and 30 times larger at 10 PeV. The use of the Earth-skimming technique is complementary to very-high-energy Earth-traversing neutrino searches [38] and the fact that it relies on the Cherenkov effect, rather than the higher energy threshold Askaryan effect, gives it unique potential to constrain dark matter in the tens of PeV mass range. Depending on the final geometry of TAMBO its sensitivity to dark matter ranges from $10^{-22}$ cm$^3$ s$^{-1}$ to $4 \times 10^{-21}$ cm$^3$ s$^{-1}$ for a 1 PeV dark matter mass. Sensitivities shown here are recast from the diffuse flux sensitivity presented by [399].

10. ***CTA***: The Cherenkov Telescope Array is a planned network of 99 air Cherenkov telescopes in the southern hemisphere and 19 in the northern hemisphere that will collectively provide full-sky coverage of the gamma ray sky over an energy range from 20 GeV to 300 TeV [400]. Several CTA prototypes have been built and some have already seen first light. The telescopes are projected to have an angular resolution down to 0.1 degrees and a duty cycle of $\sim 15\%$. For high-mass dark matter annihilation into neutrinos, electroweak final-state radiation can also lead to the production of gamma rays, despite a completely "invisible" $\nu\bar{\nu}$ final state, and can thus be constrained by gamma ray observations of the Galactic center with CTA; see Sec. 6.3 for more details. The expected limits from CTA were computed in [271], and shown as a dashed silver line assuming 100 hours of observation.

We note that the 10 MeV – 1 GeV range can in principle be covered by future tonne-scale dark matter direct detection experiments such as DARWIN and ARGO [401]. However, these are still in their planning phases, meaning that construction is still decades away, and very long ($\gtrsim$ 10 years) exposure times are required to be competitive with HyperK. For this reason we do not show them here.

Fig. 6.5 shows the extension of available constraints to larger masses, above the "unitarity bound," accessible *e.g.* for composite DM models [402]. These bounds are calculated by converting either the detected flux or reported upper limits, from observatories sensitive to these mass range, into a conservative upper bound on the DM annihilation to neutrinos. The following experiments are sensitive to this regime:

1. ***Auger***: The Pierre Auger Observatory is a hybrid detector consisting of both an array of water Cherenkov surface detectors and atmospheric fluorescence detectors. Located in Malargüe, Argentina [332] and operational since 2004, the collaboration has made a multitude of measurements of the highest energy cosmic rays. This includes measurements of the spectral distribution of cosmic rays beyond the GZK limit, anisotropy searches, as well as fits to their mass composition. Beyond the extensive cosmic ray program, Auger is able to probe extremely-high-energy neutrinos by searching for showers developing deep in the atmosphere, since showers induced by cosmic rays are likely to develop much earlier. Another possible detection channel is upgoing tau lepton showers, which are induced by Earth-skimming tau neutrino interactions near Earth's surface. In 2017, the collaboration reported a limit on the diffuse flux of high energy neutrinos between $10^8 - 10^{11}$ GeV [403] which we use to set a background-agnostic bound on $\langle \sigma v \rangle$ and $\tau$ for such energies (purple line in Figs. 6.5 and 6.3).

2. ***IceCube-EHE***: Beyond the astrophysical neutrino flux, IceCube performs searches for GZK neutrinos using a dedicated sample of events that deposit extremely high energies (EHE) in the detector. The most recent search used nine years of data and set limits on the GZK flux. We use these limits [167] to derive an bounds on the DM annihilation cross section and decay lifetime between $10^7 - 10^{11}$ GeV, represented by a light brown line in Figs. 6.5 and 6.3).

3. ***GRAND***: The Giant Radio Array for Neutrino Detection is a proposed large-scale observatory consisting of 200,000 radio antennas covering 200,000 km$^2$ near a mountain range in China. This experiment plans to use the surrounding mountains as a target for Earth-skimming tau neutrinos. After the neutrinos interact in the mountain, a tau lepton should be observed exiting the mountain and subsequently decaying in the atmosphere. The immense coverage will allow GRAND to probe GZK neutrino fluxes that are at least an order of magnitude below current limits [279]. We convert their 3-year sensitivity to the GZK neutrino flux between $10^8 - 10^{11}$ GeV into sensitivities on $\langle \sigma v \rangle$ and $\tau$ shown as a dashed navy blue line in Figs. 6.5 and 6.3.

4. ***RNO-G***: The Radio Neutrino Observatory in Greenland aims to measure the neutrino flux above $10^{16}$ eV [154]. The array of antennas to be deployed in the ice are designed to detect the Askaryan radio emission from extremely high-energy neutrinos traversing the Earth and atmosphere. The design and deployment of RNO relies upon the experience and expertise obtained in successful deployment and operation of ARA and ARIANNA [404, 162]. The plan

is to deploy 35 stations such that each station will consists of a surface array and a deep array. The surface array is going to be used for cosmic-ray detection while the deep array, benefiting from a large effective volume, will detect neutrinos.

| **Energy Range** | **Experimental Analysis** | **Directionality** | **Detected Flavor** |
|---|---|---|---|
| 2.5 – 15 MeV | Borexino [345] | × | $\bar{\nu}_e$ (IBD) |
| 8.3 – 18.3 MeV | KamLAND [349] | ✓ | $\bar{\nu}_e$ (IBD) |
| 10 – 40 MeV | JUNO [350] | ✓ | $\bar{\nu}_e$ (IBD) |
| 15 – $10^3$ MeV | SK [241] | × | $\bar{\nu}_e$ (IBD) |
| | DARWIN [401] | × | All Flavors (Coherent) |
| 0.1 – 30 GeV | DUNE [320] HK [361] | × | $\nu_e, \bar{\nu}_e, \nu_\tau, \bar{\nu}_\tau$ (CC) |
| 1 – $10^4$ GeV | SK [363, 405] | ✓ | All Flavors |
| 20 – $10^4$ GeV | IceCube [339] | ✓ | All Flavors |
| 50 – $10^5$ GeV | ANTARES [272] | ✓ | $\nu_\mu$, $\bar{\nu}_\mu$ (CC) |
| 0.2 – 100 TeV | CTA [271] | ✓ | All Flavors (Bremsstrahlung) |
| 10 – $10^4$ GeV | IC-Upgrade [388] | ✓ | All Flavors |
| > 10 PeV | IC Gen-2 [389] | ✓ | All Flavors |
| 10 – $10^4$ TeV | KM3Net [393] | ✓ | All Flavors |
| 1 – 100 PeV | TAMBO [399] | ✓ | $\nu_\tau$, $\bar{\nu}_\tau$ (CC) |
| > 100 PeV | GRAND [279] | ✓ | $\nu_\tau$, $\bar{\nu}_\tau$ (CC) |

Table 6.4: Summary of current and future experiments used to derive constraints on $\langle \sigma v \rangle$ and $\tau$ for different energy ranges. The table also indicates whether the experimental analysis used directional information and which neutrino flavors it relied on.

### Velocity-dependent annihilation

Fig. 6.6 shows the corresponding limits for $p$-wave annihilation, and Fig. 6.7 provides limits on $d$-wave annihilation. In these cases, we follow the procedures outlined in Sec. 6.4, to reweight the astrophysical portion of the flux prediction (Eqs. (6.4) and (6.3)) to account for the dark matter velocity dispersion. We do this for all-sky searches since analyses where the angular distribution of the neutrinos has been taken into account are not easily re-scaled when considering the velocity distribution of DM particles within the halo. Similarly, all the constraints taken from the literature are re-scaled using our choice of halo parameters (see Tbl. 6.1 for halo parameters and $J$-factor for

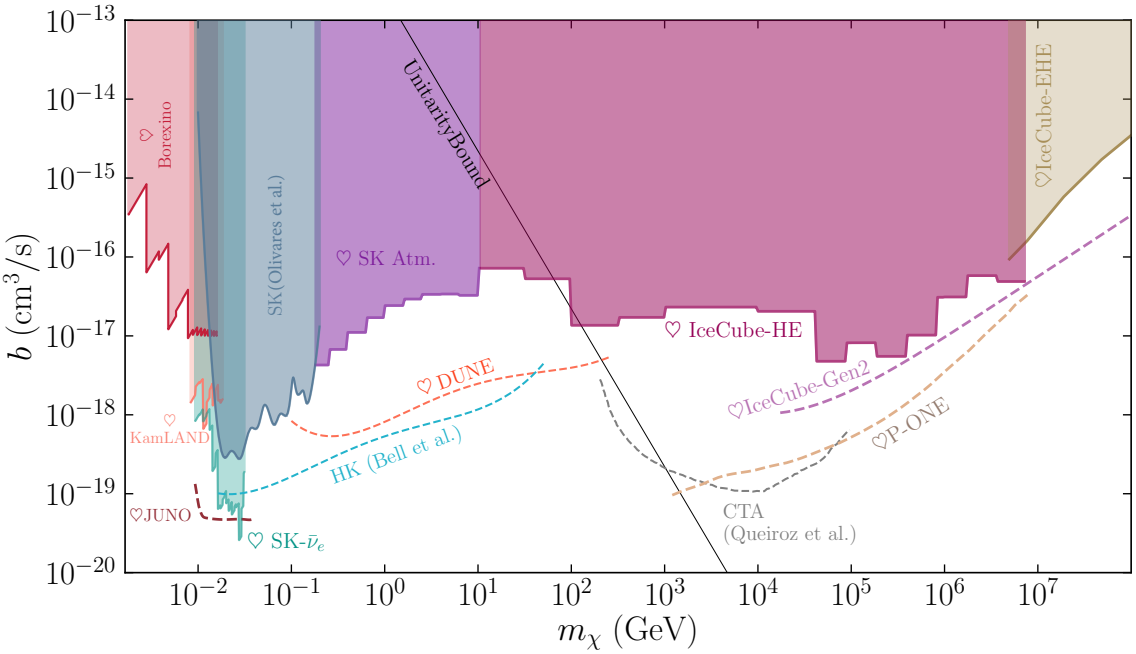



Figure 6.6: Limits on $p$-wave, $\langle\sigma v\rangle = b(v/c)^2$, velocity-dependent annihilation cross-section of dark matter to two neutrinos. The cross section needed to explain the observed abundance for thermal DM is $\langle\sigma v_r\rangle = 6 \times 10^{-26}\ \mathrm{cm}^3/\mathrm{s}$.

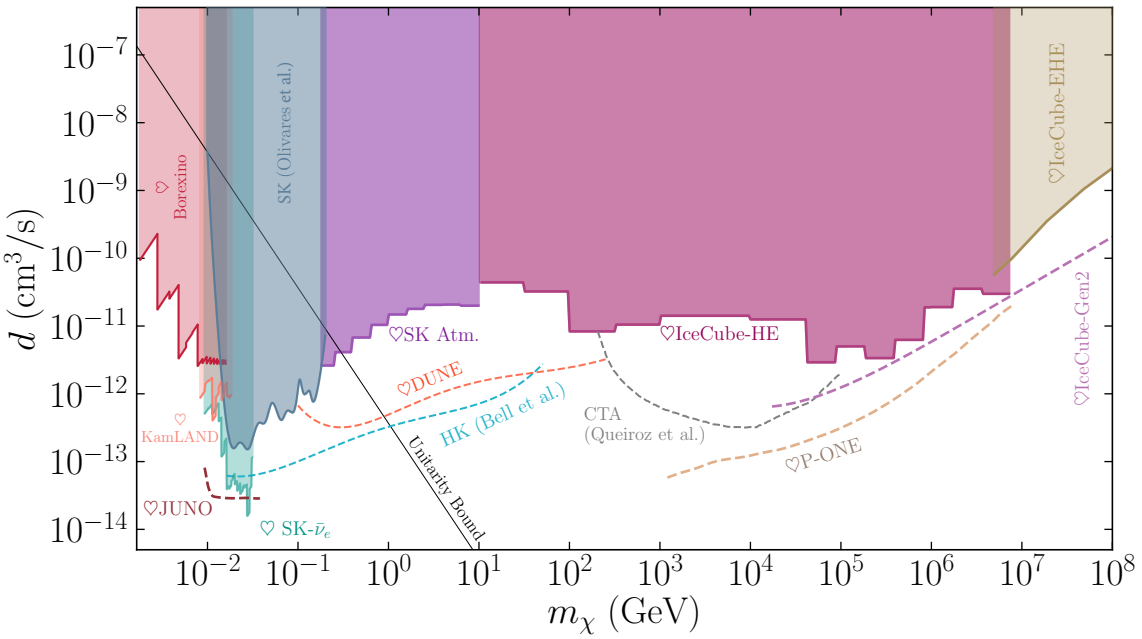



Figure 6.7: Limits on the annihilation of neutrinos to dark matter through a $d$-wave process $\langle\sigma v\rangle = d(v/c)^4$.

the different analyses in the literature). Unsurprisingly, the limits on $\langle\sigma v\rangle$ are much weaker for $p-$ and $d-$wave processes due to the strong velocity suppression. In contrast to the $s-$wave case, where the smallest halos tend to dominate the expected signal, velocity-suppressed annihilation is strongest in the largest DM halos where dispersion velocities are higher. These limits are thus insensitive to the value of the minimum halo mass $M_{min}$. However, the constraints from annihilation in the Milky Way halo remain dominant over the extragalactic contribution.

### Dark matter halo uncertainties

As previously mentioned, a major source of uncertainty comes from the spatial dark matter distribution, because of the $n_\chi^2$ dependence in the annihilation signal. For Galactic constraints, this is mainly reflected by uncertainties in the Milky Way dark matter distribution. For extragalactic constraints, we focus on the shape of the halo mass function and the minimum dark matter mass, which determines how far down extrapolations of the HMF must go to account for the total DM contribution.

***Milky Way halo shape parameters:*** To quantify the effect of the uncertainty on the MW halo shape parameters, we use the code provided by the authors of [250], which computes the log-likelihood as a function of halo shape parameters $\{\rho_0, r_s, R_0, \gamma\}$, given observed stellar kinematics data. We profile over the 4 degrees of freedom, modifying the code to account for GRAVITY measurements of $R_0$, and obtain 68% and 95% C.L. ranges on the $J$-factors which we propagate to a range on $\langle\sigma v\rangle$ for the Borexino, SK, and IceCube analyses. These are shown as dark and light bands, respectively, in Fig. 6.8.

***Halo Mass Function uncertainties:*** The largest contributions to uncertainties in the cosmological limits come from 1) the choice of HMF parametrization, and 2) the choice of minimum halo mass, $M_{min}$. In our analyses we have employed the simulation-driven HMF fit by Watson et al. [255]. Fig. 6.9 shows the boost factor $G(z)$ defined in Eq. (6.3), for four different parametrizations from the literature: the analytic Press & Schechter formalism [406, 407], Sheth & Tormen [408, 409], and Tinker [410]. The width of the bands comes from varying the minimum halo mass from $10^{-3}$ to $10^{-9}\,M_\odot$. The band labeled "Extragalactic" in Fig. 6.8 shows how this range propagates through to the cross section constraints. Since there is no way of statistically quantifying the error on the HMF and minimum halo mass, we choose the most conservative scenario $M_{min} = 10^{-3}\,M_\odot$ for our choice of HMF, corresponding to the solid magenta line in Fig. 6.8.

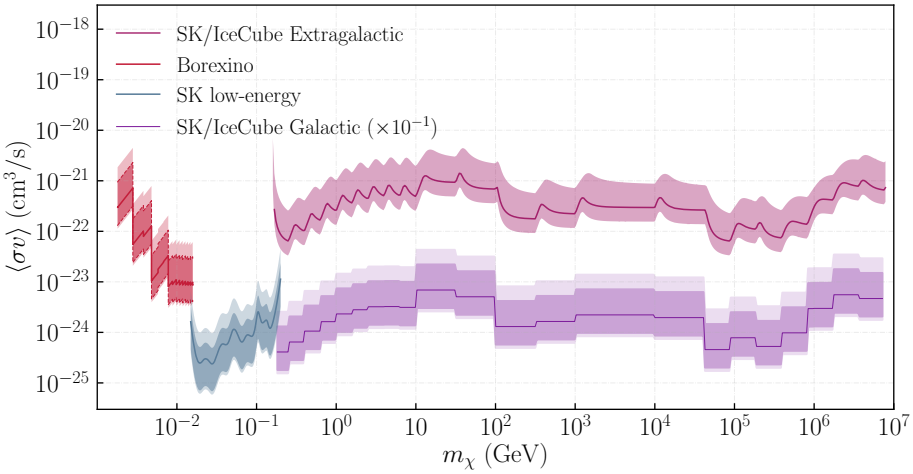



Figure 6.8: Uncertainties on the $s$-wave annihilation cross section for a subset of our results. Solid lines correspond to the limits discussed in Sec. 6.7. For all Galactic limits, namely Borexino (red, leftmost), Super-Kamiokande low-energy (grey, scond region from left), Super-Kamiokande and IceCube (lower, rightmost), the 68% (dark bands) and 95% (light bands) uncertainties arise from the allowed variation on the dark matter distribution in the Milky Way, assuming a generalized NFW profile. The width of the uncertainty band for the extragalactic limits (upper, rightmost), obtained by comparing to the unfolded neutrino flux from IceCube and Super-Kamiokande, is dominated by the choice of the minimum halo mass, $M_{min}$, although it includes the uncertainty in the choice of HMF $dn/dM$, see Fig. 6.9. For our nominal choice of HMF, we choose the value of $M_{min}$ that yields the weakest constraint.

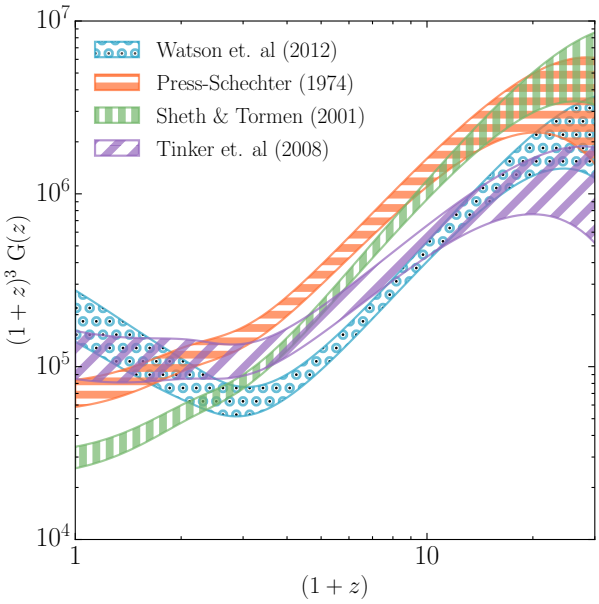



Figure 6.9: The halo boost factor $G(z)$ as a function of redshift for several parametrizations of the HMF $dn/dM$. Our extragalactic constraints use Watson *et. al* [255]. The bands represent varying choices of minimum halo mass, from $10^{-3}$ to $10^{-9}$ solar masses. Fig. 6.8 shows the effect of choosing a different parametrization on the limits.

# Chapter 7

# Conclusion

"It's still all good..
And if you don't know, now you know"

The Notorious B.I.G

The work has only just begun, and new physics in the neutrino sector remains elusive. But we are getting ever-closer, and discoveries are undoubtedly around the corner. The mechanism by which neutrinos obtain their mass remains the most promising avenue to discovering physics beyond the standard model, and astrophysical neutrinos offer a unique probe of likely scenarios. Flavor composition measurements are a smoking gun signature to BSM effects. And while the current generation of neutrino observatories lacks the definitive sensitivity, the next generation of astrophysical neutrino detectors will be able to make definitive claims using astrophysical flavor composition. We have shown in Chapters 3-5 that better modeling of the tau flavor significantly improves the sensitivity to transient sources, GZK fluxes, and allows accurate background predictions for the next generation.

The standard sources of astrophysical neutrinos are also becoming clearer, with several hints pointing to gamma-ray obscured AGN cores, as we have shown in Chapter 2. Identifying the sources will open up a multitude of new physics tests like *e.g.* neutrino decay, non-standard neutrino interactions, a fourth neutrino state, right-handed neutrinos, lorentz-invariance violation, CPT violation and more.

Beyond that, subtracting the known sources from the total measured astrophysical flux improves our sensitivity to neutrinos produced by exotic sources like dark matter. If dark matter is of a corpuscular nature, it is not an illogical leap to assume that the mechanism responsible for dark matter and neutrino masses could be shared. Again, astrophysical neutrinos allow us to ask definitive

questions about connections between the two. As we have shown in Chapter 6, indirect dark matter searches with neutrinos are promising. Indeed, we reach the thermal relic abundance cross section in the neutrino sector for the first time at dark matter masses of $\sim$ 10 MeV using data from Super-Kamiokande. We are entering an era of precision neutrino astrophysics, and we look forward to the day when questions asked in this thesis are mundane and the answers are taken for granted, and in retrospect seem quite obvious.

# List of Figures

1.1 Astrophysical neutrino flavor potential. . . . . 3

2.1 Variability measures for all sources in the CGRaBS catalog above −20° declination. . . 10
2.2 Skymaps of CGRaBS sources. . . . . 11
2.3 Stacking analysis weights. . . . . 13
2.4 Background TS distributions . . . . . 14
2.5 Analysis sensitivity and discovery potential . . . . . 14
2.6 Unblinded test-statistic distribution. . . . . 15
2.7 Unblinded log-likelihood scan. . . . . 15

3.1 Schematic of lepton propagation through Earth. . . . . 18
3.2 The neutrino-proton cross section as a function of the neutrino energy. . . . . 20
3.3 Distribution of final tau energies and total distance traveled before decay for several initial tau energies. . . . . 22
3.4 Initial polarization of taus produced from charged-current scattering. . . . . 24
3.5 Average polarization of outgoing tau after a single interaction. . . . . 25
3.6 Artistic rendition of relevant interactions for tau appearance. . . . . 26
3.7 Neutrino-nucleus cross sections per nucleon as a function of the neutrino energy assuming an isoscalar target. . . . . 27
3.8 Components of tau-neutrino flux at the detector integrated over different angular regions. 28
3.9 Primary neutrino fluxes and appearing tau-neutrino component. . . . . 30
3.10 Rates of tau-neutrino events in ten years of HESE. . . . . 31
3.11 Tau-lepton fluxes arriving at Earth-skimming experiments. . . . . 33

4.1 Flowchart of the `TauRunner` propagation algorithm . . . . . 36

4.2 Schematic of `TauRunner` geometry as contained within the `Track` class. . . . . 38
4.3 Runtime per $\nu_\tau$ event in `TauRunner`. . . . . 41
4.4 Outgoing $\nu_\tau$ distributions for an incident $E^{-1}$ power-law flux. . . . . 43
4.5 Outgoing energies of secondary tau neutrinos as a function of nadir angle for EeV neutrinos 44
4.6 Comparison of `TauRunner` and `NuPropEarth` output. . . . . 44
4.7 Tau exit probability with `TauRunner` compared with other software packages. . . . . 45
4.8 Distribution of final tau polarization before decay, for monoenergetic taus in rock. . . . 46
4.9 Effect of tau depolarization on outgoing tau neutrinos for multiple starting energies. . . 47
4.10 Impact of tau depolarization on the tau exit probability. . . . . 48
4.11 Final energy of Earth-emerging taus and their polarization starting from a neutrino beam. 49

5.1 Per flavor neutrino fluxes from 1 TeV to 10 EeV, integrated over the northern sky. . . . 52
5.2 2D signal of GZK neutrinos at IceCube as a function of energy and zenith angle. . . . . 54
5.3 Per flavor neutrino fluxes from 1 TeV to 10 EeV, integrated over various zenith bands in the Northern Sky. . . . . 55
5.4 Improved point source differential sensitivity using $\nu_\tau$. . . . . 58
5.5 Tau neutrino exit probability at or above ANITA detection threshold. . . . . 61
5.6 ANITA constraints using HESE. . . . . 62
5.7 Effective area of the IceCube event selection to muons from the northern sky. . . . . 65
5.8 Normalized cumulative distributions for Earth-emerging tau-leptons. . . . . 66
5.9 Transient point-source upper limits on ANITA anomalous events. . . . . 67

6.1 Examples of neutrino fluxes produced by dark matter annihilation overlayed on the observed neutrino distributions. . . . . 76
6.2 The landscape of dark matter annihilation into neutrinos up to $10^8$ GeV. . . . . 79
6.3 Limits on the lifetime of dark matter decay to neutrino pairs . . . . . 80
6.4 The landscape of sub-GeV dark matter annihilation into neutrinos. . . . . 93
6.5 The landscape of supra-TeV dark matter annihilation into neutrinos. . . . . 94
6.6 Limits on $p$-wave, $\langle\sigma v\rangle = b(v/c)^2$, velocity-dependent annihilation cross-section of dark matter to two neutrinos. . . . . 103
6.7 Limits on the annihilation of neutrinos to dark matter through a $d$-wave process $\langle\sigma v\rangle = d(v/c)^4$. . . . . 103

6.8 Uncertainties on the $s$-wave annihilation cross section for a subset of our results. . . . . 105

6.9 The halo boost factor $G(z)$ as a function of redshift for several parametrizations of the HMF $dn/dM$. . . . . 105

# List of Tables

3.1 Impact of appearing tau-neutrino component on cosmic tau-neutrino discovery. . . . . 30

4.1 Software comparison table . . . . . 45

5.1 Sensitivity improvement for GZK neutrino searches. . . . . 56
5.2 Summary of the neutrino candidate event properties from the first three flights of ANITA. 60
5.3 Summary of ANITA followup analysis with IceCube. . . . . 64

6.1 J-factors used to constrain DM annihilation for different experiments and their associated halo parameters. . . . . 75
6.2 Coefficients of the polynomial fit to velocity dependent halo boost factors. . . . . 78
6.3 D-factors for different experiments used to constrain dark matter decay and their associated halo parameters. . . . . 81
6.4 Summary of current and future experiments used to derive constraints on $\langle \sigma v \rangle$ and $\tau$ for different energy ranges. . . . . 102

# Bibliography


[1] W. Pauli, *Dear radioactive ladies and gentlemen*, *Phys. Today* **31N9** (1978) 27.

[2] C.L. Cowan, F. Reines, F.B. Harrison, H.W. Kruse and A.D. McGuire, *Detection of the free neutrino: A Confirmation*, *Science* **124** (1956) 103.

[3] J.N. Bahcall, *Solar neutrinos. I: Theoretical*, *Phys. Rev. Lett.* **12** (1964) 300.

[4] R. Davis, D.S. Harmer and K.C. Hoffman, *Search for neutrinos from the sun*, *Phys. Rev. Lett.* **20** (1968) 1205.

[5] B.T. Cleveland, T. Daily, R. Davis, Jr., J.R. Distel, K. Lande, C.K. Lee et al., *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, *Astrophys. J.* **496** (1998) 505.

[6] Super-Kamiokande collaboration, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562 [hep-ex/9807003].

[7] SNO collaboration, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **89** (2002) 011301 [nucl-ex/0204008].

[8] LSND collaboration, *Evidence for neutrino oscillations from the observation of $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam*, *Phys. Rev. D* **64** (2001) 112007 [hep-ex/0104049].

[9] MiniBooNE collaboration, *A Search for Electron Neutrino Appearance at the $\Delta m^2 \sim 1eV^2$ Scale*, *Phys. Rev. Lett.* **98** (2007) 231801 [0704.1500].

[10] IceCube collaboration, *Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data*, *Phys. Rev. Lett.* **113** (2014) 101101 [1405.5303].

[11] IceCube collaboration, *Constraints on Galactic Neutrino Emission with Seven Years of IceCube Data*, *Astrophys. J.* **849** (2017) 67 [1707.03416].

[12] JUNO collaboration, *Neutrino Physics with JUNO*, *J. Phys. G* **43** (2016) 030401 [1507.05613].

[13] Hyper-Kamiokande collaboration, *Hyper-Kamiokande Design Report*, 1805.04163.

[14] DUNE collaboration, *Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE): Conceptual Design Report, Volume 2: The Physics Program for DUNE at LBNF*, 1512.06148.

[15] IceCube collaboration, *The IceCube Upgrade – Design and Science Goals*, in *HAWC Contributions to the 36th International Cosmic Ray Conference (ICRC2019)*, 2019 [1908.09441].

[16] N. Song, S.W. Li, C.A. Argüelles, M. Bustamante and A.C. Vincent, *The Future of High-Energy Astrophysical Neutrino Flavor Measurements*, *JCAP* **04** (2021) 054 [2012.12893].

[17] E. Ma, *Verifiable radiative seesaw mechanism of neutrino mass and dark matter*, *Phys. Rev. D* **73** (2006) 077301 [hep-ph/0601225].

[18] M. Blennow, E. Fernandez-Martinez, A.O.-D. Campo, S. Pascoli, S. Rosauro-Alcaraz and A.V. Titov, *Neutrino Portals to Dark Matter*, 1903.00006.

[19] C. Boehm, Y. Farzan, T. Hambye, S. Palomares-Ruiz and S. Pascoli, *Is it possible to explain neutrino masses with scalar dark matter?*, *Phys. Rev.* **D77** (2008) 043516 [hep-ph/0612228].

[20] Y. Farzan and E. Ma, *Dirac neutrino mass generation from dark matter*, *Phys. Rev.* **D86** (2012) 033007 [1204.4890].

[21] M. Escudero, N. Rius and V. Sanz, *Sterile neutrino portal to Dark Matter I: The $U(1)_{B-L}$ case*, *JHEP* **02** (2017) 045 [1606.01258].

[22] M. Escudero, N. Rius and V. Sanz, *Sterile Neutrino portal to Dark Matter II: Exact Dark symmetry*, *Eur. Phys. J.* **C77** (2017) 397 [1607.02373].

[23] C. Hagedorn, J. Herrero-García, E. Molinaro and M.A. Schmidt, *Phenomenology of the Generalised Scotogenic Model with Fermionic Dark Matter*, *JHEP* **11** (2018) 103 [1804.04117].

[24] J.B.G. Alvey and M. Fairbairn, *Linking Scalar Dark Matter and Neutrino Masses with IceCube 170922A*, 1902.01450.

[25] H.H. Patel, S. Profumo and B. Shakya, *Loop Dominated Signals from Neutrino Portal Dark Matter*, 1912.05581.

[26] S. Baumholzer, V. Brdar, P. Schwaller and A. Segner, *Shining Light on the Scotogenic Model: Interplay of Colliders, Cosmology and Astrophysics*, 1912.08215.

[27] E. Bertuzzo, S. Jana, P.A.N. Machado and R. Zukanovich Funchal, *Dark Neutrino Portal to Explain MiniBooNE excess*, *Phys. Rev. Lett.* **121** (2018) 241801 [1807.09877].

[28] P. Ballett, S. Pascoli and M. Ross-Lonergan, *U(1)' mediated decays of heavy sterile neutrinos in MiniBooNE*, *Phys. Rev.* **D99** (2019) 071701 [1808.02915].

[29] P. Ballett, M. Hostert and S. Pascoli, *Neutrino Masses from a Dark Neutrino Sector below the Electroweak Scale*, *Phys. Rev.* **D99** (2019) 091701 [1903.07590].

[30] P. Ballett, M. Hostert and S. Pascoli, *Dark Neutrinos and a Three Portal Connection to the Standard Model*, 1903.07589.

[31] F. Halzen and A.D. Martin, *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS* (1984).

[32] Particle Data Group collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.

[33] IceCube collaboration, *The IceCube Neutrino Observatory: Instrumentation and Online Systems*, *JINST* **12** (2017) P03012 [1612.05093].

[34] A.G. Soto, P. Zhelnin, I. Safa and C.A. Argüelles, *Tau Appearance from High-Energy Neutrino Interactions*, 2112.06937.

[35] C.A. Argüelles, D. Garg, S. Patel, M.H. Reno and I. Safa, *Tau depolarization at very high energies for neutrino telescopes*, 2205.05629.

[36] I. Safa, J. Lazar, A. Pizzuto, O. Vasquez, C.A. Argüelles and J. Vandenbroucke, *TauRunner: A public Python program to propagate neutral and charged leptons*, *Comput. Phys. Commun.* **278** (2022) 108422 [2110.14662].

[37] C.A. Argüelles, F. Halzen, A. Kheirandish and I. Safa, *PeV Tau Neutrinos to Unveil Ultra-High-Energy Sources*, 2203.13827.

[38] I. Safa, A. Pizzuto, C.A. Argüelles, F. Halzen, R. Hussain, A. Kheirandish et al., *Observing EeV neutrinos through Earth: GZK and the anomalous ANITA events*, *JCAP* **01** (2020) 012 [1909.10487].

[39] M.G.A. et al, *A search for IceCube events in the direction of ANITA neutrino candidates*, *The Astrophysical Journal* **892** (2020) 53.

[40] C.A. Argüelles, A. Diaz, A. Kheirandish, A. Olivares-Del-Campo, I. Safa and A.C. Vincent, *Dark matter annihilation to neutrinos*, *Rev. Mod. Phys.* **93** (2021) 035007 [1912.09486].

[41] IceCube collaboration, *Search for Multi-flare Neutrino Emissions in 10 yr of IceCube Data from a Catalog of Sources*, *Astrophys. J. Lett.* **920** (2021) L45 [2109.05818].

[42] IceCube collaboration, *Search for steady point-like sources in the astrophysical muon neutrino flux with 8 years of IceCube data*, *Eur. Phys. J. C* **79** (2019) 234 [1811.07979].

[43] R. Abbasi et al., *Search for Astrophysical Neutrinos from 1FLE Blazars with IceCube*, 2207.04946.

[44] IceCube collaboration, *The contribution of Fermi-2LAC blazars to the diffuse TeV-PeV neutrino flux*, *Astrophys. J.* **835** (2017) 45 [1611.03874].

[45] IceCube collaboration, *Searches for steady neutrino emission from 3FHL blazars using eight years of IceCube data from the Northern hemisphere*, *PoS* **ICRC2019** (2020) 916 [1908.08458].

[46] IceCube collaboration, *The IceCube Realtime Alert System*, *Astropart. Phys.* **92** (2017) 30 [1612.06028].

[47] IceCube, Fermi-LAT, MAGIC, AGILE, ASAS-SN, HAWC, H.E.S.S., INTEGRAL, Kanata, Kiso, Kapteyn, Liverpool Telescope, Subaru, Swift NuSTAR, VERITAS, VLA/17B-403 collaboration, *Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A*, *Science* **361** (2018) eaat1378 [1807.08816].

[48] IceCube collaboration, *Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert*, *Science* **361** (2018) 147 [1807.08794].

[49] Fermi-LAT, ASAS-SN, IceCube collaboration, *Investigation of two Fermi-LAT gamma-ray blazars coincident with high-energy neutrinos detected by IceCube*, *Astrophys. J.* **880** (2019) 880:103 [1901.10806].

[50] E. Kun, I. Bartos, J. Becker Tjus, P.L. Biermann, F. Halzen and G. Mező, *Cosmic Neutrinos from Temporarily Gamma-suppressed Blazars*, *Astrophys. J. Lett.* **911** (2021) L18 [2009.09792].

[51] L. Nellen, K. Mannheim and P.L. Biermann, *Neutrino production through hadronic cascades in AGN accretion disks*, *Phys. Rev. D* **47** (1993) 5270 [hep-ph/9211257].

[52] O. de Bruijn, I. Bartos, P.L. Biermann and J. Becker Tjus, *Recurrent Neutrino Emission from Supermassive Black Hole Mergers*, *Astrophys. J. Lett.* **905** (2020) L13 [2006.11288].

[53] K. Murase, S.S. Kimura, B.T. Zhang, F. Oikonomou and M. Petropoulou, *High-Energy Neutrino and Gamma-Ray Emission from Tidal Disruption Events*, *Astrophys. J.* **902** (2020) 108 [2005.08937].

[54] S.E. Healey, R.W. Romani, G. Cotter, P.F. Michelson, E.F. Schlafly, A.C.S. Readhead et al., *CGRaBS: An All-Sky Survey of Gamma-Ray Blazar Candidates*, *Astrophys. J. Suppl.* **175** (2008) 97 [0709.1735].

[55] J.L. Richards et al., *Blazars in the Fermi Era: The OVRO 40-m Telescope Monitoring Program*, *Astrophys. J. Suppl.* **194** (2011) 29 [1011.3111].

[56] K. Murase, Y. Inoue and C.D. Dermer, *Diffuse Neutrino Intensity from the Inner Jets of Active Galactic Nuclei: Impacts of External Photon Fields and the Blazar Sequence*, *Phys. Rev. D* **90** (2014) 023007 [1403.4089].

[57] R. D'Abrusco, F. Massaro, A. Paggi, H.A. Smith, N. Masetti, M. Landoni et al., *The WISE Blazar-Like Radio-Loud Sources: An All-Sky Catalog Of Candidate Gamma-Ray Blazars*, *Astrophys. J. Suppl.* **215** (2014) 14 [1410.0029].

[58] XMM collaboration, *XMM-Newton observatory. I. The spacecraft and operations.*, *Astron. Astrophys.* **365** (2001) L1.

[59] J. Braun, J. Dumm, F. De Palma, C. Finley, A. Karle and T. Montaruli, *Methods for point source analysis in high energy neutrino telescopes*, *Astropart. Phys.* **29** (2008) 299 [0801.1604].

[60] IceCube collaboration, *Search for neutrino emission from cores of active galactic nuclei*, *Phys. Rev. D* **106** (2022) 022005 [2111.10169].

[61] A. Plavin, Y.Y. Kovalev, Y.A. Kovalev and S. Troitsky, *Observational Evidence for the Origin of High-energy Neutrinos in Parsec-scale Nuclei of Radio-bright Active Galaxies*, *Astrophys. J.* **894** (2020) 101 [2001.00930].

[62] A.V. Plavin, Y.Y. Kovalev, Y.A. Kovalev and S.V. Troitsky, *Directional Association of TeV to PeV Astrophysical Neutrinos with Radio Blazars*, *Astrophys. J.* **908** (2021) 157 [2009.08914].

[63] M.C. Gonzalez-Garcia, F. Halzen and M. Maltoni, *Physics reach of high-energy and high-statistics icecube atmospheric neutrino data*, *Phys. Rev.* **D71** (2005) 093010 [hep-ph/0502223].

[64] R. Gandhi, C. Quigg, M.H. Reno and I. Sarcevic, *Neutrino interactions at ultrahigh-energies*, *Phys. Rev. D* **58** (1998) 093009 [hep-ph/9807264].

[65] M. Cirelli, N. Fornengo, T. Montaruli, I.A. Sokalski, A. Strumia and F. Vissani, *Spectra of neutrinos from dark matter annihilations*, *Nucl. Phys.* **B727** (2005) 99 [hep-ph/0506298].

[66] M. Blennow, J. Edsjo and T. Ohlsson, *Neutrinos from WIMP annihilations using a full three-flavor Monte Carlo*, *JCAP* **0801** (2008) 021 [0709.3898].

[67] J. Becker Tjus, B. Eichmann, F. Halzen, A. Kheirandish and S.M. Saba, *High-energy neutrinos from radio galaxies*, *Phys. Rev.* **D89** (2014) 123005 [1406.0506].

[68] J.H. Koehne, K. Frantzen, M. Schmitz, T. Fuchs, W. Rhode, D. Chirkin et al., *PROPOSAL: A tool for propagation of charged leptons*, *Comput. Phys. Commun.* **184** (2013) 2070.

[69] A.C. Vincent, C.A. Argüelles and A. Kheirandish, *High-energy neutrino attenuation in the Earth and its associated uncertainties*, *JCAP* **1711** (2017) 012 [1706.09895].

[70] J. Alvarez-Muñiz, W.R. Carvalho, A.L. Cummings, K. Payet, A. Romero-Wolf, H. Schoorlemmer et al., *Comprehensive approach to tau-lepton production by high-energy tau neutrinos propagating through the Earth*, *Phys. Rev.* **D97** (2018) 023021 [1707.00334].

[71] M.H. Reno, J.F. Krizmanic and T.M. Venters, *Cosmic tau neutrino detection via Cherenkov signals from air showers from Earth-emerging taus*, 1902.11287.

[72] Particle Data Group collaboration, *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) 100001.

[73] MiniBooNE collaboration, *First Measurement of the Muon Neutrino Charged Current Quasielastic Double Differential Cross Section*, *Phys. Rev.* **D81** (2010) 092005 [1002.2680].

[74] NuTeV collaboration, *Precise measurement of neutrino and anti-neutrino differential cross sections*, *Phys. Rev.* **D74** (2006) 012008 [hep-ex/0509010].

[75] P. Vogel and J.F. Beacom, *Angular distribution of neutron inverse beta decay, anti-neutrino(e) + p —> e+ + n*, *Phys. Rev.* **D60** (1999) 053003 [hep-ph/9903554].

[76] A. Kurylov, M.J. Ramsey-Musolf and P. Vogel, *Radiative corrections to low-energy neutrino reactions*, *Phys. Rev.* **C67** (2003) 035502 [hep-ph/0211306].

[77] BOREXINO collaboration, *Comprehensive measurement of pp-chain solar neutrinos*, *Nature* **562** (2018) 505.

[78] Super-Kamiokande collaboration, *Measurement of the tau neutrino cross section in atmospheric neutrino oscillations with Super-Kamiokande*, *Phys. Rev.* **D98** (2018) 052006 [1711.09436].

[79] IceCube collaboration, *Measurement of the multi-TeV neutrino cross section with IceCube using Earth absorption*, *Nature* **551** (2017) 596 [1711.08119].

[80] IceCube collaboration, *Measurement of the high-energy all-flavor neutrino-nucleon cross section with IceCube*, in *36th International Cosmic Ray Conference (ICRC 2019) Madison, Wisconsin, USA, July 24-August 1, 2019*, 2019 [1908.07027].

[81] J.A. Formaggio and G.P. Zeller, *From eV to EeV: Neutrino Cross Sections Across Energy Scales*, *Rev. Mod. Phys.* **84** (2012) 1307 [1305.7513].

[82] T. Katori and M. Martini, *Neutrino–nucleus cross sections for oscillation experiments*, *J. Phys.* **G45** (2018) 013001 [1611.07770].

[83] J.L. Feng, I. Galon, F. Kling and S. Trojanowski, *ForwArd Search ExpeRiment at the LHC*, *Phys. Rev.* **D97** (2018) 035001 [1708.09389].

[84] FASER collaboration, *Detecting and Studying High-Energy Collider Neutrinos with FASER at the LHC*, 1908.02310.

[85] A. Cooper-Sarkar, P. Mertsch and S. Sarkar, *The high energy neutrino cross-section in the Standard Model and its uncertainty*, *JHEP* **08** (2011) 042 [1106.3723].

[86] A. Garcia, R. Gauld, A. Heijboer and J. Rojo, *Complete predictions for high-energy neutrino propagation in matter*, *JCAP* **09** (2020) 025 [2004.04756].

[87] M. Froissart, *Asymptotic behavior and subtractions in the Mandelstam representation*, Phys. Rev. **123** (1961) 1053.

[88] C.A. Argüelles, F. Halzen, L. Wille, M. Kroll and M.H. Reno, *High-energy behavior of photon, neutrino, and proton cross sections*, Phys. Rev. D **92** (2015) 074040 [1504.06639].

[89] M.M. Block and F. Halzen, *Experimental Confirmation that the Proton is Asymptotically a Black Disk*, Phys. Rev. Lett. **107** (2011) 212002 [1109.2041].

[90] Pierre Auger collaboration, *Measurement of the proton-air cross-section at $\sqrt{s} = 57$ TeV with the Pierre Auger Observatory*, Phys. Rev. Lett. **109** (2012) 062002 [1208.1520].

[91] TOTEM collaboration, *Measurement of proton-proton elastic scattering and total cross-section at S**(1/2) = 7-TeV*, EPL **101** (2013) 21002.

[92] Particle Data Group collaboration, *Review of particle properties. Particle Data Group*, Phys. Rev. D **50** (1994) 1173.

[93] J.-H. Koehne, K. Frantzen, M. Schmitz, T. Fuchs, W. Rhode, D. Chirkin et al., *Proposal: A tool for propagation of charged leptons*, Computer Physics Communications **184** (2013) 2070.

[94] S.I. Dutta, M.H. Reno and I. Sarcevic, *Secondary neutrinos from tau neutrino interactions in earth*, Phys. Rev. D **66** (2002) 077302 [hep-ph/0207344].

[95] S. Dutta, M. Reno, I. Sarcevic and D. Seckel, *Propagation of muons and taus at high-energies*, Phys. Rev. D **63** (2001) 094020 [hep-ph/0012350].

[96] K. Hagiwara, K. Mawatari and H. Yokoya, *Tau polarization in tau neutrino nucleon scattering*, Nucl. Phys. **B668** (2003) 364 [hep-ph/0305324].

[97] L. Pasquali and M.H. Reno, *Tau-neutrino fluxes from atmospheric charm*, Phys. Rev. D **59** (1999) 093003 [hep-ph/9811268].

[98] A. Bhattacharya, R. Enberg, Y.S. Jeong, C.S. Kim, M.H. Reno, I. Sarcevic et al., *Prompt atmospheric neutrino fluxes: perturbative QCD models and nuclear effects*, JHEP **11** (2016) 167 [1607.00193].

[99] K. Payet, *Polarization of tau leptons produced in ultra-high energy neutrino-nucleon scattering*, 0807.1236.

[100] D. Seckel, *Neutrino photon reactions in astrophysics and cosmology*, Phys. Rev. Lett. **80** (1998) 900 [hep-ph/9709290].

[101] I. Alikhanov, *Hidden Glashow resonance in neutrino–nucleus collisions*, Phys. Lett. B **756** (2016) 247 [1503.08817].

[102] V. Barger, E. Basso, Y. Gao and W.-Y. Keung, *Neutrino signals in IceCube from weak production of top and charm quarks*, Phys. Rev. D **95** (2017) 093002 [1611.00773].

[103] B. Zhou and J.F. Beacom, *Neutrino-nucleus cross sections for W-boson and trident production*, Phys. Rev. D **101** (2020) 036011 [1910.08090].

[104] B. Zhou and J.F. Beacom, *W-boson and trident production in TeV–PeV neutrino observatories*, Phys. Rev. D **101** (2020) 036010 [1910.10720].

[105] S.L. Glashow, *Resonant Scattering of Antineutrinos*, Phys. Rev. **118** (1960) 316.

[106] R. Gauld, *Precise predictions for multi-TeV and PeV energy neutrino scattering rates*, Phys. Rev. D **100** (2019) 091301 [1905.03792].

[107] G.-y. Huang and Q. Liu, *Hunting the Glashow Resonance with PeV Neutrino Telescopes*, *JCAP* **03** (2020) 005 [1912.02976].

[108] V. Bertone, R. Gauld and J. Rojo, *Neutrino Telescopes as QCD Microscopes*, *JHEP* **01** (2019) 217 [1808.02034].

[109] A. Cooper-Sarkar, P. Mertsch and S. Sarkar, *The high energy neutrino cross-section in the Standard Model and its uncertainty*, *JHEP* **08** (2011) 042 [1106.3723].

[110] A. Cooper-Sarkar and S. Sarkar, *Predictions for high energy neutrino cross-sections from the ZEUS global PDF fits*, *JHEP* **01** (2008) 075 [0710.5303].

[111] M. Gluck, P. Jimenez-Delgado and E. Reya, *On the charged current neutrino-nucleon total cross section at high energies*, *Phys. Rev. D* **81** (2010) 097501 [1003.3168].

[112] A. Connolly, R.S. Thorne and D. Waters, *Calculation of High Energy Neutrino-Nucleon Cross Sections and Uncertainties Using the MSTW Parton Distribution Functions and Implications for Future Experiments*, *Phys. Rev.* **D83** (2011) 113009 [1102.0691].

[113] J.L. Albacete, J.I. Illana and A. Soto-Ontoso, *Neutrino-nucleon cross section at ultrahigh energy and its astrophysical implications*, *Phys. Rev. D* **92** (2015) 014027 [1505.06583].

[114] V.P. Goncalves and P. Hepp, *A comparative study of the neutrino-nucleon cross section at ultra high energies*, *Phys. Rev. D* **83** (2011) 014014 [1011.2718].

[115] H1, ZEUS collaboration, *Proton Structure from HERA to LHC*, in *40th International Symposium on Multiparticle Dynamics*, 12, 2010 [1012.1438].

[116] IceCube collaboration, *The IceCube high-energy starting event sample: Description and flux characterization with 7.5 years of data*, *Phys. Rev. D* **104** (2021) 022002 [2011.03545].

[117] IceCube collaboration, *Measurement of the Diffuse Astrophysical Muon-Neutrino Spectrum with Ten Years of IceCube Data*, *PoS* **ICRC2019** (2020) 1017 [1908.09551].

[118] A. Fedynitch, R. Engel, T.K. Gaisser, F. Riehn and T. Stanev, *Calculation of conventional and prompt lepton fluxes at very high energy*, *EPJ Web Conf.* **99** (2015) 08001 [1503.00544].

[119] IceCube collaboration, *Measurements using the inelasticity distribution of multi-TeV neutrino interactions in IceCube*, *Phys. Rev. D* **99** (2019) 032004 [1808.07629].

[120] IceCube collaboration, *A combined maximum-likelihood analysis of the high-energy astrophysical neutrino flux measured with IceCube*, *Astrophys. J.* **809** (2015) 98 [1507.03991].

[121] IceCube collaboration, *Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector*, *Science* **342** (2013) 1242856 [1311.5238].

[122] IceCube collaboration, *Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data*, *Phys. Rev. Lett.* **113** (2014) 101101 [1405.5303].

[123] IceCube collaboration, *Atmospheric and astrophysical neutrinos above 1 TeV interacting in IceCube*, *Phys. Rev. D* **91** (2015) 022001 [1410.1749].

[124] IceCube collaboration, *The IceCube Neutrino Observatory - Contributions to ICRC 2015 Part II: Atmospheric and Astrophysical Diffuse Neutrino Searches of All Flavors*, in *34th International Cosmic Ray Conference*, 10, 2015 [1510.05223].

[125] IceCube collaboration, *The IceCube Neutrino Observatory - Contributions to ICRC 2017 Part II: Properties of the Atmospheric and Astrophysical Neutrino Flux*, 1710.01191.

[126] IceCube collaboration, *Characteristics of the diffuse astrophysical electron and tau neutrino flux with six years of IceCube high energy cascade data*, *Phys. Rev. Lett.* **125** (2020) 121104 [2001.09520].

[127] IceCube collaboration, *Improved Characterization of the Astrophysical Muon-Neutrino Flux with 9.5 Years of IceCube Data*, 2111.10299.

[128] T. Gaisser, R. Engel and E. Resconi, *Cosmic Rays and Particle Physics*, Cambridge University Press (2016).

[129] J. STACHURSKA, *New measurement of the flavor composition of high- energy neutrino events with contained vertices in IceCube*, June, 2018. 10.5281/zenodo.1301122.

[130] IceCube collaboration, *Measurement of Astrophysical Tau Neutrinos in IceCube's High-Energy Starting Events*, 2011.03561.

[131] IceCube-Gen2 collaboration, *IceCube-Gen2: the window to the extreme Universe*, *J. Phys. G* **48** (2021) 060501 [2008.04323].

[132] IceCube collaboration, *Detection of a particle shower at the Glashow resonance with IceCube*, *Nature* **591** (2021) 220 [2110.15051].

[133] IceCube collaboration, *Differential limit on the extremely-high-energy cosmic neutrino flux in the presence of astrophysical background from nine years of IceCube data*, *Phys. Rev. D* **98** (2018) 062003 [1807.01820].

[134] Pierre Auger collaboration, *Probing the origin of ultra-high-energy cosmic rays with neutrinos in the EeV energy range using the Pierre Auger Observatory*, *JCAP* **10** (2019) 022 [1906.07422].

[135] ANITA collaboration, *Constraints on the ultrahigh-energy cosmic neutrino flux from the fourth flight of ANITA*, *Phys. Rev. D* **99** (2019) 122001 [1902.04005].

[136] ARA collaboration, *Performance of two Askaryan Radio Array stations and first results in the search for ultrahigh energy neutrinos*, *Phys. Rev. D* **93** (2016) 082003 [1507.08991].

[137] M. Sasaki and T. Kifune, *Ashra Neutrino Telescope Array (NTA): Combined Imaging Observation of Astroparticles — For Clear Identification of Cosmic Accelerators and Fundamental Physics Using Cosmic Beams —*, *JPS Conf. Proc.* **15** (2017) 011013.

[138] GRAND collaboration, *The Giant Radio Array for Neutrino Detection (GRAND): Science and Design*, *Sci. China Phys. Mech. Astron.* **63** (2020) 219501 [1810.09994].

[139] S. Wissel et al., *Prospects for high-elevation radio detection of >100 PeV tau neutrinos*, *JCAP* **11** (2020) 065 [2004.12718].

[140] A. Romero-Wolf et al., *An Andean Deep-Valley Detector for High-Energy Tau Neutrinos*, in *Latin American Strategy Forum for Research Infrastructure*, 2, 2020 [2002.06475].

[141] POEMMA collaboration, *The POEMMA (Probe of Extreme Multi-Messenger Astrophysics) observatory*, *JCAP* **06** (2021) 007 [2012.07945].

[142] A. Wang, C. Lin, N. Otte, M. Doro, E. Gazda, I. Taboada et al., *Trinity's Sensitivity to Isotropic and Point-Source Neutrinos*, *PoS* **ICRC2021** (2021) 1234 [2108.02751].

[143] T.M. Venters, M.H. Reno, J.F. Krizmanic, L.A. Anchordoqui, C. Guépin and A.V. Olinto, *POEMMA's Target of Opportunity Sensitivity to Cosmic Neutrino Transient Sources*, 1906.07209.

[144] PUEO collaboration, *The Payload for Ultrahigh Energy Observations (PUEO): a white paper*, *JINST* **16** (2021) P08035 [2010.02892].

[145] A. van Vliet, R. Alves Batista and J.R. Hörandel, *Current constraints from cosmogenic neutrinos on the fraction of protons in UHECRs*, *PoS* **ICRC2019** (2021) [1909.01932].

[146] A.L. Cummings, R. Aloisio and J.F. Krizmanic, *Modeling of the Tau and Muon Neutrino-induced Optical Cherenkov Signals from Upward-moving Extensive Air Showers*, *Phys. Rev. D* **103** (2021) 043017 [2011.09869].

[147] G.A. Askar'yan, *Excess negative charge of an electron-photon shower and its coherent radio emission*, *Zh. Eksp. Teor. Fiz.* **41** (1961) 616.

[148] RNO-G collaboration, *Design and Sensitivity of the Radio Neutrino Observatory in Greenland (RNO-G)*, *JINST* **16** (2021) P03025 [2010.12279].

[149] P-ONE collaboration, *The Pacific Ocean Neutrino Experiment*, *Nature Astron.* **4** (2020) 913 [2005.09493].

[150] A.M. Dziewonski and D.L. Anderson, *Preliminary reference earth model*, *Phys. Earth Planet. Interiors* **25** (1981) 297.

[151] G. de Wasseige. personal communication, 8, 2020.

[152] M.M. Block, L. Durand and P. Ha, *Connection of the virtual $\gamma^* p$ cross section of ep deep inelastic scattering to real $\gamma p$ scattering, and the implications for $\nu n$ and ep total cross sections*, *Physical Review D* **89** (2014) 094027.

[153] A. Neronov, D.V. Semikoz, L.A. Anchordoqui, J. Adams and A.V. Olinto, *Sensitivity of a proposed space-based Cherenkov astrophysical-neutrino telescope*, *Phys. Rev. D* **95** (2017) 023004 [1606.03629].

[154] J.A. Aguilar et al., *The Next-Generation Radio Neutrino Observatory – Multi-Messenger Neutrino Astrophysics at Extreme Energies*, 1907.12526.

[155] NuSpaceSim collaboration, *Monte Carlo simulations of neutrino and charged lepton propagation in the Earth with nuPyProp*, *PoS* **ICRC2021** (2021) 1203 [2109.08198].

[156] J. Alvarez-Muñiz, W.R. Carvalho, A.L. Cummings, K. Payet, A. Romero-Wolf, H. Schoorlemmer et al., *Comprehensive approach to tau-lepton production by high-energy tau neutrinos propagating through the Earth*, *Phys. Rev. D* **97** (2018) 023021 [1707.00334].

[157] HiRes collaboration, *First observation of the Greisen-Zatsepin-Kuzmin suppression*, *Phys. Rev. Lett.* **100** (2008) 101101 [astro-ph/0703099].

[158] P. Sokolsky, *Observation of the GZK cutoff by the HiRes experiment and the Pierre Auger Observatory*, *Mod. Phys. Lett.* **A23** (2008) 1290.

[159] Pierre Auger collaboration, *Observation of the suppression of the flux of cosmic rays above $4 \times 10^{19} eV$*, *Phys. Rev. Lett.* **101** (2008) 061101 [0806.4302].

[160] Telescope Array collaboration, *The Cosmic Ray Energy Spectrum Observed with the Surface Detector of the Telescope Array Experiment*, *Astrophys. J.* **768** (2013) L1 [1205.5067].

[161] IceCube collaboration, *The first search for extremely-high energy cosmogenic neutrinos with the IceCube Neutrino Observatory*, *Phys. Rev.* **D82** (2010) 072003 [1009.1442].

[162] ARIANNA collaboration, *A First Search for Cosmogenic Neutrinos with the ARIANNA Hexagonal Radio Array*, *Astropart. Phys.* **70** (2015) 12 [1410.7352].

[163] Pierre Auger collaboration, *Improved limit to the diffuse flux of ultrahigh energy neutrinos from the Pierre Auger Observatory*, *Phys. Rev.* **D91** (2015) 092008 [1504.05397].

[164] IceCube collaboration, *Constraints on Ultrahigh-Energy Cosmic-Ray Sources from a Search for Neutrinos above 10 PeV with IceCube*, *Phys. Rev. Lett.* **117** (2016) 241101 [1607.05886].

[165] ARA collaboration, *Performance of two Askaryan Radio Array stations and first results in the search for ultrahigh energy neutrinos*, *Phys. Rev.* **D93** (2016) 082003 [1507.08991].

[166] ANITA collaboration, *Constraints on the diffuse high-energy neutrino flux from the third flight of ANITA*, *Phys. Rev.* **D98** (2018) 022001 [1803.02719].

[167] IceCube collaboration, *Differential limit on the extremely-high-energy cosmic neutrino flux in the presence of astrophysical background from nine years of IceCube data*, *Phys. Rev. D* **98** (2018) 062003 [1807.01820].

[168] Pierre Auger collaboration, *Probing the origin of ultra-high-energy cosmic rays with neutrinos in the EeV energy range using the Pierre Auger Observatory*, 1906.07422.

[169] K. Kotera, D. Allard and A. Olinto, *Cosmogenic Neutrinos: parameter space and detectabilty from PeV to ZeV*, *JCAP* **10** (2010) 013 [1009.1382].

[170] M. Ahlers, L.A. Anchordoqui, M.C. Gonzalez-Garcia, F. Halzen and S. Sarkar, *GZK Neutrinos after the Fermi-LAT Diffuse Photon Flux Measurement*, *Astropart. Phys.* **34** (2010) 106 [1005.2620].

[171] M. Honda, T. Kajita, K. Kasahara, S. Midorikawa and T. Sanuki, *Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data*, *Phys. Rev.* **D75** (2007) 043006 [astro-ph/0611418].

[172] R. Enberg, M.H. Reno and I. Sarcevic, *Prompt neutrino fluxes from atmospheric charm*, *Phys. Rev.* **D78** (2008) 043005 [0806.0418].

[173] IceCube collaboration, *A measurement of the diffuse astrophysical muon neutrino flux using eight years of IceCube data.*, *PoS* **ICRC2017** (2018) 1005.

[174] N. Globus, D. Allard, E. Parizot and T. Piran, *Probing the Extragalactic Cosmic Rays origin with gamma-ray and neutrino backgrounds*, *Astrophys. J.* **839** (2017) L22 [1703.04158].

[175] R. Aloisio, D. Boncioli, A. di Matteo, A.F. Grillo, S. Petrera and F. Salamida, *Cosmogenic neutrinos and ultra-high energy cosmic ray models*, *JCAP* **1510** (2015) 006 [1505.04020].

[176] A. Schneider, *Characterization of the Astrophysical Diffuse Neutrino Flux with IceCube High-Energy Starting Events*, in *36th International Cosmic Ray Conference (ICRC 2019) Madison, Wisconsin, USA, July 24-August 1, 2019*, 2019 [1907.11266].

[177] Pierre Auger collaboration, *Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory*, *JCAP* **1704** (2017) 038 [1612.07155].

[178] IceCube collaboration, *The IceCube high-energy starting event sample: Description and flux characterization with 7.5 years of data*, 2011.03545.

[179] IceCube collaboration, *Observation and Characterization of a Cosmic Muon Neutrino Flux from the Northern Hemisphere using six years of IceCube data*, *Astrophys. J.* **833** (2016) 3 [1607.08006].

[180] IceCube collaboration, *Energy Reconstruction Methods in the IceCube Neutrino Telescope*, *JINST* **9** (2014) P03009 [1311.4767].

[181] S.S. Kimura, K. Murase, P. Mészáros and K. Kiuchi, *High-Energy Neutrino Emission from Short Gamma-Ray Bursts: Prospects for Coincident Detection with Gravitational Waves*, *Astrophys. J. Lett.* **848** (2017) L4 [1708.07075].

[182] K. Fang, B.D. Metzger, K. Murase, I. Bartos and K. Kotera, *Multimessenger Implications of AT2018cow: High-energy Cosmic-Ray and Neutrino Emissions from Magnetar-powered Superluminous Transients*, *Astrophys. J.* **878** (2019) 34 [1812.11673].

[183] K. Murase, P. Meszaros and B. Zhang, *Probing the birth of fast rotating magnetars through high-energy neutrinos*, *Phys. Rev. D* **79** (2009) 103001 [0904.2509].

[184] J.A. Carpio, K. Murase, M.H. Reno, I. Sarcevic and A. Stasto, *Charm contribution to ultrahigh-energy neutrinos from newborn magnetars*, *Phys. Rev. D* **102** (2020) 103001 [2007.07945].

[185] K. Fang, K. Kotera, K. Murase and A.V. Olinto, *Testing the Newborn Pulsar Origin of Ultrahigh Energy Cosmic Rays with EeV Neutrinos*, *Phys. Rev. D* **90** (2014) 103005 [1311.2044].

[186] C. Yuan, K. Murase, S.S. Kimura and P. Mészáros, *High-energy neutrino emission subsequent to gravitational wave radiation from supermassive black hole mergers*, *Phys. Rev. D* **102** (2020) 083013 [2008.05616].

[187] K. Murase, *High energy neutrino early afterglows gamma-ray bursts revisited*, *Phys. Rev. D* **76** (2007) 123001 [0707.1140].

[188] E. Guarini, I. Tamborra, D. Bégué, T. Pitik and J. Greiner, *Multi-messenger detection prospects of gamma-ray burst afterglows with optical jumps*, 2112.07690.

[189] M. Petropoulou, S. Coenders and S. Dimitrakoudis, *Time-dependent neutrino emission from Mrk 421 during flares and predictions for IceCube*, *Astropart. Phys.* **80** (2016) 115 [1603.06954].

[190] A. Keivani et al., *A Multimessenger Picture of the Flaring Blazar TXS 0506+056: implications for High-Energy Neutrino Emission and Cosmic Ray Acceleration*, *Astrophys. J.* **864** (2018) 84 [1807.04537].

[191] IceCube collaboration, *Follow-up of Astrophysical Transients in Real Time with the IceCube Neutrino Observatory*, *Astrophys. J.* **910** (2021) 4 [2012.04577].

[192] ANITA collaboration, *The Antarctic Impulsive Transient Antenna Ultra-high Energy Neutrino Detector Design, Performance, and Sensitivity for 2006-2007 Balloon Flight*, *Astropart. Phys.* **32** (2009) 10 [0812.1920].

[193] ANITA collaboration, *Observation of Ultra-high-energy Cosmic Rays with the ANITA Balloon-borne Radio Interferometer*, *Phys. Rev. Lett.* **105** (2010) 151101 [1005.0035].

[194] ANITA collaboration, *Characteristics of Four Upward-pointing Cosmic-ray-like Events Observed with ANITA*, *Phys. Rev. Lett.* **117** (2016) 071101 [1603.05218].

[195] ANITA collaboration, *Observation of an Unusual Upward-going Cosmic-ray-like Event in the Third Flight of ANITA*, *Phys. Rev. Lett.* **121** (2018) 161102 [1803.05088].

[196] A. Romero-Wolf et al., *Comprehensive analysis of anomalous ANITA events disfavors a diffuse tau-neutrino flux origin*, *Phys. Rev. D* **99** (2019) 063011 [1811.07261].

[197] D.B. Fox, S. Sigurdsson, S. Shandera, P. Mészáros, K. Murase, M. Mostafá et al., *The ANITA Anomalous Events as Signatures of a Beyond Standard Model Particle, and Supporting Observations from IceCube*, 1809.09615.

[198] S. Chipman, R. Diesing, M.H. Reno and I. Sarcevic, *Anomalous ANITA air shower events and tau decays*, 1906.11736.

[199] I. Esteban, J. Lopez-Pavon, I. Martinez-Soler and J. Salvado, *Looking at the axionic dark sector with ANITA*, 1905.10372.

[200] G. Huang, *Sterile neutrinos as a possible explanation for the upward air shower events at ANITA*, *Phys. Rev.* **D98** (2018) 043019 [1804.05362].

[201] L.A. Anchordoqui, V. Barger, J.G. Learned, D. Marfatia and T.J. Weiler, *Upgoing ANITA events as evidence of the CPT symmetric universe*, *LHEP* **1** (2018) 13 [1803.11554].

[202] J.F. Cherry and I.M. Shoemaker, *Sterile neutrino origin for the upward directed cosmic ray showers detected by ANITA*, *Phys. Rev.* **D99** (2019) 063016 [1802.01611].

[203] A. Connolly, P. Allison and O. Banerjee, *On ANITA's sensitivity to long-lived, charged massive particles*, 1807.08892.

[204] J.H. Collins, P.S. Bhupal Dev and Y. Sui, *R-parity Violating Supersymmetric Explanation of the Anomalous Events at ANITA*, *Phys. Rev.* **D99** (2019) 043009 [1810.08479].

[205] L.A. Anchordoqui and I. Antoniadis, *Supersymmetric sphaleron configurations as the origin of the perplexing ANITA events*, *Phys. Lett.* **B790** (2019) 578 [1812.01520].

[206] L. Heurtier, Y. Mambrini and M. Pierre, *Dark matter interpretation of the ANITA anomalous events*, *Phys. Rev.* **D99** (2019) 095014 [1902.04584].

[207] D. Hooper, S. Wegsman, C. Deaconu and A. Vieregg, *Superheavy dark matter and ANITA's anomalous events*, *Phys. Rev.* **D100** (2019) 043019 [1904.12865].

[208] J.M. Cline, C. Gross and W. Xue, *Can the ANITA anomalous events be due to new physics?*, *Phys. Rev.* **D100** (2019) 015031 [1904.13396].

[209] L. Heurtier, D. Kim, J.-C. Park and S. Shin, *Explaining the ANITA Anomaly with Inelastic Boosted Dark Matter*, *Phys. Rev.* **D100** (2019) 055004 [1905.13223].

[210] D. Borah, A. Dasgupta, K. Dey and G. Tomar, *Connecting ANITA Anomalous Events to a Non-thermal Dark Matter Scenario*, 1907.02740.

[211] A. Esmaili and Y. Farzan, *Explaining the ANITA events by a $L_e - L_\tau$ gauge model*, *JCAP* **12** (2019) 017 [1909.07995].

[212] Van Santen, Jakob, *Neutrino Interactions in IceCube above 1 TeV: Constraints on Atmospheric Charmed-Meson Production and Investigation of the Astrophysical Neutrino Flux with 2 Years of IceCube Data taken 2010–2012.*, Ph.D. thesis, University of Wisconsin Madison, 2014.

[213] IceCube collaboration, *Observation of Astrophysical Neutrinos in Six Years of IceCube Data*, *PoS* **ICRC2017** (2018) 981.

[214] ANTARES, IceCube, Telescope Array collaboration, *Search for correlations of high-energy neutrinos and ultrahigh- energy cosmic rays*, *EPJ Web Conf.* **207** (2019) 02010 [1905.10111].

[215] IceCube collaboration, *A Search for Neutrino Emission from Fast Radio Bursts with Six Years of IceCube Data*, *Astrophys. J.* **857** (2018) 117 [1712.06277].

[216] IceCube collaboration, *Search for Prompt Neutrino Emission from Gamma-Ray Bursts with IceCube*, *Astrophys. J.* **805** (2015) L5 [1412.6510].

[217] IceCube collaboration, *The IceCube Neutrino Observatory Part II: Atmospheric and Diffuse UHE Neutrino Searches of All Flavors*, in *Proceedings, 33rd International Cosmic Ray Conference (ICRC2013): Rio de Janeiro, Brazil, July 2-9, 2013*, 2013 [1309.7003].

[218] V.C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, *Astrophysical Journal* **159** (1970) 379.

[219] M. Persic, P. Salucci and F. Stel, *The Universal rotation curve of spiral galaxies: 1. The Dark matter connection*, *Mon. Not. Roy. Astron. Soc.* **281** (1996) 27 [astro-ph/9506004].

[220] F. Zwicky, *On the Masses of Nebulae and of Clusters of Nebulae*, *Astrophysical Journal* **86** (1937) 217.

[221] S. Smith, *The Mass of the Virgo Cluster*, *Astrophysical Journal* **83** (1936) 23.

[222] M. Jee and J. Tyson, *Dark Matter in the Galaxy Cluster CL J1226+3332 at Z=0.89*, *Astrophys. J.* **691** (2009) 1337 [0810.0709].

[223] M.J. Jee et al., *Discovery of a Ringlike Dark Matter Structure in the Core of the Galaxy Cluster Cl 0024+17*, *Astrophys. J.* **661** (2007) 728 [0705.2171].

[224] G. Bertone and D. Hooper, *History of dark matter*, *Rev. Mod. Phys.* **90** (2018) 045002 [1605.04909].

[225] Planck collaboration, *Planck 2018 results. VI. Cosmological parameters*, 1807.06209.

[226] Fermi-LAT, DES collaboration, *Searching for Dark Matter Annihilation in Recently Discovered Milky Way Satellites with Fermi-LAT*, *Astrophys. J.* **834** (2017) 110 [1611.03184].

[227] S. Hoof, A. Geringer-Sameth and R. Trotta, *A Global Analysis of Dark Matter Signals from 27 Dwarf Spheroidal Galaxies using Ten Years of Fermi-LAT Observations*, 1812.06986.

[228] M. Boudaud, Y. Génolini, L. Derome, J. Lavalle, D. Maurin, P. Salati et al., *AMS-02 antiprotons are consistent with a secondary astrophysical origin*, *Phys. Rev. Res.* **2** (2020) 023022 [1906.07119].

[229] C. Cappiello and J.F. Beacom, *Strong New Limits on Light Dark Matter from Neutrino Experiments*, *Phys. Rev.* **D100** (2019) 103011 [1906.11283].

[230] Y. Farzan and S. Palomares-Ruiz, *Dips in the Diffuse Supernova Neutrino Background*, *JCAP* **06** (2014) 014 [1401.7019].

[231] J.H. Davis and J. Silk, *Spectral and Spatial Distortions of PeV Neutrinos from Scattering with Dark Matter*, 1505.01843.

[232] J.F. Cherry, A. Friedland and I.M. Shoemaker, *Short-baseline neutrino oscillations, Planck, and IceCube*, 1605.06506.

[233] C.A. Argüelles, A. Kheirandish and A.C. Vincent, *Imaging Galactic Dark Matter with High-Energy Cosmic Neutrinos*, *Phys. Rev. Lett.* **119** (2017) 201801 [1703.00451].

[234] K.J. Kelly and P.A.N. Machado, *Multimessenger Astronomy and New Neutrino Physics*, *JCAP* **1810** (2018) 048 [1808.02889].

[235] Y. Farzan and S. Palomares-Ruiz, *Flavor of cosmic neutrinos preserved by ultralight dark matter*, *Phys. Rev.* **D99** (2019) 051702 [1810.00892].

[236] S. Pandey, S. Karmakar and S. Rakshit, *Interactions of Astrophysical Neutrinos with Dark Matter: A model building perspective*, *JHEP* **01** (2019) 095 [1810.04203].

[237] K.-Y. Choi, J. Kim and C. Rott, *Constraining dark matter-neutrino interactions with IceCube-170922A*, *Phys. Rev.* **D99** (2019) 083018 [1903.03302].

[238] F. Capozzi, I.M. Shoemaker and L. Vecchi, *Neutrino Oscillations in Dark Backgrounds*, *JCAP* **1807** (2018) 004 [1804.05117].

[239] K. Murase and I.M. Shoemaker, *Neutrino Echoes from Multimessenger Transient Sources*, *Phys. Rev. Lett.* **123** (2019) 241102 [1903.08607].

[240] F. Capozzi, I.M. Shoemaker and L. Vecchi, *Solar Neutrinos as a Probe of Dark Matter-Neutrino Interactions*, *JCAP* **1707** (2017) 021 [1702.08464].

[241] A. Olivares-Del Campo, C. Boehm, S. Palomares-Ruiz and S. Pascoli, *Dark matter-neutrino interactions through the lens of their cosmological implications*, *Phys. Rev.* **D97** (2018) 075039 [1711.05283].

[242] G. Barenboim, P.B. Denton and I.M. Oldengott, *Constraints on inflation with an extended neutrino sector*, *Phys. Rev.* **D99** (2019) 083515 [1903.02036].

[243] MiniBooNE collaboration, *Dark Matter Search in a Proton Beam Dump with MiniBooNE*, *Phys. Rev. Lett.* **118** (2017) 221803 [1702.02688].

[244] C.A. Argüelles, M. Hostert and Y.-D. Tsai, *Testing New Physics Explanations of MiniBooNE Anomaly at Neutrino Scattering Experiments*, 1812.08768.

[245] M. Hostert, *Hidden Physics at the Neutrino Frontier: Tridents, Dark Forces, and Hidden Particles*, Ph.D. thesis, Durham U., 2019.

[246] R. Primulando and P. Uttayarat, *Dark Matter-Neutrino Interaction in Light of Collider and Neutrino Telescope Data*, *JHEP* **06** (2018) 026 [1710.08567].

[247] G. Arcadi, M. Dutra, P. Ghosh, M. Lindner, Y. Mambrini, M. Pierre et al., *The waning of the WIMP? A review of models, searches, and constraints*, *Eur. Phys. J.* **C78** (2018) 203 [1703.07364].

[248] Particle Data Group collaboration, *Review of Particle Physics*, *Phys. Rev.* **D98** (2018) 030001.

[249] GRAVITY collaboration, *Detection of the gravitational redshift in the orbit of the star S2 near the Galactic centre massive black hole*, *Astron. Astrophys.* **615** (2018) L15 [1807.09409].

[250] M. Benito, A. Cuoco and F. Iocco, *Handling the Uncertainties in the Galactic Dark Matter Distribution for Particle Dark Matter Searches*, *JCAP* **1903** (2019) 033 [1901.02460].

[251] F. Prada, A.A. Klypin, A.J. Cuesta, J.E. Betancort-Rijo and J. Primack, *Halo concentrations in the standard $\Lambda$ cold dark matter cosmology*, *Monthly Notices of the RAS* **423** (2012) 3018 [1104.5130].

[252] L. Lopez-Honorez, O. Mena, S. Palomares-Ruiz and A.C. Vincent, *Constraints on dark matter annihilation from CMB observationsbefore Planck*, *JCAP* **1307** (2013) 046 [1303.5094].

[253] I.M. Shoemaker, *Constraints on Dark Matter Protohalos in Effective Theories and Neutrinophilic Dark Matter*, *Phys. Dark Univ.* **2** (2013) 157 [1305.1936].

[254] J.M. Cornell, S. Profumo and W. Shepherd, *Kinetic Decoupling and Small-Scale Structure in Effective Theories of Dark Matter*, *Phys. Rev.* **D88** (2013) 015027 [1305.4676].

[255] W.A. Watson, I.T. Iliev, A. D'Aloisio, A. Knebe, P.R. Shapiro and G. Yepes, *The halo mass function through the cosmic ages*, *Monthly Notices of the RAS* **433** (2013) 1230 [1212.0095].

[256] R. Diamanti, L. Lopez-Honorez, O. Mena, S. Palomares-Ruiz and A.C. Vincent, *Constraining Dark Matter Late-Time Energy Injection: Decays and P-Wave Annihilations*, *JCAP* **1402** (2014) 017 [1308.2578].

[257] J.F. Beacom, N.F. Bell and G.D. Mack, *General Upper Bound on the Dark Matter Total Annihilation Cross Section*, *Phys. Rev. Lett.* **99** (2007) 231301 [astro-ph/0608090].

[258] H. Yüksel, S. Horiuchi, J.F. Beacom and S. Ando, *Neutrino Constraints on the Dark Matter Total Annihilation Cross Section*, *Phys. Rev.* **D76** (2007) 123506 [0707.0196].

[259] M. Carena, A. Daleo, B.A. Dobrescu and T.M.P. Tait, *$Z'$ gauge bosons at the Tevatron*, *Phys. Rev.* **D70** (2004) 093009 [hep-ph/0408098].

[260] BaBar collaboration, *Search for a Dark Photon in $e^+e^-$ Collisions at BaBar*, *Phys. Rev. Lett.* **113** (2014) 201801 [1406.2980].

[261] APEX collaboration, *Search for a New Gauge Boson in Electron-Nucleus Fixed-Target Scattering by the APEX Experiment*, *Phys. Rev. Lett.* **107** (2011) 191804 [1108.2750].

[262] C. Blanco, M. Escudero, D. Hooper and S.J. Witte, *$Z'$ Mediated WIMPs: Dead, Dying, or Soon to be Detected?*, 1907.05893.

[263] X.-G. He, G.C. Joshi, H. Lew and R.R. Volkas, *NEW Z-prime PHENOMENOLOGY*, *Phys. Rev.* **D43** (1991) 22.

[264] X.-G. He, G.C. Joshi, H. Lew and R.R. Volkas, *Simplest Z-prime model*, *Phys. Rev.* **D44** (1991) 2118.

[265] S. Profumo, F.S. Queiroz, J. Silk and C. Siqueira, *Searching for Secluded Dark Matter with H.E.S.S., Fermi-LAT, and Planck*, *JCAP* **1803** (2018) 010 [1711.03133].

[266] C.W. Bauer, N.L. Rodd and B.R. Webber, *Dark Matter Spectra from the Electroweak to the Planck Scale*, 2007.15001.

[267] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [1410.3012].

[268] T. Sjöstrand, *The PYTHIA Event Generator: Past, Present and Future*, *Comput. Phys. Commun.* **246** (2020) 106910 [1907.09874].

[269] Q. Liu, J. Lazar, C.A. Argüelles and A. Kheirandish, *$\chi$aro$\nu$: a tool for neutrino flux generation from WIMPs*, *JCAP* **10** (2020) 043 [2007.15010].

[270] K. Murase and J.F. Beacom, *Constraining Very Heavy Dark Matter Using Diffuse Backgrounds of Neutrinos and Cascaded Gamma Rays*, *JCAP* **1210** (2012) 043 [1206.2595].

[271] F.S. Queiroz, C.E. Yaguna and C. Weniger, *Gamma-ray Limits on Neutrino Lines*, *JCAP* **1605** (2016) 050 [1602.05966].

[272] ANTARES collaboration, *Search of Dark Matter Annihilation in the Galactic Centre using the ANTARES Neutrino Telescope*, *JCAP* **1510** (2015) 068 [1505.04866].

[273] T. Cohen, K. Murase, N.L. Rodd, B.R. Safdi and Y. Soreq, $\gamma$ *-ray Constraints on Decaying Dark Matter and Implications for IceCube*, *Phys. Rev. Lett.* **119** (2017) 021102 [1612.05638].

[274] K. Murase, R. Laha, S. Ando and M. Ahlers, *Testing the Dark Matter Scenario for PeV Neutrinos Observed in IceCube*, *Phys. Rev. Lett.* **115** (2015) 071301 [1503.04663].

[275] M. Chianese, D.F.G. Fiorillo, G. Miele, S. Morisi and O. Pisanti, *Decaying dark matter at IceCube and its signature on High Energy gamma experiments*, *JCAP* **11** (2019) 046 [1907.11222].

[276] M. Pato, F. Iocco and G. Bertone, *Dynamical constraints on the dark matter distribution in the Milky Way*, *JCAP* **1512** (2015) 001 [1504.06324].

[277] M. Benito, N. Bernal, N. Bozorgnia, F. Calore and F. Iocco, *Particle Dark Matter Constraints: the Effect of Galactic Uncertainties*, *JCAP* **1702** (2017) 007 [1612.02010].

[278] E.V. Karukes, M. Benito, F. Iocco, R. Trotta and A. Geringer-Sameth, *Bayesian reconstruction of the Milky Way dark matter distribution*, 1901.02463.

[279] GRAND collaboration, *The Giant Radio Array for Neutrino Detection (GRAND): Science and Design*, 1810.09994.

[280] Super-Kamiokande collaboration, *Measurements of the atmospheric neutrino flux by Super-Kamiokande: energy spectra, geomagnetic effects, and solar modulation*, *Phys. Rev.* **D94** (2016) 052001 [1510.08127].

[281] IceCube collaboration, *Measurement of the Atmospheric $\nu_e$ Spectrum with IceCube*, *Phys. Rev.* **D91** (2015) 122004 [1504.03753].

[282] K.K. Boddy, J. Kumar and L.E. Strigari, *The Effective J-Factor of the Galactic Center for Velocity-Dependent Dark Matter Annihilation*, 1805.08379.

[283] N.F. Bell, Y. Cai, J.B. Dent, R.K. Leane and T.J. Weiler, *Enhancing Dark Matter Annihilation Rates with Dark Bremsstrahlung*, *Phys. Rev. D* **96** (2017) 023011 [1705.01105].

[284] J.S. Gainer, J. Lykken, K.T. Matchev, S. Mrenna and M. Park, *Exploring Theory Space with Monte Carlo Reweighting*, *JHEP* **10** (2014) 078 [1404.7129].

[285] T. Glüsenkamp, *Probabilistic treatment of the uncertainty from the finite size of weighted Monte Carlo data*, *Eur. Phys. J. Plus* **133** (2018) 218 [1712.01293].

[286] C.A. Argüelles, A. Schneider and T. Yuan, *A binned likelihood for stochastic models*, *JHEP* **06** (2019) 030 [1901.04645].

[287] T. Glüsenkamp, *A unified perspective on modified Poisson likelihoods for limited Monte Carlo data*, *JINST* **15** (2020) P01035 [1902.08831].

[288] R.J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, *Comput. Phys. Commun.* **77** (1993) 219.

[289] G. Bohm and G. Zech, *Statistics of weighted Poisson events and its applications*, *Nucl. Instrum. Meth. A* **748** (2014) 1 [1309.1287].

[290] D. Chirkin, *Likelihood description for comparing data with simulation of limited statistics*, 1304.0735.

[291] J. Heinrich and L. Lyons, *Systematic errors*, *Annual Review of Nuclear and Particle Science* **57** (2007) 145 [https://doi.org/10.1146/annurev.nucl.57.090506.123052].

[292] R. Trotta, *Bayesian Methods in Cosmology*, 1, 2017 [1701.01467].

[293] R.D. Cousins and V.L. Highland, *Incorporating systematic uncertainties into an upper limit*, *Nucl. Instrum. Meth. A* **320** (1992) 331.

[294] A. Diaz, C. Arguelles, G. Collin, J. Conrad and M. Shaevitz, *Where Are We With Light Sterile Neutrinos?*, 1906.00045.

[295] D.Z. Freedman, *Coherent Neutrino Nucleus Scattering as a Probe of the Weak Neutral Current*, *Phys. Rev. D* **9** (1974) 1389.

[296] COHERENT collaboration, *Observation of Coherent Elastic Neutrino-Nucleus Scattering*, *Science* **357** (2017) 1123 [1708.01294].

[297] DARWIN collaboration, *DARWIN: towards the ultimate dark matter detector*, *JCAP* **1611** (2016) 017 [1606.07001].

[298] Super-Kamiokande collaboration, *Measurements of the solar neutrino flux from Super-Kamiokande's first 300 days*, *Phys. Rev. Lett.* **81** (1998) 1158 [hep-ex/9805021].

[299] SNO collaboration, *Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by $^8B$ solar neutrinos at the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **87** (2001) 071301 [nucl-ex/0106015].

[300] Borexino collaboration, *Science and technology of BOREXINO: A Real time detector for low-energy solar neutrinos*, *Astropart. Phys.* **16** (2002) 205 [hep-ex/0012030].

[301] Borexino collaboration, *Direct Measurement of the Be-7 Solar Neutrino Flux with 192 Days of Borexino Data*, *Phys. Rev. Lett.* **101** (2008) 091302 [0805.3843].

[302] A.M. Ankowski, *Improved estimate of the cross section for inverse beta decay*, 1601.06169.

[303] W. Leo, *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach*, Springer (1994).

[304] J.F. Beacom and M.R. Vagins, *GADZOOKS! Anti-neutrino spectroscopy with large water Cherenkov detectors*, *Phys. Rev. Lett.* **93** (2004) 171101 [hep-ph/0309300].

[305] Super-Kamiokande collaboration, *First Study of Neutron Tagging with a Water Cherenkov Detector*, *Astropart. Phys.* **31** (2009) 320 [0811.0735].

[306] D. Besson, N. Doolin, M. Stockham and I. Kravchenko, *Radio-frequency probes of Antarctic ice birefringence at South Pole vs. East Antarctica; evidence for a changing ice fabric*, *The Cryosphere Discussions* **6** (2012) 4695.

[307] J. Conrad, A. de Gouvea, S. Shalgar and J. Spitz, *Atmospheric Tau Neutrinos in a Multi-kiloton Liquid Argon Detector*, *Phys. Rev. D* **82** (2010) 093012 [1008.2984].

[308] MINERvA collaboration, *Design, Calibration, and Performance of the MINERvA Detector*, *Nucl. Instrum. Meth. A* **743** (2014) 130 [1305.5199].

[309] MINOS, MINOS+ collaboration, *First MINOS+ Data and New Results from MINOS*, *AIP Conf. Proc.* **1666** (2015) 110004 [1502.07715].

[310] A. Sousa and O. U., *Neutral current interactions in minos*, .

[311] C. Backhouse and R. Patterson, *Library Event Matching event classification algorithm for electron neutrino interactions in the NO$\nu$A detectors*, *Nucl. Instrum. Meth. A* **778** (2015) 31 [1501.00968].

[312] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. Messier, E. Niner et al., *A convolutional neural network neutrino event classifier*, *Journal of Instrumentation* **11** (2016) P09001.

[313] F. Psihas, E. Niner, M. Groh, R. Murphy, A. Aurisano, A. Himmel et al., *Context-Enriched Identification of Particles with a Convolutional Network for Neutrino Events*, *Phys. Rev. D* **100** (2019) 073005 [1906.00713].

[314] F. Cavanna, A. Ereditato and B. Fleming, *Advances in liquid argon detectors*, *Nucl. Instrum. Meth. A* **907** (2018) 1.

[315] MicroBooNE collaboration, *The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector*, *Eur. Phys. J. C* **78** (2018) 82 [1708.03135].

[316] MicroBooNE collaboration, *Deep neural network for pixel-level electromagnetic particle identification in the MicroBooNE liquid argon time projection chamber*, *Phys. Rev. D* **99** (2019) 092001 [1808.07269].

[317] MicroBooNE collaboration, *First Deep Learning based Event Reconstruction for Low-Energy Excess Searches with MicroBooNE*, .

[318] MicroBooNE collaboration, *Design and Construction of the MicroBooNE Detector*, *JINST* **12** (2017) P02017 [1612.05824].

[319] B. Ali-Mohammadzadeh et al., *Design and implementation of the new scintillation light detection system of ICARUS T600*, 2006.05261.

[320] DUNE collaboration, *Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II DUNE Physics*, 2002.03005.

[321] R. Gandhi, C. Quigg, M.H. Reno and I. Sarcevic, *Ultrahigh-energy neutrino interactions*, *Astropart. Phys.* **5** (1996) 81 [hep-ph/9512364].

[322] J.G. Learned and S. Pakvasa, *Detecting tau-neutrino oscillations at PeV energies*, *Astropart. Phys.* **3** (1995) 267 [hep-ph/9405296].

[323] IceCube collaboration, *Tau neutrinos in IceCube*, *J. Phys. Conf. Ser.* **60** (2007) 227.

[324] IceCube collaboration, *First Double Cascade Tau Neutrino Candidates in IceCube and a New Measurement of the Flavor Composition*, *PoS* **ICRC2019** (2020) 1015 [1908.05506].

[325] A. Loewy, S. Nussinov and S.L. Glashow, *The Effect of Doppler Broadening on the* 6.3 *PeV* $W^-$ *Resonance in* $\bar{\nu}_e e^-$ *Collisions*, 1407.4415.

[326] B. Zhou and J.F. Beacom, *W -boson and trident production in TeV–PeV neutrino observatories*, *Phys. Rev. D* **101** (2020) 036010 [1910.10720].

[327] L. Lu, *A novel method of rejecting muon backgrounds for the detection of the highest energy neutrinos*, in *36th International Cosmic Ray Conference (ICRC2019)*, vol. 36 of *International Cosmic Ray Conference*, p. 945, July, 2019.

[328] G.A. Gusev and I.M. Zheleznykh, *On the possibility of detection of neutrinos and muons on the basis of radio radiation of cascades in natural dielectric media (antarctic ice sheet and so forth)*, *Soviet Physics Uspekhi* **27** (1984) 550.

[329] M. Markov and I. Zheleznykh, *Large Scale Cherenkov Detectors in Ocean, Atmosphere and Ice*, *Nucl. Instrum. Meth. A* **248** (1986) 242.

[330] G. Askar'yan, *Excess negative charge of an electron-photon shower and its coherent radio emission*, *Sov. Phys. JETP* **14** (1962) 441.

[331] E. Zas, F. Halzen and T. Stanev, *Electromagnetic pulses from high-energy showers: Implications for neutrino detection*, *Phys. Rev. D* **45** (1992) 362.

[332] Pierre Auger collaboration, *The Pierre Auger Cosmic Ray Observatory*, *Nucl. Instrum. Meth.* **A798** (2015) 172 [1502.01323].

[333] Y.S. Jeong, M.V. Luu, M.H. Reno and I. Sarcevic, *Tau energy loss and ultrahigh energy skimming tau neutrinos*, *Phys. Rev. D* **96** (2017) 043003 [1704.00050].

[334] POEMMA collaboration, *A new calculation of Earth-skimming very- and ultra-high energy tau neutrinos*, *PoS* **ICRC2019** (2020) 989 [1908.03603].

[335] ANITA collaboration, *Observational Constraints on the Ultra-high Energy Cosmic Neutrino Flux from the Second Flight of the ANITA Experiment*, *Phys. Rev. D* **82** (2010) 022004 [1003.2961].

[336] A. Anker et al., *White Paper: ARIANNA-200 high energy neutrino telescope*, 2004.09841.

[337] ARA collaboration, *Constraints on the Diffuse Flux of Ultra-High Energy Neutrinos from Four Years of Askaryan Radio Array Data in Two Stations*, 1912.00987.

[338] Super-Kamiokande collaboration, *Dark matter searches with the Super-Kamiokande detector*, *J. Phys. Conf. Ser.* **888** (2017) 012210.

[339] IceCube collaboration, *All-flavour Search for Neutrinos from Dark Matter Annihilations in the Milky Way with IceCube/DeepCore*, *Eur. Phys. J.* **C76** (2016) 531 [1606.00209].

[340] G. Steigman, B. Dasgupta and J.F. Beacom, *Precise Relic WIMP Abundance and its Impact on Searches for Dark Matter Annihilation*, *Phys. Rev.* **D86** (2012) 023506 [1204.3622].

[341] K. Griest and M. Kamionkowski, *Unitarity Limits on the Mass and Radius of Dark Matter Particles*, *Phys. Rev. Lett.* **64** (1990) 615.

[342] J. Smirnov and J.F. Beacom, *TeV-Scale Thermal WIMPs: Unitarity and its Consequences*, *Phys. Rev.* **D100** (2019) 043029 [1904.11503].

[343] Borexino collaboration, *The Borexino detector at the Laboratori Nazionali del Gran Sasso*, *Nucl. Instrum. Meth.* **A600** (2009) 568 [0806.2400].

[344] Borexino collaboration, *Observation of Geo-Neutrinos*, *Phys. Lett.* **B687** (2010) 299 [1003.0284].

[345] Borexino collaboration, *Study of solar and other unknown anti-neutrino fluxes with Borexino at LNGS*, *Phys. Lett.* **B696** (2011) 191 [1010.0029].

[346] Borexino collaboration, *Search for low-energy neutrinos from astrophysical sources with Borexino*, 1909.02422.

[347] KamLAND collaboration, *Measurement of the 8B Solar Neutrino Flux with the KamLAND Liquid Scintillator Detector*, *Phys. Rev.* **C84** (2011) 035804 [1106.0861].

[348] KamLAND collaboration, *Reactor On-Off Antineutrino Measurement with KamLAND*, *Phys. Rev.* **D88** (2013) 033001 [1303.4667].

[349] KamLAND collaboration, *A study of extraterrestrial antineutrino sources with the KamLAND detector*, *Astrophys. J.* **745** (2012) 193 [1105.3516].

[350] JUNO collaboration, *Neutrino Physics with JUNO*, *J. Phys.* **G43** (2016) 030401 [1507.05613].

[351] Super-Kamiokande collaboration, *The Super-Kamiokande detector*, *Nucl. Instrum. Meth.* **A501** (2003) 418.

[352] W. Linyan, *Experimental Studies on Low Energy Electron Antineutrinos and Related Physics*, Ph.D. thesis, Tsinghua University, 2018.

[353] IceCube collaboration, *The IceCube Neutrino Observatory: Instrumentation and Online Systems*, *JINST* **12** (2017) P03012 [1612.05093].

[354] IceCube collaboration, *Measurement of the $\nu_\mu$ energy spectrum with IceCube-79*, *Eur. Phys. J.* **C77** (2017) 692 [1705.07780].

[355] C. El Aisati, C. Garcia-Cely, T. Hambye and L. Vanderheyden, *Prospects for discovering a neutrino line induced by dark matter annihilation*, *JCAP* **1710** (2017) 021 [1706.06600].

[356] C. El Aisati, *Gamma-ray and Neutrino Lines from Dark Matter: multi-messenger and dedicated smoking-gun searches*, Ph.D. thesis, U. Brussels (main), 2018-01-18.

[357] Super-Kamiokande collaboration, *Solar neutrino measurements in super-Kamiokande-I*, *Phys. Rev.* **D73** (2006) 112001 [hep-ex/0508053].

[358] Super-Kamiokande collaboration, *Solar neutrino measurements in Super-Kamiokande-II*, *Phys. Rev.* **D78** (2008) 032002 [0803.4312].

[359] Super-Kamiokande collaboration, *Solar neutrino results in Super-Kamiokande-III*, *Phys. Rev.* **D83** (2011) 052010 [1010.0118].

[360] S.W. Li and J.F. Beacom, *First calculation of cosmic-ray muon spallation backgrounds for MeV astrophysical neutrino signals in Super-Kamiokande*, *Phys. Rev.* **C89** (2014) 045801 [1402.4687].

[361] A. Olivares-Del Campo, S. Palomares-Ruiz and S. Pascoli, *Implications of a Dark Matter-Neutrino Coupling at Hyper-Kamiokande*, in *53rd Rencontres de Moriond on Electroweak Interactions and Unified Theories (Moriond EW 2018) La Thuile, Italy, March 10-17, 2018*, 2018 [1805.09830].

[362] A. Olivares-Del-Campo, *Dark Matter and Neutrinos: A Love-Hate Relationship*, Ph.D. thesis, Durham U., 2019.

[363] Super-Kamiokande collaboration, *Searching for Dark Matter Annihilation into Neutrinos with Super-Kamiokande*, in *Proceedings, Meeting of the APS Division of Particles and Fields (DPF 2015): Ann Arbor, Michigan, USA, 4-8 Aug 2015*, 2015 [1510.07999].

[364] IceCube collaboration, *Search for Neutrinos from Dark Matter Self-Annihilations in the center of the Milky Way with 3 years of IceCube/DeepCore*, *Eur. Phys. J.* **C77** (2017) 627 [1705.08103].

[365] A. Bhattacharya, A. Esmaili, S. Palomares-Ruiz and I. Sarcevic, *Update on decaying and annihilating heavy dark matter with the 6-year IceCube HESE data*, *JCAP* **1905** (2019) 051 [1903.12623].

[366] A. Albert et al., *Results from the search for dark matter in the Milky Way with 9 years of data of the ANTARES neutrino telescope*, *Phys. Lett.* **B769** (2017) 249 [1612.04595].

[367] BAIKAL collaboration, *The Baikal underwater neutrino telescope: Design, performance and first results*, *Astropart. Phys.* **7** (1997) 263.

[368] V.M. Aynutdinov et al., *Baikal neutrino telescope*, *Phys. Atom. Nucl.* **69** (2006) 1914.

[369] BAIKAL collaboration, *A search for neutrino signal from dark matter annihilation in the center of the Milky Way with Baikal NT200*, *Astropart. Phys.* **81** (2016) 12 [1512.01198].

[370] DUNE collaboration, *Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE)*, 1512.06148.

[371] DUNE collaboration, *Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume I Introduction to DUNE*, *JINST* **15** (2020) T08008 [2002.02967].

[372] F. Capozzi, S.W. Li, G. Zhu and J.F. Beacom, *DUNE as the Next-Generation Solar Neutrino Experiment*, *Phys. Rev. Lett.* **123** (2019) 131803 [1808.08232].

[373] M. Honda, M.S. Athar, T. Kajita, K. Kasahara and S. Midorikawa, *Atmospheric neutrino flux calculation using the nrlmsise-00 atmospheric model*, *Phys. Rev. D* **92** (2015) 023004.

[374] C.A. Argüelles, J. Salvado and C.N. Weaver, *A simple quantum integro-differential solver (squids)*, 1412.3832.

[375] C.A. Argüelles, J. Salvado and C.N. Weaver, *A Simple Quantum Integro-Differential Solver (SQuIDS)*, *Comput. Phys. Commun.* **255** (2020) 107405.

[376] C.A. Argüelles, J. Salvado and C.N. Weaver, "nuSQuIDS." https://github.com/Arguelles/nuSQuIDS, 2015.

[377] Hyper-Kamiokande collaboration, *Hyper-Kamiokande Design Report*, 1805.04163.

[378] Hyper-Kamiokande Proto collaboration, *Astroparticle Physics in Hyper-Kamiokande*, *PoS* **EPS-HEP2017** (2017) 020 [1710.08345].

[379] S. Horiuchi, J.F. Beacom and E. Dwek, *The Diffuse Supernova Neutrino Background is detectable in Super-Kamiokande*, *Phys. Rev.* **D79** (2009) 083013 [0812.3157].

[380] R. Laha and J.F. Beacom, *Gadolinium in water Cherenkov detectors improves detection of supernova $\nu_e$*, *Phys. Rev.* **D89** (2014) 063007 [1311.6407].

[381] N.F. Bell, M.J. Dolan and S. Robles, *Searching for Sub-GeV Dark Matter in the Galactic Centre using Hyper-Kamiokande*, 2005.01950.

[382] Super-Kamiokande collaboration, *Atmospheric Neutrino Oscillation Analysis with Improved Event Reconstruction in Super-Kamiokande IV*, *PTEP* **2019** (2019) 053F01 [1901.03230].

[383] Super-Kamiokande collaboration, *Kinematic reconstruction of atmospheric neutrino events in a large water Cherenkov detector with proton identification*, *Phys. Rev. D* **79** (2009) 112010 [0901.1645].

[384] J.F. Beacom and J. Candia, *Shower power: Isolating the prompt atmospheric neutrino flux using electron neutrinos*, *JCAP* **0411** (2004) 009 [hep-ph/0409046].

[385] S. Palomares-Ruiz and S. Pascoli, *Testing MeV dark matter with neutrino detectors*, *Phys. Rev.* **D77** (2008) 025025 [0710.5420].

[386] IceCube collaboration, *Electronics Development for the New Photo-Detectors (PDOM and D-Egg) for IceCube-Upgrade*, in *HAWC Contributions to the 36th International Cosmic Ray Conference (ICRC2019)*, 2019 [1908.11564].

[387] IceCube collaboration, *Calibration LEDs in the IceCube Upgrade D-Egg Modules*, in *HAWC Contributions to the 36th International Cosmic Ray Conference (ICRC2019)*, 2019 [1908.10780].

[388] IceCube collaboration, *Dark matter searches with the IceCube Upgrade*, in *HAWC Contributions to the 36th International Cosmic Ray Conference (ICRC2019)*, 2019 [1908.08236].

[389] IceCube collaboration, *IceCube-Gen2: A Vision for the Future of Neutrino Astronomy in Antarctica*, 1412.5106.

[390] IceCube collaboration, *Neutrino astronomy with the next generation IceCube Neutrino Observatory*, 1911.02561.

[391] Baikal-GVD collaboration, *The Baikal-GVD neutrino telescope: First results of multi-messenger studies*, in *HAWC Contributions to the 36th International Cosmic Ray Conference (ICRC2019)*, 2019 [1908.05450].

[392] A.D. Avrorin et al., *Sensitivity of the Baikal-GVD neutrino telescope to neutrino emission toward the center of the galactic dark matter halo*, *JETP Lett.* **101** (2015) 289 [1412.3672].

[393] KM3Net collaboration, *Letter of intent for KM3NeT 2.0*, *J. Phys.* **G43** (2016) 084001 [1601.07459].

[394] KM3NeT collaboration, *Sensitivity of the KM3NeT/ARCA neutrino telescope to point-like neutrino sources*, *Astropart. Phys.* **111** (2019) 100 [1810.08499].

[395] F. Halzen, *Lectures on high-energy neutrino astronomy*, in *International WE - Heraeus Summer School: Physics with Cosmic Accelerators Bad Honnef, Germany, July 5-16, 2004*, 2005 [astro-ph/0506248].

[396] S.R. Gozzini, *Search for dark matter with the antares and km3net neutrino telescopes*, .

[397] M. Agostini et al., *The Pacific Ocean Neutrino Experiment*, 2005.09493.

[398] STRAW collaboration, *STRAW (STRings for Absorption length in Water): pathfinder for a neutrino telescope in the deep Pacific Ocean*, *JINST* **14** (2019) P02013 [1810.13265].

[399] S. Wissel, J. Alvarez-Muñiz, W.R. Carvalho, A. Romero-Wolf, H. Schoorlemmer and E. Zas, *A New Concept for High-Elevation Radio Detection of Tau Neutrinos*, *EPJ Web Conf.* **216** (2019) 04007.

[400] CTA Consortium collaboration, *Science with the cherenkov telescope array*, 1709.07997.

[401] D. McKeen and N. Raj, *Monochromatic dark neutrinos and boosted dark matter in noble liquid direct detection*, 1812.05102.

[402] M. Frigerio, A. Pomarol, F. Riva and A. Urbano, *Composite Scalar Dark Matter*, *JHEP* **07** (2012) 015 [1204.2808].

[403] Pierre Auger collaboration, *Searches for neutrino fluxes in the EeV regime with the Pierre Auger Observatory*, *PoS* **ICRC2017** (2018) 972.

[404] Ara Collaboration collaboration, *Design and initial performance of the Askaryan Radio Array prototype EeV neutrino detector at the South Pole*, *Astroparticle Physics* **35** (2012) 457 [1105.2854].

[405] Super-Kamiokande collaboration, *Indirect Search for Dark Matter from the Galactic Center and Halo with the Super-Kamiokande Detector*, 2005.05109.

[406] W.H. Press and P. Schechter, *Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation*, *Astrophysical Journal* **187** (1974) 425.

[407] J.R. Bond, S. Cole, G. Efstathiou and N. Kaiser, *Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations*, *Astrophysical Journal* **379** (1991) 440.

[408] R.K. Sheth and G. Tormen, *Large scale bias and the peak background split*, *Mon. Not. Roy. Astron. Soc.* **308** (1999) 119 [astro-ph/9901122].

[409] R.K. Sheth, H.J. Mo and G. Tormen, *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes*, *Mon. Not. Roy. Astron. Soc.* **323** (2001) 1 [astro-ph/9907024].

[410] J.L. Tinker, A.V. Kravtsov, A. Klypin, K. Abazajian, M.S. Warren, G. Yepes et al., *Toward a halo mass function for precision cosmology: The Limits of universality*, *Astrophys. J.* **688** (2008) 709 [0803.2706].

# Revisiting $O(N)$ $\sigma$ model at unphysical pion masses and high temperatures

Yuan-Lin Lyu,[1,*] Qu-Zhi Li,[2,†] Zhiguang Xiao,[2,‡] and Han-Qing Zheng[2,§]

[1] *School of Physics, Peking University, Beijing 100871, P. R. China*
[2] *Institute for Particle and Nuclear Physics, College of Physics, Sichuan University, Chengdu 610065, P. R. China*

(Dated: March 1, 2024)

Roy-equation analyses on lattice data of $\pi\pi$ scattering phase shifts at $m_\pi = 391$MeV reveals that the lowest $f_0$ meson becomes a bound state under this condition. In addition, there is a pair of complex poles below threshold generated by crossing symmetry [1]. We use the $N/D$ method to partially recover crossing symmetry of the $O(N)$ $\sigma$ model amplitude at leading order of $1/N$ expansion, and qualitatively reproduce the pole structure and pole trajectories with varying pion masses as revealed by Roy-equation analyses. The $\sigma$ pole trajectory with varying temperature is also discussed and found to be similar to its properties when varying $m_\pi$. As the temperature increases, the complex $\sigma$ poles firstly move from the second Riemann sheet to the real axis becoming two virtual state poles, and then one virtual state pole moves to the first sheet turning into a bound state pole and finally tends to the pion pole position at high temperature which is as expected from the chiral symmetry restoration. Our results provide further evidences that the lowest $f_0$ state extracted from experiments and lattice data plays the role of $\sigma$ meson in the spontaneous breaking of chiral symmetry. Finally, we also briefly discuss the problems of the effective potential in the situation when $m_\pi$ and temperature get large.

arXiv:2402.19243v1 [hep-ph] 29 Feb 2024

## I. INTRODUCTION

Chiral symmetry breaking plays an important role in the QCD low energy dynamics. It is already well-known that due to the smallness of the $u$ and $d$ quark masses, QCD possesses an approximate $SU(2)_L \times SU(2)_R$ chiral symmetry, and it is also well accepted that this symmetry is broken by the nonzero $\langle 0|\bar{q}q|0\rangle$ and three pesudo-Goldstone bosons are generated which are identified as the $\pi$ mesons observed in the low energy hadron scatterings. Historically, the famous linear sigma model firstly proposed by Gell-Mann and Lévy in 1960 [2] could provide an effective field theory description of this symmetry. In this model, another scalar field $\sigma$ is combined with the three pions to form a linear realization of an $O(4)$ symmetry and acquires a vacuum expectation value to break the $O(4)$ to $O(3)$, where the $O(4) \simeq SU(2)_L \times SU(2)_R$ can be identified as the previous chiral symmetry and the remaining $O(3)$ corresponds to the preserved $SU(2)_V$. For a long time, the existence of the $\sigma$ particle was in controversy. The mild rise of the $\pi\pi$ phase shift can hardly be recognized as generated from a typical resonance. A broad resonance was proposed to describe the $\pi\pi$ scattering phase shift in 1960s, see for example [3–5]. However, such a broad resonance appeared and disappeared from the PDG table several times from the 1960s until 2000s. Another description using a nonlinear realization of the chiral symmetry [6, 7] in which the scalar-isoscalar particle is totally abandoned from the Lagrangian is the nowadays very popular chiral perturbation theory ($\chi$PT) [8, 9], which is regarded as the low energy effective theory of QCD. Within this formalism, the low energy properties of the pion-pion scattering, such as the scattering length, effective range, and phase shifts near the threshold can be reproduced. The low energy coupling constants can be saturated by integrating out vector resonances [10, 11] (see however [12, 13]). Thus, there seems to be no need to have a scalar-isoscalar particle in describing the low energy pion-pion scatterings. However, with energy going up, $\chi$PT blows up quickly. Fortunately, unitarity and dispersive techniques come to its rescue. After unitarization, the $IJ = 00$ channel $\pi\pi$ scattering amplitude dynamically generates a resonance state represented as a pair of conjugate poles on the second Riemann sheet, see for example [14, 15]. However, this kind of unitarized method always generates more poles than physically expected [16], especially spurious poles on the physical sheet, which cast doubts on the reliability of the results, not to mention the violation of crossing symmetry (for a recent review, see Ref. [17]). On the other hand, a novel model-independent analysis representing the partial wave $S$-matrix as a product of pole and the left-hand cut integral terms, showed that the left-hand cut estimated from $\chi$PT always produces a negative contribution to the phase shift while the data show a positive trend near the threshold, which demonstrates the necessity of a subthreshold resonance pole on the second sheet of the amplitude [18]. This method was further developed into the

* yllyu@stu.pku.edu.cn
† liquzhi@scu.edu.cn, corresponding author
‡ xiaozg@scu.edu.cn, corresponding author
§ zhenghq@pku.edu.cn

so-called PKU-representation of the partial wave $S$-matrix [19–23], and more precise mass and width for the particle was obtained by fitting the data with the constraints from Balachandran-Nuyts-Roskies (BNR) relations derived from crossing symmetry [20]. A major step forward in analyzing the pion-pion scattering data is the so called Roy-equation analysis [24] which incorporates crossing symmetry and unitarity into a set of integral equations involving only the partial wave amplitudes in the physical region. By solving these integral equations with the low energy constraints from $\chi$PT [25, 26] and extending the solution to the complex plane, a broad resonance pole can be found around $\sqrt{s_{\rm pole}} = 441^{16}_{-8} - i272^{+9}_{-13}$ MeV [27] (see also [28] and [29, 30]). From these model-independent efforts, the existence of a scalar-isoscalar resonance in the low energy $\pi\pi$ scattering was firmly established, and up to now PDG lists the particle as $f_0(500)$ with a pole mass in $400 \sim 550$ MeV and the half-width around $200 \sim 350$ MeV.

Although these model-independent methods confirm the existence of this particle, it is obscure what the role of this scalar-isoscalar particle plays in the low energy QCD spectrum and chiral symmetry breaking. An argument showing that this $f_0(500)$ is not a usual pure $\bar{q}q$ state is that with large number of colors, $N_c$, the pole generated in the unitarized $\chi$PT goes away from the real axis which is different from the usual meson behavior in the large $N_c$ limit [14]. However, for large enough $N_c$, the pole position still moves towards the real axis [31]. Another problem is whether it really corresponds to the original $\sigma$ particle in the Lagrangian models with linearly realized chiral symmetry [32], which is still not clearly understood (see for example, Ref. [33]). Within the unitarized $\chi$PT framework, the $f_0(500)$ particle is only a dynamically generated resonance without any useful information about the role it plays in the chiral symmetry breaking. This can be understood since $\chi$PT starts off from the broken phase of QCD with a nonzero vacuum expectation value (VEV), it may not provide much information about the global property of the effective potential which is responsible for symmetry breakdown. Moreover, a study of $\chi$PT at high temperatures reveals its drastic difference compared with linear $\sigma$ model: the former simply cannot restore the wanted $O(4)$ symmetry explicitly, though there are some implicit evidences [34–38]. So to analyze the role played by this particle, it is desirable to look back at the linear $\sigma$ model to see whether the sigma particle in this model is consistent with $f_0(500)$ in model-independent analyses. One advantage of $O(N)$ linear $\sigma$ model is that in the large $N$ limit (here $N$ denotes the number of flavors), the model is exactly solvable [39–41] and has been used to study the possible relation between $f_0(500)$ and the $\sigma$ [42].

With the recent development of lattice QCD [43–53], phase shifts of $\pi\pi$ scattering can be reproduced from the first principle at various unphysical pion masses. This provides additional valuable information for the resonances in $\pi\pi$ scattering by extracting pole positions from those phase shifts. In Ref. [51], the scattering phase shifts for $IJ = 00$ channel are calculated at $m_\pi \sim 236$ and 391 MeV, and $K$-matrix parameterization was used to extract the poles in this channel. The result shows that when $m_\pi \sim 236$ MeV the lowest $f_0$ state is still a resonance while at 391 MeV it becomes a bound state. More recent results by HadSpec collaboration using the similar $K$-matrix method shows that at $m_\pi \sim 330$ MeV, $\sigma$ already becomes a shallow bound state, whereas at $m_\pi \sim 283$ MeV it may become a virtual state or a subthreshold resonance [53]. In fact, it has been suggested that in the unitarized $\chi$PT amplitudes with larger pion masses, the dynamically generated $f_0$ particle moves towards the real axis below threshold and finally becomes a bound state [54–56]. However, it is well known that the $K$-matrix approach does not satisfy crossing symmetry [12, 13], which is important in the low energy pion-pion scattering and is crucial in determining the properties of $f_0(500)$ resonance [20, 29, 30]. The first attempt to incorporate crossing symmetry in the analysis is in [57], where the PKU representation combined with BNR relations and a virtual state accompanying the $f_0$ bound state was found at $m_\pi = 391$ MeV. However, it was noticed that the left-hand cut considered in that paper is not complete: the left-hand cut contribution introduced by the $f_0$ bound state from crossed-channels was not taken into account [58]. Not only in Ref. [57], the situation in general is also unsatisfactory. Model-independent lattice data were always analysed using rough models. A precise model-independent analysis of lattice data using Roy equation which incorporates crossing symmetry from the startup was done in Ref. [1]. It is also found that on the second sheet there is a pair of conjugate subthreshold poles generated, which is related to the left-hand cut originated from the $f_0$ bound state in crossed-channels. A general argument of why this subthreshold pole is present was also given in Ref. [1]. Recently, a further lattice study using Roy-equation analysis at $m_\pi \sim 239$ and 283 MeV found that for the former, the lowest $f_0$ state remains a resonance, whereas in the latter case the state was claimed (though not definitively, see Ref. [59] for details) to become a virtual state, accompanied by a "noisy" pole close to the left-hand cut on the second Riemann sheet. However, this additional virtual state pole is generated before the $f_0$ particle turns into a bound state, a situation not considered in Ref. [1]. To one's surprise, this phenomenon is exactly what happens in $N/D$ modified $O(N)$ model and we will explain the details in Sec. III.

The phenomena at unphysical pion masses provide a new test ground of whether the lowest $f_0$ state in model-independent analyses corresponds to the $\sigma$ particle in linear $\sigma$ model. The purpose of this paper is to use the solvable $O(N)$ linear $\sigma$ model with a crossing symmetry improvement at unphysical pion masses and compare it with the results from Roy-equation analyses.

The paper is organized as follows. In Sec. II, a review on standard results of $O(N)$ model is given, especially the behavior of the lowest $f_0$ particle is discussed and compared with similar results obtained by different model analyses

of the lattice data. The comparison necessitates the effort of going beyond the lowest order calculation of $O(N)$ model. In Sec. III, the modified $O(N)$ amplitude incorporating crossing symmetry is introduced, aided by the use of $N/D$ method. The results from Roy-equation analyses on the $\sigma$ particle can be reproduced in this approach at a qualitative level. In particular, the $f_0$ state becomes a bound state and the lower subthreshold pole generated through cross-channel effects also emerges at $m_\pi = 391$ MeV after partially imposing crossing symmetry, by tuning the parameters in $N/D$ method. It is of course well-known that the linear $\sigma$ model is not QCD [8, 9], though this paper try to demonstrate that it provides a qualitative description of low energy QCD in $IJ = 00$ channel at phenomenological level, even for unphysical pion masses – from this observation we take the perspective that $f_0(500)$ particle plays the role of the $\sigma$ particle. In Sec. IV, we also investigate thermal properties of the scattering amplitudes with leading order $1/N$ expansion. We reproduce the widely accepted results that $O(N)$ symmetry is restored at high temperature, irrespective of different $m_\pi$ values. Finally Sec. V is for discussions and conclusions, where we shortly discuss the problem that the effective potential no longer provides a local minimum at high temperature, as well as when $m_\pi$ gets large. Future improvements on the related issues are also outlined there.

## II. $\sigma$-POLE WITH VARYING $m_\pi$ IN $O(N)$ MODEL

In this section, we will review the $O(N)$ linear $\sigma$ model and look at the $\sigma$-pole trajectory with varying $m_\pi$, at leading order of $1/N$ expansion. The Lagrangian for this model is

$$\mathcal{L}_{O(N)} = \frac{1}{2}\partial_\mu\phi_a\partial^\mu\phi_a - \frac{1}{2}\mu_0^2\phi_a\phi_a - \frac{\lambda_0}{8N}(\phi_a\phi_a)^2 + \alpha\phi_N\,, \tag{1}$$

where $a = 1, 2, \cdots, N$. When $\mu_0^2 < 0$, without the linear $\alpha$ term, the system has a spontaneous symmetry breaking of $O(N) \to O(N-1)$, $\langle\phi\rangle \neq 0$. With the linear $\phi_N$ term, the VEV is aligned with the $N$-th direction, namely, $\langle\phi_N\rangle = v$, with $N-1$ pseudo-Goldstone particles, $\pi^a \equiv \phi^a$, $a = 1, \ldots, N-1$ and we define the shifted field $\sigma \equiv \phi_N - v$. For convenience when counting $1/N$ orders in the calculation of the effective action, we introduce an auxiliary field $\chi$ to the Lagrangian [41],

$$\mathcal{L} \to \mathcal{L} + \frac{N}{2\lambda_0}\left(\chi - \frac{\lambda_0}{2N}\phi_a\phi_a - \mu_0^2\right)^2 = \frac{1}{2}\partial_\mu\phi_a\partial^\mu\phi_a + \alpha\phi_N + \frac{N}{2\lambda_0}\chi^2 - \frac{1}{2}\chi\phi_a\phi_a - \frac{N\mu_0^2}{\lambda_0}\chi\,, \tag{2}$$

with an irrelevant constant omitted. The effective action can be obtained by standard procedures,

$$\begin{aligned}\Gamma(\phi,\chi) = &\int \mathrm{d}^4x \left(\frac{1}{2}\partial_\mu\phi_a\partial^\mu\phi_a + \alpha\phi_N + \frac{N}{2\lambda_0}\chi^2 - \frac{1}{2}\chi\phi_a\phi_a - \frac{N\mu_0^2}{\lambda_0}\chi\right)\\ &+ \frac{i}{2}N\mathrm{Tr}\log(\partial^2 + \chi - i\epsilon)\,,\end{aligned} \tag{3}$$

where Tr denotes trace taken in 4-dimensional Minkowski spacetime and $\epsilon \to 0^+$. For convenience we will not distinguish the notation for the classical fields and the original fields in the Lagrangian, which can be understood in the context. Then the effective potential can be obtained as

$$V(\phi,\chi) = -\alpha\phi_N - \frac{N}{2\lambda_0}\chi^2 + \frac{1}{2}\chi\phi_a\phi_a + \frac{N\mu_0^2}{\lambda_0}\chi - \frac{i}{2}N\int\frac{d^4\ell}{(2\pi)^4}\log(-\ell^2 + \chi - i\epsilon)\,, \tag{4}$$

where $\phi_a$, $\chi$ are regarded as their constant expectation values, respectively. The renormalization conditions are chosen to be [41, 60]

$$\frac{\mu(M)^2}{\lambda(M)} = \frac{\mu_0^2}{\lambda_0} + \frac{i}{2}\int\frac{d^4\ell}{(2\pi)^4}\frac{1}{\ell^2 + i\epsilon}\,, \tag{5}$$

$$\frac{1}{\lambda(M)} = \frac{1}{\lambda_0} - \frac{i}{2}\int\frac{d^4\ell}{(2\pi)^4}\frac{1}{(\ell^2 + i\epsilon)(\ell^2 - M^2 + i\epsilon)}\,. \tag{6}$$

With these conditions, the minimum of the effective potential can be obtained by solving $\frac{\partial V}{\partial\chi} = 0$ and $\frac{\partial V}{\partial\phi_a} = 0$ which reduce to relations of the VEVs for the corresponding fields

$$\phi_a\phi_a = \frac{2N}{\lambda}\chi - \frac{2N\mu^2}{\lambda} - \frac{N}{16\pi^2}\chi\log\frac{\chi}{M^2}\,, \tag{7}$$

$$\chi\phi_a = 0\ (a < N), \quad \chi\phi_N - \alpha = 0\,. \tag{8}$$

Thus, the VEV for each $\pi_a$ is still zero, i.e. $\langle\pi_a\rangle = 0$ and $\phi_N$ gets expectation value $\langle\phi_N\rangle = v = \alpha/\langle\chi\rangle$. Since the correction to the pion mass term only appears at higher $1/N$ order, at leading order we have $\langle\chi\rangle = m_\pi^2$. At zero-temperature, from the definition of the pion decay constant $f_\pi$, $\langle 0|A_a^\mu(x)|\pi\rangle = ip^\mu f_\pi e^{-ip\cdot x}$, where $A_a^\mu$ is the axial vector current, and the relation of partially conserved axial current (PCAC), $\partial_\mu A_a^\mu = \alpha\pi_a$, we have $v = f_\pi$. For phenomenological calculations, we always set $f_\pi = 92.4\,\mathrm{MeV}$ and $N = 4$.

With these preparations, now the $\pi\pi$ scattering amplitude in the leading $1/N$ order can be expressed according to the external isospin structure as follows

$$\mathcal{T}_{\pi_a\pi_b\to\pi_c\pi_d} = iD_{\tau\tau}(s)\delta_{ab}\delta_{cd} + iD_{\tau\tau}(t)\delta_{ac}\delta_{bd} + iD_{\tau\tau}(u)\delta_{ad}\delta_{bc}\,, \tag{9}$$

with $D_{\tau\tau}$ the propagator for $\tau \equiv \chi - \langle\chi\rangle$ field from the effective action. Since there are mixing terms of $\tau$ and $\sigma$, a $2\times 2$ inverse propagator matrix needs to be considered [41, 60], which can be expressed in the momentum space as

$$D^{-1}(p^2) = -i\begin{pmatrix} p^2 - m_\pi^2 & -f_\pi \\ -f_\pi & N/\lambda_0 + NB_0(p^2, m_\pi)\end{pmatrix}, \tag{10}$$

with

$$B_0(p^2, m_\pi) = \frac{-i}{2}\int \frac{d^4\ell}{(2\pi)^4}\frac{1}{(\ell^2 - m_\pi^2 + i\epsilon)((\ell+p)^2 - m_\pi^2 + i\epsilon)}\,. \tag{11}$$

The propagators can then be obtained as

$$D_{\tau\tau}(p^2) = \frac{i(p^2 - m_\pi^2)}{(p^2 - m_\pi^2)(N/\lambda_0 + NB_0(p^2, m_\pi)) - f_\pi^2}, \tag{12}$$

$$D_{\sigma\sigma}(p^2) = \frac{i(1/\lambda_0 + B_0(p^2, m_\pi))}{(p^2 - m_\pi^2)(1/\lambda_0 + B_0(p^2, m_\pi)) - f_\pi^2/N}\,. \tag{13}$$

Using the previous renormalization conditions Eqs.(5,6), we have

$$\frac{1}{\lambda_0} + B_0(p^2, m_\pi) = \frac{1}{\lambda(M)} + B(p^2, m_\pi, M)\,, \tag{14}$$

$$B(s, m_\pi, M) = \frac{1}{32\pi^2}\left(1 + \rho(s)\log\frac{\rho(s)-1}{\rho(s)+1} - \log\frac{m_\pi^2}{M^2}\right), \tag{15}$$

where $\rho(s) = \sqrt{1 - 4m_\pi^2/s}$. For our purpose, we need only $IJ = 00$ amplitude in the leading order of $1/N$ expansion,

$$\mathcal{T}_{00}^{LO}(s) = \frac{iND_{\tau\tau}(s)}{32\pi}\,, \tag{16}$$

which is of $\mathcal{O}(1)$.

The $\sigma$ resonance corresponds to poles on the second Riemann sheet, which can be obtained by solving the zero points of the denominator of $D_{\tau\tau}$:

$$(s - m_\pi^2)(1/\lambda(M) + B^{II}(s, m_\pi, M)) - f_\pi^2/N = 0\,, \tag{17}$$

where $B^{II}$ represents the analytically continued $B(s, m_\pi, M)$ function onto the second sheet, which is obtainable by changing the sign of $\rho(s)$.

Since the coupling constant $\lambda$ and the renormalization scale $M$ are related, we can define $M$ to be the intrinsic scale of $O(N)$ model, when regarded as an effective field theory, at which the coupling blows up, i.e. $1/\lambda(M) = 0$ [60]. Then the coupling $\lambda$ would not appear in the scattering amplitude:

$$\mathcal{T}_{00}^{LO}(s) = -\frac{1}{32\pi}\frac{s - m_\pi^2}{(s - m_\pi^2)B(s, m_\pi, M) - f_\pi^2/N}\,, \tag{18}$$

and Eq.(17) can be recast into

$$(s - m_\pi^2)B^{II}(s, m_\pi, M) - f_\pi^2/N = 0\,. \tag{19}$$

The leading order amplitude automatically satisfies the exact partial wave unitarity, i.e. $\mathrm{Im}\,\mathcal{T}_{00}(s) = \rho(s)|\mathcal{T}_{00}(s)|^2$. It also has an Adler zero [61] at $s = m_\pi^2$. Notice that the isospin projection for $O(N)$ singlet channel is $\mathcal{T}_{I=0}(s,t,u) =$

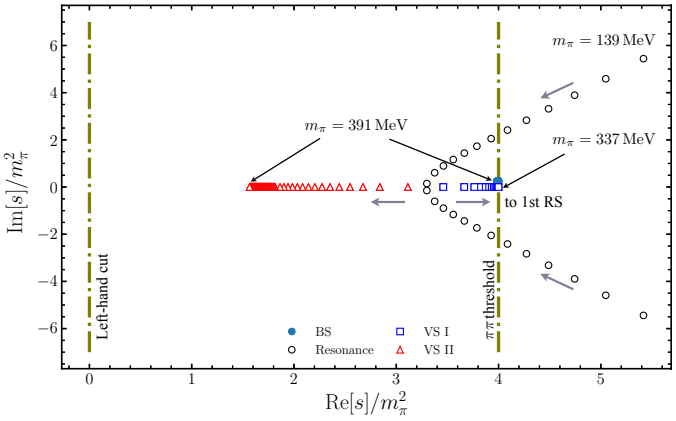



FIG. 1. The pole trajectories for the $\sigma$ poles of the leading order $O(N)$ amplitude. The left-hand cut branch point is still at $s_L = 0$ after the virtual state pole VS I moves across threshold to the physical Riemann sheet (RS), becoming a bound state. There is neither additional virtual state pole generated from the left-hand cut (suggested by Refs. [1, 59]), nor subthreshold poles similar to those found in Ref. [1].

| $m_\pi$ (MeV) | 139 | 239 | 283 | 330 | 391 |
|---|---|---|---|---|---|
| $O(N)$ (LO) | $356 - i148$ | $448 - i57$ | 558(VS I)<br>438(VS II) | 660(VS I)<br>451(VS II) | 780(BS)<br>489(VS II) |
| $N/D$ modified $O(N)$ | $348 - i180$ | 469(BS)<br>426(VS II)<br>168(VS III) | 527(BS)<br>422(VS II)<br>264(VS III) | 585(BS)<br>$396 - i28$<br>(Sub. pole) | 658(BS)<br>$466 - i77$<br>(Sub. pole) |
| lattice + $K$-matrix | | $(487 \sim 809$<br>$-i136 \sim 304)$[51, 53] | $(476 \sim 579$<br>$-i0 \sim 129)$[53] | $657^{+3}_{-4}$(BS)[53] | $758 \pm 4$(BS)[51] |
| lattice + Roy Eq. | | $(416 \sim 644$<br>$-i176 \sim 307)$[1, 59] | $522 \sim 562$<br>(VS I&II)[59] [a] | | $759^{+7}_{-16}$(BS)[1]<br>$269^{+40}_{-25} - i211^{+26}_{-23}$<br>(Sub. pole)[1] |

[a] Additionally, there is a third though "noisy" pole close to the left-hand cut on the second sheet, which could correspond to the virtual state pole VS III appearing in the $N/D$ modified $O(N)$ model analysis. However, a definitive conclusion about whether the $\sigma$ is a virtual state or a subthreshold resonance at this $m_\pi$ value cannot be made, owing to large statistical uncertainties in the results, see Ref. [59] for details.

TABLE I. Comparison of the pole positions ($\sqrt{s_{\rm pole}}$) for $O(N)$ model, lattice + $K$-matrix [51, 53] and lattice + Roy equation [1, 59]. When $m_\pi = 391$ MeV, the subthreshold (Sub.) pole close to the left-hand cut in Ref. [1] can also be found (qualitatively) within the $N/D$ modified $O(N)$ model discussed in this work.

$(N-1)iD_{\tau\tau}(s) + iD_{\tau\tau}(t) + iD_{\tau\tau}(u)$. Thus the $t$- and $u$-channel amplitudes in Eq. (9) contribute to the $\mathcal{O}(1/N)$ amplitude, and only the $s$-channel amplitude contributes to the leading order $IJ = 00$ partial wave amplitude. This already breaks crossing symmetry.

With the leading order $\mathcal{T}$-matrix, we can study the $\sigma$ pole trajectory with varying $m_\pi$. The intrinsic scale $M$ is chosen at 550 MeV. See Table I for comparison of the pole positions obtained in $O(N)$ model and the lattice results analyzed using $K$-matrix and Roy equation respectively. The pole trajectory is shown in Fig. 1. When $m_\pi$ increases from the physical mass, the $\sigma$ poles move towards the real axis, and then become two virtual state poles (VS I&II) after they meet at the real axis. One virtual state pole (VS II) moves down away and the other (VS I) moves towards the threshold, crossing it to the first sheet, and becomes a bound state pole (BS).

From Eq. (19) one can work out the condition for the critical pion mass $m_\pi = m_c$ at which the $\sigma$ pole is located

exactly at the threshold,

$$\log\frac{m_c^2}{M^2} = 1 - \frac{32\pi^2 f_\pi^2}{3m_c^2 N}\,. \tag{20}$$

The numerical result is $m_c \simeq 337\,\mathrm{MeV}$. This is a little bit different from the lattice results which may be located at somewhere between 283 and 330 MeV [51, 53, 59]. The difference is not surprising since the $m_c$ obtained here is only the leading $1/N$ order result of $O(N)$ model. When $m_\pi > m_c$, the $\sigma$ particle becomes a bound state. At the same time, it also appears in the $t$- and $u$-channel amplitudes due to crossing symmetry, which are not included in the leading order amplitude since they are of $\mathcal{O}(1/N)$ after isospin projection. If we add the $t$- and $u$-channel contributions and do the isospin and partial wave projection, then the $\mathcal{T}_{00}$ matrix is expressed as

$$\mathcal{T}_{00}(s) = \frac{N-1}{32\pi}\mathcal{A}^{LO}(s) + I_{tu}(s)\,, \tag{21}$$

where

$$\mathcal{A}^{LO}(s) = \frac{m_\pi^2 - s}{(s-m_\pi^2)NB(s,m_\pi,M) - f_\pi^2}\,, \tag{22}$$

$$I_{tu}(s) = \frac{1}{16\pi(s-4m_\pi^2)}\int_{4m_\pi^2-s}^{0}\mathrm{d}t\mathcal{A}^{LO}(t)\,. \tag{23}$$

Function $I_{tu}(s)$ in Eq. (21) comes from the partial wave projection of the crossed $t$- and $u$-channel amplitudes. The bound state in the crossed-channels also generates a left-hand cut with a branch point at $4m_\pi^2 - m_\sigma^2$. However, this amplitude does not satisfy the exact unitarity any more. Thus, to partially recover crossing symmetry and restore unitarity, we need to resort to some unitarization methods, e.g. inverse amplitude method (IAM, see Ref. [62] and references therein) and $N/D$ method [63]. The situation here is not suitable for direct application of IAM, because the partial wave amplitude Eq. (21) is not a complete calculation at $\mathcal{O}(1/N)$ and thus will break even the perturbative version of unitarity relation, namely

$$\mathrm{Im}\,\mathcal{T}_{00}^{NLO} = \rho\left[\mathcal{T}_{00}^{LO}(\mathcal{T}_{00}^{NLO})^* + (\mathcal{T}_{00}^{LO})^*\mathcal{T}_{00}^{NLO}\right] = \frac{2\rho\,\mathrm{Re}\,T_{00}^{LO}\,\mathrm{Re}\,T_{00}^{NLO}}{1-2\rho\,\mathrm{Im}\,T_{00}^{LO}}\,, \tag{24}$$

where we also represent $\mathrm{Im}\,\mathcal{T}_{00}^{NLO}$, i.e. the imaginary part of the next-to-leading order amplitude, in terms of $\mathrm{Im}\,\mathcal{T}_{00}^{LO}$ and $\mathrm{Re}\,\mathcal{T}_{00}^{NLO}$ after simple algebraic calculations. Nevertheless, it is still possible to acquire an approximation of the $\mathcal{O}(1/N)$ partial wave (denoted by $T_{00}^{NLO\prime}$) that does satisfy the above relation by setting $\mathrm{Re}\,\mathcal{T}_{00}^{NLO} = -\frac{1}{N}\mathrm{Re}\,\mathcal{T}_{00}^{LO} + I_{tu}$ and obtaining $\mathrm{Im}\,\mathcal{T}_{00}^{NLO}$ using Eq. (24). After the analytic continuation of the modified amplitude $T_{00}^{NLO\prime}$, the approximated "IAM unitarized amplitude" $\mathcal{T}_{00}^{\mathrm{IAM}}$ can be constructed as usual:

$$\mathcal{T}_{00}^{\mathrm{IAM}} = \frac{\left(\mathcal{T}_{00}^{LO}\right)^2}{\mathcal{T}_{00}^{LO} - \mathcal{T}_{00}^{NLO\prime}}\,. \tag{25}$$

However, this unitarized amplitude is found to have spurious poles both on the 1st and 2nd sheets that may not be simply removed. Thus in the following we will use the $N/D$ method.

## III. $\sigma$ POLE IN $N/D$ MODIFIED $O(N)$ MODEL WITH VARYING $m_\pi$

It has been seen that in $O(N)$ model the isospin projection to $I = 0$ channel breaks crossing symmetry at the leading order of $1/N$ expansion. The $O(N)$ amplitude must be modified in order to at least partially recover crossing symmetry while preserving unitarity. The key point is to generate the left-hand cut from crossed-channels in the scattering amplitude, which should be consistent with the bound state the $\sigma$ generated. It is well known that the $N/D$ method can introduce the left-hand cut contribution to the amplitude and at the same time preserve partial wave unitarity, which is suitable for the purposes here.

In the spirit of $N/D$ method, the $\mathcal{T}$-matrix can be expressed as

$$\mathcal{T}(s) = \frac{N(s)}{D(s)}\,, \tag{26}$$

where the singularity of $N(s)$ contains only the left-hand cut ($L$), while $D(s)$ only contains right-hand cut ($R$) and Castillejo-Dalitz-Dyson (CDD) poles [64], thus $\mathrm{Im}_R\, N(s) = \mathrm{Im}_L\, D(s) = 0$. To satisfy the partial wave unitarity relation, i.e. $\mathrm{Im}_R\, \mathcal{T}^{-1} = -\rho$, the relation between $N(s)$ and $D(s)$ should be

$$\mathrm{Im}_R\, D(s) = -\rho(s) N(s)\,, \tag{27}$$

$$\mathrm{Im}_L\, N(s) = D(s)\, \mathrm{Im}_L\, \mathcal{T}(s)\,. \tag{28}$$

Using the Cauchy integral formula, one can write down dispersion relations for $N(s)$ and $D(s)$, and then solve $N(s)$ and $D(s)$ numerically to obtain the scattering amplitude.

Our strategy is to extract the $\mathrm{Im}_L\mathcal{T}$ from $O(N)$ model, use $N/D$ method to obtain the scattering amplitude, which at the leading $1/N$ order recovers the original $O(N)$ model amplitude, and require the position of the $\sigma$ bound state to be consistent with the left-hand cut branch point at $s_L = 4m_\pi^2 - m_\sigma^2$ for large unphysical $m_\pi$. Since the numerator and denominator of the amplitude in Eq. (18) both have an $s$ power less than $s^2$ as $s \to \infty$, it is natural to use *twice subtracted* dispersion relations for $N(s)$ and $D(s)$ without CDD poles, and as a bonus, the zero point of the $\mathcal{T}$-matrix can be dynamically generated, which could correspond to the Adler zero. The twice subtracted dispersion relations for $N(s)$ and $D(s)$ can be expressed as

$$D(s) = \frac{s - s_A}{s_0 - s_A} + g_D \frac{s - s_0}{s_A - s_0} - \frac{(s - s_0)(s - s_A)}{\pi} \int_R \frac{\rho(s') N(s')}{(s' - s)(s' - s_0)(s' - s_A)} \mathrm{d}s'\,, \tag{29}$$

$$N(s) = b_0 \frac{s - s_A}{s_0 - s_A} + g_N \frac{s - s_0}{s_A - s_0} + \frac{(s - s_0)(s - s_A)}{\pi} \int_L \frac{D(s') \mathrm{Im}_L \mathcal{T}(s')}{(s' - s)(s' - s_0)(s' - s_A)} \mathrm{d}s'\,, \tag{30}$$

where the subtraction points $s_A = m_\pi^2$, $s_0 = s_{th} = 4m_\pi^2$, and $D(s_0) = 1$ are chosen for convenience. With these choices, the subtraction constants are $N(s_0) = b_0$, $N(s_A) = g_N$, $D(s_A) = g_D$. Noticing that at $s_A = m_\pi^2$, the leading order partial wave amplitude in Eq. (18) is zero, corresponding to the Adler zero. Then with $\mathcal{T}(s_A) = N(s_A)/D(s_A) \sim \mathcal{O}(1/N)$, one can choose $g_N$ to be $\mathcal{O}(1/N)$ and $g_D$ to be $\mathcal{O}(1)$. Thus to the leading order, $N(s) = b_0 \frac{s - s_A}{s_0 - s_A}$. By substituting this into Eq. (29) one recovers the leading order scattering amplitude by choosing

$$b_0 = -\frac{1}{32\pi} \frac{s_0 - s_A}{(s_0 - s_A) B(s_0, m_\pi, M) - f_\pi^2/N}\,, \tag{31}$$

$$g_D = \frac{32\pi f_\pi^2 b_0}{N(s_0 - s_A)}\,. \tag{32}$$

It is obvious that $b_0$ is just the leading order amplitude which is consistent with our prescription $N(s_0) = b_0$ and $D(s_0) = 1$. After including $t$- and $u$-channel contributions in the $N/D$ construction, the amplitude recovers unitarity and has a left-hand cut with a branch point at $4m_\pi^2 - m_\sigma^2$ when $m_\pi > m_c$, which is as expected. However, since the $\sigma$ pole in the crossed-channels is still located at the same position in the leading order amplitude, the branch point at $4m_\pi^2 - m_\sigma^2$ is determined by the leading order $m_\sigma$ in $1/N$ expansion. But with the $\mathcal{O}(1/N)$ left-hand cut contribution in $D(s)$, the mass of the $\sigma$ bound state may be shifted from the pole position of $\mathcal{T}_{00}^{LO}$. To be consistent with the left-hand cut, demanded by crossing symmetry, we need to require the $\sigma$ bound state pole position (solved from $D(s) = 0$) to be the same as the one in leading order amplitude. In general, this can be achieved by properly choosing the higher order corrections to the parameters $b_0$ and $g_D$. For $m_\pi > m_c$ this can only be done numerically, our prescription is as follows:

$$b_0 = \mathcal{T}_{00}(s_0) = \frac{N - 1}{32\pi} \mathcal{A}^{LO}(s_0) + I_{tu}(s_0)\,, \tag{33}$$

$$D(m_\sigma^2) = 0\,, \tag{34}$$

$$\frac{g_N}{g_D} = I_{tu}(s_A)\,. \tag{35}$$

where $b_0$ is chosen as the amplitude evaluated at $s_0$ with the crossed-channels added, $g_D$ is determined by requiring the bound state pole position to be the same as the leading order result $s_\sigma = m_\sigma^2$, and Eq. (35) results directly from the requirement $\mathcal{T}(s_A) = \mathcal{T}_{00}(s_A)$. When $m_\pi < m_c$, Eq. (34) is replaced by Eq. (32) with Eq. (33) and Eq. (35) not changed. This prescription for the subtraction constants captures the most important features of the inputted $O(N)$ model. Anyway, it is of course not the "unique solution" for $N/D$ modified $O(N)$ model. We also tried several different sets of subtraction constants, and found that the qualitative results for the $\sigma$ pole trajectory are quite robust.

The numerical results of the pole structure as $m_\pi$ increases are shown in Fig. 2. In order to obtain a $\sigma$ pole close to the one in the leading order $O(N)$ amplitude at physical pion mass, the intrinsic scale $M$ is set to 1.5 GeV with

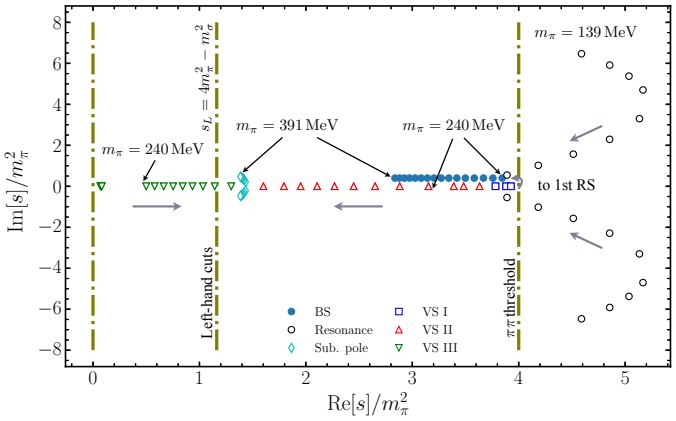



FIG. 2. The pole trajectories for $N/D$ modified $O(N)$ model. The left-hand cut branch point extends to $s_L = 4m_\pi^2 - m_\sigma^2$ after the virtual state pole VS I moves across threshold to the physical Riemann sheet (RS), turning into a bound state. Qualitatively, the $N/D$ modified $O(N)$ model reproduces the picture derived from $K$-matrix [51, 53] and Roy-equation [1, 59] analyses of the lattice data.

$m_c \simeq 214\,\mathrm{MeV}$ accordingly. When pion mass increases from the physical value $m_\pi = 139$ MeV, the $\sigma$ poles move towards the real axis and at some point they hit the real axis and separate into two virtual state poles, with one moving up (VS I) and the other moving down (VS II) along the real axis. The upper virtual state, VS I, moves across the threshold to the physical Riemann sheet and then moves down along the real axis, hence becoming a bound state. At the same time, there will be a left-hand cut generated with the branch point located at $s_L = 4m_\pi^2 - m_\sigma^2 > 0$ in the partial wave amplitude, which comes from the $\sigma$-exchange in the $t$- and $u$-channels from crossing symmetry. During this course but before VS I turns into a bound state, an additional virtual state pole (VS III) is generated from the left-hand cut, which was also implied in Ref. [59]. With $m_\pi$ growing larger, VS II and VS III hit each other, move into the complex plane, and become a pair of subthreshold poles. In fact, in Ref. [1], it was shown that besides a $\sigma$ bound state below the threshold, there is also a pair of conjugate poles below the threshold when $m_\pi = 391$ MeV, which corresponds to the pair of subthreshold poles in Fig. 2. It was argued that such subthreshold poles are inevitable due to the behavior of the left-hand cut generated from the $\sigma$ exchange in crossed $t$- and $u$-channels. A more careful argument goes as follows. Considering that the $S$-matrix near $s_L$ is mainly contributed by the singular behavior of the branch point of the log cut generated from the cross-channel $\sigma$ bound state pole, if the residue of the $\sigma$ pole is positive [65], which is always the case in practice, the sign of the partial wave $S$-matrix immediately above $s_L$ on the real axis can be proved to be negative (opposite to the sign of the residue). If $m_\pi$ is not large enough such that VS II is still on the real axis, the $S$-matrix below threshold between $s_L$ and the $\sigma$ pole should have two zero points as shown in the middle graph of Fig. 3. The two zeros of $S$-matrix mean that besides VS II generated from the original $\sigma$ poles another virtual state pole VS III is generated from the branch point of the left-hand cut [66]. As the pion mass increases further, the $S$-matrix between the branch point and the bound state becomes smaller such that VS II and VS III move towards each other, and then they hit each other at some point and move into the complex plane, becoming a new resonance. The final situation is illustrated in the right graph of Fig. 3 and is what found at $m_\pi = 391\,\mathrm{MeV}$. Admittedly, both the above argument and the one in Ref. [1] can only ensure the existence of VS III when there is a $\sigma$ bound state.

However as pointed out above, the situation is a little different in $N/D$ modified $O(N)$ model, where the additional virtual state pole, VS III, is generated before the $\sigma$ becomes a bound state (see the left graph of Fig. 3), i.e. without the left-hand cut generated by the $\sigma$-bound-state exchange in the crossed-channels. The origin of this VS III state is found to be related with the interplay of Adler zero and the left-hand cut. As $m_\pi$ increases, we found that the real Adler zero, moving towards the left-hand cut, will hit the branch point $s_L = 0$ and go into the complex plane, becoming a pair of conjugate zeros. Remarkably, complex Adler zeros at $m_\pi \sim 391$ MeV were first found in Ref. [1] and real Adler zero can not be found in Ref. [59] for $m_\pi \sim 283$ MeV, which may indicate that the above behavior of the Adler zero could be general. Then, it is straightforward to prove that VS III is generated when Adler zero passes $s_L$ as follows. The partial wave $S$-matrix is defined as $S(s) = 1 + 2i\rho(s)\mathcal{T}(s)$, thus $S(s) \to -\mathrm{sign}(\mathcal{T}(s))\infty$ when $s \to s_L + 0^+$, as long as $\lim_{s\to s_L+0^+} \mathcal{T}(s) \neq 0$. [67] Then the Adler zero, i.e. the simple zero point of $\mathcal{T}(s)$, passing $s_L$, causes a sign change of $\mathcal{T}(s)$ from negative to positive and hence also a sign change of $S(s)$, at $s \to s_L + 0^+$. This phenomenon has to occur before the $\sigma$ turns into a bound state, since according to the argument of the previous paragraph, the sign of $S$-matrix close to the branch point can not be flipped when $\sigma$ remains a bound state. Considering that the $S$-matrix equals to 1 both at the threshold and at the Adler zero, there can only be even number of $S$-matrix zeros between

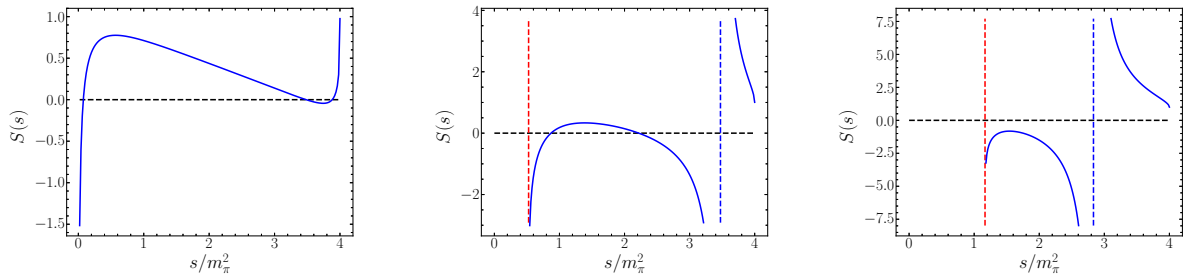



FIG. 3. Blue solid line: physical-sheet S-matrix below threshold obtained in the $N/D$ modified $O(N)$ model. Blue dashed line: position of the $\sigma$ bound state. Red dashed line: branch point of the left-hand cut ($s_L = 4m_\pi^2 - m_\sigma^2$). The left graph ($m_\pi = 207$ MeV) : two virtual states in the near threshold region and one additional virtual state pole generated close to the left-hand cut. The middle graph ($m_\pi = 283$ MeV) : one bound state with two virtual states. The right graph ($m_\pi = 391$ MeV): two virtual state poles have become a pair of resonance poles.

these two points. Thus before the Adler zero passes $s_L$, if there are two virtual states, i.e. VS I&II, generated after the $\sigma$ resonance poles hit the real axis, they must both be within that region. For smaller $m_\pi$ values, if there is no VS III and the Adler zero has not hit the left-hand cut branch point yet, thus $S(s)$ is zero-free between the Adler zero and $s_L$. Then $S(s)$ has to be positive on that interval. When Adler zero moves through the branch point, it causes the sign of $S$-matrix near the left-hand cut changing from positive to negative. This sign flip of $S(s)$ in the vicinity of $s_L$ indicates the generation of a $S$-matrix zero point from the branch point of the left-hand cut as shown in the left graph in Fig. 3, which exactly results in the appearance of VS III. As the pion mass grows further, the situation is similar to the description in the previous paragraph where VS I moves to the first sheet becoming a bound state, and VS II hit VS III then both moving into the complex plane. Thus, the existence of VS III and hence the subthreshold poles at large pion mass, e.g. $m_\pi \sim 391$ MeV, is actually the combined result of unitarity, crossing symmetry, Adler zero (which is a significant feature for low energy chiral dynamics) and the analyticity of the $S$-matrix in this region.

The $N/D$ modified $O(N)$ model reproduces the picture derived from Roy-equation analyses, which demonstrates that the $\sigma$ particle in the $O(N)$ model really can represent, *at the qualitative level*, the $f_0$ state extracted by Roy equation from the lattice data. Conversely, this also means that the lowest $f_0$ state in the low energy $\pi\pi$ scattering really plays the similar role of the $\sigma$ in $O(N)$ linear $\sigma$ model — providing the vacuum expectation value for spontaneous breaking of chiral symmetry.

## IV. SIGMA POLE TRAJECTORY WITH TEMPERATURE FOR DIFFERENT $m_\pi$

It is also instructive to look at the $\sigma$ pole trajectory in a finite temperature environment. It is well known that in the chiral limit, i.e., when the pion is massless at zero-temperature, under high temperature the system goes through a phase transition from the chiral symmetry broken phase to the chiral symmetric phase at a critical temperature $T_c$ [68–71] (e.g. for $O(N)$ model without explicit symmetry breaking [72, 73], $T_c = \sqrt{12/N} f_\pi \simeq 160\,\text{MeV}$, which can be easily read out in Fig. 4). One would expect that above the phase transition temperature the massless pion gets massive and the sigma particle would be degenerate with pions. With explicit chiral symmetry breaking where the pion has a mass at zero-temperature, there is no explicit phase transition point. But with temperature going higher and higher, the system asymptotically approaches the chiral symmetric phase, where the mass of the sigma tends to the pion mass. Notice that this whole picture can not be explicitly realized in $\chi$PT, since it is constructed intrinsically in the broken phase. This means that $\chi$PT is valid only for an energy or temperature far below the chiral symmetry breaking scale, which is determined by the VEV of the scalar field or $\bar{q}q$. Under each of the circumstances: (a) the VEV approaches zero; (b) the energy scale or (c) the temperature, becomes comparable with or goes beyond the chiral symmetry breaking scale, the whole theory breaks down. For example, in Refs. [36, 38], the sigma pole was found to be still in the complex plane of the second Riemann sheet around $T = 200$ MeV, which seems to be hopeless to restore the $O(4)$ symmetry in $\chi$PT around $T_c$.

The above picture can be easily seen in $O(N)$ linear $\sigma$ model. The leading order effective potential at finite temperature $T$ can be obtained with imaginary-time formalism [72–75] by Wick rotation to Euclidean space and substituting the momentum integral with a sum over Matsubara frequencies $\omega_n = 2\pi n T$, i.e. $\int \frac{\mathrm{d}^4 k}{(2\pi)^4} f(k_0, \mathbf{k}) \to iT \sum_n \int \frac{\mathrm{d}^3 \mathbf{k}}{(2\pi)^3} f(k_0 = i\omega_n, \mathbf{k})$ (see, e.g. Ref. [76] for details). The renormalization condition is chosen the same as the

$T = 0$ case. Then by minimizing the effective potential, we can obtain the gap equations for $v$ and $m_\pi^2$ as functions of temperature [73]:

$$v^2(T) = f_\pi^2 + \frac{N}{16\pi^2}\left(m_\pi^2 \log\frac{m_\pi^2}{M^2} - m_\pi^2(T)\log\frac{m_\pi^2(T)}{M^2}\right) - N A^{T\neq 0}\left(m_\pi^2(T)\right), \tag{36}$$

$$\alpha = v(T) m_\pi^2(T), \tag{37}$$

where $v(0) = f_\pi$ and $m_\pi(0) = m_\pi$ are set to zero-temperature values. The function $A^{T\neq 0}$ is defined as the finite temperature contribution to the tadpole integral encountered in the derivation of Eq. (7),

$$A^{T\neq 0}\left(m_\pi^2(T)\right) = \int_0^\infty \frac{\mathrm{d}k}{2\pi^2}\frac{k^2 n_B(\omega_k)}{\omega_k}, \tag{38}$$

where $\beta = 1/T$, $\omega_k = \sqrt{k^2 + m_\pi^2(T)}$ and $n_B(\omega_k) = (e^{\beta\omega_k} - 1)^{-1}$ is the Bose-Einstein distribution. With different values of $m_\pi(0)$, denoting the magnitude of explicit breaking, the solutions of $v(T)$ and $m_\pi(T)$ are shown in Fig. 4.

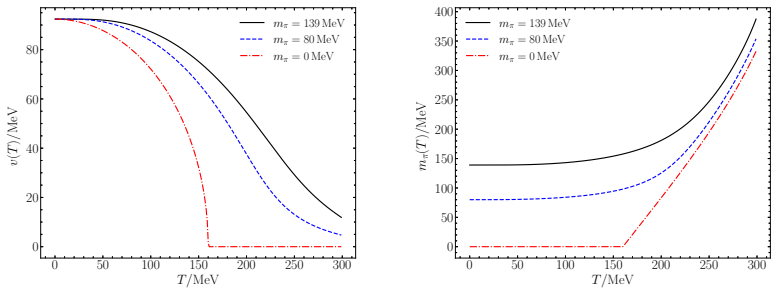



FIG. 4. $v(T)$ (left) and $m_\pi(T)$ (right). The three cases are identified by $m_\pi(0) = 139$, 80 and 0 MeV. In the chiral limit, there is a second-order phase transition with the critical temperature $T_c \simeq 160\,\mathrm{MeV}$. With a nonzero $m_\pi(0)$, there is no explicit transition point.

To study the spectrum at finite temperature in the center of mass (CM) frame, the scattering amplitude is generalized to be the amputated four-point Green's function with the external momenta analytically continued back to on-shell momenta in Minkowski space after the Matsubara sum [38, 77–79]. The leading $1/N$ order $\pi\pi$ scattering amplitude with finite temperature can be expressed as,

$$\mathcal{T}_{00}^T(s) = -\frac{1}{32\pi}\frac{s - m_\pi^2(T)}{(s - m_\pi^2(T))\; B^T\left(s, m_\pi(T), M\right) - v^2(T)/N}, \tag{39}$$

where $B^T\left(s, m_\pi(T), M\right)$, the finite temperature version of $B\left(s, m_\pi, M\right)$ defined in Eq. (15), can be obtained by standard calculations,

$$B^T\left(s, m_\pi(T), M\right) = B\left(s, m_\pi(T), M\right) + B^{T\neq 0}\left(s, m_\pi(T)\right), \tag{40}$$

$$B^{T\neq 0}\left(s, m_\pi(T)\right) = \int_0^\infty \frac{\mathrm{d}k\, k^2}{8\pi^2\omega_k^2} n_B(\omega_k)\left(\frac{1}{E + 2\omega_k} - \frac{1}{E - 2\omega_k}\right), \tag{41}$$

with $B^{T\neq 0}$ evaluated in the CM frame and $s = E^2$. The $\sigma$ resonance pole can be obtained from the zero point of the denominator of $\mathcal{T}_{00}^T$ on the second Riemann sheet. The mass and width for the $\sigma$ pole with varying temperature and with $m_\pi(0) = 200$, 139 and 80 MeV respectively are illustrated in Fig. 5, in which $m_\sigma$ and $\Gamma_\sigma$ are also compared to the behavior of $m_\pi$ at finite temperature for each case. As temperature increases, $\sigma$ resonance firstly becomes even broader due to the growth in phase space caused by the Bose-Einstein distribution $n_B(\omega_k)$, and then $\Gamma_\sigma$ drops rapidly to zero when $\sigma$ turns into a pair of virtual state poles. For the present purpose, to demonstrate the asymptotic degeneration of $\sigma$ and $\pi$'s, we only keep track of the virtual state pole, VS I (named in the same way as the zero-temperature case), which moves up through threshold to the physical sheet and becomes a $\sigma$ bound state at around $T_c$. Then $m_\sigma$ changes gradually and moves closer to $m_\pi(T)$ when $T > T_c$. Furthermore, $m_\sigma$ and $m_\pi(T)$ asymptotically tends to converge at $T \gg T_c$, which is as expected by the restoration of chiral symmetry at high temperature limit.

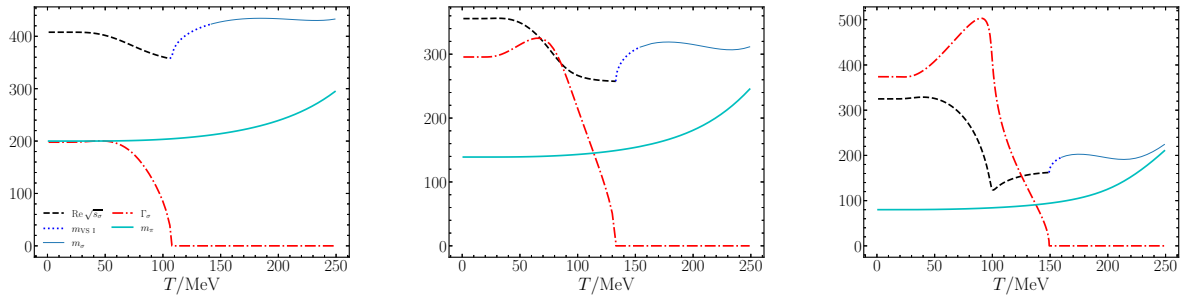



FIG. 5. The mass and width of $\sigma$ pole with varying temperature compared with $m_\pi(T)$. From left to right, $m_\pi(0) = 200$, 139 and 80 MeV respectively. The lower virtual state pole VS II is not shown above.

## V. CONCLUSIONS AND DISCUSSIONS

In this paper, we have investigated the $O(N)$ linear $\sigma$ model with varying $m_\pi$ to reproduce the subthreshold pole structure of the $IJ = 00$ channel $\pi\pi$ scattering amplitude, recently found using the Roy equation in analyzing the lattice phase shifts at several unphysical $m_\pi$ values. By using the $N/D$ method to partially recover crossing symmetry of the partial wave amplitude in $O(N)$ model, the Roy-equation analysis results can be roughly reproduced. The pole trajectory illustrates the picture that with large $m_\pi$, as the Adler zero goes into the complex plane before $\sigma$ becomes a bound state, besides the virtual state accompanying the $\sigma$ bound state, another virtual state pole will originate from the left-hand cut. As $m_\pi$ grows larger, the two virtual state poles hit each other and scattered into the complex plane and become a pair of resonance poles.

The consistency of $O(N)$ model with the Roy-equation analyses reveals that the lowest $f_0$ state from the Roy-equation analyses of the lattice data can be fairly described by the $\sigma$ field in $O(N)$ linear $\sigma$ model, thus plays the same role in the chiral symmetry breaking as the $\sigma$ particle in $O(N)$ model.

The $\sigma$ pole behavior with varying temperature is also discussed in the $O(N)$ model with different pion masses. It is shown that with higher temperature the sigma resonance pole also moves to the real axis and becomes a pair of virtual states. Then the upper one moves across threshold to the first Riemann sheet and turns into a bound state. With much higher temperatures, the mass of the sigma and pion would tend to come closer and closer, which is as expected by chiral symmetry restoration in the high temperature limit.

It is worth mentioning that $O(N)$ model itself suffers from the vacuum stability problem. In short, the original $O(N)$ model effective potential is obtained by solving the auxiliary field $\chi(\phi^2)$ as a function of $\phi_a$ from one of the gap equations $\frac{\partial V}{\partial \chi} = 0$, which has two branches of solutions: one with ordinary chiral symmetry breaking vacuum (i.e. $m_\pi \to 0$ when $\alpha \to 0$), the other with a larger pion mass which restores chiral symmetry ($v = 0$ but $m_\pi \neq 0$) in the $\alpha \to 0$ limit [41, 80–84]. In perturbation theory, $V(\phi)$ is expanded around the vacuum which is a stationary point of the effective potential. Whether it is a local minimum, local maximum or saddle point can be determined by the Hessian matrix of $V(\phi)$. After a careful investigation we found that the vacuum chosen at $v = f_\pi$ with $\chi(v^2) = m_\pi^2$ remains a local minimum on the first branch when $m_\pi < M/\sqrt{e}$. However for larger pion mass, the stable minimum exists only on the second branch. This is a warning that when $m_\pi$ gets large, the vacuum of the $O(N)$ model becomes unstable – a phenomenon not known in our knowledge of QCD [85]. A similar problem also happens at high temperature: when $T \gg T_c$, the effective potential no longer provides a local minimum on its first branch. Instead, the local minimum will move to the second branch and become a saddle point. In spite of these deficiencies, we insist on the opinion that the linear $\sigma$ model on the first branch provides the correct picture for describing low energy QCD, since high dimensional terms and other resonance terms, missed in the present discussion based on the toy linear sigma model, may alleviate the vacuum stability problem. After all, it is worthwhile to look at the $m_\pi$ dependence of the effective potential which is not discussed in the literature yet, to the best of our knowledge. We will present the details of these discussions elsewhere.

Another direction to be explored is to look at the behavior of the $N^*(920)$ recently found in the $\pi N$ scattering from the Roy-equation analysis [86–88] under different pion masses or temperatures. The role $N^*(920)$ plays in the $\pi N$ scattering is in some sense similar to the $\sigma$ particle in $\pi\pi$ scattering. If it shares similar properties at higher temperatures, one would expect that $N^*(920)$ would become a bound state and the parity partner of the nucleon, and thus may play an important role in physics.

## ACKNOWLEDGMENTS

This work is supported by China National Natural Science Foundation under contract No. 12335002, 12375078, 11975028. This work is also supported by "the Fundamental Research Funds for the Central Universities".

[1] X.-H. Cao, Q.-Z. Li, Z.-H. Guo, and H.-Q. Zheng, Phys. Rev. D **108**, 034009 (2023), arXiv:2303.02596 [hep-ph].
[2] M. Gell-Mann and M. Lévy, Nuovo Cim. **16**, 705 (1960).
[3] M. M. Islam and R. Pinon, Phys. Rev. Lett. **12**, 310 (1964).
[4] S. H. Patil, Phys. Rev. Lett. **13**, 261 (1964).
[5] V. Hagopian, W. Selove, J. Alitti, and J. P. Baton, Phys. Rev. **145**, 1128 (1966).
[6] S. R. Coleman, J. Wess, and B. Zumino, Phys. Rev. **177**, 2239 (1969).
[7] C. G. Callan, Jr., S. R. Coleman, J. Wess, and B. Zumino, Phys. Rev. **177**, 2247 (1969).
[8] J. Gasser and H. Leutwyler, Annals Phys. **158**, 142 (1984).
[9] J. Gasser and H. Leutwyler, Nucl. Phys. B **250**, 465 (1985).
[10] G. Ecker, J. Gasser, A. Pich, and E. de Rafael, Nucl. Phys. B **321**, 311 (1989).
[11] J. F. Donoghue, C. Ramirez, and G. Valencia, Phys. Rev. D **39**, 1947 (1989).
[12] Z. H. Guo, J. J. Sanz Cillero, and H. Q. Zheng, JHEP **06**, 030, arXiv:hep-ph/0701232.
[13] Z. H. Guo, J. J. Sanz-Cillero, and H. Q. Zheng, Phys. Lett. B **661**, 342 (2008), arXiv:0710.2163 [hep-ph].
[14] J. R. Pelaez, Mod. Phys. Lett. A **19**, 2879 (2004), arXiv:hep-ph/0411107.
[15] A. Gomez Nicola and J. R. Pelaez, Phys. Rev. D **65**, 054009 (2002), arXiv:hep-ph/0109056.
[16] G.-Y. Qin, W. Z. Deng, Z. Xiao, and H. Q. Zheng, Phys. Lett. B **542**, 89 (2002), arXiv:hep-ph/0205214.
[17] D.-L. Yao, L.-Y. Dai, H.-Q. Zheng, and Z.-Y. Zhou, Rept. Prog. Phys. **84**, 076201 (2021), arXiv:2009.13495 [hep-ph].
[18] Z. Xiao and H. Q. Zheng, Nucl. Phys. A **695**, 273 (2001), arXiv:hep-ph/0011260.
[19] H. Q. Zheng, Z. Y. Zhou, G. Y. Qin, Z. Xiao, J. J. Wang, and N. Wu, Nucl. Phys. A **733**, 235 (2004), arXiv:hep-ph/0310293.
[20] Z. Y. Zhou, G. Y. Qin, P. Zhang, Z. Xiao, H. Q. Zheng, and N. Wu, JHEP **02**, 043, arXiv:hep-ph/0406271.
[21] H. Q. Zheng, Z. Y. Zhou, G. Y. Qin, and Z. Xiao, AIP Conf. Proc. **717**, 322 (2004), arXiv:hep-ph/0309242.
[22] J.-j. Wang, Z. Y. Zhou, and H. Q. Zheng, JHEP **12**, 019, arXiv:hep-ph/0508040.
[23] Z. Y. Zhou and H. Q. Zheng, Nucl. Phys. A **775**, 212 (2006), arXiv:hep-ph/0603062.
[24] S. M. Roy, Phys. Lett. B **36**, 353 (1971).
[25] G. Colangelo, J. Gasser, and H. Leutwyler, Nucl. Phys. B **603**, 125 (2001), arXiv:hep-ph/0103088.
[26] B. Ananthanarayan, G. Colangelo, J. Gasser, and H. Leutwyler, Phys. Rept. **353**, 207 (2001), arXiv:hep-ph/0005297.
[27] I. Caprini, G. Colangelo, and H. Leutwyler, Phys. Rev. Lett. **96**, 132001 (2006), arXiv:hep-ph/0512364.
[28] R. Garcia-Martin, R. Kaminski, J. R. Pelaez, J. Ruiz de Elvira, and F. J. Yndurain, Phys. Rev. D **83**, 074004 (2011), arXiv:1102.2183 [hep-ph].
[29] G. Mennessier, S. Narison, and W. Ochs, Phys. Lett. B **665**, 205 (2008), arXiv:0804.4452 [hep-ph].
[30] G. Mennessier, S. Narison, and X. G. Wang, Phys. Lett. B **688**, 59 (2010), arXiv:1002.1402 [hep-ph].
[31] Z. X. Sun, L. Y. Xiao, Z. Xiao, and H. Q. Zheng, Mod. Phys. Lett. A **22**, 711 (2007), arXiv:hep-ph/0503195.
[32] It is well known that the original renormalizable toy $\sigma$ model of Gell-Mann and Lévy is not QCD, as proven in Refs. [8, 9].
[33] S. Weinberg, Phys. Rev. Lett. **110**, 261601 (2013), arXiv:1303.0342 [hep-ph].
[34] J. Gasser and H. Leutwyler, Phys. Lett. B **188**, 477 (1987).
[35] P. Gerber and H. Leutwyler, Nucl. Phys. B **321**, 387 (1989).
[36] S. Cortés, A. Gómez Nicola, and J. Morales, Phys. Rev. D **93**, 036001 (2016), arXiv:1511.00031 [hep-ph].
[37] A. Gomez Nicola and J. Ruiz de Elvira, Phys. Rev. D **97**, 074016 (2018), arXiv:1704.05036 [hep-ph].
[38] R. Gao, Z.-H. Guo, and J.-Y. Pang, Phys. Rev. D **100**, 114028 (2019), arXiv:1907.01787 [hep-ph].
[39] L. Dolan and R. Jackiw, Phys. Rev. D **9**, 3320 (1974).
[40] H. J. Schnitzer, Phys. Rev. D **10**, 1800 (1974).
[41] S. R. Coleman, R. Jackiw, and H. D. Politzer, Phys. Rev. D **10**, 2491 (1974).
[42] Z.-H. Guo, L. Y. Xiao, and H. Q. Zheng, Int. J. Mod. Phys. A **22**, 4603 (2007), arXiv:hep-ph/0610434.
[43] M. Luscher, Commun. Math. Phys. **105**, 153 (1986).
[44] M. Luscher and U. Wolff, Nucl. Phys. B **339**, 222 (1990).
[45] M. Luscher, Nucl. Phys. B **354**, 531 (1991).
[46] Y. Kuramashi, M. Fukugita, H. Mino, M. Okawa, and A. Ukawa, Phys. Rev. Lett. **71**, 2387 (1993).
[47] S. He, X. Feng, and C. Liu, JHEP **07**, 011, arXiv:hep-lat/0504019.
[48] N. Mathur, A. Alexandru, Y. Chen, S. J. Dong, T. Draper, I. Horvath, F. X. Lee, K. F. Liu, S. Tamhankar, and J. B. Zhang, Phys. Rev. D **76**, 114505 (2007), arXiv:hep-ph/0607110.
[49] X. Feng, K. Jansen, and D. B. Renner, Phys. Lett. B **684**, 268 (2010), arXiv:0909.3255 [hep-lat].
[50] Z. Fu, Phys. Rev. D **85**, 014506 (2012), arXiv:1110.0319 [hep-lat].
[51] R. A. Briceno, J. J. Dudek, R. G. Edwards, and D. J. Wilson, Phys. Rev. Lett. **118**, 022002 (2017), arXiv:1607.05900 [hep-ph].

[52] H.-W. Lin *et al.* (Hadron Spectrum), Phys. Rev. D **79**, 034502 (2009), arXiv:0810.3588 [hep-lat].
[53] A. Rodas, J. J. Dudek, and R. G. Edwards (Hadron Spectrum), Phys. Rev. D **108**, 034513 (2023), arXiv:2303.10701 [hep-lat].
[54] C. Hanhart, J. R. Pelaez, and G. Rios, Phys. Rev. Lett. **100**, 152001 (2008), arXiv:0801.2871 [hep-ph].
[55] J. R. Pelaez and G. Rios, Phys. Rev. D **82**, 114002 (2010), arXiv:1010.6008 [hep-ph].
[56] C. Hanhart, J. R. Pelaez, and G. Rios, Phys. Lett. B **739**, 375 (2014), arXiv:1407.7452 [hep-ph].
[57] X.-L. Gao, Z.-H. Guo, Z. Xiao, and Z.-Y. Zhou, Phys. Rev. D **105**, 094002 (2022), arXiv:2202.03124 [hep-ph].
[58] X.-L. Gao, Z.-H. Guo, Z. Xiao, and Z.-Y. Zhou, Phys. Rev. D **107**, 058502 (2023), arXiv:2204.01562 [hep-ph].
[59] A. Rodas, J. J. Dudek, and R. G. Edwards (Hadron Spectrum), Phys. Rev. D **109**, 034513 (2024).
[60] R. S. Chivukula and M. Golden, Phys. Lett. B **267**, 233 (1991).
[61] S. L. Adler, Phys. Rev. **137**, B1022 (1965).
[62] A. Gomez Nicola, J. R. Pelaez, and G. Rios, Phys. Rev. D **77**, 056006 (2008), arXiv:0712.2763 [hep-ph].
[63] G. F. Chew and S. Mandelstam, Phys. Rev. **119**, 467 (1960).
[64] L. Castillejo, R. H. Dalitz, and F. J. Dyson, Phys. Rev. **101**, 453 (1956).
[65] In the nonrelativistic scattering theory, the residual of the scattering amplitude is related to $-\gamma^2$ for $s$ wave, where $\gamma \in \mathbb{R}$ is defined using the asymptotic behavior of the wave function $\eta(r) \xrightarrow{r\to\infty} \gamma e^{-\alpha r}$ [89], thus the residue for the $S$-matrix at the $\sigma$ bound state is positive.
[66] The second sheet $S$-matrix is the inverse of the one defined on the physical sheet, owing to unitarity and analyticity.
[67] The situation here is similar to but much more complicated than that of physical $m_\pi$ [20].
[68] R. D. Pisarski and F. Wilczek, Phys. Rev. D **29**, 338 (1984).
[69] A. Bazavov *et al.*, Phys. Rev. D **85**, 054503 (2012), arXiv:1111.1710 [hep-lat].
[70] A. Bazavov *et al.* (HotQCD), Phys. Lett. B **795**, 15 (2019), arXiv:1812.08235 [hep-lat].
[71] H. T. Ding *et al.* (HotQCD), Phys. Rev. Lett. **123**, 062002 (2019), arXiv:1903.04801 [hep-lat].
[72] A. Bochkarev and J. I. Kapusta, Phys. Rev. D **54**, 4066 (1996), arXiv:hep-ph/9602405.
[73] J. O. Andersen, D. Boer, and H. J. Warringa, Phys. Rev. D **70**, 116007 (2004), arXiv:hep-ph/0408033.
[74] H. Meyers-Ortmanns, H. J. Pirner, and B. J. Schaefer, Phys. Lett. B **311**, 213 (1993).
[75] H. Meyer-Ortmanns, Rev. Mod. Phys. **68**, 473 (1996), arXiv:hep-lat/9608098.
[76] M. L. Bellac, *Thermal Field Theory*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 2011).
[77] E. Quack, P. Zhuang, Y. Kalinovsky, S. P. Klevansky, and J. Hufner, Phys. Lett. B **348**, 1 (1995), arXiv:hep-ph/9410243.
[78] N. Kaiser, Phys. Rev. C **59**, 2945 (1999).
[79] A. Gomez Nicola, F. J. Llanes-Estrada, and J. R. Pelaez, Phys. Lett. B **550**, 55 (2002), arXiv:hep-ph/0203134.
[80] M. Kobayashi and T. Kugo, Prog. Theor. Phys. **54**, 1537 (1975).
[81] L. F. Abbott, J. S. Kang, and H. J. Schnitzer, Phys. Rev. D **13**, 2212 (1976).
[82] A. D. Linde, Nucl. Phys. B **125**, 369 (1977).
[83] W. A. Bardeen and M. Moshe, Phys. Rev. D **28**, 1372 (1983).
[84] W. A. Bardeen and M. Moshe, Phys. Rev. D **34**, 1229 (1986).
[85] The situation of $\chi$PT is even worse due to the Oströgradski instability, such that the $\chi$PT Hamiltonian is unbounded from below.
[86] X.-H. Cao, Q.-Z. Li, and H.-Q. Zheng, JHEP **12**, 073, arXiv:2207.09743 [hep-ph].
[87] Y.-F. Wang, D.-L. Yao, and H.-Q. Zheng, Eur. Phys. J. C **78**, 543 (2018), arXiv:1712.09257 [hep-ph].
[88] M. Hoferichter, J. R. de Elvira, B. Kubis, and U.-G. Meißner, (2023), arXiv:2312.15015 [hep-ph].
[89] J. R. Taylor, *Scattering Theory: The Quantum Theory of Nonrelativistic Collisions* (John Wiley & Sons, Inc., New York, 1972).

# Minimal SU(5) theory on the edge: the importance of being effective

Goran Senjanović*
*Arnold Sommerfeld Center, Ludwig-Maximilians University, Munich, Germany and International Centre for Theoretical Physics, Trieste, Italy*

Michael Zantedeschi†
*Tsung-Dao Lee Institute and School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, China*

It is well known that the minimal renormalizable $SU(5)$ grand unified theory is ruled out: it predicts same masses of down quarks and charged leptons, the gauge couplings do not unify and neutrinos are massless. We show here that all this can be cured simultaneously by the addition of higher-dimensional effective operators. However, the theory lives on the edge since the unification scale turns out as low as roughly $10^{14}$ GeV, threatening proton longevity. If the lower bound on the proton lifetime was to be increased by an order of magnitude, the usual desert in energies between the weak and unification scales would be populated. We also revisit two minimal extensions of this theory that offer a dynamical seesaw origin of neutrino mass, and discuss the resulting consequences.

arXiv:2402.19224v1 [hep-ph] 29 Feb 2024

## I. INTRODUCTION

By unifying electro-weak and strong interactions in a single compact gauge group, grand unified theories (GUTs) explain the mystery of charge conjugation in nature and in turn predict the existence of magnetic monopoles [1, 2]. Moreover, they imply proton decay and provide a rough estimate of its lifetime, tantalizingly close to the present experimental limits [3].

While there are many models on the market, the original $SU(5)$ theory of Georgi and Glashow [4] stands out due to its simplicity and predictivity. Actually, in its simplest renormalizable form, it is predictive enough to be ruled out by experiment due to two fundamental failures: it predicts equal masses for charged leptons and down quarks, and gauge couplings do not unify - $\alpha_1$ meets $\alpha_2$ too early, contrary to what seemed originally [5]. Furthermore, neutrinos end up being massless, but that can be cured by adding fermion singlets - without altering the gauge structure of the theory - which through the seesaw mechanism [6–10], naturally provide small neutrino masses.

There are two distinctive ways of potentially salvaging the theory. One is to add more fields, and take a road of model building, a field interesting in itself - something, however, we will not pursue here. Alternatively, one can employ higher-dimensional effective operators to correct the bad fermion mass relations [11] and provide non-vanishing neutrino masses [12], and even ensure gauge coupling unification [13].

We follow here the latter road and reanalyze the minimal $SU(5)$ theory, augmented by dimension $d = 5$ operators. A rather low unification scale and seemingly too rapid proton decay, prompted a belief that the theory was ruled out, repeated often over the years. The point, however, is that proton could be stable in the limit of the tiny third generation CKM angle going to zero [14], in which case, proton lifetime would be enhanced by some five orders of magnitude. This has been discussed at length [15–19], but still, no serious attempt was made to verify the validity of the theory. After all, it was failing on three fronts as we mentioned above, and it all indicated that it could not survive experimental challenges. And so, the theory kept being sentenced to death by experts, including the present authors (especially one of them [20]). By today it became a gospel - the most recent review of proton decay [21] even cites an extension of this theory as the minimal one.

We show, however, that there exists a region of parameter space where this theory is still phenomenologically viable. It lives on the edge though - an experimental improvement on proton lifetime limit by a factor 10 (20) would point towards the existence of a new light scalar state, the color octet or the weak triplet, below 100 TeV (10 TeV) energies. In other words, the infamous desert at energies between the weak and the GUT scales would not exist in this case, and further improvements on proton lifetime limits could finally rule out the remaining region of the parameter space. Until that happens, we believe that desires to bury what is arguably the minimal grand unified theory, should be put to rest.

The rest of this work, devoted to demonstrating our claim, is organized as follows. In the next Section, the central features of the minimal $SU(5)$ theory are summarized. In Sec. III, its gauge-coupling running is analyzed, with particular focus on the impact of particle thresholds, as well as the effect of $d = 5$ operators on unification conditions. Sec. IV discusses the flavour of proton decay, and its constraints on the unification scale. In Sec. V, we revisit the predictions of two minimal models that generate neutrino mass through renormalisable interactions. Our findings are summarized in Sec. VI.

* goran.senjanovic@physik.uni-muenchen.de
† zantedeschim@sjtu.edu.cn

## II. THE MINIMAL $SU(5)$ THEORY

The minimal $SU(5)$ theory contains three fermion generations of the following representations

$$\overline{5}^i_{\rm F}, \quad 10^i_{\rm F}, \qquad i=1,2,3\,, \tag{1}$$

and the adjoint and fundamental Higgs scalar representations

$$24_{\rm H}, \quad 5_{\rm H}\,,. \tag{2}$$

The adjoint Higgs is responsible for the GUT scale symmetry breaking $M_{\rm GUT}$, whereas the fundamental one provides the electro-weak scale $M_{\rm W}$, with

$$\begin{aligned}\langle 24_{\rm H}\rangle &= M_{\rm GUT}\,{\rm diag}\,(1,1,1,-3/2,-3/2)\,;\\ \langle 5_{\rm H}\rangle^T &= M_{\rm W}\,(0,0,0,0,1)\,.\end{aligned} \tag{3}$$

Strictly speaking, $\langle 24_{\rm H}\rangle$ receives a correction after the electro-weak breaking $\langle 5_{\rm H}\rangle$ is turned on. This has to be small, and thus we omit it here; for a discussion with the phenomenological implications for the $W$-boson mass [22], see the last part of Sec. IV.

The renormalizable $d=4$ Yukawa interaction is given by

$$\mathcal{L}^{d=4}_{\rm y} = \overline{5}_{\rm F} Y_{\rm d}\, 5^*_{\rm H} 10_{\rm F} + 10_{\rm F}\, Y_{\rm u}\, 5_{\rm H} 10_{\rm F}\,, \tag{4}$$

where we ignore the generation indices, as in the rest of the paper. Since $\langle 5_{\rm H}\rangle$, *per se*, keeps unbroken $SU(4)$ symmetry between charged leptons and down quarks, their masses end up the same. This fails badly, ruling out the theory at the renormalizable level.

This can be easily corrected by the addition of $d=5$ Yukawa couplings [11]

$$\mathcal{L}^{d=5}_{\rm y} = \frac{1}{\Lambda}\overline{5}_{\rm F} 24_{\rm H}\, 5^*_{\rm H} 10_{\rm F} + \dots\,, \tag{5}$$

where $\Lambda$ denotes the cutoff due to some unknown new physics, and we assume $\Lambda \gtrsim 10\, M_{\rm GUT}$ in order to keep the $SU(5)$ symmetry well defined.

What happens is the following. Since $\langle 24_{\rm H}\rangle$ breaks the accidental $SU(4)$ of $\langle 5_{\rm H}\rangle$, the effective Yukawa couplings and fermion mass matrices end up being arbitrary, implying arbitrary unitary matrices that diagonalize them. This is why we stop at the leading $d=5$ term - the additional terms are unnecessary. The bottom line is that the bad predictions are gone but this has a dramatic impact on proton decay, as we will see in Sec. IV.

It turns out that at the GUT scale, one has $m_\tau \simeq 2\, m_b$ (see e.g. [23]) instead of them being equal as (4) would suggest. This implies that the cutoff cannot lie far from the unification scale - $\Lambda \lesssim 100 M_{\rm GUT}$. Remarkably, as we will see below, a similar bound emerges from the requirement of gauge coupling unification.

The more stringent limit on the cutoff seemingly arises from neutrino mass considerations. Since neutrino mass vanishes at the renormalisable level, its leading contribution stems from the $d=5$ operator [12]

$$\mathcal{L}_{\not L} = \frac{c_\nu}{\Lambda}\overline{5}_{\rm F} 5_{\rm H} 5_{\rm H} \overline{5}_{\rm F}, \tag{6}$$

where $c_\nu$ is a dimensionless parameter. One obtains in turn

$$m_\nu \simeq \frac{c_\nu}{2}\frac{v^2}{\Lambda}\,, \tag{7}$$

where a factor $1/2$ arises from the renormalization from high to low energies [24]. From $m_\nu \lesssim 10^{-1}$eV, one gets $\Lambda \lesssim 10^{14}$GeV for $c_\nu \sim \mathcal{O}(1)$, which, at face value, would imply a too low unification scale. If one were to live on the edge of perturbativity with $c_\nu \sim \mathcal{O}(4\pi)$, one would increase the limit on the cutoff scale by an order of magnitude $\Lambda \lesssim 10^{15}$GeV - but more about it later.

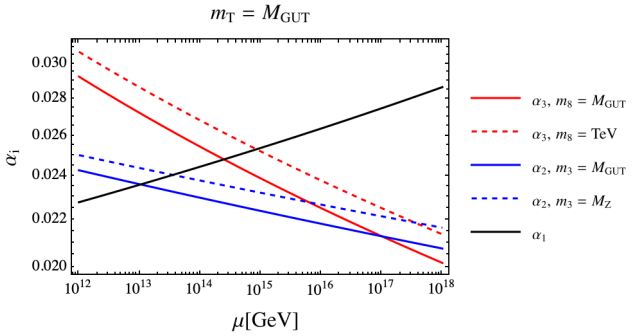



FIG. 1. The failure of $m_3$ and $m_8$ at unifying the gauge couplings.

## III. GAUGE COUPLING UNIFICATION

It is known that the standard model particle content alone does not suffice to unify the gauge couplings of electro-weak and strong interactions. In particular, $\alpha_1$ meets $\alpha_2$ at around $10^{12-13}$GeV, while $\alpha_2$ and $\alpha_3$ meet much later, at $10^{16-17}$GeV. In principle, though, new massive states can affect the running and potentially ensure unification.

In the minimal $SU(5)$, besides the new heavy gauge bosons which do not run below $M_{\rm GUT}$, there are only three additional states, all scalars: a weak triplet and a colour octet from $24_{\rm H}$, and a colour triplet T from $5_{\rm H}$. We denote their masses respectively as $m_3$, $m_8$ and $m_{\rm T}$. Since T induces proton decay it is customarily required to be heavier than $10^{12}$GeV. However, it could be light, in principle, with its Yukawa couplings partially canceled by higher-dimensional operators [25]. In any case, its impact on the gauge coupling running is marginal.

Moreover, when the dust settles (see Section IV), proton longevity requires specific mixing angles in the heavy gauge sector, thus eliminating the freedom of the Yukawa

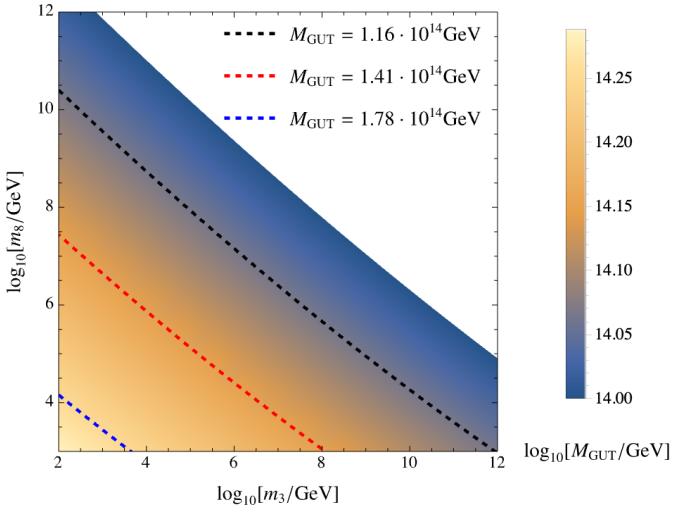



FIG. 2. Dependence of $M_{\rm GUT}$ on $m_3$ and $m_8$. As discussed in the text, $m_{\rm T} \simeq M_{\rm GUT}$.

couplings of T - forcing it to lie close to the GUT scale. In what follows, we thus take $m_{\rm T} \simeq M_{\rm GUT}$, and vary $m_3$ and $m_8$ between $M_{\rm Z}$ and $M_{\rm GUT}$ to explore all of the phenomenologically allowed parameter space of the model.

In Fig. 1 it is shown explicitly how freedom of $m_3$ and $m_8$ does not suffice to obtain unification. Note that taking $m_{\rm T}$ away from $M_{\rm GUT}$ only makes things (slightly) worse. It is clear that the best case scenario is to have $m_3$ ($m_8$) as small (large) as possible. Notice that a small $m_3$ could modify the SM W−boson mass through an induced $\langle 3_H \rangle$ [26]. The triplet decay rates are then uniquely determined by W−mass deviation and $m_3$ [22, 27].

Since the particle thresholds are not sufficient, one must include the additional $d = 5$ gauge boson kinetic energy [13]

$$\Delta\mathcal{L}_{\rm kin} = \frac{c_{\rm F}}{\Lambda} \,{\rm Tr}\, F_{\mu\nu} \langle 24_{\rm H} \rangle F^{\mu\nu} \,, \tag{8}$$

which modifies unification conditions to [13]

$$\begin{aligned} (1-\epsilon)\,\alpha_3(M_{\rm GUT}) &= \left(1+\frac{3}{2}\epsilon\right)\alpha_2(M_{\rm GUT}) \\ &= \left(1+\frac{1}{2}\epsilon\right)\alpha_1(M_{\rm GUT})\,, \end{aligned} \tag{9}$$

with the small expansion parameter $\epsilon = c_{\rm F} M_{\rm GUT}/\Lambda$. The sign of $d = 5$ term in (8) is arbitrary, and for the unification to work $\epsilon$ must be positive, so that (9) can accommodate the problem of $\alpha_1$ being bigger than $\alpha_2$ at the GUT scale, as suggested by the running. While this is achieved by choice, positive $\epsilon$ automatically ensures that the $\alpha_3$ coupling remains large enough at the unification scale.

It turns out [28] that the $\epsilon$ correction is sufficient *per se* (i.e., with $m_3 \simeq m_8 \simeq M_{\rm GUT}$, in accord with the survival principle which states that particle masses should lie at the largest possible scale allowed by symmetries in question [29]) to achieve unification, with $\epsilon \sim 0.04$ and $M_{\rm GUT} \simeq 5 \cdot 10^{13}$GeV. It is noteworthy that this works precisely in the physically allowed region $10 M_{\rm GUT} \lesssim \Lambda \lesssim M_{\rm pl}/\sqrt{N}$, where the upper bond stems from the fact that $\Lambda$ can be at most as big as the perturbative gravity cutoff [30, 31]. Here $N$ is the number of degrees of freedom of the theory, on the order of 100.

As shown in the next section, this would be in tension with the limits on the proton lifetime. However, by utilizing the freedom of $m_3$ and $m_8$, one can increase the GUT scale to actually make the theory viable. This can be understood from the following simple relation

$$\begin{aligned} \frac{M_{\rm GUT}}{M_{\rm Z}} = \exp\Bigg\{ \frac{\pi}{21} \Bigg[ 5\left(1+\frac{\epsilon}{4}\right)\alpha_1^{-1} - 3\left(1+\frac{5\,\epsilon}{4}\right)\alpha_2^{-1} \\ -2\left(1-\frac{5\,\epsilon}{4}\right)\alpha_3^{-1} \Bigg] \Bigg\} \left(\frac{M_{\rm Z}^2}{m_3\, m_8}\right)^{\frac{1}{42}}, \end{aligned} \tag{10}$$

where the gauge couplings are evaluated at $M_{\rm Z}$ and $\epsilon$ turns out to be in the range $0.02 \lesssim \epsilon \lesssim 0.06$.

It is interesting that, once unification is achieved, the color octet needs to be as light as possible in order for the unification scale to be large enough to keep the proton sufficiently long lived. In fact, for $m_3 \simeq m_8 \simeq M_{\rm Z}$, the unification scale becomes the highest, providing its absolute upper limit

$$M_{GUT} \lesssim 2 \cdot 10^{14}\text{GeV}\,. \tag{11}$$

This can be seen explicitly in Fig. 2, where it is shown how $M_{\rm GUT}$ depends on the mass scales $m_3$ and $m_8$. For completeness, a 2-loop RG analysis was performed, taking into account proper matching conditions for thresholds [32, 33].

The reader is probably worried that proton decay is still too fast, even for the above maximal value (11) of the unification scale. However, as we show in the next section, we wish to reassure them that all is well - the freedom in fermion flavor mixing angles [14–19] can keep the proton sufficiently long lived.

## IV. THE FLAVOUR OF PROTON DECAY

The proton decay in this theory stems from two different sources: the heavy X and Y gauge bosons and the new color triplet scalar T, the partner of the SM Higgs doublet. As we remarked above, since T has generically small couplings to the first generation of fermions, its contribution to proton decay becomes sub-leading as soon as its mass exceeds $10^{12}$ GeV or so, and if it were on the order of the GUT scale, it would be completely negligible. In other words, the contribution from T can easily be suppressed, and so we concentrate on the one from the new

| Channel | Lifetime ($10^{30}$yrs) |
|---|---|
| $N \to e^+ \pi$ | 5300 (n), 16000 (p) |
| $N \to \mu^+ \pi$ | 3500 (n), 7700 (p) |
| $N \to \nu \pi$ | 1100 (n), 390 (p) |
| $N \to e^+ K$ | 17 (n), 1000 (p) |
| $N \to \mu^+ K$ | 26 (n), 1600 (p) |
| $N \to \nu K$ | 86 (n), 5900 (p) |

TABLE I. Bounds on proton and neutron decay [34].

gauge bosons. We should stress that some of our results were announced in [27].

The effective 6−dimensional baryon number violating operators in $SU(5)$, due to X and Y gauge boson exchange, are schematically given by (all the fermions and anti-fermions are left-handed in our notation)

$$\mathcal{L}_{\not B} = (\overline{u}^c u)\left(\overline{e}^c d + \overline{d}^c e\right) + (\overline{u}^c d)\left(\overline{e}^c u + \overline{d}^c \nu\right) . \quad (12)$$

We don't care here about the relative strength of order one between these terms. They induce a proton lifetime given by

$$\tau_{\rm p} \simeq C\,\frac{M_{\rm GUT}^4}{\alpha_{\rm GUT}^2} m_{\rm p}^{-5}, \quad (13)$$

where $m_{\rm p}$ is the proton mass, $M_{\rm GUT}$ is the unification scale, $\alpha_{\rm GUT}$ is the value of gauge couplings at that scale and $C$ denotes the physical effects of running from the GUT scale to the QCD scale and the non-perturbative effects of the confinement of quarks in mesons. When the dust settles (see Appendix), one typically claims $C \simeq 1$. The most stringent bound on proton lifetime from Super-Kamiokande is about $\tau_{\rm p} \gtrsim 10^{34}$yrs for the channel ${\rm p} \to \pi^0\, e^+$ [3]. For $\alpha_{\rm GUT} \simeq 40^{-1}$ (obtained from the running), and if $C$ were really of order one, one would get $M_{\rm GUT} \gtrsim 4 \cdot 10^{15}$GeV which would be too high to be compatible with the values obtained in the previous section.

This, however, is not enough to rule out minimal $SU(5)$, not before studying carefully the value of $C$, which turns out to be intimately connected with the flavour structure of the above baryon decay operators. What happens is that $C$ could in principle be quite suppressed by small flavor mixings. This was studied carefully in Ref. [19] and we briefly recapitulate here their arguments and update their results.

The crucial point, as we discussed in the Sec. II, is that due to the presence of higher-dimensional operators, Yukawa couplings become arbitrary, which implies arbitrary fermion mass matrices and, in turn, arbitrary unitary transformations from the weak to the physical mass eigenstates base.

Rotating the fermions into the mass basis, (12) becomes

$$\begin{aligned}\mathcal{L}_{\not B} = &\left(\overline{u}^c\, U_c^\dagger U\, u\right)\left(\overline{e}^c\, E_c^\dagger D\, d + \overline{d}^c\, D_c^\dagger E\, e\right)\\ &+ \left(\overline{u}^c\, U_c^\dagger D\, d\right)\left(\overline{e}^c\, E_c^\dagger U\, u + \overline{d}^c\, D_c^\dagger N\, \nu\right) .\end{aligned} \quad (14)$$

The capital symbols denote the unitary transformations

$$f \to F\, f\,; \quad f^c \to F_c\, f^c\,; \quad F F^\dagger = F_c F_c{}^\dagger = 1\,, \quad (15)$$

needed to diagonalize fermion mass matrices in the obvious notation ($N$ being rotation on neutrinos). Recall that the CKM matrix is given by $V_{\rm CKM} = U^\dagger D$.

### A. Analytical study: preliminaries

One could even imagine rotating proton decay completely away, however, unitarity forbids it. In fact, for this to happen one would need [14]

$$\left(U_c^\dagger U\right)_{11} = 0\,, \quad \left(E_c^\dagger U\right)_{a1} = 0\,, \quad \left(U_c^\dagger D\right)_{1a} = 0\,, \quad (16)$$

for $a = 1, 2$. It is readily seen that the third equations in (16) cannot be simultaneously satisfied for both $a = 1, 2$ as long as $(V_{\rm CKM})_{13} \neq 0$. Still, one can exploit the available freedom in these mixings in order to maximize the proton decay lifetime.

In order to set the stage, we report in Table I the experimental bounds on proton and neutron decays [34] relevant for the analysis. We are best off with killing the pion and positron or anti-muon modes, in order to try to keep the GUT scale as low as possible and salvage the theory. With that in mind, a natural thing to try would be $\left(U_c^\dagger U\right)_{11} = 0$, $\left(E_c^\dagger U\right)_{a1} = 0$, $\left(U_c^\dagger D\right)_{11} = 0$ for $a = 1, 2$. In this case, proton (neutron) could decay into pions and kaons, and antineutrino and the decay rates would be suppressed by small $(V_{\rm CKM})_{13}$. In this case, $M_{\rm GUT} \geq \sqrt{\alpha_{GUT}/40^{-1}}\, 3 \cdot 10^{14}$GeV, too large to be compatible with gauge coupling unification, c.f., (11).

A better choice is

$$\begin{aligned}\left(U_c^\dagger D\right)_{1a} = \left(E_c^\dagger D\right)_{1a} = \left(E_c^\dagger D\right)_{a1} = 0\,,\\ \left(D_c^\dagger E\right)_{1a} = \left(D_c^\dagger E\right)_{a1} = 0\,,\end{aligned} \quad (17)$$

for $a = 1, 2$. This leads to the following lower bound on the GUT scale[1]

$$M_{\rm GUT} \geq \sqrt{\frac{\alpha_{\rm GUT}}{40^{-1}}}\, 1.3 \cdot 10^{14}{\rm GeV}\,. \quad (18)$$

Here the relevant channel is $p \to K^0 + \mu^+$, again with $(V_{\rm CKM})_{13}$ suppression. Table I shows that this mode is not so well measured, which helps the theory pass the unification test (11).

Fig. 2 shows that (18) implies $m_3, m_8 \lesssim 10^8\,{\rm GeV}$. Moreover, an improvement in $K^0$ and $\mu^+$ lifetime by a factor of about 4-5 would require $m_3$ and $m_8$ below 10 TeV - in the right ballpark for hadron colliders.

[1] Chiral Lagrangian parameters were taken from [35].

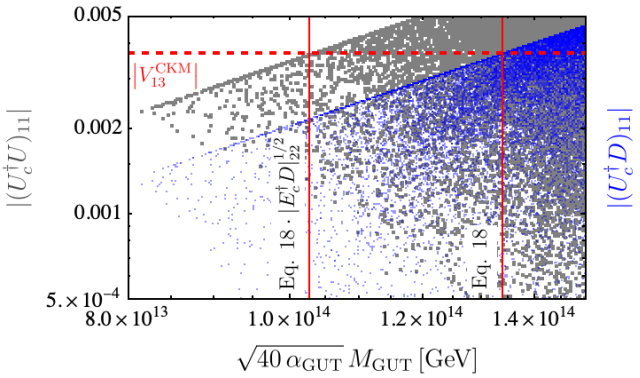



FIG. 3. Randomized values of $|(U_c^\dagger U)_{11}|$ (gray) and corresponding $|(U_c^\dagger D)_{11}|$ (blue) vs $M_{\rm GUT}$ compatible with experimental decay rate bounds. Dashed red line corresponds to $|(V_{\rm CKM})_{13}|$. The red left line corresponds to $|E_c^\dagger D|_{22} = 1$ leading to Eq. (18), while the red right line corresponds to its relaxation discussed in the appendix, c.f. (41).

It is noteworthy that (18) differs from analogous result of [19] by roughly a factor of 2. Over the years, both the nucleon lifetime limits and the chiral Lagrangian parameters changed, which could explain the discrepancy. Furthermore, [19] seems to have omitted the running of the proton decay amplitude from the weak to the unification scale.

### B. Numerical study: maximizing proton decay rotation

Naively, it would seem that the theory favours particular channels mentioned above, and moreover, that it could soon be ruled out. The trouble is that the zeroes in mixing matrices need not be exact, a small leakage would allow for other decay rates, such as the pion and charged anti-leptons one. In order to ensure a complete exhaustion of the parameter space, we performed a Monte Carlo exploration - whose details can be found in Appendix - in which the rotation conditions were allowed to be more general.

The resulting analysis is summarized in Fig. 3, in which the absolute values of $(U_c^\dagger U)_{11}$ (gray) and the corresponding $(U_c^\dagger D)_{11}$ (blue) are shown as a function of $M_{\rm GUT}$. A similar plot holds for the $(U_c^\dagger D)_{12}$ matrix element, which we omit here. The result of Eq. (18), is automatically found by the numerical analysis, which is shown with the vertical red line in Fig. 3. All points to its left - with matrix elements whose values are smaller than $(V_{\rm CKM})_{13}$ - correspond to effective realizations compatible with experimental limits on nucleon lifetimes with a lower $M_{\rm GUT}$.

It is not surprising that such points correspond to matrix elements whose absolute values are below $|(V_{\rm CKM})_{13}|$, as in the case of $|(U_c^\dagger D)_{11}|$. The lower bound on $M_{\rm GUT}$ compatible with current proton decay lifetime is then found to be

$$M_{\rm GUT} \gtrsim \sqrt{\frac{\alpha_{\rm GUT}}{40^{-1}}}\, 8\cdot 10^{13}\,{\rm GeV}. \tag{19}$$

Due to QCD low-energy inputs, affected by uncertainties (see, e.g., [35]), our results are numerically reliable up to corrections of about 10%. Eq. (19) shows that the rotation of all nucleon decay amplitudes into $(V_{\rm CKM})_{13}$, discussed above, does not exhaust all of the parameter space.

If the reader feels that there is something unnatural in having a theory live on what looks like a convoluted point, we remind them of the way the GIM suppression of $K_L - K_S$ works as a conspiracy of a small charm quark mass, and then a huge top mass contribution suppressed by a tiny $(V_{\rm CKM})_{13}$. Good theories live dangerously.

In our analysis we exploited the freedom in different proton decay channels, in order to obtain the lowest possible unification scale compatible with proton lifetime. Therefore, an optimistic scenario of an improvement in proton lifetime by a factor of $10-25$ in future experiments for all channels, would still not rule out the theory. It would, however, indicate the presence of either $m_3$ or $m_8$ below 100 TeV. An ulterior improvement roughly up to a factor $25-50$ in proton lifetime would require these particles to be lighter than 10 TeV. The principal uncertainty in these numbers stems from the QCD computation of the nucleon decay matrix elements [35]. In fact, a 10% uncertainty on the $M_{\rm GUT}$ bound, implies, roughly, a 50% uncertainty in the necessary improvement on proton lifetime.

The necessary flavor rotation in heavy gauge boson interactions implies, however, the impossibility of doing the same for the Yukawa interactions of the T scalar, forcing it to lie roughly close to the GUT scale.

**Consequences for neutrino mass.** As we discusses in the Sec. II, when defining the minimal realistic $SU(5)$, one resorts to higher dimensional operators, implying that neutrino mass ought to be given by (7). The crucial question here is how perturbative one has to be. If one stays at its boundary, one gets $\Lambda \sim 10 M_{\rm GUT} \sim 10^{15}$ GeV and all is well. If one wanted, though, to keep the couplings at worst of order one, this would require $\Lambda \lesssim 10^{14}$ GeV, which would seemingly invalidate the theory. In such a case the reader may wonder why we we would go through all of this tedium, if the theory was already ruled out by neutrino mass consideration. We wanted to show, however, that the often repeated claim of the minimal $SU(5)$ not being able to provide the unification of gauge couplings with the sufficient proton longevity is simply wrong. There is more to it, though.

The point is that most completions of the $d=5$ operator in (6) tend to be based on new fermions or scalars, with potentially small Yukawa couplings. The most popular example of this realization is the seesaw mechanism,

which in its original form consists of the addition of RH neutrino singlets. Even if they were much lighter than the GUT scale, and even if they were at the weak scale itself, there would be almost no change of the original theory due to the associated smallness of their Yukawa couplings - except for the emerging neutrino mass. In other words, the physical cutoff scale in (7) could be arbitrarily small without affecting the unification scale and proton decay.

## V. SU(5) EXTENSIONS AND NEUTRINO MASS

In principle, the minimal theory could be defined even with the addition of RH neutrinos, due do their inoffensive nature. There is something rather unappealing, though, in adding gauge singlets in a theory whose predictions depend profoundly on its gauge nature. There are two particularly attractive alternatives, based on the so-called type II and type III seesaw mechanisms (for a review, see e.g. [36]), with minimal changes of the theory and possibly predictive outcomes, both obtained by an addition of a single field augmented with higher-dimensional operators as in the original minimal theory. We recapitulate here their salient features in view of our findings.

*a. Type II seesaw* Instead of fermion singlets, the RH neutrinos, one may opt for the so-called type II seesaw [37–39] through the addition of a symmetric scalar $15_{\rm H}$ field [10, 40]. This induces the following new relevant couplings

$$\mathcal{L}_{type\,II} = \overline{5}_{\rm F} Y_\nu 15_{\rm H} 10_{\rm F} + \mu\, 5_{\rm H}\, 15^*_{\rm H}\, 5_{\rm H} + \dots\,, \qquad (20)$$

which gives neutrino a Majorana mass through the small vev of $15_{\rm H}$. While it may be considered somewhat ad-hoc, an outcome of model building, it has an appeal of being potentially verifiable as the origin of neutrino mass. Namely, the weak triplet in $15_{\rm H}$ contains doubly charged scalars, whose branching ratio into charged leptons are dictated by the neutrino mass matrix [41, 42] - of course, assuming these scalars can be produced at hadron colliders. The minimal $SU(5)$ model is a sub-case of this one, and therefore there exists a region of the parameter space where all of the sub-multiplet of $15_{\rm H}$ has a mass around the unification scale.

The problem is that the states in $15_{\rm H}$ do not suffice to sufficiently increase the GUT scale [43], therefore still requiring cancellations in proton decay amplitudes. In other words, from the point of view of constructing a minimal realistic theory, one might as well stick to the original theory, as we amply discussed above.

*b. Type III seesaw* Another simple alternative is based on a fermionic $24_{\rm F}$ field [44, 45] instead of $15_{\rm H}$. A naive analysis which ignores (8) and the flavour of proton decay would imply the presence of the weak triplet fermions and scalars, within the reach of the LHC [44] - an exciting prediction, if true (the detailed study of the resulting phenomenology can be found in [46]). The fermion triplet variation of the seesaw mechanism is coined type III [47], with today's experimental limit on the triplet mass being roughly TeV [48].

However, the additional freedom of Eq. (8) leads to a straighforward generalization of (10)

$$\frac{M_{\rm GUT}}{M_{\rm Z}} = e^{\frac{\pi}{21}\left[5\left(1+\frac{1}{4}\epsilon\right)\alpha_1^{-1} - 3\left(1+\frac{5}{4}\epsilon\right)\alpha_2^{-1} - 2\left(1-\frac{5}{4}\epsilon\right)\alpha_3^{-1}\right]} \left(\frac{m_{\rm LQ}^4\, M_{\rm Z}}{m_3^{1/2}\, m_8^{1/2} m_{\rm 3F}^2 m_{\rm 8F}^2}\right)^{\frac{1}{21}\left(1-\frac{1}{4}\epsilon\right)}, \qquad (21)$$

where the labels 3F, 8F and LQ denote respectively the weak triplet, the colour octet and the leptoquark component of $24_{\rm F}$. In the previous section, a low $m_3$ and $m_8$ could increase unification up to $10^{14}\,$GeV. It is not surprising then that lowering also the masses of their fermionic counterpart can give $M_{\rm GUT} \gtrsim 10^{15}\,$GeV.

While for $\epsilon = 0$, both $m_{3\rm F}$ and $m_3$ lie close to TeV energies [44], it is a straightforward numerical exercise to show that for $\epsilon \simeq 0.05$, all the particle masses are above $10^5$GeV, with a unification scale of $4 \cdot 10^{15}\,$GeV, compatible with the naive proton lifetime estimate.

A comment is noteworthy here. Ref. [44] made two assumptions: (i) no flavour cancellations in proton decay amplitudes and (ii) that the cutoff is as large as possible, in order to ensure precise perturbative predictions. The latter implies $\Lambda \simeq 100\, M_{\rm GUT}$ or $\epsilon \simeq 0.01$ (a larger cutoff would not suffice to correct the bottom-tau mass relation), which then justifies the predictions quoted.

A similar situation emerges also in minimal $SO(10)$ theory with higher dimensional operators [49]. In fact, also there the presence of new light particle states is intimately tied to the assumption of no cancellations in proton decay amplitudes. It is precisely in this sense that the minimal $SU(5)$ theory stands - it is predictive without any further assumption.

## VI. SUMMARY AND OUTLOOK

In the last twenty years or so, people - including one of the authors - have been regularly writing necrologies for the the minimal $SU(5)$ theory, even in its non-renormalisable version. And so, people - now including both authors - have been arguing in favor of going beyond the minimal theory. As we showed here, in spite of these morbid desires, the minimal $SU(5)$ theory is alive and kicking - as long as one is willing to accept higher-dimensional operators as the simultaneous cure of the wrong fermion mass relations and the failure of the gauge coupling unification. However, it does live dangerously and an improvement in proton lifetime bounds could require new light scalar states, the weak triplet and the

colour octet. This strengthens a case for proton decay searches and future hadron colliders.

Moreover, a light scalar weak triplet would be expected to change the Standard Model W-mass relation, paving a window for an indirect observation of new physics. Remarkably, its low energy effective theory is unambiguously determined by its mass and the W-mass deviation [22].

Admittedly, higher-dimensional operators may not be that appealing, but the vicinity of the non-perturbative gravitational effects to the GUT scale makes them potentially sizeable and thus hard to ignore. In this minimal scenario they are actually a must, being the only possible source of neutrino masses through the $d = 5$ Weinberg operator. Moreover, the actual scale of this operator could be independent of the cutoff and naturally much lower, due to small Yukawa couplings of particles that provide UV completions of the theory.

Otherwise, in order to allow for a higher cutoff scale, one would have to add fermionic singlets (the RH neutrinos), an admittedly ugly solution. For this reason we discussed more natural extensions that utilize the inclusion of $15_{\rm H}$ or $24_{\rm F}$ states [43, 44] This allows for possible direct collider probes of the neutrino mass origin - an exciting possibility - but there is no guarantee to have the new particles accessible at the LHC or a new hadron collider. The case of the adjoint fermion $24_{\rm F}$ is free from a necessity of suppressing the proton decay by tuning the flavour structure and at the first glance seemed to predict a weak fermion triplet at TeV energies. The higher-dimensional operators, though, in the gauge field sector could push it above 100 TeV, and render it inaccessible even at the next hadron collider. While this model remains a natural extension of the minimal theory, our results show that it is less predictive than imagined originally.

We wish to end on a more sober note regarding the whole idea of grand unification. The main message of our work, besides the fact that the minimal $SU(5)$ theory with higher dimensional operators is still very much viable, is that grand unified theories have a generic problem of a lack of clear predictions when it comes to nucleon decay. But nucleon decay is the essence of grand unification, just as the neutral currents were the essence of the SM. The main problem is the fact that the GUT scale is astronomically large compared to the weak one, and one is often forced to make assumptions - non testable in any foreseeable future - at $M_{\rm GUT}$ in order to claim verifiable consequences at accessible energies. However, it is precisely the enormity of the GUT scale that allows for a successful effective $d = 6$ theory of baryon violation [12, 50], with a number of clear predictions (for an overview, see [20]) that hold true if $M_{\rm GUT} \gg M_{\rm Z}$.

Establishing the nature of the effective theory of baryon number violation through experiment would be equivalent to the historic determination of the V-A character of the effective Fermi theory of weak interaction - which then allowed to create the fundamental renormalisable theory, the Standard Model. As Weinberg put it nicely [51]: "V-A was the key". One perhaps ought to wait for an analogous thing to happen in the nucleon decay interaction in order to find a fundamental grand unified theory?

***Acknowledgements.*** We are grateful to Anca Preda for participating at the initial stages of this projects. We are indebted to Borut Bajc for important comments and a careful reading of the manuscript, and to Kaladi Babu for a correspondence regarding proton decay rates. We also wish to thank Alejandra Melfo for her help in improving the presentation of our work.

## APPENDIX: DECAY RATE ANALYSIS

This appendix relaxes conditions (18). Namely, we look for a lower bound on $M_{\rm GUT}$ by allowing for arbitrary matrix elements in the $d = 6$ effective operators (14). To do so, a numerical sampling of the parameter space has been performed via a Monte Carlo method that will be explained below. For simplicity and completeness, the notation of Ref. [18] will be adopted and adapted to the following.

### 1. Matrix Elements

The $d = 6$ operators leading to proton decay in $SU(5)$ grand unification can be derived from integrating out the heavy gauge bosons $X, Y$ with electric charge $4/3$ and $1/3$ respectively. The resulting Standard Model operators are $B - L$ conserving and are given by [12, 20, 50, 52]

$$O_I^{B-L} = \frac{g^2}{2M_X^2}\,\epsilon_{ijk}\,\epsilon_{\alpha\beta}\,\overline{u_{ia}^C}\,\gamma^\mu\,Q_{j\alpha a}\,\overline{e_b^C}\,\gamma_\mu\,Q_{k\beta b} \quad (22)$$

$$O_{II}^{B-L} = \frac{g^2}{2M_X^2}\,\epsilon_{ijk}\,\epsilon_{\alpha\beta}\,\overline{u_{ia}^C}\,\gamma^\mu\,Q_{j\alpha a}\,\overline{d_{kb}^C}\,\gamma_\mu\,L_{\beta b} \quad (23)$$

where $g$ denotes the gauge coupling at $M_X \simeq M_{\rm GUT}$ and $Q = (u, d)$, $L = (\nu, e)$; $i$, $j$ and $k$ are the color indices, $a$,$b$ are the family indices, and $\alpha, \beta = 1, 2$.
When going to the physical basis, the effective operators (22,23) become

$$O\ (e_\alpha^C, d_\beta) = \frac{g^2}{2M_X^2} c(e_\alpha^C, d_\beta)\epsilon_{ijk}\overline{u_i^C}\gamma^\mu u_j \overline{e_\alpha^C}\gamma_\mu d_{k\beta} \quad (24)$$

$$O\ (e_\alpha, d_\beta^C) = \frac{g^2}{2M_X^2} c(e_\alpha, d_\beta^C)\epsilon_{ijk}\overline{u_i^C}\gamma^\mu u_j \overline{d_{k\beta}^C}\gamma_\mu e_\alpha \quad (25)$$

$$O\ (\nu_l, d_\alpha, d_\beta^C) = \frac{g^2}{2M_X^2} c(\nu_l, d_\alpha, d_\beta^C)\epsilon_{ijk}\overline{u_i^C}\gamma^\mu d_{j\alpha} \overline{d_{k\beta}^C}\gamma_\mu \nu_l \quad (26)$$

where

$$c\ (e^C_\alpha, d_\beta) = \left(U_c^\dagger U\right)_{11} \left(E_c^\dagger D\right)_{\alpha\beta} + (U_c^\dagger D)_{1\beta}(E_c^\dagger U)_{\alpha 1} \quad (27)$$

$$c\ (e_\alpha, d^C_\beta) = \left(U_c^\dagger U\right)_{11} \left(D_c^\dagger E\right)_{\beta\alpha} \quad (28)$$

$$c\ (\nu_l, d_\alpha, d^C_\beta) = \left(U_c^\dagger D\right)_{1\alpha} (D_c^\dagger N)_{\beta l} \quad (29)$$

This is the precise form of (14).

### 2. Two body decay channels of the nucleon

In proton decay experiments, no distinction is made between the flavour of neutrino, nor the chirality of the charged lepton in the decay products. In this case, the relevant decay channel can be derived with the appropriate chiral techniques [18, 53, 54] and are given by

$$\Gamma(p \to K^+ \bar{\nu}) = \frac{g^4}{4M_{\rm X}^4} \frac{(m_p^2 - m_K^2)^2}{8\pi m_p^3 f_\pi^2} A_L^2 \left|\alpha\right|^2 \sum_{i=1}^{3} \left| \frac{2m_p}{3m_B} D\ c(\nu_i, d, s^C) + [1 + \frac{m_p}{3m_B}(D+3F)] c(\nu_i, s, d^C) \right|^2 \quad (30)$$

$$\Gamma(p \to \pi^+ \bar{\nu}) = \frac{g^4}{4M_{\rm X}^4} \frac{m_p}{8\pi f_\pi^2} A_L^2 \left|\alpha\right|^2 (1+D+F)^2 \sum_{i=1}^{3} \left| c(\nu_i, d, d^C) \right|^2 \quad (31)$$

$$\Gamma(p \to \eta e^+_\beta) = \frac{g^4}{4M_{\rm X}^4} \frac{(m_p^2 - m_\eta^2)^2}{48\pi f_\pi^2 m_p^3} A_L^2 \left|\alpha\right|^2 (1+D-3F)^2 \{ \left| c(e_\beta, d^C) \right|^2 + \left| c(e^C_\beta, d) \right|^2 \} \quad (32)$$

$$\Gamma(p \to K^0 e^+_\beta) = \frac{g^4}{4M_{\rm X}^4} \frac{(m_p^2 - m_K^2)^2}{8\pi f_\pi^2 m_p^3} A_L^2 \left|\alpha\right|^2 [1 + \frac{m_p}{m_B}(D-F)]^2 \{ \left| c(e_\beta, s^C) \right|^2 + \left| c(e^C_\beta, s) \right|^2 \} \quad (33)$$

$$\Gamma(p \to \pi^0 e^+_\beta) = \frac{g^4}{4M_{\rm X}^4} \frac{m_p}{16\pi f_\pi^2} A_L^2 \left|\alpha\right|^2 (1+D+F)^2 \{ \left| c(e_\beta, d^C) \right|^2 + \left| c(e^C_\beta, d) \right|^2 \} \quad (34)$$

$$\Gamma(n \to K^0 \bar{\nu}) = \frac{g^4}{4M_{\rm X}^4} \frac{(m_n^2 - m_K^2)^2}{8\pi m_n^3 f_\pi^2} A_L^2 \left|\alpha\right|^2 \cdot \cdot \sum_{i=1}^{3} \left| c(\nu_i, d, s^C)[1 + \frac{m_n}{3m_B}(D-3F)] - c(\nu_i, s, d^C)[1 + \frac{m_n}{3m_B}(D+3F)] \right|^2 \quad (35)$$

$$\Gamma(n \to \pi^0 \bar{\nu}) = \frac{g^4}{4M_{\rm X}^4} \frac{m_n}{16\pi f_\pi^2} A_L^2 \left|\alpha\right|^2 (1+D+F)^2 \sum_{i=1}^{3} \left| c(\nu_i, d, d^C) \right|^2 \quad (36)$$

$$\Gamma(n \to \eta \bar{\nu}) = \frac{g^4}{4M_{\rm X}^4} \frac{(m_n^2 - m_\eta^2)^2}{48\pi m_n^3 f_\pi^2} A_L^2 \left|\alpha\right|^2 (1+D-3F)^2 \sum_{i=1}^{3} \left| c(\nu_i, d, d^C) \right|^2 \quad (37)$$

$$\Gamma(n \to \pi^- e^+_\beta) = \frac{g^4}{4M_{\rm X}^4} \frac{m_n}{8\pi f_\pi^2} A_L^2 \left|\alpha\right|^2 (1+D+F)^2 \{ \left| c(e_\beta, d^C) \right|^2 + \left| c(e^C_\beta, d) \right|^2 \} \quad (38)$$

where the average baryon mass $m_B \approx m_\Sigma \approx m_\Lambda \simeq 1.15\,\text{GeV}$, $D \simeq 0.8$, $F \simeq 0.47$ and $\alpha \simeq 0.012\,\text{GeV}^3$ are the parameters of the chiral lagrangian taken from [35]. Here all coefficients of four-fermion operators are evaluated at $M_Z$ scale. $A_L$ takes into account renormalization effects from $M_Z$ to GeV - corresponding to a factor of about 1.4 [18]. A further correction arises due to the gauge coupling running from $M_{\rm GUT}$ to $M_Z$ [55] which gives an additional factor 3.4. Finally, $\nu_i = \nu_e, \nu_\mu, \nu_\tau$ and $e_\beta = e, \mu$.

Given the above values, direct inspection of (34) fixes the prefactor coefficient $C \simeq 0.73$ in (13), as mentioned in the main text.

### 3. On the numerical analysis

The most stringent bound found by Ref. [19] is obtained with

$$\begin{aligned} &\left(U_c^\dagger D\right)_{1a} = 0 \\ &\left(E_c^\dagger D\right)_{1a} = \left(E_c^\dagger D\right)_{a1} = 0, \\ &\left(D_c^\dagger E\right)_{1a} = \left(D_c^\dagger E\right)_{a1} = 0, \\ &\left(E_c^\dagger U\right)_{a1} = 0 \end{aligned} \quad (39)$$

for $a = 1, 2$. The first condition of Eq. (39) immediately leads to

$$|\left(U_c^\dagger U\right)_{11}| = |\left(V_{\rm CKM}\right)_{13}|. \quad (40)$$

Second and third condition in Eq. (39) cannot be accommodated for all four matrix elements. In particular,

bound (18) is obtained by setting all matrix elements to vanish, except for the $|\left(E_c^\dagger D\right)_{22}| = 1$. This effectively suppresses the decay into $K\mu$ by $(V_{\rm CKM})_{13}$.

Indeed, a less stringent condition can be imposed. Namely, all the elements $(E_c^\dagger D)_{ab}$, $a, b = 1, 2$ can be chosen to minimize equally proton decay into the different charged lepton channels. An inspection of (27,28), suggests it would be natural to take $|(E_c^\dagger D)| = |(D_c^\dagger E)^T|$. Given current values of proton lifetime [34], the $2 \times 2$ block matrix in the first two families minimizing decay rates in Table I - while preserving unitarity of the total $3 \times 3$ matrix - is given by

$$|(E_c^\dagger D)| \simeq |(D_c^\dagger E)^T| \simeq \begin{pmatrix} 0.11 & 0.77 \\ 0.16 & 0.61 \end{pmatrix}. \qquad (41)$$

In order to relax the first condition in (39), leading to (40), we numerically randomized the choice of $(U_c^\dagger U)_{11}$ and used $U^\dagger D = K_1 V_{\rm CKM} K_2$ (where $K_{1,2}$ are two diagonal matrices with respectively three and two phases) to compute the matrix $U_c^\dagger D$. Each sampled solution was then used to derive the corresponding $M_{\rm GUT}$.


[1] Alexander M. Polyakov, "Particle Spectrum in Quantum Field Theory," JETP Lett. **20**, 194–195 (1974).

[2] Gerard 't Hooft, "Magnetic Monopoles in Unified Gauge Theories," Nucl. Phys. B **79**, 276–284 (1974).

[3] A. Takenaka *et al.* (Super-Kamiokande), "Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ with an enlarged fiducial volume in Super-Kamiokande I-IV," Phys. Rev. D **102**, 112011 (2020), arXiv:2010.16098 [hep-ex].

[4] H. Georgi and S. L. Glashow, "Unity of All Elementary Particle Forces," Phys. Rev. Lett. **32**, 438–441 (1974).

[5] H. Georgi, Helen R. Quinn, and Steven Weinberg, "Hierarchy of Interactions in Unified Gauge Theories," Phys. Rev. Lett. **33**, 451–454 (1974).

[6] P. Minkowski, "$\mu \to e\gamma$ at a Rate of One Out of $10^9$ Muon Decays?" Phys. Lett. B **67**, 421–428 (1977).

[7] R.N. Mohapatra and G. Senjanović, "Neutrino Mass and Spontaneous Parity Nonconservation," Phys. Rev. Lett. **44**, 912 (1980).

[8] T. Yanagida, "Horizontal gauge symmetry and masses of neutrinos," Conf. Proc. C **7902131**, 95–99 (1979).

[9] M. Gell-Mann, P. Ramond, and R. Slansky, "Complex Spinors and Unified Theories," Conf. Proc. C **790927**, 315–321 (1979), arXiv:1306.4669 [hep-th].

[10] S.L. Glashow, "The Future of Elementary Particle Physics," NATO Sci. Ser. B **61**, 687 (1980).

[11] John R. Ellis and Mary K. Gaillard, "Fermion Masses and Higgs Representations in SU(5)," Phys. Lett. B **88**, 315–319 (1979).

[12] Steven Weinberg, "Baryon and Lepton Nonconserving Processes," Phys. Rev. Lett. **43**, 1566–1570 (1979).

[13] Q. Shafi and C. Wetterich, "Modification of GUT Predictions in the Presence of Spontaneous Compactification," Phys. Rev. Lett. **52**, 875 (1984).

[14] S. Nandi, A. Stern, and E. C. G. Sudarshan, "CAN PROTON DECAY BE ROTATED AWAY?" Phys. Lett. B **113**, 165–169 (1982).

[15] V. S. Berezinsky and A. Yu. Smirnov, "HOW TO SAVE MINIMAL SU(5)," Phys. Lett. B **140**, 49–52 (1984).

[16] Borut Bajc, Pavel Fileviez Perez, and Goran Senjanović, "Proton decay in minimal supersymmetric SU(5)," Phys. Rev. D **66**, 075005 (2002), arXiv:hep-ph/0204311.

[17] Borut Bajc, Pavel Fileviez Perez, and G. Senjanović, (2002), unpublished.

[18] Pavel Fileviez Perez, "Fermion mixings versus d = 6 proton decay," Phys. Lett. B **595**, 476–483 (2004), arXiv:hep-ph/0403286.

[19] Ilja Doršner and Pavel Fileviez Perez, "How long could we live?" Phys. Lett. B **625**, 88–95 (2005), arXiv:hep-ph/0410198.

[20] Goran Senjanović, "Proton decay and grand unification," AIP Conf. Proc. **1200**, 131–141 (2010), arXiv:0912.5375 [hep-ph].

[21] Tommy Ohlsson, "Proton decay," Nucl. Phys. B **993**, 116268 (2023), arXiv:2306.02401 [hep-ph].

[22] Goran Senjanović and Michael Zantedeschi, "$SU(5)$ grand unification and $W$-boson mass," (2022), arXiv:2205.05022 [hep-ph].

[23] K. S. Babu, Borut Bajc, and Shaikh Saad, "Yukawa Sector of Minimal SO(10) Unification," JHEP **02**, 136 (2017), arXiv:1612.04329 [hep-ph].

[24] K. S. Babu and Chung Ngoc Leung, "Classification of effective neutrino mass operators," Nucl. Phys. B **619**, 667–689 (2001), arXiv:hep-ph/0106054.

[25] G. R. Dvali, "Can 'doublet - triplet splitting' problem be solved without doublet - triplet splitting?" Phys. Lett. B **287**, 101–108 (1992).

[26] A. J. Buras, John R. Ellis, M. K. Gaillard, and Dimitri V. Nanopoulos, "Aspects of the Grand Unification of Strong, Weak and Electromagnetic Interactions," Nucl. Phys. B **135**, 66–92 (1978).

[27] Goran Senjanović and Michael Zantedeschi, "Grand unification, small vs large representations, hadron colliders and all that," in *CORFU2022: 22th Hellenic School and Workshops on Elementary Particle Physics and Gravity* (2023) arXiv:2304.07932 [hep-ph].

[28] Joydeep Chakrabortty and Amitava Raychaudhuri, "A Note on dimension-5 operators in GUTs and their impact," Phys. Lett. B **673**, 57–62 (2009), arXiv:0812.2783 [hep-ph].

[29] Rabindra N. Mohapatra and Goran Senjanovic, "Higgs Boson Effects in Grand Unified Theories," Phys. Rev. D **27**, 1601 (1983).

[30] Gia Dvali, "Black Holes and Large N Species Solution to the Hierarchy Problem," Fortsch. Phys. **58**, 528–536 (2010), arXiv:0706.2050 [hep-th].

[31] Gia Dvali and Michele Redi, "Black Hole Bound on the Number of Species and Quantum Gravity at LHC," Phys. Rev. D **77**, 045027 (2008), arXiv:0710.4344 [hep-th].

[32] Lawrence J. Hall, "Grand Unification of Effective Gauge Theories," Nucl. Phys. B **178**, 75–124 (1981).

[33] Steven Weinberg, "Effective Gauge Theories," Phys. Lett. B **91**, 51–55 (1980).

[34] P.A. Zyla *et al.* (Particle Data Group), "Review of Particle Physics," PTEP **2020**, 083C01 (2020).

[35] Jun-Sik Yoo, Yasumichi Aoki, Peter Boyle, Taku Izubuchi, Amarjit Soni, and Sergey Syritsyn, "Proton decay matrix elements on the lattice at physical pion mass," (2021), arXiv:2111.01608 [hep-lat].

[36] G. Senjanovic, "Neutrino mass: From LHC to grand unification," Riv. Nuovo Cim. **34**, 1–68 (2011).

[37] M. Magg and C. Wetterich, "Neutrino Mass Problem and Gauge Hierarchy," Phys. Lett. B **94**, 61–64 (1980).

[38] Rabindra N. Mohapatra and Goran Senjanović, "Neutrino Masses and Mixings in Gauge Models with Spontaneous Parity Violation," Phys. Rev. D **23**, 165 (1981).

[39] George Lazarides, Q. Shafi, and C. Wetterich, "Proton Lifetime and Fermion Masses in an SO(10) Model," Nucl. Phys. B **181**, 287–300 (1981).

[40] Goran Senjanović, "Neutrino mass and unification," (1981).

[41] Julia Garayoa and Thomas Schwetz, "Neutrino mass hierarchy and Majorana CP phases within the Higgs triplet model at the LHC," JHEP **03**, 009 (2008), arXiv:0712.1453 [hep-ph].

[42] M. Kadastik, M. Raidal, and L. Rebane, "Direct determination of neutrino mass parameters at future colliders," Phys. Rev. D **77**, 115023 (2008), arXiv:0712.3912 [hep-ph].

[43] Ilja Dorsner and Pavel Fileviez Perez, "Unification without supersymmetry: Neutrino mass, proton decay and light leptoquarks," Nucl. Phys. B **723**, 53–76 (2005), arXiv:hep-ph/0504276.

[44] Borut Bajc and Goran Senjanović, "Seesaw at LHC," JHEP **08**, 014 (2007), arXiv:hep-ph/0612029.

[45] Borut Bajc, Miha Nemevsek, and Goran Senjanovic, "Probing seesaw at LHC," Phys. Rev. D **76**, 055011 (2007), arXiv:hep-ph/0703080.

[46] Abdesslam Arhrib, Borut Bajc, Dilip Kumar Ghosh, Tao Han, Gui-Yu Huang, Ivica Puljak, and Goran Senjanovic, "Collider Signatures for Heavy Lepton Triplet in Type I+III Seesaw," Phys. Rev. D **82**, 053004 (2010), arXiv:0904.2390 [hep-ph].

[47] Robert Foot, H. Lew, X. G. He, and Girish C. Joshi, "Seesaw Neutrino Masses Induced by a Triplet of Leptons," Z. Phys. C **44**, 441 (1989).

[48] Tadej Novak, *Search for type-III seesaw heavy leptons with the ATLAS detector at the LHC*, Ph.D. thesis, Ljubljana U. (2020).

[49] Anca Preda, Goran Senjanović, and Michael Zantedeschi, "SO(10): a Case for Hadron Colliders," (2022), arXiv:2201.02785 [hep-ph].

[50] Frank Wilczek and A. Zee, "Operator Analysis of Nucleon Decay," Phys. Rev. Lett. **43**, 1571–1573 (1979).

[51] Steven Weinberg, "V-A was the key," J. Phys. Conf. Ser. **196**, 012002 (2009).

[52] Steven Weinberg, "Varieties of Baryon and Lepton Nonconservation," Phys. Rev. D **22**, 1694 (1980).

[53] Mark Claudson, Mark B. Wise, and Lawrence J. Hall, "Chiral Lagrangian for Deep Mine Physics," Nucl. Phys. B **195**, 297–307 (1982).

[54] S. Chadha and M. Daniel, "Chiral Lagrangian Calculation of Nucleon Decay Modes Induced by $d = 5$ Supersymmetric Operators," Nucl. Phys. B **229**, 105–114 (1983).

[55] L. F. Abbott and Mark B. Wise, "The Effective Hamiltonian for Nucleon Decay," Phys. Rev. D **22**, 2208 (1980).

Prepared for submission to JHEP

# FKS subtraction for quarkonium production at NLO

**Ajjath A H, Hua-Sheng Shao, and Lukas Simon**

*Laboratoire de Physique Théorique et Hautes Energies (LPTHE), UMR 7589, Sorbonne Université et CNRS, 4 place Jussieu, 75252 Paris Cedex 05, France*

*E-mail:* aabdulhameed@lpthe.jussieu.fr, huasheng.shao@lpthe.jussieu.fr, lsimon@lpthe.jussieu.fr

Abstract: We extend the local infrared-divergence subtraction formalism, originally proposed by Frixione, Kunszt and Signer (FKS), to calculate short-distance (differential) cross section for any inclusive process involving a quarkonium particle in non-relativistic QCD (NRQCD) factorisation at next-to-leading order (NLO) accuracy in the strong coupling constant $\alpha_s$. The new formulas are generally applicable to the production of an S- or P-wave quarkonium state in association with any number of elementary particles. The main new ingredients derived in this paper are the local and integrated soft counterterms for the colour-singlet and colour-octet P-wave bound states. It, therefore, paves the way to the automation of the NLO calculations for heavy quarkonium inclusive and associated production processes.

Keywords: NLO Computations, IR divergences, Quarkonium, QCD, NRQCD

arXiv:2402.19221v1 [hep-ph] 29 Feb 2024

## Contents


**1 Introduction** **1**

**2 Quarkonium production in NRQCD factorisation** **3**

**3 Soft limit** **7**
3.1 Soft limit at the amplitude level 7
3.1.1 Colour projection 8
3.1.2 Spin and orbital angular momentum projections 10
3.2 Soft limit of real matrix elements 15
3.2.1 Colour-singlet S-wave state 16
3.2.2 Colour-octet S-wave state 17
3.2.3 Colour-singlet spin-singlet P-wave state 17
3.2.4 Colour-octet spin-singlet P-wave state 19
3.2.5 Colour-singlet spin-triplet P-wave state 20
3.2.6 Colour-octet spin-triplet P-wave state 21

**4 FKS subtraction for single quarkonium production** **22**
4.1 FKS pairs and partition functions 22
4.2 Local subtraction counterterms 24
4.3 Integrated soft counterterm 27
4.3.1 Soft counterterms of LDMEs 28
4.3.2 Colour-singlet S-wave state 29
4.3.3 Colour-octet S-wave state 30
4.3.4 Colour-singlet spin-singlet P-wave state 30
4.3.5 Colour-octet spin-singlet P-wave state 31
4.3.6 Colour-singlet spin-triplet P-wave state 31
4.3.7 Colour-octet spin-triplet P-wave state 32
4.4 Integrated collinear and soft-collinear counterterms 32

**5 Validations** **35**

**6 Summary** **36**

**A Eikonal tensor integrals** **36**
A.1 Two-massless case 37
A.2 One-massive-one-massless case 38
A.3 Massive self-eikonal case 39
A.4 Two-massive case 40

**B Infrared poles of one-loop matrix elements** **42**
B.1 Colour-singlet S-wave state 42
B.2 Colour-octet S-wave state 43
B.3 Colour-singlet spin-singlet P-wave state 43
B.4 Colour-octet spin-singlet P-wave state 44
B.5 Colour-singlet spin-triplet P-wave state 45
B.6 Colour-octet spin-triplet P-wave state 46

## 1 Introduction

The theoretical interpretations of the analysed LHC data nowadays heavily rely on precision calculations of short-distance cross sections, as well as precise Monte Carlo event simulations in the context of both the Standard Model (SM) and its extensions. As of today, next-to-leading order (NLO) QCD calculations and their interface to general-purpose parton-shower Monte Carlo programmes have been automated, as seen, *e.g.*, in the MadGraph5_aMC@NLO [1] framework, for elementary-particle production processes [1] in the SM and in a large class of new physics models. NLO electroweak corrections have also been automated in recent years by several collaborations, such as MadGraph5_aMC@NLO [2] and Sherpa [3] along with external one-loop matrix element providers like Recola [4], OpenLoops [5] or GoSam [6]. Some low-particle-multiplicity processes, solved using customised methods, have been extended to next-to-NLO (NNLO) and even next-to-NNLO ($N^3LO$) accuracies. However, these significant theoretical developments are currently restricted to point-like elementary particles, and they cannot be directly applied to non-relativistic bound states like heavy quarkonia. This limitation can be roughly understood as the latter case intrinsically involving multiple scales and thus requiring simultaneous consideration in relativistic quantum field theories (QFT), such as QCD, and their non-relativistic low-energy effective field theories (EFT), *e.g.*, non-relativistic QCD (NRQCD) [7]. This introduces additional conceptual and technical challenges on the theory side.

Therefore, theoretical progress in perturbative calculations for heavy quarkonium is far less advanced. Fairly speaking, automation has been achieved only for tree-level quarkonium processes (single quarkonium in MadOnia [8] and one or more quarkonia in HELAC-Onia [9, 10]). NLO and even higher-order calculations, in many cases, are indispensable not only for precision or accuracy but also for a qualitative understanding. Due to conservation laws at the quantum level, short-distance cross sections of quarkonium production often receive giant $K$ factors [2] from high-order radiative corrections (see, *e.g.*, ref. [12] and

[1] Strictly speaking, this statement only applies to processes with not-too-high particle multiplicity, given the limitations of computing resources. It is also possible that special issues may arise for particular problems that have not yet been addressed in the automated codes. Additionally, this consideration does not cover processes involving loops at their lowest order. In the context of this paper, when we refer to a similar statement, we mean it in a loose sense.

[2] Thanks to recent advancements in parton showers [11], there is a chance that the problem of giant $K$ factors can be addressed through the matching and merging of matrix elements and parton showers.

references therein). This places the theoretical interpretations of measured quarkonium data on shaky ground if higher-order radiative corrections are not well under control.

The physics that we can learn from quarkonium is, however, no less interesting. In fact, quarkonia provide powerful and sometimes even unique tools that allow us to conduct rich particle and nuclear physics studies [13]. For instance, they can be used to determine the structures of free nucleons [14–25] and nuclei [26–29]. Due to their sequential binding energies, quarkonia are widely used as a thermometer of quark-gluon plasma produced in heavy-ion collisions [30, 31] to probe the hot-and-dense QCD. They signify the presence of a QCD phase transition by either disappearing [30, 31] or being abundantly produced, hinting at collective heavy-quark effects [32, 33]. A golden channel of searching for QCD instantons, arising from the non-trivial topological structure of QCD vacuum which is believed to be crucial in understanding quark confinement, was suggested to study charmonium decays [34, 35]. Quarkonia were also proposed as a good system to investigate the non-linear dynamics of QCD, also known as parton saturation [36], in addition to the well-known DGLAP and BFKL dynamics. They have been readily used to extract the fundamental SM parameters, *e.g.*, the strong coupling constant $\alpha_s$ [37], the Higgs-charm Yukawa coupling [38–40], the CKM matrix elements [41, 42], as well as the masses of the charm and bottom quarks [43]. Some exotic QCD hadrons, such as the fully-charmed tetraquark $X(6900)$ [44–46] and the first-observed pentaquark states $P_c^+(4380)$ and $P_c^+(4450)$ [47], were also discovered in final states with quarkonia.

It has been well understood that the perturbative calculations at NLO and beyond in QFT encounter ultraviolet (UV) and infrared (soft and collinear) divergences in the intermediate steps. While the UV divergences are removed through the renormalisation procedure, handling infrared (IR) divergences is more intricate. First, IR divergences can only be cancelled for so-called IR-safe observables, thanks to the Kinoshita-Lee-Nauenberg (KLN) theorem [48, 49] and factorisation theorems/conjectures. Second, in a generic situation, phase-space integration must be carried out numerically using Monte Carlo importance sampling methods. However, this is hindered by the IR singularities present in real radiative corrections. Both IR subtraction and phase-space slicing approaches are employed to overcome such complications in the real-emission contributions, with IR subtraction methods known to outperform slicing approaches. Therefore, the former serves as the backbone of contemporary NLO automation codes. The two widely adopted NLO subtraction methods were originally proposed by Frixione, Kunszt and Signer (FKS) [50, 51], and Catani and Seymour [52, 53]. They are usually referred to as the FKS and dipole subtraction schemes, respectively. Both methods work well for both simple and complicated processes involving elementary particles. Initially devised for massless coloured particles, they have been generalised to include massive coloured particles [54–56]. Both schemes have been implemented in various public computer programmes [56–61].

On the other hand, regarding the problem of heavy quarkonium production in NRQCD, NLO calculations in the literature are almost exclusively carried out using the slicing methods [62], with only a few exceptions. The earliest exception pertains to inclusive colour-singlet S-wave quarkonium production at hadron colliders [63], where the dipole counterterms for massless quarks and gluons suffice, as the colour-singlet S-wave quarkonium does

not exhibit any IR singularities. The remaining exceptional NLO calculations [64, 65] employ the dipole formalism [3] developed for processes featuring an S- or P-wave quarkonium alongside massless quarks and gluons, as outlined in ref. [66, 67]. If one considers a process involving both a P-wave quarkonium and massive partons, such as the associated production processes of quarkonium and heavy quarks, new (yet unknown) dipole terms may be required. The aim of this paper is to incorporate heavy quarkonium into the FKS subtraction scheme. As we demonstrate later, the formalism is general enough to be applied to arbitrary processes involving a quarkonium and massless/massive partons. Therefore, the scope of the phenomenological applications using our formalism is anticipated to be broader than that of the dipole formalism derived in refs. [66, 67].

Since our ultimate goal is to automate NLO computations for quarkonium production processes within the MadGraph5_aMC@NLO framework, we will closely adhere to the notations and conventions of the original FKS formulation [50] and the MadFKS paper [56]. For the sake of completeness and self-consistency, we will reproduce some known equations from the literature. We hope this will aid in improving the readability of the article, especially for readers who may not be familiar with the two aforementioned papers.

The remaining context of this paper is organised as follows. In section 2, we elucidate how to obtain short-distance cross sections for quarkonium production within NRQCD factorisation. This enables us to establish a few notations used throughout the paper. We derive the soft limit of the (squared) amplitudes in the real radiative corrections for a single quarkonium production in section 3. The local and integrated FKS subtraction counterterms are given in section 4. We perform a few cross-checks to ensure the validity of our formalism in section 5, and finally draw our conclusions in section 6. Appendix A presents the analytic expressions for eikonal tensor integrals that appear in the integrated counterterms. The universal IR poles of one-loop matrix elements can be found in appendix B.

## 2 Quarkonium production in NRQCD factorisation

In NRQCD factorisation [7], the inclusive production of a heavy quarkonium factorises into the perturbative short-distance cross section and the non-perturbative long-distance matrix elements (LDMEs):

$$d\sigma(AB \to H + X) = \sum_n \Big( \sum_{a,b,X} \int dx_a dx_b f_{a/A}(x_a) f_{b/B}(x_b) d\hat{\sigma}(ab \to Q\bar{Q}'[n] + X) \Big) \langle \mathcal{O}_n^H \rangle \,, \tag{2.1}$$

where $f_{a/A}$ and $f_{b/B}$ are parton distribution functions (PDFs) of partons $a$ and $b$ in the initial hadrons $A$ and $B$. $d\hat{\sigma}(ab \to Q\bar{Q}'[n]+X)$ describes the short distance production of a $Q\bar{Q}'$ pair [4] in a specific colour irreducible representation $C$, with spin $S$ and orbital angular momentum state $L$ denoted as $n = {}^{2S+1}L_J^{[C]}$ following the usual spectroscopic notation, and

[3] The formalism was derived under the circumstance of a $2 \to 2$ underlying Born process.

[4] Note that the flavours of the two constituent heavy quarks do not have to be identical. For instance, the constituent quarks of $B_c^+$ are a charm quark and a bottom antiquark.

the LDME, $\langle \mathcal{O}_n^H \rangle$, represents the hadronisation of the heavy quark pair into the physical quarkonium state $H$. An important consequence of NRQCD factorisation is the prediction that the LDMEs do not depend on the details of the hard process, and their values can be extracted from experiments, lattice QCD calculations [68, 69] or potential models [70].

In principle, for a specific quarkonium, there is an infinite number of Fock states $n$ and an infinite number of LDMEs $\langle \mathcal{O}_n^H \rangle$ to be determined, which limits the prediction power. Thanks to the power counting rules in NRQCD, only a limited number of Fock states should be involved in the calculations up to a specific order of $v$, where $v$ ($v \ll 1$) is the relative velocity of the heavy quark pair $Q\bar{Q}'$. We express a given Fock state using the spectroscopic notation ${}^{2S+1}L_J^{[C]}$.

Our focus is to evaluate the perturbative short-distance coefficients, which can be determined from the amplitudes of $Q$ and $\bar{Q}'$ production with necessary operations in order to constrain the heavy quark pair $Q\bar{Q}'$ into a specific quantum state $n$. A convenient way to do so is by performing projections. Let us consider a general $2 \to n$ process involving only open $Q$ and $\bar{Q}'$ quarks [5], denoted as $\mathcal{I}_1\mathcal{I}_2 \to \mathcal{I}_3\mathcal{I}_4 \cdots \mathcal{I}_{n+2}$, where the identity of the $k$-th particle is denoted by $\mathcal{I}_k$, and $\mathcal{I}_1 = a, \mathcal{I}_2 = b, \mathcal{I}_3 = Q$, and $\mathcal{I}_4 = \bar{Q}'$. Following the same notation as ref. [56], we can write the process as $r = (\mathcal{I}_1, \ldots, \mathcal{I}_{n+2})$. We denote the corresponding (tree-level) amplitude as $\mathcal{A}^{(n,0)}(r)$. Let us also define the amputated amplitude $\Gamma^{(n,0)}(r)$ by removing the external wavefunctions of $Q$ and $\bar{Q}'$ that will form a bound state $Q\bar{Q}'$, *i.e.*,

$$\mathcal{A}^{(n,0)}(r) = \bar{u}_{\lambda_Q}(k_Q)\Gamma^{(n,0)}(r)v_{\lambda_{\bar{Q}'}}(k_{\bar{Q}'}), \tag{2.2}$$

where $u_\lambda$ and $v_\lambda$ are Dirac spinors and $\lambda_{Q/\bar{Q}'}$ are helicities of $Q$ and $\bar{Q}'$, respectively.

Since $\mathbf{3} \otimes \overline{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$, we only have colour singlet $C = 1$ and colour octet $C = 8$ in the decomposition of $\mathcal{A}^{(n,0)}(r)$. The colour projectors are :

$$\begin{aligned} \mathbb{P}_{C=1} &= \frac{\delta_{c_4 c_3}}{\sqrt{N_c}}, \\ \mathbb{P}_{C=8} &= \sqrt{2} t^{c_{34}}_{c_4 c_3}, \end{aligned} \tag{2.3}$$

where $c_3, c_4$ are the colour indices of $Q$ and $\bar{Q}'$, and $t^{c_{34}}$ is the Gell-Mann matrix. In other words, we can define the following two amplitudes from $\mathcal{A}^{(n,0)}(r)$:

$$\begin{aligned} \mathcal{A}^{(n,0)}_{\{[C=1]\}}(r) &= \sum_{c_3,c_4} \mathbb{P}_{C=1}\mathcal{A}^{(n,0)}(r), \\ \mathcal{A}^{(n,0)}_{\{[C=8]\}}(r) &= \sum_{c_3,c_4} \mathbb{P}_{C=8}\mathcal{A}^{(n,0)}(r), \end{aligned} \tag{2.4}$$

where we have explicitly summed over the colour indices $c_3, c_4$.

Similarly, we only have spin singlet $S = 0$ and spin triplet $S = 1$ for the $Q\bar{Q}'$ pair. The spin projectors for the heavy quark momenta $k_Q^\mu = \frac{m_Q}{m_Q + m_{\bar{Q}'}} K^\mu + q^\mu$ and $k_{\bar{Q}'}^\mu =$

[5] We do not restrict the numbers of $Q$ and $\bar{Q}'$ appearing in the process.

$\frac{m_{\bar{Q}'}}{m_Q+m_{\bar{Q}'}}K^\mu - q^\mu$ are given by

$$\begin{aligned}\mathbb{P}_{S=0} &= \frac{1}{2\sqrt{2m_Q m_{\bar{Q}'}}}\bar{v}_{\lambda_{\bar{Q}'}}(k_{\bar{Q}'})\gamma_5 u_{\lambda_Q}(k_Q),\\ \mathbb{P}_{S=1} &= \frac{1}{2\sqrt{2m_Q m_{\bar{Q}'}}}\bar{v}_{\lambda_{\bar{Q}'}}(k_{\bar{Q}'})\not{\varepsilon}^*_{\lambda_s}(K) u_{\lambda_Q}(k_Q),\end{aligned} \tag{2.5}$$

where $\varepsilon^*_{\lambda_s}(K)$ is the polarisation vector for the spin-1 $Q\bar{Q}'$ with its spin quantum number as $\lambda_s = \pm 1, 0$. Here, $K$ is the four-momentum of the $Q\bar{Q}'$ pair, $q$ is the relative momentum between the two constituent heavy quarks, and $m_Q, m_{\bar{Q}'}$ are the heavy quark masses of $Q$ and $\bar{Q}'$, respectively. For simplicitly, we just denote $\tilde{\gamma}_0 = \gamma_5$ and $\tilde{\gamma}_1 = \not{\varepsilon}^*_{\lambda_s}(K)$. Thus, the amplitudes can be further decomposed into two spin configurations

$$\begin{aligned}\mathcal{A}^{(n,0)}_{\{S\}}(r) &= \sum_{\lambda_Q,\lambda_{\bar{Q}'}} \mathbb{P}_S \mathcal{A}^{(n,0)}(r)\\ &= \sum_{\lambda_Q,\lambda_{\bar{Q}'}} \frac{1}{2\sqrt{2m_Q m_{\bar{Q}'}}}\bar{v}_{\lambda_{\bar{Q}'}}(k_{\bar{Q}'})\tilde{\gamma}_S u_{\lambda_Q}(k_Q)\bar{u}_{\lambda_Q}(k_Q)\Gamma^{(n,0)}(r) v_{\lambda_{\bar{Q}'}}(k_{\bar{Q}'})\\ &= \frac{1}{2\sqrt{2m_Q m_{\bar{Q}'}}}\mathrm{Tr}_\gamma\left[\left(\not{k}_{\bar{Q}'} - m_{\bar{Q}'}\right)\tilde{\gamma}_S\left(\not{k}_Q + m_Q\right)\Gamma^{(n,0)}(r)\right].\end{aligned} \tag{2.6}$$

Together with both spin and colour configurations, the amplitude takes the form:

$$\begin{aligned}\mathcal{A}^{(n,0)}_{\{[C],S\}}(r) &= \sum_{\lambda_Q,\lambda_{\bar{Q}'}} \mathbb{P}_S \mathcal{A}^{(n,0)}_{\{[C]\}}(r)\\ &= \sum_{\lambda_Q,\lambda_{\bar{Q}'}}\sum_{c_3,c_4} \mathbb{P}_S \mathbb{P}_{[C]} \mathcal{A}^{(n,0)}(r)\\ &= \sum_{c_3,c_4} \mathbb{P}_{[C]} \frac{1}{2\sqrt{2m_Q m_{\bar{Q}'}}}\mathrm{Tr}_\gamma\left[\left(\not{k}_{\bar{Q}'} - m_{\bar{Q}'}\right)\tilde{\gamma}_S\left(\not{k}_Q + m_Q\right)\Gamma^{(n,0)}(r)\right].\end{aligned} \tag{2.7}$$

Here, the spin projection operator commutes with the colour projection operator. Note that the trace $\mathrm{Tr}_\gamma$ in the Dirac spinor space does not necessarily represent a real trace that we need to compute. Its evaluation depends on how the other fermion lines are organised in the amputated amplitude $\Gamma^{(n,0)}(r)$.

The non-relativistic nature, *i.e.*, in the rest frame of $Q\bar{Q}'$, $q \ll \sqrt{m_Q m_{\bar{Q}'}}$, allows us to expand the amplitudes into the series of $v \sim q/\sqrt{m_Q m_{\bar{Q}'}} \ll 1$. This gives us the eigenfunctions of the orbital angular momentum operator. The projection on a state with orbital angular momentum $L$ is obtained by differentiating $L = 0, 1, \ldots$ (*a la* $S, P, \ldots$ waves) times the spin-colour projected amplitude with respect to the relative momentum $q$ of the heavy quarks in the $Q\bar{Q}'$ rest frame, and then setting $q \to 0$. Considering only $L = 0, 1$ states, which we are only interested in at this stage, the amplitude takes the form:

$$\mathcal{A}^{(n,0)}_{\{[C],S,L\}}(r) = \left[\left(\varepsilon^{\mu,*}_{\lambda_l}(K)\frac{d}{dq^\mu}\right)^L \mathcal{A}^{(n,0)}_{[C],S}(r)\right]_{q=0}, \tag{2.8}$$

where $\varepsilon^{\mu,*}_{\lambda_l}(K)$ is the polarisation vector for $L = 1$ orbital angular momentum with $\lambda_l = \pm 1, 0$. Since the spin projectors depend on the relative momentum $q$, the orbital angular

momentum expansion must be carried out after projecting onto the given spin configuration.

Finally, the total angular momentum $J$ is uniquely determined by $L$ or $S$ unless $L \neq 0$ and $S \neq 0$. In our specific case of interest, this can only occur when $L = 1$ and $S = 1$. In the latter case, we know how to determine $J = 0, 1, 2$ and $\lambda_j = -J, -J+1, \ldots, J-1, J$ from quantum mechanics, *i.e.*,

$$\varepsilon^{\mu\nu,*}_{J,\lambda_j}(K) = \sum_{\lambda_s,\lambda_l} \langle J, \lambda_j | 1, \lambda_l; 1, \lambda_s \rangle \varepsilon^{\mu *}_{\lambda_l}(K) \varepsilon^{\nu *}_{\lambda_s}(K), \tag{2.9}$$

where $\langle J, \lambda_j | 1, \lambda_l; 1, \lambda_s \rangle$ is the Clebsch-Gordan coefficient. For $J = 0, 1$, the expressions are

$$\begin{aligned} \varepsilon^{\mu\nu,*}_{0,0}(K) &= \frac{1}{\sqrt{3}} \left( g^{\mu\nu} - \frac{K^\mu K^\nu}{K^2} \right), \\ \varepsilon^{\mu\nu,*}_{1,\lambda_j}(K) &= -\frac{i}{\sqrt{2}} \epsilon^{\mu\nu\alpha\beta} \frac{K_\alpha}{\sqrt{K^2}} \varepsilon^*_{\lambda_j,\beta}(K), \end{aligned} \tag{2.10}$$

where $\epsilon^{\mu\nu\alpha\beta}$ is the Levi-Civita tensor. Thus, we obtain the amplitude $\mathcal{A}^{(n,0)}_{\{[C],S,L,J\}}(r)$ for a given quantum number ${}^{2S+1}L^{[C]}_J$. When the product $LS = 0$, we find $\mathcal{A}^{(n,0)}_{\{[C],S,L,J\}}(r) = \mathcal{A}^{(n,0)}_{\{[C],S,L\}}(r)$. On the other hand, if $L = S = 1$, we have

$$\begin{aligned} \mathcal{A}^{(n,0)}_{\{[C],1,1,J\}}(r) &= \sum_{\lambda_s,\lambda_l} \langle J, \lambda_j | 1, \lambda_l; 1, \lambda_s \rangle \mathcal{A}^{(n,0)}_{\{[C],1,1\}}(r) \\ &= \frac{1}{2\sqrt{2 m_Q m_{\bar{Q}'}}} \sum_{c_3,c_4} \mathbb{P}_{[C]} \varepsilon^{\mu\nu,*}_{J,\lambda_j}(K) \\ &\quad \times \frac{d}{dq^\mu} \mathrm{Tr}_\gamma \left[ \left( \not{k}_{\bar{Q}'} - m_{\bar{Q}'} \right) \gamma_\nu \left( \not{k}_Q + m_Q \right) \Gamma^{(n,0)}(r) \right] \Big|_{q=0} . \end{aligned} \tag{2.11}$$

After all of the above preparations, we can now glue $\mathcal{I}_3 = Q$ and $\mathcal{I}_4 = \bar{Q}'$ as a new single particle $\mathcal{I}_{3\oplus 4} = Q\bar{Q}'[{}^{2S+1}L^{[C]}_J]$ with four-momentum $K^\mu$ and invariant mass $\sqrt{K^2} = m_{Q\bar{Q}'} = m_Q + m_{\bar{Q}'}$ in the non-relativistic limit. The new process is denoted as

$$\dot{r} = r^{3\oplus 4, \not{4}} = (\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_{3\oplus 4}, \not{\mathcal{I}}_4, \ldots \mathcal{I}_{n+2}) \ . \tag{2.12}$$

The amplitude for the process $\dot{r} = r^{3\oplus 4, \not{4}}$ is

$$\mathbb{A}^{(n-1,0)}(\dot{r}) = \mathcal{A}^{(n,0)}_{\{[C],S,L,J\}}(r). \tag{2.13}$$

Note that the final state symmetry must be applied at the level of $\dot{r}$ not the initial $r$, while the phase space integration should also be carried out at the level of $\dot{r}$. The partonic cross section can be written as

$$d\hat{\sigma}(\dot{r}) = \frac{1}{\mathcal{N}(\dot{r})} \underbrace{\frac{1}{(2J+1)N_{[C]}} \frac{m_Q + m_{\bar{Q}'}}{2 m_Q m_{\bar{Q}'}}}_{=\mathcal{G}(\dot{r})} \left( \mathbb{M}^{(n-1,0)}(\dot{r}) J^{n^{(B)}_L} \right) d\phi_{n-1}(\dot{r}), \tag{2.14}$$

where the amplitude square is given by

$$\mathbb{M}^{(n-1,0)}(\dot{r}) = \frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)}\sum_{\substack{\text{colour}\\\text{spin}}}\left|\mathbb{A}^{(n-1,0)}(\dot{r})\right|^2$$
$$= \mathcal{M}^{(n,0)}_{\{[C],S,L,J\}}(r) = \frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)}\sum_{\substack{\text{colour}\\\text{spin}}}\left|\mathcal{A}^{(n,0)}_{\{[C],S,L,J\}}(r)\right|^2, \tag{2.15}$$

$\mathcal{N}(\dot{r})$ is the final state symmetry factor, $N_{[C=1]} = 2N_c, N_{[C=8]} = N_c^2 - 1$ with $N_c = 3$ in QCD, and $d\phi_{n-1}(\dot{r})$ is the $(n-1)$-body phase space measure. The Mandelstam variable $s = (k_1 + k_2)^2 = 2k_1 \cdot k_2$, and $\omega(\mathcal{I})$ is the product of spin and colour degrees of freedom for the particle $\mathcal{I}$. The condition with the measurement $n_L^{(B)}$-jet function $J^{n_L^{(B)}}$, where $n_L^{(B)}$ is the number of the light partons in the underlying Born, is sufficient to prevent the appearance of phase-space singularities in the Born-like quantities. Without losing generality, we can always assume these cuts to be equivalent to the request of having either $n_L^{(B)}$ and $n_L^{(B)}+1$ jets in the final state for an NLO computation. The same procedure can be iterated if we have more-than-one quarkonia.

## 3 Soft limit

At NLO, we have contributions coming from one-loop virtual corrections and real emissions besides Born. Due to the complexities introduced by bound states, it is necessary to derive new local and integrated FKS counterterms to handle the IR divergences in real contributions at NLO. We observe that, since the constituent quarks $Q$ and $\bar{Q}'$ are massive, we can recycle the counterterms for collinear and soft-collinear origins. What we need to deal with beyond elementary particle production is the soft but non-collinear part. This includes two new components. The first involves the usual soft counterterms that locally cancel singularities of real emissions and their one-body phase space integrated counterparts. The second consists of additional integrated counterterms resulting from the renormalisation of LDMEs, analogous to the usual initial/final collinear counterterms but originating from a soft origin. Similar to the latter, which are necessary to cancel the remaining IR divergences protected by the collinear factorisation in perturbative QCD, the former are a consequence of the NRQCD factorisation formalism. In this section, we begin by considering the soft limit of the quarkonium real emission amplitudes and their squares, following the procedure outlined before.

### 3.1 Soft limit at the amplitude level

Let us consider a $2 \to n+1$ real process denoted as $r = (\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, \mathcal{I}_4, \ldots, \mathcal{I}_{n+3})$. Following the approach in ref. [56], and without loss of generality, we can always reorder the final-state particles such that the final massless (anti-)quarks and gluons are $3 \leq i \leq n_L^{(R)} + 2$ and the strongly-interacting particles are $n_I \leq j \leq n_L^{(R)} + n_H + 2$, where $n_I = 1, 2, 3$ for hadron-hadron, lepton-hadron, and lepton-lepton collisions. Here, $n_H$ represents the number of massive coloured partons in the final state. Simultaneously, we have $\mathcal{I}_{n_L^{(R)}+n_H+1} = Q$ and

$\mathcal{I}_{n_L^{(R)}+n_H+2} = \bar{Q}'$. Consequently, we always have $n_H \geq 2$. For simplicity, in the following discussion, we denote $j_Q = n_L^{(R)} + n_H + 1$ and $j_{\bar{Q}'} = n_L^{(R)} + n_H + 2$. In the soft limit of $\mathcal{I}_i = g$ with $3 \leq i \leq n_L^{(R)} + 2$, the real amplitude can be expressed in terms of the reduced Born amplitude multiplied with an additional eikonal factor. If the soft gluon with four-momentum $k_i$ is emitted from an external leg $\mathcal{I}_j$ ($j \neq i$) with four-momentum $k_j$, the real emission amplitude takes the form in the soft or eikonal approximation:

$$\lim_{k_i \to 0} \mathcal{A}^{(n+1,0)}(r) = g_s \frac{k_j \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_j \cdot k_i} \vec{Q}(\mathcal{I}_j) \mathcal{A}^{(n,0)}(r^{\not{i}}), \tag{3.1}$$

where the reduced process is $r^{\not{i}} = (\mathcal{I}_1, \ldots \not{\mathcal{I}}_i, \ldots \mathcal{I}_j, \ldots \mathcal{I}_{n+3})$, and $\vec{Q}(\mathcal{I})$ represents the colour generator associated with the particle $\mathcal{I}$, and $g_s = \sqrt{4\pi\alpha_s}$ is the strong coupling. Depending on the particle species $\mathcal{I}$, we have

$$\vec{Q}(\mathcal{I}) = \{t^a\}_{a=1}^8, \quad \{-t^{aT}\}_{a=1}^8, \quad \{T^a\}_{a=1}^8 \quad \mathcal{I} \in \mathbf{3}, \overline{\mathbf{3}}, \mathbf{8} \tag{3.2}$$

with $t^a$ and $T^a$ being the SU(3) generators in the fundamental and adjoint representations, respectively. We take the final quark or initial antiquark as $\mathbf{3}$, while the initial quark and final antiquark are taken as $\overline{\mathbf{3}}$. The matrix element of the adjoint representation is $T^a_{bc} = -if_{abc}$ with $f_{abc}$ being the anti-symmetric structure constants.

After performing the quantum number projection, as described in sect. 2, we obtain a similar eikonal decomposition as long as $j \neq j_Q, j_{\bar{Q}'}$, *i.e.*,

$$\lim_{k_i \to 0} \mathcal{A}^{(n+1,0)}_{\{[C],S,L,J\}}(r) = g_s \frac{k_j \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_j \cdot k_i} \vec{Q}(\mathcal{I}_j) \mathcal{A}^{(n,0)}_{\{[C],S,L,J\}}(r^{\not{i}}). \tag{3.3}$$

However, when $j = j_Q$ or $j = j_{\bar{Q}'}$, we should pay special attention to it. It will be always convenient to sum the contributions of $j = j_Q$ and $j = j_{\bar{Q}'}$ together, which we will adopt in the following. Now, let us consider the case of $j = j_Q, j_{\bar{Q}'}$.

### 3.1.1 Colour projection

With the procedure outlined in sect. 2, the colour projected amplitudes are

$$\begin{aligned}
&\lim_{k_i \to 0} \mathcal{A}^{(n+1,0)}_{\{[C=1]\}}(r) \\
&= g_s \sum_{c_{j_Q}, c_{j_{\bar{Q}'}}} \mathbb{P}_{[C=1]} \left[ \frac{k_{j_Q} \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q} \cdot k_i} \vec{Q}(\mathcal{I}_{j_Q}) + \frac{k_{j_{\bar{Q}'}} \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}} \cdot k_i} \vec{Q}(\mathcal{I}_{j_{\bar{Q}'}}) \right] \mathcal{A}^{(n,0)}(r^{\not{i}}) \\
&= g_s \sum_{c_{j_Q}, c_{j_{\bar{Q}'}}} \frac{\delta_{c_{j_Q} c_{j_{\bar{Q}'}}}}{\sqrt{N_c}} \left[ \frac{k_{j_Q} \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q} \cdot k_i} t^a_{c_{j_Q} c'_{j_Q}} \mathcal{A}^{(n,0)}_{c'_{j_Q} c_{j_{\bar{Q}'}}}(r^{\not{i}}) - \frac{k_{j_{\bar{Q}'}} \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}} \cdot k_i} t^a_{c'_{j_{\bar{Q}'}} c_{j_{\bar{Q}'}}} \mathcal{A}^{(n,0)}_{c_{j_Q} c'_{j_{\bar{Q}'}}}(r^{\not{i}}) \right] \\
&= g_s \left[ \frac{k_{j_Q} \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q} \cdot k_i} - \frac{k_{j_{\bar{Q}'}} \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}} \cdot k_i} \right] \frac{\delta_{a c'_{j_Q j_{\bar{Q}'}}}}{\sqrt{2N_c}} \mathcal{A}^{(n,0)}_{\{[C=8]\}}(r^{\not{i}}),
\end{aligned} \tag{3.4}$$

and

$$
\begin{aligned}
&\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[C=8]\}}(r)\\
&= g_s\sum_{c_{j_Q},c_{j_{\bar{Q}'}}}\mathbb{P}_{[C=8]}\left[\frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i}\vec{Q}(\mathcal{I}_{j_Q})+\frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i}\vec{Q}(\mathcal{I}_{j_{\bar{Q}'}})\right]\mathcal{A}^{(n,0)}(r^{\not i})\\
&= g_s\sum_{c_{j_Q},c_{j_{\bar{Q}'}}}\sqrt{2}t^{c_{j_Q}j_{\bar{Q}'}}_{c_{j_{\bar{Q}'}}c_{j_Q}}\left[\frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i}t^a_{c_{j_Q}c'_{j_{\bar{Q}'}}}\mathcal{A}^{(n,0)}_{c'_{j_Q}c_{j_{\bar{Q}'}}}(r^{\not i})-\frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i}t^a_{c'_{j_{\bar{Q}'}}c_{j_{\bar{Q}'}}}\mathcal{A}^{(n,0)}_{c_{j_Q}c'_{j_{\bar{Q}'}}}(r^{\not i})\right]\\
&= g_s\left[\frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i}-\frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i}\right]\left[\frac{\delta_{ac_{j_Q}j_{\bar{Q}'}}}{\sqrt{2N_c}}\mathcal{A}^{(n,0)}_{\{[C=1]\}}(r^{\not i})+\frac{1}{2}d_{ac_{j_Q}j_{\bar{Q}'}c'_{j_Q}j_{\bar{Q}'}}\mathcal{A}^{(n,0)}_{\{[C=8]\}}(r^{\not i})\right]\\
&\quad+g_s\left[\frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i}+\frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i}\right]\left(-\frac{i}{2}f_{ac_{j_Q}j_{\bar{Q}'}c'_{j_Q}j_{\bar{Q}'}}\right)\mathcal{A}^{(n,0)}_{\{[C=8]\}}(r^{\not i}),
\end{aligned}
\tag{3.5}
$$

where $d_{abc}$'s are the symmetric structure constants. Moreover, we have used the following relation of Gell-Mann matrices

$$
(t^at^b)_{kl}=\frac{\delta_{ab}}{2N_c}\delta_{kl}+\frac{1}{2}(d_{abc}+if_{abc})t^c_{kl}
\tag{3.6}
$$

and have assumed the colour index for $Q\bar{Q}'$ in the real process $r$ (the reduced Born process $r^{\not i}$) to be $c_{j_Qj_{\bar{Q}'}}$ ($c'_{j_Qj_{\bar{Q}'}}$). We can put the two equations into the compact matrix form by using the colour nonet **9** index $\mathbf{b}=0,1,2,\ldots,8$, where $\mathbf{b}=0$ corresponds to $\mathcal{A}^{(n+1,0)}_{\{[C=1]\}}(r)$ and $\mathcal{A}^{(n,0)}_{\{[C=1]\}}(r^{\not i})$, while $\mathbf{b}=1,\ldots,8$ are $\mathcal{A}^{(n+1,0)}_{\{[C=8]\}}(r)$ and $\mathcal{A}^{(n,0)}_{\{[C=8]\}}(r^{\not i})$ with the colour index of the $Q\bar{Q}'$ as $\mathbf{b}$. In other words, with the emitters being $Q$ and $\bar{Q}'$, we have

$$
\begin{aligned}
\lim_{k_i\to 0}\begin{pmatrix}\mathcal{A}^{(n+1,0)}_{\{[C=1]\}}(r)\\ \mathcal{A}^{(n+1,0)}_{\{[C=8]\},\mathbf{b}=1}(r)\\ \vdots\\ \mathcal{A}^{(n+1,0)}_{\{[C=8]\},\mathbf{b}=8}(r)\end{pmatrix} &= g_s\left\{\left[\frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i}-\frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i}\right]\vec{Q}_1(Q\bar{Q}')\right.\\
&\quad\left.+\left[\frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i}+\frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i}\right]\vec{Q}_2(Q\bar{Q}')\right\}\\
&\quad\times\begin{pmatrix}\mathcal{A}^{(n,0)}_{\{[C=1]\}}(r^{\not i})\\ \mathcal{A}^{(n,0)}_{\{[C=8]\},\mathbf{b}=1}(r^{\not i})\\ \vdots\\ \mathcal{A}^{(n,0)}_{\{[C=8]\},\mathbf{b}=8}(r^{\not i})\end{pmatrix},
\end{aligned}
\tag{3.7}
$$

where the colour generators are

$$\vec{Q}_1(Q\bar{Q}') = \begin{pmatrix} 0 & \frac{\delta_{a1}}{\sqrt{2N_c}} & \cdots & \frac{\delta_{a8}}{\sqrt{2N_c}} \\ \frac{\delta_{a1}}{\sqrt{2N_c}} & \ddots & & \iddots \\ \vdots & & \frac{D^a}{2} & \\ \frac{\delta_{a8}}{\sqrt{2N_c}} & \iddots & & \ddots \end{pmatrix}, \tag{3.8}$$

$$\vec{Q}_2(Q\bar{Q}') = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \ddots & & \iddots \\ \vdots & & \frac{T^a}{2} & \\ 0 & \iddots & & \ddots \end{pmatrix}, \tag{3.9}$$

with the elements of the matrix $D^a$ as $D^a_{bc} = d_{abc}$.

### 3.1.2 Spin and orbital angular momentum projections

The next step is to perform the projection for the spin and the orbital angular momentum via

$$\begin{aligned} &\lim_{k_i\to 0} \begin{pmatrix} \mathcal{A}^{(n+1,0)}_{\{[C=1],S,L\}}(r) \\ \mathcal{A}^{(n+1,0)}_{\{[C=8],S,L\},\mathbf{b}=1}(r) \\ \vdots \\ \mathcal{A}^{(n+1,0)}_{\{[C=8],S,L\},\mathbf{b}=8}(r) \end{pmatrix} \\ &= \frac{g_s}{2\sqrt{2m_Q m_{\bar{Q}'}}} \left\{ \left( \varepsilon^{\mu,*}_{\lambda_l}(K) \frac{d}{dq^\mu} \right)^L \left[ \left( \frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i} - \frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i} \right) \vec{Q}_1(Q\bar{Q}') \right.\right. \\ &\left.\left. + \left( \frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i} + \frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i} \right) \vec{Q}_2(Q\bar{Q}') \right] \right. \\ &\left. \times \sum_{c_{j_Q}, c_{j_{\bar{Q}'}}} \begin{pmatrix} \frac{\delta_{c_{j_Q} c_{j_{\bar{Q}'}}}}{\sqrt{N_c}} \\ \sqrt{2} t^1_{c_{j_{\bar{Q}'}} c_{j_Q}} \\ \vdots \\ \sqrt{2} t^8_{c_{j_{\bar{Q}'}} c_{j_Q}} \end{pmatrix} \mathrm{Tr}_\gamma \left( \left( \not{k}_{\bar{Q}'} - m_{\bar{Q}'} \right) \tilde{\gamma}_S \left( \not{k}_Q + m_Q \right) \Gamma^{(n,0)}(r^{\not{\perp}}) \right) \right\}_{q=0} . \end{aligned} \tag{3.10}$$

When $L = 0$, we just set the relative momentum $q$ to be zero. The coefficient of $\vec{Q}_1(Q\bar{Q}')$ vanishes. This implies the following two consequences:

- For the colour singlet $C = 1$ with $L = 0$ (S-wave), regardless of the value of the spin $S$, there are no soft divergences. Thus, we can treat a colour-singlet S-wave state as any other elementary colour-singlet particle, such as $Z$ and $H$ bosons, from the IR perspective.

- For the colour octet $C = 8$ S-wave states, we have the following soft limit relation

$$\lim_{k_i\to 0} \mathcal{A}^{(n+1,0)}_{\{[8],S,0\}}(r) = g_s \frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i} \vec{Q}(Q\bar{Q}'[{}^{2S+1}S^{[8]}_J]) \mathcal{A}^{(n,0)}_{\{[8],S,0\}}(r^{\not{\perp}}), \tag{3.11}$$

where now $Q\bar{Q}'[^{2S+1}S_J^{[8]}] \in \mathbf{8}$ and $\vec{Q}(Q\bar{Q}'[^{2S+1}S_J^{[8]}]) = \{T^a\}_{a=1}^8$. It means that, in the soft limit, a colour-octet S-wave state behaves as if $Q\bar{Q}'$ were an elementary colour-octet particle, akin to a sgluon or a Kaluza-Klein massive gluon.

For P-wave ($L = 1$) Fock states, we have

$$\begin{aligned}\lim_{k_i\to 0}\begin{pmatrix}\mathcal{A}^{(n+1,0)}_{\{[C=1],S,1\}}(r)\\ \mathcal{A}^{(n+1,0)}_{\{[C=8],S,1\},\mathbf{b}=1}(r)\\ \vdots\\ \mathcal{A}^{(n+1,0)}_{\{[C=8],S,1\},\mathbf{b}=8}(r)\end{pmatrix} &= g_s\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}2\vec{Q}_2(Q\bar{Q}')\begin{pmatrix}\mathcal{A}^{(n,0)}_{\{[C=1],S,1\}}(r^{\nwarrow})\\ \mathcal{A}^{(n,0)}_{\{[C=8],S,1\},\mathbf{b}=1}(r^{\nwarrow})\\ \vdots\\ \mathcal{A}^{(n,0)}_{\{[C=8],S,1\},\mathbf{b}=8}(r^{\nwarrow})\end{pmatrix}\\ &\quad+g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right]\\ &\quad\times\left[\left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)\vec{Q}_1(Q\bar{Q}')+\left(\frac{m_{\bar{Q}'}}{m_Q}-\frac{m_Q}{m_{\bar{Q}'}}\right)\vec{Q}_2(Q\bar{Q}')\right]\\ &\quad\times\begin{pmatrix}\mathcal{A}^{(n,0)}_{\{[C=1],S,0\}}(r^{\nwarrow})\\ \mathcal{A}^{(n,0)}_{\{[C=8],S,0\},\mathbf{b}=1}(r^{\nwarrow})\\ \vdots\\ \mathcal{A}^{(n,0)}_{\{[C=8],S,0\},\mathbf{b}=8}(r^{\nwarrow})\end{pmatrix},\end{aligned}\tag{3.12}$$

because of

$$\begin{aligned}\varepsilon^{\mu,*}_{\lambda_l}(K)\frac{d}{dq^\mu}\frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i} &= \frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_Q}\cdot k_i}-\frac{k_{j_Q}\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{\left(k_{j_Q}\cdot k_i\right)^2}\\ &\overset{q=0}{=}\frac{m_Q+m_{\bar{Q}'}}{m_Q}\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right],\\ \varepsilon^{\mu,*}_{\lambda_l}(K)\frac{d}{dq^\mu}\frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i} &= -\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_{j_{\bar{Q}'}}\cdot k_i}+\frac{k_{j_{\bar{Q}'}}\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{\left(k_{j_{\bar{Q}'}}\cdot k_i\right)^2}\\ &\overset{q=0}{=}-\frac{m_Q+m_{\bar{Q}'}}{m_{\bar{Q}'}}\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right].\end{aligned}\tag{3.13}$$

This essentially means that

- For the colour-singlet ($C = 1$) P-wave states, we get

$$\begin{aligned}\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[C=1],S,1\}}(r) &= g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right]\\ &\quad\times\left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)\frac{\delta_{ac'_{j_Q j_{\bar{Q}'}}}}{\sqrt{2N_c}}\mathcal{A}^{(n,0)}_{\{[C=8],S,0\}}(r^{\nwarrow}).\end{aligned}\tag{3.14}$$

  The object $\mathcal{I}_{j_Q\oplus j_{\bar{Q}'}\oplus i}$ may be interpreted as a new colour-octet S-wave particle (with colour index $c'_{j_Q j_{\bar{Q}'}} = a$) with a non-standard eikonal factor.

- For the colour-octet ($C = 8$) P-wave states, the soft limit yields a more complicated expression

$$\begin{aligned}\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[C=8],S,1\}}(r) = & g_s\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}\vec{Q}(Q\bar{Q}'[^{2S+1}P^{[8]}_J])\mathcal{A}^{(n,0)}_{\{[C=8],S,1\}}(r\!\!\!/) \\ & +g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right] \\ & \times\left[\left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)\left(\frac{\delta_{ac_{j_Qj_{\bar{Q}'}}}}{\sqrt{2N_c}}\mathcal{A}^{(n,0)}_{\{[C=1],S,0\}}(r\!\!\!/)\right.\right. \\ & \left.+\frac{d_{ac_{j_Qj_{\bar{Q}'}}c'_{j_Qj_{\bar{Q}'}}}}{2}\mathcal{A}^{(n,0)}_{\{[C=8],S,0\}}(r\!\!\!/)\right) \\ & \left.+\frac{1}{2}\left(\frac{m_{\bar{Q}'}}{m_Q}-\frac{m_Q}{m_{\bar{Q}'}}\right)\vec{Q}(Q\bar{Q}'[^{2S+1}S^{[8]}_J])\mathcal{A}^{(n,0)}_{\{[C=8],S,0\}}(r\!\!\!/)\right].\end{aligned} \tag{3.15}$$

If we define the following effective colour generators

$$\begin{aligned}\vec{Q}_{\rm eff}(Q\bar{Q}'_{[18]}) &= \left\{\left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)\frac{\Delta^a}{\sqrt{2N_c}}\right\}^8_{a=1}, \\ \vec{Q}_{\rm eff}(Q\bar{Q}'_{[81]}) &= \left\{\left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)\frac{\Delta^{aT}}{\sqrt{2N_c}}\right\}^8_{a=1}, \\ \vec{Q}_{\rm eff}(Q\bar{Q}'_{[88]}) &= \left\{\left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)\frac{D^a}{2}+\left(\frac{m_{\bar{Q}'}}{m_Q}-\frac{m_Q}{m_{\bar{Q}'}}\right)\frac{T^a}{2}\right\}^8_{a=1},\end{aligned} \tag{3.16}$$

where $\Delta^a_{ob}=\delta_{ab}$ and $\circ$ represents no colour index for a colour-singlet state, the soft-limit expressions can be written in a more compact form

$$\begin{aligned}\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[C=1],S,1\}}(r) = & g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right] \\ & \times\vec{Q}_{\rm eff}(Q\bar{Q}'_{[18]})\mathcal{A}^{(n,0)}_{\{[C=8],S,0\}}(r\!\!\!/),\end{aligned} \tag{3.17}$$

$$\begin{aligned}\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[C=8],S,1\}}(r) = & g_s\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}\vec{Q}(Q\bar{Q}'[^{2S+1}P^{[8]}_J])\mathcal{A}^{(n,0)}_{\{[C=8],S,1\}}(r\!\!\!/) \\ & +g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right] \\ & \times\left[\vec{Q}_{\rm eff}(Q\bar{Q}'_{[81]})\mathcal{A}^{(n,0)}_{\{[C=1],S,0\}}(r\!\!\!/)+\vec{Q}_{\rm eff}(Q\bar{Q}'_{[88]})\mathcal{A}^{(n,0)}_{\{[C=8],S,0\}}(r\!\!\!/)\right].\end{aligned} \tag{3.18}$$

Here, we have defined the effective Casimir constants through the following relations

$$
\begin{aligned}
C_{\text{eff}}(Q\bar{Q}'_{[18]}) &= \vec{Q}_{\text{eff}}(Q\bar{Q}'_{[18]})\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[18]}) = \left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)^2\frac{1}{2N_c}\,,\\
C_{\text{eff}}(Q\bar{Q}'_{[81]}) &= \vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]})\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]}) = \left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)^2 C_F\,,\\
C_{\text{eff}}(Q\bar{Q}'_{[88]}) &= \vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]})\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]}) = \left(2+\frac{m_Q}{m_{\bar{Q}'}}+\frac{m_{\bar{Q}'}}{m_Q}\right)\\
&\qquad\times\left(-\frac{2}{N_c}+\frac{N_c^2-2}{2N_c}\left(\frac{m_{\bar{Q}'}}{m_Q}+\frac{m_Q}{m_{\bar{Q}'}}\right)\right).
\end{aligned}
\tag{3.19}
$$

In addition, we also have

$$
\begin{aligned}
\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]})\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]}) &= \vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]})\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]}) = 0,\\
\vec{Q}(Q\bar{Q}'[^{2S+1}P_J^{[8]}])\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]}) &= \vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]})\cdot\vec{Q}(Q\bar{Q}'[^{2S+1}P_J^{[8]}]) = 0\,,\\
\vec{Q}(Q\bar{Q}'[^{2S+1}P_J^{[8]}])\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]}) &= \vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]})\cdot\vec{Q}(Q\bar{Q}'[^{2S+1}P_J^{[8]}]) = \frac{1}{2}\left(\frac{m_{\bar{Q}'}}{m_Q}-\frac{m_Q}{m_{\bar{Q}'}}\right)C_A\,.
\end{aligned}
\tag{3.20}
$$

Finally, in order to get a given total angular momentum $J$ when $L=S=1$, we need to use eq.(2.9) at the amplitude level:

$$
\begin{aligned}
\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[C=1],1,1,J\}}(r) &= \sum_{\lambda_l,\lambda_s}\langle J,\lambda_j|1,\lambda_l;1,\lambda_s\rangle g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right]\\
&\quad\times\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[18]})\mathcal{A}^{(n,0)}_{\{[C=8],1,0\}}(r\!\!\!/),
\end{aligned}
\tag{3.21}
$$

and

$$
\begin{aligned}
\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[C=8],1,1,J\}}(r) &= g_s\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}\vec{Q}(Q\bar{Q}'[^3P_J^{[8]}])\mathcal{A}^{(n,0)}_{\{[C=8],1,1,J\}}(r\!\!\!/)\\
&\quad+\sum_{\lambda_l,\lambda_s}\langle J,\lambda_j|1,\lambda_l;1,\lambda_s\rangle g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right]\\
&\quad\times\left[\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]})\mathcal{A}^{(n,0)}_{\{[C=1],S,0\}}(r\!\!\!/)+\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]})\mathcal{A}^{(n,0)}_{\{[C=8],S,0\}}(r\!\!\!/)\right].
\end{aligned}
\tag{3.22}
$$

To rewrite this in a unified compact form, we can further define the following effective identity (called "flavour") operators

$$
\begin{aligned}
\vec{F}_{\text{eff}}(Q\bar{Q}'_{[18]}) &: Q\bar{Q}'[^{2S+1}P_J^{[1]}]\to Q\bar{Q}'[^{2S+1}S_S^{[8]}],\\
\vec{F}_{\text{eff}}(Q\bar{Q}'_{[81]}) &: Q\bar{Q}'[^{2S+1}P_J^{[8]}]\to Q\bar{Q}'[^{2S+1}S_S^{[1]}],\\
\vec{F}_{\text{eff}}(Q\bar{Q}'_{[88]}) &: Q\bar{Q}'[^{2S+1}P_J^{[8]}]\to Q\bar{Q}'[^{2S+1}S_S^{[8]}].
\end{aligned}
\tag{3.23}
$$

The eikonal current operator for an elementary coloured particle $\mathcal{I}_j$ is

$$\vec{J}(\mathcal{I}_j) = \frac{k_j \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_j \cdot k_i} \vec{Q}(\mathcal{I}_j), \tag{3.24}$$

while for bound states it is

$$\begin{aligned}
\vec{J}(Q\bar{Q}'[^{2S+1}S^{[1]}_S]) &= 0, \\
\vec{J}(Q\bar{Q}'[^{2S+1}S^{[8]}_S]) &= \frac{K \cdot \varepsilon^*_{\lambda_i}(k_i)}{K \cdot k_i} \vec{Q}(Q\bar{Q}'[^{2S+1}S^{[8]}_S]), \\
\vec{J}(Q\bar{Q}'[^{1}P^{[1]}_1]) &= \left[ \frac{\varepsilon^*_{\lambda_l}(K) \cdot \varepsilon^*_{\lambda_i}(k_i)}{K \cdot k_i} - \frac{K \cdot \varepsilon^*_{\lambda_i}(k_i) k_i \cdot \varepsilon^*_{\lambda_l}(K)}{(K \cdot k_i)^2} \right] \vec{Q}_{\rm eff}(Q\bar{Q}'_{[18]}) \vec{F}_{\rm eff}(Q\bar{Q}'_{[18]}), \\
\vec{J}(Q\bar{Q}'[^{1}P^{[8]}_1]) &= \frac{K \cdot \varepsilon^*_{\lambda_i}(k_i)}{K \cdot k_i} \vec{Q}(Q\bar{Q}'[^{2S+1}P^{[8]}_J]) \\
&\quad + \left[ \frac{\varepsilon^*_{\lambda_l}(K) \cdot \varepsilon^*_{\lambda_i}(k_i)}{K \cdot k_i} - \frac{K \cdot \varepsilon^*_{\lambda_i}(k_i) k_i \cdot \varepsilon^*_{\lambda_l}(K)}{(K \cdot k_i)^2} \right] \\
&\quad \times \left[ \vec{Q}_{\rm eff}(Q\bar{Q}'_{[81]}) \vec{F}_{\rm eff}(Q\bar{Q}'_{[81]}) + \vec{Q}_{\rm eff}(Q\bar{Q}'_{[88]}) \vec{F}_{\rm eff}(Q\bar{Q}'_{[88]}) \right], \\
\vec{J}(Q\bar{Q}'[^{3}P^{[1]}_J]) &= \sum_{\lambda_l, \lambda_s} \langle J, \lambda_j | 1, \lambda_1; 1, \lambda_s \rangle \left[ \frac{\varepsilon^*_{\lambda_l}(K) \cdot \varepsilon^*_{\lambda_i}(k_i)}{K \cdot k_i} - \frac{K \cdot \varepsilon^*_{\lambda_i}(k_i) k_i \cdot \varepsilon^*_{\lambda_l}(K)}{(K \cdot k_i)^2} \right] \\
&\quad \times \vec{Q}_{\rm eff}(Q\bar{Q}'_{[18]}) \vec{F}_{\rm eff}(Q\bar{Q}'_{[18]}), \\
\vec{J}(Q\bar{Q}'[^{3}P^{[8]}_J]) &= \frac{K \cdot \varepsilon^*_{\lambda_i}(k_i)}{K \cdot k_i} \vec{Q}(Q\bar{Q}'[^{2S+1}P^{[8]}_J]) \\
&\quad + \sum_{\lambda_l, \lambda_s} \left[ \frac{\varepsilon^*_{\lambda_l}(K) \cdot \varepsilon^*_{\lambda_i}(k_i)}{K \cdot k_i} - \frac{K \cdot \varepsilon^*_{\lambda_i}(k_i) k_i \cdot \varepsilon^*_{\lambda_l}(K)}{(K \cdot k_i)^2} \right] \\
&\quad \times \langle J, \lambda_j | 1, \lambda_l; 1, \lambda_s \rangle \left[ \vec{Q}_{\rm eff}(Q\bar{Q}'_{[81]}) \vec{F}_{\rm eff}(Q\bar{Q}'_{[81]}) + \vec{Q}_{\rm eff}(Q\bar{Q}'_{[88]}) \vec{F}_{\rm eff}(Q\bar{Q}'_{[88]}) \right].
\end{aligned} \tag{3.25}$$

In such a case, the amplitude in the soft limit of $k_i \to 0$ can be rewritten as

$$\lim_{k_i \to 0} \mathbb{A}^{(n,0)}(\dot{r}) = g_s \sum_{\substack{j=n_I \\ j \neq i}}^{n^{(R)}_L + n_H + 1} \vec{J}(\mathcal{I}_j) \mathbb{A}^{(n-1,0)}(\dot{r}^{\not{i}}), \tag{3.26}$$

where we have glued the $Q\bar{Q}'$ pair into a single particle in processes with a dot, *i.e.*,

$$\begin{aligned}
\dot{r} &= r^{j_Q \oplus j_{\bar{Q}'}, \not{j}_{\bar{Q}'}} = \left( \mathcal{I}_1, \ldots, \mathcal{I}_i, \ldots, \mathcal{I}_j, \ldots, \mathcal{I}_{n^{(R)}_L + n_H}, Q\bar{Q}'[^{2S+1}L^{[C]}_J], \not{\bar{Q}}', \ldots, \mathcal{I}_{n+3} \right), \\
\dot{r}^{\not{i}} &= \left( \mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, \mathcal{I}_{n^{(R)}_L + n_H}, Q\bar{Q}'[^{2S+1}L^{[C]}_J], \not{\bar{Q}}', \ldots, \mathcal{I}_{n+3} \right).
\end{aligned} \tag{3.27}$$

Note that $\dot{r}$ is essentially equivalent to eq.(2.12) except that we have reordered the final particles, and the total number of external parton legs is $n+3$ instead of $n+2$.

### 3.2 Soft limit of real matrix elements

We can now examine the soft limit of the real amplitude square by considering the $\ell$-loop amplitude $\mathcal{A}^{(n,\ell)}(r)$ for a generic $2 \to n$ process $r$. To facilitate our later discussion, we introduce the following amplitude squares:

$$\mathbb{M}^{(n,0)}(\dot{r}) = \mathcal{M}^{(n+1,0)}_{\{[C],S,L,J\}}(r) = \frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)} \sum_{\substack{\text{colour}\\\text{spin}}} \left|\mathcal{A}^{(n+1,0)}_{\{[C],S,L,J\}}(r)\right|^2, \tag{3.28}$$

$$\begin{aligned}\mathbb{M}^{(n,0)}_J(\dot{r})^{\mu\nu} = {}& \frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)} \\ & \times \sum_{\substack{\text{colour}\\\text{spin}}} \left[ \left( \sum_{\lambda_l,\lambda_s} \langle J,\lambda_j|1,\lambda_l;1,\lambda_s\rangle \varepsilon^{\mu,*}_{\lambda_l}(K) \mathbb{A}^{(n-1,0)}(\dot{r}) \right) \right. \\ & \left. \times \left( \sum_{\lambda'_l,\lambda'_s} \langle J,\lambda_j|1,\lambda'_l;1,\lambda'_s\rangle \varepsilon^{\nu,*}_{\lambda'_l}(K) \mathbb{A}^{(n-1,0)}(\dot{r}) \right)^\star \right],\end{aligned} \tag{3.29}$$

$$\mathbb{M}^{(n-1,0)}(\dot{r}) = \mathcal{M}^{(n,0)}_{\{[C],S,L,J\}}(r) = \frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)} \sum_{\substack{\text{colour}\\\text{spin}}} \left|\mathcal{A}^{(n,0)}_{\{[C],S,L,J\}}(r)\right|^2, \tag{3.30}$$

$$\begin{aligned}\mathcal{M}^{(n,0)}_{\{[C],S,L,J\},kl}(r) = {}& -\frac{1}{2s}\frac{2-\delta_{kl}}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)} \\ & \times \sum_{\substack{\text{colour}\\\text{spin}}} \mathcal{A}^{(n,0)}_{\{[C],S,L,J\}}(r) \vec{Q}(\mathcal{I}_k)\cdot\vec{Q}(\mathcal{I}_l) \mathcal{A}^{(n,0)}_{\{[C],S,L,J\}}(r)^\star,\end{aligned} \tag{3.31}$$

$$\begin{aligned}\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) = {}& -\frac{1}{2s}\frac{2-\delta_{kl}}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)} \\ & \times \sum_{\substack{\text{colour}\\\text{spin}}} \mathbb{A}^{(n-1,0)}(\dot{r}) \vec{Q}(\mathcal{I}_k)\cdot\vec{Q}(\mathcal{I}_l) \mathbb{A}^{(n-1,0)}(\dot{r})^\star,\end{aligned} \tag{3.32}$$

$$\begin{aligned}\mathbb{M}^{(n-1,0)}_{k[C_1C_2]}(\dot{r}_1,\dot{r}_2)^\mu = {}& -\frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)} \\ & \sum_{\substack{\text{colour}\\\text{spin}}} 2\Re\Big\{ \varepsilon^{*,\mu}_{\lambda_l}(K) \mathbb{A}^{(n-1,0)}(\dot{r}_1)^\star \vec{Q}(\mathcal{I}_k)\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[C_1C_2]}) \mathbb{A}^{(n-1,0)}(\dot{r}_2)\Big\},\end{aligned} \tag{3.33}$$

$$\begin{aligned}\mathbb{M}^{(n-1,0)}_{J,k[C_1C_2]}(\dot{r}_1,\dot{r}_2)^\mu = {}& -\frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)} \\ & \sum_{\substack{\text{colour}\\\text{spin}}} 2\Re\Big\{ \mathbb{A}^{(n-1,0)}(\dot{r}_1)^\star \vec{Q}(\mathcal{I}_k)\cdot\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[C_1C_2]}) \\ & \times \sum_{\lambda_l,\lambda_s} \langle J,\lambda_j|1,\lambda_l;1,\lambda_s\rangle \varepsilon^{*,\mu}_{\lambda_l}(K) \mathbb{A}^{(n-1,0)}(\dot{r}_2) \Big\},\end{aligned} \tag{3.34}$$

$$\mathcal{M}^{(n,1)}(r) = \frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)} \sum_{\substack{\text{colour}\\\text{spin}}} 2\Re\left\{\mathcal{A}^{(n,0)}(r)\mathcal{A}^{(n,1)}(r)^\star\right\}. \tag{3.35}$$

Once again, in these equations, $s = (k_1 + k_2)^2 = 2k_1 \cdot k_2$, and $\omega(\mathcal{I})$ represents the product of spin and colour degrees of freedom for particle $\mathcal{I}$. In $d = 4 - 2\epsilon$ dimensions, the average factors are given by $\omega(q) = \omega(\bar{q}) = 2N_c = 6$ and $\omega(g) = 2(1-\epsilon)(N_c^2 - 1) = 16(1-\epsilon)$. These average factors are utilised to fully specify the divergent part of the one-loop contribution within the conventional dimensional regularisation (CDR) scheme. It is important to note that while these factors include $\epsilon$ for completeness, in numerical calculations, $d = 4$ tree-level amplitudes are typically evaluated, and the $\epsilon$ dependence is dropped.

In general, we can write down the soft limit of the amplitude square in the form

$$\lim_{k_i \to 0} \mathbb{M}^{(n,0)}(\dot{r}) = g_s^2 \sum_{\substack{k,l=n_I \\ k,l\neq i, k\leq l}}^{n_L^{(R)}+n_H+1} \mathsf{M}^{(n-1,0)}_{\text{soft},kl}(\dot{r}^{\not{i}}). \tag{3.36}$$

It is well known that if both $\mathcal{I}_k$ and $\mathcal{I}_l$ are elementary particles, we have

$$\mathsf{M}^{(n-1,0)}_{\text{soft},kl}(\dot{r}^{\not{i}}) = \frac{k_k \cdot k_l}{k_k \cdot k_i k_l \cdot k_i} \mathbb{M}^{(n-1,0)}_{kl}(\dot{r}^{\not{i}}). \tag{3.37}$$

In the following, we will try to derive $\mathsf{M}^{(n-1,0)}_{\text{soft},kl}(\dot{r}^{\not{i}})$ when at least $\mathcal{I}_k$ or $\mathcal{I}_l$ is a quarkonium state.

### 3.2.1 Colour-singlet S-wave state

We have derived the soft limit of colour-singlet S-wave states at the amplitude level in sect. 3.1 as

$$\lim_{k_i \to 0} \mathcal{A}^{(n+1,0)}_{\{[1],S,0,J=S\}}(r) = \sum_{\substack{j=n_I \\ j\neq i}}^{n_L^{(R)}+n_H} g_s \frac{k_j \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_j \cdot k_i} \vec{Q}(\mathcal{I}_j) \mathcal{A}^{(n,0)}_{\{[1],S,0,J=S\}}(r^{\not{i}}). \tag{3.38}$$

It follows that the amplitude square in the soft limit is given by

$$\begin{aligned} \lim_{k_i \to 0} \mathbb{M}^{(n,0)}(\dot{r}) = \lim_{k_i \to 0} \mathcal{M}^{(n+1,0)}_{\{[1],S,0,J=S\}}(r) &= g_s^2 \sum_{\substack{k,l=n_I \\ k,l\neq i, k\leq l}}^{n_L^{(R)}+n_H} \frac{k_k \cdot k_l}{k_k \cdot k_i k_l \cdot k_i} \mathcal{M}^{(n,0)}_{\{[1],S,0,J=S\},kl}(r^{\not{i}}) \\ &= g_s^2 \sum_{\substack{k,l=n_I \\ k,l\neq i, k\leq l}}^{n_L^{(R)}+n_H} \frac{k_k \cdot k_l}{k_k \cdot k_i k_l \cdot k_i} \mathbb{M}^{(n-1,0)}_{kl}(\dot{r}^{\not{i}}), \end{aligned} \tag{3.39}$$

where

$$\dot{r}^{\not{i}} = \left(\mathcal{I}_1, \ldots, \not{\mathcal{I}_i}, \ldots, \mathcal{I}_j, \ldots, \mathcal{I}_{n_L^{(R)}+n_H}, Q\bar{Q}'[{}^{2S+1}S_J^{[1]}], \bar{Q}', \ldots, \mathcal{I}_{n+3}\right). \tag{3.40}$$

In eq.(3.39), we have used the light-cone axial gauge

$$\sum_{\lambda_i} \varepsilon^{\mu *}_{\lambda_i}(k_i) \varepsilon^{\nu}_{\lambda_i}(k_i) = -g^{\mu\nu} + \frac{k_i^\mu n^\nu + k_i^\nu n^\mu}{k_i \cdot n}, \tag{3.41}$$

where $n^\mu$ is a light-like auxiliary vector ($n^2 = 0$). The terms involving $n^\mu$ cancel out in eq.(3.39) as a result of colour conservation. Therefore, we have, for $n_I \leq k \leq n_L^{(R)} + n_H + 1$,

$$\mathsf{M}^{(n-1,0)}_{\text{soft},kj_Q}(\dot{r}^{\not{k}}) = 0. \tag{3.42}$$

### 3.2.2 Colour-octet S-wave state

For colour-octet S-wave states, we have derived the soft limit at the amplitude level in sect. 3.1 (*cf.* eq.(3.11)) as

$$\begin{aligned}\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[8],S,0,J=S\}}(r) = &\sum_{\substack{j=n_I\\ j\neq i}}^{n_L^{(R)}+n_H} g_s\frac{k_j\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_j\cdot k_i}\vec{Q}(\mathcal{I}_j)\mathcal{A}^{(n,0)}_{\{[8],S,0,J=S\}}(r^{\not{k}})\\ &+g_s\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}\vec{Q}(Q\bar{Q}'[^{2S+1}S^{[8]}_J])\mathcal{A}^{(n,0)}_{\{[8],S,0,J=S\}}(r^{\not{k}})\,.\end{aligned} \tag{3.43}$$

The amplitude square in the soft limit is

$$\lim_{k_i\to 0}\mathbb{M}^{(n,0)}(\dot{r}) = \lim_{k_i\to 0}\mathcal{M}^{(n+1,0)}_{\{[8],S,0,J=S\}}(r) = g_s^2\sum_{\substack{k,l=n_I\\ k,l\neq i,k\leq l}}^{n_L^{(R)}+n_H+1}\frac{\tilde{k}_k\cdot\tilde{k}_l}{\tilde{k}_k\cdot k_i\tilde{k}_l\cdot k_i}\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}^{\not{k}})\,, \tag{3.44}$$

where $\tilde{k}_{n_L^{(R)}+n_H+1} = K$ and $\tilde{k}_j = k_j$ when $j \neq j_Q, j_{\bar{Q}'}$ and

$$\dot{r}^{\not{k}} = \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[^{2S+1}S^{[8]}_J],\bar{\mathcal{Q}}',\ldots,\mathcal{I}_{n+3}\right). \tag{3.45}$$

Thanks to colour conservation, the $n^\mu$ dependent terms stemming from eq.(3.41) vanish on the right-hand side (r.h.s) of eq.(3.44). Thus, we have, for $n_I \leq k \leq n_L^{(R)} + n_H$,

$$\begin{aligned}\mathsf{M}^{(n-1,0)}_{\text{soft},kj_Q}(\dot{r}^{\not{k}}) &= \frac{k_k\cdot K}{k_k\cdot k_i K\cdot k_i}\mathbb{M}^{(n-1,0)}_{kj_Q}(\dot{r}^{\not{k}})\,,\\ \mathsf{M}^{(n-1,0)}_{\text{soft},j_Qj_Q}(\dot{r}^{\not{k}}) &= -\frac{K^2}{(K\cdot k_i)^2}C_A\mathbb{M}^{(n-1,0)}(\dot{r}^{\not{k}})\,.\end{aligned} \tag{3.46}$$

### 3.2.3 Colour-singlet spin-singlet P-wave state

With eq.(3.17) in sect. 3.1, the soft limit of the amplitude for a colour-singlet spin-singlet P-wave quarkonium is

$$\begin{aligned}\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[1],0,1,1\}}(r) = &\sum_{\substack{j=n_I\\ j\neq i}}^{n_L^{(R)}+n_H} g_s\frac{k_j\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_j\cdot k_i}\vec{Q}(\mathcal{I}_j)\mathcal{A}^{(n,0)}_{\{[1],0,1,1\}}(r^{\not{k}})\\ &+g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right]\\ &\times\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[18]})\mathcal{A}^{(n,0)}_{\{[8],0,0,0\}}(r^{\not{k}}).\end{aligned} \tag{3.47}$$

Then, the amplitude square in the soft limit becomes

$$\begin{aligned}
&\lim_{k_i\to 0}\mathbb{M}^{(n,0)}(\dot{r}) = \lim_{k_i\to 0}\mathcal{M}^{(n+1,0)}_{\{[1],0,1,1\}}(r)\\
&= g_s^2\sum_{\substack{k,l=n_I\\ k,l\neq i,k\leq l}}^{n_L^{(R)}+n_H}\frac{k_k\cdot k_l}{k_k\cdot k_i k_l\cdot k_i}\mathbb{M}^{(n-1,0)}_{kl}(\dot{\not{r}})\\
&+g_s^2\sum_{\substack{k=n_I\\ k\neq i}}^{n_L^{(R)}+n_H}\left[\frac{k_{k,\mu}}{k_k\cdot k_i K\cdot k_i}-\frac{K\cdot k_k k_{i,\mu}}{k_k\cdot k_i\left(K\cdot k_i\right)^2}\right]\mathbb{M}^{(n-1,0)}_{k[18]}(\dot{\not{r}},\dot{\not{r}}_1)^\mu\\
&-g_s^2\frac{2\epsilon-2}{\left(K\cdot k_i\right)^2}C_{\text{eff}}(Q\bar{Q}'_{[18]})\mathbb{M}^{(n-1,0)}(\dot{\not{r}}_1),
\end{aligned}\tag{3.48}$$

where

$$\begin{aligned}
\dot{\not{r}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^1P_1^{[1]}],\bar{Q}',\ldots,\mathcal{I}_{n+3}\right),\\
\dot{\not{r}}_1 &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^1S_0^{[8]}],\bar{Q}',\ldots,\mathcal{I}_{n+3}\right),
\end{aligned}\tag{3.49}$$

and we have used the relations

$$K\cdot\varepsilon^{(*)}_{\lambda_l}(K) = 0,\tag{3.50}$$

$$\sum_{\lambda_l}\varepsilon^{\mu,*}_{\lambda_l}(K)\varepsilon^{\nu}_{\lambda_l}(K) = -g^{\mu\nu}+\frac{K^\mu K^\nu}{K^2}.\tag{3.51}$$

In numerical calculations, setting the dimensional regulator $\epsilon$ to zero in the prefactor of the soft matrix element ensures consistency with the use of 4-dimensional external wave functions. The terms dependent on $n^\mu$ from eq.(3.41) are absent on the r.h.s of eq.(3.48) due to colour conservation and

$$\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{\left(K\cdot k_i\right)^2}\right]_{\varepsilon^*_{\lambda_i}(k_i)\to k_i} = 0.\tag{3.52}$$

The same observation holds for the other P-wave states discussed later, and there is no need to reiterate this point.

Thus, we have, for $n_I\leq k\leq n_L^{(R)}+n_H$,

$$\begin{aligned}
\mathsf{M}^{(n-1,0)}_{\text{soft},kj_Q}(\dot{\not{r}}) &= \left[\frac{k_{k,\mu}}{k_k\cdot k_i K\cdot k_i}-\frac{K\cdot k_k k_{i,\mu}}{k_k\cdot k_i\left(K\cdot k_i\right)^2}\right]\mathbb{M}^{(n-1,0)}_{k[18]}(\dot{\not{r}},\dot{\not{r}}_1)^\mu,\\
\mathsf{M}^{(n-1,0)}_{\text{soft},j_Q j_Q}(\dot{\not{r}}) &= -\frac{2\epsilon-2}{\left(K\cdot k_i\right)^2}C_{\text{eff}}(Q\bar{Q}'_{[18]})\mathbb{M}^{(n-1,0)}(\dot{\not{r}}_1).
\end{aligned}\tag{3.53}$$

### 3.2.4 Colour-octet spin-singlet P-wave state

For the production of a colour-octet spin-singlet P-wave quarkonium, we have established the soft limit at the amplitude level (*cf.* eq.(3.18) in sect. 3.1) as follows:

$$\begin{aligned}\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[8],0,1,1\}}(r) = &\sum_{\substack{j=n_I\\ j\neq i}}^{n_L^{(R)}+n_H} g_s\frac{k_j\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_j\cdot k_i}\vec{Q}(\mathcal{I}_j)\mathcal{A}^{(n,0)}_{\{[8],0,1,1\}}(\dot{r})\\ &+g_s\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}\vec{Q}(Q\bar{Q}'[^1P_1^{[8]}])\mathcal{A}^{(n,0)}_{\{[8],0,1,1\}}(\dot{r})\\ &+g_s\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right]\\ &\times\left[\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]})\mathcal{A}^{(n,0)}_{\{[1],0,0,0\}}(\dot{r})+\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]})\mathcal{A}^{(n,0)}_{\{[8],0,0,0\}}(\dot{r})\right].\end{aligned} \tag{3.54}$$

The amplitude square in the soft limit is

$$\begin{aligned}&\lim_{k_i\to 0}\mathbb{M}^{(n,0)}(\dot{r}) = \lim_{k_i\to 0}\mathcal{M}^{(n+1,0)}_{\{[8],0,1,1\}}(r)\\ &= g_s^2\sum_{\substack{k,l=n_I\\ k,l\neq i,k\leq l}}^{n_L^{(R)}+n_H+1}\frac{\tilde{k}_k\cdot\tilde{k}_l}{\tilde{k}_k\cdot k_i\tilde{k}_l\cdot k_i}\mathbb{M}^{(n-1,0)}_{kl}(\dot{r})\\ &+g_s^2\sum_{\substack{k=n_I\\ k\neq i}}^{n_L^{(R)}+n_H+1}\left[\frac{\tilde{k}_{k,\mu}}{\tilde{k}_k\cdot k_i K\cdot k_i}-\frac{K\cdot\tilde{k}_k k_{i,\mu}}{\tilde{k}_k\cdot k_i\,(K\cdot k_i)^2}\right]\left[\mathbb{M}^{(n-1,0)}_{k[88]}(\dot{r},\dot{r}_1)^\mu+\mathbb{M}^{(n-1,0)}_{k[81]}(\dot{r},\dot{r}_2)^\mu\right]\\ &-g_s^2\frac{2\epsilon-2}{(K\cdot k_i)^2}\left[C_{\text{eff}}(Q\bar{Q}'_{[88]})\mathbb{M}^{(n-1,0)}(\dot{r}_1)+C_{\text{eff}}(Q\bar{Q}'_{[81]})\mathbb{M}^{(n-1,0)}(\dot{r}_2)\right],\end{aligned} \tag{3.55}$$

where

$$\begin{aligned}\dot{r} &= \left(\mathcal{I}_1,\ldots,\cancel{\mathcal{I}_i},\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[^1P_1^{[8]}],\bar{Q}',\ldots,\mathcal{I}_{n+3}\right),\\ \dot{r}_1 &= \left(\mathcal{I}_1,\ldots,\cancel{\mathcal{I}_i},\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[^1S_0^{[8]}],\bar{Q}',\ldots,\mathcal{I}_{n+3}\right),\\ \dot{r}_2 &= \left(\mathcal{I}_1,\ldots,\cancel{\mathcal{I}_i},\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[^1S_0^{[1]}],\bar{Q}',\ldots,\mathcal{I}_{n+3}\right),\end{aligned} \tag{3.56}$$

and $\tilde{k}_{n_L^{(R)}+n_H+1} = K$ and $\tilde{k}_j = k_j$ when $j\neq j_Q, j_{\bar{Q}'}$. Additionally, we have used the relations eq.(3.51).

Thus, we have, for $n_I \leq k \leq n_L^{(R)} + n_H$,

$$\begin{aligned}
\mathsf{M}^{(n-1,0)}_{\text{soft},kj_Q}(\dot{r}^{\not{1}}) &= \frac{k_k \cdot K}{k_k \cdot k_i K \cdot k_i} \mathbb{M}^{(n-1,0)}_{kj_Q}(\dot{r}^{\not{1}}) \\
&\quad + \left[ \frac{k_{k,\mu}}{k_k \cdot k_i K \cdot k_i} - \frac{K \cdot k_k k_{i,\mu}}{k_k \cdot k_i \left(K \cdot k_i\right)^2} \right] \left[ \mathbb{M}^{(n-1,0)}_{k[88]}(\dot{r}^{\not{1}}, \dot{r}^{\not{1}}_1)^\mu + \mathbb{M}^{(n-1,0)}_{k[81]}(\dot{r}^{\not{1}}, \dot{r}^{\not{1}}_2)^\mu \right], \\
\mathsf{M}^{(n-1,0)}_{\text{soft},j_Q j_Q}(\dot{r}^{\not{1}}) &= -\frac{K^2}{\left(K \cdot k_i\right)^2} C_A \mathbb{M}^{(n-1,0)}(\dot{r}^{\not{1}}) \\
&\quad + \left[ \frac{K_\mu}{\left(K \cdot k_i\right)^2} - \frac{K^2 k_{i,\mu}}{\left(K \cdot k_i\right)^3} \right] \left[ \mathbb{M}^{(n-1,0)}_{j_Q[88]}(\dot{r}^{\not{1}}, \dot{r}^{\not{1}}_1)^\mu + \mathbb{M}^{(n-1,0)}_{j_Q[81]}(\dot{r}^{\not{1}}, \dot{r}^{\not{1}}_2)^\mu \right] \\
&\quad - \frac{2\epsilon - 2}{\left(K \cdot k_i\right)^2} \left[ C_{\text{eff}}(Q\bar{Q}'_{[88]}) \mathbb{M}^{(n-1,0)}(\dot{r}^{\not{1}}_1) + C_{\text{eff}}(Q\bar{Q}'_{[81]}) \mathbb{M}^{(n-1,0)}(\dot{r}^{\not{1}}_2) \right]. \qquad (3.57)
\end{aligned}$$

### 3.2.5 Colour-singlet spin-triplet P-wave state

We have derived the soft limit at the amplitude level for a colour-singlet spin-triplet P-wave state in sect. 3.1 (*cf.* eq.(3.21)) as

$$\begin{aligned}
\lim_{k_i \to 0} \mathcal{A}^{(n+1,0)}_{\{[1],1,1,J\}}(r) &= \sum_{\substack{j=n_I \\ j \neq i}}^{n_L^{(R)}+n_H} g_s \frac{k_j \cdot \varepsilon^*_{\lambda_i}(k_i)}{k_j \cdot k_i} \vec{Q}(\mathcal{I}_j) \mathcal{A}^{(n,0)}_{\{[1],1,1,J\}}(r^{\not{1}}) \\
&\quad + g_s \sum_{\lambda_l, \lambda_s} \langle J, \lambda_j | 1, \lambda_l; 1, \lambda_s \rangle \left[ \frac{\varepsilon^*_{\lambda_l}(K) \cdot \varepsilon^*_{\lambda_i}(k_i)}{K \cdot k_i} - \frac{K \cdot \varepsilon^*_{\lambda_i}(k_i) k_i \cdot \varepsilon^*_{\lambda_l}(K)}{\left(K \cdot k_i\right)^2} \right] \\
&\quad \times \vec{Q}_{\text{eff}}(Q\bar{Q}'_{[18]}) \mathcal{A}^{(n,0)}_{\{[8],1,0,1\}}(r^{\not{1}}). \qquad (3.58)
\end{aligned}$$

The amplitude square in the soft limit is

$$\begin{aligned}
\lim_{k_i \to 0} \mathbb{M}^{(n,0)}(\dot{r}) &= \lim_{k_i \to 0} \mathcal{M}^{(n+1,0)}_{\{[1],1,1,J\}}(r) \\
&= g_s^2 \sum_{\substack{k,l=n_I \\ k,l \neq i, k \leq l}}^{n_L^{(R)}+n_H} \frac{k_k \cdot k_l}{k_k \cdot k_i k_l \cdot k_i} \mathbb{M}^{(n-1,0)}_{kl}(\dot{r}^{\not{1}}) \\
&\quad + g_s^2 \sum_{\substack{k=n_I \\ k \neq i}}^{n_L^{(R)}+n_H} \left[ \frac{k_{k,\mu}}{k_k \cdot k_i K \cdot k_i} - \frac{K \cdot k_k k_{i,\mu}}{k_k \cdot k_i \left(K \cdot k_i\right)^2} \right] \mathbb{M}^{(n-1,0)}_{J,k[18]}(\dot{r}, \dot{r}_1)^\mu \\
&\quad - g_s^2 \left[ \frac{g_{\mu\nu}}{\left(K \cdot k_i\right)^2} + \frac{K^2 k_{i,\mu} k_{i,\nu}}{\left(K \cdot k_i\right)^4} \right] C_{\text{eff}}(Q\bar{Q}'_{[18]}) \mathbb{M}^{(n-1,0)}_J(\dot{r}^{\not{1}}_1)^{\mu\nu}, \qquad (3.59)
\end{aligned}$$

where

$$\begin{aligned}
\dot{r}^{\not{1}} &= \left( \mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, \mathcal{I}_{n_L^{(R)}+n_H}, Q\bar{Q}'[{}^3P_J^{[1]}], \bar{Q}', \ldots, \mathcal{I}_{n+3} \right), \\
\dot{r}^{\not{1}}_1 &= \left( \mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, \mathcal{I}_{n_L^{(R)}+n_H}, Q\bar{Q}'[{}^3S_1^{[8]}], \bar{Q}', \ldots, \mathcal{I}_{n+3} \right), \qquad (3.60)
\end{aligned}$$

and we have used the relations

$$\begin{aligned}
K\cdot\varepsilon^{(*)}_{\lambda_l}(K) &= K\cdot\varepsilon^{(*)}_{\lambda_s}(K)=0\,,\\
\sum_{\lambda_l}\varepsilon^{\mu,*}_{\lambda_l}(K)\varepsilon^{\nu}_{\lambda_l}(K) &= -g^{\mu\nu}+\frac{K^\mu K^\nu}{K^2}=\Pi^{\mu\nu}(K)\,,\\
\varepsilon^{\mu\nu,*}_{0,0}(K)\varepsilon^{\alpha\beta}_{0,0}(K) &= \frac{1}{3}\Pi^{\mu\nu}(K)\Pi^{\alpha\beta}(K)\,,\\
\sum_{\lambda_j=-1}^{1}\varepsilon^{\mu\nu,*}_{1,\lambda_j}(K)\varepsilon^{\alpha\beta}_{1,\lambda_j}(K) &= \frac{1}{2}\left[\Pi^{\mu\alpha}(K)\Pi^{\nu\beta}(K)-\Pi^{\mu\beta}(K)\Pi^{\nu\alpha}(K)\right]\,,\\
\sum_{\lambda_j=-2}^{2}\varepsilon^{\mu\nu,*}_{2,\lambda_j}(K)\varepsilon^{\alpha\beta}_{2,\lambda_j}(K) &= \frac{1}{2}\left[\Pi^{\mu\alpha}(K)\Pi^{\nu\beta}(K)+\Pi^{\mu\beta}(K)\Pi^{\nu\alpha}(K)\right]-\frac{1}{3}\Pi^{\mu\nu}(K)\Pi^{\alpha\beta}(K)\,.
\end{aligned}\tag{3.61}$$

Then, we can easily derive,

$$\begin{aligned}
\mathsf{M}^{(n-1,0)}_{\text{soft},kj_Q}(\dot{r}^{\natural}) &= \left[\frac{k_{k,\mu}}{k_k\cdot k_i K\cdot k_i}-\frac{K\cdot k_k k_{i,\mu}}{k_k\cdot k_i\,(K\cdot k_i)^2}\right]\mathbb{M}^{(n-1,0)}_{J,k[18]}(\dot{r}^{\natural},\dot{r}^{\natural}_1)^\mu\,,\\
\mathsf{M}^{(n-1,0)}_{\text{soft},j_Qj_Q}(\dot{r}^{\natural}) &= -\left[\frac{g_{\mu\nu}}{(K\cdot k_i)^2}+\frac{K^2k_{i,\mu}k_{i,\nu}}{(K\cdot k_i)^4}\right]C_{\text{eff}}(Q\bar{Q}'_{[18]})\mathbb{M}^{(n-1,0)}_{J}(\dot{r}^{\natural}_1)^{\mu\nu}\,,
\end{aligned}\tag{3.62}$$

with $n_I\le k\le n_L^{(R)}+n_H$.

### 3.2.6 Colour-octet spin-triplet P-wave state

Finally, for the colour-octet spin-triplet P-wave state production, the soft limit of the amplitude from sect. 3.1 (*cf.* eq.(3.22)) is

$$\begin{aligned}
\lim_{k_i\to 0}\mathcal{A}^{(n+1,0)}_{\{[8],1,1,J\}}(r) = &\sum_{\substack{j=n_I\\ j\neq i}}^{n_L^{(R)}+n_H} g_s\frac{k_j\cdot\varepsilon^*_{\lambda_i}(k_i)}{k_j\cdot k_i}\vec{Q}(\mathcal{I}_j)\mathcal{A}^{(n,0)}_{\{[8],1,1,J\}}(r^{\natural})\\
&+g_s\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}\vec{Q}(Q\bar{Q}'[^3P^{[8]}_J])\mathcal{A}^{(n,0)}_{\{[8],1,1,J\}}(r^{\natural})\\
&+g_s\sum_{\lambda_l,\lambda_s}\langle J,\lambda_j|1,\lambda_l;1,\lambda_s\rangle\left[\frac{\varepsilon^*_{\lambda_l}(K)\cdot\varepsilon^*_{\lambda_i}(k_i)}{K\cdot k_i}-\frac{K\cdot\varepsilon^*_{\lambda_i}(k_i)k_i\cdot\varepsilon^*_{\lambda_l}(K)}{(K\cdot k_i)^2}\right]\\
&\times\left[\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[81]})\mathcal{A}^{(n,0)}_{\{[1],1,0,1\}}(r^{\natural})+\vec{Q}_{\text{eff}}(Q\bar{Q}'_{[88]})\mathcal{A}^{(n,0)}_{\{[8],1,0,1\}}(r^{\natural})\right].
\end{aligned}\tag{3.63}$$

The amplitude square in the soft limit is

$$\begin{aligned}
&\lim_{k_i\to 0}\mathbb{M}^{(n,0)}(\dot{r}) = \lim_{k_i\to 0}\mathcal{M}^{(n+1,0)}_{\{[8],1,1,J\}}(r)\\
&= g_s^2\sum_{\substack{k,l=n_I\\ k,l\neq i,k\leq l}}^{n_L^{(R)}+n_H+1}\frac{\tilde{k}_k\cdot\tilde{k}_l}{\tilde{k}_k\cdot k_i\tilde{k}_l\cdot k_i}\mathbb{M}^{(n-1,0)}_{kl}(\dot{\not{r}})\\
&\quad+g_s^2\sum_{\substack{k=n_I\\ k\neq i}}^{n_L^{(R)}+n_H+1}\left[\frac{\tilde{k}_{k,\mu}}{\tilde{k}_k\cdot k_i K\cdot k_i}-\frac{K\cdot\tilde{k}_k k_{i,\mu}}{\tilde{k}_k\cdot k_i\left(K\cdot k_i\right)^2}\right]\left[\mathbb{M}^{(n-1,0)}_{J,k[88]}(\dot{\not{r}},\dot{\not{r}}_1)^\mu+\mathbb{M}^{(n-1,0)}_{J,k[81]}(\dot{\not{r}},\dot{\not{r}}_2)^\mu\right]\\
&\quad-g_s^2\left[\frac{g_{\mu\nu}}{\left(K\cdot k_i\right)^2}+\frac{K^2k_{i,\mu}k_{i,\nu}}{\left(K\cdot k_i\right)^4}\right]\left[C_{\text{eff}}(Q\bar{Q}'_{[88]})\mathbb{M}^{(n-1,0)}_J(\dot{\not{r}}_1)^{\mu\nu}+C_{\text{eff}}(Q\bar{Q}'_{[81]})\mathbb{M}^{(n-1,0)}_J(\dot{\not{r}}_2)^{\mu\nu}\right],
\end{aligned}\tag{3.64}$$

where

$$\begin{aligned}
\dot{\not{r}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^3P_J^{[8]}],\bar{Q}',\ldots,\mathcal{I}_{n+3}\right),\\
\dot{\not{r}}_1 &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^3S_1^{[8]}],\bar{Q}',\ldots,\mathcal{I}_{n+3}\right),\\
\dot{\not{r}}_2 &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^3S_1^{[1]}],\bar{Q}',\ldots,\mathcal{I}_{n+3}\right),
\end{aligned}\tag{3.65}$$

and $\tilde{k}_{n_L^{(R)}+n_H+1}=K$ and $\tilde{k}_j=k_j$ when $j\neq j_Q,j_{\bar{Q}'}$. Once again, we have used the relations eq.(3.51).

Thus, we can easily get, for $n_I\leq k\leq n_L^{(R)}+n_H$,

$$\begin{aligned}
\mathsf{M}^{(n-1,0)}_{\text{soft},kj_Q}(\dot{\not{r}}) &= \frac{k_k\cdot K}{k_k\cdot k_i K\cdot k_i}\mathbb{M}^{(n-1,0)}_{kj_Q}(\dot{\not{r}})\\
&\quad+\left[\frac{k_{k,\mu}}{k_k\cdot k_i K\cdot k_i}-\frac{K\cdot k_k k_{i,\mu}}{k_k\cdot k_i\left(K\cdot k_i\right)^2}\right]\left[\mathbb{M}^{(n-1,0)}_{J,k[88]}(\dot{\not{r}},\dot{\not{r}}_1)^\mu+\mathbb{M}^{(n-1,0)}_{J,k[81]}(\dot{\not{r}},\dot{\not{r}}_2)^\mu\right],\\
\mathsf{M}^{(n-1,0)}_{\text{soft},j_Qj_Q}(\dot{\not{r}}) &= -\frac{K^2}{\left(K\cdot k_i\right)^2}C_A\mathbb{M}^{(n-1,0)}(\dot{\not{r}})\\
&\quad+\left[\frac{K_\mu}{\left(K\cdot k_i\right)^2}-\frac{K^2k_{i,\mu}}{\left(K\cdot k_i\right)^3}\right]\left[\mathbb{M}^{(n-1,0)}_{J,j_Q[88]}(\dot{\not{r}},\dot{\not{r}}_1)^\mu+\mathbb{M}^{(n-1,0)}_{J,j_Q[81]}(\dot{\not{r}},\dot{\not{r}}_2)^\mu\right]\\
&\quad-\left[\frac{g_{\mu\nu}}{\left(K\cdot k_i\right)^2}+\frac{K^2k_{i,\mu}k_{i,\nu}}{\left(K\cdot k_i\right)^4}\right]\left[C_{\text{eff}}(Q\bar{Q}'_{[88]})\mathbb{M}^{(n-1,0)}_J(\dot{\not{r}}_1)^{\mu\nu}\right.\\
&\quad\left.+C_{\text{eff}}(Q\bar{Q}'_{[81]})\mathbb{M}^{(n-1,0)}_J(\dot{\not{r}}_2)^{\mu\nu}\right].
\end{aligned}\tag{3.66}$$

## 4 FKS subtraction for single quarkonium production

### 4.1 FKS pairs and partition functions

With the squared real amplitudes in the soft limit at hand, we will now structure our approach using the FKS formalism. To categorise the IR divergences and facilitate their

subtraction, we introduce a set of ordered pairs for any given process $\dot{r} \in \dot{\mathcal{R}}_n$. This set is denoted as the set of FKS pairs

$$\mathcal{P}_{\rm FKS}(\dot{r}) = \Big\{(i,j) \;\Big|\; 3 \le i \le n_L^{(R)} + 2\,, n_I \le j \le n_L^{(R)} + n_H + 1\,, i \neq j\,, \\ \mathbb{M}^{(n,0)}(\dot{r}) J^{n_L^{(B)}} \to \infty \;\text{ if }\; \tilde{k}_i^0 \to 0 \;\text{ or }\; \tilde{k}_j^0 \to 0 \;\text{ or }\; \vec{\tilde{k}}_i \parallel \vec{\tilde{k}}_j \Big\}. \tag{4.1}$$

This implies that a pair of particles belongs to the set of FKS pairs if they induce soft or collinear singularities (or both) in the $n$-body real matrix elements. It is important to note that $\tilde{k}_j^0 \to 0$ is irrelevant when $j = 1, 2$. In the calculation of an NLO cross section within the FKS formalism, each pair belonging to $\mathcal{P}_{\rm FKS}$ corresponds to a set of subtractions of soft and collinear singularities.

In the FKS formalism, we multiply the real emission matrix elements with the so-called measurement/partition function $\mathcal{S}_{ij}$, so that the phase space is partitioned into different kinematic regions where each region contains at most one soft and one collinear singularity. The partitioning is accomplished through the introduction of a set of positive-definite functions

$$\mathcal{S}_{ij}(\dot{r})\,, \qquad (i,j) \in \mathcal{P}_{\rm FKS}(\dot{r})\,, \tag{4.2}$$

where the argument $\dot{r} \in \dot{\mathcal{R}}_n$ means that we can choose different $\mathcal{S}$ for different processes. As described in ref. [56], $\mathcal{S}_{ij}$ is defined in different regions as following:

$$\begin{aligned}
&\sum_{(i,j)\in\mathcal{P}_{\rm FKS}(\dot{r})} \mathcal{S}_{ij}(\dot{r}) = 1, \\
&\lim_{\vec{\tilde{k}}_i \parallel \vec{\tilde{k}}_j} \mathcal{S}_{ij}(\dot{r}) = h_{ij}\left(\frac{\tilde{k}_i^0}{\tilde{k}_i^0 + \tilde{k}_j^0}\right), \qquad \text{if } \; \tilde{m}_i = \sqrt{\tilde{k}_i^2} = \tilde{m}_j = \sqrt{\tilde{k}_j^2} = 0\,, \\
&\lim_{\tilde{k}_i^0 \to 0} \mathcal{S}_{ij}(\dot{r}) = c_{ij}\,, \text{ if } \; \mathcal{I}_i = g\,, \qquad \text{with } \; 0 < c_{ij} \le 1 \qquad \text{and} \qquad \sum_{\substack{j \\ (i,j)\in\mathcal{P}_{\rm FKS}(\dot{r})}} c_{ij} = 1\,, \\
&\lim_{\vec{\tilde{k}}_k \parallel \vec{\tilde{k}}_l} \mathcal{S}_{ij}(\dot{r}) = 0 \;\; \forall \{k,l\} \neq \{i,j\} \;\text{ with }\; (k,l) \in \mathcal{P}_{\rm FKS}(\dot{r}) \;\wedge\; \tilde{m}_k = \tilde{m}_l = 0\,, \\
&\lim_{\tilde{k}_k^0 \to 0} \mathcal{S}_{ij}(\dot{r}) = 0 \;\; \forall k \;\text{ with }\; \mathcal{I}_k = g \;\text{ and }\; \exists l \;\text{ with }\; (k,l) \in \mathcal{P}_{\rm FKS}(\dot{r}) \vee (l,k) \in \mathcal{P}_{\rm FKS}(\dot{r}).
\end{aligned} \tag{4.3}$$

In other words, $\mathcal{S}_{ij}$ goes to zero in all regions of the phase space where the real emission matrix elements diverge, except if this involves particle $i$ being soft, or particles $i$ and $j$ being collinear. The functions $h_{ij}(z)$ introduced in eq.(4.3) are defined in $0 \le z \le 1$, and have the following properties:

$$h_{ij}(z) = 1\,, \qquad \text{if } \; n_I \le j \le 2\,, \tag{4.4}$$

$$h_{ij}(z) = h(z)\,, \qquad \text{if } \; 3 \le j \le n_L^{(R)} + n_H + 1\,, \tag{4.5}$$

with $h(z)$ a positive-definite function such that

$$\lim_{z\to 0} h(z) = 1\,, \qquad \lim_{z\to 1} h(z) = 0\,, \qquad h(z) + h(1-z) = 1\,. \tag{4.6}$$

Note that in $\dot{r}$, we have glued $Q$ and $\bar{Q}'$ to a single particle $Q\bar{Q}'[n]$. This means that the number of strongly interacting massive particles (except $n$ is a colour-singlet S-wave state) is $n_H - 1$ instead of $n_H$.

The real matrix elements can be rewritten

$$\mathbb{M}^{(n,0)}(\dot{r}) = \sum_{(i,j)\in\mathcal{P}_{\rm FKS}(\dot{r})} \mathcal{S}_{ij}(\dot{r})\mathbb{M}^{(n,0)}(\dot{r}), \tag{4.7}$$

which will only be singular for a given term on the r.h.s if particle $i$ is soft and/or particles $i$ and $j$ are collinear after applying $J^{n_L^{(B)}}$.

### 4.2 Local subtraction counterterms

The IR divergence subtracted real cross sections can be formulated in the center of mass frame of the incoming partons:

$$k_1 = \tilde{k}_1 = \frac{\sqrt{s}}{2}(1,0,0,1)\,, \qquad k_2 = \tilde{k}_2 = \frac{\sqrt{s}}{2}(1,0,0,-1)\,. \tag{4.8}$$

In this frame, for each pair $(i,j) \in \mathcal{P}_{\rm FKS}(\dot{r})$, we can introduce the variables $\xi_i$ and $y_{ij}$, where

$$\tilde{k}_i^0 = \frac{\sqrt{s}}{2}\xi_i\,, \tag{4.9}$$

$$\vec{\tilde{k}}_i\cdot\vec{\tilde{k}}_j = \left|\vec{\tilde{k}}_i\right|\left|\vec{\tilde{k}}_j\right| y_{ij}\,. \tag{4.10}$$

Thus, $\xi_i$ is the rescaled energy of the FKS parton $i$, and $y_{ij}$ is the cosine of the angle between the FKS parton $i$ and its sister $j$. The soft and collinear singularities of $\mathcal{S}_{ij}(\dot{r})\mathbb{M}^{(n,0)}(\dot{r})J^{n_L^{(B)}}$ correspond to $\xi_i = 0$ and to $y_{ij} = 1$, respectively. The IR-divergence locally subtracted partonic cross section is

$$d\hat{\sigma}_{\rm FKS}(\dot{r}) = \sum_{(i,j)\in\mathcal{P}_{\rm FKS}(\dot{r})} d\hat{\sigma}_{ij}(\dot{r}), \tag{4.11}$$

where the IR-divergence locally subtracted real partonic cross section is

$$d\hat{\sigma}_{ij}(\dot{r}) = \left(\frac{1}{\xi_i}\right)_c \left(\frac{1}{1-y_{ij}}\right)_\delta \Big((1-y_{ij})\xi_i^2\mathbb{M}^{(n,0)}(\dot{r})\Big)\mathcal{S}_{ij}(\dot{r})\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r})}\mathcal{G}(\dot{r})\,d\xi_i dy_{ij} d\varphi_i d\widetilde{\phi}^{ij}_{n-1}(\dot{r})\,. \tag{4.12}$$

The variable $\varphi_i$ is the azimuthal direction of the FKS parton. The quantity $d\widetilde{\phi}^{ij}_{n-1}$ is the reduced $(n-1)$-body phase space via the following relation:

$$d\phi_n(\dot{r}) = \xi_i^{1-2\epsilon}d\xi_i(1-y_{ij}^2)^{-\epsilon}dy_{ij}d\Omega_i^{(2-2\epsilon)}d\widetilde{\phi}^{ij}_{n-1}(\dot{r})\,. \tag{4.13}$$

The reduced phase space measure has the following limits:

$$\lim_{\xi_i\to 0} d\widetilde{\phi}^{ij}_{n-1}(\dot{r}) = \frac{s^{1-\epsilon}}{(4\pi)^{3-2\epsilon}}d\phi_{n-1}(\dot{r}^{\not{i}})\,, \qquad \text{if} \;\; \tilde{m}_i = 0\,, \tag{4.14}$$

$$\lim_{y_{ij}\to 1} d\widetilde{\phi}^{ij}_{n-1}(\dot{r}) = \frac{s^{1-\epsilon}}{(4\pi)^{3-2\epsilon}}d\phi_{n-1}(\dot{r}^{j\oplus i,\not{i}})\,, \qquad \text{if} \;\; \tilde{m}_i = \tilde{m}_j = 0\,. \tag{4.15}$$

In $d = 4$ dimensions, we can simply set $\epsilon = 0$ and $d\Omega_i^{(2-2\epsilon)} = d\varphi_i$. The distributions entering eq.(4.12) are defined as follows, for any test functions $f()$ and $g()$:

$$\int_0^{\xi_{\max}} d\xi_i f(\xi_i) \left(\frac{1}{\xi_i}\right)_c = \int_0^{\xi_{\max}} d\xi_i \frac{f(\xi_i) - f(0)\Theta(\xi_{cut} - \xi_i)}{\xi_i}, \tag{4.16}$$

$$\int_{-1}^{1} dy_{ij} g(y_{ij}) \left(\frac{1}{1-y_{ij}}\right)_\delta = \int_{-1}^{1} dy_{ij} \frac{g(y_{ij}) - g(1)\Theta(y_{ij} - 1 + \delta)}{1 - y_{ij}}, \tag{4.17}$$

where

$$\xi_{\max} = 1 - \frac{1}{s}\left(\sum_{k=3}^{n+2} \tilde{m}_k\right)^2, \tag{4.18}$$

and $\Theta()$ is the Heaviside theta function. Note that $\tilde{k}_j = k_j, \tilde{m}_j = \sqrt{\tilde{k}_j^2} = \sqrt{k_j^2} = m_j$ when $j \leq n_L^{(R)} + n_H$, $\tilde{k}_{n_L^{(R)}+n_H+1} = k_{n_L^{(R)}+n_H+1} + k_{n_L^{(R)}+n_H+2} = K, \tilde{m}_{n_L^{(R)}+n_H+1} = m_Q + m_{\bar{Q}'}$, and $\tilde{k}_j = k_{j+1}, \tilde{m}_j = m_{j+1}$ when $j \geq n_L^{(R)} + n_H + 3$. In eqs.(4.16) and (4.17), $\xi_{cut}$ and $\delta$ are free parameters, that can be chosen in the ranges

$$0 < \xi_{cut} \leq \xi_{\max}, \qquad 0 < \delta \leq 2. \tag{4.19}$$

In MadFKS [56], $\delta = \delta_I$ for the initial state collinear singularities $(i,j) \in \mathcal{P}_{\rm FKS}(\dot{r}), j \leq 2$, and $\delta = \delta_O$ for the final state collinear singularities $(i,j) \in \mathcal{P}_{\rm FKS}(\dot{r}), j \geq 3$. Now, we can introduce the quantity in $d = 4$ dimensions

$$\Sigma_{ij}(\dot{r}; \xi_i, y_{ij}) = \Big((1-y_{ij})\xi_i^2 \mathbb{M}^{(n,0)}(\dot{r})\Big) \mathcal{S}_{ij}(\dot{r}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r})} \mathcal{G}(\dot{r})\, d\varphi_i d\widetilde{\phi}_{n-1}^{ij}(\dot{r}), \tag{4.20}$$

so that

$$d\hat{\sigma}_{ij}(\dot{r}) = \left(\frac{1}{\xi_i}\right)_c \left(\frac{1}{1-y_{ij}}\right)_\delta \Sigma_{ij}(\dot{r}; \xi_i, y_{ij}) d\xi_i dy_{ij}. \tag{4.21}$$

If we expand the plus distributions, we have

$$\begin{aligned} d\hat{\sigma}_{ij}(\dot{r}) = \frac{1}{\xi_i(1-y_{ij})} \Big[ & \Sigma_{ij}(\dot{r}; \xi_i, y_{ij}) - \Sigma_{ij}(\dot{r}; \xi_i, 1)\Theta(y_{ij} - 1 + \delta) \\ & - \Sigma_{ij}(\dot{r}; 0, y_{ij})\Theta(\xi_{cut} - \xi_i) + \Sigma_{ij}(\dot{r}; 0, 1)\Theta(\xi_{cut} - \xi_i)\Theta(y_{ij} - 1 + \delta)\Big] d\xi_i dy_{ij}. \end{aligned} \tag{4.22}$$

The first term in the integrand, called "events", contributes to the initial partonic real cross section, while the remaining three terms are the local subtraction counterterms, which are called "collinear counterevent", "soft counterevent", and "soft-collinear counterevent", respectively. The collinear and soft-collinear local counterterms can be inferred from the

collinear limit of the real-emission matrix elements

$$\lim_{\vec{k}_i\|\vec{k}_j}(1-y_{ij})\xi_i^2\mathbb{M}^{(n,0)}(\dot{r}) = \frac{4}{s}g_s^2\mu^{2\epsilon}\xi_i P^{<}_{\mathcal{I}_{j\oplus\bar{i}}\mathcal{I}_j}(1-\xi_i,\epsilon)\mathbb{M}^{(n-1,0)}(\dot{r}^{j\oplus\bar{i},\not{i}}) + \frac{4}{s}g_s^2\mu^{2\epsilon}\xi_i \underbrace{Q_{\mathcal{I}^\star_{j\oplus\bar{i}}\mathcal{I}_j}(1-\xi_i)\tilde{\mathbb{M}}^{(n-1,0)}_{ij}(\dot{r}^{j\oplus\bar{i},\not{i}})}_{\equiv\Delta_{ij}}, \quad j=n_I,\dots,2, \tag{4.23}$$

$$\lim_{\vec{k}_i\|\vec{k}_j}(1-y_{ij})\xi_i^2\mathbb{M}^{(n,0)}(\dot{r})\mathcal{S}_{ij}(\dot{r}) = \frac{4}{s}g_s^2\mu^{2\epsilon}\frac{1-z}{z}h(z)P^{<}_{\mathcal{I}_j\mathcal{I}_{j\oplus i}}(z,\epsilon)\mathbb{M}^{(n-1,0)}(\dot{r}^{j\oplus i,\not{i}}) + \frac{4}{s}g_s^2\mu^{2\epsilon}\frac{1-z}{z}h(z)\underbrace{Q_{\mathcal{I}_j\mathcal{I}^\star_{j\oplus i}}(z)\tilde{\mathbb{M}}^{(n-1,0)}_{ij}(\dot{r}^{j\oplus i,\not{i}})}_{\equiv\Delta_{ij}}, \quad j=3,\dots,n_L^{(R)}+2, j\neq i, \tag{4.24}$$

where $\bar{i}$ denotes the antiparticle of the particle $i$, $P^{<}_{\mathcal{I}_k\mathcal{I}_l}(z,\epsilon)$ is the unregularised Altarelli-Parisi kernel for $z<1$ in $d=4-2\epsilon$ dimensions that can be found in the literature (see, *e.g.*, eqs.(D.15-D.18) in ref. [56]), and $Q_{\mathcal{I}^\star_{j\oplus\bar{i}}\mathcal{I}_j}$ and $Q_{\mathcal{I}_j\mathcal{I}^\star_{j\oplus i}}$ are given in eqs.(D.3-D.10) of ref. [56]. The reduced matrix element is defined as [50, 56]

$$\tilde{\mathbb{M}}^{(n-1,0)}_{ij}(\dot{r}^{j\oplus i,\not{i}}) = \frac{1}{2s}\frac{1}{\omega(\mathcal{I}_1)\omega(\mathcal{I}_2)}\Re\left\{\frac{\langle ij\rangle}{[ij]}\sum_{\substack{\text{colour}\\\text{spin}}}\tilde{}\,\mathbb{A}^{(n-1,0)}_{+}(\dot{r}^{j\oplus i,\not{i}})\mathbb{A}^{(n-1,0)}_{-}(\dot{r}^{j\oplus i,\not{i}})^\star\right\}, \tag{4.25}$$

where the spinor-helicity formalism takes the conventions of ref. [71], $\mathbb{A}^{(n-1,0)}_{\pm}$ represents the helicity amplitude with the helicity of the parton $j\oplus i$ being $\pm$, and the sum with a tilde has summed over the colour and spin of the external states except the spin of the particle $j\oplus i$. The $\Delta_{ij}$ term vanishes upon the integration of the azimuthal variable $d\varphi_i$ of the parton $i$. In eq.(4.24), we have assumed $\tilde{k}_i=(1-z)\left(\tilde{k}_i+\tilde{k}_j\right)$ and $\tilde{k}_j=z\left(\tilde{k}_i+\tilde{k}_j\right)$. This essentially gives us for the initial-state collinear singularities ($n_I\leq j\leq 2$)

$$\Sigma_{ij}(\dot{r};\xi_i,1) = \frac{4}{(4\pi)^3}g_s^2\xi_i\left[P^{<}_{\mathcal{I}_{j\oplus\bar{i}}\mathcal{I}_j}(1-\xi_i,0)\mathbb{M}^{(n-1,0)}(\dot{r}^{j\oplus\bar{i},\not{i}})+\Delta_{ij}\right] \times\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r})}\mathcal{G}(\dot{r})\,d\varphi_i d\phi_{n-1}(\dot{r}^{j\oplus\bar{i},\not{i}}), \tag{4.26}$$

and, for the final state collinear singularities ($3\leq j\leq n_L^{(R)}+2, j\neq i$), we have

$$\Sigma_{ij}(\dot{r};\xi_i,1) = \frac{4}{(4\pi)^3}g_s^2\frac{1-z}{z}h(z)\left[P^{<}_{\mathcal{I}_j\mathcal{I}_{j\oplus i}}(z,0)\mathbb{M}^{(n-1,0)}(\dot{r}^{j\oplus i,\not{i}})+\Delta_{ij}\right] \times\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r})}\mathcal{G}(\dot{r})\,d\varphi_i d\phi_{n-1}(\dot{r}^{j\oplus i,\not{i}})\,. \tag{4.27}$$

For the soft-collinear counterparts, we need to take $\xi_i\to 0$ and $z\to 1$ on the r.h.s of the above two equations. The soft local counterterm is

$$\Sigma_{ij}(\dot{r};0,y_{ij}) = \frac{s}{(4\pi)^3}(1-y_{ij})\Big(\lim_{\xi_i\to 0}\xi_i^2\mathbb{M}^{(n,0)}(\dot{r})\Big)c_{ij}\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r})}\mathcal{G}(\dot{r})\,d\varphi_i d\phi_{n-1}(\dot{r}^{\not{i}}), \tag{4.28}$$

where the soft limit of the real matrix element has been derived in sect. 3.2.

### 4.3 Integrated soft counterterm

We can split the integrated soft counterterms into two parts:

$$d\hat{\sigma}^{(S)}(\dot{r}) = d\hat{\sigma}^{(S,1)}(\dot{r}) + d\hat{\sigma}^{(S,2)}(\dot{r}). \tag{4.29}$$

The first one is simply stemming from the $(4-2\epsilon)$-dimensional counterpart of the function $\Sigma_{ij}(\dot{r};0,y_{ij})$ in the local subtraction counterterms appearing in the former section. It is

$$\begin{aligned}
&d\hat{\sigma}^{(S,1)}(\dot{r}) \\
&= d\phi_{n-1}(\dot{r}^{\not i}) \left(\frac{\mu^2}{s}\right)^{\epsilon} \frac{s}{(4\pi)^{3-2\epsilon}} \sum_{(i,j)\in\mathcal{P}_{\rm FKS}(\dot{r})} c_{ij} \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r})} \mathcal{G}(\dot{r}) \\
&\quad \times \int \xi_i^{-1-2\epsilon}(1-y_{ij}^2)^{-\epsilon} \left(\lim_{\xi_i\to 0} \xi_i^2 \mathbb{M}^{(n,0)}(\dot{r})\right) \Theta(\xi_{cut}-\xi_i) d\xi_i dy_{ij} d\Omega_i^{(2-2\epsilon)} \\
&= \frac{d\phi_{n-1}(\dot{r}^{\not i})}{8\pi^2} \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r})} \mathcal{G}(\dot{r}) \sum_{\substack{i=3 \\ \mathcal{I}_i=g}}^{n_L^{(R)}+2} \left[ -\frac{\xi_{cut}^{-2\epsilon}}{2\epsilon} \frac{2^{2\epsilon}}{(2\pi)^{1-2\epsilon}} \left(\frac{s}{\mu^2}\right)^{-\epsilon} \int \left(\lim_{\xi_i\to 0} \left(\tilde{k}_i^0\right)^2 \mathbb{M}^{(n,0)}(\dot{r})\right) d\Omega_i \right] \\
&= \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}^{\not i}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r})} \mathcal{G}(\dot{r}) \\
&\quad \times \sum_{\substack{i=3 \\ \mathcal{I}_i=g}}^{n_L^{(R)}+2} \sum_{\substack{k,l=n_I \\ k,l\neq i, k\leq l}}^{n_L^{(R)}+n_H+1} \left[ -\frac{\xi_{cut}^{-2\epsilon}}{2\epsilon} \frac{2^{2\epsilon}}{(2\pi)^{1-2\epsilon}} \left(\frac{s}{\mu^2}\right)^{-\epsilon} \int \left(\lim_{\xi_i\to 0} \left(\tilde{k}_i^0\right)^2 \mathsf{M}_{{\rm soft},kl}^{(n-1,0)}(\dot{r}^{\not i})\right) d\Omega_i \right] \\
&= \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}^{\not i}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not i})} \mathcal{G}(\dot{r}^{\not i}) \\
&\quad \times \sum_{k=n_I}^{n_L^{(B)}+n_H+1} \sum_{l=k}^{n_L^{(B)}+n_H+1} \left[ -\frac{\xi_{cut}^{-2\epsilon}}{2\epsilon} \frac{2^{2\epsilon}}{(2\pi)^{1-2\epsilon}} \left(\frac{s}{\mu^2}\right)^{-\epsilon} \int \left(\lim_{\xi_i\to 0} \left(\tilde{k}_i^0\right)^2 \mathsf{M}_{{\rm soft},kl}^{(n-1,0)}(\dot{r}^{\not i})\right) d\Omega_i \right],
\end{aligned} \tag{4.30}$$

where the $d-1 = 3-2\epsilon$ dimensional solid angle measure is

$$d\Omega_i = (1-y_{ij}^2)^{-\epsilon} dy_{ij} d\Omega_i^{(2-2\epsilon)}, \tag{4.31}$$

and we have used eq.(3.36) as well as the fact that $\lim_{\xi_i\to 0} \xi_i^2 \mathbb{M}^{(n,0)}(\dot{r})$ is independent of $\xi_i$. In the last equation, we have removed any final state gluon $\mathcal{I}_i = g$ and have changed the final state symmetry factor accordingly.

The second one is from the soft singularities of the LDMEs in NRQCD, which we will give explicitly in the following.

### 4.3.1 Soft counterterms of LDMEs

The second term originates from the renormalisation of LDMEs and is analogous to the initial state collinear counterterm at hadron colliders or the final state collinear counterterm due to the presence of fragmentation functions [2]. The soft counterterms of the LDMEs can be incorporated into either the real or the virtual matrix elements, depending on the preference. In this paper, we include them as part of the integrated soft subtraction terms for the real matrix element. We have the following perturbative corrections for the LDMEs [72] in the $\overline{\text{MS}}$ scheme, by defining $1/\bar{\epsilon} = 1/\epsilon + \log(4\pi) - \gamma_E$,

$$\begin{aligned}
\langle \mathcal{O}^H_{{}^3_{S_1^{[8]}}} \rangle(\mu) &= \langle \mathcal{O}^H_{{}^3_{S_1^{[8]}}} \rangle + \frac{4\alpha_s}{3\pi m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log \frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right) \left[ \frac{C_F}{2N_c} \sum_{J=0}^{2} \langle \mathcal{O}^H_{{}^3_{P_J^{[1]}}} \rangle + B_F \sum_{J=0}^{2} \langle \mathcal{O}^H_{{}^3_{P_J^{[8]}}} \rangle \right], \\
\langle \mathcal{O}^H_{{}^3_{S_1^{[1]}}} \rangle(\mu) &= \langle \mathcal{O}^H_{{}^3_{S_1^{[1]}}} \rangle + \frac{4\alpha_s}{3\pi m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log \frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right) \left[ \sum_{J=0}^{2} \langle \mathcal{O}^H_{{}^3_{P_J^{[8]}}} \rangle \right], \\
\langle \mathcal{O}^H_{{}^1_{S_0^{[8]}}} \rangle(\mu) &= \langle \mathcal{O}^H_{{}^1_{S_0^{[8]}}} \rangle + \frac{4\alpha_s}{3\pi m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log \frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right) \left[ \frac{C_F}{2N_c} \langle \mathcal{O}^H_{{}^1_{P_1^{[1]}}} \rangle + B_F \langle \mathcal{O}^H_{{}^1_{P_1^{[8]}}} \rangle \right], \\
\langle \mathcal{O}^H_{{}^1_{S_0^{[1]}}} \rangle(\mu) &= \langle \mathcal{O}^H_{{}^1_{S_0^{[1]}}} \rangle + \frac{4\alpha_s}{3\pi m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log \frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right) \left[ \langle \mathcal{O}^H_{{}^1_{P_1^{[8]}}} \rangle \right],
\end{aligned} \tag{4.32}$$

where $\gamma_E$ is the Euler-Mascheroni constant, $\mu_{\text{NRQCD}}$ is the NRQCD cutoff scale and $B_F = (N_c^2 - 4)/(4N_c) = 5/12$. If we use the Binoth Les Houches Accord convention by taking out the global prefactor $(4\pi)^\epsilon/\Gamma(1-\epsilon)$, $1/\bar{\epsilon}$ simply becomes $1/\epsilon$ in eq.(4.32). Therefore, we have the additional integrated soft counterterms for P-waves:

- $Q\bar{Q}'[{}^1P_1^{[1]}]$:

$$\begin{aligned}
d\hat{\sigma}^{(S,2)}(\dot{r}) = \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}_1) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}_1)} \mathcal{G}(\dot{r}_1) \left[ \frac{1}{2N_c} \mathbb{M}^{(n-1,0)}(\dot{r}_1) \right] \\
\times \frac{8}{m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log \frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right),
\end{aligned} \tag{4.33}$$

  where

$$\begin{aligned}
\dot{r}_1 &= \left( \mathcal{I}_1, \ldots, \mathcal{I}_i, \ldots, \mathcal{I}_j, \ldots, \mathcal{I}_{n_L^{(R)}+n_H}, Q\bar{Q}'[{}^1P_1^{[1]}], \bar{Q}', \ldots, \mathcal{I}_{n+3} \right), \\
\dot{r}_1 &= \left( \mathcal{I}_1, \ldots, \mathcal{I}_i, \ldots, \mathcal{I}_j, \ldots, \mathcal{I}_{n_L^{(R)}+n_H}, Q\bar{Q}'[{}^1S_0^{[8]}], \bar{Q}', \ldots, \mathcal{I}_{n+3} \right),
\end{aligned} \tag{4.34}$$

  and $\mathcal{I}_i = g$.

- $Q\bar{Q}'[{}^1P_1^{[8]}]$:

$$\begin{aligned}
d\hat{\sigma}^{(S,2)}(\dot{r}) = \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}_1) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}_1)} \mathcal{G}(\dot{r}_1) \left[ B_F \mathbb{M}^{(n-1,0)}(\dot{r}_1) + C_F \mathbb{M}^{(n-1,0)}(\dot{r}_2) \right] \\
\times \frac{8}{m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log \frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right),
\end{aligned} \tag{4.35}$$

where

$$\begin{aligned}
\dot{r}^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^1P_1^{[8]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\\
\dot{r}_1^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^1S_0^{[8]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\\
\dot{r}_2^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^1S_0^{[1]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right).
\end{aligned} \tag{4.36}$$

- $Q\bar{Q}'[{}^3P_J^{[1]}]$:

$$\begin{aligned}
d\hat{\sigma}^{(S,2)}(\dot{r}) = \frac{\alpha_s}{2\pi}d\phi_{n-1}(\dot{r}^{\not{k}})\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}})\left[\frac{2J+1}{2N_c}\mathbb{M}^{(n-1,0)}(\dot{r}_1^{\not{k}})\right]\\
\times\frac{8}{9m_Qm_{\bar{Q}'}}\left(\frac{1}{\bar{\epsilon}}+\log\frac{\mu^2}{\mu^2_{\rm NRQCD}}\right),
\end{aligned} \tag{4.37}$$

where

$$\begin{aligned}
\dot{r}^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^3P_J^{[1]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\\
\dot{r}_1^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^3S_1^{[8]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right).
\end{aligned} \tag{4.38}$$

- $Q\bar{Q}'[{}^3P_J^{[8]}]$:

$$\begin{aligned}
d\hat{\sigma}^{(S,2)}(\dot{r}) = \frac{\alpha_s}{2\pi}d\phi_{n-1}(\dot{r}^{\not{k}})\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}})\left[B_F\mathbb{M}^{(n-1,0)}(\dot{r}_1^{\not{k}})+C_F\mathbb{M}^{(n-1,0)}(\dot{r}_2^{\not{k}})\right]\\
\times\frac{8(2J+1)}{9m_Qm_{\bar{Q}'}}\left(\frac{1}{\bar{\epsilon}}+\log\frac{\mu^2}{\mu^2_{\rm NRQCD}}\right),
\end{aligned} \tag{4.39}$$

where

$$\begin{aligned}
\dot{r}^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^3P_J^{[8]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\\
\dot{r}_1^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^3S_1^{[8]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\\
\dot{r}_2^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,\mathcal{I}_{n_L^{(R)}+n_H},Q\bar{Q}'[{}^3S_1^{[1]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right).
\end{aligned} \tag{4.40}$$

It is clear that for S-waves $d\hat{\sigma}^{(S,2)}(\dot{r})=0$.

### 4.3.2 Colour-singlet S-wave state

The soft integrated partonic cross section for a colour-singlet S-wave quarkonium production is

$$d\hat{\sigma}^{(S)}(\dot{r}) = \frac{\alpha_s}{2\pi}d\phi_{n-1}(\dot{r}^{\not{k}})\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}})\sum_{k=n_I}^{n_L^{(B)}+n_H}\sum_{l=k}^{n_L^{(B)}+n_H}\bar{\mathcal{E}}(\{1,1\},\{k_k,k_l\})\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}^{\not{k}}), \tag{4.41}$$

where the eikonal integrals $\bar{\mathcal{E}}()$ can be found in app. A, and

$$\dot{r}^{\not{k}} = \left(\mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, Q\bar{Q}'[{}^{2S+1}S_J^{[1]}], \bar{\mathcal{Q}}', \ldots, \mathcal{I}_{n+3}\right) \tag{4.42}$$

with $\mathcal{I}_i$ being a final state gluon. Now, the momenta $k_k, k_l$ are defined in the reduced Born process $\dot{r}^{\not{k}} \in \dot{\mathcal{R}}_{n-1}$ with $k_{n_L^{(B)}+n_H+1} = K$ (the quarkonium momentum).

### 4.3.3 Colour-octet S-wave state

The integrated soft counterterm for a colour-octet S-wave quarkonium is

$$d\hat{\sigma}^{(S)}(\dot{r}) = \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}^{\not{k}}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})} \mathcal{G}(\dot{r}^{\not{k}}) \sum_{k=n_I}^{n_L^{(B)}+n_H+1} \sum_{l=k}^{n_L^{(B)}+n_H+1} \bar{\mathcal{E}}(\{1,1\},\{k_k,k_l\}) \mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\not{k}}), \tag{4.43}$$

where the eikonal integrals can be found in app. A, and

$$\dot{r}^{\not{k}} = \left(\mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, Q\bar{Q}'[{}^{2S+1}S_J^{[8]}], \bar{\mathcal{Q}}', \ldots, \mathcal{I}_{n+3}\right) \tag{4.44}$$

with $\mathcal{I}_i$ being a gluon. The momenta $k_k, k_l$ are defined in the Born process $\dot{r}^{\not{k}} \in \dot{\mathcal{R}}_{n-1}$ with $k_{n_L^{(B)}+n_H+1} = K$ (the quarkonium momentum).

### 4.3.4 Colour-singlet spin-singlet P-wave state

For a colour-singlet spin-singlet P-wave state, the integrated soft counterterm is

$$\begin{aligned} d\hat{\sigma}^{(S)}(\dot{r}) = \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}^{\not{k}}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})} \mathcal{G}(\dot{r}^{\not{k}}) \Bigg[ & \sum_{k=n_I}^{n_L^{(B)}+n_H} \sum_{l=k}^{n_L^{(B)}+n_H} \bar{\mathcal{E}}(\{1,1\},\{k_k,k_l\}) \mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\not{k}}) \\ & + \sum_{k=n_I}^{n_L^{(B)}+n_H} \left( \bar{\mathcal{E}}(\{1,1\},\{k_k,K\}) \frac{k_{k,\mu}}{K\cdot k_k} - \bar{\mathcal{E}}_\mu(\{1,2\},\{k_k,K\}) \right) \mathbb{M}_{k[18]}^{(n-1,0)}(\dot{r}^{\not{k}}, \dot{r}_1^{\not{k}})^\mu \\ & + \left( \frac{1}{2N_c} \frac{8}{m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log \frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right) - \frac{2\epsilon-2}{K^2} \bar{\mathcal{E}}(\{1,1\},\{K,K\}) C_{\text{eff}}(Q\bar{Q}'_{[18]}) \right) \\ & \times \mathbb{M}^{(n-1,0)}(\dot{r}_1^{\not{k}}) \Bigg], \end{aligned} \tag{4.45}$$

where the eikonal integrals can be found in app. A, and

$$\begin{aligned} \dot{r}^{\not{k}} &= \left(\mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, Q\bar{Q}'[{}^1P_1^{[1]}], \bar{\mathcal{Q}}', \ldots, \mathcal{I}_{n+3}\right), \\ \dot{r}_1^{\not{k}} &= \left(\mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, Q\bar{Q}'[{}^1S_0^{[8]}], \bar{\mathcal{Q}}', \ldots, \mathcal{I}_{n+3}\right), \end{aligned} \tag{4.46}$$

with $\mathcal{I}_i = g$. Once again, we remind the readers that the momenta $k_k, k_l$ are defined in the reduced Born process $\dot{r}^{\not{k}} \in \dot{\mathcal{R}}_{n-1}$ with the quarkonium momentum $k_{n_L^{(B)}+n_H+1} = K$. Now, besides the usual scalar eikonal integrals, we also need to introduce the rank-1 tensor integrals $\bar{\mathcal{E}}_\mu(\{1,2\},\{k_k,K\})$.

### 4.3.5 Colour-octet spin-singlet P-wave state

The soft integrated counterterm for a colour-octet spin-singlet P-wave quarkonium is

$$
\begin{aligned}
d\hat{\sigma}^{(S)}(\dot{r}) = \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}^{\not{1}}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{1}})} \mathcal{G}(\dot{r}^{\not{1}}) \Bigg[ & \sum_{k=n_I}^{n_L^{(B)}+n_H+1} \sum_{l=k}^{n_L^{(B)}+n_H+1} \bar{\mathcal{E}}(\{1,1\},\{k_k,k_l\}) \mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\not{1}}) \\
& + \sum_{k=n_I}^{n_L^{(B)}+n_H+1} \left( \bar{\mathcal{E}}(\{1,1\},\{k_k,K\}) \frac{k_{k,\mu}}{K\cdot k_k} - \bar{\mathcal{E}}_\mu(\{1,2\},\{k_k,K\}) \right) \\
& \times \left( \mathbb{M}_{k[88]}^{(n-1,0)}(\dot{r}^{\not{1}},\dot{r}_1^{\not{1}})^\mu + \mathbb{M}_{k[81]}^{(n-1,0)}(\dot{r}^{\not{1}},\dot{r}_2^{\not{1}})^\mu \right) \\
& + \left( B_F \frac{8}{m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log\frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right) - \frac{2\epsilon-2}{K^2} \bar{\mathcal{E}}(\{1,1\},\{K,K\}) C_{\text{eff}}(Q\bar{Q}'_{[88]}) \right) \\
& \times \mathbb{M}^{(n-1,0)}(\dot{r}_1^{\not{1}}) \\
& + \left( C_F \frac{8}{m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log\frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right) - \frac{2\epsilon-2}{K^2} \bar{\mathcal{E}}(\{1,1\},\{K,K\}) C_{\text{eff}}(Q\bar{Q}'_{[81]}) \right) \\
& \times \mathbb{M}^{(n-1,0)}(\dot{r}_2^{\not{1}}) \Bigg],
\end{aligned}
\tag{4.47}
$$

where the expressions of the eikonal integrals can be found in app. A, and

$$
\begin{aligned}
\dot{r}^{\not{1}} &= \left( \mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, Q\bar{Q}'[{}^1P_1^{[8]}], \bar{Q}', \ldots, \mathcal{I}_{n+3} \right), \\
\dot{r}_1^{\not{1}} &= \left( \mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, Q\bar{Q}'[{}^1S_0^{[8]}], \bar{Q}', \ldots, \mathcal{I}_{n+3} \right), \\
\dot{r}_2^{\not{1}} &= \left( \mathcal{I}_1, \ldots, \not{\mathcal{I}}_i, \ldots, \mathcal{I}_j, \ldots, Q\bar{Q}'[{}^1S_0^{[1]}], \bar{Q}', \ldots, \mathcal{I}_{n+3} \right),
\end{aligned}
\tag{4.48}
$$

and $\mathcal{I}_i$ is a final state gluon.

### 4.3.6 Colour-singlet spin-triplet P-wave state

The integrated soft partonic cross section of a process with a colour-singlet spin-triplet P-wave quarkonium is

$$
\begin{aligned}
d\hat{\sigma}^{(S)}(\dot{r}) = \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}^{\not{1}}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{1}})} \mathcal{G}(\dot{r}^{\not{1}}) \Bigg[ & \sum_{k=n_I}^{n_L^{(B)}+n_H} \sum_{l=k}^{n_L^{(B)}+n_H} \bar{\mathcal{E}}(\{1,1\},\{k_k,k_l\}) \mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\not{1}}) \\
& + \sum_{k=n_I}^{n_L^{(B)}+n_H} \left( \bar{\mathcal{E}}(\{1,1\},\{k_k,K\}) \frac{k_{k,\mu}}{K\cdot k_k} - \bar{\mathcal{E}}_\mu(\{1,2\},\{k_k,K\}) \right) \mathbb{M}_{J,k[18]}^{(n-1,0)}(\dot{r}^{\not{1}},\dot{r}_1^{\not{1}})^\mu \\
& - \left( \frac{g_{\mu\nu}}{K^2} \bar{\mathcal{E}}(\{1,1\},\{K,K\}) + \bar{\mathcal{E}}_{\mu\nu}(\{2,2\},\{K,K\}) \right) C_{\text{eff}}(Q\bar{Q}'_{[18]}) \mathbb{M}_J^{(n-1,0)}(\dot{r}_1^{\not{1}})^{\mu\nu} \\
& + \frac{2J+1}{2N_c} \frac{8}{9 m_Q m_{\bar{Q}'}} \left( \frac{1}{\bar{\epsilon}} + \log\frac{\mu^2}{\mu^2_{\text{NRQCD}}} \right) \mathbb{M}^{(n-1,0)}(\dot{r}_1^{\not{1}}) \Bigg],
\end{aligned}
\tag{4.49}
$$

where the eikonal integrals can be found in app. A, and

$$\begin{aligned}\dot{r}^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,Q\bar{Q}'[^3P_J^{[1]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\\ \dot{r}_1^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,Q\bar{Q}'[^3S_1^{[8]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\end{aligned} \tag{4.50}$$

and $\mathcal{I}_i$ is a final state gluon. In this case, we also need the rank-2 eikonal tensor integrals $\bar{\mathcal{E}}_{\mu\nu}(\{2,2\},\{K,K\})$.

### 4.3.7 Colour-octet spin-triplet P-wave state

Finally, the colour-octet spin-triplet P-wave quarkonium production has the following soft integrated counterterm

$$\begin{aligned}d\hat{\sigma}^{(S)}(\dot{r}) = \frac{\alpha_s}{2\pi}d\phi_{n-1}(\dot{r}^{\not{k}})\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}})\Bigg[&\sum_{k=n_I}^{n_L^{(B)}+n_H+1}\sum_{l=k}^{n_L^{(B)}+n_H+1}\bar{\mathcal{E}}(\{1,1\},\{k_k,k_l\})\mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\not{k}})\\ &+\sum_{k=n_I}^{n_L^{(B)}+n_H+1}\left(\bar{\mathcal{E}}(\{1,1\},\{k_k,K\})\frac{k_{k,\mu}}{K\cdot k_k}-\bar{\mathcal{E}}_\mu(\{1,2\},\{k_k,K\})\right)\\ &\times\left(\mathbb{M}_{J,k[88]}^{(n-1,0)}(\dot{r}^{\not{k}},\dot{r}_1^{\not{k}})^\mu+\mathbb{M}_{J,k[81]}^{(n-1,0)}(\dot{r}^{\not{k}},\dot{r}_2^{\not{k}})^\mu\right)\\ &-\left(\frac{g_{\mu\nu}}{K^2}\bar{\mathcal{E}}(\{1,1\},\{K,K\})+\bar{\mathcal{E}}_{\mu\nu}(\{2,2\},\{K,K\})\right)\\ &\times\left(C_{\text{eff}}(Q\bar{Q}'_{[88]})\mathbb{M}_J^{(n-1,0)}(\dot{r}_1^{\not{k}})^{\mu\nu}+C_{\text{eff}}(Q\bar{Q}'_{[81]})\mathbb{M}_J^{(n-1,0)}(\dot{r}_2^{\not{k}})^{\mu\nu}\right)\\ &+\frac{8(2J+1)}{9m_Q m_{\bar{Q}'}}\left(\frac{1}{\bar{\epsilon}}+\log\frac{\mu^2}{\mu^2_{\text{NRQCD}}}\right)\left(B_F\mathbb{M}^{(n-1,0)}(\dot{r}_1^{\not{k}})+C_F\mathbb{M}^{(n-1,0)}(\dot{r}_2^{\not{k}})\right)\Bigg],\end{aligned} \tag{4.51}$$

where the eikonal integrals can be found in app. A, and

$$\begin{aligned}\dot{r}^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,Q\bar{Q}'[^3P_J^{[8]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\\ \dot{r}_1^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,Q\bar{Q}'[^3S_1^{[8]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\\ \dot{r}_2^{\not{k}} &= \left(\mathcal{I}_1,\ldots,\not{\mathcal{I}}_i,\ldots,\mathcal{I}_j,\ldots,Q\bar{Q}'[^3S_1^{[1]}],\bar{\not{Q}}',\ldots,\mathcal{I}_{n+3}\right),\end{aligned} \tag{4.52}$$

with $\mathcal{I}_i$ being a final state gluon.

## 4.4 Integrated collinear and soft-collinear counterterms

As explained earlier, the integrated collinear and soft-collinear counterterms should be identical to the original case since quarkonia are massive and do not exhibit any collinear divergences. However, for the completeness, we still discuss them here and adjust to our notations with the risk that they are actually well understood.

Following the explanation of ref. [50], it is straightforward to show that the sum of the integrated counterterms for collinear and soft-collinear emissions from an incoming initial

state parton takes the form

$$d\hat{\sigma}^{(in)}(\dot{r}) = -\frac{\alpha_s}{2\pi}\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}})\sum_{k=n_I}^{2}\int_0^{\xi_{\max}} d\xi_i \left(\frac{1}{\bar{\epsilon}} - \log\left(\frac{s\delta_I}{2\mu^2}\right)\right)\left[\left(\frac{1}{\xi_i}\right)_c - 2\epsilon\left(\frac{\log(\xi_i)}{\xi_i}\right)_c\right]$$
$$\times\xi_i P^{<}_{\mathcal{I}_{k\oplus\bar{i}}\mathcal{I}_k}(1-\xi_i,\epsilon)\mathbb{M}^{(n-1,0)}(\dot{r}^{k\oplus\bar{i},\not{k}})d\phi_{n-1}(\dot{r}^{k\oplus\bar{i},\not{k}})\,, \tag{4.53}$$

and that for emissions from an outgoing final state parton for a given underlying Born process $\dot{r}^{\not{k}}$ the counterterms can be written as

$$d\hat{\sigma}^{(out)}(\dot{r}) = \frac{\alpha_s}{2\pi}\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}})\sum_{k=3}^{n_L^{(B)}+2}\left\{\frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q^2_{\rm ES}}\right)^\epsilon\frac{1}{\epsilon}\left[\gamma(\mathcal{I}_k) + C(\mathcal{I}_k)\log\left(\frac{\xi_{cut}^2 s}{4E_k^2}\right)\right]\right.$$
$$+\left[\gamma'(\mathcal{I}_k) - \log\left(\frac{s\delta_O}{2Q^2_{\rm ES}}\right)\left(\gamma(\mathcal{I}_k) - 2C(\mathcal{I}_k)\log\left(\frac{2E_k}{\xi_{cut}\sqrt{s}}\right)\right)\right.$$
$$\left.\left.+2C(\mathcal{I}_k)\left(\log^2\left(\frac{2E_k}{\sqrt{s}}\right) - \log^2(\xi_{cut})\right) - 2\gamma(\mathcal{I}_k)\log\left(\frac{2E_k}{\sqrt{s}}\right)\right]\right\}$$
$$\times\mathbb{M}^{(n-1,0)}(\dot{r}^{\not{k}})d\phi_{n-1}(\dot{r}^{\not{k}})\,, \tag{4.54}$$

where $Q_{\rm ES}$ is the Ellis-Sexton scale [73], the Casimir factors are

$$C(\mathcal{I}) = \begin{cases} C_F, & \text{if } \mathcal{I}\in\mathbf{3},\overline{\mathbf{3}} \\ C_A, & \text{if } \mathcal{I}\in\mathbf{8} \end{cases}. \tag{4.55}$$

The collinear anomalous dimensions are

$$\gamma(\mathcal{I}) = \begin{cases} \frac{3}{2}C_F\,, & \text{if } \mathcal{I} = q,\bar{q} \\ \frac{11}{6}C_A - \frac{2}{3}T_F n_f\,, & \text{if } \mathcal{I} = g \end{cases}, \tag{4.56}$$

with $T_F = \frac{1}{2}$ and $n_f$ being the number of massless quark flavours, and

$$\gamma'(\mathcal{I}) = \begin{cases} \left(\frac{13}{2} - \frac{2\pi^2}{3}\right)C_F\,, & \text{if } \mathcal{I} = q,\bar{q} \\ \left(\frac{67}{9} - \frac{2\pi^2}{3}\right)C_A - \frac{23}{9}T_F n_f\,, & \text{if } \mathcal{I} = g \end{cases}. \tag{4.57}$$

These integrated counterterms have to be combined with the counterterms stemming from the collinear renormalisation of PDFs which read

$$d\hat{\sigma}^{(cnt)}(\dot{r}) = \frac{\alpha_s}{2\pi}\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}})\sum_{k=n_I}^{2}\int_{1-\xi_{\max}}^{1} dz\left(\frac{1}{\bar{\epsilon}}P_{\mathcal{I}_{k\oplus\bar{i}}\mathcal{I}_k}(z,0) - K_{\mathcal{I}_{k\oplus\bar{i}}\mathcal{I}_k}(z)\right)$$
$$\times\mathbb{M}^{(n-1,0)}(\dot{r}^{k\oplus\bar{i},\not{k}})d\phi_{n-1}(\dot{r}^{k\oplus\bar{i},\not{k}}) \tag{4.58}$$

with the four dimensional regularised Altarelli-Parisi splitting kernels

$$P_{ab}(z,0) = \frac{(1-z)P^{<}_{ab}(z,0)}{(1-z)_+} + \gamma(a)\delta_{ab}\,\delta(1-z) \tag{4.59}$$

and the subtraction scheme dependent function $K_{ab}(z)$ [6]. A change of the integration variable, $z \to 1-\xi_i$, results in [7]

$$
\begin{aligned}
d\hat{\sigma}^{(cnt)}(\dot{r}) = \frac{\alpha_s}{2\pi}\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}}) \sum_{k=n_I}^{2} \Bigg\{ & \frac{1}{\bar{\epsilon}}\left[\gamma(\mathcal{I}_k) + C(\mathcal{I}_k)\log\left(\frac{\xi_{cut}^2 s}{4E_k^2}\right)\right] \mathbb{M}^{(n-1,0)}(\dot{r}^{\not{k}}) d\phi_{n-1}(\dot{r}^{\not{k}}) \\
& + \int_0^{\xi_{\max}} d\xi_i \left[\frac{1}{\bar{\epsilon}}\left(\frac{1}{\xi_i}\right)_c \xi_i P^{<}_{\mathcal{I}_{k\oplus\bar{i}}\mathcal{I}_k}(1-\xi_i,0) - K_{\mathcal{I}_{k\oplus\bar{i}}\mathcal{I}_k}(1-\xi_i)\right] \mathbb{M}^{(n-1,0)}(\dot{r}^{k\oplus\bar{i},\not{k}}) \\
& \times d\phi_{n-1}(\dot{r}^{k\oplus\bar{i},\not{k}})\Bigg\}\,,
\end{aligned}
\tag{4.60}
$$

where we made use of the identity

$$
\delta(\xi_i)\xi_i P^{<}_{ab}(1-\xi_i,0) = 2C(a)\delta_{ab}\,\delta(\xi_i)\,. \tag{4.61}
$$

If we sum up eqs.(4.53), (4.54), and (4.60), we obtain the sum of the integrated collinear and soft-collinear counterterms

$$
\begin{aligned}
d\hat{\sigma}^{(C)}(\dot{r}) &= d\hat{\sigma}^{(in)}(\dot{r}) + d\hat{\sigma}^{(out)}(\dot{r}) + d\hat{\sigma}^{(cnt)}(\dot{r}) \\
&= d\hat{\sigma}^{(C)}_{\text{sing.}}(\dot{r}) + d\hat{\sigma}^{(C,n-1)}_{\text{FIN}}(\dot{r}) + d\hat{\sigma}^{(C,n)}_{\text{FIN}}(\dot{r})\,,
\end{aligned}
\tag{4.62}
$$

which we have decomposed into a singular term

$$
\begin{aligned}
d\hat{\sigma}^{(C)}_{\text{sing.}}(\dot{r}) = \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}^{\not{k}}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}}) & \sum_{k=n_I}^{n_L^{(B)}+2} \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q^2_{\text{ES}}}\right)^\epsilon \frac{1}{\epsilon} \\
& \times \left[\gamma(\mathcal{I}_k) + C(\mathcal{I}_k)\log\left(\frac{\xi_{cut}^2 s}{4E_k^2}\right)\right] \mathbb{M}^{(n-1,0)}(\dot{r}^{\not{k}})\,,
\end{aligned}
\tag{4.63}
$$

a finite $(n-1)$-body contribution (equivalent to eqs.(4.5-4.6) in ref. [56])

$$
\begin{aligned}
d\hat{\sigma}^{(C,n-1)}_{\text{FIN}}(\dot{r}) = \frac{\alpha_s}{2\pi} d\phi_{n-1}(\dot{r}^{\not{k}}) \frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{k}})}\mathcal{G}(\dot{r}^{\not{k}}) \Bigg\{ & -\log\left(\frac{\mu^2}{Q^2_{\text{ES}}}\right) \sum_{k=n_I}^{2} \Big(\gamma(\mathcal{I}_k) + 2C(\mathcal{I}_k)\log(\xi_{cut})\Big) \\
& + \sum_{k=3}^{n_L^{(B)}+2} \Bigg[\gamma'(\mathcal{I}_k) - \log\left(\frac{s\delta_O}{2Q^2_{\text{ES}}}\right)\left(\gamma(\mathcal{I}_k) - 2C(\mathcal{I}_k)\log\left(\frac{2E_k}{\xi_{cut}\sqrt{s}}\right)\right) \\
& + 2C(\mathcal{I}_k)\left(\log^2\left(\frac{2E_k}{\sqrt{s}}\right) - \log^2(\xi_{cut})\right) - 2\gamma(\mathcal{I}_k)\log\left(\frac{2E_k}{\sqrt{s}}\right)\Bigg]\Bigg\} \mathbb{M}^{(n-1,0)}(\dot{r}^{\not{k}})\,,
\end{aligned}
\tag{4.64}
$$

and a finite contribution with a degenerated $n$-body phase space (equivalent to eqs.(4.40-4.42) in ref. [56]), where the angular dependence of the collinear emitted parton is integrated

[6] $K_{ab}(z)$ is trivially zero for $\overline{\text{MS}}$ PDFs.

[7] Note that $E_1 = E_2 = \sqrt{s}/2$ here.

out, but its energy integration remains left,

$$d\hat{\sigma}^{(C,n)}_{\rm FIN}(\dot{r}) = \frac{\alpha_s}{2\pi}\frac{J^{n_L^{(B)}}}{\mathcal{N}(\dot{r}^{\not{i}})}\mathcal{G}(\dot{r}^{\not{i}})\sum_{k=n_I}^{2}\int_0^{\xi_{\rm max}} d\xi_i \left\{\left[\left(\frac{1}{\xi_i}\right)_c \log\left(\frac{s\delta_I}{2\mu^2}\right) + 2\left(\frac{\log(\xi_i)}{\xi_i}\right)_c\right]\right.$$
$$\times\xi_i P^{<}_{\mathcal{I}_{k\oplus\bar{i}}\mathcal{I}_k}(1-\xi_i,0) - \left(\frac{1}{\xi_i}\right)_c \xi_i P'^{<}_{\mathcal{I}_{k\oplus\bar{i}}\mathcal{I}_k}(1-\xi_i,0) - K_{\mathcal{I}_{k\oplus\bar{i}}\mathcal{I}_k}(1-\xi_i)\Big\}$$
$$\times\mathbb{M}^{(n-1,0)}(\dot{r}^{k\oplus\bar{i},\not{i}})d\phi_{n-1}(\dot{r}^{k\oplus\bar{i},\not{i}})\,. \tag{4.65}$$

$P'^{<}_{ab}(z,0)$ is the $\epsilon$ part of the $d=4-2\epsilon$ dimensional (unregularised) Altarelli-Parisi splitting kernels

$$P^{<}_{ab}(z,\epsilon) = P^{<}_{ab}(z,0) + \epsilon P'^{<}_{ab}(z,0). \tag{4.66}$$

## 5 Validations

We present a few cross-checks of our new formalism derived in the last sections. Concerning the local FKS counterterms, we have verified the initial collinear and soft limits of all $2\to 1$ partonic Born processes $gg\to c\bar{c}[n]$ and $q\bar{q}\to c\bar{c}[n]$ with $n$ covering all S- and P-wave charmonium states, including both colour-singlet and colour-octet. One of the most non-trivial checks we performed involves the $2\to 2$ Born processes illustrated in fig. 1. We consider the real processes $gg\to c\bar{c}[n]gg$, $q\bar{q}\to c\bar{c}[n]gg$ and $bg\to c\bar{c}[n]bg$, where both the charm quark $c$ and the bottom quark $b$ are massive, and the light quark $q$ is massless. The real matrix elements are numerically evaluated with HELAC-Onia [9, 10], while the (reduced) Born and colour-linked Born matrix elements are computed analytically. The relative difference of the matrix elements is defined as

$$\Delta_{\rm soft} = \left|\frac{\mathbb{M}^{(n,0)}(\dot{r}) - \lim_{k_i\to 0}\mathbb{M}^{(n,0)}(\dot{r})}{\mathbb{M}^{(n,0)}(\dot{r})}\right| \propto \xi_i + \mathcal{O}(\xi_i^2), \tag{5.1}$$

where the expression of $\lim_{k_i\to 0}\mathbb{M}^{(n,0)}(\dot{r})$ corresponds to the local soft counterterm of the real-emission counterpart $\mathbb{M}^{(n,0)}(\dot{r})$ defined on the r.h.s of eq.(3.36). In fig. 1, we demonstrate the soft gluon limit as $\xi_5\to 0$ for all 8 P-wave Fock states. The relative differences $\Delta_{\rm soft}$ linearly vanish with $\xi_5$ (the rescaled final gluon energy defined in eq.(4.9)) asymptotically approaching zero. For S-wave states, similar soft limit tests were conducted, and the same behaviour was observed. These non-trivial checks confirm the correctness of our local soft counterterms defined in sect. 3.1. Additionally, we have performed initial and final collinear tests for some single quarkonium production processes.

In order to test the integrated counterterms, we have verified the universal IR poles of the one-loop virtual matrix elements outlined in app. B, stemming from the derived integrated counterterms due to the KLN theorem. These checks are carried out explicitly by computing the analytic UV-renormalised one-loop matrix elements for the processes $q\bar{q}\to c\bar{c}[n]g$ with all eight P-wave and four S-wave Fock states $c\bar{c}[n]$. The IR poles of the virtual matrix elements for these 12 processes cancel perfectly with the formulas given in app. B at the analytic level.

## 6 Summary

In this paper, we have generalised the FKS subtraction formalism to processes involving S- or P-wave quarkonium and elementary particles at NLO QCD in NRQCD factorisation. Our main new results are the local soft counterterms in sect. 3.2 and the integrated soft counterterms in sect. 4.3. As a byproduct, we have also derived the universal IR poles of one-loop matrix elements in app. B. This serves as a cornerstone to achieve the NLO automation of cross section computations of quarkonium production processes [8]. Our next steps would be to implement the new formulas in the MadGraph5_aMC@NLO framework, and to generalise the formulation presented here to processes involving more-than-one quarkonia. The latter case, however, involves additional subtleties, such as the breakdown of NRQCD factorisation for processes with two P-wave bound states elucidated in ref. [75].


### Acknowledgments

We thank Michelangelo Mangano for reading the article and the comments. This work is supported by the grants from the ERC (grant 101041109 'BOSON'), the French ANR (grant ANR-20-CE31-0015 'PrecisOnium'), and the French LIA FCPPN.


## A Eikonal tensor integrals

From the integrated soft counterterms, we need to solve the general eikonal tensor integral

$$\bar{\mathcal{E}}^{\alpha_1\dots\alpha_{n_1-1}\beta_1\dots\beta_{n_2-1}}(\{n_1,n_2\},\{k_k,k_l\}) = -\frac{\xi_{\rm cut}^{-2\epsilon}}{2\epsilon}\frac{2^{2\epsilon}}{(2\pi)^{1-2\epsilon}}\left(\frac{s}{\mu^2}\right)^{-\epsilon}\left(k_i^0\right)^2 k_k\cdot k_l \times \underbrace{\int d\Omega_i \frac{k_i^{\alpha_1}\dots k_i^{\alpha_{n_1-1}}k_i^{\beta_1}\dots k_i^{\beta_{n_2-1}}}{(k_k\cdot k_i)^{n_1}(k_l\cdot k_i)^{n_2}}}_{\equiv I^{\alpha_1\dots\alpha_{n_1-1}\beta_1\dots\beta_{n_2-1}}(\{n_1,n_2\},\{k_k,k_l\})}, \quad n_1,n_2\geq 1, \tag{A.1}$$

where we have $m_k^2 = k_k^2$ and $m_l^2 = k_l^2$. We have followed the conventions in ref. [56]. The energy of the parton $i$ and the measure (in $d-1 = 3-2\epsilon$ dimensions) over its angular variables $d\Omega_i$ (*cf.* eq.(4.31)) are defined in the center of mass frame of the colliding partons. We can rewrite the eikonal integral into the phase-space measure as

$$\begin{aligned}\bar{\mathcal{E}}^{\alpha_1\dots\alpha_{n_1-1}\beta_1\dots\beta_{n_2-1}}(\{n_1,n_2\},\{k_k,k_l\}) &= 8\pi^2\mu^{2\epsilon}k_k\cdot k_l \\ &\times\int\frac{d^{3-2\epsilon}\mathbf{k}_i}{(2\pi)^{3-2\epsilon}2k_i^0}\frac{k_i^{\alpha_1}\dots k_i^{\alpha_{n_1-1}}k_i^{\beta_1}\dots k_i^{\beta_{n_2-1}}}{(k_k\cdot k_i)^{n_1}(k_l\cdot k_i)^{n_2}} \\ &\times\theta(\xi_{\rm cut}-\xi_i)\,.\end{aligned} \tag{A.2}$$

[8] Note that the factorisation-breaking effect due to colour transfer between a quarkonium and a heavy quark [74] only occurs in a particular phase-space corner, *i.e.* the mass threshold region, where the relative velocity of the heavy quark and the quarkonium tends to zero. Therefore, it would be harmless for our general purpose unless we are probing that particular phase-space region, where, in any case, we need additional theoretical care.

The introduction of $\theta(\xi_{\rm cut}-\xi_i)$ spoils the Lorentz covariance.

In order to solve these tensor integrals, we are based on the observation that [9]

$$\begin{aligned}
&\frac{\partial}{\partial k_{k,\alpha_{n_1}}} I^{\alpha_1\ldots\alpha_{n_1-1}\beta_1\ldots\beta_{n_2-1}}(\{n_1,n_2\},\{k_k,k_l\}) \\
=\;& -n_1 I^{\alpha_1\ldots\alpha_{n_1-1}\alpha_{n_1}\beta_1\ldots\beta_{n_2-1}}(\{n_1+1,n_2\},\{k_k,k_l\})\,, \\
&\frac{\partial}{\partial k_{l,\beta_{n_2}}} I^{\alpha_1\ldots\alpha_{n_1-1}\beta_1\ldots\beta_{n_2-1}}(\{n_1,n_2\},\{k_k,k_l\}) \\
=\;& -n_2 I^{\alpha_1\ldots\alpha_{n_1-1}\beta_1\ldots\beta_{n_2-1}\beta_{n_2}}(\{n_1,n_2+1\},\{k_k,k_l\})\,.
\end{aligned} \tag{A.3}$$

Then, we can derive the tensor integrals from the scalar integral via

$$\begin{aligned}
I^{\alpha_1}(\{2,1\},\{k_k,k_l\}) &= -\frac{\partial}{\partial k_{k,\alpha_1}} I(\{1,1\},\{k_k,k_l\}), \\
I^{\beta_1}(\{1,2\},\{k_k,k_l\}) &= -\frac{\partial}{\partial k_{l,\beta_1}} I(\{1,1\},\{k_k,k_l\}), \\
I^{\alpha_1\beta_1}(\{2,2\},\{k_k,k_l\}) &= \frac{\partial^2}{\partial k_{k,\alpha_1}\partial k_{l,\beta_1}} I(\{1,1\},\{k_k,k_l\}).
\end{aligned} \tag{A.4}$$

The analytic expressions for $I(\{1,1\},\{k_k,k_l\})$ are known in literature and are given in appendix A of ref. [56].

In the following, we will present the concrete analytic expressions of the eikonal tensor integrals that we need. In general, it would be convenient to further split the tensor integrals into the pole and finite parts:

$$\begin{aligned}
\bar{\mathcal{E}}^{\alpha_1\ldots\alpha_{n_1-1}\beta_1\ldots\beta_{n_2-1}}(\{n_1,n_2\},\{k_k,k_l\}) = \hat{\mathcal{E}}^{\alpha_1\ldots\alpha_{n_1-1}\beta_1\ldots\beta_{n_2-1}}(\{n_1,n_2\},\{k_k,k_l\}) \\
+\mathcal{E}^{\alpha_1\ldots\alpha_{n_1-1}\beta_1\ldots\beta_{n_2-1}}(\{n_1,n_2\},\{k_k,k_l\}),
\end{aligned} \tag{A.5}$$

where $\hat{\mathcal{E}}$ and $\mathcal{E}$ represent the IR poles and the finite component of $\bar{\mathcal{E}}$, respectively. We discuss their expressions in four categories.

### A.1 Two-massless case

We first consider the case of two massless external legs ($m_k = m_l = 0$). Since neither $\mathcal{I}_k$ nor $\mathcal{I}_l$ can be a quarkonium, we only need to consider the scalar integral ($n_1 = n_2 = 1$), which has been known (*e.g.* eqs.(A.5-A.6) in ref. [56]). In this case, if $l = k$, the tensor integrals are zero. If $l \neq k$, for the completeness, the expressions are

$$\hat{\mathcal{E}}(\{1,1\},\{k_k,k_l\}) = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q_{\rm ES}^2}\right)^\epsilon\left[\frac{1}{\epsilon^2}-\frac{1}{\epsilon}\left(\log\left(\frac{2k_k\cdot k_l}{Q_{\rm ES}^2}\right)-\log\left(\frac{4E_kE_l}{\xi_{cut}^2 s}\right)\right)\right], \tag{A.6}$$

$$\begin{aligned}
\mathcal{E}(\{1,1\},\{k_k,k_l\}) = \;& \frac{1}{2}\log^2\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)+\log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)\log\left(\frac{k_k\cdot k_l}{2E_kE_l}\right)-{\rm Li}_2\left(\frac{k_k\cdot k_l}{2E_kE_l}\right) \\
&+\frac{1}{2}\log^2\left(\frac{k_k\cdot k_l}{2E_kE_l}\right)-\log\left(1-\frac{k_k\cdot k_l}{2E_kE_l}\right)\log\left(\frac{k_k\cdot k_l}{2E_kE_l}\right),
\end{aligned} \tag{A.7}$$

where ${\rm Li}_2()$ is the dilogarithm.

[9] Note that in the case $l = k$ the derivative acts on both momenta.

### A.2 One-massive-one-massless case

The second case we consider involves one-massive and one-massless external leg ($m_k = 0$ and $m_l \neq 0$). The scalar case has been given in eqs.(A.7-A.8) in ref. [56]:

$$\hat{\mathcal{E}}(\{1,1\},\{k_k,k_l\}) = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q^2_{\rm ES}}\right)^\epsilon\left[\frac{1}{2\epsilon^2}-\frac{1}{\epsilon}\left(\log\left(\frac{2k_k\cdot k_l}{Q^2_{\rm ES}}\right)-\frac{1}{2}\log\left(\frac{4m_l^2E_k^2}{\xi^2_{cut}sQ^2_{\rm ES}}\right)\right)\right], \tag{A.8}$$

$$\begin{aligned}\mathcal{E}(\{1,1\},\{k_k,k_l\}) = {}& \log(\xi_{cut})\left(\log\left(\frac{\xi_{cut}s}{Q^2_{\rm ES}}\right)+2\log\left(\frac{k_k\cdot k_l}{m_lE_k}\right)\right)-\frac{\pi^2}{12}+\frac{1}{4}\log^2\left(\frac{s}{Q^2_{\rm ES}}\right)\\ &-\frac{1}{4}\log^2\left(\frac{1+\beta_l}{1-\beta_l}\right)+\frac{1}{2}\log^2\left(\frac{k_k\cdot k_l}{(1-\beta_l)E_kE_l}\right)+\log\left(\frac{s}{Q^2_{\rm ES}}\right)\log\left(\frac{k_k\cdot k_l}{m_lE_k}\right)\\ &-{\rm Li}_2\left(1-\frac{(1+\beta_l)E_kE_l}{k_k\cdot k_l}\right)+{\rm Li}_2\left(1-\frac{k_k\cdot k_l}{(1-\beta_l)E_kE_l}\right),\end{aligned} \tag{A.9}$$

where

$$\beta_l = \sqrt{1-\frac{m_l^2}{E_l^2}}. \tag{A.10}$$

The IR pole part of the rank-1 tensor integral needed in the P-wave soft integrated counterterms is

$$\begin{aligned}&\hat{\mathcal{E}}^\mu(\{1,2\},\{k_k,k_l\})\\ &= \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q^2_{\rm ES}}\right)^\epsilon\left\{\frac{1}{2\epsilon^2}\frac{k_k^\mu}{k_k\cdot k_l}+\frac{1}{\epsilon}\left[\frac{k_k^\mu}{k_k\cdot k_l}\left(1-\log\left(\frac{k_k\cdot k_l}{E_km_l}\right)-\frac{1}{2}\log\left(\frac{\xi^2_{cut}s}{Q^2_{\rm ES}}\right)\right)-\frac{k_l^\mu}{m_l^2}\right]\right\},\end{aligned} \tag{A.11}$$

while its finite part can be decomposed into 3 tensor structures

$$\mathcal{E}^\mu(\{1,2\},\{k_k,k_l\}) = \sum_{i=1}^{3} T^{(0,m_l),\mu}_{i,12}, \tag{A.12}$$

where

$$\begin{aligned}T^{(0,m_l),\mu}_{1,12} = {}& \frac{k_k^\mu}{k_k\cdot k_l}\left\{-\frac{\pi^2}{12}-\log\left(\frac{\xi^2_{cut}s}{Q^2_{\rm ES}}\right)+\frac{1}{4}\log^2\left(\frac{\xi^2_{cut}s}{Q^2_{\rm ES}}\right)+\log\left(\frac{k_k\cdot k_l}{E_km_l}\right)\log\left(\frac{\xi^2_{cut}s}{Q^2_{\rm ES}}\right)\right.\\ &-\log\left(\frac{k_k\cdot k_l}{E_kE_l\,(1-\beta_l)}\right)+\frac{1}{2}\log^2\left(\frac{k_k\cdot k_l}{E_kE_l\,(1-\beta_l)}\right)-\frac{1}{4}\log^2\left(\frac{1+\beta_l}{1-\beta_l}\right)\\ &+{\rm Li}_2\left(1-\frac{k_k\cdot k_l}{E_kE_l\,(1-\beta_l)}\right)-{\rm Li}_2\left(1-\frac{E_kE_l\,(1+\beta_l)}{k_k\cdot k_l}\right)\\ &+\frac{k_k\cdot k_l}{k_k\cdot k_l-E_kE_l(1-\beta_l)}\log\left(\frac{k_k\cdot k_l}{E_kE_l\,(1-\beta_l)}\right)+\frac{E_kE_l(1+\beta_l)}{k_k\cdot k_l-E_kE_l(1+\beta_l)}\\ &\left.\times\log\left(\frac{k_k\cdot k_l}{E_kE_l\,(1+\beta_l)}\right)\right\},\end{aligned}$$

$$
\begin{aligned}
T_{2,12}^{(0,m_l),\mu} &= \frac{k_l^\mu}{m_l^2}\left\{\log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)+\frac{m_l^2}{\beta_l E_l^2}\left[-\frac{1}{1-\beta_l^2}\log\left(\frac{1+\beta_l}{1-\beta_l}\right)-\frac{E_k E_l}{k_k\cdot k_l - E_k E_l(1-\beta_l)}\right.\right.\\
&\left.\left.\times\log\left(\frac{k_k\cdot k_l}{E_k E_l(1-\beta_l)}\right)+\frac{E_k E_l}{k_k\cdot k_l - E_k E_l(1+\beta_l)}\log\left(\frac{k_k\cdot k_l}{E_k E_l(1+\beta_l)}\right)\right]\right\},\\
T_{3,12}^{(0,m_l),\mu} &= \frac{\delta^{\mu 0}}{k_k\cdot k_l - E_k E_l\,(1-\beta_l)}\left\{E_k\left[\log\left(\frac{E_k E_l\,(1+\beta_l)}{k_k\cdot k_l}\right)-\log\left(\frac{k_k\cdot k_l}{E_k E_l\,(1-\beta_l)}\right)\right]\right.\\
&+\frac{k_k\cdot k_l\, m_l^2}{E_l^3\beta_l\left(1-\beta_l^2\right)}\log\left(\frac{1+\beta_l}{1-\beta_l}\right)+\frac{E_k\left(E_l^2\beta_l\,(1+\beta_l)+m_l^2\right)}{E_l^2\left(k_k\cdot k_l - E_k E_l\,(1+\beta_l)\right)}\\
&\left.\times\left[E_k E_l\left(\log\left(\frac{E_k E_l\,(1+\beta_l)}{k_k\cdot k_l}\right)-\log\left(\frac{k_k\cdot k_l}{E_k E_l\,(1-\beta_l)}\right)\right)+\frac{k_k\cdot k_l}{1+\beta_l}\log\left(\frac{1+\beta_l}{1-\beta_l}\right)\right]\right\}.
\end{aligned}
\tag{A.13}
$$

### A.3 Massive self-eikonal case

In the case of the massive self-eikonal integrals, we have $l = k$ and $m_k = m_l \neq 0$. The scalar integral corresponds to eqs.(A.9-A.10) in ref. [56]:

$$
\hat{\mathcal{E}}(\{1,1\},\{k_k,k_k\}) = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q_{\rm ES}^2}\right)^\epsilon\left(-\frac{1}{\epsilon}\right), \tag{A.14}
$$

$$
\mathcal{E}(\{1,1\},\{k_k,k_k\}) = \log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)-\frac{1}{\beta_k}\log\left(\frac{1+\beta_k}{1-\beta_k}\right). \tag{A.15}
$$

For the rank-1 tensor integral, the poles are

$$
\hat{\mathcal{E}}^\mu(\{1,2\},\{k_k,k_k\}) = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q_{\rm ES}^2}\right)^\epsilon\frac{k_k^\mu}{m_k^2}\left(-\frac{1}{\epsilon}\right), \tag{A.16}
$$

and the finite term is

$$
\begin{aligned}
\mathcal{E}^\mu(\{1,2\},\{k_k,k_k\}) &= \frac{k_k^\mu}{2\beta_k^3 m_k^2}\left[\log\left(\frac{1+\beta_k}{1-\beta_k}\right)-2\beta_k-3\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\beta_k^2+2\log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)\beta_k^3\right]\\
&+\frac{\delta^{\mu 0}}{2\beta_k^3 E_k}\left[-\log\left(\frac{1+\beta_k}{1-\beta_k}\right)+2\beta_k+\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\beta_k^2\right].
\end{aligned}
\tag{A.17}
$$

The pole of the rank-2 integral is

$$
\hat{\mathcal{E}}^{\mu\nu}(\{2,2\},\{k_k,k_k\}) = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q_{\rm ES}^2}\right)^\epsilon\frac{1}{3m_k^2}\left[-\frac{1}{\epsilon}\left(-g^{\mu\nu}+\frac{4k_k^\mu k_k^\nu}{m_k^2}\right)\right]. \tag{A.18}
$$

The finite piece is

$$
\mathcal{E}^{\mu\nu}(\{2,2\},\{k_k,k_k\}) = \sum_{i=1}^{4} T_{i,22}^{(m_k,m_k),\mu\nu}, \tag{A.19}
$$

where

$$
\begin{aligned}
T_{1,22}^{(m_k,m_k),\mu\nu} &= \frac{\delta^{\mu 0}\delta^{\nu 0}}{6\beta_k^5 m_k^2}(1-\beta_k^2)\left[-3\log\left(\frac{1+\beta_k}{1-\beta_k}\right)+6\beta_k+3\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\beta_k^2-4\beta_k^3\right],\\
T_{2,22}^{(m_k,m_k),\mu\nu} &= \frac{g^{\mu\nu}}{6\beta_k^3 m_k^2}\left[-\log\left(\frac{1+\beta_k}{1-\beta_k}\right)+2\beta_k+3\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\beta_k^2-2\log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)\beta_k^3\right],\\
T_{3,22}^{(m_k,m_k),\mu\nu} &= \frac{\delta^{\mu 0}k_k^\nu+k_k^\mu\delta^{\nu 0}}{6\beta_k^5 E_k m_k^2}\left[3\log\left(\frac{1+\beta_k}{1-\beta_k}\right)-6\beta_k-6\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\beta_k^2\right.\\
&\quad \left.+10\beta_k^3+3\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\beta_k^4\right],\\
T_{4,22}^{(m_k,m_k),\mu\nu} &= \frac{k_k^\mu k_k^\nu}{6\beta_k^5 m_k^4}\left[-3\log\left(\frac{1+\beta_k}{1-\beta_k}\right)+6\beta_k+10\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\beta_k^2\right.\\
&\quad \left.-18\beta_k^3-15\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\beta_k^4+8\log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)\beta_k^5\right].
\end{aligned}
\tag{A.20}
$$

### A.4 Two-massive case

Finally, let us consider the most complicated case with $l \neq k$, $m_k \neq 0$ and $m_l \neq 0$. The scalar integral case can be referred to eqs.(A.11-A.12) in ref. [56]. Its expression is

$$
\hat{\mathcal{E}}(\{1,1\},\{k_k,k_l\}) = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q_{\rm ES}^2}\right)^\epsilon\left(-\frac{1}{2\epsilon}\frac{1}{v_{kl}}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\right), \tag{A.21}
$$

$$
\begin{aligned}
\mathcal{E}(\{1,1\},\{k_k,k_l\}) &= \frac{1}{2v_{kl}}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)\\
&\quad +\frac{(1+v_{kl})(k_k\cdot k_l)^2}{2m_k^2}\left(\mathrm{J}^{(A)}\left(\alpha_{kl}E_k,\alpha_{kl}E_k\beta_k\right)-\mathrm{J}^{(A)}\left(E_l,E_l\beta_l\right)\right),
\end{aligned}
\tag{A.22}
$$

where the introduced function is

$$
\mathrm{J}^{(A)}(x,y) = \frac{1}{2\lambda\nu}\left[\log^2\left(\frac{x-y}{x+y}\right)+4\mathrm{Li}_2\left(1-\frac{x+y}{\nu}\right)+4\mathrm{Li}_2\left(1-\frac{x-y}{\nu}\right)\right] \tag{A.23}
$$

and

$$
\begin{aligned}
v_{kl} &= \sqrt{1-\left(\frac{m_k m_l}{k_k\cdot k_l}\right)^2}, \quad \alpha_{kl} = \frac{1+v_{kl}}{m_k^2}k_k\cdot k_l,\\
\lambda &= \alpha_{kl}E_k - E_l, \quad \nu = \frac{\alpha_{kl}^2 m_k^2 - m_l^2}{2\lambda}.
\end{aligned}
\tag{A.24}
$$

For the two-massive rank-1 eikonal tensor integral, its pole is

$$
\begin{aligned}
\hat{\mathcal{E}}^\mu(\{1,2\},\{k_k,k_l\}) &= \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q_{\rm ES}^2}\right)^\epsilon\frac{1}{2\epsilon}\frac{m_k^2}{(k_k\cdot k_l)^2 v_{kl}^3}\left\{-k_k^\mu\left[\frac{k_k\cdot k_l}{m_k^2}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)v_{kl}^2\right.\right.\\
&\quad \left.\left.+\frac{m_l^2}{k_k\cdot k_l}\left(\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)-\frac{2v_{kl}}{1-v_{kl}^2}\right)\right]+k_l^\mu\left(\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)-\frac{2v_{kl}}{1-v_{kl}^2}\right)\right\}.
\end{aligned}
\tag{A.25}
$$

The finite term can be decomposed into 8 tensorial components

$$\mathcal{E}^{\mu}(\{1,2\},\{k_k,k_l\}) = \sum_{i=1}^{8} T_{i,12}^{(m_k,m_l),\mu}, \tag{A.26}$$

where

$$\begin{aligned}
T_{1,12}^{(m_k,m_l),\mu} &= k_k^{\mu}\left[\frac{1}{2(k_k\cdot k_l)v_{kl}}\log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\right.\\
&\left.-\frac{(1+v_{kl})(k_k\cdot k_l)}{2m_k^2}\left(\mathrm{J}^{(A)}(\alpha_{kl}E_k,\alpha_{kl}\beta_k E_k)-\mathrm{J}^{(A)}(E_l,\beta_l E_l)\right)\right],\\
T_{2,12}^{(m_k,m_l),\mu} &= \frac{1}{2v_{kl}}\left(k_k^{\mu}\frac{m_l^2}{(k_k\cdot k_l)}-k_l^{\mu}\right)\left\{\frac{m_k^2}{(k_k\cdot k_l)^2 v_{kl}^2}\log\left(\frac{\xi_{cut}^2 s}{Q_{\rm ES}^2}\right)\left[\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)-\frac{2v_{kl}}{1-v_{kl}^2}\right]\right.\\
&\left.-\left(\mathrm{J}^{(A)}(\alpha_{kl}E_k,\alpha_{kl}\beta_k E_k)-\mathrm{J}^{(A)}(E_l,\beta_l E_l)\right)\right\},\\
T_{3,12}^{(m_k,m_l),\mu} &= \frac{(1+v_{kl})E_k(k_k\cdot k_l)}{\lambda\nu\left[(\nu-\alpha_{kl}E_k)^2-(\alpha_{kl}\beta_k E_k)^2\right]v_{kl}m_k^2}\left(-k_k^{\mu}\left(\alpha_{kl}v_{kl}+\frac{m_l^2}{k_k\cdot k_l}\right)+k_l^{\mu}\right)\\
&\left\{\left(\nu-\alpha_{kl}E_k(1-\beta_k^2)\right)\left[\log\left(\frac{\alpha_{kl}E_k(1-\beta_k)}{\nu}\right)+\log\left(\frac{\alpha_{kl}E_k(1+\beta_k)}{\nu}\right)\right]\right.\\
&\left.+\nu\beta_k\log\left(\frac{1+\beta_k}{1-\beta_k}\right)\right\},\\
T_{4,12}^{(m_k,m_l),\mu} &= \frac{1+v_{kl}}{2\lambda^2\nu^2 v_{kl}m_k^2}\left[-k_k^{\mu}\left(\nu E_k-\alpha_{kl}m_k^2\right)\left(\alpha_{kl}v_{kl}(k_k\cdot k_l)+m_l^2\right)\right.\\
&\left.+k_l^{\mu}(k_k\cdot k_l)(\nu E_k-v_{kl}(k_k\cdot k_l)-\alpha_{kl}m_k^2)+\delta^{\mu 0}\nu v_{kl}(k_k\cdot k_l)^2\right]\\
&\times\left[\frac{2\alpha_{kl}E_k(1-\beta_k)}{\nu-\alpha_{kl}E_k(1-\beta_k)}\log\left(\frac{\alpha_{kl}E_k(1-\beta_k)}{\nu}\right)+\frac{2\alpha_{kl}E_k(1+\beta_k)}{\nu-\alpha_{kl}E_k(1+\beta_k)}\right.\\
&\left.\times\log\left(\frac{\alpha_{kl}E_k(1+\beta_k)}{\nu}\right)+\lambda\nu\mathrm{J}^{(A)}(\alpha_{kl}E_k,\alpha_{kl}\beta_k E_k)\right],\\
T_{5,12}^{(m_k,m_l),\mu} &= \frac{1+v_{kl}}{2\lambda v_{kl}m_k^2}\left[k_k^{\mu}E_k(\alpha_{kl}v_{kl}(k_k\cdot k_l)+m_l^2)-k_l^{\mu}E_k(k_k\cdot k_l)-\delta^{\mu 0}v_{kl}(k_k\cdot k_l)^2\right]\\
&\times\left[\mathrm{J}^{(A)}(\alpha_{kl}E_k,\alpha_{kl}\beta_k E_k)-\mathrm{J}^{(A)}(E_l,\beta_l E_l)\right],\\
T_{6,12}^{(m_k,m_l),\mu} &= \frac{(1+v_{kl})(k_k\cdot k_l)^2}{\lambda\nu m_k^2\left[(\nu-E_l)^2-(\beta_l E_l)^2\right]}\delta^{\mu 0}\left\{\left(\nu-E_l(1-\beta_l^2)\right)\left[\log\left(\frac{E_l(1-\beta_l)}{\nu}\right)\right.\right.\\
&\left.\left.+\log\left(\frac{E_l(1+\beta_l)}{\nu}\right)\right]+\nu\beta_l\log\left(\frac{1+\beta_l}{1-\beta_l}\right)\right\},\\
T_{7,12}^{(m_k,m_l),\mu} &= \frac{(1+v_{kl})(k_k\cdot k_l)^2}{\lambda\nu\beta_l E_l^2 m_k^2}\left(k_l^{\mu}E_l-\delta^{\mu 0}m_l^2\right)\left\{\frac{1}{\nu-E_l(1-\beta_l)}\log\left(\frac{E_l(1-\beta_l)}{\nu}\right)\right.\\
&\left.-\frac{1}{\nu-E_l(1+\beta_l)}\log\left(\frac{E_l(1+\beta_l)}{\nu}\right)-\frac{1}{E_l(1-\beta_l^2)}\log\left(\frac{1+\beta_l}{1-\beta_l}\right)\right\},
\end{aligned}$$

$$
\begin{aligned}
T_{8,12}^{(m_k,m_l),\mu} = & \frac{1+v_{kl}}{2\lambda^2\nu^2 v_{kl} m_k^2}\left[k_k^\mu\left(\nu E_k - \alpha_{kl} m_k^2\right)\left(\alpha_{kl} v_{kl}(k_k\cdot k_l) + m_l^2\right)\right. \\
& \left. -k_l^\mu (k_k\cdot k_l)(\nu E_k - v_{kl}(k_k\cdot k_l) - \alpha_{kl} m_k^2) - \delta^{\mu 0}\nu v_{kl}(k_k\cdot k_l)^2\right] \\
& \times\left[\frac{2E_l(1-\beta_l)}{\nu - E_l(1-\beta_l)}\log\left(\frac{E_l(1-\beta_l)}{\nu}\right) + \frac{2E_l(1+\beta_l)}{\nu - E_l(1+\beta_l)}\log\left(\frac{E_l(1+\beta_l)}{\nu}\right)\right. \\
& \left. +\lambda\nu \mathrm{J}^{(A)}(E_l, \beta_l E_l)\right]. \qquad \text{(A.27)}
\end{aligned}
$$

We have verified that if we take $k = l$, the two-massive eikonal integrals are reduced to the massive self-eikonal case.

## B Infrared poles of one-loop matrix elements

We present here the general IR poles of the UV renormalised one-loop matrix element at NLO QCD for arbitrary process that involves a quarkonium and elementary particles by assuming the validity of the KLN theorem. The one-loop virtual matrix element can be written as

$$
\mathbb{M}^{(n-1,1)}(\dot{r}^{\dot{\wr}}) = \frac{\alpha_s}{2\pi}\frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{Q_{\rm ES}^2}\right)^\epsilon \mathbb{V}(\dot{r}^{\dot{\wr}})\,. \qquad \text{(B.1)}
$$

We can derive the IR poles of $\mathbb{V}(\dot{r}^{\dot{\wr}})$ generally from our integrated counterterms with its finite term denoted as $\mathbb{V}_{\rm FIN}^{(n-1,1)}(\dot{r}^{\dot{\wr}})$. We remind readers that in the process $\dot{r}^{\dot{\wr}}$, the particles with their indices from $n_I$ to $n_L^{(B)}+2$ are coloured and massless, while those from $n_L^{(B)}+3$ to $n_L^{(B)}+n_H$ ($n_H \geq 2$) are massive coloured elementary particles. The index of the quarkonium is $n_L^{(B)} + n_H + 1$.

### B.1 Colour-singlet S-wave state

For a colour-singlet S-wave state with $\dot{r}^{\dot{\wr}}$ being eq.(4.42), the IR poles can be obtained by taking the bound state as a colour-singlet elementary particle, which amount to

$$
\begin{aligned}
\mathbb{V}(\dot{r}^{\dot{\wr}}) = & -\left(\frac{1}{\epsilon^2}\sum_{k=n_I}^{n_L^{(B)}+2} C(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\gamma(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H} C(\mathcal{I}_k)\right)\mathbb{M}^{(n-1,0)}(\dot{r}^{\dot{\wr}}) \\
& +\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\sum_{l=k+1}^{n_L^{(B)}+n_H}\log\left(\frac{2k_k\cdot k_l}{Q_{\rm ES}^2}\right)\mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\dot{\wr}}) \\
& +\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H-1}\sum_{l=k+1}^{n_L^{(B)}+n_H}\frac{1}{v_{kl}}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\dot{\wr}}) \\
& -\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\log\left(\frac{m_k^2}{Q_{\rm ES}^2}\right)\sum_{l=n_I}^{n_L^{(B)}+2}\mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\dot{\wr}}) + \mathbb{V}_{\rm FIN}^{(n-1,1)}(\dot{r}^{\dot{\wr}})\,. \qquad \text{(B.2)}
\end{aligned}
$$

This equation is equivalent to eq.(B.2) in ref. [56].

### B.2 Colour-octet S-wave state

The colour-octet S-wave case with $\dot{r}$ being eq.(4.44) has the IR poles as follows:

$$\begin{aligned}
\mathbb{V}(\dot{r}) = &-\left(\frac{1}{\epsilon^2}\sum_{k=n_I}^{n_L^{(B)}+2} C(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\gamma(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+1} C(\mathcal{I}_k)\right)\mathbb{M}^{(n-1,0)}(\dot{r}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\sum_{l=k+1}^{n_L^{(B)}+n_H+1}\log\left(\frac{2k_k\cdot k_l}{Q^2_{\rm ES}}\right)\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\sum_{l=k+1}^{n_L^{(B)}+n_H+1}\frac{1}{v_{kl}}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) \\
&-\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+1}\log\left(\frac{m_k^2}{Q^2_{\rm ES}}\right)\sum_{l=n_I}^{n_L^{(B)}+2}\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) + \mathbb{V}^{(n-1,1)}_{\rm FIN}(\dot{r})\,.
\end{aligned} \tag{B.3}$$

### B.3 Colour-singlet spin-singlet P-wave state

The IR poles of the virtual matrix elements for a colour-singlet spin-singlet P-wave quarkonium production are

$$\begin{aligned}
\mathbb{V}(\dot{r}) = &-\left(\frac{1}{\epsilon^2}\sum_{k=n_I}^{n_L^{(B)}+2} C(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\gamma(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H} C(\mathcal{I}_k)\right)\mathbb{M}^{(n-1,0)}(\dot{r}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\sum_{l=k+1}^{n_L^{(B)}+n_H}\log\left(\frac{2k_k\cdot k_l}{Q^2_{\rm ES}}\right)\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H-1}\sum_{l=k+1}^{n_L^{(B)}+n_H}\frac{1}{v_{kl}}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) \\
&-\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\log\left(\frac{m_k^2}{Q^2_{\rm ES}}\right)\sum_{l=n_I}^{n_L^{(B)}+2}\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\frac{k_{k,\mu}}{K\cdot k_k}\mathbb{M}^{(n-1,0)}_{k[18]}(\dot{r},\dot{r}_1)^\mu \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\frac{k_{k,\mu}}{K\cdot k_k}\frac{m_k^2K^2}{v_k^3(K\cdot k_k)^2}\left[\frac{2v_k}{1-v_k^2}-\log\left(\frac{1+v_k}{1-v_k}\right)\right]\mathbb{M}^{(n-1,0)}_{k[18]}(\dot{r},\dot{r}_1)^\mu \\
&-\frac{1}{\epsilon}\left(\frac{1}{2N_c}\frac{8}{m_Q m_{\bar{Q}'}}-\frac{2}{K^2}C_{\rm eff}(Q\bar{Q}'_{[18]})\right)\mathbb{M}^{(n-1,0)}(\dot{r}_1)^{\mu\nu} + \mathbb{V}^{(n-1,1)}_{\rm FIN}(\dot{r})\,,
\end{aligned} \tag{B.4}$$

where we have defined the abbreviation for $v_{kl}$ with $l = n_L^{(B)} + n_H + 1$ (quarkonium)

$$v_k = \sqrt{1 - \frac{k_k^2 K^2}{(k_k \cdot K)^2}}. \tag{B.5}$$

Note that, the two reduced Born processes $\dot{r}$ and $\dot{r}_1$ have been given in eq.(4.46).

### B.4 Colour-octet spin-singlet P-wave state

For the colour-octet spin-singlet P-wave state, the IR poles of the virtual corrections are

$$\begin{aligned}
\mathbb{V}(\dot{r}) = &-\left(\frac{1}{\epsilon^2}\sum_{k=n_I}^{n_L^{(B)}+2} C(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2} \gamma(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+1} C(\mathcal{I}_k)\right)\mathbb{M}^{(n-1,0)}(\dot{r}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\sum_{l=k+1}^{n_L^{(B)}+n_H+1} \log\left(\frac{2k_k\cdot k_l}{Q^2_{\rm ES}}\right)\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\sum_{l=k+1}^{n_L^{(B)}+n_H+1} \frac{1}{v_{kl}}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) \\
&-\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+1} \log\left(\frac{m_k^2}{Q^2_{\rm ES}}\right)\sum_{l=n_I}^{n_L^{(B)}+2}\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\frac{k_{k,\mu}}{K\cdot k_k}\left(\mathbb{M}^{(n-1,0)}_{k[88]}(\dot{r},\dot{r}_1)^\mu + \mathbb{M}^{(n-1,0)}_{k[81]}(\dot{r},\dot{r}_2)^\mu\right) \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\frac{k_{k,\mu}}{K\cdot k_k}\frac{m_k^2K^2}{v_k^3(K\cdot k_k)^2}\left[\frac{2v_k}{1-v_k^2} - \log\left(\frac{1+v_k}{1-v_k}\right)\right] \\
&\times\left(\mathbb{M}^{(n-1,0)}_{k[88]}(\dot{r},\dot{r}_1)^\mu + \mathbb{M}^{(n-1,0)}_{k[81]}(\dot{r},\dot{r}_2)^\mu\right) \\
&-\frac{1}{\epsilon}\left(B_F\frac{8}{m_Q m_{\bar{Q}'}} - \frac{2}{K^2}C_{\rm eff}(Q\bar{Q}'_{[88]})\right)\mathbb{M}^{(n-1,0)}(\dot{r}_1) \\
&-\frac{1}{\epsilon}\left(C_F\frac{8}{m_Q m_{\bar{Q}'}} - \frac{2}{K^2}C_{\rm eff}(Q\bar{Q}'_{[81]})\right)\mathbb{M}^{(n-1,0)}(\dot{r}_2) + \mathbb{V}^{(n-1,1)}_{\rm FIN}(\dot{r})\,.
\end{aligned} \tag{B.6}$$

The three reduced Born processes $\dot{r}$, $\dot{r}_1$ and $\dot{r}_2$ have been given in eq.(4.48).

### B.5 Colour-singlet spin-triplet P-wave state

The IR poles of the one-loop virtual corrections for the processes involving a colour-singlet spin-triple P-wave quarkonium are

$$
\begin{aligned}
\mathbb{V}(\dot{r}^{\dot{\lambda}}) = &-\left(\frac{1}{\epsilon^2}\sum_{k=n_I}^{n_L^{(B)}+2} C(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2} \gamma(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H} C(\mathcal{I}_k)\right)\mathbb{M}^{(n-1,0)}(\dot{r}^{\dot{\lambda}}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\sum_{l=k+1}^{n_L^{(B)}+n_H} \log\left(\frac{2k_k\cdot k_l}{Q^2_{\rm ES}}\right)\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}^{\dot{\lambda}}) \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H-1}\sum_{l=k+1}^{n_L^{(B)}+n_H} \frac{1}{v_{kl}}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}^{\dot{\lambda}}) \\
&-\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H} \log\left(\frac{m_k^2}{Q^2_{\rm ES}}\right)\sum_{l=n_I}^{n_L^{(B)}+2}\mathbb{M}^{(n-1,0)}_{kl}(\dot{r}^{\dot{\lambda}}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\frac{k_{k,\mu}}{K\cdot k_k}\mathbb{M}^{(n-1,0)}_{J,k[18]}(\dot{r}^{\dot{\lambda}},\dot{r}_1^{\dot{\lambda}})^{\mu} \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\frac{k_{k,\mu}}{K\cdot k_k}\frac{m_k^2K^2}{v_k^3(K\cdot k_k)^2}\left[\frac{2v_k}{1-v_k^2}-\log\left(\frac{1+v_k}{1-v_k}\right)\right]\mathbb{M}^{(n-1,0)}_{J,k[18]}(\dot{r}^{\dot{\lambda}},\dot{r}_1^{\dot{\lambda}})^{\mu} \\
&-\frac{1}{\epsilon}\frac{2}{3K^2}g_{\mu\nu}C_{\rm eff}(Q\bar{Q}'_{[18]})\mathbb{M}^{(n-1,0)}_{J}(\dot{r}_1^{\dot{\lambda}})^{\mu\nu} - \frac{1}{\epsilon}\frac{2J+1}{2N_c}\frac{8}{9m_Qm_{\bar{Q}'}}\mathbb{M}^{(n-1,0)}(\dot{r}_1^{\dot{\lambda}}) \\
&+\mathbb{V}^{(n-1,1)}_{\rm FIN}(\dot{r}^{\dot{\lambda}})\,,
\end{aligned}
\tag{B.7}
$$

where the processes $\dot{r}^{\dot{\lambda}}$ and $\dot{r}_1^{\dot{\lambda}}$ are eq.(4.50).

### B.6 Colour-octet spin-triplet P-wave state

Similarly, the IR poles of the virtual corrections for the colour-octet spin-triplet P-wave quarkonium are

$$
\begin{aligned}
\mathbb{V}(\dot{r}^{\dagger}) = &-\left(\frac{1}{\epsilon^2}\sum_{k=n_I}^{n_L^{(B)}+2} C(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\gamma(\mathcal{I}_k) + \frac{1}{\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+1} C(\mathcal{I}_k)\right)\mathbb{M}^{(n-1,0)}(\dot{r}^{\dagger}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\sum_{l=k+1}^{n_L^{(B)}+n_H+1}\log\left(\frac{2k_k\cdot k_l}{Q_{\rm ES}^2}\right)\mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\dagger}) \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\sum_{l=k+1}^{n_L^{(B)}+n_H+1}\frac{1}{v_{kl}}\log\left(\frac{1+v_{kl}}{1-v_{kl}}\right)\mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\dagger}) \\
&-\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H+1}\log\left(\frac{m_k^2}{Q_{\rm ES}^2}\right)\sum_{l=n_I}^{n_L^{(B)}+2}\mathbb{M}_{kl}^{(n-1,0)}(\dot{r}^{\dagger}) \\
&+\frac{1}{\epsilon}\sum_{k=n_I}^{n_L^{(B)}+2}\frac{k_{k,\mu}}{K\cdot k_k}\left(\mathbb{M}_{J,k[88]}^{(n-1,0)}(\dot{r}^{\dagger},\dot{r}_1^{\dagger})^{\mu}+\mathbb{M}_{J,k[81]}^{(n-1,0)}(\dot{r}^{\dagger},\dot{r}_2^{\dagger})^{\mu}\right) \\
&+\frac{1}{2\epsilon}\sum_{k=n_L^{(B)}+3}^{n_L^{(B)}+n_H}\frac{k_{k,\mu}}{K\cdot k_k}\frac{m_k^2K^2}{v_k^3(K\cdot k_k)^2}\left[\frac{2v_k}{1-v_k^2}-\log\left(\frac{1+v_k}{1-v_k}\right)\right] \\
&\times\left(\mathbb{M}_{J,k[88]}^{(n-1,0)}(\dot{r}^{\dagger},\dot{r}_1^{\dagger})^{\mu}+\mathbb{M}_{J,k[81]}^{(n-1,0)}(\dot{r}^{\dagger},\dot{r}_2^{\dagger})^{\mu}\right) \\
&-\frac{1}{\epsilon}\frac{2}{3K^2}g_{\mu\nu}\left(C_{\rm eff}(Q\bar{Q}'_{[88]})\mathbb{M}_J^{(n-1,0)}(\dot{r}_1^{\dagger})^{\mu\nu}+C_{\rm eff}(Q\bar{Q}'_{[81]})\mathbb{M}_J^{(n-1,0)}(\dot{r}_2^{\dagger})^{\mu\nu}\right) \\
&-\frac{1}{\epsilon}\frac{8(2J+1)}{9m_Qm_{\bar{Q}'}}\left(B_F\mathbb{M}^{(n-1,0)}(\dot{r}_1^{\dagger})+C_F\mathbb{M}^{(n-1,0)}(\dot{r}_2^{\dagger})\right)+\mathbb{V}_{\rm FIN}^{(n-1,1)}(\dot{r}^{\dagger})\,. \qquad \text{(B.8)}
\end{aligned}
$$

The processes $\dot{r}^{\dagger}$, $\dot{r}_1^{\dagger}$ and $\dot{r}_2^{\dagger}$ are eq.(4.52).

## References


[1] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [1405.0301].

[2] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. S. Shao and M. Zaro, *The automation of next-to-leading order electroweak calculations*, *JHEP* **07** (2018) 185, [1804.10017].

[3] M. Schönherr, *An automated subtraction of NLO EW infrared divergences*, *Eur. Phys. J. C* **78** (2018) 119, [1712.07975].

[4] B. Biedermann, S. Bräuer, A. Denner, M. Pellen, S. Schumann and J. M. Thompson, *Automation of NLO QCD and EW corrections with Sherpa and Recola*, *Eur. Phys. J. C* **77** (2017) 492, [1704.05783].

[5] S. Kallweit, J. M. Lindert, S. Pozzorini and M. Schönherr, *NLO QCD+EW predictions for $2\ell 2\nu$ diboson signatures at the LHC*, *JHEP* **11** (2017) 120, [1705.00598].

[6] M. Chiesa, N. Greiner, M. Schönherr and F. Tramontano, *Electroweak corrections to diphoton plus jets*, *JHEP* **10** (2017) 181, [1706.09022].

[7] G. T. Bodwin, E. Braaten and G. P. Lepage, *Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium*, *Phys. Rev. D* **51** (1995) 1125–1171, [hep-ph/9407339].

[8] P. Artoisenet, F. Maltoni and T. Stelzer, *Automatic generation of quarkonium amplitudes in NRQCD*, *JHEP* **02** (2008) 102, [0712.2770].

[9] H.-S. Shao, *HELAC-Onia: An automatic matrix element generator for heavy quarkonium physics*, *Comput. Phys. Commun.* **184** (2013) 2562–2570, [1212.5293].

[10] H.-S. Shao, *HELAC-Onia 2.0: an upgraded matrix-element and event generator for heavy quarkonium physics*, *Comput. Phys. Commun.* **198** (2016) 238–259, [1507.03435].

[11] N. Cooke, P. Ilten, L. Lönnblad and S. Mrenna, *Non-Relativistic Quantum Chromodynamics in Parton Showers*, 2312.05203.

[12] H.-S. Shao, *Boosting perturbative QCD stability in quarkonium production*, *JHEP* **01** (2019) 112, [1809.02369].

[13] E. Chapon et al., *Prospects for quarkonium studies at the high-luminosity LHC*, *Prog. Part. Nucl. Phys.* **122** (2022) 103906, [2012.14161].

[14] D. Boer et al., *Gluons and the quark sea at high energies: Distributions, polarization, tomography*, 1108.1713.

[15] W. J. den Dunnen, J. P. Lansberg, C. Pisano and M. Schlegel, *Accessing the Transverse Dynamics and Polarization of Gluons inside the Proton at the LHC*, *Phys. Rev. Lett.* **112** (2014) 212001, [1401.7611].

[16] J.-P. Lansberg, C. Pisano, F. Scarpa and M. Schlegel, *Pinning down the linearly-polarised gluons inside unpolarised protons using quarkonium-pair production at the LHC*, *Phys. Lett. B* **784** (2018) 217–222, [1710.01684].

[17] S. P. Jones, A. D. Martin, M. G. Ryskin and T. Teubner, *Exclusive $J/\psi$ and $\Upsilon$ photoproduction and the low $x$ gluon*, *J. Phys. G* **43** (2016) 035002, [1507.06942].

[18] S. P. Jones, A. D. Martin, M. G. Ryskin and T. Teubner, *The exclusive $J/\psi$ process at the LHC tamed to probe the low $x$ gluon*, *Eur. Phys. J. C* **76** (2016) 633, [1610.02272].

[19] J.-P. Lansberg and H.-S. Shao, *$J/\psi$ -pair production at large momenta: Indications for double parton scatterings and large $\alpha_s^5$ contributions*, *Phys. Lett. B* **751** (2015) 479–486, [1410.8822].

[20] H.-S. Shao and Y.-J. Zhang, *Complete study of hadroproduction of a $\Upsilon$ meson associated with a prompt $J/\psi$*, *Phys. Rev. Lett.* **117** (2016) 062001, [1605.03061].

[21] J.-P. Lansberg and H.-S. Shao, *Associated production of a quarkonium and a $Z$ boson at one loop in a quark-hadron-duality approach*, *JHEP* **10** (2016) 153, [1608.03198].

[22] J.-P. Lansberg and H.-S. Shao, *Phenomenological analysis of associated production of $Z^0 + b$ in the $b \to J/\psi X$ decay channel at the LHC*, *Nucl. Phys. B* **916** (2017) 132–142, [1611.09303].

[23] J.-P. Lansberg, H.-S. Shao and N. Yamanaka, *Indication for double parton scatterings in $W+$ prompt $J/\psi$ production at the LHC*, *Phys. Lett. B* **781** (2018) 485–491, [1707.04350].

[24] H.-S. Shao and Y.-J. Zhang, *Triple prompt $J/\psi$ hadroproduction as a hard probe of multiple-parton scatterings*, *Phys. Rev. Lett.* **122** (2019) 192002, [1902.04949].

[25] C. A. Flett, A. D. Martin, M. G. Ryskin and T. Teubner, *Very low x gluon density determined by LHCb exclusive $J/\psi$ data*, *Phys. Rev. D* **102** (2020) 114021, [2006.13857].

[26] A. Kusina, J.-P. Lansberg, I. Schienbein and H.-S. Shao, *Gluon Shadowing in Heavy-Flavor Production at the LHC*, *Phys. Rev. Lett.* **121** (2018) 052004, [1712.07024].

[27] H.-S. Shao, *Probing impact-parameter dependent nuclear parton densities from double parton scatterings in heavy-ion collisions*, *Phys. Rev. D* **101** (2020) 054036, [2001.04256].

[28] H.-S. Shao, *$J/\psi$ meson production in association with an open charm hadron at the LHC: A reappraisal*, *Phys. Rev. D* **102** (2020) 034023, [2005.12967].

[29] V. Guzey, E. Kryshen, M. Strikman and M. Zhalov, *Evidence for nuclear gluon shadowing from the ALICE measurements of PbPb ultraperipheral exclusive $J/\psi$ production*, *Phys. Lett. B* **726** (2013) 290–295, [1305.1724].

[30] T. Matsui and H. Satz, *$J/\psi$ Suppression by Quark-Gluon Plasma Formation*, *Phys. Lett. B* **178** (1986) 416–422.

[31] S. Digal, P. Petreczky and H. Satz, *Quarkonium feed down and sequential suppression*, *Phys. Rev. D* **64** (2001) 094015, [hep-ph/0106017].

[32] P. Braun-Munzinger and J. Stachel, *(Non)thermal aspects of charmonium production and a new look at J / psi suppression*, *Phys. Lett. B* **490** (2000) 196–202, [nucl-th/0007059].

[33] R. L. Thews, M. Schroedter and J. Rafelski, *Enhanced $J/\psi$ production in deconfined quark matter*, *Phys. Rev. C* **63** (2001) 054905, [hep-ph/0007323].

[34] J. D. Bjorken, *Intersections 2000: What's new in hadron physics*, *AIP Conf. Proc.* **549** (2000) 211–229, [hep-ph/0008048].

[35] V. Zetocha and T. Schäfer, *Instanton contribution to scalar charmonium and glueball decays*, *Phys. Rev. D* **67** (2003) 114003, [hep-ph/0212125].

[36] D. Kharzeev, E. Levin, M. Nardi and K. Tuchin, *Gluon saturation effects on J/Psi production in heavy ion collisions*, *Phys. Rev. Lett.* **102** (2009) 152301, [0808.2954].

[37] N. Brambilla, X. Garcia i Tormo, J. Soto and A. Vairo, *Extraction of alpha(s) from radiative Upsilon(1S) decays*, *Phys. Rev. D* **75** (2007) 074014, [hep-ph/0702079].

[38] G. T. Bodwin, F. Petriello, S. Stoynev and M. Velasco, *Higgs boson decays to quarkonia and the $H\bar{c}c$ coupling*, *Phys. Rev. D* **88** (2013) 053003, [1306.5770].

[39] ATLAS collaboration, G. Aad et al., *Search for Higgs and Z Boson Decays to $J/\psi\gamma$ and $\Upsilon(nS)\gamma$ with the ATLAS Detector*, *Phys. Rev. Lett.* **114** (2015) 121801, [1501.03276].

[40] ATLAS collaboration, M. Aaboud et al., *Searches for exclusive Higgs and Z boson decays into $J/\psi\gamma$, $\psi(2S)\gamma$, and $\Upsilon(nS)\gamma$ at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 134–155, [1807.00802].

[41] Belle collaboration, K. Abe et al., *Observation of large CP violation in the neutral B meson system*, *Phys. Rev. Lett.* **87** (2001) 091802, [hep-ex/0107061].

[42] BaBar collaboration, B. Aubert et al., *Observation of CP violation in the $B^0$ meson system*, *Phys. Rev. Lett.* **87** (2001) 091801, [hep-ex/0107013].

[43] V. Mateu and P. G. Ortega, *Bottom and Charm Mass determinations from global fits to $Q\bar{Q}$ bound states at $N^3LO$*, *JHEP* **01** (2018) 122, [1711.05755].

[44] LHCb collaboration, R. Aaij et al., *Observation of structure in the $J/\psi$ -pair mass spectrum*, *Sci. Bull.* **65** (2020) 1983–1993, [2006.16957].

[45] ATLAS collaboration, G. Aad et al., *Observation of an Excess of Dicharmonium Events in the Four-Muon Final State with the ATLAS Detector*, *Phys. Rev. Lett.* **131** (2023) 151902, [2304.08962].

[46] CMS collaboration, A. Hayrapetyan et al., *Observation of new structure in the $J/\psi J/\psi$ mass spectrum in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, 2306.07164.

[47] LHCb collaboration, R. Aaij et al., *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, *Phys. Rev. Lett.* **115** (2015) 072001, [1507.03414].

[48] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650–677.

[49] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Phys. Rev.* **133** (1964) B1549–B1562.

[50] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys. B* **467** (1996) 399–442, [hep-ph/9512328].

[51] S. Frixione, *A General approach to jet cross-sections in QCD*, *Nucl. Phys. B* **507** (1997) 295–314, [hep-ph/9706545].

[52] S. Catani and M. H. Seymour, *The Dipole formalism for the calculation of QCD jet cross-sections at next-to-leading order*, *Phys. Lett. B* **378** (1996) 287–301, [hep-ph/9602277].

[53] S. Catani and M. H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291–419, [hep-ph/9605323].

[54] L. Phaf and S. Weinzierl, *Dipole formalism with heavy fermions*, *JHEP* **04** (2001) 006, [hep-ph/0102207].

[55] S. Catani, S. Dittmaier, M. H. Seymour and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys. B* **627** (2002) 189–265, [hep-ph/0201036].

[56] R. Frederix, S. Frixione, F. Maltoni and T. Stelzer, *Automation of next-to-leading order computations in QCD: The FKS subtraction*, *JHEP* **10** (2009) 003, [0908.4272].

[57] T. Gleisberg and F. Krauss, *Automating dipole subtraction for QCD NLO calculations*, *Eur. Phys. J. C* **53** (2008) 501–523, [0709.2881].

[58] R. Frederix, T. Gehrmann and N. Greiner, *Automation of the Dipole Subtraction Method in MadGraph/MadEvent*, *JHEP* **09** (2008) 122, [0808.2128].

[59] M. Czakon, C. G. Papadopoulos and M. Worek, *Polarizing the Dipoles*, *JHEP* **08** (2009) 085, [0905.0883].

[60] K. Hasegawa, S. Moch and P. Uwer, *AutoDipole: Automated generation of dipole subtraction terms*, *Comput. Phys. Commun.* **181** (2010) 1802–1817, [0911.4371].

[61] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [1002.2581].

[62] B. W. Harris and J. F. Owens, *The Two cutoff phase space slicing method*, *Phys. Rev. D* **65** (2002) 094032, [hep-ph/0102128].

[63] J. M. Campbell, F. Maltoni and F. Tramontano, *QCD corrections to J/psi and Upsilon production at hadron colliders*, *Phys. Rev. Lett.* **98** (2007) 252002, [hep-ph/0703113].

[64] J.-W. Qiu, X.-P. Wang and H. Xing, *Exploring $J/\psi$ Production Mechanism at the Future Electron-Ion Collider*, *Chin. Phys. Lett.* **38** (2021) 041201, [2005.10832].

[65] M. Butenschoen and B. A. Kniehl, *Constraints on Nonrelativistic-QCD Long-Distance Matrix Elements from $J/\psi$ Plus $W/Z$ Production at the LHC*, *Phys. Rev. Lett.* **130** (2023) 041901, [2207.09366].

[66] M. Butenschoen and B. A. Kniehl, *Dipole subtraction at next-to-leading order in nonrelativistic-QCD factorization*, *Nucl. Phys. B* **950** (2020) 114843, [1909.03698].

[67] M. Butenschoen and B. A. Kniehl, *Dipole Subtraction vs. Phase Space Slicing in NLO NRQCD Heavy-Quarkonium Production Calculations*, *Nucl. Phys. B* **957** (2020) 115056, [2003.01014].

[68] G. T. Bodwin, D. K. Sinclair and S. Kim, *Quarkonium decay matrix elements from quenched lattice QCD*, *Phys. Rev. Lett.* **77** (1996) 2376–2379, [hep-lat/9605023].

[69] N. Brambilla, H. S. Chung and A. Vairo, *Inclusive production of heavy quarkonia in pNRQCD*, *JHEP* **09** (2021) 032, [2106.09417].

[70] E. J. Eichten and C. Quigg, *Quarkonium wave functions at the origin*, *Phys. Rev. D* **52** (1995) 1726–1728, [hep-ph/9503356].

[71] M. L. Mangano and S. J. Parke, *Multiparton amplitudes in gauge theories*, *Phys. Rept.* **200** (1991) 301–367, [hep-th/0509223].

[72] A. Petrelli, M. Cacciari, M. Greco, F. Maltoni and M. L. Mangano, *NLO production and decay of quarkonium*, *Nucl. Phys. B* **514** (1998) 245–309, [hep-ph/9707223].

[73] R. K. Ellis and J. C. Sexton, *QCD Radiative Corrections to Parton Parton Scattering*, *Nucl. Phys. B* **269** (1986) 445–484.

[74] G. C. Nayak, J.-W. Qiu and G. F. Sterman, *Color transfer in associated heavy-quarkonium production*, *Phys. Rev. Lett.* **99** (2007) 212001, [0707.2973].

[75] Z.-G. He, B. A. Kniehl and X.-P. Wang, *Breakdown of Nonrelativistic QCD Factorization in Processes Involving Two Quarkonia and its Cure*, *Phys. Rev. Lett.* **121** (2018) 172001, [1809.07993].

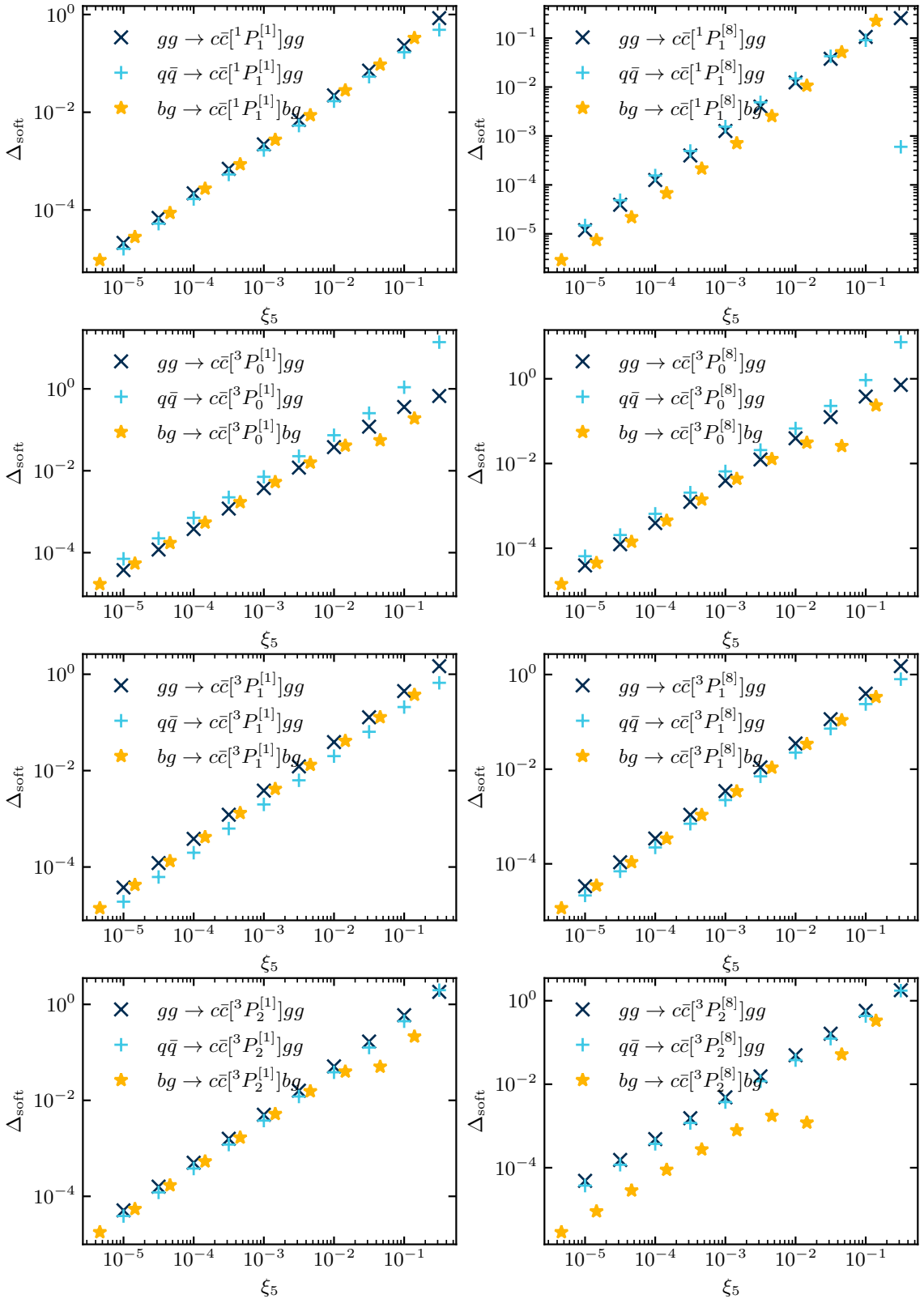



**Figure 1**: Soft limit tests of real processes $gg \to c\bar{c}[n]gg$ (black cross), $q\bar{q} \to c\bar{c}[n]gg$ (blue plus) and $bg \to c\bar{c}[n]bg$ (orange star), where the Fock states $n$ include 8 P-wave states.

arXiv:2402.19208v1 [hep-ph] 29 Feb 2024

# Deciphering the Belle II data on $B \to K\nu\bar{\nu}$ decay in the (dark) SMEFT with minimal flavour violation

**Biao-Feng Hou,**[a] **Xin-Qiang Li,**[a,b] **Meng Shen,**[a] **Ya-Dong Yang**[a,c] **and Xing-Bo Yuan**[a]

[a] *Institute of Particle Physics and Key Laboratory of Quark and Lepton Physics (MOE), Central China Normal University, Wuhan, Hubei 430079, China*

[b] *Center for High Energy Physics, Peking University, Beijing 100871, China*

[c] *Institute of Particle and Nuclear Physics, Henan Normal University, Xinxiang 453007, China*

*E-mail:* resonhou@zknu.edu.cn, xqli@ccnu.edu.cn, shenmeng@mails.ccnu.edu.cn, yangyd@ccnu.edu.cn, y@ccnu.edu.cn

Abstract: Recently, the Belle II collaboration announced the first measurement of the branching ratio $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$, which is found to be about $2.7\sigma$ higher than the Standard Model (SM) prediction. We decipher the data with two new physics scenarios: the underlying quark-level $b \to s\nu\bar{\nu}$ transition is, besides the SM contribution, further affected by heavy new mediators that are much heavier than the electroweak scale, or amended by an additional decay channel with undetected light final states like dark matter or axion-like particles. These two scenarios can be most conveniently analyzed in the SM effective field theory (SMEFT) and the dark SMEFT (DSMEFT) framework, respectively. We consider the flavour structures of the resulting effective operators to be either generic or satisfy the minimal flavour violation (MFV) hypothesis, both for the quark and lepton sectors. In the first scenario, once the MFV assumption is made, only one SM-like low-energy effective operator induced by the SMEFT dimension-six operators can account for the Belle II excess, the parameter space of which is, however, excluded by the Belle upper bound of the branching ratio $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})$. In the second scenario, it is found that the Belle II excess can be accommodated by 22 of the DSMEFT operators involving one or two scalar, fermionic, or vector dark matters as well as axion-like particles. These operators also receive dominant constraints from the $B^0 \to K^{*0} + \text{inv}$ and $B_s \to \text{inv}$ decays. Once the MFV hypothesis is assumed, the number of viable operators is reduced to 14, and the $B^+ \to \pi^+ + \text{inv}$ and $K^+ \to \pi^+ + \text{inv}$ decays start to put further constraints on them. Within the parameter space allowed by all the current experimental data, the $q^2$ distributions of the $B \to K^{(*)} + \text{inv}$ decays are then studied for each viable operator. We find that the resulting prediction of the operator $\mathcal{Q}_{q\chi} = \left(\bar{q}_p\gamma_\mu q_r\right)\left(\bar{\chi}\gamma^\mu\chi\right)$ with a fermionic dark matter mass $m_\chi \approx 700\,\text{MeV}$ can closely match the Belle II event distribution in the bins $2 \le q^2 \le 7\,\text{GeV}^2$. In addition, the future prospects at Belle II, CEPC and FCC-ee are also discussed for some of these FCNC processes.

## Contents


**1 Introduction** **1**

**2 Minimal Flavour Violation** **4**
2.1 Quark sector 4
2.2 Lepton sector 7

**3 $M_1 \to M_2 + \nu\bar{\nu}$** **9**
3.1 SMEFT 9
3.2 Observables 10
3.2.1 $b \to s(d)\nu\bar{\nu}$ decay 10
3.2.2 $b \to s\ell\ell$ decay 12
3.2.3 $s \to d\nu\bar{\nu}$ decay 13

**4 $M_1 \to M_2 +$ DM** **13**
4.1 DSMEFT 13
4.2 DLEFT 16
4.3 Observables 18

**5 Numerical analysis** **19**
5.1 $M_1 \to M_2 + \nu\bar{\nu}$ 20
5.2 $M_1 \to M_2 +$ DM 22
5.2.1 General flavour structure 24
5.2.2 MFV 29

**6 Conclusion** **36**

**A Hadronic matrix elements** **38**
A.1 Decay constants 38
A.2 Transition form factors 38


# 1 Introduction

The flavour-changing neutral current (FCNC) processes are highly suppressed by the Glashow-Iliopoulos-Maiani (GIM) mechanism in the Standard Model (SM) [1], making them sensitive probes of new physics (NP) beyond the SM. In particular, the $B \to K^{(*)}\nu\bar{\nu}$ decays are one of the cleanest channels within the SM and most suitable for indirect NP searches [2–4] (see also refs. [5–8] for recent studies on these

as well as other related decays in the context of NP, but before the recent Belle II announcement [9]).

Recently, using the dataset with an integrated luminosity of $362\,\text{fb}^{-1}$, the Belle II collaboration reported the first evidence of the $B^+ \to K^+ + \nu\bar{\nu}$ decay with a branching ratio [9]

$$\mathcal{B}(B^+ \to K^+\nu\bar{\nu})_{\text{Belle II}} = \left[23 \pm 5(\text{stat})^{+5}_{-4}(\text{syst})\right] \times 10^{-6}, \tag{1.1}$$

which is obtained by combining the inclusive and hadronic tagging results. Interestingly, this measurement shows about $2.7\sigma$ higher than the SM prediction $(4.16 \pm 0.57) \times 10^{-6}$, which is based on the calculations in refs. [2–4] but with updated input parameters. In the future, the Belle II with just $5\,\text{ab}^{-1}$ integrated luminosity is expected to measure this decay with $20 \sim 30\%$ uncertainty relative to the SM value [10, 11], and could therefore confirm or exclude such an excess. Nevertheless, it has already motivated numerous studies of possible NP explanations [12–27].

As an interesting feature, the Belle II result, once the $2\sigma$ range in eq. (1.1) is considered, would be higher than the SM prediction by a factor of $3 \sim 8$. Such a large excess motivates us to ask the following questions: why has such a large NP effect not shown up or will it be found in other $b \to s$ processes, such as the $B \to K^*\nu\bar{\nu}$ and $B_s \to \nu\bar{\nu}$ decays [28, 29]? In this paper, we aim to answer these questions by performing a detailed model-independent analysis of possible NP effects on these and related FCNC processes in the Effective Field Theory (EFT) framework.

The $B^+ \to K^+\nu\bar{\nu}$ excess reported by Belle II could be explained by introducing heavy NP particles that are much heavier than the electroweak scale and hence contribute to the $b \to s\nu\bar{\nu}$ transition as virtual mediators. In this case, the Standard Model Effective Field Theory (SMEFT) provides a well-suited model-independent framework to study such kinds of NP effects [30, 31]. Here, all the NP effects are encoded in the Wilson coefficients of the higher-dimensional operators that are invariant under the SM gauge group. Below the electroweak scale, the heavy SM particles like the massive bosons and the top quark also decouple, and the dynamics is then described by the Low-energy Effective Field Theory (LEFT) [32, 33], which consists of, besides the QCD and QED Lagrangians for light SM fermions, a set of higher-dimensional operators compatible with the QCD and QED gauge symmetries. The Wilson coefficients of these higher-dimensional operators encode all the physics related to the heavy SM and NP degrees of freedom. Starting with the SMEFT, one can also fix these low-energy Wilson coefficients by perform a matching between these two EFTs at the electroweak scale [33–35]. Such a procedure has been extensively applied to various $B$-meson and kaon decays [32, 33]. In this paper, we will interpret the Belle II excess in terms of the SMEFT operators and consider the LEFT as the low-energy limit of SMEFT to calculate the relevant $B$-meson and kaon decays.

The Belle II excess could also be accommodated by amending the $B^+ \to K^+\nu\bar{\nu}$ decay by an additional decay channel with undetected light final states like dark

matter (DM) or axion-like particles (ALP). To mimic the signal, these new light states must couple to the SM quark sector and be sufficiently long-lived to escape the detector. The potential candidates are new light particles that are singlet under the SM gauge group. In order to study the interactions of such new states with the SM quark sector model-independently, it is also convenient to work in an EFT framework, such as the DM EFT [36, 37] without any stabilizing symmetry on the DM fields.[1] In this respect, the SMEFT can be extended by including the scalar, fermionic or vector DM fields, resulting in the so-called Dark SMEFT (DSMEFT) [38–43]. In the DSMEFT, the new operators involving the DM fields are invariant under the SM $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge group and the electroweak symmetry is broken by the usual Higgs mechanism. The general non-redundant operator basis can be found in ref. [42] up to dim= 6, and in ref. [43] up to dim= 8. In order to describe the low-energy processes below the electroweak scale, such as $d_i \to d_j + \text{DM}$ decays, a general framework is the generalization of the LEFT to include the scalar, fermionic and vector DM fields, referred to as the Dark LEFT (DLEFT) [42]. In the DLEFT, the DM particles should be lighter than the electroweak scale and the effective operators are invariant under the $SU(3)_C \otimes U(1)_{\text{em}}$ gauge group. A complete basis of the DLEFT operators can be found in refs. [42, 44] for dim $\leq 6$, and in ref. [45] for dim $\leq 7$. In this work, starting from the DSMEFT and considering the DLEFT as the low-energy limit of DSMEFT, we investigate the parameter space of each DSMEFT operator required to explain the Belle II excess, together under the constraints from the other $b \to s+\text{inv}$ decays. The future prospects at Belle II, CEPC and FCC-ee will also be discussed for some of these FCNC processes. In addition, for completeness, our analyses also include the ALP case [46–49] in the EFT framework [50–55].

The above two NP scenarios proposed to account for the Belle II excess could also affect the quark-level $b \to d$ and $s \to d$ processes. In order to correlate the $b \to s$ to the $b \to d$ and $s \to d$ transitions, we have to specify the underlying flavour structures of the EFT considered [56, 57]. To this end, we assume the minimal flavour violation (MFV) hypothesis [58–60] and implement it into the SMEFT and DSMEFT frameworks. According to the hypothesis, all the flavour violating currents are controlled by the SM Yukawa couplings, so that all the FCNC interactions in the quark sector are naturally suppressed by the CKM factors. Therefore, any potential large FCNCs in the SMEFT and DSMEFT can be avoided. Considering the current experimental data on the $b \to s + \text{inv}$, $b \to d + \text{inv}$ and $s \to d + \text{inv}$ processes, such as the $B \to \pi + \text{inv}$ and $K \to \pi + \text{inv}$ decays, we derive the constraints on each operator in the SMEFT and DSMEFT with the MFV hypothesis. Implications of the resulting parameter space of each viable operator for the differential distributions of these decays are also investigated in detail, which can be tested at the Belle II

[1] In this work, we do not impose any symmetry to stabilize the DM fields in the EFT framework. Thus, the DM fields we are considering are actually just new singlets under the SM gauge group.

with more statistics.

The remainder of this paper is organized as follows. In section 2, we introduce the MFV hypothesis. In sections 3 and 4, we investigate the various $d_i \to d_j \nu\bar{\nu}$ and $d_i \to d_j + \text{DM}$ decays in the SMEFT and DSMEFT, respectively. In these EFT frameworks, the flavour structures of the effective operators are taken to be either of the most general form or satisfy the MFV hypothesis. We present our detailed numerical analyses and discussions in section 5. Our conclusion is finally made in section 6. The hadronic matrix elements involved throughout this work, such as the transition form factors and decay constants, are summarized in appendix A.

## 2 Minimal Flavour Violation

In this section, we discuss the idea of MFV hypothesis and its application in the SMEFT and DSMEFT.

### 2.1 Quark sector

In the interaction eigenbasis, the SM Yukawa interactions in the quark sector are described by the Lagrangian

$$-\mathcal{L}_Y = \bar{q}\, Y_d H d + \bar{q}\, Y_u \tilde{H} u + \text{h.c.}, \tag{2.1}$$

where $\tilde{H} \equiv i\sigma_2 H^*$ with $H$ being the SM Higgs doublet. The fields $q$ and $u, d$ denote the left-handed quark doublet and the right-handed quark singlets in the SM, respectively. The Yukawa coupling matrices $Y_{u,d}$ are $3 \times 3$ complex matrices in flavour space. In the SM, these Yukawa interactions violate the global flavour symmetry [58]

$$G_{\text{QF}} = SU(3)_q \otimes SU(3)_u \otimes SU(3)_d. \tag{2.2}$$

We can formally recover this flavour symmetry by promoting the Yukawa matrices $Y_{u,d}$ to spurion fields [60]. The Yukawa spurions should transform under $G_{\text{QF}}$ as

$$Y_u \sim \left(\mathbf{3}, \bar{\mathbf{3}}, \mathbf{1}\right), \qquad Y_d \sim \left(\mathbf{3}, \mathbf{1}, \bar{\mathbf{3}}\right). \tag{2.3}$$

Then, we can construct two useful basic building blocks $\mathsf{A} \equiv Y_u Y_u^\dagger$ and $\mathsf{B} \equiv Y_d Y_d^\dagger$, which transform as $(\mathbf{1} \oplus \mathbf{8}, \mathbf{1}, \mathbf{1})$ under the flavour group $G_{\text{QF}}$. As will be seen later, these building blocks are useful to parameterize the quark flavour transitions in effective theories.

In the SMEFT and DSMEFT, there are four types of quark currents relevant for the down-type FCNC transitions, which read

$$\bar{q}\gamma^\mu \mathcal{C} q, \qquad \bar{d}\gamma^\mu \mathcal{C} d, \qquad \bar{q}\mathcal{C} d, \qquad \bar{q}\sigma^{\mu\nu}\mathcal{C} d, \tag{2.4}$$

where $\mathcal{C}$ denotes the Wilson coefficient of the operator $\mathcal{Q}$ and is a $3 \times 3$ matrix in quark flavour space. Here and in the following, we suppress the operator index $i$,

i.e., $\mathcal{C}$ denoting $\mathcal{C}_i$. In the MFV hypothesis, it is assumed that at low energy the Yukawa couplings are the only source of flavour violation in the SM and beyond [60]. Technically, this hypothesis implies that the quark currents in eq. (2.4) should be invariant under the flavour group $G_{\rm QF}$. Therefore, the Wilson coefficient $\mathcal{C}$ can be written in the following form:

$$\mathcal{C}^{\rm MFV} = \begin{cases} f(\mathsf{A},\mathsf{B}) & \text{for } \bar{q}\gamma^\mu\mathcal{C}q, \\ f(\mathsf{A},\mathsf{B})Y_d & \text{for } \bar{q}\mathcal{C}d,\ \bar{q}\sigma^{\mu\nu}\mathcal{C}d, \\ \epsilon_0\mathbb{1} + Y_d^\dagger g(\mathsf{A},\mathsf{B})Y_d & \text{for } \bar{d}\gamma^\mu\mathcal{C}d, \end{cases} \tag{2.5}$$

where $\epsilon_0$ is a real constant. The functions $f(\mathsf{A},\mathsf{B})$ and $g(\mathsf{A},\mathsf{B})$ are infinite series of the building blocks $\mathsf{A}$ and $\mathsf{B}$, e.g., explicitly $f(\mathsf{A},\mathsf{B}) \equiv \xi_{ijk\cdots}\mathsf{A}^i\mathsf{B}^j\mathsf{A}^k\cdots$ with real coefficients $\xi_{ijk\cdots}$ to ensure no new sources of CP violation beyond the Yukawa couplings. By application of the Cayley-Hamilton identity, the infinite series in $f(\mathsf{A},\mathsf{B})$ and $g(\mathsf{A},\mathsf{B})$ can be resummed into a finite number of terms. Taking $f(\mathsf{A},\mathsf{B})$ as an example, its explicit summed expression can be written as [61–63]

$$\begin{aligned} f(\mathsf{A},\mathsf{B}) = \epsilon_0\mathbb{1} &+ \epsilon_1\mathsf{A} + \epsilon_3\mathsf{A}^2 + \epsilon_5\mathsf{AB} + \epsilon_7\mathsf{ABA} + \epsilon_{10}\mathsf{AB}^2 + \epsilon_{12}\mathsf{A}^2\mathsf{B}^2 + \epsilon_{14}\mathsf{B}^2\mathsf{AB} + \epsilon_{15}\mathsf{AB}^2\mathsf{A}^2 \\ &+ \epsilon_2\mathsf{B} + \epsilon_4\mathsf{B}^2 + \epsilon_6\mathsf{BA} + \epsilon_9\mathsf{BAB} + \epsilon_8\mathsf{BA}^2 + \epsilon_{13}\mathsf{B}^2\mathsf{A}^2 + \epsilon_{11}\mathsf{ABA}^2 + \epsilon_{16}\mathsf{B}^2\mathsf{A}^2\mathsf{B}. \end{aligned}$$

During the resummation, higher-order polynomials of $\mathsf{A}$ and $\mathsf{B}$ contribute to the coefficients $\epsilon_i$, and make them complex generally. However, it is found that the imaginary parts of the coefficients are tiny (e.g., ${\rm Im}\epsilon_i \propto |{\rm Tr}(\mathsf{A}^2\mathsf{BAB}^2)| \ll 1$) and can be therefore neglected in the numerical analysis [61, 62, 64–66]. In the following, we assume the coefficients $\epsilon_i$ are of the same orders of magnitude. Therefore, only the terms built by $\mathsf{A}$ are kept, while all the terms involving $\mathsf{B}$ are neglected, as the spurion $\mathsf{B}$ is suppressed by $\mathcal{O}(\hat{\lambda}_d^2)$, where $\hat{\lambda}_d$ denotes the diagonal Yukawa coupling matrix for down-type quarks defined by eq. (2.8). Finally, we obtain the approximation [67]

$$f(\mathsf{A},\mathsf{B}) \approx \epsilon_0\mathbb{1} + \epsilon_1\mathsf{A} + \epsilon_2\mathsf{A}^2, \tag{2.6}$$

where the coefficients $\epsilon_{0,1,2}$ are free real parameters. Similarly, being suppressed by $\mathcal{O}(\hat{\lambda}_d^2)$, the approximation $Y_d^\dagger g(\mathsf{A},\mathsf{B})Y_d \approx 0$ can be taken.

In the MFV framework discussed above, the quark currents are all given in the interaction eigenbasis. In order to rotate to the fermion mass eigenbasis, the following transforms should be performed:

$$u_L \to U_u u_L, \qquad u_R \to W_u u_R, \qquad d_L \to U_d d_L, \qquad d_R \to W_d d_R, \tag{2.7}$$

where $U_{u,d}$ and $W_{u,d}$ are $3\times 3$ unitary matrices introduced to diagonalize the Yukawa matrices

$$U_d^\dagger Y_d W_d = \hat{\lambda}_d, \qquad U_u^\dagger Y_u W_u = \hat{\lambda}_u. \tag{2.8}$$

Therefore, the CKM matrix is given by $V = U_u^\dagger U_d$. In the following, we turn to the fermion mass eigenbasis and derive explicit expressions of the MFV couplings.

The current $\bar{q}\gamma^\mu \mathcal{C} q$ involves the following down-type quark interactions in the mass eigenbasis:

$$\bar{d}_L U_d^\dagger \gamma_\mu \mathcal{C}^{\rm MFV} U_d d_L \approx \bar{d}_L \gamma^\mu V^\dagger \big(\epsilon_0 + \epsilon_1 \hat{\lambda}_u^2 + \epsilon_2 \hat{\lambda}_u^4\big) V d_L, \tag{2.9}$$

where $d_L$ denotes now the left-handed down-type quark in the mass eigenbasis, and $\hat{\lambda}_u$ denotes the diagonal Yukawa coupling matrix for up-type quarks, whose diagonal elements are given by $\lambda_{u_i} = \sqrt{2} m_{u_i}/v$ with the vacuum expectation value $v = 246\,\text{GeV}$. It is noted that, due to the large hierarchy among the diagonal elements $\lambda_{u_i}$, the matrices $\epsilon_1 \hat{\lambda}_u^2$ and $\epsilon_2 \hat{\lambda}_u^4$ have almost the same structures. Therefore, one can obtain the following approximation [68]:

$$V^\dagger (\epsilon_1 \hat{\lambda}_u^2 + \epsilon_2 \hat{\lambda}_u^4) V \approx V^\dagger (\epsilon_1 \hat{\lambda}_u^2) V \equiv \epsilon_1 \Delta_q, \tag{2.10}$$

which is equivalent to taking the approximation $\lambda_t^4 - \lambda_{u,c}^4 \approx \lambda_t^2 (\lambda_t^2 - \lambda_{u,c}^2)$ and redefining $\epsilon_1 + \epsilon_2 \lambda_t^2 \to \epsilon_1$.[2] For the observables investigated in this paper, we have verified that the numerical differences caused by this approximation are negligible. Finally, by defining the basic FCNC couplings

$$\Delta_q = V^\dagger \hat{\lambda}_u^2 V, \tag{2.11}$$

we arrive at

$$\bar{d}_L U_d^\dagger \gamma_\mu \mathcal{C}^{\rm MFV} U_d d_L \approx \bar{d}_L \gamma^\mu \big(\epsilon_0 \mathbb{1} + \epsilon_1 \Delta_q\big) d_L. \tag{2.12}$$

Similarly, from the current $\bar{q}\mathcal{C} d$, the following scalar interactions among the down-type quarks are obtained in the mass eigenbasis:

$$\bar{d}_L U_d^\dagger \mathcal{C}^{\rm MFV} W_d d_R \approx \bar{d}_L V^\dagger \big(\epsilon_0 \mathbb{1} + \epsilon_1 \hat{\lambda}_u^2 + \epsilon_2 \hat{\lambda}_u^4\big) V \hat{\lambda}_d d_R \approx \bar{d}_L \big(\epsilon_0 \hat{\lambda}_d + \epsilon_1 \Delta_q \hat{\lambda}_d\big) d_R, \tag{2.13}$$

where $d_R$ is the right-handed down-type quark in the mass eigenbasis, and the diagonal elements of the diagonal Yukawa coupling matrix for down-type quarks are given by $\lambda_d^i = \sqrt{2} m_{d_i}/v$. Here the approximation in eq. (2.10) has been used. In addition, due to the same flavour structure as of the scalar interactions, the tensor interactions $\bar{d}_L U_d^\dagger \sigma^{\mu\nu} \mathcal{C}^{\rm MFV} W_d d_R$ resulting from the current $\bar{q}\sigma^{\mu\nu}\mathcal{C}^{\rm MFV} d$ have the same MFV coupling as in eq. (2.13).

For the vector interactions originating from the current $\bar{d}\gamma^\mu \mathcal{C} d$, the MFV coupling in the mass eigenbasis is given by

$$\bar{d}_R W_d^\dagger \gamma^\mu \mathcal{C}^{\rm MFV} W_d d_R \approx \bar{d}_R \gamma^\mu (\epsilon_0 \mathbb{1}) d_R. \tag{2.14}$$

[2] Using the unitarity of the CKM matrix, one can obtain $[V^\dagger(\epsilon_1 \hat{\lambda}_u^2 + \epsilon_2 \hat{\lambda}_u^4) V]_{ij} = \lambda_t^2(\epsilon_1 + \epsilon_2 \lambda_t^2)\delta_{ij} - (\lambda_t^2 - \lambda_u^2)[\epsilon_1 + \epsilon_2(\lambda_t^2 + \lambda_u^2)] V_{ui}^* V_{uj} - (\lambda_t^2 - \lambda_c^2)[\epsilon_1 + \epsilon_2(\lambda_t^2 + \lambda_c^2)] V_{ci}^* V_{cj} \approx \lambda_t^2(\epsilon_1 + \epsilon_2 \lambda_t^2)\delta_{ij} - (\lambda_t^2 - \lambda_u^2)[\epsilon_1 + \epsilon_2 \lambda_t^2] V_{ui}^* V_{uj} - (\lambda_t^2 - \lambda_c^2)[\epsilon_1 + \epsilon_2 \lambda_t^2] V_{ci}^* V_{cj} = (\epsilon_1 + \epsilon_2 \lambda_t^2)[V^\dagger \hat{\lambda}_u^2 V]_{ij}$, where the approximation $\lambda_t^2 + \lambda_{u,c}^2 \approx \lambda_t^2$ have been used.

This implies that the vector interactions among the right-handed down-type quarks are approximately flavour diagonal and universal. Thus, the FCNC vector interactions among the right-handed down-type quarks are forbidden in the MFV scenario.

As a summary, the MFV couplings for various down-type quark currents are given in the mass eigenbasis by

$$\mathcal{C}_i^{\rm MFV} = \begin{cases} \epsilon_0^i \mathbb{1} + \epsilon_1^i \Delta_q & \text{for } \bar{d}_L \gamma^\mu \mathcal{C}_i d_L, \\ \epsilon_0^i \hat{\lambda}_d + \epsilon_1^i \Delta_q \hat{\lambda}_d & \text{for } \bar{d}_L \mathcal{C}_i d_R,\ \bar{d}_L \sigma^{\mu\nu} \mathcal{C}_i d_R, \\ \epsilon_0^i \mathbb{1} & \text{for } \bar{d}_R \gamma^\mu \mathcal{C}_i d_R. \end{cases} \tag{2.15}$$

It is then clear that only two types of MFV couplings, $\Delta_q$ and $\Delta_q \hat{\lambda}_d$, can generate FCNC interactions among the down-type quarks and their numerical values are given, respectively, by

$$\Delta_q = \begin{pmatrix} 0.8 & -3.3 - 1.5i & 79.3 + 35.4i \\ -3.3 + 1.5i & 16.6 & -397.5 + 8.1i \\ 79.3 - 35.4i & -397.5 - 8.1i & 9839.0 \end{pmatrix} \times 10^{-4}, \tag{2.16}$$

and

$$\Delta_q \hat{\lambda}_d = \begin{pmatrix} 0.0021 & -0.18 - 0.08i & 191.3 + 85.4i \\ -0.009 + 0.004i & 0.88 & -958.7 + 19.6i \\ 0.21 - 0.10i & -21.1 - 0.4i & 23728.1 \end{pmatrix} \times 10^{-6}. \tag{2.17}$$

One can see that a large hierarchy exists among the down-type FCNC interactions, i.e., $\mathcal{C}_{bs}^{\rm MFV} \gg \mathcal{C}_{bd}^{\rm MFV} \gg \mathcal{C}_{sd}^{\rm MFV}$. It is also seen that the flavour-conserving couplings $\mathcal{C}_{bb}^{\rm MFV}$ is more than one order of magnitude larger than all the other couplings.

### 2.2 Lepton sector

One can also apply the MFV hypothesis to the lepton sector. However, since the mechanism of neutrino mass generation is still unknown, there are different approaches to formulate the leptonic MFV [69–75]. Here, we consider the realization of leptonic MFV within the so-called minimal field content [69, 70], in which the neutrino masses are generated by the Weinberg operator. In this case, the Yukawa interactions in the lepton sector can be written as

$$-\Delta\mathcal{L} = \bar{e}\, Y_e H^\dagger l + \frac{1}{2\Lambda_{\rm LN}} \big(\bar{l}^c \tau_2 H\big) Y_\nu \big(H^T \tau_2 l\big) + \text{h.c.}, \tag{2.18}$$

where $l$ denotes the left-handed lepton doublet with the charge conjugated field given by $l^c = -i\gamma_2 l^*$, and $e$ is the right-handed charged lepton singlet. $\Lambda_{\rm LN}$ denotes the breaking scale of the lepton number symmetry $U(1)_{\rm LN}$. $Y_e$ and $Y_\nu$ stand for the $3 \times 3$ Yukawa coupling matrices in flavour space. In the absence of these Yukawa couplings, the lepton sector respects the flavour symmetry

$$G_{\rm LF} = SU(3)_l \otimes SU(3)_e. \tag{2.19}$$

Analogous to the quark sector, this flavour symmetry can be restored by promoting the Yukawa couplings $Y_{e,\nu}$ to spurion fields. Then, the relevant building blocks are $\mathsf{A}_\ell = Y_\nu^\dagger Y_\nu$ and $\mathsf{B}_\ell = Y_e^\dagger Y_e$, which transform as $(\mathbf{1} \oplus \mathbf{8}, \mathbf{1})$ under the lepton flavour group $G_{\rm LF}$.

For the three SMEFT operators, $\mathcal{Q}_{lq}^{(1,3)}$ and $\mathcal{Q}_{ld}$, relevant to the $B^+ \to K^+\nu\bar{\nu}$ decay, they contain the lepton current

$$\bar{l}\gamma^\mu \mathcal{C} l, \tag{2.20}$$

where $\mathcal{C}$ denotes the Wilson coefficient of the operator $\mathcal{Q}$ and is a $3 \times 3$ matrix in lepton flavour space. In the MFV hypothesis, analogous to the quark current, the Wilson coefficient $\mathcal{C}$ should take the form $\mathcal{C}_{\rm MFV} = h(\mathsf{A}_\ell, \mathsf{B}_\ell)$, where $h(\mathsf{A}_\ell, \mathsf{B}_\ell)$ is a finite polynomial of $\mathsf{A}_\ell$ and $\mathsf{B}_\ell$. After neglecting all the terms involving $\mathsf{B}_\ell$, which are suppressed by the small lepton Yukawa couplings $Y_e$, we obtain

$$\mathcal{C}_{\rm MFV} \approx \kappa_0 + \kappa_1 \mathsf{A}_\ell + \kappa_2 \mathsf{A}_\ell^2, \tag{2.21}$$

where the coefficients $\kappa_{0,1,2}$ are free real parameters. In the numerical analysis, we keep only the leading lepton flavour violation term $\mathsf{A}_\ell$ for simplicity, i.e., $\kappa_2 = 0$. Turning to the lepton mass eigenbasis, the current $\bar{l}\gamma^\mu \mathcal{C} l$ gives in the MFV hypothesis the following interactions:

$$\bar{e}_L \gamma^\mu (\kappa_0 \mathbb{1} + \kappa_0 \Delta_\ell) e_L + \bar{\nu}_L \gamma^\mu (\kappa_0 \mathbb{1} + \kappa_0 \hat{\lambda}_\nu^2) \nu_L, \tag{2.22}$$

where the basic LFV coupling $\Delta_\ell$ can be obtained from $\mathsf{A}_\ell$ and takes the form

$$\Delta_\ell = U \hat{\lambda}_\nu^2 U^\dagger, \tag{2.23}$$

where $U$ is the PMNS matrix. Here $e_L$ and $\nu_L$ denote the left-handed charged lepton and neutrino in the mass eigenbasis, respectively. $\hat{\lambda}_\nu \equiv 2 m_\nu \Lambda_{\rm LN}/v^2$ stands for the diagonal effective neutrino Yukawa coupling and $m_\nu$ the diagonal neutrino mass matrix. As can be seen from eq. (2.22), the lepton vector current does not induce neutral LFV in the leptonic MFV with minimal field content.[3]

As discussed above, the charged LFV interactions are governed by the MFV coupling $\Delta_\ell$. Numerically, we obtain

$$\Delta_\ell^{\rm NO} = \begin{pmatrix} -0.19 - 0.01i & -0.25 - 0.02i & 0.31 - 0.04i \\ 0.12 + 0.01i & 0.28 - 0.00i & 0.29 + 0.04i \\ -0.37 - 0.01i & 0.21 - 0.05i & -0.03 + 0.01i \end{pmatrix}, \tag{2.24}$$

[3]This is, however, not true for the most generic leptonic MFV. For example, in the leptonic MFV with type-I seesaw mechanism, the basic LFV coupling $\Delta_\ell = U\hat{\lambda}_\nu U^\dagger$ should be replaced by $U\hat{\lambda}_\nu^{1/2} O O^\dagger \hat{\lambda}_\nu^{1/2} U^\dagger$, where $O$ is a general complex orthogonal matrix satisfying $OO^T = \mathbb{1}$ [67]. Therefor, any non-diagonal matrix $OO^\dagger$ can induce neutral LFV.

for normal ordering (NO), and

$$\Delta_\ell^{\rm IO} = \begin{pmatrix} 0.21+0.09i & -0.34+0.05i & 0.03+0.11i \\ 0.31+0.12i & 0.19+0.00i & -0.15-0.14i \\ 0.12-0.02i & 0.04-0.19i & 0.34-0.10i \end{pmatrix}, \tag{2.25}$$

for inverted ordering (IO) of neutrino masses. Here $\Lambda_{\rm LN} = 10^{14}\,\text{GeV}$ and the lightest neutrino mass $m_{\nu_1} = 0.2\,\text{eV}$ have been taken. One can see that the magnitudes of all the matrix elements of $\Delta_\ell$ are of similar size except for $(\Delta_\ell^{\rm NO})_{33}$, which has much smaller magnitude than the other couplings.

Let us finally make a comment on the DSMEFT with MFV. In general, similar to the SM fermions, the DM particles can also appear in a form of multiple generations. In this case, the SM flavour symmetry could be extended to include the global symmetry associated with the DM multiplet and the MFV hypothesis can be generalized to the DM sector [76–80]. In this paper, we focus on a minimal DM sector, which contains only one or two DM generations, and leave a detained study of the DM with MFV for future work.

## 3 $M_1 \to M_2 + \nu\bar{\nu}$

In this section, we first summarize the SMEFT framework with the most general flavour structure as well as in the MFV hypothesis. Then, the NP effects on various rare FCNC decays are discussed.

### 3.1 SMEFT

If new particles are much heavier than the electroweak scale and the electroweak symmetry breaking is realized linearly, their effects can be parameterized by a series of higher dimensional gauge-invariant operators in the SMEFT (see refs. [30, 31] for recent reviews). The SMEFT Lagrangian at dim-6 takes the form [81, 82]

$$\mathcal{L}_{\rm SMEFT} = \sum_i \frac{\mathcal{C}_i}{\Lambda^2}\mathcal{Q}_i, \tag{3.1}$$

where $\Lambda$ denotes the NP scale. The dim-6 effective operators $\mathcal{Q}_i$ are built out of the SM fields and respect the SM gauge symmetry $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, with $\mathcal{C}_i$ the corresponding dimensionless Wilson coefficients.

For the $b \to s\nu\bar{\nu}$ transition, the relevant operators are given by

$$\begin{aligned} \mathcal{Q}_{lq}^{(1)} &= \left(\bar{l}_p\gamma^\mu l_r\right)\left(\bar{q}_s\gamma_\mu q_t\right), & \mathcal{Q}_{Hq}^{(1)} &= \left(H^\dagger i\overleftrightarrow{D}_\mu H\right)\left(\bar{q}_p\gamma^\mu q_r\right), \\ \mathcal{Q}_{lq}^{(3)} &= \left(\bar{l}_p\gamma^\mu\tau^I l_r\right)\left(\bar{q}_s\tau^I\gamma_\mu q_t\right), & \mathcal{Q}_{Hq}^{(3)} &= \left(H^\dagger i\overleftrightarrow{D}^I_\mu H\right)\left(\bar{q}_p\tau^I\gamma^\mu q_r\right), \\ \mathcal{Q}_{ld} &= \left(\bar{l}_p\gamma^\mu l_r\right)\left(\bar{d}_s\gamma_\mu d_t\right), & \mathcal{Q}_{Hd} &= \left(H^\dagger i\overleftrightarrow{D}_\mu H\right)\left(\bar{d}_p\gamma^\mu d_r\right), \end{aligned} \tag{3.2}$$

where $D_\mu = \partial_\mu + ig_s T^a G^a_\mu + ig_2 t^I W^I_\mu + ig_1 Y B_\mu$ denotes the covariant derivative under the SM gauge symmetry. Here $T^a = \lambda^a/2$ and $t^I = \tau^I/2$ are the $SU(3)$ and $SU(2)$ generators, where $\lambda^a$ and $\tau^I$ stand for the Gell-Mann and Pauli matrices, respectively. $l$ and $q$ denote the left-handed lepton and quark doublets respectively, while $d$ is the right-handed down-type quark singlet, all being given in the mass eigenbasis. The generation indices of the SM fields are characterized by $p$, $r$, $s$, $t$.

The operators $\mathcal{Q}^{(1)}_{Hq}$, $\mathcal{Q}^{(3)}_{Hq}$ and $\mathcal{Q}_{Hd}$ can induce a tree-level $\bar{s}\gamma^\mu b Z_\mu$ coupling, and thus contribute to the $b \to s\nu\bar{\nu}$ transition. However, the same $\bar{s}\gamma^\mu b Z_\mu$ interaction also affects the $b \to s\ell^+\ell^-$ processes [2, 3]. Due to the $SU(2)_L$ gauge invariance, its contributions to these two processes are of equal magnitudes and lepton flavour universal. Thus, it is impossible to explain the Belle II data on the branching ratio $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ while satisfying the stringent constraints from the precisely measured $b \to s\ell^+\ell^-$ observables, e.g., the branching ratio $\mathcal{B}(B_s \to \mu^+\mu^-)$. Explicitly, it is found that $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ can only be enhanced by about 20% after considering other relevant experimental constraints [14]. Therefore, we will not consider these operators and focus on the remaining three four-fermion operators in the following.

In the MFV hypothesis, by using eqs. (2.12), (2.14) and (2.22), the relevant four-fermion operators take the following forms:

$$\mathcal{C}_{ld}\mathcal{Q}_{ld} = \Big[\bar{e}_L\gamma_\mu\left(\kappa_0\mathbb{1} + \kappa_1\Delta_\ell\right)e_L + \bar{\nu}_L\gamma_\mu\big(\kappa_0\mathbb{1} + \kappa_1\hat{\lambda}^2_\nu\big)\nu_L\Big]\Big[\bar{d}_R\gamma^\mu(\epsilon_0\mathbb{1})d_R\Big], \tag{3.3}$$

$$\mathcal{C}^{(1,3)}_{lq}\mathcal{Q}^{(1,3)}_{lq} \supset \Big[\bar{e}_L\gamma_\mu\left(\kappa_0\mathbb{1} + \kappa_1\Delta_\ell\right)e_L \pm \bar{\nu}_L\gamma_\mu\big(\kappa_0\mathbb{1} + \kappa_1\hat{\lambda}^2_\nu\big)\nu_L\Big]\Big[\bar{d}_L\gamma^\mu\left(\epsilon_0\mathbb{1} + \epsilon_1\Delta_q\right)d_L\Big],$$

in the mass eigenbasis. Here we have suppressed the operator indices for the MFV parameters $\epsilon_{0,1}$ and $\kappa_{0,1}$ are suppressed. It is noted that $\mathcal{Q}_{ld}$ does not involve any quark FCNC interactions, and thus has no contribution to the $b \to s\nu\bar{\nu}$ transition.

### 3.2 Observables

#### 3.2.1 $b \to s(d)\nu\bar{\nu}$ decay

The $B \to K^{(*)}\nu\bar{\nu}$ decays are induced by the $b \to s\nu_i\bar{\nu}_j$ transitions. In the most general case, the effective weak Hamiltonian contains only two operators and can be written as [2, 3]

$$\mathcal{H}^{b\to s\nu\nu}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}}\frac{\alpha_e}{4\pi}V_{tb}V^*_{ts}\sum_{i,j=e,\mu,\tau}\left(C^{\nu_i\nu_j}_L\mathcal{O}^{\nu_i\nu_j}_L + C^{\nu_i\nu_j}_R\mathcal{O}^{\nu_i\nu_j}_R\right) + \text{h.c.}, \tag{3.4}$$

with the effective operators defined by

$$\mathcal{O}^{\nu_i\nu_j}_L = \big(\bar{s}\gamma_\mu P_L b\big)\big(\bar{\nu}_i\gamma^\mu P_L\nu_j\big), \qquad \mathcal{O}^{\nu_i\nu_j}_R = \big(\bar{s}\gamma_\mu P_R b\big)\big(\bar{\nu}_i\gamma^\mu P_L\nu_j\big). \tag{3.5}$$

In the SM, $C^{\nu_i\nu_j}_R$ is negligible and $C^{\nu_i\nu_j}_L = C^{\text{SM}}_L\delta_{ij}$ with $C^{\text{SM}}_L = -12.64 \pm 0.14$ [3] after including the NLO QCD corrections [83–85] and the two-loop electroweak contributions [86]. In the SMEFT, by comparing eqs. (3.1) and (3.4), one can derive that

the NP contributions to these two operators are given, respectively, by

$$C^{\nu_i\nu_j}_{L,\mathrm{NP}} = c_\nu\big(\big[\mathcal{C}^{(1)}_{lq}\big]_{ij23} - \big[\mathcal{C}^{(3)}_{lq}\big]_{ij23}\big), \qquad C^{\nu_i\nu_j}_{R,\mathrm{NP}} = c_\nu\big[\mathcal{C}_{ld}\big]_{ij23}, \tag{3.6}$$

with the normalization factor $c_\nu \equiv (2\pi/(\alpha_e V_{tb}V^*_{ts}))\cdot(v^2/\Lambda^2)$. In the MFV framework, by using eq. (3.3), we find that

$$\begin{aligned} C^{\nu_i\nu_j}_{L,\mathrm{MFV}} &= c_\nu \Delta^{23}_q\big[\epsilon^{(1)}_1\big(\kappa^{(1)}_0 + \kappa^{(1)}_1\lambda^2_{\nu_i}\big) - \big((1)\to(3)\big)\big]\delta_{ij}, \\ C^{\nu_i\nu_j}_{R,\mathrm{MFV}} &= 0, \end{aligned} \tag{3.7}$$

where $\epsilon^{(n)}_1$ and $\kappa^{(n)}_{0,1}$ denote the MFV parameters associated with the operator $\mathcal{Q}^{(n)}_{lq}$, with $n = 1, 3$. Thus, one can see that the transition $b \to s\nu\bar{\nu}$ in the MFV hypothesis is only induced by the effective operator $\mathcal{O}^{\nu_i\nu_i}_L$, which also conserves the lepton flavour. In addition, it is straightforward to obtain the operators and Wilson coefficients for the $b \to d\nu_i\bar{\nu}_j$ transitions by changing the corresponding flavour indices.

Starting with the effective weak Hamiltonian in eq. (3.4), we can write the dineutrino invariant mass spectrum of the $B \to K^*\nu\bar{\nu}$ decay as [2, 3]

$$\frac{d\Gamma(B\to K^*\nu\bar{\nu})}{ds_B} = m^2_B \sum_{i,j=e,\mu,\tau}\Big(\big|A^{ij}_\perp(s_B)\big|^2 + \big|A^{ij}_\parallel(s_B)\big|^2 + \big|A^{ij}_0(s_B)\big|^2\Big), \tag{3.8}$$

where $s_B = q^2/m^2_B$, and $q^2$ denotes the dineutrino invariant mass squared in the $B$-meson rest frame. The three polarization amplitudes $A^{ij}_{\perp,\parallel,0}$ are functions of the $B \to K^*$ transition form factors and the Wilson coefficients, and take the form

$$\begin{aligned} A^{ij}_\perp(s_B) &= 2\sqrt{2}N\,\lambda^{1/2}(1,\tilde{m}^2_{K^*},s_B)\,(C^{\nu_i\nu_j}_L + C^{\nu_i\nu_j}_R)\frac{V(s_B)}{1+\tilde{m}_{K^*}}, \\ A^{ij}_\parallel(s_B) &= -2\sqrt{2}N\,(1+\tilde{m}_{K^*})\,(C^{\nu_i\nu_j}_L - C^{\nu_i\nu_j}_R)\,A_1(s_B), \\ A^{ij}_0(s_B) &= -\frac{N(C^{\nu_i\nu_j}_L - C^{\nu_i\nu_j}_R)}{\tilde{m}_{K^*}\sqrt{s_B}}\Bigg[(1-\tilde{m}^2_{K^*}-s_B)(1+\tilde{m}_{K^*})A_1(s_B) \\ &\qquad - \lambda(1,\tilde{m}^2_{K^*},s_B)\frac{A_2(s_B)}{1+\tilde{m}_{K^*}}\Bigg], \end{aligned} \tag{3.9}$$

with $\tilde{m}_{K^*} = m_{K^*}/m_B$ and $N = 1/2\,V_{tb}V^*_{ts}\left[G^2_F\alpha^2_e m^3_B s_B\lambda^{1/2}(1,\tilde{m}^2_{K^*},s_B)/(3\cdot 2^{10}\pi^5)\right]^{1/2}$. Similarly, the dineutrino invariant mass distribution of the $B \to K\nu\bar{\nu}$ decay can be written as [2, 3, 87]

$$\begin{aligned} \frac{d\Gamma(B\to K\nu\bar{\nu})}{ds_B} = \frac{G^2_F\alpha^2_e}{3\cdot 2^{10}\pi^5}|V_{tb}V^*_{ts}|^2 m^5_B\lambda^{3/2}\left(1,\tilde{m}^2_K,s_B\right)\left[f^{B\to K}_+(s_B)\right]^2 \\ \times \sum_{i,j=e,\mu,\tau}\left|C^{\nu_i\nu_j}_L + C^{\nu_i\nu_j}_R\right|^2. \end{aligned} \tag{3.10}$$

For convenience, all the relevant form factors present in the above expressions are given in appendix A.2. In addition, it is straightforward to obtain the distributions

of other $b \to s(d)\nu\bar{\nu}$ decays, such as $B \to \pi\nu\bar{\nu}$ and $B_s \to \phi\nu\bar{\nu}$, from the above expressions with some replacements. Finally, one should notice that for charged $B$-meson decays, there is a long-distance contribution arising from the tree-level weak annihilation mediated by the on-shell $\tau$ lepton [88]. For example, the $B^+ \to \tau^+(\to K^+\bar{\nu})\nu$ process provides an additional $\mathcal{O}(10\%)$ contribution to the $B^+ \to K^+\nu\bar{\nu}$ decay, and constitutes a background in the experimental extraction of the branching ratio [9]. Here we do not include the tree-level contribution, which has already been subtracted away during the experimental data analysis [9].

#### 3.2.2 $b \to s\ell\ell$ decay

The quark-level $b \to s\ell^+\ell^-$ transitions, such as the rare $B_s \to \ell^+\ell^-$, $B \to K^{(*)}\ell^+\ell^-$ and $B_s \to \phi\ell^+\ell^-$ decays, can provide promising probes of various NP effects (see refs. [89, 90] for a recent review). For a general $b \to s\ell_i^-\ell_j^+$ transition, the effective weak Hamiltonian can be written as [32]

$$\mathcal{H}_{\text{eff}}^{b\to s\ell\ell} = -\frac{4G_F}{\sqrt{2}}\frac{e^2}{16\pi^2}V_{tb}V_{ts}^*\sum_k C_k\mathcal{O}_k + \text{h.c.}, \tag{3.11}$$

where the most relevant operators are $\mathcal{O}_9^{(\prime)}$ and $\mathcal{O}_{10}^{(\prime)}$. They are defined, respectively, by

$$\begin{aligned} \mathcal{O}_9^{ij} &= \big(\bar{s}\gamma_\mu P_L b\big)\big(\bar{\ell}_i\gamma^\mu\ell_j\big), & \mathcal{O}_{10}^{ij} &= \big(\bar{s}\gamma_\mu P_L b\big)\big(\bar{\ell}_i\gamma^\mu\gamma_5\ell_j\big), \\ \mathcal{O}_9^{\prime ij} &= \big(\bar{s}\gamma_\mu P_R b\big)\big(\bar{\ell}_i\gamma^\mu\ell_j\big), & \mathcal{O}_{10}^{\prime ij} &= \big(\bar{s}\gamma_\mu P_R b\big)\big(\bar{\ell}_i\gamma^\mu\gamma_5\ell_j\big). \end{aligned} \tag{3.12}$$

In the SMEFT, the NP contributions to their Wilson coefficients are given by

$$\begin{aligned} C_{9,\text{NP}}^{ij} &= +c_\ell\big(\big[\mathcal{C}_{lq}^{(1)}\big]_{ij23} + \big[\mathcal{C}_{lq}^{(3)}\big]_{ij23}\big), & C_{9,\text{NP}}^{\prime ij} &= +c_\ell\big[\mathcal{C}_{ld}\big]_{ij23}, \\ C_{10,\text{NP}}^{ij} &= -c_\ell\big(\big[\mathcal{C}_{lq}^{(1)}\big]_{ij23} + \big[\mathcal{C}_{lq}^{(3)}\big]_{ij23}\big), & C_{10,\text{NP}}^{\prime ij} &= -c_\ell\big[\mathcal{C}_{ld}\big]_{ij23}, \end{aligned} \tag{3.13}$$

with the normalization constant $c_\ell \equiv (\pi/(\alpha_e V_{tb}V_{ts}^*))\cdot(v^2/\Lambda^2)$. In the MFV framework, from eq. (3.3), we can further obtain

$$\begin{aligned} C_{9,\text{MFV}}^{ij} &= -C_{10,\text{MFV}}^{ij} = c_\ell\Delta_q^{23}\big[\epsilon_1^{(1)}\big(\kappa_0^{(1)}\delta^{ij} + \kappa_1^{(1)}\Delta_\ell^{ij}\big) + \big((1)\to(3)\big)\big], \\ C_{9,\text{MFV}}^{\prime ij} &= C_{10,\text{MFV}}^{\prime ij} = 0, \end{aligned} \tag{3.14}$$

where $\epsilon_1^{(n)}$ and $\kappa_{0,1}^{(n)}$ denote the MFV parameters for the operator $\mathcal{Q}_{lq}^{(n)}$ $(n = 1, 3)$. For the $B \to \pi\ell_i\ell_j$, $B \to K^{(*)}\ell_i\ell_j$ and $B_s \to \ell_i\ell_j$ decays, which are all induced by the quark-level $b \to s(d)\ell_i\ell_j$ transitions and will be included in our numerical analysis, the explicit expressions of their branching ratios can be found in, e.g., refs. [68, 91, 92].

### 3.2.3 $s \to d\nu\bar{\nu}$ decay

Both the $K^+ \to \pi^+\nu\bar{\nu}$ and $K_L \to \pi^0\nu\bar{\nu}$ decays are induced by the $s \to d\nu_i\bar{\nu}_j$ transition. The corresponding effective weak Hamiltonian, including the effective operators and their Wilson coefficients, can be obtained from the ones of the $b \to s\nu\bar{\nu}$ transition in section 3.2.1 by replacing the related flavour indices. In addition, for the $K^+ \to \pi^+\nu\bar{\nu}$ decay, a non-negligible charm contribution must be taken into account [84, 85].

In absence of the operator $O_R^{\nu_i\nu_j}$, the branching ratio of the $K^+ \to \pi^+\nu\bar{\nu}$ decay is given by [93–95]

$$\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu}) = \kappa_+(1+\Delta_{\rm EM}) \sum_{i,j=e,\mu,\tau} \frac{1}{3}\left[\left(\sin^2\theta_W \frac{\mathrm{Im}[\lambda_t C_L^{\nu_i\nu_j}]}{2\lambda^5}\right)^2 + \left(\frac{\mathrm{Re}\lambda_c}{\lambda}P_c(X)\delta^{ij} - \sin^2\theta_W \frac{\mathrm{Re}[\lambda_t C_L^{\nu_i\nu_j}]}{2\lambda^5}\right)^2\right], \quad (3.15)$$

where $\lambda_i = V_{is}^* V_{id}$, $\lambda = |V_{us}|$, and $\sin^2\theta_W = 0.23116$ is the weak mixing angle. The parameter $\kappa_+ = (5.173 \pm 0.025) \times 10^{-11}(\lambda/0.225)^8$ contains, besides some other factors, the hadronic matrix element of the weak current $\bar{s}\gamma_\mu P_L d$ that can be extracted from the $K_{\ell 3}$ data with the help of isospin symmetry [96]. $\Delta_{\rm EM} = -0.003$ accounts for the isospin-breaking electromagnetic corrections [94, 95]. $P_c(X) = 0.404 \pm 0.024$ comprises the short- and long-distance charm contributions [96–99].

The rare decay $K_L \to \pi^0\nu\bar{\nu}$ proceeds almost entirely through direct CP violation in the SM [100, 101], and the charm contribution can be fully neglected [93]. The branching ratio of this decay can be written as [93, 102]

$$\mathcal{B}(K_L \to \pi^0\nu\bar{\nu}) = \kappa_L(1-\delta_\epsilon) \sum_{i,j=e,\mu,\tau} \frac{1}{3}\left(\frac{\sin^2\theta_W}{2\lambda^5}\mathrm{Im}\left[\lambda_t C_L^{\nu_i\nu_j}\right]\right)^2, \quad (3.16)$$

where $\kappa_L = (2.231 \pm 0.013) \times 10^{-10}(\lambda/0.225)^8$ encodes, besides some other factors, the hadronic matrix element that can be again extracted from the $K_{\ell 3}$ decay with the help of isospin symmetry [94, 103]. The parameter $\delta_\epsilon$ encodes, on the other hand, the highly suppressed indirect CP-violating contribution [101].

# 4 $M_1 \to M_2 + \mathrm{DM}$

In this section, we first recapitulate the theoretical framework of the DSMEFT and DLEFT, and then detail the calculation of $M_1 \to M_2 + \mathrm{DM}$ decays.

## 4.1 DSMEFT

It is known that EFT can provide a systematic framework for parameterizing various effects induced by the DM particles. The DSMEFT is a generalization of

the SMEFT [81, 82] to include the DM particles, which can be either lighter or heavier than the electroweak scale and are gauge singlets under the SM gauge group [42, 43, 45]. It is also assumed that electroweak symmetry breaking must be implemented via the usual Higgs mechanism. Following the convention of ref. [42], the DSMEFT Lagrangian takes the form

$$\mathcal{L}_{\rm DSMEFT} \supset \frac{1}{\Lambda}\sum_i \mathcal{C}_i \mathcal{Q}_i^{(5)} + \frac{1}{\Lambda^2}\sum_j \mathcal{C}_j \mathcal{Q}_j^{(6)}, \tag{4.1}$$

where $\Lambda$ is the NP scale above which new degrees of freedom appear and interact with the SM and DM particles. The dim-$d$ effective operators $\mathcal{Q}_i^{(d)}$ involve at least one DM field and respect the SM $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry. $\mathcal{C}_i$ denote the corresponding dimensionless Wilson coefficients. We will define the DSMEFT at the electroweak scale $\mu_{\rm EW}$.

In the DSMEFT framework, three types of DM particles can be added: spin-0 scalar $\phi$, spin-1/2 fermion $\chi$ and spin-1 vector $X$. Generally, each type of the DM particles can have multiple generations. For simplicity, the case of one generation DM will be assumed in this paper. For the operators vanishing in the case of one generation, however, two generations of DM particles will be considered. Furthermore, in order to affect the rare $B$-meson and kaon FCNC decays, only the operators that can induce tree-level FCNC interactions among the down-type quarks are included in our analysis. Finally, motivated by the large excess of $\mathcal{B}(B^+ \to K^+ + \nu\bar{\nu})$ reported by Belle II [9], we will focus on the dim-5 and dim-6 operators, which are less suppressed than the ones with dim $\geq 7$. Consequently, the DSMEFT operators we are considering must involve two down-type quarks and at least one DM field. Furthermore, in order to mimic the rare $B$-meson and kaon FCNC decays involving a dineutrino final state, the DM particles must be light enough to be produced in these decays. For each type of the DM particles, the relevant effective operators are given in the following.

For the scalar DM particles, there are four operators,

$$\begin{aligned} \mathcal{Q}_{d\phi} &= \big(\bar{q}_p d_r H\big)\phi + \text{h.c.}, & \mathcal{Q}_{d\phi^2} &= \big(\bar{q}_p d_r H\big)\phi^2 + \text{h.c.}, \\ \mathcal{Q}_{\phi q} &= \big(\bar{q}_p \gamma_\mu q_r\big)\big(i\phi_1 \overleftrightarrow{\partial^\mu} \phi_2\big), & \mathcal{Q}_{\phi d} &= \big(\bar{d}_p \gamma_\mu d_r\big)\big(i\phi_1 \overleftrightarrow{\partial^\mu} \phi_2\big), \end{aligned} \tag{4.2}$$

where $\phi_1 \overleftrightarrow{\partial^\mu} \phi_2 = \phi_1(\partial^\mu \phi_2) - (\partial^\mu \phi_1)\phi_2$, and $\phi_{(i)}$ is the real scalar DM field with the generation index characterized by $i$. For the operators which are not hermitian, "+h.c." should be added to them to indicate that the hermitian conjugated operators should be included in the DSMEFT Lagrangian eq. (4.1). In addition, constructions of the operators $\mathcal{Q}_{\phi q}$ and $\mathcal{Q}_{\phi d}$ requires at least two generations of scalar DM fields.

For the fermionic DM particles, there are only two operators,

$$\mathcal{Q}_{q\chi} = \big(\bar{q}_p \gamma_\mu q_r\big)(\bar{\chi}\gamma^\mu \chi), \qquad \mathcal{Q}_{d\chi} = \big(\bar{d}_p \gamma_\mu d_r\big)(\bar{\chi}\gamma^\mu \chi), \tag{4.3}$$

where $\chi$ denotes the fermionic DM field. In this paper, all the observables considered are insensitive to the chirality of the fermionic DM. Without loss of generality, $\chi$ will be therefore chosen as the right-handed Dirac field.

For the vector DM particles, by using the field strength tensor as the building block, there is only one operator,

$$\mathcal{Q}_{dHX} = \big(\bar{q}_p \sigma_{\mu\nu} d_r\big) H X^{\mu\nu} + \text{h.c.}, \tag{4.4}$$

where $X_{\mu\nu} = \partial_\mu X_\nu - \partial_\nu X_\mu$, and $X_\mu$ denotes the real vector DM field. Taking $X_\mu$ an additional building block, the following operators also appear [43]:

$$\begin{aligned}
\mathcal{Q}_{dX} &= \big(\bar{d}_p \gamma_\mu d_r\big) X^\mu, & \mathcal{Q}_{HdX} &= \big(H^\dagger H\big)\big(\bar{d}_p \gamma^\mu d_r\big) X_\mu, \\
\mathcal{Q}_{qX} &= \big(\bar{q}_p \gamma_\mu q_r\big) X^\mu, & \mathcal{Q}^{(1)}_{HqX} &= \big(H^\dagger H\big)\big(\bar{q}_p \gamma^\mu q_r\big) X_\mu, \\
\mathcal{Q}_{dX^2} &= \big(\bar{q}_p d_r H\big) X_\mu X^\mu + \text{h.c.}, & \mathcal{Q}^{(3)}_{HqX} &= \big(H^\dagger \tau^I H\big)\big(\bar{q}_p \tau^I \gamma^\mu q_r\big) X_\mu, \\
\mathcal{Q}_{qXX} &= \big(\bar{q}_p \gamma_\mu q_r\big) X^{\mu\nu} X_\nu\,, & \mathcal{Q}_{dXX} &= \big(\bar{d}_p \gamma_\mu d_r\big) X^{\mu\nu} X_\nu, \\
\mathcal{Q}_{q\tilde{X}X} &= \big(\bar{q}_p \gamma_\mu q_r\big) \tilde{X}^{\mu\nu} X_\nu\,, & \mathcal{Q}_{d\tilde{X}X} &= \big(\bar{d}_p \gamma_\mu d_r\big) \tilde{X}^{\mu\nu} X_\nu\,, \\
\mathcal{Q}_{DqX^2} &= \big(\bar{q}_p \gamma^\mu D^\nu q_r\big) X_\mu X_\nu + \text{h.c.}, & \mathcal{Q}_{DdX^2} &= \big(\bar{d}_p \gamma^\mu D^\nu d_r\big) X_\mu X_\nu + \text{h.c.}, \\
\mathcal{Q}_{dHX^2} &= (\bar{q}_p \sigma_{\mu\nu} d_r H) X_1^\mu X_2^\nu + \text{h.c.},
\end{aligned} \tag{4.5}$$

where $\tilde{X}^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} X_{\rho\sigma}/2$. It is noted that the rare $B$-meson and kaon decay rates induced by the operators involving the vector DM field $X_\mu$ are divergent in the $m_X \to 0$ limit. Such a divergence is caused by the longitudinal polarization of the vector field and related to the phenomenon that it is impossible to consistently define a massless limit for a vector field without an active gauge symmetry [39]. As in refs. [39, 44, 104], this singularity is treated by assuming some kind of Higgs mechanism in the dark sector. As a result, the Wilson coefficients of the operators in eq. (4.5) should be proportional to some powers of $(m_X/\Lambda)$ and can be redefined as

$$\mathcal{C}_i = \tilde{\mathcal{C}}_i \cdot \begin{cases} (m_X/\Lambda)^2 & \text{for } \mathcal{Q}_i = \mathcal{Q}_{dX^2}, \mathcal{Q}_{DdX^2}, \mathcal{Q}_{DqX^2}, \mathcal{Q}_{dHX^2}, \\ (m_X/\Lambda) & \text{for } \mathcal{Q}_i = \text{others}. \end{cases} \tag{4.6}$$

These redefined Wilson coefficients $\tilde{\mathcal{C}}_i$ will be used in the following analysis.

We can also include the axion or axion-like particles (ALPs) [46–49] in the EFT approach [50–55]. They are pseudo Nambu-Goldstone bosons with non-trivial properties, and can emerge from spontaneous breaking of some global symmetries. Up to dim-5, there are only two operators,

$$\mathcal{Q}_{qa} = \big(\bar{q}_p \gamma_\mu q_r\big) \partial^\mu a, \qquad \mathcal{Q}_{da} = \big(\bar{d}_p \gamma_\mu d_r\big) \partial^\mu a, \tag{4.7}$$

where $a$ denotes the ALP field.

In the MFV hypothesis discussed in section 2, the Wilson coefficients of the DSMEFT operators should take special flavour structures. Explicitly, by using the quark currents in eqs. (2.12)-(2.14), we obtain

$$\mathcal{C}_i^{\rm MFV} = \begin{cases} \epsilon_0^i \hat{\lambda}_d + \epsilon_1^i \Delta_q \hat{\lambda}_d & \text{for } \mathcal{Q}_i = \mathcal{Q}_{d\phi}, \mathcal{Q}_{d\phi^2}, \mathcal{Q}_{dHX}, \mathcal{Q}_{dHX^2}, \mathcal{Q}_{dX^2}, \\ \epsilon_0^i \mathbb{1} + \epsilon_1^i \Delta_q & \text{for } \mathcal{Q}_i = \mathcal{Q}_{\phi q}, \mathcal{Q}_{q\chi}, \mathcal{Q}_{qXX}, \mathcal{Q}_{q\tilde{X}X}, \mathcal{Q}_{DqX^2}, \mathcal{Q}_{qX}, \mathcal{Q}_{HqX}^{(1,3)}, \mathcal{Q}_{qa}, \\ \epsilon_0^i \mathbb{1} & \text{for } \mathcal{Q}_i = \mathcal{Q}_{\phi d}, \mathcal{Q}_{d\chi}, \mathcal{Q}_{dXX}, \mathcal{Q}_{d\tilde{X}X}, \mathcal{Q}_{DdX^2}, \mathcal{Q}_{dX}, \mathcal{Q}_{HdX}, \mathcal{Q}_{da}, \end{cases} \tag{4.8}$$

for the DSMEFT operators discussed above. Here the Wilson coefficient $\mathcal{C}_i^{\rm MFV}$ are given for the down-type quark currents expressed in the mass eigenbasis, and they are $3 \times 3$ matrices in flavour space. Generally, the real parameters $\epsilon_{0,1}^i$ are different from each other for different operators $\mathcal{Q}_i$. We can see that the operators listed in the last line of the above equation do not generate tree-level FCNC interactions among the down-type quarks.

### 4.2 DLEFT

Below the electroweak scale, the NP effects induced by the DM particles can be described by the DLEFT [42]. This EFT is obtained by integrating out all the heavy particles with masses at or above the electroweak scale and can be considered as an extension of the LEFT [32, 33] by including besides the light SM fields also the light DM particles. The DLEFT respects the SM $SU(3)_C \otimes U(1)_{\rm em}$ gauge symmetry, and does not contain the Higgs boson. Thus, it can be considered without reference to the DSMEFT for the light DM and light SM particles interacting at energies below the electroweak scale. Following the same convention as in ref. [42], we can write the DLEFT Lagrangian as

$$\mathcal{L}_{\rm DLEFT} \supset \frac{1}{\Lambda} \sum_i L_i \mathcal{O}_i^{(5)} + \frac{1}{\Lambda^2} \sum_j L_j \mathcal{O}_j^{(6)}, \tag{4.9}$$

up to dim-6, where $\mathcal{O}_i^{(d)}$ denote the dim-$d$ effective operators and $L_i$ the corresponding Wilson coefficients.

Analogous to the DSMEFT case, we can write down all the relevant DLEFT operators involving the scalar, fermionic and vector DM fields. For the scalar DM, we have

$$\begin{aligned} \mathcal{O}_{d\phi} &= (\bar{d}_{Lp} d_{Rr})\phi + \text{h.c.}, & \mathcal{O}_{\phi d}^{L} &= (\bar{d}_{Lp} \gamma_\mu d_{Lr}) \big(i\phi_1 \overleftrightarrow{\partial^\mu} \phi_2\big), \\ \mathcal{O}_{d\phi^2} &= (\bar{d}_{Lp} d_{Rr})\phi^2 + \text{h.c.}, & \mathcal{O}_{\phi d}^{R} &= (\bar{d}_{Rp} \gamma_\mu d_{Rr}) \big(i\phi_1 \overleftrightarrow{\partial^\mu} \phi_2\big), \end{aligned} \tag{4.10}$$

where $d_L$ and $d_R$ denote the left-handed and right-handed down-type quarks, respectively. For the fermionic DM, the relevant operators are given by

$$\mathcal{O}_{d\chi}^{V,LR} = (\bar{d}_{Lp} \gamma_\mu d_{Lr})(\bar{\chi}_a \gamma^\mu \chi_b), \qquad \mathcal{O}_{d\chi}^{V,RR} = (\bar{d}_{Rp} \gamma_\mu d_{Rr})(\bar{\chi}_a \gamma^\mu \chi_b),$$

$$\mathcal{O}_{d\chi}^{S,LR} = (\bar{d}_{Rp} d_{Lr})(\chi_a^T C \chi_b) + \text{h.c.}, \qquad \mathcal{O}_{d\chi}^{S,RR} = (\bar{d}_{Lp} d_{Rr})(\chi_a^T C \chi_b) + \text{h.c.},$$
$$\mathcal{O}_{d\chi}^{T,RR} = (\bar{d}_{Lp} \sigma^{\mu\nu} d_{Rr})(\chi_a^T C \sigma_{\mu\nu} \chi_b) + \text{h.c.}, \tag{4.11}$$

with the charge conjugate operator defined by $C = i\gamma^2\gamma^0$. For the vector DM, there is only one operator,

$$\mathcal{O}_{dX}^{T} = (\bar{d}_{Lp} \sigma_{\mu\nu} d_{Rr}) X_a^{\mu\nu} + \text{h.c.}, \tag{4.12}$$

by using the field strength $X_{\mu\nu}$ as the building block. However, including $X_\mu$ as additional building block, the following operators are obtained:

$$\begin{aligned}
\mathcal{O}_{dX}^{L} &= \big(\bar{d}_{Lp} \gamma_\mu d_{Lr}\big) X^\mu, & \mathcal{O}_{dX}^{R} &= \big(\bar{d}_{Rp} \gamma_\mu d_{Rr}\big) X^\mu, \\
\mathcal{O}_{dXX}^{L} &= \big(\bar{d}_{Lp} \gamma_\mu d_{Lr}\big) X^{\mu\nu} X_\nu, & \mathcal{O}_{dXX}^{R} &= \big(\bar{d}_{Rp} \gamma_\mu d_{Rr}\big) X^{\mu\nu} X_\nu, \\
\mathcal{O}_{d\tilde{X}X}^{L} &= \big(\bar{d}_{Lp} \gamma_\mu d_{Lr}\big) \tilde{X}^{\mu\nu} X_\nu, & \mathcal{O}_{d\tilde{X}X}^{R} &= \big(\bar{d}_{Rp} \gamma_\mu d_{Rr}\big) \tilde{X}^{\mu\nu} X_\nu, \\
\mathcal{O}_{DdX^2}^{L} &= \big(\bar{d}_{Lp} \gamma^\nu D^\mu d_{Lr}\big) X_\mu X_\nu + \text{h.c.}, & \mathcal{O}_{DdX^2}^{R} &= \big(\bar{d}_{Rp} \gamma^\nu D^\mu d_{Rr}\big) X_\mu X_\nu + \text{h.c.}, \\
\mathcal{O}_{dX^2} &= \big(\bar{d}_{Lp} d_{Rr}\big) X_\mu X^\mu + \text{h.c.}, & \mathcal{O}_{dX^2}^{T} &= \big(\bar{d}_{Lp} \sigma_{\mu\nu} d_{Rr}\big) X_1^\mu X_2^\nu + \text{h.c.},
\end{aligned} \tag{4.13}$$

where $D_\mu = \partial_\mu + i g_s T^a G_\mu^a + ieQA_\mu$ denotes the covariant derivative of the $SU(3)_C \otimes U(1)_{\text{em}}$ gauge symmetry. Here we have chosen the basis where the combinations of $X_{\mu\nu}$ and $X_\mu$ are similar as in the DSMEFT operators of eq. (4.5). One can prove that this basis is equivalent to the one adopted in ref. [44]. In addition, there are two operators involving the ALP field, which take the form

$$\mathcal{O}_{da}^{L} = \big(\bar{d}_{Lp} \gamma_\mu d_{Lr}\big) \partial^\mu a, \qquad \mathcal{O}_{da}^{R} = \big(\bar{d}_{Rp} \gamma_\mu d_{Rr}\big) \partial^\mu a. \tag{4.14}$$

Although the DLEFT can be considered without reference to the DSMEFT, the former should arise as the low-energy limit of the latter, if we start from the DSMEFT. At the electroweak scale $\mu_{\text{EW}}$, the tree-level matching conditions between these two EFTs are summarized below:

- Scalar DM:

$$L_{\phi d}^{L} = \mathcal{C}_{\phi q}, \quad L_{\phi d}^{R} = \mathcal{C}_{\phi d}, \quad L_{d\phi^2} = \frac{v}{\sqrt{2}\Lambda} \mathcal{C}_{d\phi^2}, \quad L_{d\phi} = \frac{v}{\sqrt{2}\Lambda} \mathcal{C}_{d\phi}. \tag{4.15}$$

- Fermionic DM:

$$L_{d\chi}^{V,LR} = \mathcal{C}_{q\chi}, \quad L_{d\chi}^{V,RR} = \mathcal{C}_{d\chi}, \quad L_{d\chi}^{S,LR} = 0, \quad L_{d\chi}^{S,RR} = 0, \quad L_{d\chi}^{T,RR} = 0. \tag{4.16}$$

- Vector DM:

$$L_{dX}^{T} = \frac{v}{\sqrt{2}\Lambda} \mathcal{C}_{dHX}, L_{dX}^{L} = \mathcal{C}_{qX} + \frac{v^2}{2\Lambda^2} \big(\mathcal{C}_{HqX}^{(1)} + \mathcal{C}_{HqX}^{(3)}\big), L_{dX}^{R} = \mathcal{C}_{dX} + \frac{v^2}{2\Lambda^2} \mathcal{C}_{HdX},$$

$$\begin{aligned}
L_{dXX}^{L} &= \mathcal{C}_{qXX}, \quad L_{d\tilde{X}X}^{L} = \mathcal{C}_{q\tilde{X}X}, \quad L_{DdX^2}^{L} = \mathcal{C}_{DqX^2}, \quad L_{dX^2} = \frac{v}{\sqrt{2}\Lambda} \mathcal{C}_{dX^2}, \\
L_{dXX}^{R} &= \mathcal{C}_{dXX}, \quad L_{d\tilde{X}X}^{R} = \mathcal{C}_{d\tilde{X}X}, \quad L_{DdX^2}^{R} = \mathcal{C}_{DdX^2}, \quad L_{dX^2}^{T} = \frac{v}{\sqrt{2}\Lambda} \mathcal{C}_{dHX^2}.
\end{aligned} \tag{4.17}$$

- ALP:

$$L_{da}^{L} = \mathcal{C}_{qa}, \qquad L_{da}^{R} = \mathcal{C}_{da}. \tag{4.18}$$

Here all the Wilson coefficients are given in the mass eigenbasis. In the MFV hypothesis, the Wilson coefficients in DSMEFT take the form of eq. (4.8).

Since the above matching conditions are given at the electroweak scale, i.e., $L_i \equiv L_i(\mu_{\rm EW})$ and $\mathcal{C}_i \equiv \mathcal{C}_i(\mu_{\rm EW})$, the renormalization group (RG) evolution should be performed in the DLEFT to obtain the corresponding Wilson coefficients at the typical scales of $B$-meson and kaon decays. In the DLEFT, as the DM particles are color singlet, we need only consider the QCD RG evolution of the quark current in each DLEFT operator. Furthermore, the different DLEFT operators do not mix under the QCD RG running. Accordingly, the RG equations take the form

$$\frac{dL_i(\mu)}{d\ln\mu} = \frac{\alpha_s}{4\pi}\gamma_i\, L_i(\mu), \tag{4.19}$$

where $L_i(\mu)$ denote the Wilson coefficients of the operators $\mathcal{O}_i$ at the scale $\mu$. $\gamma_i$ denote the anomalous dimension matrices (ADM) for the operators $\mathcal{O}_i$. They can be obtained from the ADM of the semi-leptonic operators in the LEFT [32, 105]. For the DLEFT operators with a non-zero matching condition, the values of these ADM are given by

$$\gamma_i = \begin{cases} -6C_F & \text{for } \mathcal{O}_i = \mathcal{O}_{d\phi}, \mathcal{O}_{d\phi^2}, \mathcal{O}_{dX^2}, \\ +2C_F & \text{for } \mathcal{O}_i = \mathcal{O}_{dX}^{T}, \mathcal{O}_{dX^2}^{T}, \\ 0 & \text{for } \mathcal{O}_i = \mathcal{O}_{\phi d}^{L,R}, \mathcal{O}_{d\chi}^{V,LR}, \mathcal{O}_{d\chi}^{V,RR}, \mathcal{O}_{dX}^{L,R}, \mathcal{O}_{dXX}^{L,R}, \mathcal{O}_{d\tilde{X}X}^{L,R}, \mathcal{O}_{DdX^2}^{L,R}, \end{cases} \tag{4.20}$$

with $C_F = (N_c^2 - 1)/(2N_c)$ and $N_c = 3$. Numerically, we find

$$\begin{aligned} L_i(4.8\,\text{GeV}) &= L_i(m_Z) \cdot \begin{cases} 1.37 & \text{for } \mathcal{O}_i = \mathcal{O}_{d\phi}, \mathcal{O}_{d\phi^2}, \mathcal{O}_{dX^2}, \\ 0.90 & \text{for } \mathcal{O}_i = \mathcal{O}_{dX}^{T}, \mathcal{O}_{dX^2}^{T}, \end{cases} \\ L_i(2.0\,\text{GeV}) &= L_i(m_Z) \cdot \begin{cases} 1.59 & \text{for } \mathcal{O}_i = \mathcal{O}_{d\phi}, \mathcal{O}_{d\phi^2}, \mathcal{O}_{dX^2}, \\ 0.86 & \text{for } \mathcal{O}_i = \mathcal{O}_{dX}, \end{cases} \end{aligned} \tag{4.21}$$

which can be used in the $B$-meson and kaon decays, respectively.

### 4.3 Observables

Within the DLEFT, by using the similar treatment as for the $d_i \to d_j \nu\bar{\nu}$ transitions discussed in section 3.2, one can calculate the various $d_i \to d_j + \text{DM}$ processes. Generally, the amplitudes of these decays can be factorized into the product of the Wilson coefficient and the hadronic matrix element of the corresponding effective operator that can be parameterized form factors or decay constants. In appendix A,

| Input | Value | Unit | Reference |
|---|---|---|---|
| $m_t^{\rm pole}$ | $172.69 \pm 0.30$ | GeV | [28] |
| $\lvert V_{cb}\rvert$(semi-leptonic) | $41.15 \pm 0.34 \pm 0.45$ | $10^{-3}$ | [118] |
| $\lvert V_{ub}\rvert$(semi-leptonic) | $3.88 \pm 0.08 \pm 0.21$ | $10^{-3}$ | [118] |
| $\lvert V_{us}\rvert f_+^{K\to\pi}(0)$ | $0.2165 \pm 0.0004$ | | [118] |
| $\gamma$ | $72.1^{+5.4}_{-5.7}$ | $\circ$ | [118] |
| $f_+^{K\to\pi}(0)$ | $0.9675 \pm 0.0009 \pm 0.0023$ | | [118] |
| $\Delta m^2_{21}$ | $7.41^{+0.21}_{-0.20}$ $(7.41^{+0.21}_{-0.20})$ | $10^{-5}\cdot \text{eV}^2$ | [28] |
| $\Delta m^2_{32}$ | $2.437^{+0.028}_{-0.027}$ $(-2.498^{+0.032}_{-0.025})$ | $10^{-3}\cdot \text{eV}^2$ | [28] |
| $\theta_{12}$ | $33.41^{+0.75}_{-0.72}$ $(33.41^{+0.75}_{-0.72})$ | $\circ$ | [28] |
| $\theta_{23}$ | $49.1^{+1.0}_{-1.3}$ $(49.5^{+0.9}_{-1.2})$ | $\circ$ | [28] |
| $\theta_{13}$ | $8.54^{+0.11}_{-0.12}$ $(8.57^{+0.12}_{-0.11})$ | $\circ$ | [28] |
| $\delta_{\rm CP}$ | $197^{+42}_{-25}$ $(286^{+27}_{-32})$ | $\circ$ | [28] |

**Table 1**: Main input parameters used in our numerical analysis. The values (in brackets) of the parameters in the lepton sector correspond to the NO (IO) of neutrino masses.

we summarize the relevant hadronic matrix elements used in this paper. For the explicit expressions of the decay rates, we refer the readers to refs. [39, 106–111] for the $d_i \to d_j + \phi\,(+\phi)$ decays, refs. [5, 39, 112, 113] for the $d_i \to d_j + \chi + \chi$ decay, refs. [39, 44, 110] for the $d_i \to d_j + X\,(+X)$ decays, and refs. [114–116] for the $d_i \to d_i + a$ decays. All the relevant analytical expressions have been implemented in our recently developed package `HadronToNP` [117], which will be used in the following numerical analysis.

## 5 Numerical analysis

In this section, we proceed to present our numerical results and discussions. In table 1, we list the main input parameters used in our numerical analysis. Table 2 summaries our SM predictions and the up-to-date experimental measurements of the relevant observables. We take $\Lambda_{\rm NP} = 1\,\text{TeV}$ in the SMEFT and DSMEFT. In the following, we investigate two explanations to the recent Belle II measurement of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ [9]. In section 5.1, we consider the $b \to s + \nu\bar{\nu}$ transition in the SMEFT. In section 5.2, the missing energy signals in the Belle II measurement are interpreted in the DSMEFT as contributions from DM particles.

| Observable | SM | Exp | Unit | Sensitivity |
|---|---|---|---|---|
| $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ | $4.16 \pm 0.57$ | $23 \pm 5^{+5}_{-4}$ [119] | $10^{-6}$ | 19% [11] |
| $\mathcal{B}(B^0 \to K^0\nu\bar{\nu})$ | $3.85 \pm 0.52$ | $< 26$ [120] | $10^{-6}$ | 87% [11] |
| $\mathcal{B}(B^+ \to K^{*+}\nu\bar{\nu})$ | $9.70 \pm 0.94$ | $< 61$ [120] | $10^{-6}$ | 75% [11] |
| $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})$ | $9.00 \pm 0.87$ | $< 18$ [120] | $10^{-6}$ | 40% [11] |
| $\mathcal{B}(B_s \to \phi\nu\bar{\nu})$ | $9.93 \pm 0.72$ | $< 5400$ [121] | $10^{-6}$ | 2% [122] |
| $\mathcal{B}(B^+ \to \pi^+\nu\bar{\nu})$ | $1.40 \pm 0.18$ | $< 140$ [120] | $10^{-7}$ | |
| $\mathcal{B}(B^0 \to \pi^0\nu\bar{\nu})$ | $6.52 \pm 0.85$ | $< 900$ [120] | $10^{-8}$ | |
| $\mathcal{B}(B^+ \to \rho^+\nu\bar{\nu})$ | $4.06 \pm 0.79$ | $< 300$ [120] | $10^{-7}$ | |
| $\mathcal{B}(B^0 \to \rho^0\nu\bar{\nu})$ | $1.89 \pm 0.36$ | $< 400$ [120] | $10^{-7}$ | |
| $\mathcal{B}(K^+ \to \pi^+\nu\bar{\nu})$ | $8.42 \pm 0.61$ | $10.6^{+4.0}_{-3.4} \pm 0.9$ [123] | $10^{-11}$ | |
| $\mathcal{B}(K_L \to \pi^0\nu\bar{\nu})$ | $3.41 \pm 0.45$ | $< 300$ [124] | $10^{-11}$ | |
| $\mathcal{B}(B_s \to \nu\bar{\nu})$ | $\approx 0$ | $< 5.9$ [29] | $10^{-4}$ | $< 1.1 \times 10^{-5}$ [10] |
| $\mathcal{B}(B^0 \to \nu\bar{\nu})$ | $\approx 0$ | $< 1.4$ [29] | $10^{-4}$ | $< 5.0 \times 10^{-6}$ [10] |

**Table 2**: Our SM predictions and the experimental measurements of the relevant observables. Upper limits are all given at 90% confidence level (CL). The last column shows the expected sensitivities at the Belle II ($5\,\text{ab}^{-1}$) or at the CEPC (Tera-$Z$ phase). The sensitivities at the FCC-ee are very similar to that at the CEPC, and hence are not given here.

### 5.1 $M_1 \to M_2 + \nu\bar{\nu}$

The recent Belle II measurement of the $B^+ \to K^+ + \text{inv}$ decay could be explained by NP contributions to the $b \to s\nu\bar{\nu}$ transition. In this case, the theoretical prediction of the related $M_1 \to M_2 + \text{inv}$ processes takes the form

$$\mathcal{B}_{\text{th}}(M_1 \to M_2 + \text{inv}) = \sum_{i,j=e,\mu,\tau} \mathcal{B}(M_1 \to M_2 \nu_i \bar{\nu}_j). \tag{5.1}$$

The NP contributions to these processes can be generally described in the framework of SMEFT. We refer to refs. [13, 14, 25] for the recent relevant studies. In the following, we also focus on the SMEFT but with the MFV hypothesis.

In the SMEFT, as mentioned in section 3.2.1, one interesting implication of the MFV hypothesis is that the $d_i \to d_j\nu\bar{\nu}$ transitions are all governed by only one effective operator at low energy. As a consequence, the ratio of the branching fractions of any two $b \to s\nu\bar{\nu}$ decays should be independent of the NP contribution.

Considering as an example the $B^+ \to K^+\nu\bar{\nu}$ and $B^0 \to K^{*0}\nu\bar{\nu}$ decays, we obtain

$$\frac{\mathcal{B}(B^+ \to K^+\nu\bar{\nu})}{\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})} = \frac{\mathcal{B}(B^+ \to K^+\nu\bar{\nu})_{\rm SM}}{\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})_{\rm SM}} = 0.46 \pm 0.07, \tag{5.2}$$

where summation over the neutrino flavours has been taken into account. In this ratio, all the relevant SMEFT operators, i.e. $\mathcal{Q}_{lq}^{(1)}$ and $\mathcal{Q}_{lq}^{(3)}$, have been included simultaneously.[4] When deriving this ratio, it is noted that only the MFV hypothesis in the quark sector is required to eliminate the right-handed operator $\mathcal{O}_R$. Therefore, the ratio still holds even by assuming generic couplings in the lepton sector, i.e., without the leptonic MFV. In addition, as both the CKM and Wilson coefficients involved are cancelled in this ratio, the theoretical uncertainty is reduced compared to that of the individual branching ratios. If future experimental measurements find any deviation from the theoretical value, it could imply new sources of flavour violation beyond the Yukawa couplings. By using the above ratio, the Belle 90% CL bound $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu}) < 18 \times 10^{-6}$ would imply that

$$\mathcal{B}(B^+ \to K^+\nu\bar{\nu})_{\rm MFV} < 10.5 \times 10^{-6}, \tag{5.3}$$

which is obviously lower that the 90% CL lower bound of the Belle II measurement, $\mathcal{B}(B^+ \to K^+\nu\bar{\nu}) > 12.5 \times 10^{-6}$. Or inversely, the Belle II measurement $\mathcal{B}(B^+ \to K^+\nu\bar{\nu}) = (23^{+7}_{-6}) \times 10^{-6}$ would indicate that

$$\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})_{\rm MFV} = \left(50^{+17}_{-16}\right) \times 10^{-6}, \tag{5.4}$$

which is incompatible with the Belle 90% CL bound $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu}) < 18 \times 10^{-6}$. As a conclusion, the recent Belle II measurement of the $B^+ \to K^+\nu\bar{\nu}$ decay cannot be explained in the SMEFT with MFV hypothesis. However, the MFV prediction of $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})$ is well within the reach of Belle II at $5\,\text{ab}^{-1}$ (cf. table 2), which can be used as a further test of the MFV hypothesis considered here.

Similarly, the ratio of the branching fractions of a $b \to s\nu\bar{\nu}$ and a $b \to d\nu\bar{\nu}$ decay is also independent of the NP contribution in the SMEFT with MFV hypothesis. As an illustration, let us consider the $B^+ \to K^+\nu\bar{\nu}$ and $B^+ \to \pi^+\nu\bar{\nu}$ decays, and obtain

$$\frac{\mathcal{B}(B^+ \to K^+\nu\bar{\nu})}{\mathcal{B}(B^+ \to \pi^+\nu\bar{\nu})} = \frac{\mathcal{B}(B^+ \to K^+\nu\bar{\nu})_{\rm SM}}{\mathcal{B}(B^+ \to \pi^+\nu\bar{\nu})_{\rm SM}} = 29.7 \pm 5.6, \tag{5.5}$$

where summation over the neutrino flavours has been included. Taking as input the Belle II measurement in eq. (1.1), this ratio implies that

$$\mathcal{B}(B^+ \to \pi^+\nu\bar{\nu})_{\rm MFV} = \left(7.8^{+2.8}_{-2.6}\right) \times 10^{-7}, \tag{5.6}$$

[4] The ratio still holds after including the operators $\mathcal{Q}_{Hq}^{(1,3)}$, although they are not considered in our analysis as mentioned in section 3.1.

which shows a $2.3\sigma$ deviation from the SM prediction $\mathcal{B}(B^+ \to \pi^+\nu\bar{\nu})_{\rm SM} = (1.40 \pm 0.18) \times 10^{-7}$, but is still compatible with the current 90% CL experimental bound $\mathcal{B}(B^+ \to \pi^+\nu\bar{\nu}) < 140 \times 10^{-7}$. Of course, similar ratios can also be constructed for other processes to test the framework used here.

In the MFV framework, all the quark FCNC or the LFV processes are related with each other. Explicitly, for the operators $\mathcal{Q}_{lq}^{(i)}$ $(i = 1, 3)$, all the quark FCNC transitions are governed by the two parameter products $\epsilon_1^{(i)}\kappa_0^{(i)}$ and $\epsilon_1^{(i)}\kappa_1^{(i)}$. In order to constrain them, we consider several quark (and lepton) FCNC processes, including the $B \to K^{(*)}\nu\bar{\nu}$, $B \to K^{(*)}e\mu$, $B \to \pi e\mu$, $B_s \to e\mu$, $B_s \to \mu^+\mu^-$ as well as $K \to \pi\nu\bar{\nu}$ decays. Besides the input parameters listed in table 1, the lepton number breaking scale $\Lambda_{\rm LN} = 10^{14}\,{\rm GeV}$ and the lightest neutrino mass $m_{\nu_1} = 0.2\,{\rm eV}$ are used in the leptonic MFV. Numerically, we find that the dominant constraints come from the $B^+ \to K^+\nu\bar{\nu}$, $B^0 \to K^{*0}\nu\bar{\nu}$, $B^0 \to K^{*0}e^-\mu^+$ and $B_s \to \mu^+\mu^-$ decays. The allowed parameter regions of $\epsilon_1^{(i)}\kappa_0^{(i)}$ and $\epsilon_1^{(i)}\kappa_1^{(i)}$ are shown in figure 1. We can see that there is a small gap between the allowed regions of the $B^+ \to K^+\nu\bar{\nu}$ and $B^0 \to K^{*0}\nu\bar{\nu}$ decays. This reflects the discrepancy between the MFV prediction in eq. (5.4) and the current Belle bound on $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})$.

Since the uncertainty of the recent Belle II measurement is still quite large, we use instead all the processes except the $B^+ \to K^+\nu\bar{\nu}$ decay to derive the allowed parameter space. As shown in figure 1, the upper bounds on the magnitudes of the products $\epsilon_1^{(i)}\kappa_0^{(i)}$ and $\epsilon_1^{(i)}\kappa_1^{(i)}$ for the two operators $\mathcal{Q}_{lq}^{(1,3)}$ are all around 0.05 in the cases of NO and IO of neutrino masses, which are much smaller than the parameter regions required to explain the $B^+ \to K^+\nu\bar{\nu}$ excess. Using the allowed parameter regions for $\mathcal{Q}_{lq}^{(i)}$, we can derive the following numerical predictions:

$$\begin{aligned}
&3.4\,(3.5) \times 10^{-6} < \mathcal{B}(B^+ \to K^+\nu\bar{\nu})_{\rm MFV} < 5.0\,(4.9) \times 10^{-6},\\
&7.4\,(7.5) \times 10^{-6} < \mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})_{\rm MFV} < 10.8\,(10.6) \times 10^{-6},\\
&1.1\,(1.2) \times 10^{-7} < \mathcal{B}(B^+ \to \pi^+\nu\bar{\nu})_{\rm MFV} < 1.7\,(1.7) \times 10^{-7},
\end{aligned} \tag{5.7}$$

which correspond the NO (IO) of neutrino masses. The numerical results show less than 0.5% difference between $\mathcal{Q}_{lq}^{(1)}$ and $\mathcal{Q}_{lq}^{(3)}$, implying that both of these two operators coincide with each other under the above bounds. From these bounds, we can see that the NP contributions to the $B \to K\nu\bar{\nu}$ and $B \to \pi\nu\bar{\nu}$ decays should be less than $\sim 20\%$ of their SM predictions. It is noted that the leptonic MFV has been used to derive the allowed parameter space and the above bounds. Therefore, the bound on $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ is more stringent than the one in eq. (5.3) derived without the leptonic MFV.

### 5.2 $M_1 \to M_2 + \mathrm{DM}$

The large excess of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ reported by Belle II can also be explained by light NP particles that contribute to the signal as $B^+ \to K^+ +$ missing energy. In

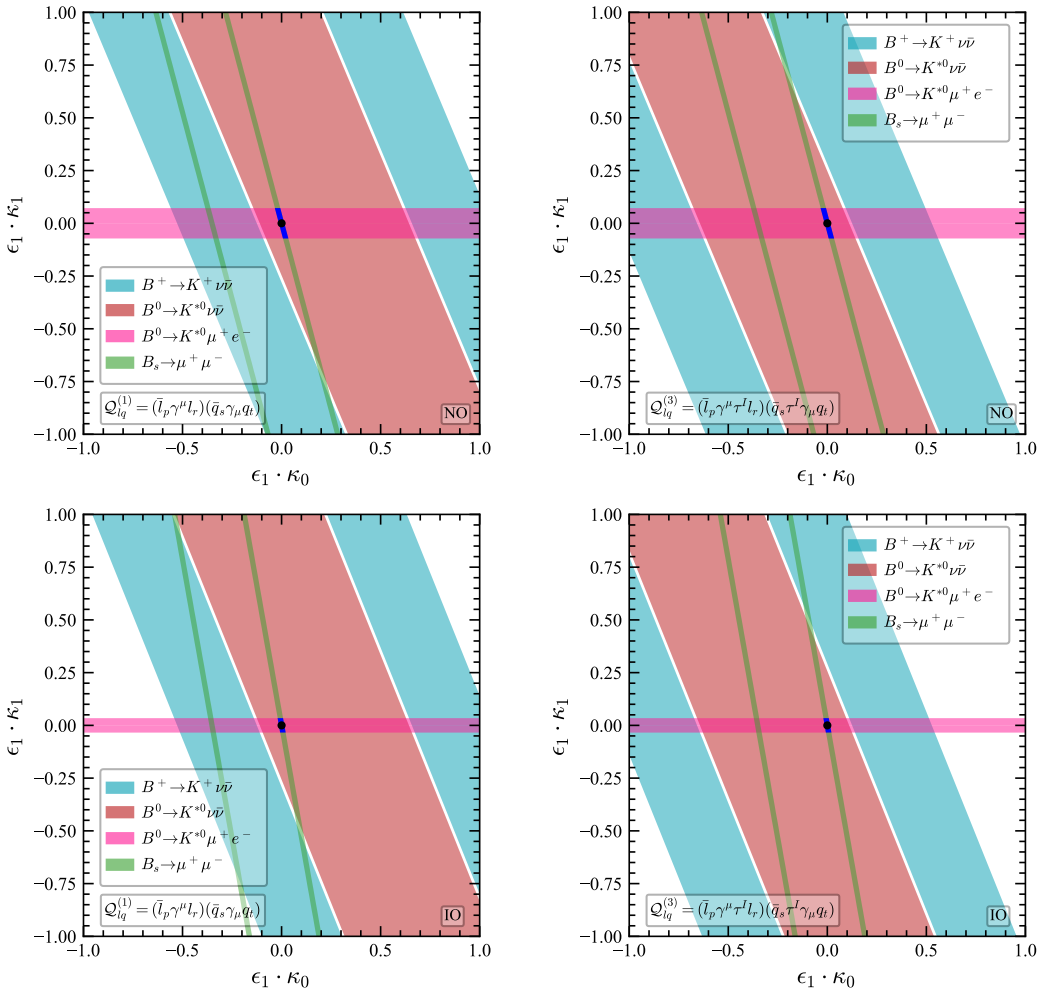



**Figure 1**: Constraints on the MFV parameter products $(\epsilon_1^{(i)}\kappa_0^{(i)}, \epsilon_1^{(i)}\kappa_1^{(i)})$ for $\mathcal{Q}_i = \mathcal{Q}_{lq}^{(1)}$ (left) and $\mathcal{Q}_i = \mathcal{Q}_{lq}^{(3)}$ (right) in the cases of NO (upper) and IO (lower) of neutrino masses. The allowed regions by the measurement of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ are at $2\sigma$ level, while the regions by $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})$, $\mathcal{B}(B^0 \to K^{*0}\mu^+e^-)$ and $\mathcal{B}(B_s \to \mu^+\mu^-)$ at 90% CL. The allowed regions by all the last three observables are shown in blue. The black points $(0, 0)$ denote the SM case.

this case, the theoretical predictions of the $M_1 \to M_2 + \text{inv}$ decay can be written as

$$\mathcal{B}_{\text{th}}(M_1 \to M_2 + \text{inv}) = \mathcal{B}_{\text{SM}}(M_1 \to M_2 + \nu\bar{\nu}) + \mathcal{B}_{\text{NP}}(M_1 \to M_2 + \text{DM}). \tag{5.8}$$

As discussed in section 4, the effects of light NP particles in the second term can be generally described in the EFT framework. In the following, we investigate this explanation in the DSMEFT framework, in which the Wilson coefficients are taken to be the most general form or satisfy the MFV hypothesis.

### 5.2.1 General flavour structure

In the DSMEFT with general flavour structure, only the $b \to s + \text{inv}$ processes are directly related to the $B^+ \to K^+ + \text{inv}$ decays, which include $B^0 \to K^0 + \text{inv}$, $B^{0,+} \to K^{*\,0,+} + \text{inv}$, $B_s \to \phi + \text{inv}$ and $B_s \to \text{inv}$ decays. In the following, we investigate constraints on the DSMEFT operators from the experimental data on these decays as summarized in table 2, and discuss future prospects at Belle II, CEPC and FCC-ee. As discussed in section 4.1, seven DSMEFT operators are not hermitian. In order to compare with the MFV analysis in the next subsection, contributions from the hermitian conjugation of these operators are neglected in the numerical analysis in this subsection, i.e., $\mathcal{C}_{ji} = 0$ is taken for $d_i \to d_j$ processes.[5] For example, contributions to the $b \to s\phi\phi$ processes from the operator $\mathcal{Q}_{d\phi^2}$ are calculated by taking $[\mathcal{C}_{d\phi^2}]_{23} \neq 0$ and $[\mathcal{C}_{d\phi^2}]_{32} = 0$. Two exceptions are the operators $\mathcal{Q}_{DqX^2}$ and $\mathcal{Q}_{DdX^2}$. Consider the former as an example, we take $[\mathcal{C}_{DqX^2}]_{ji} = [\mathcal{C}_{DqX^2}]^*_{ij}$ for the $d_i \to d_j$ processes to compare with the MFV analysis, where its MFV coupling matrix in eq. (2.16) also implies a hermitian $\mathcal{C}_{DqX^2}$. In this case, the hadronic matrix elements of the weak currents can be parameterized in terms of the transition form factors of the vector current as discussed in appendix A.2.

For the scalar DM, there are three operators ($\mathcal{Q}_{d\phi^2}$, $\mathcal{Q}_{\phi q}$ and $\mathcal{Q}_{\phi d}$, cf. eq. (4.2)) contributing to the three-body $d_i \to d_j\phi\phi$ and one operator ($\mathcal{Q}_{d\phi}$, cf. eq. (4.2)) to the two-body $d_i \to d_j\phi$ decay. Allowed regions of the Wilson coefficients for these four operators are shown in figure 2. From this figure, we make the following observations:

- The $B^0 \to K^{*0} + \text{inv}$ decay provides the strongest constraints on the parameter regions allowed by the Belle II measurement of the $B^+ \to K^+ + \text{inv}$ decay. Especially, for the operators $\mathcal{Q}_{\phi q}$ and $\mathcal{Q}_{\phi d}$, the regions of $m_{\phi_2} \lesssim 1.5\,\text{GeV}$ ($m_{\phi_1} < m_{\phi_2}$ is assumed without loss of generality) are almost excluded by the decay. For the operator $\mathcal{Q}_{d\phi}$, on the other hand, large part of the parameter region with $m_\phi \lesssim 3.8\,\text{GeV}$ is excluded.

- Another stronger constraint arises from the $B_s \to \text{inv}$ decay, which benefits from the large phase space. As a result, it excludes the parameter region for large DM mass, e.g., the region with $m_\phi \gtrsim 2.3\,\text{GeV}$ is excluded for the operator $\mathcal{Q}_{d\phi^2}$. For the operators $\mathcal{Q}_{\phi q}$ and $\mathcal{Q}_{\phi d}$, however, the $B_s \to \phi_1\phi_2$ amplitude vanishes in the case of $m_{\phi_1} = m_{\phi_2}$ due to the derivative term $\phi_1 \overleftrightarrow{\partial^\mu} \phi_2$ in the operators. Therefore, only the parameter regions with $m_{\phi_1} \neq m_{\phi_2}$ receive constraint from the $B_s \to \text{inv}$ decay.

- Compared to the $B^0 \to K^{*0} + \text{inv}$ and $B_s \to \text{inv}$ decays, The $B^0 \to K^0 + \text{inv}$ and $B^+ \to K^{*+} + \text{inv}$ decays provide almost no further constraints. The upper

[5] Let us consider $\mathcal{Q}_{d\phi^2}$ as an example. For the $b \to s\phi\phi$ process, $[\mathcal{C}_{d\phi^2}]_{23} \neq 0$ and $[\mathcal{C}_{d\phi^2}]_{32} = 0$ are taken in the general DSMEFT analysis. In the MFV hypothesis, the coupling matrix eq. (2.17) implies that $[\mathcal{C}_{d\phi^2}]_{23} \gg [\mathcal{C}_{d\phi^2}]_{32}$, which results in the approximation $[\mathcal{C}_{d\phi^2}]_{32} \approx 0$.

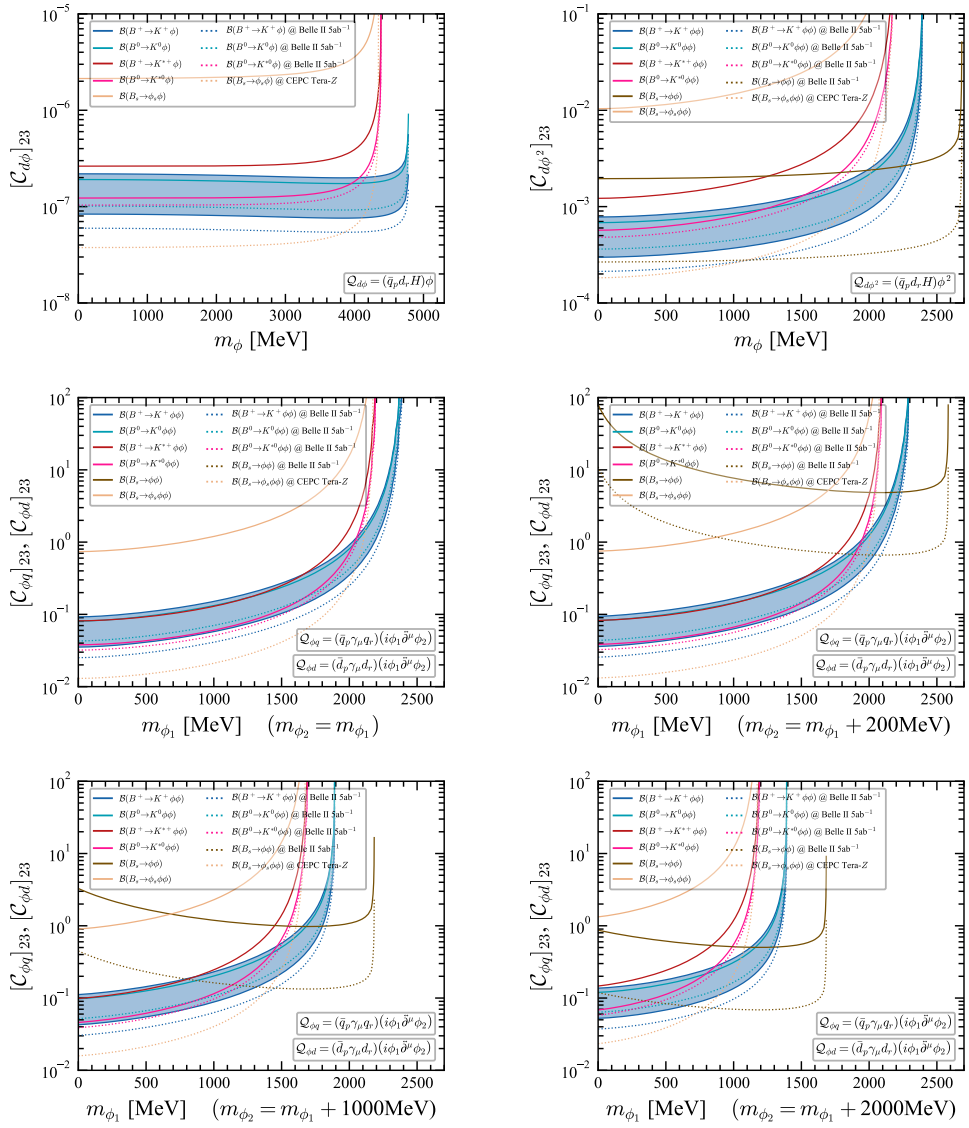



**Figure 2**: Constraints on the Wilson coefficients of the operators involving scalar DM as a function of the scalar mass $m_\phi$. The blue band denotes the $2\sigma$ allowed region by the recent Belle II measurement of the $B^+ \to K^+ + \text{inv}$ decay. The upper bounds from other measurements are shown by the solid lines, while the expected sensitivities at Belle II and CEPC by the dotted lines. The sensitivities at FCC-ee are very similar to that at CEPC and hence are not shown. For the operators $\mathcal{Q}_{\phi q}$ and $\mathcal{Q}_{\phi d}$, we have considered four different cases for the DM masses, as indicated in the last four plots.

bounds from these two decays are comparable to or weaker than the upper limits of the $2\sigma$ allowed region required to explain the $B^+ \to K^+ + \text{inv}$ excess.

- All the processes put very weak constraints on the parameter regions with

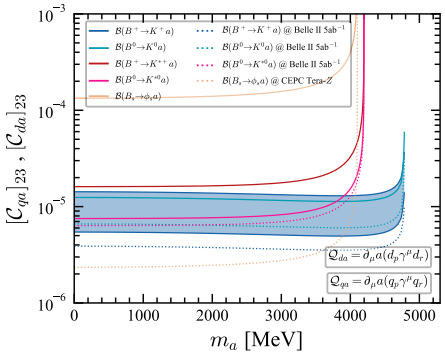



**Figure 3**: Same as in figure 2, but for the operators involving ALP.

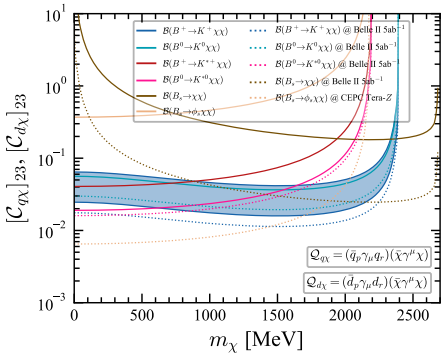



**Figure 4**: Same as in figure 2, but for the operators involving fermionic DM.

$m_\phi \lesssim 2.3\,\text{GeV}$ for the operator $\mathcal{Q}_{d\phi^2}$.

In addition, constraints on the operators involving ALP are shown in figure 3, which are very similar to the ones for the operator $\mathcal{Q}_{d\phi}$ but with much larger allowed Wilson coefficients.

For the fermionic DM, there are only two relevant operators $\mathcal{Q}_{q\chi}$ and $\mathcal{Q}_{d\chi}$ (cf. eq. (4.3)) in the DSMEFT. The experimental constraints are shown in figure 4. In the cases of $m_\chi \lesssim 0.6\,\text{GeV}$ and $m_\chi \gtrsim 2.3\,\text{GeV}$, the parameter spaces required to account for the Belle II measurement are already excluded by the $B^0 \to K^{*0} + \text{inv}$ and $B_s \to \text{inv}$ decays, respectively. In the range of $0.6 \lesssim m_\chi \lesssim 2.3\,\text{GeV}$, however, large part of the parameter space survives all the constraints. This is quite different from the case of the SMEFT operator $\mathcal{Q}_{lq}^{(1)}$ discussed in section 5.1, due to the massive fermionic DM fields in the operators $\mathcal{Q}_{q\chi}$ and $\mathcal{Q}_{d\chi}$. In the massless limit, their effects on the decay width are the same as that of the operator $\mathcal{Q}_{lq}^{(1)}$.[6]

For the vector DM, there are 8 operators ($\mathcal{Q}_{dX^2}$, $\mathcal{Q}_{qXX}$, $\mathcal{Q}_{dXX}$, $\mathcal{Q}_{q\tilde{X}X}$, $\mathcal{Q}_{d\tilde{X}X}$, $\mathcal{Q}_{DqX^2}$, $\mathcal{Q}_{DdX^2}$ and $\mathcal{Q}_{dHX^2}$, cf. eq. (4.5)) contributing to the three-body $d_i \to d_j XX$

[6] Although the leptonic and DM currents in these operators have different chiralities, the total decay width does not depend on the chirality of the currents.

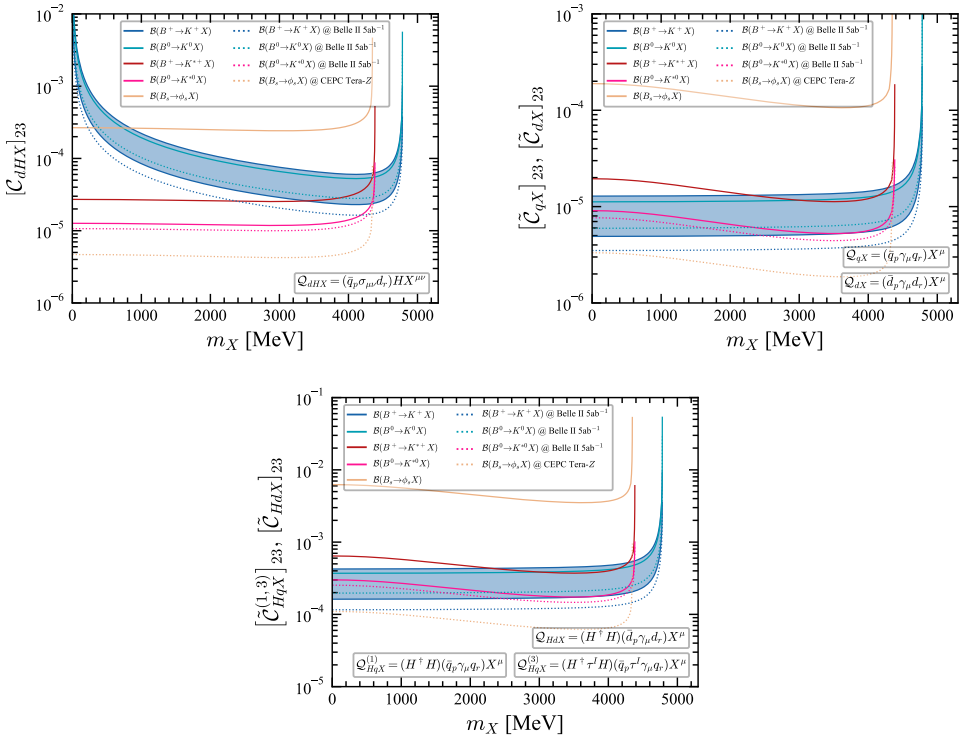



**Figure 5**: Same as in figure 2, but for the operators involving one vector DM.

and 6 operators ($\mathcal{Q}_{dHX}$, $\mathcal{Q}_{dX}$, $\mathcal{Q}_{qX}$, $\mathcal{Q}_{HdX}$ and $\mathcal{Q}^{(1,3)}_{HqX}$, cf. eqs. (4.4) and (4.5)) to the two-body $d_i \to d_j X$ decay. For the Wilson coefficient of each operator, the parameter region required to account for the Belle II measurement of the $B^+ \to K^+ + \text{inv}$ decay as well as the experimental bounds from other processes are shown in figures 5 and 6. The following observations are made:

- For the operators involving one DM field, the high-mass region of $4.3 \lesssim m_X \lesssim 4.6\,\text{GeV}$ does not suffer from any constraint, while large part of the parameter region with $m_X \lesssim 4.3\,\text{GeV}$ is excluded by the $B^0 \to K^{*0} + \text{inv}$ decay. Specifically, for the operator $\mathcal{Q}_{dHX}$, the unique one built from the field strength tensor $X_{\mu\nu}$, the parameter region with $m_X \lesssim 4.3\,\text{GeV}$ is excluded; for the other operators, parts of the parameter region still survive for $m_X \lesssim 2.5\,\text{GeV}$ or $4.3 \lesssim m_X < 4.6\,\text{GeV}$.

- For the operators involving two DM fields, except for $\mathcal{Q}_{dHX^2}$, the allowed parameter regions with $m_X \lesssim 1.7\,\text{GeV}$ suffer almost no constraints from any other processes. In the case of a heavier DM, the $B_s \to \text{inv}$ decay excludes a large part of the parameter space. Especially, the parameter spaces with $m_X \gtrsim 2\,\text{GeV}$ are excluded by the $B_s \to \text{inv}$ decay for all these operators.

- Constraints on the operator $\mathcal{Q}_{dHX^2}$ are quite different from the other ones. As the tensor hadronic matrix element $\langle 0|\bar{b}\sigma_{\mu\nu}s|B_s\rangle$ vanishes identically, this

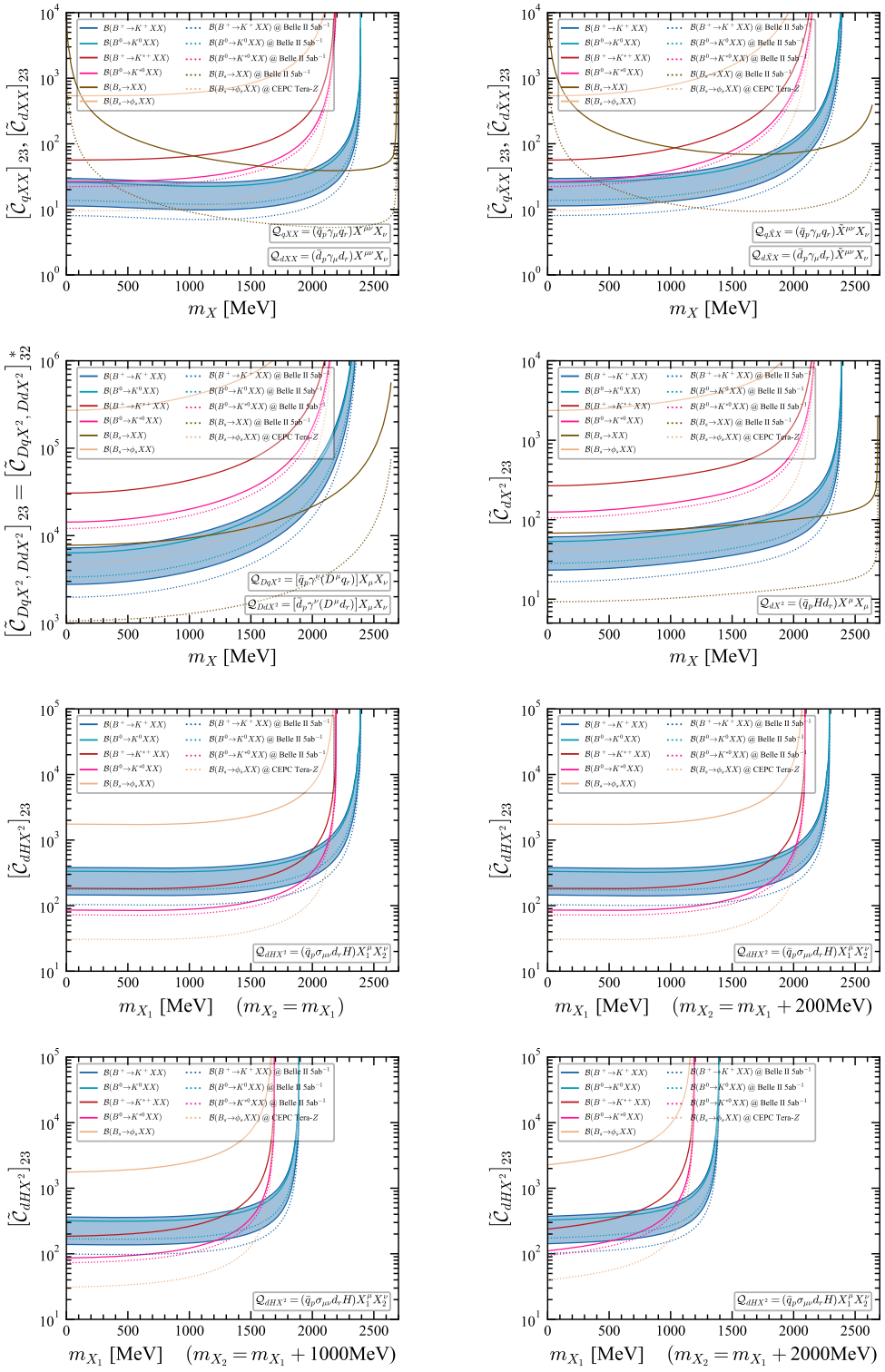



**Figure 6**: Same as in figure 2, but for the operators involving two vector DM fields. For the operator $\mathcal{Q}_{dHX^2}$, we have considered four different cases for the masses of the two DM particles, as indicated in the last four plots.

operator does not contribute to the $B_s \to$ inv decay. Therefore, the high-mass regions do not suffer from any constraint. However, most of the low-mass regions are excluded by the $B^0 \to K^{*0} +$ inv decay, e.g., the one with $m_{X_1} \lesssim 1.8\,\text{GeV}$ is excluded in the case of $m_{X_2} - m_{X_1} = 200\,\text{MeV}$.

All the data on $b \to s$ decays discussed above will be significantly improved in the future [10, 11, 122, 125]. Considering the expected sensitivities at Belle II ($5\,\text{ab}^{-1}$), CEPC and FCC-ee (Tera-$Z$ phase) listed in table 2, we also show in figures 2-6 the projected 90% CL bounds on the DSMEFT operators. It is observed that

- At the Belle II with $5\,\text{ab}^{-1}$, the relative uncertainty of the measured $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ is expected to be 19%. Using the SM prediction to estimate the future experimental central value, the projected bound covers all the parameter space explaining the recent Belle II excess. If the central value of the current Belle II data remains unchanged in the future, the future experimental measurement of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ is expected to be $(23 \pm 4) \times 10^{-6}$, which would deviate from the current SM prediction by about $4\sigma$.

- For the $B^0 \to K^0 +$ inv decay, the projected bound can cover most of the parameter regions required to explain the $B^+ \to K^+ +$ inv excess. Therefore, this decay can provide an independent probe of the proposed mechanism behind the Belle II excess.

- The $B_s \to \phi +$ inv and $B_s \to$ inv decays can strongly constrain the low- and high-mass regions of the parameter space used to explain the Belle II excess, respectively. Their combination can even exclude almost all the parameter regions.

- All the operators involving one DM (such as $\mathcal{Q}_{d\phi}$ and $\mathcal{Q}_{da}$), as well as the operator $\mathcal{Q}_{dHX^2}$, do not contribute to the $B_s \to$ inv decay. Therefore, the high-mass regions do not suffer from any constraint, even at the Belle II with $5\,\text{ab}^{-1}$ or at the CEPC and FCC-ee during the Tera-$Z$ phase.

#### 5.2.2 MFV

As demonstrated in eq. (4.8), once the MFV hypothesis is assumed, 8 out of the 22 DSMEFT operators do not induce down-type FCNC interactions. Therefore, these operators cannot account for the recent Belle II measurement of the $B^+ \to K^+ +$ inv decay. For the remaining 14 operators, their MFV couplings make the $b \to d$ and $s \to d$ processes also relevant. We consider the constraints from the experimental measurements summarized in table 2, which include the branching ratios of the $B \to K^{(*)}, \pi, \rho +$ inv, $K^+ \to \pi^+ +$ inv, $K_L \to \pi^0 +$ inv and $B_{s,d} \to$ inv decays. It is noted that, as discussed in the last subsection, the $b \to s$ transition in the MFV hypothesis approximately depends only on one Wilson coefficient, which is similar to

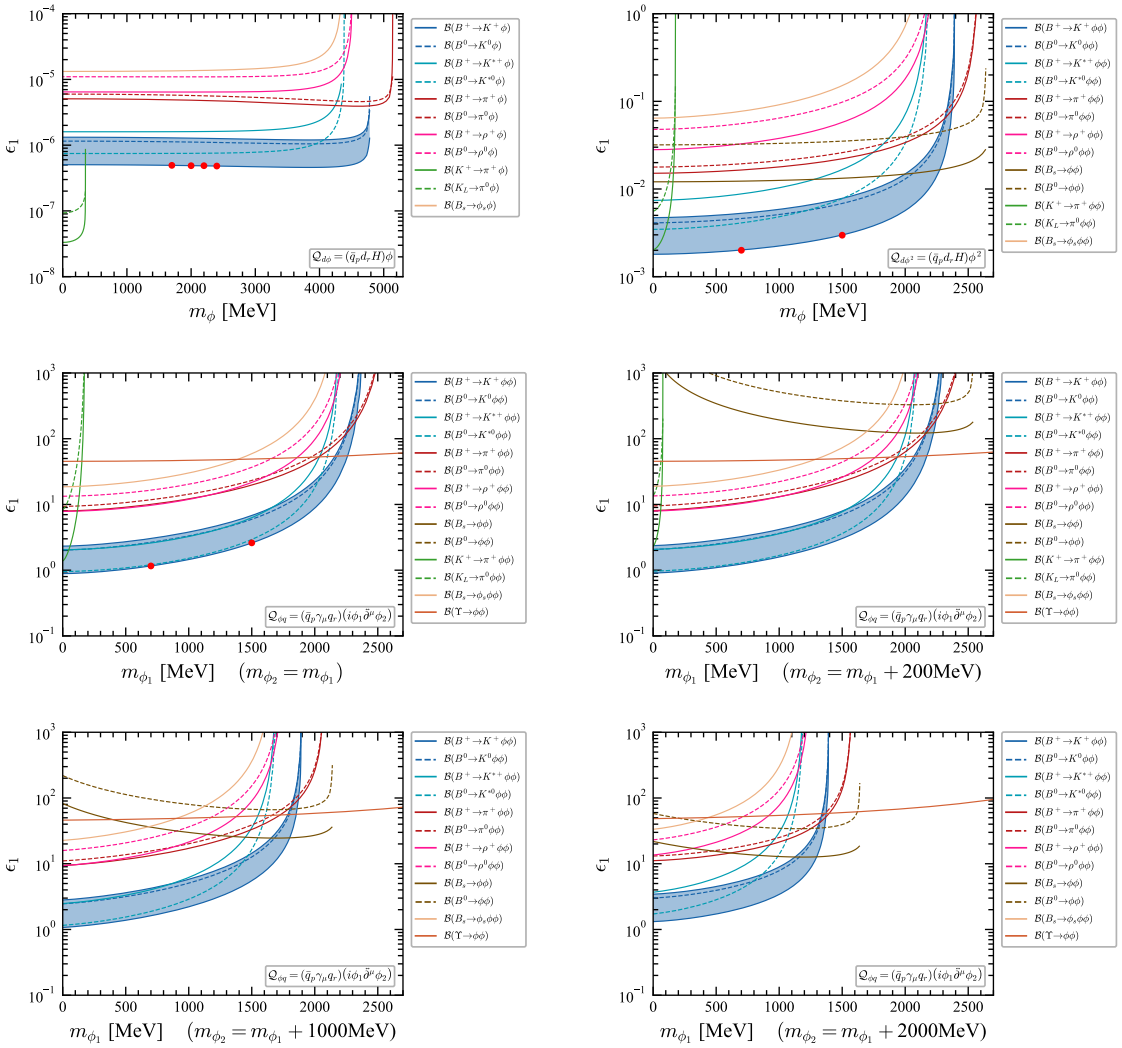



**Figure 7**: Constraints on the MFV parameter $\epsilon_1$ for the operators involving scalar DM as a function of the scalar mass $m_\phi$. The blue band denotes the $2\sigma$ allowed region by the recent Belle II measurement of the $B^+ \to K^+ + \text{inv}$ decay, while the upper bounds obtained from all the other processes are shown by the solid and dashed lines. The red points denote the benchmark points used to investigate the $q^2$ distributions in figure 12; see text for more details.

the treatment in the general DSMEFT discussed in the last section; see footnote [5] for an example. Consequently, the relative strength of the constraints from the various $b \to s$ processes are the same between the general case and the MFV hypothesis. Therefore, in the following, we focus on whether the $s \to d$ and $b \to d$ processes can provide further constraint with respect to that from the $b \to s$ processes.

For the three operators with scalar DM (cf. eq. (4.2)), the allowed regions of the MFV parameter $\epsilon_1^i$ are shown in figure 7. In the low-mass region of $m_\phi \lesssim 200\,\text{MeV}$, the $K^+ \to \pi^+ + \text{inv}$ decay puts the strongest constraint for the operators

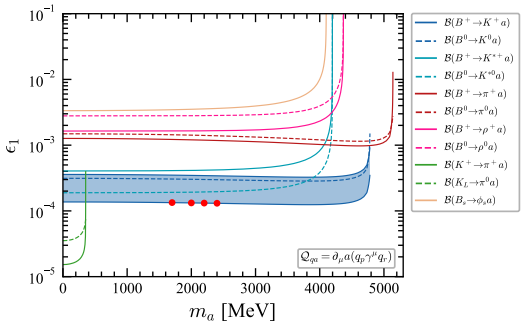



**Figure 8**: Same as in figure 7, but for the operator involving ALP.

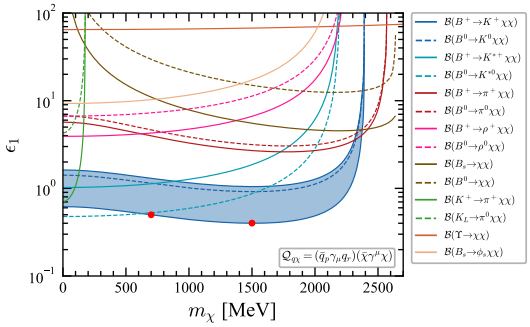



**Figure 9**: Same as in figure 7, but for the operator involving fermionic DM.

$\mathcal{Q}_{d\phi}$ and $\mathcal{Q}_{d\phi^2}$. For all the operators, the $B^+ \to \pi^+ + \text{inv}$ decay provides the strongest constraint among all the $b \to d$ processes. However, this decay provides no further constraint after considering the constraints from the $b \to s$ processes. The only exception is the constraint for the operator $\mathcal{Q}_{\phi q}$ in the case of $m_{\phi_1} = m_{\phi_2} \equiv m_\phi$. In the region of $m_\phi \gtrsim 2.3\,\text{GeV}$, the strongest constraint is provided by the $B^+ \to \pi^+ + \text{inv}$ decay (as well as the $\Upsilon(1S) \to \text{inv}$ decay; see the later discussion), because $m_{\phi_1} = m_{\phi_2}$ makes the NP contribution to the $B_s \to \phi_1\phi_2$ decay vanishing. In addition, $\mathcal{Q}_{qa}$ is the unique operator involving ALP in the MFV framework. The constraints on this operator are shown in figure 8, which are, as in the general case, very similar to the ones for $\mathcal{Q}_{d\phi}$ but with a much larger allowed Wilson coefficient.

Among the operators with fermionic DM, there is only one relevant operator ($\mathcal{Q}_{q\chi}$, cf. eq. (4.3)) in the MFV hypothesis. The experimental constraints are shown in figure 9. We can see that, after considering the constraints from the $b \to s$ processes, all the $s \to d$ and $b \to d$ decays provide no further constraints, except in a narrow window near $m_\chi \approx 2.3\,\text{GeV}$, where the $B^+ \to \pi^+ + \text{inv}$ decay gives the strongest bound.

For the vector DM, in the MFV framework, there are five operators ($\mathcal{Q}_{qXX}$, $\mathcal{Q}_{q\tilde{X}X}$, $\mathcal{Q}_{dX^2}$, $\mathcal{Q}_{DqX^2}$ and $\mathcal{Q}_{dHX^2}$, cf. eq. (4.5)) contributing to the three-body $d_i \to d_j XX$ and four operators ($\mathcal{Q}_{dHX}$, $\mathcal{Q}_{qX}$ and $\mathcal{Q}^{(1,3)}_{HqX}$, cf. eqs. (4.4) and (4.5)) to the two-

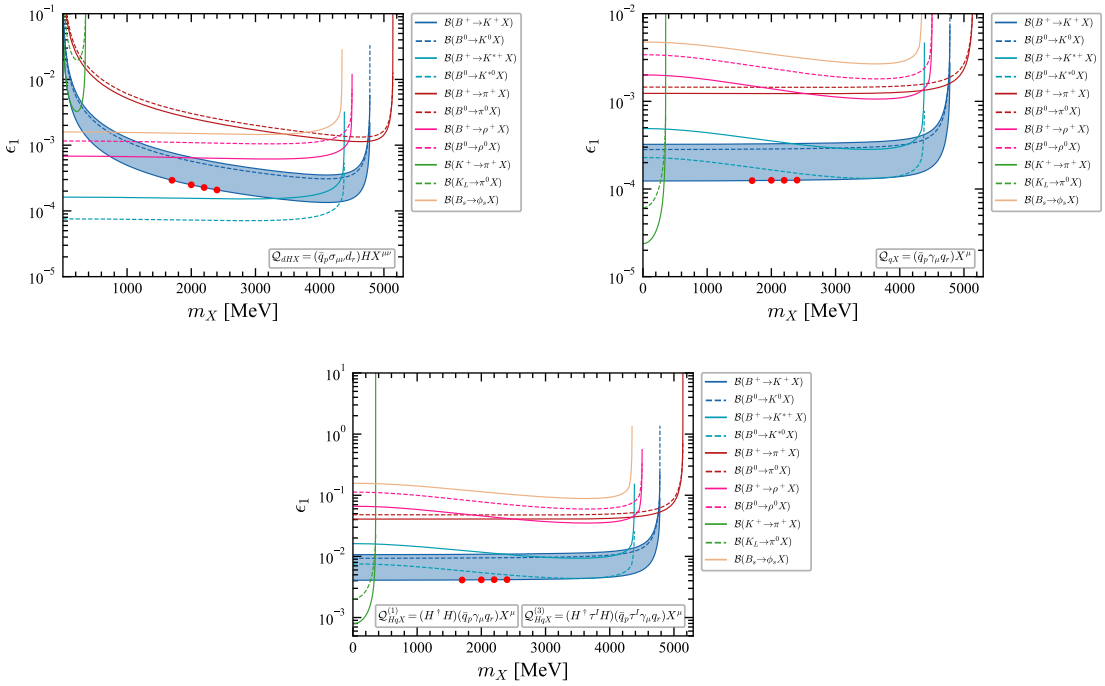



**Figure 10**: Same as in figure 7, but for the operators involving one vector DM.

body $d_i \to d_j X$ decay. The relevant experimental constraints are shown in figures 10 and 11. For the operators $\mathcal{Q}_{qX}$ and $\mathcal{Q}^{(1,3)}_{HqX}$, the $K^+ \to \pi^+ + \text{inv}$ decay provides the strongest constraint for $m_X \lesssim 0.35\,\text{GeV}$ and exclude the parameter region required to account for the Belle II excess. For the four operators involving one DM, the $B^+ \to \pi^+ + \text{inv}$ decay excludes the mass window around $m_X \approx 4.8\,\text{GeV}$. For the four operators involving two DM, neither the $s \to d$ nor the $b \to d$ processes can put further constraints after considering those from the $b \to s$ decays. One exception is the operator $\mathcal{Q}_{dHX^2}$, where the decay $B^+ \to \pi^+ + \text{inv}$ excludes the tail in the high-mass region.

In the MFV hypothesis, as indicated by eqs. (2.16) and (2.17), the flavour-conserving couplings to the third generation are much larger than all the other FCNC couplings.[7] Therefore, the $\Upsilon \to \text{DM+DM}$ decay could be highly enhanced. However, the current experimental bounds on the invisible $\Upsilon$ decays are still quite weak. For the $\Upsilon(1S) \to \text{inv}$ decay, the 90% CL upper limit on its branching ratio is $3.0 \times 10^{-4}$ [126], which is about one order of magnitude higher than the SM prediction $9.9 \times 10^{-6}$ [127–129]. Constraints from this process are also shown in figures 7-11. For

[7] As can be seen from eq. (4.8), the flavour-conserving couplings can also be altered by the MFV parameter $\epsilon_0$. However, this parameter does not affect the FCNC processes, e.g., the $B^+ \to K^+ + \text{inv}$ decay. For simplicity, we take $\epsilon_0 = 0$ to derive the bounds on $\epsilon_1$ from the invisible decays of charmoniums and bottomoniums. For a non-zero $\epsilon_0$, it is straightforward to obtain the corresponding bounds on $\epsilon_1$ by using eq. (4.8).

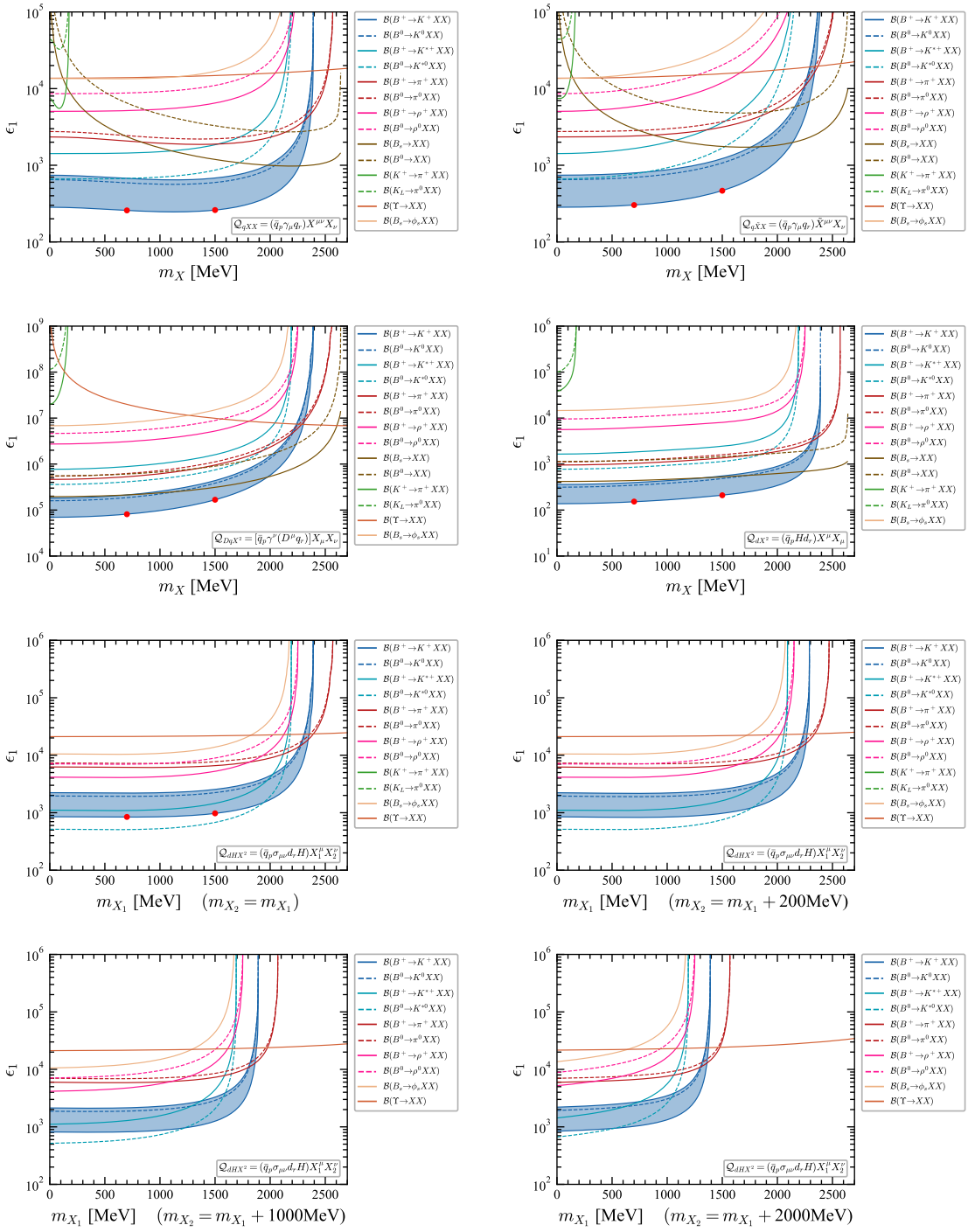



**Figure 11**: Same as in figure 7, but for the operators involving two vector DM fields. For the operator $\mathcal{Q}_{dHX^2}$, we have considered four different cases for the DM masses, as indicated in the last four plots.

the operator $\mathcal{Q}_{\phi q}$ with $m_{\phi_2} - m_{\phi_1} \leq 200\,\text{MeV}$, we can see that the constraints from the $\Upsilon(1S) \to$ inv decay are much stronger than from all the $b \to d$ decays in the high

mass region of $m_{\phi_2} \gtrsim 2\,\text{GeV}$ and, especially, exclude the regions of $m_{\phi_2} \gtrsim 2.3\,\text{GeV}$ required to explain the Belle II excess. However, the constraints for all the other operators are much weaker than from the $B$-meson and kaon decays in the range of $m_{\text{DM}} \lesssim m_B/2$. On the other hand, the constraints from invisible charmonium decays are totally negligible, because the NP contributions are suppressed by a factor of $\mathcal{O}(\lambda_c^4)$ in the MFV hypothesis.

The $q^2$ distributions of the FCNC decays can also provide useful information about the NP effects. Considering the large uncertainty of the Belle II measurement, we take a conservative estimation of the NP contributions to the $q^2$ distributions: For each DSMEFT operator, the minimal value of the MFV parameter allowed by all the experimental constraints is chosen as the benchmark point, which is also shown in figures 7-11. Furthermore, we choose two typical values of the DM masses, 700 MeV and 1500 MeV. By using these benchmark points, we show in figure 12 our predictions for the $q^2$ distributions of $B^+ \to K^+ + \text{inv}$ and $B^0 \to K^{*0} + \text{inv}$ decays.[8] In the case of $m_{\text{DM}} = 700\,\text{MeV}$, we can see that, for the operators $\mathcal{Q}_{\phi q}$ and $\mathcal{Q}_{d\phi^2}$, the distributions are close to each other in the $B^+ \to K^+ + \text{inv}$ decay, but distinguishable in the $B^0 \to K^{*0} + \text{inv}$. However, the distributions for the operators $\mathcal{Q}_{qXX}$ and $\mathcal{Q}_{q\tilde{X}X}$ are quite close in both of these two decays. In the case of $m_{\text{DM}} = 1500\,\text{MeV}$, the distributions are quite different for all the operators, and hence distinguishable from each other.

As can be seen from ref. [9], excesses of the $B^+ \to K^+ + \text{inv}$ events mostly appear in the bins $2 \leq q^2 \leq 7\,\text{GeV}^2$, although the uncertainties are still quite large. From the predictions shown in figure 12, one can see that the $q^2$ distribution resulting from the operator $\mathcal{Q}_{q\chi}$ with $m_\chi = 700\,\text{MeV}$ can closely match that of the Belle II excess. As another possibility, the two-body decay with a DM mass around 2 GeV could also contribute to the observed distribution. In this case, the DSMEFT operators involving one DM field become relevant, and there are 6 such operators in the MFV hypothesis, i.e., $\mathcal{Q}_{d\phi}$, $\mathcal{Q}_{qa}$, $\mathcal{Q}_{dHX}$, $\mathcal{Q}_{qX}$ and $\mathcal{Q}_{HqX}^{(1,3)}$. In order to study the possible excesses in the [2, 3], [3, 4], [4, 5] and [5, 6] $\text{GeV}^2$ bins, we choose $m_{\text{DM}} = 1.7$, 2.0, 2.2 and 2.4 GeV as benchmarks, respectively. For the MFV couplings of these 6 operators, we choose the minimal values allowed by all the experimental constraints, as done for the operators involving two DM fields. By using these benchmark points, we also show in figure 12 the bin-averaged branching ratios of the $B^+ \to K^+ + \text{inv}$ and $B^0 \to K^{*0} + \text{inv}$ decays. Since the minimal values allowed by the Belle II measurement are chosen, the distributions resulting from these operators are the same for the $B^+ \to K^+ + \text{inv}$ decay spectrum. The resulting $B^0 \to K^{*0} + \text{inv}$ decay

[8] The $q^2$ distributions of other $P \to P + \text{inv}$ and $P \to V + \text{inv}$ decays are similar to that of $B^+ \to K^+ + \text{inv}$ and $B^0 \to K^{*0} + \text{inv}$ respectively, and hence are not shown here anymore. In addition, for comparison, we show the $q^2$ distributions of these decays by considering the benchmark values of the Wilson coefficients of the operators $\mathcal{Q}_{dHX}$ and $\mathcal{Q}_{dHX^2}$, although they are already excluded by the $B^0 \to K^{*0} + \text{inv}$ bound.

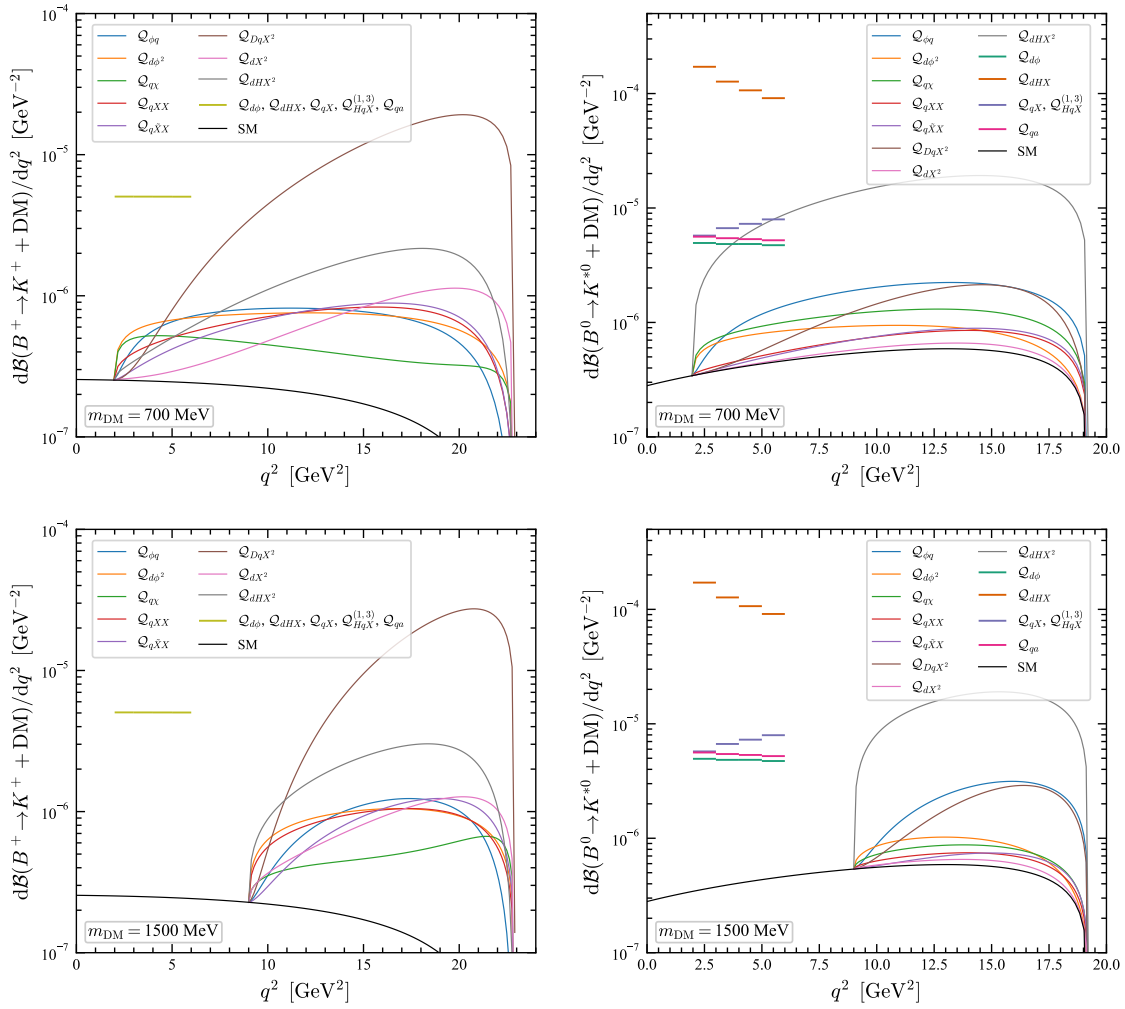



**Figure 12**: Differential branching ratios of the $B^+ \to K^+ + \text{inv}$ and $B^0 \to K^{*0} + \text{inv}$ decays both within the SM and in the DSMEFT with MFV hypothesis. For each effective operator, the DM mass is chosen either as 700 MeV or 1500 MeV, and the MFV parameters are fixed at the corresponding benchmark points shown in figures 7–11. For the operators involving one DM field, $m_{\rm DM} = 1.7$, 2.0, 2.2 and 2.4 GeV are chosen and the corresponding averaged branching ratios in the [2, 3], [3, 4], [4, 5] and [5, 6] GeV$^2$ bins are shown, respectively.

spectra are, however, distinguishable from each other for the operators $\mathcal{Q}_{dHX}$, $\mathcal{Q}_{qa}$ and $\mathcal{Q}_{d\phi}$. They can also be distinguished from that resulting from the operators $\mathcal{Q}_{qX}$ and $\mathcal{Q}^{(1,3)}_{HqX}$, which have the same distribution. As the two-body decay kinematics is quite different from that of the three-body decay, a dedicated statistical analysis can provide more insightful information, for which we refer to ref. [21] for a recent study.

From figures 7-11, we can see that the benchmark points actually correspond to the minimal allowed parameter values required to explain the $B^+ \to K^+ + \text{inv}$ excess. They are irrelevant to the $b \to d + \text{inv}$ and $s \to d + \text{inv}$ decays. Therefore, the

predicted distributions of the $B^+ \to K^+ + \text{inv}$ and $B^0 \to K^{*0} + \text{inv}$ decays in figure 12 can also be applied to the general DSMEFT case discussed in the last subsection.

## 6 Conclusion

The recent Belle II measurement of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ is about $2.7\sigma$ higher than the SM prediction. In this work, we have deciphered the data with two different NP scenarios: the underlying quark-level $b \to s\nu\bar{\nu}$ transition is, besides the SM contribution, further affected by some heavy new mediators that are much heavier than the electroweak scale, or amended by an additional decay channel with light new final states that are sufficiently long-lived to escape the detector and appear as missing-energy signals. To make our analyses model-independent as much as possible, we have studied these two scenarios in the SMEFT and the DSMEFT framework, respectively. Furthermore, the flavour structures of the resulting effective operators are taken to be either of the most generic form or satisfy the MFV hypothesis, both for the quark and lepton sectors.

In the first scenario, we have focused on the implications of the MFV hypothesis. It is found that, once the MFV hypothesis is assumed for the quark sector while without any assumption on the lepton sector, only one SM-like low-energy effective operator $\mathcal{O}_L^{\nu_i\nu_j}$ can be induced by the SMEFT dim-6 operators. As a consequence, the ratio of the branching fractions of any two $b \to s\nu\bar{\nu}$ decays remains a constant, which can be used to test the MFV hypothesis. In particular, the resulting ratio $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})_{\text{MFV}}/\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})_{\text{MFV}} = 0.46 \pm 0.07$ is inconsistent with the current 90% CL lower limit 0.70 derived by combing the recent Belle II measurement of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ and the Belle upper bound on $\mathcal{B}(B^0 \to K^{*0}\nu\bar{\nu})$. Thus, the Belle II excess cannot be explained in the SMEFT with MFV hypothesis for the quark sector. Such a conclusion is true irrespective of the flavour structure of the lepton sector. As the uncertainty of the Belle II measurement of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ is still quite large, we have also used all the relevant processes but without the $B^+ \to K^+\nu\bar{\nu}$ decay to constrain the parameter space of the MFV parameters, and then predicted that $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})_{\text{MFV}} < 10.5 \times 10^{-6}$ and $3.5 < \mathcal{B}(B^+ \to K^+\nu\bar{\nu})_{\text{MFV}} < 5.0 \times 10^{-6}$, which correspond to the cases without and with the leptonic MFV, respectively. These predictions can be further tested at the Belle II with more statistics.

In the second scenario, it is found that the Belle II excess can be accommodated by 22 DSMEFT operators with dim $\leq 6$, which include 4 operators involving scalar, 2 operators involving fermionic, 14 operators involving vector DM, as well as 2 operators involving ALP. For each operator, the current experimental constraints on the corresponding Wilson coefficient and DM (ALP) mass are derived. For most of these operators, large parts of the parameter spaces required to account for the $B^+ \to K^+ + \text{inv}$ excess are excluded by the $B^0 \to K^{*0} + \text{inv}$ and $B_s \to \text{inv}$ decays. Nevertheless, none of the operators is fully excluded by the current data. The future

Belle II ($5\,\text{ab}^{-1}$), CEPC and FCC-ee (Tera-$Z$ phase) data on $B_s \to \phi + \text{inv}$ and $B_s \to \text{inv}$ decays are, however, expected to cover almost all the parameter spaces needed for explaining the Belle II excess. Furthermore, the $B^0 \to K^0 + \text{inv}$ decay can provide an independent probe of the NP mechanism proposed for the Belle II excess. Once the MFV hypothesis is assumed, 8 out of the 22 viable DSMEFT operators do not induce the down-type FCNC interactions, and hence cannot account for the Belle II excess. For the remaining 14 operators, their MFV couplings make the $b \to d$ and $s \to d$ processes also relevant. For $\mathcal{Q}_{d\phi}$, $\mathcal{Q}_{qa}$, $\mathcal{Q}_{qX}$ and $\mathcal{Q}_{HqX}^{(1,3)}$, which all involve one DM field, the $K^+ \to \pi^+ + \text{inv}$ decay provides the strongest constraint in the low-mass region $m_{\text{DM}} \lesssim m_K - m_\pi$. For $\mathcal{Q}_{q\chi}$, $\mathcal{Q}_{dHX}$, $\mathcal{Q}_{qX}$, $\mathcal{Q}_{HqX}^{(1,3)}$ and $\mathcal{Q}_{dHX^2}$, especially for the ones (e.g., $\mathcal{Q}_{dHX^2}$) that make no contribution to the $B_s \to \text{inv}$ decay, the $B^+ \to \pi^+ + \text{inv}$ decay puts the most stringent bound in the high-mass region. In addition, the $\Upsilon(1S) \to \text{inv}$ decay provides the strongest constraint in the high-mass region for the operator $\mathcal{Q}_{\phi q}$ with $m_{\phi_2} - m_{\phi_1} < 200\,\text{MeV}$. Within the parameter space allowed by all the current experimental data, the $q^2$ distributions of the $B \to K^{(*)} + \text{inv}$ decays are then studied for each viable operator. We find that the resulting prediction of the operator $\mathcal{Q}_{q\chi} = \left(\bar{q}_p \gamma_\mu q_r\right)\left(\bar{\chi}\gamma^\mu \chi\right)$ with a fermionic dark matter mass $m_\chi \approx 700\,\text{MeV}$ can closely match the Belle II event distribution in the bins $2 \le q^2 \le 7\,\text{GeV}^2$.

At the Belle II with $5\,\text{ab}^{-1}$ dataset, the branching ratio $\mathcal{B}(B^+ \to K^+ + \text{inv})$ is expected to be measured with an uncertainty of 19%. If the central value of the current Belle II measurement remains unchanged but the uncertainty is reduced accordingly in the future, the discrepancy from the SM prediction will increase to $4\sigma$ level. At the same time, in the future, the $B^0 \to K^0 + \text{inv}$, $B^0 \to K^{*0} + \text{inv}$, $B^+ \to \pi^+ + \text{inv}$ and $B_s \to \text{inv}$ decays at Belle II [10, 11], the $B_s \to \phi + \text{inv}$ at CEPC [122] and FCC-ee [125] as well as the $K \to \pi + \text{inv}$ at NA62 [130] and KOTO [131], are all expected to provide complementary information about the (D)SMEFT operators as well as the validity of the MFV hypothesis studied here.

The new singlets introduced in the DSMEFT could be the DM candidates in the universe. In this case, the parameter space derived in our work could be further constrained by cosmological observations, such as the DM relic density and the Big Bang nucleosynthesis, as well as the DM direct detection [132–136]. For the operators involving one new particle, their couplings should also suffer from additional constraints to ensure that the particle is sufficiently long-lived to escape the detector. In addition, the DM FCNC couplings can affect the neutral-meson mixings, such as $B_s - \bar{B}_s$ mixing [114, 115]. At the LHC, the large couplings of the new singlets to the third generation in MFV can lead to the DM production associated with top and bottom quarks [137]. However, a complete and dedicated study of all these constraints should go beyond the EFT framework and specific UV completions are needed. Therefore, we leave these explorations for our future work.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant Nos. 12135006, 12075097 and 11805077, as well as by the Fundamental Research Funds for the Central Universities under Grant No. CCNU22LJ004. XY is also supported in part by the Startup Research Funding from CCNU.

## A Hadronic matrix elements

For the $B$-meson and kaon decays discussed in sections 3 and 4, as well as the invisible decays of charmoniums and bottomoniums, we need the hadronic matrix elements between the vacuum and the initial hadronic states or between the initial and the final hadronic state, which can be parameterized by the decay constants or by the transition form factors. For convenience, they are recapitulated in appendices A.1 and A.2, respectively. The matrix elements of other bilinear quark currents not mentioned explicitly are zero due to parity invariance of strong interaction.

### A.1 Decay constants

For the $B_q$ meson ($q = u, d$), the non-vanishing annihilation matrix elements are defined as

$$\langle 0|\bar{q}\gamma_\mu\gamma_5 b|\bar{B}_q(p)\rangle = if_{B_q}\, p_\mu, \qquad \langle 0|\bar{q}\gamma_5 b|\bar{B}_q(p)\rangle = -i\frac{m_{B_q}^2}{\overline{m}_q(\mu)+\overline{m}_b(\mu)} f_{B_q}, \tag{A.1}$$

with $\overline{m}_b(\mu)$ and $\overline{m}_q(\mu)$ denoting the $b$- and $q$-quark running masses at the scale $\mu$ defined in the $\overline{\rm MS}$ scheme. Here $f_{B_q}$ denote the $B_q$-meson decay constants, and their up-to-date Lattice QCD values can be found in ref. [138] and references therein. For the vector quarkonium $V$ (such as $\Upsilon$ and $J/\psi$), the relevant hadronic matrix elements are given by [139, 140]

$$\begin{aligned} \langle 0|\bar{q}\gamma^\mu q|V(p,\varepsilon)\rangle &= f_V m_V \varepsilon^\mu, \\ \langle 0|\bar{q}\sigma^{\mu\nu} q|V(p,\varepsilon)\rangle &= if_V^T(\mu)\big(\varepsilon^\mu p^\nu - \varepsilon^\nu p^\mu\big), \end{aligned} \tag{A.2}$$

where $\varepsilon^\mu$ denotes the polarization vector of the vector quarkonium, and $f_V^{(T)}$ is the decay constant. The matrix element of the axial-tensor current can be related to that of the tensor current, while the ones of all the other currents are zero. In the numerical analysis, we use the values of $f_\Upsilon$ in ref. [140], $f_{J/\psi}$ in ref. [141], and $f_V^T(2\,\text{GeV})/f_V$ in ref. [142].

### A.2 Transition form factors

The transition form factors for a $B$ meson decaying into a pseudoscalar meson $P$ (e.g., $B \to K$) are defined by [143, 144]

$$c_P\langle P(p')|\bar{q}\gamma^\mu b|\bar{B}(p)\rangle = f_+(q^2)\left(P^\mu - \frac{m_B^2-m_P^2}{q^2}q^\mu\right) + f_0(q^2)\frac{m_B^2-m_P^2}{q^2}q^\mu,$$

$$
\begin{aligned}
c_P\langle P(p')|\bar{q}b|\bar{B}(p)\rangle &= \frac{m_B^2 - m_P^2}{\overline{m}_b(\mu) - \overline{m}_q(\mu)} f_0(q^2), \\
c_P\langle P(p')|\bar{q}\sigma^{\mu\nu}b(p)|\bar{B}(p)\rangle &= i\frac{P^\mu q^\nu - P^\nu q^\mu}{m_B + m_P} f_T(q^2),
\end{aligned} \tag{A.3}
$$

where $q = p - p'$, $P = p + p'$, and $c_P = -\sqrt{2}$ for $\pi^0$ and $c_P = 1$ otherwise. For a $B$ meson decaying into a vector meson $V$, the relevant transition form factors are introduced by the following Lorentz decompositions of the bilinear quark current matrix elements [143, 145]:

$$
\begin{aligned}
c_V\langle V(p',\varepsilon^*)|\bar{q}\gamma_\mu b|\bar{B}(p)\rangle &= -\epsilon_{\mu\nu\rho\sigma}\varepsilon^{*\nu}P^\rho q^\sigma \frac{V(q^2)}{m_B + m_V}, \\
c_V\langle V(p',\varepsilon^*)|\bar{q}\gamma^\mu\gamma^5 b|\bar{B}(p)\rangle &= i\varepsilon^{*\mu}(m_B + m_V)A_1(q^2) - iP^\mu\frac{\varepsilon^*\cdot q}{m_B + m_V}A_2(q^2) \\
&\quad - iq^\mu(\varepsilon^*\cdot q)\frac{2m_V}{q^2}\left[A_3(q^2) - A_0(q^2)\right], \\
c_V\langle V(p',\varepsilon^*)|\bar{q}\gamma^5 b|\bar{B}(p)\rangle &= -2i\frac{m_V}{\overline{m}_b(\mu) + \overline{m}_q(\mu)}(\varepsilon^*\cdot q)A_0(q^2), \\
c_V\langle V(p',\varepsilon^*)|\bar{q}\sigma_{\mu\nu}q^\nu b|\bar{B}(p)\rangle &= -i\epsilon_{\mu\nu\rho\sigma}\varepsilon^{*\nu}P^\rho q^\sigma T_1(q^2), \\
c_V\langle V(p',\varepsilon^*)|\bar{q}\sigma_{\mu\nu}q^\nu\gamma^5 b|\bar{B}(p)\rangle &= \left[(m_B^2 - m_V^2)\varepsilon_\mu^* - (\varepsilon^*\cdot q)P_\mu\right]T_2(q^2) \\
&\quad + (\varepsilon^*\cdot q)\left(q_\mu - \frac{q^2}{m_B^2 - m_V^2}P_\mu\right)T_3(q^2),
\end{aligned} \tag{A.4}
$$

where $\varepsilon$ denotes the polarization vector of the vector meson, and $c_V = -\sqrt{2}$ for $\rho^0$ and $c_V = 1$ otherwise. We adopt the convention $\epsilon^{0123} = +1$ and $\sigma_{\mu\nu} = i/2\,[\gamma_\mu, \gamma_\nu]$. With the above parameterizations, the absence of dynamical singularities at $q^2 = 0$ implies that $f_+(0) = f_0(0)$ and $A_0(0) = A_3(0)$. In addition, the algebraic relation between $\sigma_{\mu\nu}$ and $\sigma_{\mu\nu}\gamma_5$, $\sigma_{\mu\nu}\gamma_5 = i/2\epsilon_{\mu\nu\alpha\beta}\sigma^{\alpha\beta}$, gives rise to the identity $T_1(0) = T_2(0)$. It is also noted that the form factor $A_3$ is redundant and can be written as a linear combination of $A_1$ and $A_2$ [145]

$$
A_3(q^2) = \frac{(m_B + m_V)}{2m_V}A_1(q^2) - \frac{m_B - m_V}{2m_V}A_2(q^2). \tag{A.5}
$$

The hadronic matrix elements of the tensor and axial-tensor currents can also be decomposed as [146]

$$
\begin{aligned}
c_V\langle V(p',\varepsilon^*)|\bar{q}\sigma^{\mu\nu}b|\bar{B}(p)\rangle =& i\epsilon_{\mu\nu\rho\sigma}\left[P^\rho\varepsilon^{*\sigma} - \frac{m_B^2 - m_V^2}{q^2}q^\rho\varepsilon^{*\sigma} + \frac{\varepsilon^*\cdot q}{q^2}q^\rho P^\sigma\right]T_1(q^2) \\
&+ i\epsilon_{\mu\nu\rho\sigma}\left[\frac{m_B^2 - m_V^2}{q^2}q^\rho\varepsilon^{*\sigma} + \frac{\varepsilon^*\cdot q}{q^2}P^\rho q^\sigma\right]T_2(q^2) \\
&+ i\epsilon_{\mu\nu\rho\sigma}\frac{\varepsilon^*\cdot q}{m_B^2 - m_V^2}P^\rho q^\sigma T_3(q^2),
\end{aligned}
$$

$$
\begin{aligned}
c_V\langle V(p',\varepsilon^*)|\bar{q}\sigma^{\mu\nu}\gamma^5 b|\bar{B}(p)\rangle = &\Bigg[(\varepsilon^{*\mu}P^\nu - P^\mu\varepsilon^{*\nu}) + \frac{m_B^2 - m_V^2}{q^2}(q^\mu\varepsilon^{*\nu} - \varepsilon^{*\mu}q^\nu) \\
&+ \frac{\varepsilon^*\cdot q}{q^2}(P^\mu q^\nu - q^\mu P^\nu)\Bigg] T_1(q^2) \\
&+ \left[\frac{m_B^2 - m_V^2}{q^2}(\varepsilon^{*\mu}q^\nu - q^\mu\varepsilon^{*\nu}) + \frac{\varepsilon^*\cdot q}{q^2}(q^\mu P^\nu - P^\mu q^\nu)\right] T_2(q^2) \\
&+ \frac{\varepsilon^*\cdot q}{m_B^2 - m_V^2}\,(q^\mu P^\nu - P^\mu q^\nu)\, T_3(q^2), \qquad \text{(A.6)}
\end{aligned}
$$

from which the forms given in eq. (A.4) can be obtained by multiplying both sides by $q_\nu$. All the $B_s \to P\,(V)$ and $K \to \pi$ form factors can be defined similarly.

Numerically, we use the values obtained in ref. [144] for the $B \to P$ and in ref. [145] for the $B \to V$ transition form factors. As the Lattice QCD results are still not available for all of these form factors, especially for the tensor ones (see ref. [138] and references therein), we use instead the calculations based on the light-cone sum rule approach [144, 145]. For the $K \to \pi$ vector and scalar form factors, the results obtained in refs. [147, 148] are adopted, while the value in ref. [149] is used for the $K \to \pi$ tensor form factor.

For the operator $\mathcal{Q}_{DdX^2}$ defined in eq. (4.5), the relevant form factors are not available in the literature so far. However, when the Wilson coefficient $\mathcal{C}_{DdX^2}$ is assumed to be hermitian, they can be derived from the above form factors. Taking the $\bar{B} \to \bar{K}^0$ transition as an example, we have the relation $\langle\bar{K}^0(p')|(\partial^\mu\bar{s})\gamma^\nu P_L b + \bar{s}\gamma^\nu P_L(\partial^\mu b)|\bar{B}(p)\rangle = -iq^\mu\langle\bar{K}^0(p')|\bar{s}\gamma^\nu P_L b|\bar{B}(p)\rangle$.

## References


[1] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D* **2** (1970) 1285–1292.

[2] W. Altmannshofer, A. J. Buras, D. M. Straub, and M. Wick, *New strategies for New Physics search in $B \to K^*\nu\bar{\nu}$, $B \to K\nu\bar{\nu}$ and $B \to X_s\nu\bar{\nu}$ decays*, *JHEP* **04** (2009) 022, [arXiv:0902.0160].

[3] A. J. Buras, J. Girrbach-Noe, C. Niehoff, and D. M. Straub, *$B \to K^{(*)}\nu\bar{\nu}$ decays in the Standard Model and beyond*, *JHEP* **02** (2015) 184, [arXiv:1409.4557].

[4] D. Bečirević, G. Piazza, and O. Sumensari, *Revisiting $B \to K^{(*)}\nu\bar{\nu}$ decays in the Standard Model and beyond*, *Eur. Phys. J. C* **83** (2023), no. 3 252, [arXiv:2301.06990].

[5] T. Felkl, S. L. Li, and M. A. Schmidt, *A tale of invisibility: constraints on new physics in $b \to s\nu\nu$*, *JHEP* **12** (2021) 118, [arXiv:2111.04327].

[6] R. Bause, H. Gisbert, M. Golz, and G. Hiller, *Interplay of dineutrino modes with semileptonic rare B-decays*, *JHEP* **12** (2021) 061, [arXiv:2109.01675].

[7] X. G. He and G. Valencia, *$R^{\nu}_{K(*)}$ and non-standard neutrino interactions*, *Phys. Lett. B* **821** (2021) 136607, [arXiv:2108.05033].

[8] T. E. Browder, N. G. Deshpande, R. Mandal, and R. Sinha, *Impact of $B \to K\nu\bar{\nu}$ measurements on beyond the Standard Model theories*, *Phys. Rev. D* **104** (2021), no. 5 053007, [arXiv:2107.01080].

[9] **Belle-II** Collaboration, I. Adachi et al., *Evidence for $B^+ \to K^+\nu\bar{\nu}$ Decays*, arXiv:2311.14647.

[10] **Belle-II** Collaboration, W. Altmannshofer et al., *The Belle II Physics Book*, *PTEP* **2019** (2019), no. 12 123C01, [arXiv:1808.10567]. [Erratum: PTEP 2020, 029201 (2020)].

[11] **Belle-II** Collaboration, L. Aggarwal et al., *Snowmass White Paper: Belle II physics reach and plans for the next decade and beyond*, arXiv:2207.06307.

[12] P. Athron, R. Martinez, and C. Sierra, *$B$ meson anomalies and large $B^+ \to K^+\nu\bar{\nu}$ in non-universal $U(1)'$ models*, arXiv:2308.13426.

[13] R. Bause, H. Gisbert, and G. Hiller, *Implications of an enhanced $B \to K\nu\bar{\nu}$ branching ratio*, *Phys. Rev. D* **109** (2024), no. 1 015006, [arXiv:2309.00075].

[14] L. Allwicher, D. Becirevic, G. Piazza, S. Rosauro-Alcaraz, and O. Sumensari, *Understanding the first measurement of $\mathcal{B}(B \to K\nu\bar{\nu})$*, *Phys. Lett. B* **848** (2024) 138411, [arXiv:2309.02246].

[15] T. Felkl, A. Giri, R. Mohanta, and M. A. Schmidt, *When energy goes missing: new physics in $b \to s\nu\nu$ with sterile neutrinos*, *Eur. Phys. J. C* **83** (2023), no. 12 1135, [arXiv:2309.02940].

[16] M. Abdughani and Y. Reyimuaji, *Constraining light dark matter and mediator with $B^+ \to K^+\nu\bar{\nu}$ data*, arXiv:2309.03706.

[17] H. K. Dreiner, J. Y. Günther, and Z. S. Wang, *The Decay $B \to K\nu\bar{\nu}$ at Belle II and a Massless Bino in R-parity-violating Supersymmetry*, arXiv:2309.03727.

[18] X.-G. He, X.-D. Ma, and G. Valencia, *Revisiting models that enhance $B^+ \to K^+\nu\bar{\nu}$ in light of the new Belle II measurement*, arXiv:2309.12741.

[19] A. Berezhnoy and D. Melikhov, *$B \to K^* M_X$ vs $B \to K M_X$ as a probe of a scalar-mediator dark matter scenario*, *EPL* **145** (2024), no. 1 14001, [arXiv:2309.17191].

[20] A. Datta, D. Marfatia, and L. Mukherjee, *$B \to K\nu\bar{\nu}$, MiniBooNE and muon g-2 anomalies from a dark sector*, *Phys. Rev. D* **109** (2024), no. 3 L031701, [arXiv:2310.15136].

[21] W. Altmannshofer, A. Crivellin, H. Haigh, G. Inguglia, and J. Martin Camalich, *Light New Physics in $B \to K^{(*)}\nu\bar{\nu}$?*, arXiv:2311.14629.

[22] D. McKeen, J. N. Ng, and D. Tuckler, *Higgs Portal Interpretation of the Belle II $B^+ \to K^+\nu\nu$ Measurement*, arXiv:2312.00982.

[23] K. Fridell, M. Ghosh, T. Okui, and K. Tobioka, *Decoding the $B \to K\nu\nu$ excess at Belle II: kinematics, operators, and masses*, arXiv:2312.12507.

[24] S.-Y. Ho, J. Kim, and P. Ko, *Recent $B^+ \to K^+\nu\bar{\nu}$ Excess and Muon $g-2$ Illuminating Light Dark Sector with Higgs Portal*, arXiv:2401.10112.

[25] F.-Z. Chen, Q. Wen, and F. Xu, *Correlating $B \to K^{(*)}\nu\bar{\nu}$ and flavor anomalies in SMEFT*, arXiv:2401.11552.

[26] E. Gabrielli, L. Marzola, K. Müürsepp, and M. Raidal, *Explaining the $B^+ \to K^+\nu\bar{\nu}$ excess via a massless dark photon*, arXiv:2402.05901.

[27] T. Li, Z. Qian, M. A. Schmidt, and M. Yuan, *The quark flavor-violating ALPs in light of B mesons and hadron colliders*, arXiv:2402.14232.

[28] **Particle Data Group** Collaboration, R. L. Workman et al., *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01. and 2023 update.

[29] G. Alonso-Álvarez and M. Escudero, *The first limit on invisible decays of $B_s$ mesons comes from LEP*, arXiv:2310.13043.

[30] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1–98, [arXiv:1706.08945].

[31] G. Isidori, F. Wilsch, and D. Wyler, *The Standard Model effective field theory at work*, arXiv:2303.16922.

[32] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, *Rev. Mod. Phys.* **68** (1996) 1125–1144, [hep-ph/9512380].

[33] E. E. Jenkins, A. V. Manohar, and P. Stoffer, *Low-Energy Effective Field Theory below the Electroweak Scale: Operators and Matching*, *JHEP* **03** (2018) 016, [arXiv:1709.04486].

[34] W. Dekens and P. Stoffer, *Low-energy effective field theory below the electroweak scale: matching at one loop*, *JHEP* **10** (2019) 197, [arXiv:1908.05295]. [Erratum: JHEP 11, 148 (2022)].

[35] J. Aebischer, A. Crivellin, M. Fael, and C. Greub, *Matching of gauge invariant dimension-six operators for $b \to s$ and $b \to c$ transitions*, *JHEP* **05** (2016) 037, [arXiv:1512.02830].

[36] E. Del Nobile and F. Sannino, *Dark Matter Effective Theory*, *Int. J. Mod. Phys. A* **27** (2012) 1250065, [arXiv:1102.3116].

[37] M. Baumgart et al., *Snowmass White Paper: Effective Field Theories for Dark Matter Phenomenology*, arXiv:2203.08204.

[38] M. Duch, B. Grzadkowski, and J. Wudka, *Classification of effective operators for interactions between the Standard Model and dark matter*, *JHEP* **05** (2015) 116, [arXiv:1412.0520].

[39] J. F. Kamenik and C. Smith, *FCNC portals to the dark sector*, *JHEP* **03** (2012) 090, [arXiv:1111.6402].

[40] J. Brod, A. Gootjes-Dreesbach, M. Tammaro, and J. Zupan, *Effective Field Theory for Dark Matter Direct Detection up to Dimension Seven*, *JHEP* **10** (2018) 065, [arXiv:1710.10218]. [Erratum: JHEP 07, 012 (2023)].

[41] J. C. Criado, A. Djouadi, M. Perez-Victoria, and J. Santiago, *A complete effective field theory for dark matter*, *JHEP* **07** (2021) 081, [arXiv:2104.14443].

[42] J. Aebischer, W. Altmannshofer, E. E. Jenkins, and A. V. Manohar, *Dark matter effective field theory and an application to vector dark matter*, *JHEP* **06** (2022) 086, [arXiv:2202.06968].

[43] H. Song, H. Sun, and J.-H. Yu, *Complete EFT Operator Bases for Dark Matter and Weakly-Interacting Light Particle*, arXiv:2306.05999.

[44] X.-G. He, X.-D. Ma, and G. Valencia, *FCNC B and K meson decays with light bosonic Dark Matter*, *JHEP* **03** (2023) 037, [arXiv:2209.05223].

[45] J.-H. Liang, Y. Liao, X.-D. Ma, and H.-L. Wang, *Dark sector effective field theory*, *JHEP* **12** (2023) 172, [arXiv:2309.12166].

[46] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.

[47] R. D. Peccei and H. R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791–1797.

[48] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223–226.

[49] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279–282.

[50] H. Georgi, D. B. Kaplan, and L. Randall, *Manifesting the Invisible Axion at Low-energies*, *Phys. Lett. B* **169** (1986) 73–78.

[51] I. Brivio, M. B. Gavela, L. Merlo, K. Mimasu, J. M. No, R. del Rey, and V. Sanz, *ALPs Effective Field Theory and Collider Signatures*, *Eur. Phys. J. C* **77** (2017), no. 8 572, [arXiv:1701.05379].

[52] M. Chala, G. Guedes, M. Ramos, and J. Santiago, *Running in the ALPs*, *Eur. Phys. J. C* **81** (2021), no. 2 181, [arXiv:2012.09017].

[53] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, *The Low-Energy Effective Theory of Axions and ALPs*, *JHEP* **04** (2021) 063, [arXiv:2012.12272].

[54] A. M. Galda, M. Neubert, and S. Renner, *ALP — SMEFT interference*, *JHEP* **06** (2021) 135, [arXiv:2105.01078].

[55] H. Song, H. Sun, and J.-H. Yu, *Effective Field Theories of Axion, ALP and Dark Photon*, arXiv:2305.16770.

[56] G. Isidori, Y. Nir, and G. Perez, *Flavor Physics Constraints for Physics Beyond the Standard Model*, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 355, [arXiv:1002.0900].

[57] W. Altmannshofer and J. Zupan, *Snowmass White Paper: Flavor Model Building*, in *Snowmass 2021*, 3, 2022. arXiv:2203.07726.

[58] R. S. Chivukula and H. Georgi, *Composite Technicolor Standard Model*, *Phys. Lett. B* **188** (1987) 99–104.

[59] A. J. Buras, P. Gambino, M. Gorbahn, S. Jager, and L. Silvestrini, *Universal unitarity triangle and physics beyond the standard model*, *Phys. Lett. B* **500** (2001) 161–167, [hep-ph/0007085].

[60] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, *Nucl. Phys. B* **645** (2002) 155–187, [hep-ph/0207036].

[61] G. Colangelo, E. Nikolidakis, and C. Smith, *Supersymmetric models with minimal flavour violation and their running*, *Eur. Phys. J. C* **59** (2009) 75–98, [arXiv:0807.0801].

[62] L. Mercolli and C. Smith, *EDM constraints on flavored CP-violating phases*, *Nucl. Phys. B* **817** (2009) 1–24, [arXiv:0902.1949].

[63] B. Grinstein, X. Lu, L. Merlo, and P. Quílez, *Hilbert series for covariants and their applications to Minimal Flavor Violation*, arXiv:2312.13349.

[64] X.-G. He, C.-J. Lee, S.-F. Li, and J. Tandean, *Fermion EDMs with Minimal Flavor Violation*, *JHEP* **08** (2014) 019, [arXiv:1404.4436].

[65] X.-G. He, C.-J. Lee, S.-F. Li, and J. Tandean, *Large electron electric dipole moment in minimal flavor violation framework with Majorana neutrinos*, *Phys. Rev. D* **89** (2014), no. 9 091901, [arXiv:1401.2615].

[66] X.-G. He, C.-J. Lee, J. Tandean, and Y.-J. Zheng, *Seesaw Models with Minimal Flavor Violation*, *Phys. Rev. D* **91** (2015), no. 7 076008, [arXiv:1411.6612].

[67] C.-W. Chiang, X.-G. He, J. Tandean, and X.-B. Yuan, *$R_{K^{(*)}}$ and related $b \to s\ell\bar{\ell}$ anomalies in minimal flavor violation framework with $Z'$ boson*, *Phys. Rev. D* **96** (2017), no. 11 115022, [arXiv:1706.02696].

[68] J.-J. Zhang, M. He, X.-G. He, and X.-B. Yuan, *Flavor Violating Higgs Couplings in Minimal Flavor Violation*, *JHEP* **02** (2019) 007, [arXiv:1807.00921].

[69] V. Cirigliano, B. Grinstein, G. Isidori, and M. B. Wise, *Minimal flavor violation in the lepton sector*, *Nucl. Phys. B* **728** (2005) 121–134, [hep-ph/0507001].

[70] V. Cirigliano and B. Grinstein, *Phenomenology of minimal lepton flavor violation*, *Nucl. Phys. B* **752** (2006) 18–39, [hep-ph/0601111].

[71] S. Davidson and F. Palorini, *Various definitions of Minimal Flavour Violation for Leptons*, *Phys. Lett. B* **642** (2006) 72–80, [hep-ph/0607329].

[72] G. C. Branco, A. J. Buras, S. Jager, S. Uhlig, and A. Weiler, *Another look at minimal lepton flavour violation, $l_i \to l_{j\gamma}$, leptogenesis, and the ratio $M_\nu/\Lambda_{LFV}$*, *JHEP* **09** (2007) 004, [hep-ph/0609067].

[73] M. B. Gavela, T. Hambye, D. Hernandez, and P. Hernandez, *Minimal Flavour Seesaw Models*, *JHEP* **09** (2009) 038, [arXiv:0906.1461].

[74] R. Alonso, G. Isidori, L. Merlo, L. A. Munoz, and E. Nardi, *Minimal flavour violation extensions of the seesaw*, *JHEP* **06** (2011) 037, [arXiv:1103.5461].

[75] D. Aristizabal Sierra, A. Degee, and J. F. Kamenik, *Minimal Lepton Flavor Violating Realizations of Minimal Seesaw Models*, *JHEP* **07** (2012) 135, [arXiv:1205.5547].

[76] B. Batell, J. Pradler, and M. Spannowsky, *Dark Matter from Minimal Flavor Violation*, *JHEP* **08** (2011) 038, [arXiv:1105.1781].

[77] L. Lopez-Honorez and L. Merlo, *Dark matter within the minimal flavour violation ansatz*, *Phys. Lett. B* **722** (2013) 135–143, [arXiv:1303.1087].

[78] P. Agrawal, M. Blanke, and K. Gemmler, *Flavored dark matter beyond Minimal Flavor Violation*, *JHEP* **10** (2014) 072, [arXiv:1405.6709].

[79] M.-C. Chen, J. Huang, and V. Takhistov, *Beyond Minimal Lepton Flavored Dark Matter*, *JHEP* **02** (2016) 060, [arXiv:1510.04694].

[80] H. Acaroğlu and M. Blanke, *Tasting flavoured Majorana dark matter*, *JHEP* **05** (2022) 086, [arXiv:2109.10357].

[81] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621–653.

[82] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085, [arXiv:1008.4884].

[83] G. Buchalla and A. J. Buras, *QCD corrections to rare K and B decays for arbitrary top quark mass*, *Nucl. Phys. B* **400** (1993) 225–239.

[84] M. Misiak and J. Urban, *QCD corrections to FCNC decays mediated by Z penguins and W boxes*, *Phys. Lett. B* **451** (1999) 161–169, [hep-ph/9901278].

[85] G. Buchalla and A. J. Buras, *The rare decays $K \to \pi\nu\bar{\nu}$, $B \to X\nu\bar{\nu}$ and $B \to l^+l^-$: An Update*, *Nucl. Phys. B* **548** (1999) 309–327, [hep-ph/9901288].

[86] J. Brod, M. Gorbahn, and E. Stamou, *Two-Loop Electroweak Corrections for the $K \to \pi\nu\bar{\nu}$ Decays*, *Phys. Rev. D* **83** (2011) 034030, [arXiv:1009.0947].

[87] P. Colangelo, F. De Fazio, P. Santorelli, and E. Scrimieri, *Rare $B \to K^{(*)}\nu\bar{\nu}$ decays at B factories*, *Phys. Lett. B* **395** (1997) 339–344, [hep-ph/9610297].

[88] J. F. Kamenik and C. Smith, *Tree-level contributions to the rare decays $B^+ \to \pi^+\nu\bar{\nu}$, $B^+ \to K^+\nu\bar{\nu}$, and $B^+ \to K^{*+}\nu\bar{\nu}$ in the Standard Model*, *Phys. Lett. B* **680** (2009) 471–475, [arXiv:0908.1174].

[89] D. London and J. Matias, *B Flavour Anomalies: 2021 Theoretical Status Report*, *Ann. Rev. Nucl. Part. Sci.* **72** (2022) 37–68, [arXiv:2110.13270].

[90] B. Capdevila, A. Crivellin, and J. Matias, *Review of Semileptonic B Anomalies*, *Eur. Phys. J. ST* **1** (2023) 20, [arXiv:2309.01311].

[91] A. Crivellin, L. Hofer, J. Matias, U. Nierste, S. Pokorski, and J. Rosiek,

*Lepton-flavour violating B decays in generic Z′ models*, *Phys. Rev. D* **92** (2015), no. 5 054013, [arXiv:1504.07928].

[92] C.-J. Lee and J. Tandean, *Minimal lepton flavor violation implications of the $b \to s$ anomalies*, *JHEP* **08** (2015) 123, [arXiv:1505.04692].

[93] A. J. Buras, F. Schwab, and S. Uhlig, *Waiting for precise measurements of $K^+ \to \pi^+ \nu\bar{\nu}$ and $K_L \to \pi^0 \nu\bar{\nu}$*, *Rev. Mod. Phys.* **80** (2008) 965–1007, [hep-ph/0405132].

[94] F. Mescia and C. Smith, *Improved estimates of rare K decay matrix-elements from $K_{\ell 3}$ decays*, *Phys. Rev. D* **76** (2007) 034017, [arXiv:0705.2025].

[95] V. Cirigliano, G. Ecker, H. Neufeld, A. Pich, and J. Portoles, *Kaon Decays in the Standard Model*, *Rev. Mod. Phys.* **84** (2012) 399, [arXiv:1107.6001].

[96] A. J. Buras, D. Buttazzo, J. Girrbach-Noe, and R. Knegjens, *$K^+ \to \pi^+ \nu\bar{\nu}$ and $K_L \to \pi^0 \nu\bar{\nu}$ in the Standard Model: status and perspectives*, *JHEP* **11** (2015) 033, [arXiv:1503.02693].

[97] A. J. Buras, M. Gorbahn, U. Haisch, and U. Nierste, *Charm quark contribution to $K^+ \to \pi^+ \nu\bar{\nu}$ at next-to-next-to-leading order*, *JHEP* **11** (2006) 002, [hep-ph/0603079]. [Erratum: JHEP 11, 167 (2012)].

[98] J. Brod and M. Gorbahn, *Electroweak Corrections to the Charm Quark Contribution to $K^+ \to \pi^+ \nu\bar{\nu}$*, *Phys. Rev. D* **78** (2008) 034006, [arXiv:0805.4119].

[99] G. Isidori, F. Mescia, and C. Smith, *Light-quark loops in $K \to \pi\nu\bar{\nu}$*, *Nucl. Phys. B* **718** (2005) 319–338, [hep-ph/0503107].

[100] L. S. Littenberg, *The CP Violating Dacay $K_L^0 \to \pi^0 \nu\bar{\nu}$*, *Phys. Rev. D* **39** (1989) 3322–3324.

[101] G. Buchalla and G. Isidori, *The CP conserving contribution to $K_L \to \pi^0 \nu\bar{\nu}$ neutrino anti-neutrino in the standard model*, *Phys. Lett. B* **440** (1998) 170–178, [hep-ph/9806501].

[102] G. Buchalla and A. J. Buras, *$K_L \to \pi\nu\bar{\nu}$ and high precision determinations of the CKM matrix*, *Phys. Rev. D* **54** (1996) 6782–6789, [hep-ph/9607447].

[103] W. J. Marciano and Z. Parsa, *Rare kaon decays with "missing energy"*, *Phys. Rev. D* **53** (1996), no. 1 R1.

[104] M. Williams, C. P. Burgess, A. Maharana, and F. Quevedo, *New Constraints (and Motivations) for Abelian Gauge Bosons in the MeV-TeV Mass Range*, *JHEP* **08** (2011) 106, [arXiv:1103.4556].

[105] E. E. Jenkins, A. V. Manohar, and P. Stoffer, *Low-Energy Effective Field Theory below the Electroweak Scale: Anomalous Dimensions*, *JHEP* **01** (2018) 084, [arXiv:1711.05270].

[106] C. Bird, P. Jackson, R. V. Kowalewski, and M. Pospelov, *Search for dark matter in*

$b \to s$ *transitions with missing energy*, *Phys. Rev. Lett.* **93** (2004) 201803, [hep-ph/0401195].

[107] G. Li, J.-Y. Su, and J. Tandean, *Flavor-changing hyperon decays with light invisible bosons*, *Phys. Rev. D* **100** (2019), no. 7 075003, [arXiv:1905.08759].

[108] X.-G. He, X.-D. Ma, J. Tandean, and G. Valencia, *Evading the Grossman-Nir bound with* $\Delta I = 3/2$ *new physics*, *JHEP* **08** (2020), no. 08 034, [arXiv:2005.02942].

[109] C.-Q. Geng and J. Tandean, *Probing new physics with the kaon decays* $K \to \pi\pi \not{E}$, *Phys. Rev. D* **102** (2020) 115021, [arXiv:2009.00608].

[110] G. Li, T. Wang, J.-B. Zhang, and G.-L. Wang, *The light invisible boson in FCNC decays of B and* $B_c$ *mesons*, *Eur. Phys. J. C* **81** (2021), no. 6 564, [arXiv:2103.12921].

[111] F. Kling, S. Li, H. Song, S. Su, and W. Su, *Light Scalars at FASER*, *JHEP* **08** (2023) 001, [arXiv:2212.06186].

[112] J.-Y. Su and J. Tandean, *Exploring leptoquark effects in hyperon and kaon decays with missing energy*, *Phys. Rev. D* **102** (2020), no. 7 075032, [arXiv:1912.13507].

[113] G. Li, T. Wang, Y. Jiang, J.-B. Zhang, and G.-L. Wang, *Spin-1/2 invisible particles in heavy meson decays*, *Phys. Rev. D* **102** (2020), no. 9 095019, [arXiv:2004.10942].

[114] J. Martin Camalich, M. Pospelov, P. N. H. Vuong, R. Ziegler, and J. Zupan, *Quark Flavor Phenomenology of the QCD Axion*, *Phys. Rev. D* **102** (2020), no. 1 015023, [arXiv:2002.04623].

[115] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, *Flavor probes of axion-like particles*, *JHEP* **09** (2022) 056, [arXiv:2110.10698].

[116] A. W. M. Guerrera and S. Rigolin, *ALP Production in Weak Mesonic Decays*, *Fortsch. Phys.* **71** (2023), no. 2-3 2200192, [arXiv:2211.08343].

[117] B.-F. Hou, X.-Q. Li, H. Yan, Y.-D. Yang, and X.-B. Yuan, "`HadronToNP`: a package to calculate decay of hadron to new particles." in preparation.

[118] **CKMfitter Group** Collaboration, J. Charles, A. Hocker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malcles, J. Ocariz, M. Pivk, and L. Roos, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, *Eur. Phys. J.* **C41** (2005), no. 1 1–131, [hep-ph/0406184]. Updated results and plots available at: http://ckmfitter.in2p3.fr.

[119] A. Glazov, "Belle II physics highlights." plenary talk given at the EPS-HEP2023 Conference in Hamburg (Germany), 2023.

[120] **Belle** Collaboration, J. Grygier et al., *Search for* $B \to h\nu\bar{\nu}$ *decays with semileptonic tagging at Belle*, *Phys. Rev. D* **96** (2017), no. 9 091101, [arXiv:1702.03224]. [Addendum: Phys.Rev.D 97, 099902 (2018)].

[121] **DELPHI** Collaboration, W. Adam et al., *Study of rare b decays with the DELPHI detector at LEP*, *Z. Phys. C* **72** (1996) 207–220.

[122] L. Li, M. Ruan, Y. Wang, and Y. Wang, *Analysis of $B_s \to \phi\nu\bar{\nu}$ at CEPC*, *Phys. Rev. D* **105** (2022), no. 11 114036, [arXiv:2201.07374].

[123] **NA62** Collaboration, E. Cortina Gil et al., *Measurement of the very rare $K^+ \to \pi^+\nu\bar{\nu}$ decay*, *JHEP* **06** (2021) 093, [arXiv:2103.15389].

[124] **KOTO** Collaboration, J. K. Ahn et al., *Search for the $K_L \to \pi^0\nu\bar{\nu}$ and $K_L \to \pi^0 X^0$ decays at the J-PARC KOTO experiment*, *Phys. Rev. Lett.* **122** (2019), no. 2 021802, [arXiv:1810.09655].

[125] Y. Amhis, M. Kenzie, M. Reboud, and A. R. Wiederhold, *Prospects for searches of $b \to s\nu\bar{\nu}$ decays at FCC-ee*, *JHEP* **01** (2024) 144, [arXiv:2309.11353].

[126] **BaBar** Collaboration, B. Aubert et al., *A Search for Invisible Decays of the Upsilon(1S)*, *Phys. Rev. Lett.* **103** (2009) 251801, [arXiv:0908.2840].

[127] L. N. Chang, O. Lebedev, and J. N. Ng, *On the invisible decays of the Upsilon and $J/\Psi$ resonances*, *Phys. Lett. B* **441** (1998) 419–424, [hep-ph/9806487].

[128] N. Fernandez, J. Kumar, I. Seong, and P. Stengel, *Complementary Constraints on Light Dark Matter from Heavy Quarkonium Decays*, *Phys. Rev. D* **90** (2014), no. 1 015029, [arXiv:1404.6599].

[129] E. Bertuzzo, C. J. Caniu Barros, and G. Grilli di Cortona, *MeV Dark Matter: Model Independent Bounds*, *JHEP* **09** (2017) 116, [arXiv:1707.00725].

[130] **NA62** Collaboration, E. Cortina Gil et al., *The Beam and detector of the NA62 experiment at CERN*, *JINST* **12** (2017), no. 05 P05025, [arXiv:1703.08501].

[131] **KOTO** Collaboration, T. Yamanaka, *The J-PARC KOTO experiment*, *PTEP* **2012** (2012) 02B006.

[132] K. M. Nollett and G. Steigman, *BBN And The CMB Constrain Light, Electromagnetically Coupled WIMPs*, *Phys. Rev. D* **89** (2014), no. 8 083508, [arXiv:1312.5725].

[133] K. Cheung, P.-Y. Tseng, Y.-L. S. Tsai, and T.-C. Yuan, *Global Constraints on Effective Dark Matter Interactions: Relic Density, Direct Detection, Indirect Detection, and Collider*, *JCAP* **05** (2012) 001, [arXiv:1201.3402].

[134] C. Balázs, T. Li, and J. L. Newstead, *Thermal dark matter implies new physics not far above the weak scale*, *JHEP* **08** (2014) 061, [arXiv:1403.5829].

[135] S. Liem, G. Bertone, F. Calore, R. Ruiz de Austri, T. M. P. Tait, R. Trotta, and C. Weniger, *Effective field theory of dark matter: a global analysis*, *JHEP* **09** (2016) 077, [arXiv:1603.05994].

[136] **GAMBIT** Collaboration, P. Athron et al., *Thermal WIMPs and the scale of new physics: global fits of Dirac dark matter effective field theories*, *Eur. Phys. J. C* **81** (2021), no. 11 992, [arXiv:2106.02056].

[137] T. Lin, E. W. Kolb, and L.-T. Wang, *Probing dark matter couplings to top and bottom quarks at the LHC*, *Phys. Rev. D* **88** (2013), no. 6 063510, [arXiv:1303.6638].

[138] **Flavour Lattice Averaging Group (FLAG)** Collaboration, Y. Aoki et al., *FLAG Review 2021*, *Eur. Phys. J. C* **82** (2022), no. 10 869, [arXiv:2111.09849]. Updated results and plots available at: http://flag.unibe.ch/.

[139] D. Aloni, A. Efrati, Y. Grossman, and Y. Nir, $\Upsilon$ *and* $\psi$ *leptonic decays as probes of solutions to the* $R_D^{(*)}$ *puzzle*, *JHEP* **06** (2017) 019, [arXiv:1702.07356].

[140] D. Hatton, C. T. H. Davies, J. Koponen, G. P. Lepage, and A. T. Lytle, *Bottomonium precision tests from full lattice QCD: Hyperfine splitting,* $\Upsilon$ *leptonic width, and b quark contribution to* $e^+e^- \to$ *hadrons*, *Phys. Rev. D* **103** (2021), no. 5 054512, [arXiv:2101.08103].

[141] **HPQCD** Collaboration, D. Hatton, C. T. H. Davies, B. Galloway, J. Koponen, G. P. Lepage, and A. T. Lytle, *Charmonium properties from lattice* $QCD + QED$*: Hyperfine splitting,* $J/\psi$ *leptonic width, charm quark mass, and* $a_\mu^c$, *Phys. Rev. D* **102** (2020), no. 5 054511, [arXiv:2005.01845].

[142] M. König and M. Neubert, *Exclusive Radiative Higgs Decays as Probes of Light-Quark Yukawa Couplings*, *JHEP* **08** (2015) 012, [arXiv:1505.03870].

[143] M. Beneke and T. Feldmann, *Symmetry breaking corrections to heavy to light B meson form-factors at large recoil*, *Nucl. Phys. B* **592** (2001) 3–34, [hep-ph/0008255].

[144] P. Ball and R. Zwicky, *New results on* $B \to \pi, K, \eta$ *decay formfactors from light-cone sum rules*, *Phys. Rev. D* **71** (2005) 014015, [hep-ph/0406232].

[145] P. Ball and R. Zwicky, $B_{d,s} \to \rho, \omega, K^*, \phi$ *decay form-factors from light-cone sum rules revisited*, *Phys. Rev. D* **71** (2005) 014029, [hep-ph/0412079].

[146] A. Khodjamirian, *Hadron Form Factors: From Basic Phenomenology to QCD Sum Rules*. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 2020.

[147] V. Bernard, M. Oertel, E. Passemar, and J. Stern, $K^L_{\mu 3}$ *decay: A Stringent test of right-handed quark currents*, *Phys. Lett. B* **638** (2006) 480–486, [hep-ph/0603202].

[148] V. Bernard, M. Oertel, E. Passemar, and J. Stern, *Dispersive representation and shape of the K(l3) form factors: Robustness*, *Phys. Rev. D* **80** (2009) 034034, [arXiv:0903.1654].

[149] I. Baum, V. Lubicz, G. Martinelli, L. Orifici, and S. Simula, *Matrix elements of the electromagnetic operator between kaon and pion states*, *Phys. Rev. D* **84** (2011) 074503, [arXiv:1108.1021].

arXiv:2402.19193v1 [hep-ph] 29 Feb 2024

# Magnetic catalysis and diamagnetism from pion fluctuations

Jie Mei,[1,2,3,*] Rui Wen,[1] Shijun Mao,[4] Mei Huang,[1] and Kun Xu[1,†]

[1] School of Nuclear Science and Technology, University of Chinese Academy of Sciences, Beijing, 100049, P.R. China
[2] School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, P.R. China
[3] Institute of High Energy Physics, Chinese Academy of Sciences,Beijing, 100049,P.R. China
[4] School of Physics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, P.R. China

In the framework of Nambu–Jona-Lasinio model beyond mean field approximation, the effects of pion fluctuations on (inverse) magnetic catalysis and magnetic susceptibility are studied. The negative magnetic susceptibility at low temperature is observed when contributions from both neutral and charged pions are taken into account. In weak field approximation, it is observed that at finite temperature, the magnetic inhibition effect in the chiral limit, resulting from the difference between the transverse and longitudinal velocities of neutral pions, converts to weak magnetic catalysis when considering a non-zero current quark mass. Moreover, the magnetic catalysis is amplified by the charged pions.

## I. INTRODUCTION

The investigation of the response of quark matter to a uniform magnetic field background has been a hot topic for the last decade, see reviews [1–5]. In the experimental aspect, a strong but transient magnetic field can be generated in the initial stage of heavy ion collisions (HIC), which was believed to be the strongest magnetic field ever created, with the strength of $eB \sim 10^{18-20}$ Gauß and life time of $10^{-22}$ second [6–10]. In the theoretical aspect, the interplay between a magnetic field and quantum chromodynamics (QCD) can lead to various novel behaviors and can be used as a probe to investigate the inner structure of quark matter.

The surge of the interest in studying the influence of magnetic field to a equilibrium system starts from two abnormal results from a lattice QCD group's ab-initio calculation, (i) although the vacuum quark mass is enhanced by the magnetic field, the chiral critical temperature decrease with the increasing of it, which is called inverse magnetic catalysis (IMC), (ii) the magnetic susceptibility is negative at low temperature while positive at high temperature [11–16]. These results disagree with most of the effective model predictions at that time, for instance, the standard Nambu-Jona-Lasinio (NJL) model and linear-$\sigma$ model with quark (Quark-Meson model) under mean-field approximation[17, 18].

Numerous studies have been conducted to elucidate the IMC and diamagnetic effects. These investigations have explored various mechanisms, including magnetic inhibition resulting from fluctuations of neutral pions [19], chirality imbalance stemming from sphaleron transitions or instanton-anti-instanton pairings [20, 21], and the influence of the running coupling constant in the presence of a magnetic field [22]. Some groups tried to include the anomalous magnetic moment effect [23–30] or the effect of tensor channel [26] in the NJL model to reproduce the IMC or diamagnetism. By considering the running coupling with $eB$-dependence fitted by lattice QCD data, the IMC result can be successfully reproduced in [31–38]. Hadron resonance gas (HRG) model, where the hadrons are assumed as point-like particles with no interaction in between, can also reproduce diamagnetic result at low temperature region [39, 40]. Functional continuum field approaches, such as the functional renormalization group (FRG), Dyson-Schwinger equations (DSE), and holographic QCD models also have made great efforts on both the effective models [39, 41–45] and QCD theory [46–48]. In general, the understanding of the IMC and diamagnetic effects remains an open question.

Most of the work in the NJL model was done based on the mean-field approximation which only considers the lowest-order in $1/N_c$ expansion [49]. In next-to-leading order of this expansion, the feedback effect from mesons is taken into consideration [50, 51]. Generally, mean field approximation for quark together with random phase approximation for meson works well to describe the thermodynamic properties of QCD matter in absence of magnetic field. For the puzzle of IMC and diamagnetism, the feedback from mesons can be part of the solution, given that they are influenced by magnetic field in both direct (for charged mesons) and indirect (for neutral mesons) manners [38, 52–58]. In references [19, 59, 60], the feedback effect from $\pi_0$ with a physical propagating velocity is included in the chiral limit, giving an IMC result. Besides, in the (Polyakov-loop extended) quark-meson model, the meson fluctuations, especially the light pion contributions, lead to the diamagnetism [39, 42]. In our calculation, we investigated beyond mean field by including the effect from both $\pi_0$ and $\pi^{\pm}$ with and without finite propagating velocities, and consider a physical situation where chiral symmetry is explicitly broken, to see their role in (inverse) magnetic catalysis and magnetic susceptibility.

This paper is arranged as follows: In Sec. II we introduce the calculating procedure in beyond mean-field NJL model in the manner of weak-field expansion. In Sec. III, we give the numerical results for $eB$-dependence of critical temperature $T_c$ and magnetic susceptibility followed

* meijie22@mails.ucas.ac.cn
† xukun21@ucas.ac.cn

by a summary and discussion in Sec. IV.

## II. THE NJL MODEL

We start with the Lagrangian of SU(2) NJL model in the presence of a uniform magnetic field,

$$\mathcal{L} = \bar{\psi}\left(i\gamma^{\mu}D_{\mu} - \hat{m}\right)\psi + G\left[\left(\bar{\psi}\psi\right)^2 + \left(\bar{\psi}i\gamma_5\vec{\tau}\psi\right)^2\right]. \quad (1)$$

Here, $\psi$ is two-flavor quark field $\psi = (u, d)^T$, $\hat{m}$ is the quark current mass matrix $\hat{m} = \mathrm{diag}(m_u, m_d)$ with $m_u = m_d = m_0$, which explicitly breaks the chiral symmetry, and $\tau_i$ is $i$-th component of Pauli matrices. The covariant derivative $D_\mu = \partial_\mu - iQA_\mu$ coupling quarks with electric charge $Q = \mathrm{diag}(q_u, q_d) = \mathrm{diag}(2/3e, -1/3e)$ to a gauge field $\mathbf{B} = \nabla \times \mathbf{A}$. In this article, we choose the Landau gauge, where the potential $A_\mu = (0, 0, Bx_1, 0)$.

In Schwinger scheme, the translational invariance for a charged particle is broken by the magnetic field, deviding the quark propagator into two parts, namely, the Schwinger phase part and the Fourior transformation of translational invariant part.

$$S_f(x, y) = e^{i\Phi_f(x_\perp, y_\perp)}\int\frac{d^4k}{(2\pi)^4}e^{-ik(x-y)}\tilde{S}_f\left(k_\perp, k_\parallel\right). \quad (2)$$

Here, the parallel and perpendicular components of the four-dimensional coordinate and momentum are written as $x_\perp = (x_1, x_2)$, $k_\perp = (k_1, k_2)$ and $k_\parallel = (k_0, k_3)$. The quark Schwinger phase in Landau gauge reads

$$\Phi_f(x_\perp, y_\perp) = \frac{q_f B}{2}(x_1 + y_1)(x_2 - y_2), \quad (3)$$

which is gauge-dependent. The detailed forms of Schwinger phase in different gauges can be found in [4, 61].

In the weak magnetic field limit, the quark propagator in momentum space can be expanded in the power of $(q_f B)$, and reads

$$\begin{aligned} &i\tilde{S}_f(k_\perp, k_\parallel) \\ &= i\frac{m_q + \not{k}}{k^2 - m_q^2} - (q_f B)\frac{\gamma_1\gamma_2\left(m_q + \not{k}_\parallel\right)}{(k^2 - m_q^2)^2} \\ &\quad - 2i\left(q_f B\right)^2\frac{k_\perp^2(m_q + \not{k}_\parallel) + \not{k}_\perp(m_q^2 - k_\parallel^2)}{(k^2 - m_q^2)^4} + \mathcal{O}(q_f B)^3 \\ &= i\tilde{S}^{(0)} + (q_f B)i\tilde{S}^{(1)} + (q_f B)^2 i\tilde{S}^{(2)} + \mathcal{O}(q_f B)^3 \end{aligned} \quad (4)$$

where the contributions of transverse and parallel momentum are considered separately. Here, $m_q$ is the constituent quark mass, which is the same value for both $u$ and $d$ flavor. We use the notation,

$$\begin{aligned} (a\cdot b)_\parallel &= a^0b^0 - a^3b^3, \\ (a\cdot b)_\perp &= a^1b^1 + a^2b^2, \end{aligned} \quad (5)$$

hence $k^2 = k_\parallel^2 - k_\perp^2$. In the weak field approximation, the effect of Schwinger phase is weak enough so that we can neglect its contribution and assume the system is translational invariant.

### A. Mean field approximation

In mean field approximation, the thermodynamic potential of the system at finite temperature and under external magnetic field $B$ takes the form of:

$$\Omega_{\mathrm{MF}}(T, B) = \frac{(m_q - m_0)^2}{4G} + \Omega_q(T, B), \quad (6)$$

with the contribution from quarks:

$$\Omega_q(T, B) = \mathrm{Tr}_{\{c,f,s,x\}}\ln\left(\frac{1}{T}S^{-1}(x, x)\right). \quad (7)$$

The trace operation is carried out over color (c), flavor (f), spinor (s) degrees of freedom, as well as over the four-dimensional coordinate ($x$). $S^{-1}(x, x)$ is the inverse of quark propagator in coordinate space.

To determine the effective quark mass, we have to find the ground state by locating the global minimum of the thermodynamic potential

$$\begin{aligned} \frac{\partial\Omega_{\mathrm{MF}}(T, B)}{\partial m_q} &= 0, \\ \frac{\partial^2\Omega_{\mathrm{MF}}(T, B)}{\partial m_q^2} &\geq 0, \end{aligned} \quad (8)$$

which leads to the gap equation

$$m_0 = m_q - 2G\mathrm{Tr}_{\{c,f,s,k\}}[i\tilde{S}(k)]. \quad (9)$$

Inserting Eq. (2) into Eq. (9) and apply the summation over Matsubara frequency, we can obtain the weak-field expansion version of gap equation:

$$m_q\left(1 - 2GN_c\sum_f\left(I_1^{(0)} + (q_f B)^2 I_1^{(2)} + \mathcal{O}(q_f B)^4\right)\right) = m_0, \quad (10)$$

with

$$I_1^{(0)} = -4\int\frac{d^3k}{(2\pi)^3}\mathcal{F}_{(1)}(m_q^2), \quad (11)$$

$$I_1^{(2)} = 8\int\frac{d^3k}{(2\pi)^3}k_\perp^2\mathcal{F}_{(4)}(m_q^2). \quad (12)$$

There is no contribution from odd terms of $I_1^{(n)}$ after performing the trace in spinor space, which agrees with symmetry analysis. The contributions from higher orders of magnetic field are neglected. Here we define the fermionic threshold functions

$$\mathcal{F}_{(n)}(m_q^2) \equiv T\sum_{n_T}\frac{1}{(k^2 - m_q^2)^n}, \quad (13)$$

then it is straightforward to obtain:

$$\mathcal{F}_{(1)} = \frac{1}{2E_q}(1-2n_f(m_q^2)), \tag{14}$$

$$\mathcal{F}_{(n+1)} = \frac{1}{n}\frac{\partial}{\partial m_q^2}\mathcal{F}_{(n)}. \tag{15}$$

Here, $n_f(m_q^2)$ is fermionic distribution functions and quark energy dispersion relation $E_q = \sqrt{\mathbf{k}^2+m_q^2} = \sqrt{k_3^2+k_\perp^2+m_q^2}$.

Correspondingly, the quark part of thermodynamic potential can also be rewritten in this weak-field expansion manner:

$$\Omega_q = \Omega_q^{(0)} + \sum_f (q_f B)^2 \Omega_q^{(2)} + \mathcal{O}(q_f B)^4, \tag{16}$$

$$\Omega_q^{(0)} = -2N_c N_f \int \frac{d^3k}{(2\pi)^3}\left[E_q + 2T\ln\left(1+e^{\frac{-E_q}{T}}\right)\right], \tag{17}$$

$$\Omega_q^{(2)} = -\frac{4N_c}{3}\int \frac{d^3k}{(2\pi)^3} k_\perp^2 \mathcal{F}_{(3)}(m_q^2). \tag{18}$$

### B. Meson section

In the NJL model, mesons are treated as quantum fluctuations above the mean field. Through the random phase approximation (RPA) method, the meson propagator can be expressed in terms of the irreducible polarization function or quark bubble,

$$\Pi_M(q^2) = i\int \frac{d^4k}{(2\pi)^4}\mathrm{Tr}_{\{c,f,s\}}\left[\Gamma_M^* \tilde{S}(k)\Gamma_M \tilde{S}(k-q)\right], \tag{19}$$

with the meson vertex

$$\Gamma_M = \begin{cases} 1, & M=\sigma \\ i\gamma_5\tau_+, & M=\pi_+ \\ i\gamma_5\tau_-, & M=\pi_- \\ i\gamma_5\tau_3, & M=\pi_0, \end{cases}, \tag{20}$$

where $\tau_\pm$ is determined by $\tau_\pm = (\tau_1 \pm i\tau_2)/\sqrt{2}$. By inserting Eq. (2) into Eq. (19), we can get the polarization function for $\pi_0$ in weak field approximation

$$\Pi_{\pi_0}(q^2) = N_c \sum_f \left[\Pi^{00} + (q_f B)^2\left(2\Pi^{20}+\Pi^{11}\right)\right], \tag{21}$$

with definition

$$\Pi^{lm}(q^2) = i\int \frac{d^4k}{(2\pi)^4}\mathrm{Tr}_{\{s\}}\left[i\gamma_5\tilde{S}^{(l)}(k)i\gamma_5\tilde{S}^{(m)}(k-q)\right], \tag{22}$$

After performing the trace in spinor space, the contribution of $\Pi^{10}$ and $\Pi^{01}$ vanishes. The calculation for Eq. (22) is straightforward but tedious. For the polarization function for charged pions, it has the similar form as the neutral pion case,

$$\Pi_{\pi_\pm}(q^2) = 2N_c\left[\Pi^{00} + (q_u B)(q_d B)\Pi^{11} + \sum_f (q_f B)^2 \Pi^{20}\right]. \tag{23}$$

Again, the contribution of $\Pi^{01}$ vanishes after carrying out the trace in spinor space. In weak field approximation, we consider the influence of magnetic field is small enough so that we can ignore the impact of Schwinger phase and expect homogeneity of the system.

In RPA method, the effective meson propagator can be constructed by

$$U_M = \frac{2G}{1-2G\Pi_M}, \tag{24}$$

and the pole mass $m_{pole}$ (static solution, setting $q_1 = q_2 = q_3 = 0$) and the screening masses $m_{scr,i}$ in $q^i$ direction (setting $q_0 = 0$, and $q_j = 0$ for $j \neq i$) can be solved by following equations, correspondingly,

$$1 - 2G\Pi_M(q_0^2 = m_{pole}^2, 0) = 0 \tag{25}$$

and

$$1 - 2G\Pi_M(0, q_i^2 = -m_{scr,i}^2) = 0. \tag{26}$$

The detailed calculation for the polarization function in weak-field expansion are listed in appendix A. It should be noted that when meson pole (screening) mass exceeds the threshold of the mass sum of its constituent quarks, the meson undergoes a Mott transition and a finite width ($m_M \to m_M - i\frac{\Gamma}{2}$) should be taken into consideration. The meson mass and its width can be determined by the corresponding complex equations. In this paper, we work in the temperature region where pions are still bound state particles.

### C. Beyond mean-field approximation

We go beyond the mean-field approximation by using the $1/N_c$ expansion [50, 51]. Another self-consistent beyond mean-field method in the NJL model is the FRG approach [62, 63]. By including the next-to-leading order of $1/N_c$ expansion, meson degrees of freedom are self-consistently introduced. The thermodynamic potential of the quark-meson plasma [51] can be rewritten as

$$\Omega = \frac{(m_q - m_0)^2}{4G} + \Omega_q + \sum_M \Omega_M. \tag{27}$$

In pole approximation, the meson contribution in thermodynamic potential with vanishing magnetic field takes the form,

$$\Omega_{\pi_0} = \int \frac{d^3k}{(2\pi)^3}\left[\frac{E_{\pi_0}}{2} + T\ln\left(1-e^{-E_{\pi_0}/T}\right)\right], \tag{28}$$

with meson energy dispersion relation

$$E_{\pi_0} = \sqrt{m^2_{\pi_0,pole} + (\mathbf{v}\cdot\mathbf{k})^2}. \tag{29}$$

As shown in [64–68], $\mathbf{v}$ in Eq. (29) represents the propagating velocity for the corresponding meson, whose $i$-th component has the form of [65–68]

$$v_i = \frac{m_{M,pole}}{m_{M,scr,i}}. \tag{30}$$

For neutral pion, when we consider its behaviors in the hadron level, it has no interaction with magnetic field at all. However, in a model in quark level like NJL model, it's magnetic field-sensitive constituent quark will contribute to the variation of $\pi_0$ mass and the splitting of its propagating velocity in parallel and perpendicular directions, leading to the new form of energy dispersion relation

$$E_{\pi_0} = \sqrt{m^2_{pole,\pi_0} + v_\perp^2 k_\perp^2 + v_\parallel^2 k_3^2}, \tag{31}$$

which is believed to be responsible for the IMC effect in the chiral limit, as in [19, 59, 60]. In absence of a uniform magnetic field, the neutral pion $\pi_0$ and charged pions $\pi_\pm$ are isospin triplet that share the same form of thermodynamic potential. When a magnetic field is turned on, the charged pions obtain extra masses from magnetic field, and decouple from neutral pion, which means we should consider their effect seperately. For charged pions, the magnetic field not only alter their pole masses and propagating velocities, but also change the momentum in perpendicular directions from thermodynamic potential and energy dispersion relation into discrete Landau levels,

$$\Omega_{\pi_\pm} = \sum_{n=0}^{\infty} \frac{|eB|}{2\pi} \int_{-\infty}^{\infty} \frac{dk_3}{2\pi} \left[ \frac{E_{\pi_\pm}}{2} + T \ln\left(1 - e^{-E_{\pi_\pm}/T}\right)\right], \tag{32}$$

$$E_{\pi_\pm} = \sqrt{m^2_{pole,\pi_\pm} + v_\perp^2 (2n+1)|eB| + v_\parallel^2 k_3^2}. \tag{33}$$

Here we use the strong-field form of charged pion's energy dispersion relation, to take the cyclotron motion into consideration. Now with the complete form of thermodynamic potential beyond mean field, we can get the new constituent quark mass from the corresponding gap equation with the feedback effect from meson,

$$m_q \left( \frac{1}{4G} + \frac{\partial \Omega_q}{\partial m_q^2} + \sum_M \frac{\partial \Omega_M}{\partial m_q^2} \right) = \frac{m_0}{4G}. \tag{34}$$

Comparing the mean field quark mass $m_{\rm MF}$ from Eq. (9) with the newly obtained quark mass $m_q$ from Eq. (34), there is a mass difference which comes from the quantum fluctuations above the mean field. Following the procedure described in [51], the meson contribution in thermodynamic potential can be expanded in power of $(m_q^2 - m_{\rm MF}^2)$,

$$\Omega_M = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n \Omega_M}{(\partial m_q^2)^n}\Big|_{m_q^2 = m_{\rm MF}^2} (m_q^2 - m_{\rm MF}^2)^n. \tag{35}$$

To simplify the calculation, we only consider the first two terms of the above series, where the leading order term vanishes and only the next-to-leading order term has none-zero contribution, then we get the practical form of gap equation,

$$m_q \left( \frac{1}{4G} + \frac{\partial \Omega_q}{\partial m_q^2} + \sum_M \frac{\partial \Omega_M}{\partial m_q^2}\Big|_{m_q^2 = m_{\rm MF}^2} \right) = \frac{m_0}{4G}. \tag{36}$$

In the previous beyond-mean-field calculation with non-vanishing magnetic field, most works focus on the effect of neutral pions in the chiral limit, giving a conclusion that including a neutral pion with a physical propagating velocity may cause inverse magnetic catalysis effect. Those calculations stand because they are working in strong field limit, where charged pion gained a large mass from magnetic field and hence their contribution in the distribution function can be neglected. In this work, however, we work in weak field limit with $eB \sim m_\pi^2$. The mass disparity between neutral and charged pions is not significant enough, and the contributions from charged pions should not be overlooked. In the next section, we will investigate how the neutral pion and charged pions affect the chiral condensate and magnetic susceptibility.

## III. NUMERICAL RESULTS

In order to analyze the roles of neutral and charged pions in (inverse) magnetic catalysis and magnetic susceptibility, we consider the following four case: Case-0) Mean field approximation; Case-I) only neutral pion contribution; Case-II) only charged pions contribution; Case-III) both neutral and charged pions contribution, which are listed in TABLE I. We will also investigate the effect with and without finite pion propagating velocities separately.

Because of the contact interaction in the NJL model, the ultraviolet divergence cannot be eliminated through renormalization, and a proper regularization scheme is needed. In our work, we apply the gauge invariant Pauli-Villars regularization [59], which can guarantee the law of causality and effectively avoid the unphysical oscillation at finite magnetic field. Other schemes, like proper-time regularization scheme and magnetic field independent regularization (MFIR)[56, 57], which separates the vacuum and magnetic contribution, are also effective when dealing with system under a uniform magnetic field. By fitting the physical quantities, chiral condensate $\langle\bar{\psi}\psi\rangle = (-250\text{MeV})^3$, pion decay constant $f_\pi = 93\text{MeV}$ in vacuum, we fix the current mass of light quarks $m_0 = 5\text{MeV}$, and obtain the parameter $\Lambda = 1127\text{MeV}$. For different cases, the coupling constants are given in TABLE I.

The full calculating procedure is as follows: 1) we first calculate the mean field gap equation Eq. (9) to get the mean field quark mass $m_{\rm MF}$ before substituting it into the pole equation Eq. (25), where we can get the pole and screening masses of $\pi_0$ and $\pi_\pm$, and hence the corresponding propagating velocity in longitudinal and trans-

| Case | Included mesons | $G\Lambda^2$ |
|---|---|---|
| 0 | None | 4.37 |
| I | $\pi_0$ only | 4.81 |
| II | $\pi_\pm$ only | 5.36 |
| III | $\pi_0$, $\pi_\pm$ | 6.05 |

TABLE I. Coupling constants in different cases in and beyond mean field approximation.

verse directions. 2) Inserting the pion dispersion relation obtained above into the beyond-mean-field thermodynamic potential Eq. (27) and gap equation Eq. (34), we can finally get the numerical result of order parameter $m_q$ and magnetic susceptibility. In this work, we employ this systematic step-by-step approach to go beyond the mean field approximation, and more self-consistent ways to go beyond mean field in absence of magnetic field are given in [69, 70].

In the case of explicit chiral symmetry breaking, the pseudo Nambu-Goldstone modes $\pi$ have a finite mass, making them propagating with a velocity lower than the speed of light at finite magnetic field and finite temperature. Due to the qualitative similarity in the behaviors of pion masses and propagating velocities across different cases, in Fig. 1 to Fig. 3 we will choose the case that exhibits the most conclusive results to highlight the distinctions arising from a relatively weak magnetic field.

### A. Meson properties

Here we start by examining the characteristics of neutral and charged pion masses at finite temperature and magnetic field. Fig. 1 demonstrates the magnetic field dependence of $\pi_0$ and $\pi_\pm$ masses at zero temperature. We use the weak field expansion and only calculate to $eB = 3m_\pi^2$ with $m_\pi^2 = 0.0179$ GeV$^2$. For $\pi_0$ mass, it decreases with the increase of the magnetic field, while for the ground state energy of $\pi_\pm$, which is given by $E_{\pi_\pm} = \sqrt{m_{\pi_\pm}(eB)^2 + eB}$, it increases with the magnetic field, showing an opposite tendency. If we consider $\pi_0$ as a charge-neutral point-like particle, it should remain unaffected by external magnetic field, while in the NJL model where mesons are considered as composite particles made of quark anti-quark, $\pi_0$ mass changes with the changing of quark condensate. From Nambu-Goldstone theorem, since the explicit breaking of chiral symmetry will get more "explicit" with the restoration of chiral symmetry, the mass of the pseudo-Goldstone boson $\pi_0$ should generally decreases with magnetic field, which is consist with our numerical result. A uniform magnetic field breaks the isospin symmetry between $u$ and $d$ quarks, and $\pi_\pm$ gained mass from magnetic field, which is qualitatively consist with the point-like approximation (PLA) result, i.e. $E^{\rm PLA}_{\pi_\pm}(B) = \sqrt{m^2_{\pi_\pm}(B=0) + eB}$.

In Fig. 2 we show the temperature-dependence of pion's pole masses at $eB = 0$ and $eB = 3m_\pi^2$. At vanishing magnetic field, neutral and charged pions share the same mass. As the temperature gets higher, pion mass increases with the restoration of chiral symmetry, which is in agreement with the analysis of Nambu-Goldstone theorem. At $eB = 3m_\pi^2$, the pole masses of neutral and charged pions split. It should be noticed that at high temperature region, the system is thermalized and the splitting between the pole masses of neutral and charged pions becomes smaller.

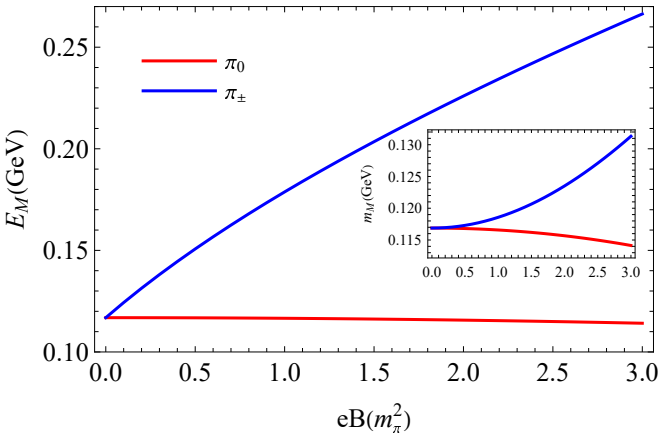



FIG. 1. Neutral (Red solid line) and charged (Blue solid line) pion masses beyond mean field approximation in Case III as a function of magnetic field, with vanishing temperature.

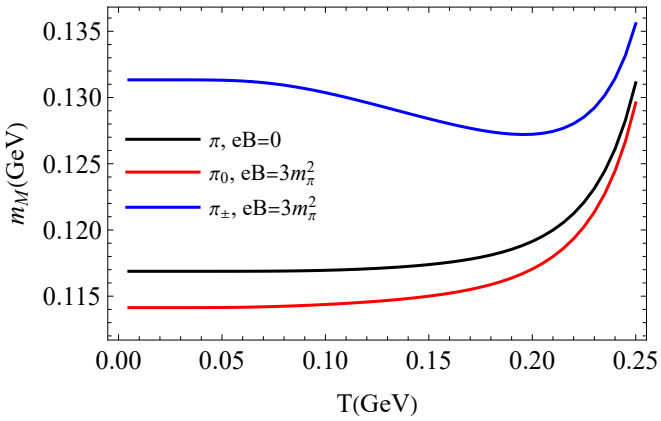



FIG. 2. Pion mass beyond mean field approximation in Case III as a function of temperature with vanishing magnetic field.

Fig. 3 shows longitudinal and transverse velocities of both $\pi_0$ and $\pi_\pm$ as a function of the temperature at $eB = 0$ and $eB = 3m_\pi^2$. At vanishing magnetic field, both $\pi_0$ and $\pi_\pm$ exhibit isotropic behavior and their velocities decrease with the increase of temperature, since the temperature leads to a breakdown of Lorentz symmetry in boost transformations. When a magnetic field is turned on, we can clearly observe the splitting between the parallel and perpendicular velocities for both $\pi_0$ and $\pi_\pm$. For perpendicular velocity, it gets lower than the vanishing magnetic field scenario, while for parallel velocity, it

is generally the same as that in the vanishing magnetic field case at low temperature region. When the temperature gets higher, the splitting between the longitudinal and transverse velocities induced by magnetic field gets melted and the system tends to become isotropic again. This thermalization phenomenon is also observed in the strong field case, e.g. [68]. When the temperature increases, the velocity of neutral pion at $eB = 3m_\pi^2$ gets larger than that at $eB = 0$, since the neutral pion becomes lighter at higher magnetic field. The behavior of $\pi_\pm$ is basically the same as the neutral pion case, except that with the increase of magnetic field strength, the propagating velocity of $\pi_\pm$ tends to be smaller, given that the charged pion mass is increased by the magnetic field.

It should be noted that in the chiral limit, the behavior of propagating velocities appears to exhibit the opposite trend, where the longitudinal velocity is always exactly the same as the speed of light, the transverse velocity gets lower as temperature increases. This discrepancy between transverse and longitudinal velocities was believed to lead to magnetic inhibition, which gives the IMC result in the chiral limit [19, 59, 60].

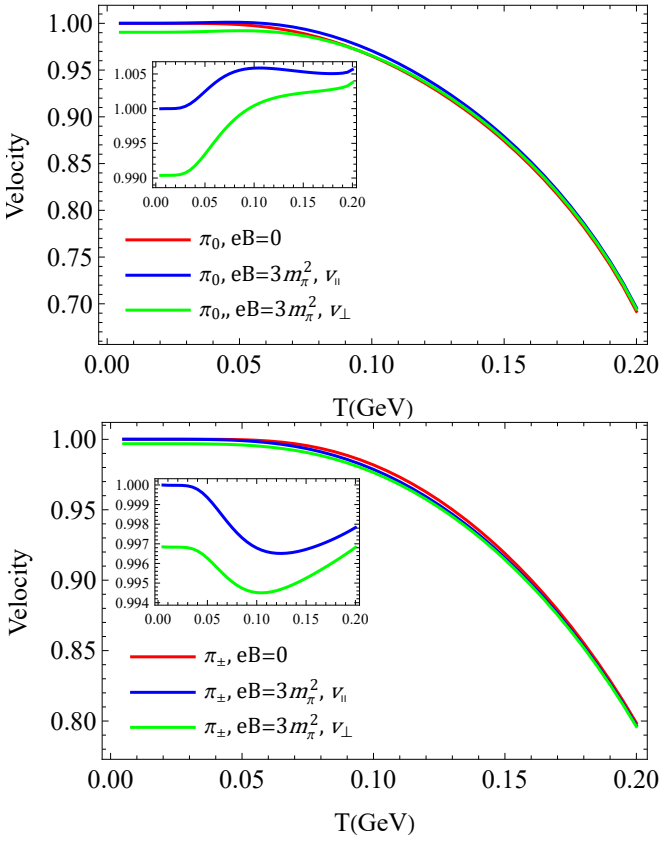



FIG. 3. Propagating velocity of $\pi_0$ (upper panel) and $\pi_\pm$ (lower pane) beyond mean field in Case I and Case II as a function of temperature at different magnetic field in the chiral symmetry breaking phase. The inlaid sub-figure shows the ratio of velocities at $eB = 3m_\pi^2$ and $eB = 0$ in corresponding directions.

### B. Chiral condensate beyond mean field

Now with the pole mass, screening masses and propagating velocities of $\pi_0$ and $\pi_\pm$ obtained in the mean-field approximation, we can insert them into the beyond-mean field thermodynamic potential and gap equation to get the quark mass in cases I,II and III. Before that, we can first calculate the effective coupling $\tilde{G}$ with

$$\frac{1}{4\tilde{G}} \equiv \frac{1}{4G} + \sum_M \frac{\partial \Omega_M}{\partial m_q^2}\Big|_{m_q^2 = m_{\rm MF}^2} \tag{37}$$

to investigate the feedback effect from different pion to the system. To demonstrate the influence of magnetic field on effective coupling in the whole temperature region, in Fig. 4 and Fig. 5 we show $\tilde{G}(eB = 0)$ and $\tilde{G}(eB = 3m_\pi^2)$ with the feedback from different mesons, with and without the consideration of the effect of finite pion propagating velocities.

Fig. 4 shows the temperature dependence of effective coupling at different magnetic field strength with velocities $v = 1$. We start with Case-I in the upper panel. The effective coupling generally increases until getting close to its Mott transition point, where a plunge is oberved. As shown in Eq. (37), the effective coupling is determined by the contribution from mesons. The mesonic energy dispersion relation is controlled by two different factors, i.e. the pole mass and propagating velocity. In Fig. 4, velocities are set to be the speed of light, and as demonstrated in Fig. 2, the pole mass of neutral pion is monotonically increasing with temperature, which leads to the increasing of mesonic energy and the decreasing of its contribution in the effective coupling. For the same reason, the effective coupling at higher magnetic field background always gets lower than that at vanishing magnetic field in Case I. For Case II and Case III in the middle and lower panels, Since we are examining the impact of charged pions, the effective coupling in these scenarios actually increases with the magnetic field. This is due to the fact that charged pions exhibit an opposing and more pronounced reaction to the magnetic field. This result suggests that neglecting the finite velocity effect may weaken magnetic catalysis in Case I, but strengthen it in Cases II and III.

Fig. 5 shows the temperature dependence of effective coupling at different magnetic field strength with finite velocities. Different from the result in Fig. 4, the effective coupling in Fig. 5 monotonically decreases with the increase of temperature. This is because the drop in propagating velocities leads to a decrease in mesonic energy, which in turn increases the overall contribution from pions. We can also notice that, although the response of effective coupling to the magnetic field is qualitatively the same as in Fig. 4, the magnitude is significantly weakened. In Case I, for example, the effect from magnetic field becomes indistinguishable at around temperature $T \sim 0.15$GeV. One can also find explanation from Fig. 2 and Fig. 3, where the changes from propa-

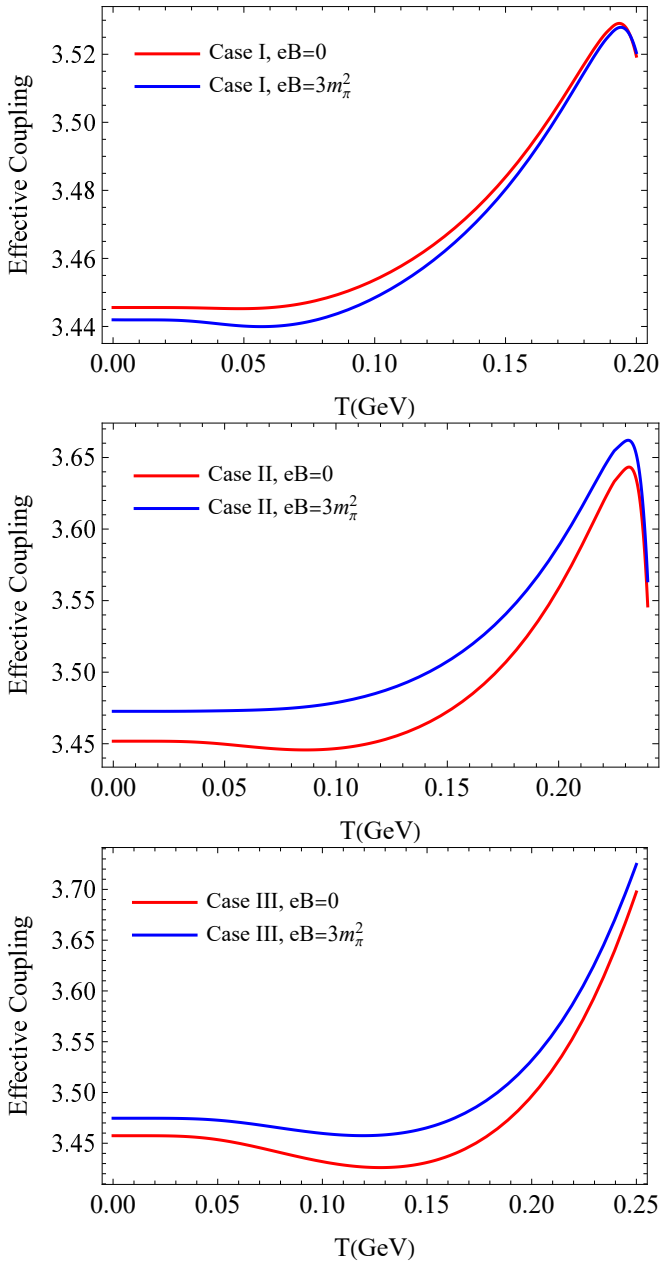



FIG. 4. Effective coupling $\tilde{G}$ at $eB = 0$ and $eB = 3m_\pi^2$ as a function of temperature in three different cases, considering constant pion propagating velocities.

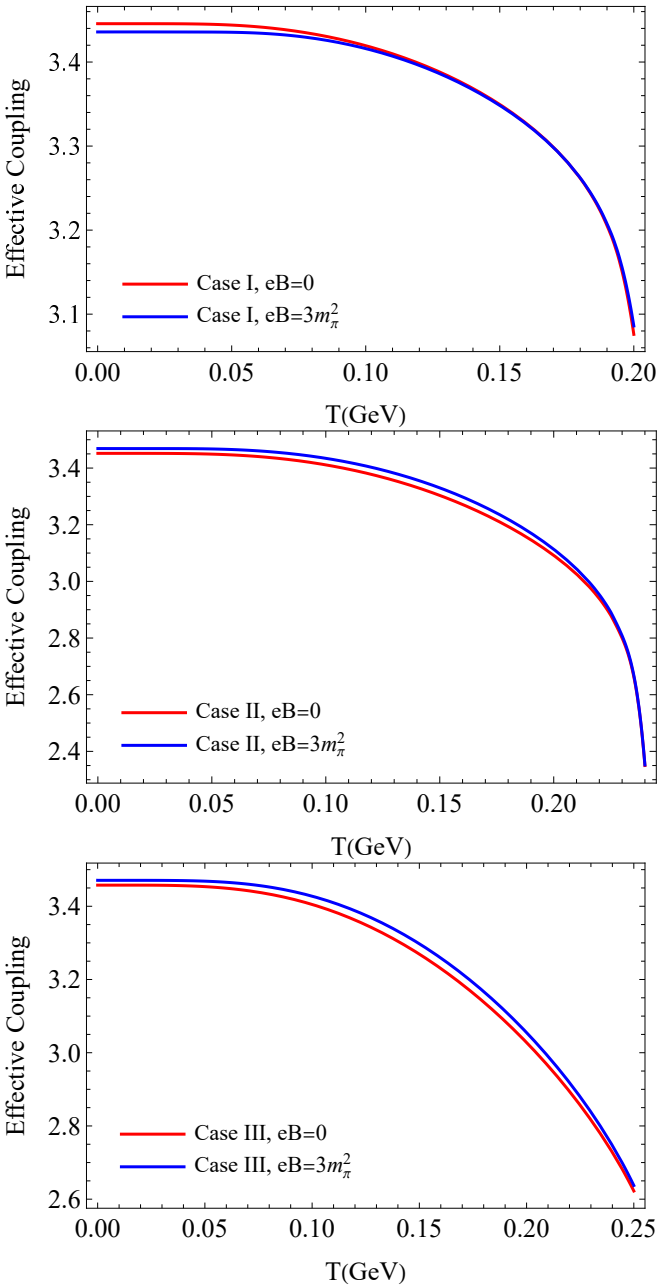



FIG. 5. Effective coupling $\tilde{G}$ at $eB = 0$ and $eB = 3m_\pi^2$ as a function of temperature in three different cases, considering pion propagating velocities dependent on temperature and magnetic field.

gating velocities counteracts the the effect of pole masses in energy dispersion relation.

In Fig. 6, we show the quark mass as a function of the magnetic field (up panel) and the temperature with pion propagating velocity $v = 1$ (middel panel) and finite $v$(below panel) for four cases listed in TABLE I. It is noticed from the upper panel that although the constituent quark mass increases with magnetic field in all cases, the increasing rate is lower in Case I with only neutral pion contribution compared to the mean-field case. The increasing rate of quark mass is significantly enhanced in both Case II and Case III in the case of considering the charged pion contribution. It is observed that in the middle figure where we consider the pionic velocity to be exactly the speed of light, the chiral symmetry restoration phase transition is catalysed in Case I and inverse-catalysed in Case III, while in Case II the behavior of quark condensate is similar to that of mean field case. However, in the lower panel, when the propagating velocities are dependent on the temperature and magnetic field strength, in all the beyond-mean-field cases, the chi-

ral phase transition is inverse-catalysed, and the more pions are considered, the more inverse-catalytic the system becomes. By observing the results for different cases in Fig. 6, it is noticed that the behaviors of these corrected quark masses agrees with the previous discussion in Fig. 4 and Fig. 5.

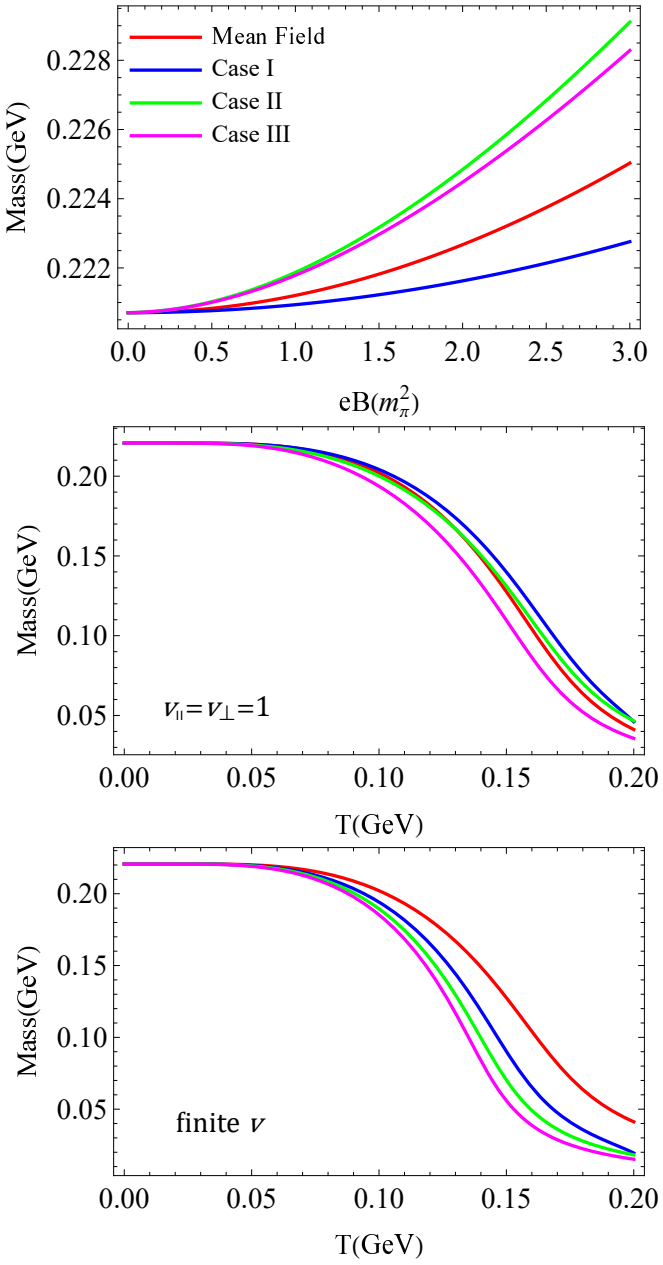



FIG. 6. Effctive quark mass $m_q$ in four different cases, as a function of magnetic field with zero temperature (upper panel) and of temperature in vanishing magnetic field with finite and constant pion propagating velocities (middle and lower panel), with $m_\pi^2 \simeq 0.0179$ GeV$^2$.

To identify the magnetic field effect on the effective quark mass at high temperature region, in Fig. 7 we show the re-scaled $T_c - eB$ phase diagram in case-I,II, and III with and without the consideration of finite pionic velocities, compared with the mean field result. In the upper panel with constant pion propagating velocities, the critical temperature in different cases is as follows: Case 0 (mean field): $T_c(0) = 157.2$MeV, Case I: $T_c(0) = 163.5$MeV, Case II: $T_c(0) = 159.8$MeV, Case III: $T_c(0) = 151.6$MeV; and in the lower panel with finite pion propagating velocities Case 0: $T_c(0) = 157.2$MeV, Case I: $T_c(0) = 145.3$MeV, Case II: $T_c(0) = 139.5$MeV, Case III: $T_c(0) = 135.6$MeV. The critical temperature $T_c$ is determined by the fastest drop of effective quark mass $\partial^2 m_q(T_c, eB)/\partial T^2 = 0$. It is noticed that in Case I where we only include the feedback effect from $\pi_0$, the increasing rate of critical temperature is weakened but the system is still magnetic catalytic. However, if we include the charged pions contribution, the increasing rate of critical temperature get enhanced. Once we consider the finite velocity effect, both the enhancement from charged pions and the weakening from neutral pion to magnetic catalysis are suppressed, just as dicussed in Fig. 5.

In the previous study of beyond-mean field calculation [19, 59, 60], IMC is achieved when a neutral pion with finite velocity is included into the system, due to the magnetic inhibition effect where the splitting between the longitudinal and transverse velocities get larger as the magnetic field and temperature are increasing. However, our calculation in this work gives an opposite result: Whether or not the influence of finite velocities is taken into account, the system remains magnetic catalytic. The main difference between the previous study and this work is that, the previous study works in the chiral limit and in this work the chiral symmetry is explicitly broken. For transverse velocity, both previous and our present result share the same form, but for longitudinal velocity, the chiral limit version is set to be the speed of light since it is a massless particle, unlike our casese, where the splitting in velocities induced by magnetic field will be thermalized once the temperature goes near the critical temperature.

### C. Magnetic susceptibility

Recently the lattice QCD calculation showed that the magnetized QCD matter exhibits diamagnetism at low temperature and paramagnetism at high temperature, so in this work we also calculate the magnetic susceptibility in the NJL model beyond mean field.

The magnetic susceptibility is defined by

$$\bar{\chi}(T) = -\frac{\partial^2 \Omega(T, eB)}{\partial eB^2}\Big|_{eB=0}. \tag{38}$$

In lattice calculation, the renormalization scale choice fixes $\bar{\chi}(0) = 0$ so that the divergence in vacuum magnetic susceptibility can be eliminated. Mimicking their procedure, we define a new magnetic susceptibility to ensure the relation $\chi(0) = 0$, where

$$\chi(T) = \bar{\chi}(T) - \bar{\chi}(0) = \frac{\partial^2 \Delta(T, eB)}{\partial eB^2}\Big|_{eB=0}, \tag{39}$$

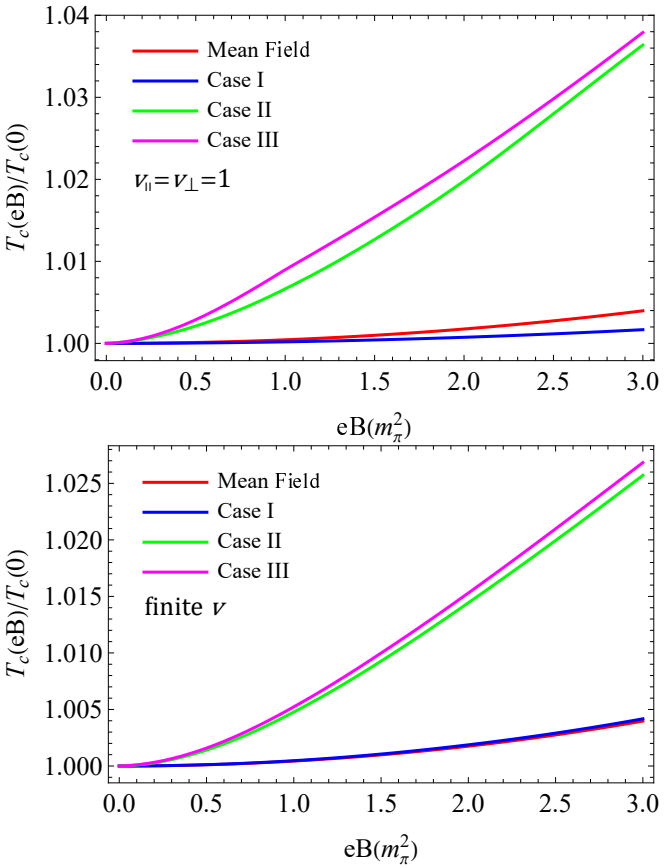



FIG. 7. Critical temperature for chiral symmetry restoration phase transition in mean field approximation and three different cases beyond mean field as a function of magnetic field, normalized by vanishing magnetic field critical temperature, considering both constant pion propagating velocities (upper panel) and finite pion propagating velocities (lower panel).

with

$$\Delta(T, eB) = \Omega(0, eB) - \Omega(T, eB). \tag{40}$$

We numerically calculate the magnetic susceptibility in the vicinity of $B = 0$, and show it in Fig. 8. In mean field approximation, the magnetic susceptibility is always positive and increases with temperature. In the Case I where $\pi_0$ is included, there is no qualitative change in the behavior of magnetic susceptibility, but instead quantitatively it is enhanced by the feedback effect of $\pi_0$. However, if we include the effect of charged pions as well like in Case II and Case III, the magnetic susceptibility is negative at low temperature and becomes positive at high temperature region, showing qualitative agreement with the lattice result.

In the mean field approximation, the system only consists of quark, a charged particle with non-zero spin, which means its magnetic susceptibility is subject to two different mechanisms, namely Pauli paramagnetism and Landau diamagnetism [71, 72]. The former originate from the spin distribution in the presence of magnetic field, and the latter is a quantum effect that is related to the cyclotron motion of charged particles. Normally the Pauli paramatnetism is stronger than Landau diamagnetism, making the total magnetic susceptibility positive in the mean field approximation. However, the pseudoscalar particles $\pi_\pm$ are spin-0 and only subject to Landau diamagnetism, giving a negative magnetic susceptibility result. When we go beyond mean field approximation, we actually consider a system composed of quarks and mesons. At low temperature, the pion meson contribution is dominant and the total magnetic susceptibility is negative, while at high temperature, the dominance of quark contribution makes the system going back to paramagnetism.

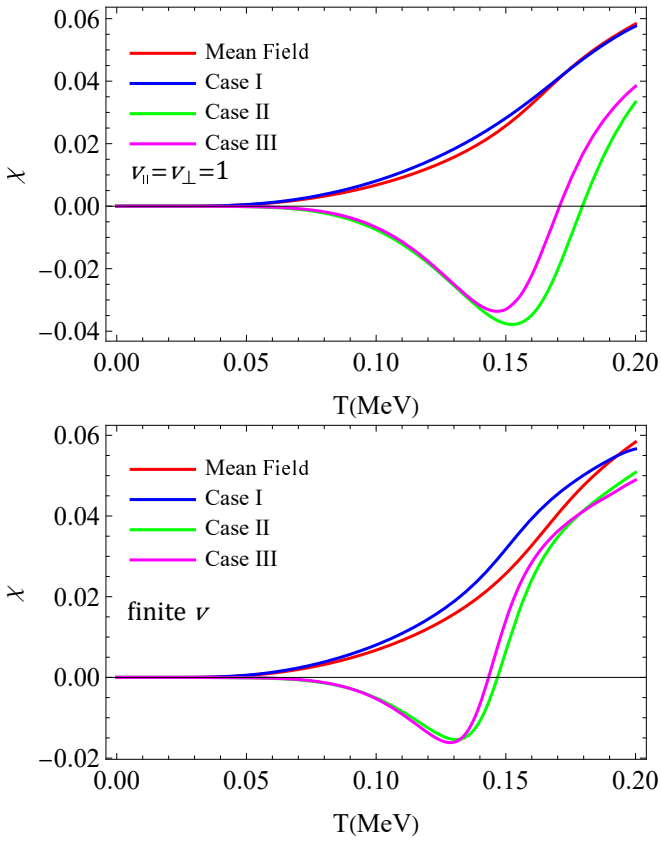



FIG. 8. Magnetic susceptibility in mean field approximation (Red solid lines), beyond mean field with the feedback from $\pi_0$ (Blue solid lines), $\pi_\pm$ (Green solid lines) and $\pi_0$, $\pi_\pm$ (Cyan solid lines) as a function of temperature with vanishing magnetic field. Propagating velocity for pions are set to be speed of light in upper panel, and finite velocities effect is considered in lower panel.

## IV. CONCLUSION

In this work, we investigated the magnetic field effect on the phase transition of chiral symmetry restoration in the framework of $SU(2)$ NJL model beyond mean field

approximation. We consider three cases, by separately including the effect of $\pi_0$ (Case I), $\pi_\pm$ (Case II), and all three pions altogether (Case III). Since sigma meson is much heavier than pions and it is not directly influenced by the magnetic field, we neglect its contribution in the calculation. To identify the effect of finite pionic propagating velocities, we consider two situations, one setting pion velocity to be exactly speed of light, and another with finite velocities determined by Eq. (30). Unlike the previous IMC results in the chiral limit due to the magnetic inhibition effect where the splitting between the longitudinal and transverse velocities of pion suppressed the critical temperature, in this work in chiral symmetry explicitly broken case, the system is always magnetic catalytic. In the discussion of effective coupling, we notice that in absence of magnetic field, both neutral and charged pions share the same contribution to the system. However, when a magnetic field is introduced, due to the opposite response of charged and neutral pion masses to it, the feedback from $\pi_0$ weakens the MC but that from $\pi_\pm$ instead enhances it. Once we consider the effect from finite velocities, it will counter the effect from the changes in pionic pole mass. When we go to finite temperature region, all the effect above will be thermalized, resulting that IMC cannot be observed in any of the above scenarios.

We also calculate the magnetic susceptibility in these three scenarios. In mean field approximation and only including $\pi_0$ scenario, the magnetic susceptibility is always positive and increases with temperature. However, if we include the effect of charged pions, the magnetic susceptibility becomes negative at low temperature and positive at high temperature region, showing qualitative agreement with the lattice result. A system composed of quarks is governed by both Landau diamagnetism and Pauli paramagnetism. The former originates from the spin distribution in the presence of magnetic field, and the latter is a quantum effect which relates to the cyclotron motion of charged particles. Normally the Pauli paramagnetism is stronger than Landau diamagnetism, making the total magnetic susceptibility positive in the mean field approximation. For charged pions which is spin zero, its magnetic susceptibility is only controlled by Landau diamagnetism. In beyond mean field approximation, the system is composed of quarks and mesons, and at low temperature, the meson contribution is dominant and the total magnetic susceptibility is negative, while at high temperature, the dominance of quark contribution makes the system going back to be paramagnetic.

## ACKNOWLEDGMENTS

In this work, J. M., R. W., M. H. and K. X. are supported in part by the National Natural Science Foundation of China (NSFC) Grant Nos: 12235016, 12221005, 12147150 and the Strategic Priority Research Program of Chinese Academy of Sciences under Grant No XDB34030000 and and Fundamental Research Funds for the Central Universities, and S. M. is supported by the NSFC Grant No. 12275204 and Fundamental Research Funds for the Central Universities.

## Appendix A: The sub-leading order of polarization functions

In this part we will derive the specific form of polarization function in weak field expansion. We start with the equation of $\Pi^{00}(q^2)$. For the pole mass, with the static condition, the $B=0$ contribution of the polarization function can be written as

$$\Pi^{00}(q_0^2,0) = -I_1^{(0)}(m_q^2) + q_0^2 I_2^{(0)}(m_q^2,q_0^2,0), \tag{A1}$$

and $I_2^{(0)}(m_q^2,q_0^2,0)$ is given by

$$I_2^{(0)}(m_q^2,q_0^2,0) = 2\int\frac{d^3k}{(2\pi)^3}\frac{\tanh\left(\frac{E_q}{2T}\right)}{E_q(4E_q^2-q_0^2)}. \tag{A2}$$

Besides, the formula of $\Pi^{00}(0,q_i^2)$ for screening mass reads

$$\Pi^{00}(0,q_i^2) = -\frac{T}{4\pi}\sum_{l=-\infty}^{\infty}\frac{1}{\sqrt{-q_i^2}}\left[\arctan\left(\frac{2m_l}{q_i}\right) - \arctan\left(\frac{2m_l-\sqrt{-q_i^2}}{q_i}\right) + \arctan\left(\frac{2q_i}{2m_l-\sqrt{-q_i^2}}\right)\right], \tag{A3}$$

with $m_l^2 = m_q^2+\omega_l^2$ and Matsubara frequency $\omega_l = (2l+1)\pi T$. Here we start with $\Pi^{20}$. Substituting Eq. (2) into Eq. (22), we have

$$\Pi^{20}(q^2) = 2i\int\frac{d^4k}{(2\pi)^4}\mathrm{Tr}_{\{s\}}\left[\gamma_5\frac{k_\perp^2(m+\not{k}_\parallel)+\not{k}_\perp(m^2-k_\parallel^2)}{(k^2-m^2)^4}\gamma_5\frac{m+\not{k}-\not{q}}{(k-q)^2-m^2}\right]$$
$$= 8i\int\frac{d^4k}{(2\pi)^4}\frac{k_\perp^2[m^2-k_\parallel\cdot(k-q)_\parallel]-k_\perp\cdot(k-q)_\perp(m^2-k_\parallel^2)}{(k^2-m^2)^4[(k-q)^2-m^2]}. \tag{A4}$$

To calculate the pole and screening mass in different directions in a form that is as simplified as possible, we should consider these three situations separately. For pole mass, we set the external momentum $q_\perp = q_3 = 0$ and induce a replacement $\int\frac{d^4k}{(2\pi)^4}\to iT\sum_{n=-\infty}^{\infty}\int\frac{d^3k}{(2\pi)^3}$, $k_0\to i\omega_n$ in finite temperature case, to have

$$\Pi^{20}(q_0^2, q_\perp^2 = q_3^2 = 0) = -8T\sum_{n=-\infty}^{\infty}\int\frac{d^3k}{(2\pi)^3}\frac{(i\omega_n)q_0k_\perp^2}{[(i\omega_n)^2-E_q^2]^4\{(i\omega_n-q_0)^2-E_q^2\}}. \tag{A5}$$

One can choose to either do the summation of Matsubara frequencies numerically, or use the package `MatsubaraSum` to do the calculation.

For screening mass, according to [68] the Matsubara summation should be done after the integration of internal momentum. Follow the similar procedure as above, we can set $q_0 = 0$ and get

$$\Pi^{20}(q_0^2 = 0, q_\perp^2, q_3^2) = 8i\int\frac{d^4k}{(2\pi)^4}\frac{k_\perp^2[m^2+k_3\cdot(k-q)_3]-k_\perp\cdot(k-q)_\perp(m^2+k_3^2)-k_\perp q_\perp k_0^2}{(k^2-m^2)^4[(k-q)^2-m^2]}. \tag{A6}$$

Above we introduce an integration over Feynman parameters, which is given by

$$\frac{1}{(k^2-m^2)^a[(k-q)^2-m^2]^b} = \int_0^1 dx\frac{x^{a-1}(1-x)^{b-1}}{\{x(k^2-m^2)+(1-x)[(k-q)^2-m^2]\}^{a+b}}\times\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$
$$= \int_0^1 dx\frac{x^{a-1}(1-x)^{b-1}}{(l^2-\Delta)^{a+b}}\times\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}, \tag{A7}$$

where

$$\Delta = m^2 - x(1-x)q^2,$$
$$l = k-(1-x)q. \tag{A8}$$

Substituting Eq. (A7) into Eq. (A6), setting $k = l+(1-x)q$ and discarding all the odd powers of $l$, we obtain

$$\Pi^{20}(q_0^2 = 0, q_\perp^2, q_3^2) = -32T\sum_{n=-\infty}^{\infty}\int\frac{d^3l}{(2\pi)^3}\int_0^1 dx\frac{x^3(1-x)(-l_\perp^2q_3^2+m^2q_\perp^2+l_3^2q_\perp^2-(i\omega_n)^2q_\perp^2)}{((i\omega_n)^2-l_3^2-l_\perp^2-\Delta)^5}. \tag{A9}$$

Separately set $q_\perp^2 = 0$ and $q_3^2 = 0$ and performing the integration in $l$, we can get the $\Pi^{20}$ in longitudinal and transverse directions, accordingly,

$$\Pi^{20}(q_0^2 = q_\perp^2 = 0, q_3^2) = -\frac{T}{8\pi}\sum_{n=-\infty}^{\infty}\int_0^1 dx\frac{x^3(1-x)q_3^2}{(\Delta-(i\omega_n)^2)^{5/2}},$$
$$\Pi^{20}(q_0^2 = q_3^2 = 0, q_\perp^2) = \frac{T}{16\pi}\sum_{n=-\infty}^{\infty}\int_0^1 dx\ x^3(1-x)\left\{\frac{q_\perp^2}{(\Delta-(i\omega_n)^2)^{5/2}}+\frac{5(m^2-(i\omega_n)^2)q_\perp^2}{(\Delta-(i\omega_n)^2)^{7/2}}\right\}. \tag{A10}$$

Following the similar procedure, we can get the expression for $\Pi^{11}$:

$$\Pi^{11}(q_\perp^2 = q_3^2 = 0, q_0^2) = 4T\sum_{n=-\infty}^{\infty}\int\frac{d^3k}{(2\pi)^3}\frac{m^2+k_3^2-(i\omega_n)(i\omega_n-q_0)}{[(i\omega_n)^2-E_q^2]^2\{(i\omega_n-q_0)^2-E_q^2\}^2},$$
$$\Pi^{11}(q_0^2 = q_\perp^2 = 0, q_3^2) = \frac{T}{8\pi}\sum_{n=-\infty}^{\infty}\int_0^1 dx\ x(1-x)\left\{\frac{1}{(\Delta-(i\omega_n)^2)^{3/2}}+\frac{3[m^2-x(1-x)q_3^2-(i\omega_n)^2]}{(\Delta-(i\omega_n)^2)^{5/2}}\right\},$$
$$\Pi^{11}(q_0^2 = q_3^2 = 0, q_\perp^2) = \frac{T}{8\pi}\sum_{n=-\infty}^{\infty}\int_0^1 dx\ x(1-x)\left\{\frac{1}{(\Delta-(i\omega_n)^2)^{3/2}}+\frac{3[m^2-(i\omega_n)^2]}{(\Delta-(i\omega_n)^2)^{5/2}}\right\}. \tag{A11}$$

For vanishing temperature case, the sub-leading order $\Pi^{02}$ and $\Pi^{11}$ have the form

$$\Pi^{20}(q_\parallel^2, q_\perp^2)(T=0) = \frac{1}{6\pi^2}\int_0^1 dx \frac{x^3(1-x)}{\Delta^3}\left(2m_q^2 q_\perp^2 + (q_\perp^2 + q_\parallel^2)\Delta\right),$$
$$\Pi^{11}(q_\parallel^2, q_\perp^2)(T=0) = \frac{1}{4\pi^2}\int_0^1 dx \frac{x(1-x)}{\Delta^2}\left(m_q^2 + x(1-x)q_\parallel^2 + \Delta\right), \tag{A12}$$

The above expression works well in calculating both pole and screening masses for pions.


[1] F. Preis, A. Rebhan, and A. Schmitt, Inverse magnetic catalysis in field theory and gauge-gravity duality, Lect. Notes Phys. **871**, 51 (2013), arXiv:1208.0536 [hep-ph].

[2] R. Gatto and M. Ruggieri, Quark Matter in a Strong Magnetic Background, Lect. Notes Phys. **871**, 87 (2013), arXiv:1207.3190 [hep-ph].

[3] M. D'Elia, Lattice QCD Simulations in External Background Fields, Lect. Notes Phys. **871**, 181 (2013), arXiv:1209.0374 [hep-lat].

[4] V. A. Miransky and I. A. Shovkovy, Quantum field theory in a magnetic field: From quantum chromodynamics to graphene and Dirac semimetals, Phys. Rept. **576**, 1 (2015), arXiv:1503.00732 [hep-ph].

[5] J. O. Andersen, W. R. Naylor, and A. Tranberg, Phase diagram of QCD in a magnetic field: A review, Rev. Mod. Phys. **88**, 025001 (2016), arXiv:1411.7176 [hep-ph].

[6] D. E. Kharzeev, L. D. McLerran, and H. J. Warringa, The Effects of topological charge change in heavy ion collisions: 'Event by event P and CP violation', Nucl. Phys. A **803**, 227 (2008), arXiv:0711.0950 [hep-ph].

[7] U. Gursoy, D. Kharzeev, and K. Rajagopal, Magnetohydrodynamics, charged currents and directed flow in heavy ion collisions, Phys. Rev. C **89**, 054905 (2014), arXiv:1401.3805 [hep-ph].

[8] V. de la Incera, Nonperturbative Physics in a Magnetic Field, AIP Conf. Proc. **1361**, 74 (2011), arXiv:1004.4931 [hep-ph].

[9] R. C. Duncan and C. Thompson, Formation of very strongly magnetized neutron stars - implications for gamma-ray bursts, Astrophys. J. Lett. **392**, L9 (1992).

[10] H. Li, X.-L. Xia, X.-G. Huang, and H. Z. Huang, Dynamic calculations of magnetic field and implications on spin polarization and spin alignment in heavy ion collisions, Phys. Rev. C **108**, 044902 (2023), arXiv:2306.02829 [nucl-th].

[11] G. S. Bali et al., Thermodynamic properties of QCD in external magnetic fields, PoS **ConfinementX**, 197 (2012), arXiv:1301.5826 [hep-lat].

[12] V. G. Bornyakov, P. V. Buividovich, N. Cundy, O. A. Kochetkov, and A. Schäfer, Deconfinement transition in two-flavor lattice QCD with dynamical overlap fermions in an external magnetic field, Phys. Rev. D **90**, 034501 (2014), arXiv:1312.5628 [hep-lat].

[13] M. D'Elia, F. Manigrasso, F. Negro, and F. Sanfilippo, QCD phase diagram in a magnetic background for different values of the pion mass, Phys. Rev. D **98**, 054509 (2018), arXiv:1808.07008 [hep-lat].

[14] G. Endrodi, Critical point in the QCD phase diagram for extremely strong background magnetic fields, JHEP **07**, 173, arXiv:1504.08280 [hep-lat].

[15] G. Endrodi, M. Giordano, S. D. Katz, T. G. Kovács, and F. Pittler, Magnetic catalysis and inverse catalysis for heavy pions, JHEP **07**, 007, arXiv:1904.10296 [hep-lat].

[16] G. S. Bali, G. Endrődi, and S. Piemonte, Magnetic susceptibility of QCD matter and its decomposition from the lattice, JHEP **07**, 183, arXiv:2004.08778 [hep-lat].

[17] S. P. Klevansky and R. H. Lemmer, Chiral symmetry restoration in the Nambu-Jona-Lasinio model with a constant electromagnetic field, Phys. Rev. D **39**, 3478 (1989).

[18] V. P. Gusynin, V. A. Miransky, and I. A. Shovkovy, Dimensional reduction and catalysis of dynamical symmetry breaking by a magnetic field, Nucl. Phys. B **462**, 249 (1996), arXiv:hep-ph/9509320.

[19] K. Fukushima and Y. Hidaka, Magnetic Catalysis Versus Magnetic Inhibition, Phys. Rev. Lett. **110**, 031601 (2013), arXiv:1209.1319 [hep-ph].

[20] J. Chao, P. Chu, and M. Huang, Inverse magnetic catalysis induced by sphalerons, Phys. Rev. D **88**, 054009 (2013), arXiv:1305.1100 [hep-ph].

[21] L. Yu, H. Liu, and M. Huang, Spontaneous generation of local CP violation and inverse magnetic catalysis, Phys. Rev. D **90**, 074009 (2014), arXiv:1404.6969 [hep-ph].

[22] E. J. Ferrer, V. de la Incera, and X. J. Wen, Quark Antiscreening at Strong Magnetic Field and Inverse Magnetic Catalysis, Phys. Rev. D **91**, 054006 (2015), arXiv:1407.3503 [nucl-th].

[23] S. Fayazbakhsh and N. Sadooghi, Anomalous magnetic moment of hot quarks, inverse magnetic catalysis, and reentrance of the chiral symmetry broken phase, Phys. Rev. D **90**, 105030 (2014), arXiv:1408.5457 [hep-ph].

[24] J. Mei and S. Mao, Inverse catalysis effect of the quark anomalous magnetic moment to chiral restoration and deconfinement phase transitions, Phys. Rev. D **102**, 114035 (2020),

arXiv:2008.12123 [hep-ph].
[25] K. Xu, J. Chao, and M. Huang, Effect of the anomalous magnetic moment of quarks on magnetized QCD matter and meson spectra, Phys. Rev. D **103**, 076015 (2021), arXiv:2007.13122 [hep-ph].
[26] F. Lin, K. Xu, and M. Huang, Magnetism of QCD matter and the pion mass from tensor-type spin polarization and the anomalous magnetic moment of quarks, Phys. Rev. D **106**, 016005 (2022), arXiv:2202.03226 [hep-ph].
[27] M. Kawaguchi and M. Huang, Restriction on the form of the quark anomalous magnetic moment from lattice QCD results, Chin. Phys. C **47**, 064103 (2023), arXiv:2205.08169 [hep-ph].
[28] W. R. Tavares, S. S. Avancini, R. L. S. Farias, and R. P. Cardoso, Artificial first-order phase transition in a magnetized Nambu–Jona-Lasinio model with a quark anomalous magnetic moment, Phys. Rev. D **109**, 016011 (2024), arXiv:2309.04055 [hep-ph].
[29] N. Chaudhuri, S. Ghosh, P. Roy, and S. Sarkar, Collective modes of a massive fermion in a magnetized medium with finite anomalous magnetic moment, Phys. Rev. D **108**, 116006 (2023), arXiv:2310.05769 [hep-ph].
[30] C. A. Islam, M. S. Ali, and M. Huang, Deciding on the anomalous magnetic moment of quarks in a framework of nonlocal NJL model, (2023), arXiv:2302.00696 [hep-ph].
[31] M. Ferreira, P. Costa, O. Lourenço, T. Frederico, and C. Providência, Inverse magnetic catalysis in the (2+1)-flavor Nambu-Jona-Lasinio and Polyakov-Nambu-Jona-Lasinio models, Phys. Rev. D **89**, 116011 (2014), arXiv:1404.5577 [hep-ph].
[32] H. Liu, L. Yu, M. Chernodub, and M. Huang, Possible formation of high temperature superconductor at an early stage of heavy-ion collisions, Phys. Rev. D **94**, 113006 (2016), arXiv:1604.06662 [hep-ph].
[33] S. S. Avancini, R. L. S. Farias, M. Benghi Pinto, W. R. Tavares, and V. S. Timóteo, $\pi_0$ pole mass calculation in a strong magnetic field and lattice constraints, Phys. Lett. B **767**, 247 (2017), arXiv:1606.05754 [hep-ph].
[34] A. Ayala, C. A. Dominguez, L. A. Hernandez, M. Loewe, A. Raya, J. C. Rojas, and C. Villavicencio, Thermomagnetic properties of the strong coupling in the local Nambu–Jona-Lasinio model, Phys. Rev. D **94**, 054019 (2016), arXiv:1603.00833 [hep-ph].
[35] R. L. S. Farias, V. S. Timoteo, S. S. Avancini, M. B. Pinto, and G. Krein, Thermo-magnetic effects in quark matter: Nambu–Jona-Lasinio model constrained by lattice QCD, Eur. Phys. J. A **53**, 101 (2017), arXiv:1603.03847 [hep-ph].
[36] X. Chen, L. Zhang, and D. Hou, Running coupling constant at finite chemical potential and magnetic field from holography, Chin. Phys. C **46**, 073101 (2022), arXiv:2108.03840 [hep-ph].
[37] B.-k. Sheng, X. Wang, and L. Yu, Impacts of inverse magnetic catalysis on screening masses of neutral pions and sigma mesons in hot and magnetized quark matter, Phys. Rev. D **105**, 034003 (2022), arXiv:2110.12811 [hep-ph].
[38] S. Mao and Y. Tian, Pion superfluid phase transition under an external magnetic field including the inverse magnetic catalysis effect, Phys. Rev. D **106**, 094017 (2022), arXiv:2209.10738 [nucl-th].
[39] K. Kamikado and T. Kanazawa, Magnetic susceptibility of a strongly interacting thermal medium with 2+1 quark flavors, JHEP **01**, 129, arXiv:1410.6253 [hep-ph].
[40] B. Sahoo, K. K. Pradhan, D. Sahu, and R. Sahoo, Effect of a magnetic field on the thermodynamic properties of a high-temperature hadron resonance gas with van der Waals interactions, Phys. Rev. D **108**, 074028 (2023), arXiv:2306.03477 [hep-ph].
[41] R. Wen, S. Yin, W.-j. Fu, and M. Huang, Functional renormalization group study of neutral and charged pions in magnetic fields in the quark-meson model, Phys. Rev. D **108**, 076020 (2023), arXiv:2306.04045 [hep-ph].
[42] X. Li, W.-j. Fu, and Y.-x. Liu, Thermodynamics of 2+1 Flavor Polyakov-Loop Quark-Meson Model under External Magnetic Field, Phys. Rev. D **99**, 074029 (2019), arXiv:1902.03866 [hep-ph].
[43] K. Kamikado and T. Kanazawa, Chiral dynamics in a magnetic field from the functional renormalization group, JHEP **03**, 009, arXiv:1312.3124 [hep-ph].
[44] K. Fukushima and J. M. Pawlowski, Magnetic catalysis in hot and dense quark matter and quantum fluctuations, Phys. Rev. D **86**, 076013 (2012), arXiv:1203.4330 [hep-ph].
[45] W.-j. Fu and Y.-x. Liu, Four-fermion interactions and the chiral symmetry breaking in an external magnetic field, Phys. Rev. D **96**, 074019 (2017), arXiv:1705.09841 [hep-ph].
[46] N. Mueller, J. A. Bonnet, and C. S. Fischer, Dynamical quark mass generation in a strong external magnetic field, Phys. Rev. D **89**, 094023 (2014), arXiv:1401.1647 [hep-ph].
[47] N. Mueller and J. M. Pawlowski, Magnetic catalysis and inverse magnetic catalysis in QCD, Phys. Rev. D **91**, 116010 (2015), arXiv:1502.08011 [hep-ph].
[48] J. Braun, W. A. Mian, and S. Rechenberger, Delayed Magnetic Catalysis, Phys. Lett. B **755**, 265 (2016), arXiv:1412.6025 [hep-ph].
[49] S. P. Klevansky, The Nambu-Jona-Lasinio model of quantum chromodynamics, Rev. Mod. Phys. **64**, 649 (1992).
[50] E. Quack and S. P. Klevansky, Effective 1/N(c) expansion in the NJL model, Phys. Rev. C **49**, 3283 (1994).
[51] P. Zhuang, J. Hufner, and S. P. Klevansky, Thermodynamics of a quark - meson plasma in the Nambu-Jona-Lasinio model, Nucl. Phys. A **576**, 525 (1994).
[52] S. Mao, Pions in magnetic field at finite temperature, Phys. Rev. D **99**, 056005 (2019), arXiv:1808.10242 [nucl-th].
[53] L. Li and S. Mao, Light mesons around critical end points in $T-\mu_B-\mu_I-eB$ space, Chin. Phys. C **46**, 094105 (2022), arXiv:2107.10067 [hep-ph].
[54] J. Mei, T. Xia, and S. Mao, Mass spectra of neutral mesons K0,$\pi$0,$\eta$,$\eta$' at finite magnetic field, temperature and quark chemical potential, Phys. Rev. D **107**, 074018 (2023), arXiv:2212.04778 [hep-ph].
[55] H. Liu, X. Wang, L. Yu, and M. Huang, Neutral and charged scalar mesons, pseudoscalar mesons, and diquarks in mag-

netic fields, Phys. Rev. D **97**, 076008 (2018), arXiv:1801.02174 [hep-ph].
[56] M. Coppola, D. Gomez Dumm, S. Noguera, and N. N. Scoccola, Neutral and charged pion properties under strong magnetic fields in the NJL model, Phys. Rev. D **100**, 054014 (2019), arXiv:1907.05840 [hep-ph].
[57] S. S. Avancini, M. Coppola, N. N. Scoccola, and J. C. Sodré, Light pseudoscalar meson masses under strong magnetic fields within the SU(3) Nambu–Jona-Lasinio model, Phys. Rev. D **104**, 094040 (2021), arXiv:2109.01911 [hep-ph].
[58] L. Li and S. Mao, Inverse magnetic catalysis effect and current quark mass effect on mass spectra and Mott transitions of pions under external magnetic field, Phys. Rev. D **108**, 054001 (2023), arXiv:2308.12491 [hep-ph].
[59] S. Mao, Inverse magnetic catalysis in Nambu–Jona-Lasinio model beyond mean field, Phys. Lett. B **758**, 195 (2016), arXiv:1602.06503 [hep-ph].
[60] S. Mao, Chiral Symmetry Restoration and Quark Deconfinement beyond Mean Field in a Magnetized PNJL Model, Phys. Rev. D **97**, 011501 (2018), arXiv:1712.06062 [nucl-th].
[61] D. Gomez Dumm, S. Noguera, and N. N. Scoccola, Charged meson masses under strong magnetic fields: Gauge invariance and Schwinger phases, Phys. Rev. D **108**, 016012 (2023), arXiv:2306.04128 [hep-ph].
[62] W.-j. Fu, C. Huang, J. M. Pawlowski, and Y.-y. Tan, Four-quark scatterings in QCD I, SciPost Phys. **14**, 069 (2023), arXiv:2209.13120 [hep-ph].
[63] W.-j. Fu, C. Huang, J. M. Pawlowski, and Y.-y. Tan, Four-quark scatterings in QCD II, (2024), arXiv:2401.07638 [hep-ph].
[64] E. V. Shuryak, Physics of the pion liquid, Phys. Rev. D **42**, 1764 (1990).
[65] R. D. Pisarski and M. Tytgat, Propagation of cool pions, Phys. Rev. D **54**, R2989 (1996), arXiv:hep-ph/9604404.
[66] S. Fayazbakhsh, S. Sadeghian, and N. Sadooghi, Properties of neutral mesons in a hot and magnetized quark matter, Phys. Rev. D **86**, 085042 (2012), arXiv:1206.6051 [hep-ph].
[67] S. Fayazbakhsh and N. Sadooghi, Weak decay constant of neutral pions in a hot and magnetized quark matter, Phys. Rev. D **88**, 065030 (2013), arXiv:1306.2098 [hep-ph].
[68] B. Sheng, Y. Wang, X. Wang, and L. Yu, Pole and screening masses of neutral pions in a hot and magnetized medium: A comprehensive study in the Nambu–Jona-Lasinio model, Phys. Rev. D **103**, 094001 (2021), arXiv:2010.05716 [hep-ph].
[69] M. Huang, P.-f. Zhuang, and W.-q. Chao, Massive meson fluctuation in NJL model, Commun. Theor. Phys. **34**, 91 (2000), arXiv:hep-ph/9903304.
[70] M. Oertel, M. Buballa, and J. Wambach, Meson loop effects in the NJL model at zero and nonzero temperature, Phys. Atom. Nucl. **64**, 698 (2001), arXiv:hep-ph/0008131.
[71] L. Landau, Diamagnetismus der Metalle, Zeitschrift für Physik **64**, 629 (1930).
[72] L. Landau and E. Lifshitz, in Statistical Physics (Third Edition), edited by L. Landau and E. Lifshitz (Butterworth-Heinemann, Oxford, 1980) third edition ed., pp. 158–190.

# Study of Long Range Force in P2SO and T2HKK

Priya Mishra,[1,*] Rudra Majhi,[2,†] Sambit Kumar Pusty,[1,‡] Monojit Ghosh,[3,§] and Rukmani Mohanta[1,¶]

[1] *School of Physics, University of Hyderabad, Hyderabad - 500046, India*
[2] *Nabarangpur College, Nabarangpur - 764063, Odisha, India*
[3] *Center of Excellence for Advanced Materials and Sensing Devices, Ruđer Bošković Institute, 10000 Zagreb, Croatia*

arXiv:2402.19178v1 [hep-ph] 29 Feb 2024

## Abstract

In this paper we have studied the sensitivity of the future long-baseline neutrino experiments P2SO and T2HKK to the long-range force (LRF). In the context of these two experiments, our aim is to study: (i) the capability to put bounds on the LRF parameters, (ii) effect of LRF in the measurement of standard oscillation parameters and (iii) capability to constrain the mass of the new gauge boson and the value of new coupling constant, that gives rise to LRF due to matter density in Sun. In our study, we find that among the different neutrino experiments, the best bound on the LRF parameters including mass of the new gauge boson and the value of new coupling constant will come from the P2SO experiment. Our study also shows that LRF has non-trivial effect on the determination of the standard neutrino oscillation parameters except the precision of $\Delta m^2_{31}$. For this parameter, the precision remains unaltered in the presence of LRF for both these experiments.

* mishpriya99@gmail
† rudra.majhi95@gmail
‡ pustysambit@gmail.com
§ mghosh@irb.hr
¶ rmsp@uohyd.ac.in

## I. INTRODUCTION

Neutrinos, being elusive subatomic particles with intriguing properties, hold great promise as a tool for discovering and understanding new phenomena in the realm of physics. Their unique characteristics, such as extremely weak interactions with matter and ability to change between different flavors, make them particularly interesting. Numerous experiments provide evidence that lepton flavor is not conserved as neutrinos propagate [1]. Neutrinos undergo oscillations, transitioning between different flavors, and the frequency of these oscillations is contingent upon the distance travelled and the energy of the neutrinos. The evolution of neutrino flavor during their traversal through matter is influenced by the Mikheev-Smirnov-Wolfenstein (MSW) mechanism [2]. This mechanism is characterized by the elastic forward scattering of neutrinos with matter. In the framework of the Standard Model (SM) of particle interactions, this effect is precisely determined. The presence of electrons in matter creates an effective potential for the neutrinos, known as the matter potential (V), plays a crucial role in influencing flavor transitions [3]. In the MSW mechanism, this potential is proportional to the electron number density at the neutrino's position:

$$V_{CC} = \sqrt{2} G_F N_e(r). \tag{1}$$

Here, $G_F$ is the Fermi constant, and $N_e(r)$ is the number density of electrons. The matter potential can be modified by introducing flavor dependent interactions and such alterations have the capacity to reshape the pattern of flavor transitions experienced by neutrinos as they traverse through matter. One example of such interactions is flavor-dependent vector-like leptonic Long Range Force (LRF) [4–6].

The SM gauge group $SU(2)_L \times U(1)_Y$, can be extended by $U(1)$ symmetries. The anomalies associated with $L_e$, $L_\mu$, $L_\tau$ and $B/3$ symmetries in the particle spectrum of the SM gauge group are equivalent, and combinations such as $L_e - L_\mu$, $L_e - L_\tau$ and $L_\mu - L_\tau$ can be consistently accommodated in a way that preserves the condition of gauge group anomaly free [7–10]. It is important to emphasize that at any given moment, any one of these three symmetries can be seamlessly incorporated into the SM gauge group, as outlined in references [8, 9]. This paves the way for them to be potentially natural and efficient extensions of the SM. An additional new neutral vector boson is introduced by gauging each one of these symmetries, which mediates novel neutrino interactions with matter. In the context of local symmetry, the Goldstone boson is absorbed by the corresponding gauge boson, resulting in the acquisition of mass by the latter. The mass scale of this gauge boson can be interpreted in two ways: either through contact interaction or long range interaction. If the new gauge boson possesses sufficient mass, it will induce contact interactions and consequently lead to the emergence of Neutral Current (NC) Non-Standard Interactions (NSI) [11]. On the flip side, if the mediator is excessively lightweight, the approximation of contact interaction becomes invalid, and the flavor-specific forces between neutrinos and matter particles extend over long distances. In such instances, neutrino propagation can still be characterized in

terms of a matter potential. However, this potential is no longer solely determined by the particle number density at the neutrino's position in the medium. Instead, it relies on the average matter density within a radius of approximately $1/M_{Z'}$ around the neutrino with $M_{Z'}$ being the mass of the new gauge boson $Z'$ [12]. With such an exceedingly small mass for $M_{Z'}$, its interaction with matter is extremely weak, resulting in a very small gauge coupling ($g'$). For this reason, this scenario can be considered as an invisible sector. There exist constraints on the LRF parameters expressed in terms of $\alpha_{LRF} = g'^2/4\pi$ derived from the measurement of differential acceleration of the earth and moon towards the Sun from lunar ranging [13–16]. The values of LRF parameter is obtained as $\alpha_{LRF} = 3.4 \times 10^{-49}$ at $2\sigma$ C.L. for the interaction range $\leq 1$ A.U.

In general, interactions induced by LRF alter the matter potential in neutrino propagation. Because of this alteration of the matter potential, the probabilities of neutrino oscillation will get modified and therefore one can study LRF in the neutrino oscillation experiments. The concept of long range interaction with the help of flavor symmetries in the context of neutrino oscillation was first introduced in Ref. [17], where the bounds on flavor dependent LRF are obtained from atmospheric neutrinos. Later, bounds from solar neutrinos are also obtained in Ref. [18]. In a study conducted in Ref. [19], bounds on the couplings of both vector and non-vector long-range forces have been reported from the oscillation of the solar neutrinos. There exist several other descriptions in the literature, for example in Refs. [6, 20–24] the effect of LRF in neutrino oscillation experiments has been illustrated. In this paper, we will study the effect of LRF in the upcoming long-baseline experiments P2SO [25] and T2HKK [26]. For the study of LRF in other two future long-baseline experiments i.e., T2HK [27] and DUNE [28], we refer to Refs. [21, 23]. In this study, our aim is to estimate the upper bounds of the LRF parameters and cosequently study their effects on the determination of the standard oscillation parameters. We will also estimate the bounds on the mass of the new gauge boson and the new gauge coupling for LRF due to the Sun.

This paper is structured in the following manner. In the next section, we will describe the theoretical overview of the long-range force and write down the expressions of the LRF potentials. Next, we will show how these potentials can be incorporated in the neutrino oscillation Hamiltonian in a model independent way. After this, we will write down a prescription on how one can calculate the neutrino oscillation probabilities in the presence of a LRF. In section V, we will provide the configurations of the experiments that we will use in our calculation. Section VI outlines simulation details of our numerical analysis. In section VII, we present our results and finally, in section VIII, we summarize our findings and conclude.

## II. THEORETICAL BACKGROUND

According to the SM, lepton family number is believed to be conserved in nature. However, certain decay processes such as $\mu \to e\gamma$, $\mu \to ee\overline{e}$, and $\tau \to \mu\gamma$ with upper limits on the branching ratios (BR), $4.2 \times 10^{-13}$ [29], $1 \times 10^{-12}$ [30] and $4.2 \times 10^{-8}$ [31] respectively, are considered as potential indications of lepton flavor violation (LFV). As these numbers are relatively small, it is challenging to conclusively confirm the existence of LFV in nature. We take advantage of small BR for LFV processes and assume that lepton flavor conservation could be a fundamental symmetry of nature and hence, the specific linear combinations involving $L_e$, $L_\mu$, and $L_\tau$ may be subject to gauging [7–10]. In this work, we consider three such combinations which are: $L_e - L_\mu$, $L_e - L_\tau$ and $L_\mu - L_\tau$ symmetries. In scenarios where the additional $U(1)$ corresponds to $L_e - L_j$ for $j = \mu, \tau$; the electrons within celestial bodies such as the Sun or the Earth create a potential that influences neutrinos in terrestrial experiments [6, 17, 18]. The flavor-specific nature of $U(1)_{L_e - L_j}$ leads to alterations in neutrino oscillations, allowing for constraints on the gauge coupling. The effective potential for neutrinos on Earth, in such case is defined as:

$$V_{ej} = g_{ej}^2 \frac{N_e}{4\pi r} e^{-M_{Z_{ej}} r}, \tag{2}$$

where, $g_{ej}$ is the new gauge coupling i.e., $g_{e\mu}$ for $L_e - L_\mu$ and $g_{e\tau}$ for $L_e - L_\tau$ symmetries, $N_e \simeq 10^{57}$ [32], is the number of electrons inside the Sun[1], $M_{Z_{ej}}$ is the mass of new gauge boson and $r$ measures the distance between the source of potential to the neutrino on Earth. The absence of muons and taus on earth matter appears to exclude parallel investigations of $L_\mu - L_\tau$, as its associated gauge boson $Z_{\mu\tau}$ does not directly couple to electrons, protons or neutrons. Nevertheless, there exists an indirect effect stemming from the mixing between the SM gauge boson Z and $Z_{\mu\tau}$ particles. Assuming that there are equal number of electrons and protons inside the Sun, so that their effect cancel each other completely and neutrinos on Earth are influenced by only neutrons inside the Sun. Such effective potential is defined as [33],

$$V_{\mu\tau} = g_{\mu\tau}(\xi - \sin\theta_w \chi) \frac{e}{4 \sin\theta_w \cos\theta_w} \frac{N_n}{4\pi r} e^{-M_{Z_{\mu\tau}} r}, \tag{3}$$

where, $g_{\mu\tau}$ is the new gauge coupling associated with the new gauge boson $Z_{\mu\tau}$ with mass $M_{Z_{\mu\tau}}$, $\chi$ is the kinetic mixing parameter between $Z$ and $Z_{\mu\tau}$ [34], $\xi$ is the rotation angle between mass and flavour bases of gauge bosons, $\theta_w$ is the Weinberg angle and $N_n \simeq N_e/4 \simeq 1.5 \times 10^{56}$ [33], is the number of neutrons in the Sun. The bound on the value of $(\xi - \sin\theta_w \chi)$ is $\leq 5 \times 10^{-24}$ [33] and $e/\sin\theta_w \cos\theta_w = 0.723$.

Our goal in this work is to study the effects of the LRF potentials $V_{e\mu}$, $V_{e\tau}$ and $V_{\mu\tau}$ in the neutrino oscillation experiments.

[1] The $e$, $p$, $n$ within the Earth can also generate a long-range potential at the neutrino site. However, it's noteworthy that this potential is approximately 10 times smaller compared to the potential generated by matter in the Sun [18].

## III. FORMALISM

In the three flavor scenario, the effective Hamiltonian (in mass basis) for the propagation of neutrinos in the vacuum is given as follows:

$$H_{vac} = \frac{1}{2E}\begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}, \tag{4}$$

where $m_1$, $m_2$, $m_3$ are the masses of the neutrinos $(\nu_1, \nu_2, \nu_3)$ and $E$ is their energy. Including the matter potential and the potential due to long range force in flavor basis one obtains,

$$H_{\nu/\overline{\nu}} = \frac{1}{2E}\left[U\begin{pmatrix} 0 & 0 & 0 \\ 0 & \Delta m_{21}^2 & 0 \\ 0 & 0 & \Delta m_{31}^2 \end{pmatrix}U^\dagger\right] \pm H_{\rm matter} \pm H_{LRF}, \tag{5}$$

with

$$H_{LRF} = \begin{cases} \text{diag}(V_{e\mu}, -V_{e\mu},\ 0) & \text{for } \ U(1)_{L_e - L_\mu}, \\ \text{diag}(V_{e\tau},\ 0, -V_{e\tau}) & \text{for } \ \mathrm{U(1)_{L_e - L_\tau}}, \\ \text{diag}(0, V_{\mu\tau}, -V_{\mu\tau}) & \text{for } \ U(1)_{L_\mu - L_\tau}, \end{cases} \tag{6}$$

where '+' sign is for neutrino and '−' sign is for antineutrino flavor states, $U$ is the unitary matrix having three mixing angles i.e., $\theta_{12}$, $\theta_{13}$ and $\theta_{23}$ and one Dirac type CP phase $\delta_{\rm CP}$ to rotate mass basis to flavor basis also called PMNS matrix for this particular purpose, and $\Delta m_{ij}^2 = m_i^2 - m_j^2$ $(i = 2, 3\ \&\ j = 1)$. Since matter contains only $e$, $u$ and $d$, the form of $H_{\rm matter}$ will have only '$ee$' term and this will be equal to the earth matter potential $V_{\rm CC}$ defined in Eqn. (1). It can be further simplified as:

$$V_{CC} \simeq 7.5\, X_e \left(\frac{\rho}{10^{14}\ (\mathrm{g/cm^3})}\right)\ \mathrm{eV}. \tag{7}$$

In this context, $X_e$ is defined as the ratio of the number of electrons $(N_e)$ to the sum of the numbers of protons $(N_p)$ and neutrons $(N_n)$. For a neutral medium, where the charges balance, $X_e$ takes the value 0.5. The value of $\rho$ is the earth matter density which is around 2.95 g/cm$^3$ for the P2SO experiment and around 2.7 g/cm$^3$ for the T2HKK experiment.

The structure of $H_{LRF}$ depends on the coupling of $Z_{\alpha\beta}$ to the leptons generation. For instance, as $Z_{e\mu}$ can only couple to the first and second generations of leptons, the resulting potential emerges solely in the '$ee$' and '$\mu\mu$' terms, while other matrix elements are zero in $H_{LRF}$.

## IV. ANALYTICAL EXPRESSIONS

The analytical expressions for the oscillation probabilities in the presence of LRF for $L_e - L_\mu$ and $L_e - L_\tau$ cases have been derived in Ref. [21, 22] whereas for the analytical

expressions due to $L_\mu - L_\tau$ case, one can follow Ref. [35]. The total Hamiltonian $H_\nu$ defined in Eqn. (5) can be diagonalized using a unitary matrix $\hat{U}$ defined as,

$$\hat{U} = \hat{R}(\theta'_{23})\hat{R}(\theta'_{13})\hat{R}(\theta'_{12}) \ , \tag{8}$$

where for simplicity, it is assumed that CP is conserved. $\theta'_{ij}$ are the modified mixing angles and can be expressed as:

$$\begin{aligned}
\tan 2\theta'_{23} &\simeq \frac{\Gamma \sin 2\theta_{23} - \zeta \cos 2\theta_{23}}{\Gamma \cos 2\theta_{23} + \zeta \sin 2\theta_{23} - 2\hat{B}} \ , \\
\tan 2\theta'_{13} &\simeq \frac{\sin 2\theta_{13}(1 - \alpha \sin^2\theta_{12}) \cos \Delta\theta_{23} - \eta \sin \Delta\theta_{23}}{(\lambda_3 - \hat{A} - \alpha \sin^2\theta_{12} \cos^2\theta_{13} - \sin^2\theta_{13})} \ , \\
\tan 2\theta'_{12} &\simeq \frac{\cos\theta'_{13}[\eta \cos \Delta\theta_{23} + \sin 2\theta_{13}(1 - \alpha \sin^2\theta_{12}) \sin \Delta\theta_{23}]}{(\lambda_2 - \lambda_1)} \ ,
\end{aligned} \tag{9}$$

with

$$\lambda_3 = \frac{1}{2}\Big[\Gamma + 2\alpha \cos^2\theta_{12} + \frac{\zeta \sin 2\theta_{23} + \Gamma \cos 2\theta_{23} - 2\hat{B}}{\cos 2\theta'_{23}}\Big] \ ,$$

$$\lambda_2 = \frac{1}{2}\Big[\Gamma + 2\alpha \cos^2\theta_{12} - \frac{\zeta \sin 2\theta_{23} + \Gamma \cos 2\theta_{23} - 2\hat{B}}{\cos 2\theta'_{23}}\Big] \ ,$$

$$\lambda_1 = \frac{1}{2}\Big[\lambda_3 + \hat{A} + \sin^2\theta_{13} + \alpha \sin^2\theta_{12} \cos^2\theta_{13} - \frac{\lambda_3 - \hat{A} - \sin^2\theta_{13} - \alpha \sin^2\theta_{12} \cos^2\theta_{13}}{\cos 2\theta'_{13}}\Big] \ , \tag{10}$$

and

$$\begin{aligned}
\Gamma &= \cos^2\theta_{13} - \alpha \cos^2\theta_{12} + \alpha \sin^2\theta_{12} \sin^2\theta_{13} \ , \\
\zeta &= \alpha \sin 2\theta_{12} \sin\theta_{13} \ , \eta = \alpha \sin 2\theta_{12} \cos\theta_{13} \ ,
\end{aligned} \tag{11}$$

$$\alpha = \frac{\Delta m^2_{21}}{\Delta m^2_{31}}, \quad \hat{A} = \frac{2EV_{CC}}{\Delta m^2_{31}}, \ \ \hat{B} = \frac{2EV_{\mu\tau}}{\Delta m^2_{31}} \ , \ \ \Delta\theta_{23} = \theta_{23} - \theta'_{23} \tag{12}$$

The diagonalised form of $H_\nu$ can be written as,

$$\hat{H}_\nu = \frac{1}{2E}\begin{pmatrix} m'^2_1 & 0 & 0 \\ 0 & m'^2_2 & 0 \\ 0 & 0 & m'^2_3 \end{pmatrix} \ , \tag{13}$$

with $m'_i$ $(i = 1, 2, 3)$ being the modified masses of neutrinos defined as,

$$m'^2_1 \simeq \frac{\Delta m^2_{31}}{2}\Big[\lambda_1 + \lambda_2 + \frac{\lambda_1 - \lambda_2}{\cos 2\theta'_{12}}\Big] \ , \tag{14}$$

$$m_2'^2 \simeq \frac{\Delta_{31}}{2}\Big[\lambda_1 + \lambda_2 - \frac{\lambda_1 - \lambda_2}{\cos 2\theta'_{12}}\Big], \tag{15}$$

$$m_3'^2 \simeq \frac{\Delta m^2_{31}}{2}\Big[\lambda_3 + \hat{A} + \sin^2\theta_{13} + \alpha \sin^2\theta_{12}\cos^2\theta_{13} + \frac{\lambda_3 - \hat{A} - \sin^2\theta_{13} - \alpha\sin^2\theta_{12}\cos^2\theta_{13}}{\cos 2\theta'_{13}}\Big], \tag{16}$$

The probability of $\nu_\mu \to \nu_e$ transition can be given as [21, 36],

$$\begin{aligned} P_{\nu_\mu\to\nu_e} = & \; 4\,\hat{U}^2_{\mu 2}\hat{U}^2_{e2}\sin^2\frac{\Delta m'^2_{21}L}{4E} + 4\,\hat{U}^2_{\mu 3}\hat{U}^2_{e3}\sin^2\frac{\Delta m'^2_{31}L}{4E} \\ & + \; 2\,\hat{U}_{\mu 3}\hat{U}_{e3}\hat{U}_{\mu 2}\hat{U}_{e2}\left(4\sin^2\frac{\Delta m'^2_{21}L}{4E}\sin^2\frac{\Delta m'^2_{31}L}{4E}\right) \\ & + \; 2\,\hat{U}_{\mu 3}\hat{U}_{e3}\hat{U}_{\mu 2}\hat{U}_{e2}\left(\sin\frac{\Delta m'^2_{21}L}{2E}\sin\frac{\Delta m'^2_{31}L}{2E}\right), \end{aligned} \tag{17}$$

and the expression for $\nu_\mu \to \nu_\mu$ transition probability can be given as,

$$\begin{aligned} P_{\nu_\mu\to\nu_\mu} = 1 - \Bigg( & 4\,\hat{U}^2_{\mu 2}\left(1 - \hat{U}^2_{\mu 2}\right)\sin^2\frac{\Delta m'^2_{21}L}{4E} \\ & + 4\,\hat{U}^2_{\mu 3}\left(1 - \hat{U}^2_{\mu 3}\right)\sin^2\frac{\Delta m'^2_{31}L}{4E} \\ & - 2\,\hat{U}^2_{\mu 2}\hat{U}^2_{\mu 3}\left(4\sin^2\frac{\Delta m'^2_{21}L}{4E}\sin^2\frac{\Delta m'^2_{31}L}{4E}\right) \\ & - 2\,\hat{U}^2_{\mu 2}\hat{U}^2_{\mu 3}\left(\sin\frac{\Delta m'^2_{21}L}{2E}\sin\frac{\Delta m'^2_{31}L}{2E}\right)\Bigg). \end{aligned} \tag{18}$$

Where, $L$ represents the baseline which is 2595 km for P2SO and 1100 km for T2HKK and $\Delta m'^2_{ij}$ is the effective mass square difference.

## V. EXPERIMENTAL DETAILS

In this study, we mainly focus on the two upcoming long-baseline experiments P2SO and T2HKK. Here are the relevant experimental details:

### A. P2SO

The Protvino to Super-ORCA (P2SO) is an impending long-baseline experiment where neutrinos will be produced at a U-70 synchrotron located at Protvino, Russia, and will propagate towards the detector situated at a distance of 2595 km in the Mediterranean Sea 40 km offshore Toulon, France. We refer to Refs. [25, 37–39] for the detailed configuration of the P2SO experiment. The accelerator will produce a 450 KW beam corresponding to $4 \times 10^{20}$ protons on target annually for P2SO configuration. The Super-ORCA detector will

use ten times more denser detector compared to the ORCA. Energy window ranges from 0.2 GeV to 10 GeV for P2SO experiment. We have considered total run period of six years consisting of three years in neutrino and three years in antineutrino modes.

### B. T2HKK

In our computational analysis for T2HKK experiment, we adhere to the configuration outlined in Ref [26]. The neutrino source is situated at J-PARC, featuring a beam power of 1.3 MW and a total exposure of $27 \times 10^{21}$ protons on target (POT), corresponding to a comprehensive 10-year operational period. This operational timeline is divided, with 2.5 years dedicated to neutrino mode and the subsequent 7.5 years to anti-neutrino mode, maintaining a balanced 1 : 3 ratio between neutrinos and anti-neutrinos. This setup will have a detector of 187 kt fiducial volume positioned at 295 km from the source, exposed to a 2.5° off-axis flux. Additionally, there will be another detector with the same volume, situated 1100 km from the source, encountering a 1.5° off-axis flux. The energy window under consideration spans from 0-3 GeV. This is an alternate setup of the T2HK (Tokai to Hyper Kamiokande) experiment where both the detectors will be placed at 295 km.

## VI. SIMULATION DETAILS

We have simulated P2SO and T2HKK experiments using GLoBES [40, 41] software package. In order to implement LRF we have modified the probability engine of GLoBES. We have estimated the sensitivity in terms of $\chi^2$ analysis, where we use the Poisson log-likelihood and assume that it is $\chi^2$-distributed:

$$\chi^2_{\rm stat} = 2\sum_{i=1}^{n}\left[N_i^{\rm test} - N_i^{\rm true} - N_i^{\rm true}\log\left(\frac{N_i^{\rm test}}{N_i^{\rm true}}\right)\right], \tag{19}$$

where $N^{\rm test}$ and $N^{\rm true}$ are the number of events in the test and true spectra respectively, and $n$ is the number of energy bins. The systematic error is incorporated by the method of pull [42, 43]. The values of the oscillation parameters are taken from NuFit 5.2 and are listed in Tab. I. While calculating the $\chi^2$, the true values of the oscillation parameters are always kept at their best-fit values as shown in Tab. I. The relevant oscillation parameters are minimized in the test using the current uncertainties associated with these parameters. We will present all our results for the normal ordering of the neutrino masses i.e., $\Delta m^2_{31} > 0$.

## VII. RESULTS

In this section, we will present our results. First, we will show how the appearance and disappearance channel probabilities depend on the LRF parameters. Then we will compute

| Parameters | True values $\pm$ $1\sigma$ |
|---|---|
| $\sin^2\theta_{12}$ | $0.303^{+0.012}_{-0.012}$ |
| $\sin^2\theta_{13}$ | $0.02225^{+0.00056}_{-0.00059}$ |
| $\sin^2\theta_{23}$ | $0.451^{+0.019}_{-0.016}$ |
| $\delta_{\rm CP}[^\circ]$ | $232^{+36}_{-26}$ |
| $\Delta m^2_{21}$ $[10^{-5}$ eV$^2]$ | $7.41^{+0.21}_{-0.20}$ |
| $\Delta m^2_{31}$ $[10^{-3}$ eV$^2]$ | $2.507^{+0.026}_{-0.027}$ |

TABLE I: Values of oscillation parameters with their $1\sigma$ errors used for the study [1].

the bounds on the LRF parameters in P2SO and T2HKK. After that, we will study the effect of the LRF parameters in the determination of CP violation, octant of $\theta_{23}$ and mass ordering of the neutrinos for these two experiments. Then we will study whether the precision of $\delta_{\rm CP}$, $\Delta m^2_{31}$ and $\theta_{23}$ will be affected if LRF exists in nature. Finally, we will show the capability of the above mentioned experiments to constrain the mass and coupling of the new gauge boson for the long-range force due to Sun.

### A. Effect on the probabilities

First, let us discuss how LRF affects the oscillation probabilities in the long-baseline experiments. Fig. 1 shows the probabilities for standard interaction (SI) and LRF potentials for P2SO and T2HKK. Left and right columns represent the appearance and disappearance probabilities, respectively. Upper (lower) panels are for neutrino (antineutrino) cases. In each panel, the black line is for the probability in SI case. Solid lines are for P2SO and dashed lines are for T2HKK. Orange, cyan, and purple curves represent the variation of probabilities with respect to LRF parameters $V_{e\mu}, V_{e\tau}$, and $V_{\mu\tau}$, respectively. These panels are generated for the neutrino energy of 4.9 GeV for P2SO and 0.75 GeV for T2HKK. These are the energies where the event spectrum peaks in neutrinos for these experiments. Here it is important to note that for the T2HKK baseline, the first oscillation maxima does not occur at 0.75 GeV. Rather, this energy corresponds somewhere between the first and the second oscillation maximum. For this reason, the values of the neutrino probabilities and the antineutrino probabilities for T2HKK are not very different at $V_{\alpha\beta} = 0$. Whereas for P2SO, at $V_{\alpha\beta} = 0$, the neutrino probabilities are higher than antineutrino probabilities as for this baseline, 4.9 GeV is close to the first oscillation maximum.

From the figure, we see that, for both P2SO and T2HKK, LRF parameters of $e\tau$ and $\mu\tau$ sectors significantly affect the appearance probability in neutrino case, whereas in the antineutrino appearance probabilities, the significant change occurs due to $V_{e\mu}$. In the disappearance channel, all the three LRF parameters significantly affect the probabilities in P2SO in both neutrinos and antineutrinos. However, for T2HKK, the effect of $V_{e\mu}$ and $V_{\mu\tau}$

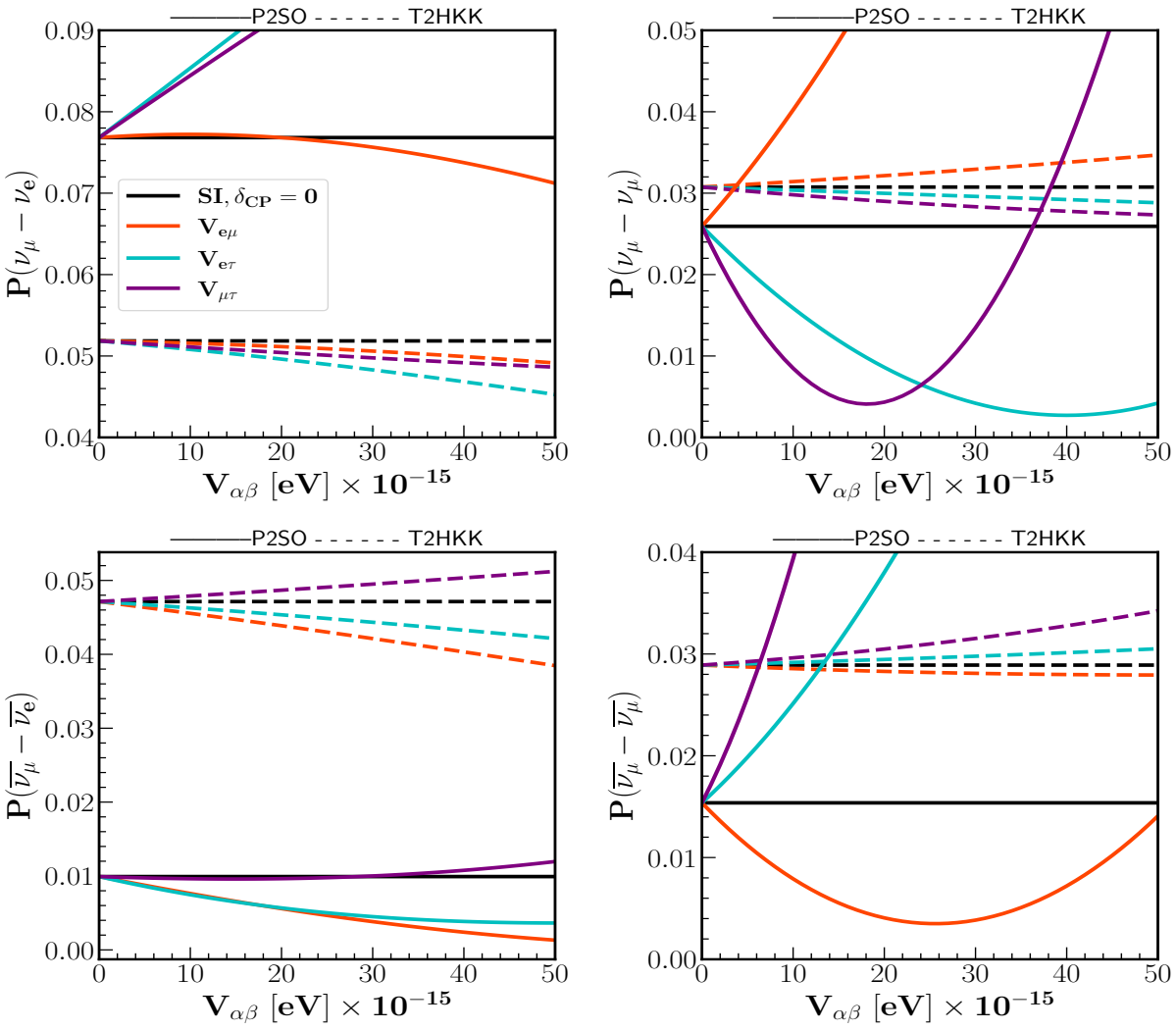



FIG. 1: Appearance and disappearance probabilities for neutrino and antineutrino cases as a function of LRF potentials for P2SO and T2HKK. Neutrino energy is taken where the neutrino events peak.

is higher in neutrinos and the effect of $V_{\mu\tau}$ is higher in antineutrinos.

From the panels, we also see certain differences in the probabilities for P2SO and T2HKK. For example, in the top left panel, the probabilities are increasing functions of $V_{e\tau}$ and $V_{\mu\tau}$ for P2SO, but they are decreasing functions in T2HKK. Some differences in the nature of the probability curves in P2SO and T2HKK can also be observed in the disappearance channel probabilities for both neutrinos and antineutrinos. To understand these differences, one way is to look at the effective values of the mixing parameters in presence of LRF. This we have shown in Fig. 2.

In the top left, top right and bottom panels of Fig. 2, we present the variation of the effective parameters $\Delta m'^2_{31}$ for neutrinos, $\theta'_{12}$ for neutrinos and $\theta'_{12}$ for antineutrinos, respec-

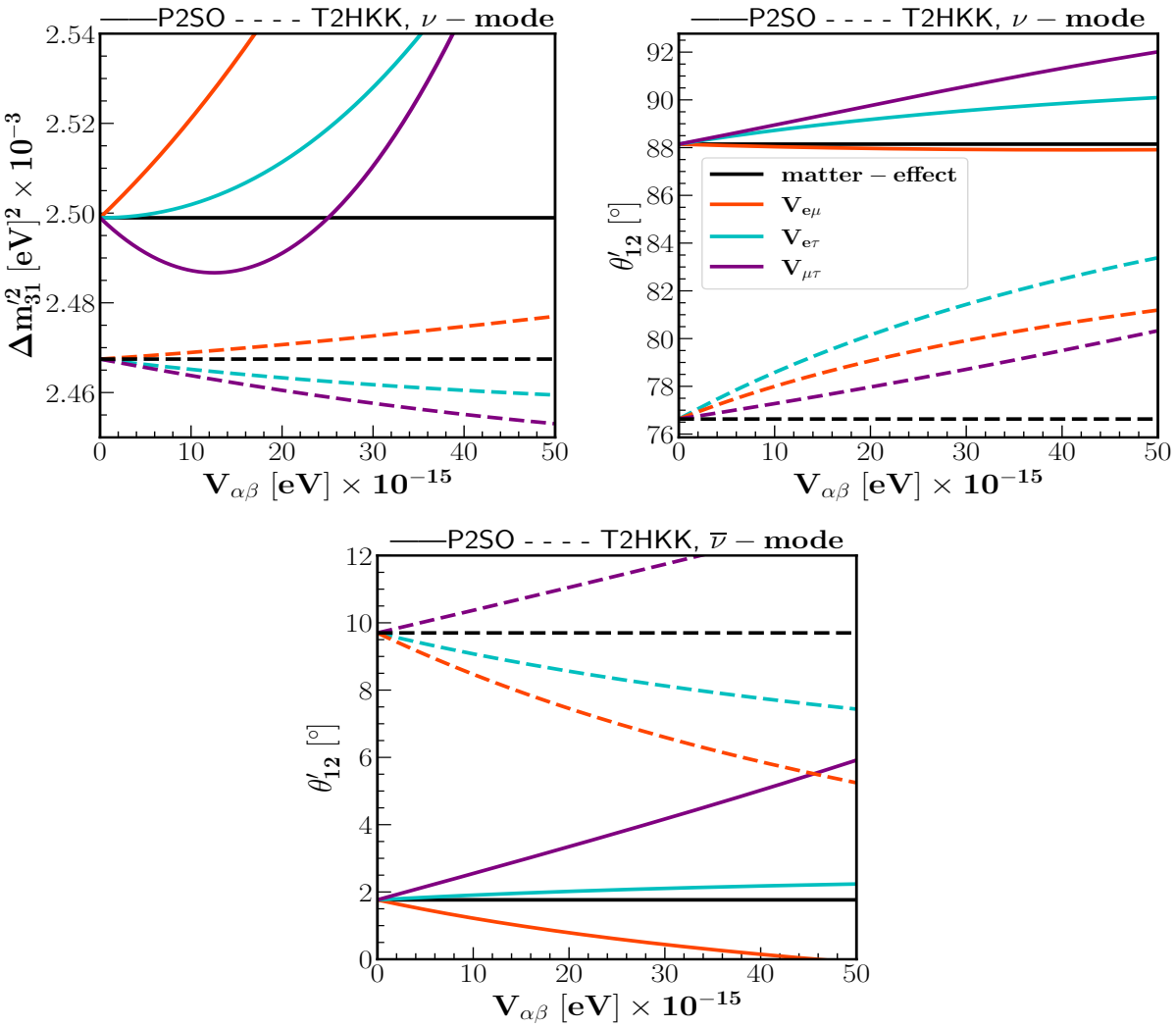



FIG. 2: Presenting the relevant modified oscillation parameters with the variation in the LRF parameter $V_{\alpha\beta}$, which mainly affect the oscillation probabilities presented in Fig. 1.

tively, as a function of the LRF parameter $V_{\alpha\beta}$. The other specifications of these panels are exactly same as Fig. 1. From these panels, we see that for some of these curves, the behaviour in P2SO and T2HKK are quite different. For example, in the top left panel, the cyan curve is an increasing function of $V_{e\tau}$ in P2SO, whereas it is a decreasing function of $V_{e\tau}$ in T2HKK. Further, the nature of the purple curve is also different in P2SO and T2HKK. Similarly, the differences between P2SO and T2HKK can also be seen in the orange curve in the top right panel and the cyan curve in the bottom panel. Because of these differences in the behaviour of the effective mixing parameters in P2SO and T2HKK, the probability curves also become different in these two experiments. We have checked that, the behaviour of the other effective mixing parameters, which are not shown in Fig. 2 are similar in P2SO and T2HKK.

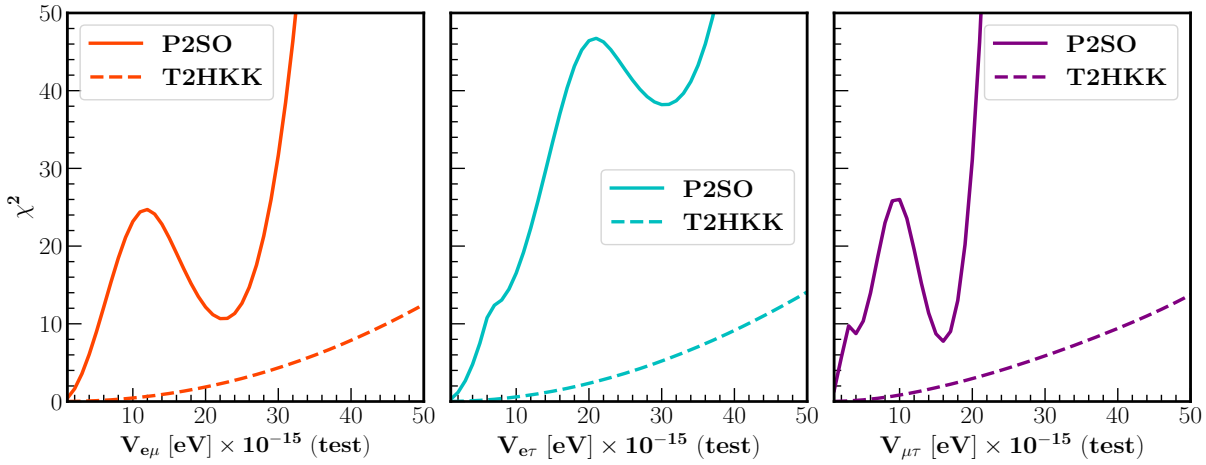



FIG. 3: Sensitivity limits on LRF parameters $V_{e\mu}$, $V_{e\tau}$, and $V_{\mu\tau}$ for P2SO and T2HKK experiments.

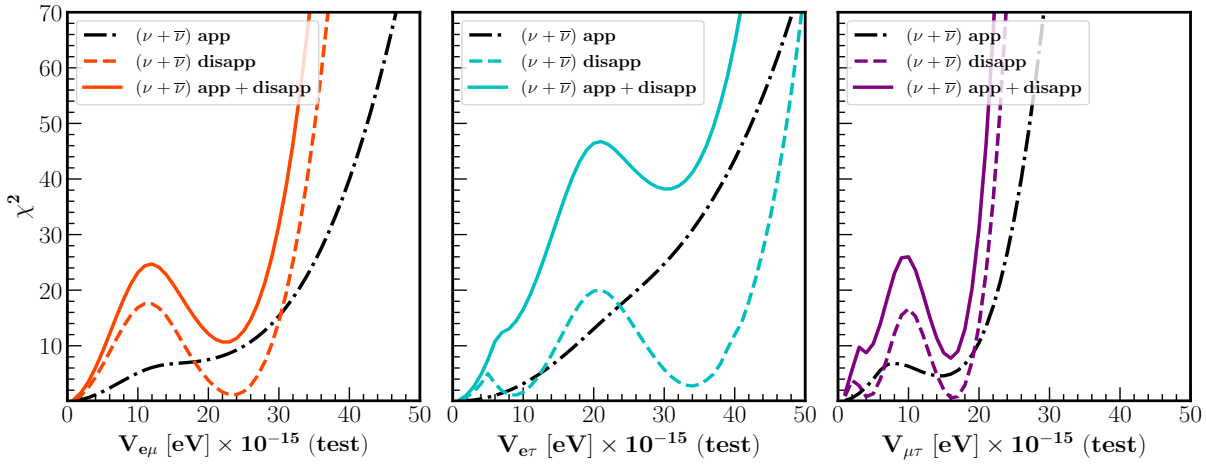



FIG. 4: Sensitivity limits on LRF parameters $V_{e\mu}$, $V_{e\tau}$, and $V_{\mu\tau}$ in appearance, disappearance and in combined case for P2SO experiment, for $\delta_{\rm CP}(\text{true}) = 232°$.

### B. Sensitivity limits on the LRF parameters

In this section, we study the capability of P2SO and T2HKK to put limits on the LRF parameters. In Fig. 3, we show the bounds on the LRF parameters corresponding to $e\mu$, $e\tau$, and $\mu\tau$ sectors from P2SO and T2HKK experiments. This has been obtained by taking the standard scenario in the true spectrum and the LRF scenario in the test spectrum for calculating the $\chi^2$.

Left, middle, and right panels are for $V_{e\mu}$, $V_{e\tau}$, and $V_{\mu\tau}$, respectively. In each panel solid curve represents the sensitivity limit for P2SO and dashed curve is for T2HKK experiment.

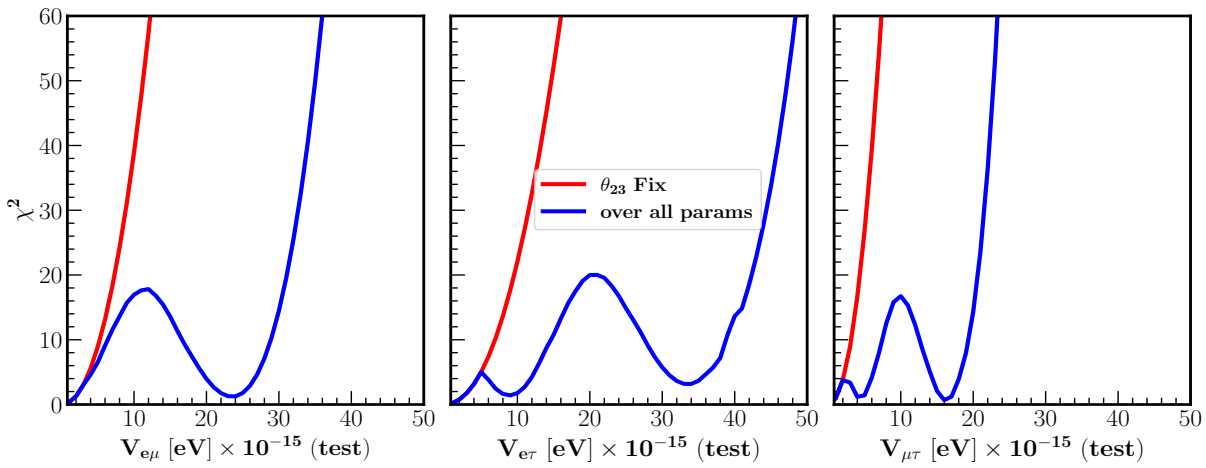



FIG. 5: In each of the three panels, the red curve illustrates the sensitivity limits of LRF parameters while maintaining $\theta_{23}$ value fixed at both true and test values. Conversely, the blue curve depicts the sensitivity bounds of LRF parameters when the true values of oscillation parameters, as specified in Tab. I, are retained, and marginalization is performed across all parameters.

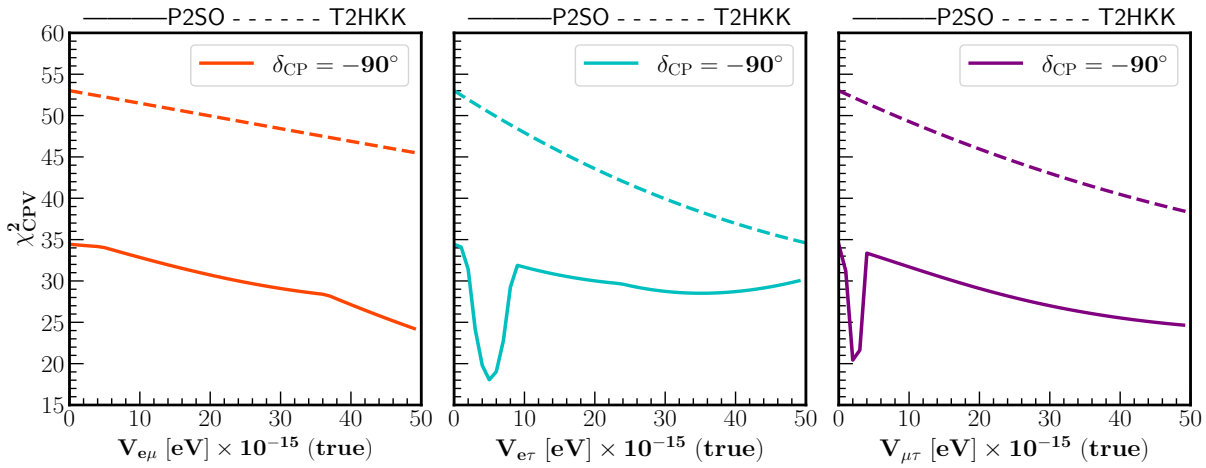



FIG. 6: CPV sensitivities as a function of $V_{\alpha\beta}$ for different true value of $\delta_{CP} = -90°$ for P2SO and T2HKK experiment.

More constrained bounds are obtained from P2SO experiment as compared to T2HKK for all the LRF parameters due to the longer baseline and higher statistics of P2SO experiment. Sensitivity limits on the LRF parameters are shown in Tab. II at 90% C.L. from P2SO and T2HKK. In that table, we also list the bounds from Super-Kamiokande (SK), INO, DUNE and T2HK in order to compare our result with the bounds form other experiments. From the table, we see that P2SO experiment can put the strongest bounds on the LRF parameters.

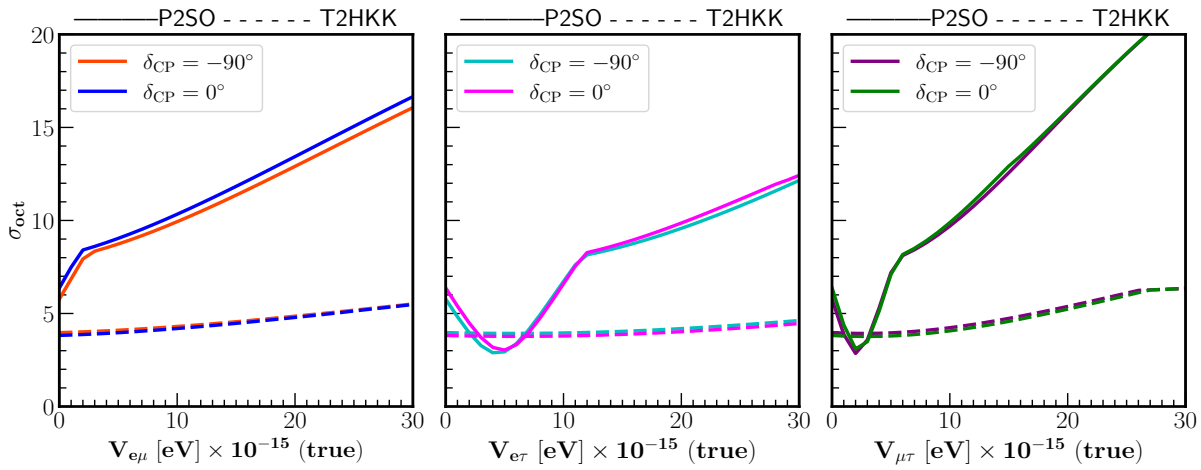



FIG. 7: Octant sensitivity as a function of $V_{\alpha\beta}$ for two different true values of $\delta_{CP}$ for P2SO and T2HKK experiments. We have considered the parameter $\theta_{23}$ to be LO for the analysis.

| **LRF Potential** $[eV]$ | **SK** [17] | **INO** [22] | **DUNE** [23] | **T2HK** [23] | **P2SO** (This work) | **T2HKK** (This work) |
|---|---|---|---|---|---|---|
| $V_{e\mu}(\times 10^{-14})$ | 71.5 | 1.56 | 1.46 | 3.45 | 0.23 | 2.40 |
| $V_{e\tau}(\times 10^{-14})$ | 83.2 | 1.56 | 1.03 | 3.43 | 0.23 | 2.15 |
| $V_{\mu\tau}(\times 10^{-14})$ | - | - | 0.67 | 1.84 | 0.13 | 1.5 |

TABLE II: Sensitivity limits at 90% C.L. on LRF parameters from several experiments.

The bounds obtained from T2HKK are better than T2HK but poor than DUNE. The bounds obtained from SK are somewhat poor and the future atmospheric experiment INO is expected to put stronger bounds on the LRF parameters than T2HKK.

In Fig. 3, we see a dip in the sensitivity curves for P2SO in all three panels. However, these dips are not present in the T2HKK curves. To understand this, in Fig. 4, we show the contribution from the individual probability channels for the P2SO experiment.

In each panel solid curve is for the combination of appearance and disappearance events. Dashed (dash-dotted) curves in each panel are the sensitivity considering only disappearance (appearance) events. It is clear from the panels that, the dip is mainly the contribution of disappearance neutrino events.

To understand the origin of the dip coming from the disappearance channel, we tried to look at the effect of marginalization. In Fig. 5, we show how the sensitivity of the disappearance channel changes with respect to the marginalization of the parameter $\theta_{23}$. From the panels, we see that when $\theta_{23}$ is kept fixed in the test, the dip disappears in all three cases. This implies the presence of octant degeneracy (i.e., $\theta_{23} \to 90° - \theta_{23}$) in the disappearance channel in the case of P2SO. As this degeneracy is not present for T2HKK, we do not see any dip for this experiment.

### C. Effects of LRF parameters on CP violation, Octant and mass ordering sensitivities

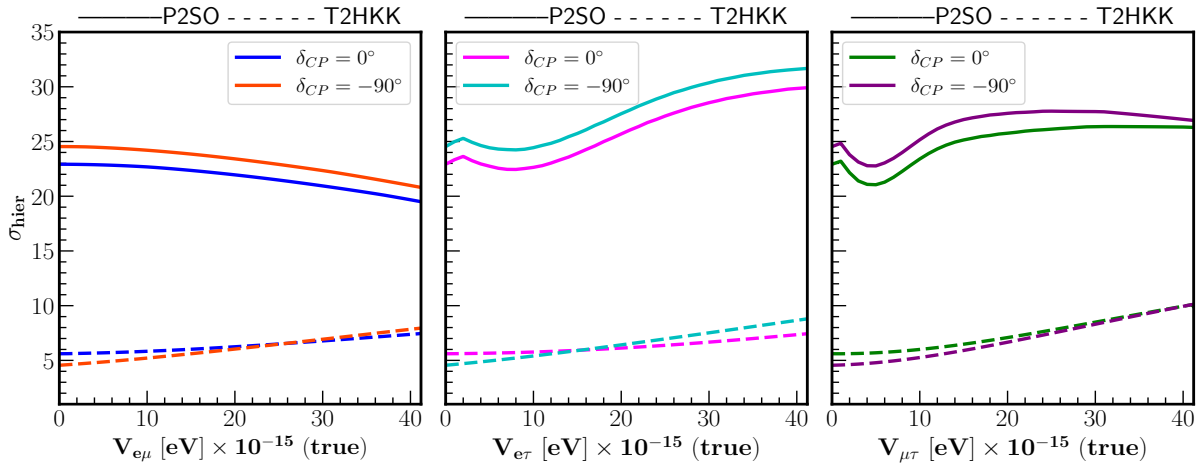



FIG. 8: Variation of neutrino mass ordering sensitivities with respect to $V_{\alpha\beta}$ for both T2HKK and P2SO experiments.

In this section, we discuss how LRF can affect the sensitivity of P2SO and T2HKK in determining the unknowns in the neutrino oscillation. We do this by keeping $V_{\alpha\beta}$ fixed in both true and test spectrum of the $\chi^2$. At present, there are three unknowns in the neutrino oscillations which are: (i) true ordering of the neutrino masses, which can be either normal or inverted, (ii) true octant of $\theta_{23}$ which can be lower or higher and (iii) the value of $\delta_{\rm CP}$ which can lead to CP violation (CPV) in the neutrino sector. Let us begin with the CPV sensitivity. Fig. 6 shows the CPV sensitivity as a function of $V_{\alpha\beta}$. Left/middle/right panel shows the sensitivity for $V_{e\mu}/V_{e\tau}/V_{\mu\tau}$. We have obtained CPV sensitivities by excluding CP conserving values of $\delta_{\rm CP}$ for true value of $\delta_{\rm CP}$ as $-90°$. Solid curves represent the sensitivities for P2SO while dashed curves are for T2HKK experiment. CPV sensitivities of T2HKK are higher compared to P2SO experiment for all the LRF parameters as well as SI. In general, we see that as $V_{\alpha\beta}$ increases, the CPV sensitivity decreases. Additionally, we observe kinks in the $V_{e\tau}$ and $V_{\mu\tau}$ curves. We have checked that these kinks appear due to the degeneracy associated with the parameter $\theta_{23}$. If we keep $\theta_{23}$ fixed in our calculation, then these kinks disappear.

In Fig. 7, we show the same as Fig. 6 but for octant sensitivities. We obtain the octant sensitivity by considering the true $\theta_{23}$ in lower octant (LO) and varying the test $\theta_{23}$ in higher octant (HO). We show this for two values of $\delta_{\rm CP}$ i.e., $0°$ and $-90°$. From the panels we see that, in general, octant sensitivity is higher in P2SO as compared to T2HKK. Further, we see that sensitivity is almost similar for both the values of $\delta_{\rm CP}$. For T2HKK, we see that octant sensitivity continuously increases as $V_{\alpha\beta}$ increases. Whereas for P2SO, octant

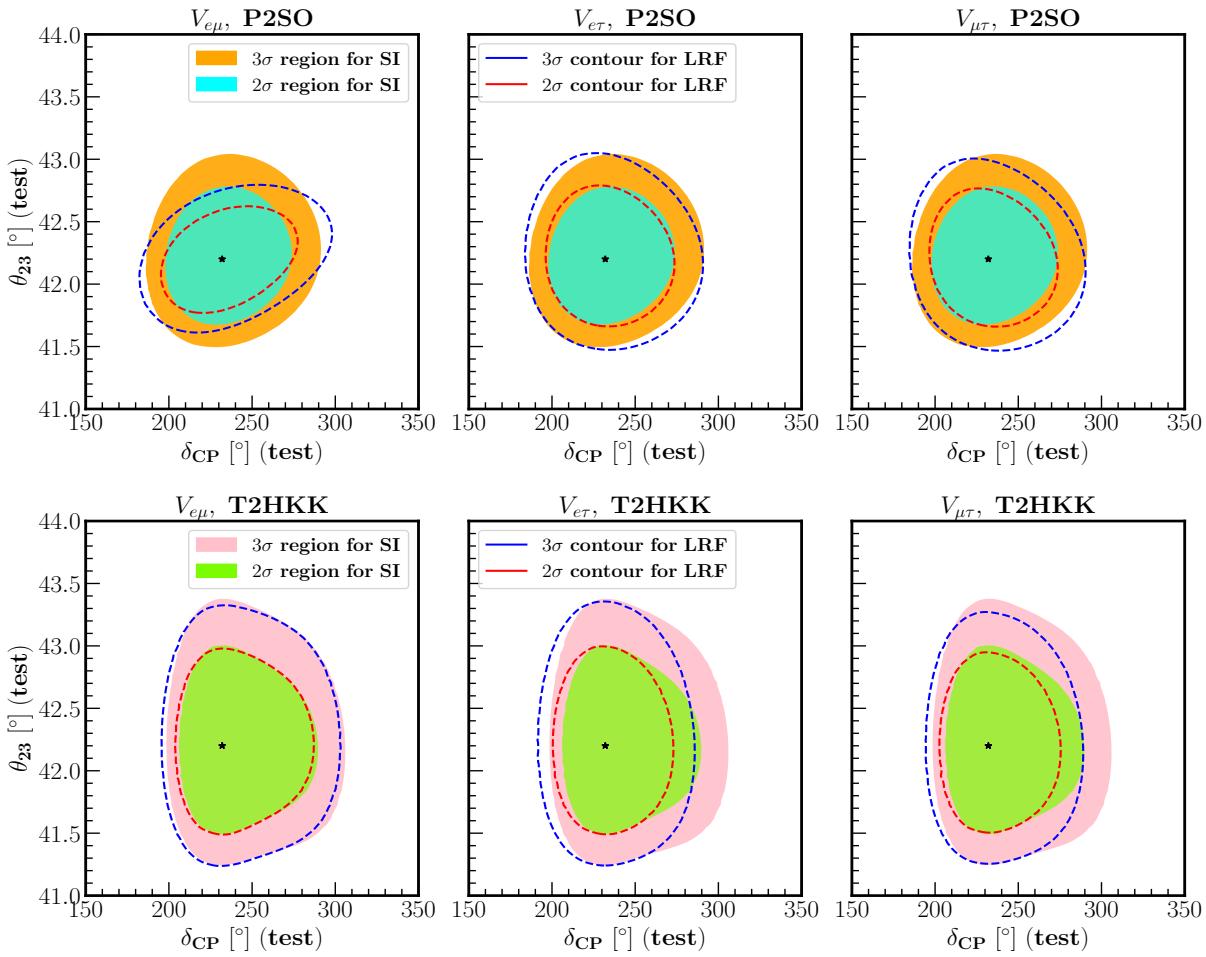



FIG. 9: Allowed parameter space between $\theta_{23}$ (test) $\delta_{\rm CP}$ (test) with $V_{\alpha\beta}$ fixed to their upper bound values at 5 $\sigma$ C.L. for P2SO and T2HKK.

sensitivity increases continuously for $V_{e\mu}$ and for the other two $V_{e\tau}$ and $V_{\mu\tau}$, the sensitivity curve has a dip. We have checked that this dip occurs because the $\chi^2$ minimum appears with different values of $\theta_{23}$ for different $V_{\alpha\beta}$. The change of sensitivity with respect to $V_{\alpha\beta}$ is much higher in P2SO as compared to T2HKK. Next, we study the effect of LRF parameters on the neutrino mass ordering sensitivities. Mass ordering sensitivity signifies the ability of the experiment to determine the true ordering of the neutrino masses. Fig. 8 shows the mass ordering sensitivities as a function of $V_{\alpha\beta}$ for two different true values of $\delta_{\rm CP}$ i.e., 0° and −90°. The labelling of the panels is the same as Figs. 6 and 7. Here also, we see that the sensitivity of P2SO is higher than T2HKK and sensitivities are similar for both the values of $\delta_{\rm CP}$. However, it is interesting to see that for $V_{e\mu}$, the sensitivity increases for T2HKK and decreases for P2SO as $V_{e\mu}$ increases. For the other two $V_{\alpha\beta}$, sensitivity increases for both the experiments as $V_{\alpha\beta}$ increases. We also observe a kink in P2SO for $V_{\mu\tau}$ and $V_{e\tau}$. We have verified that these kinks are due to the large backgrounds of the P2SO experiment [44].

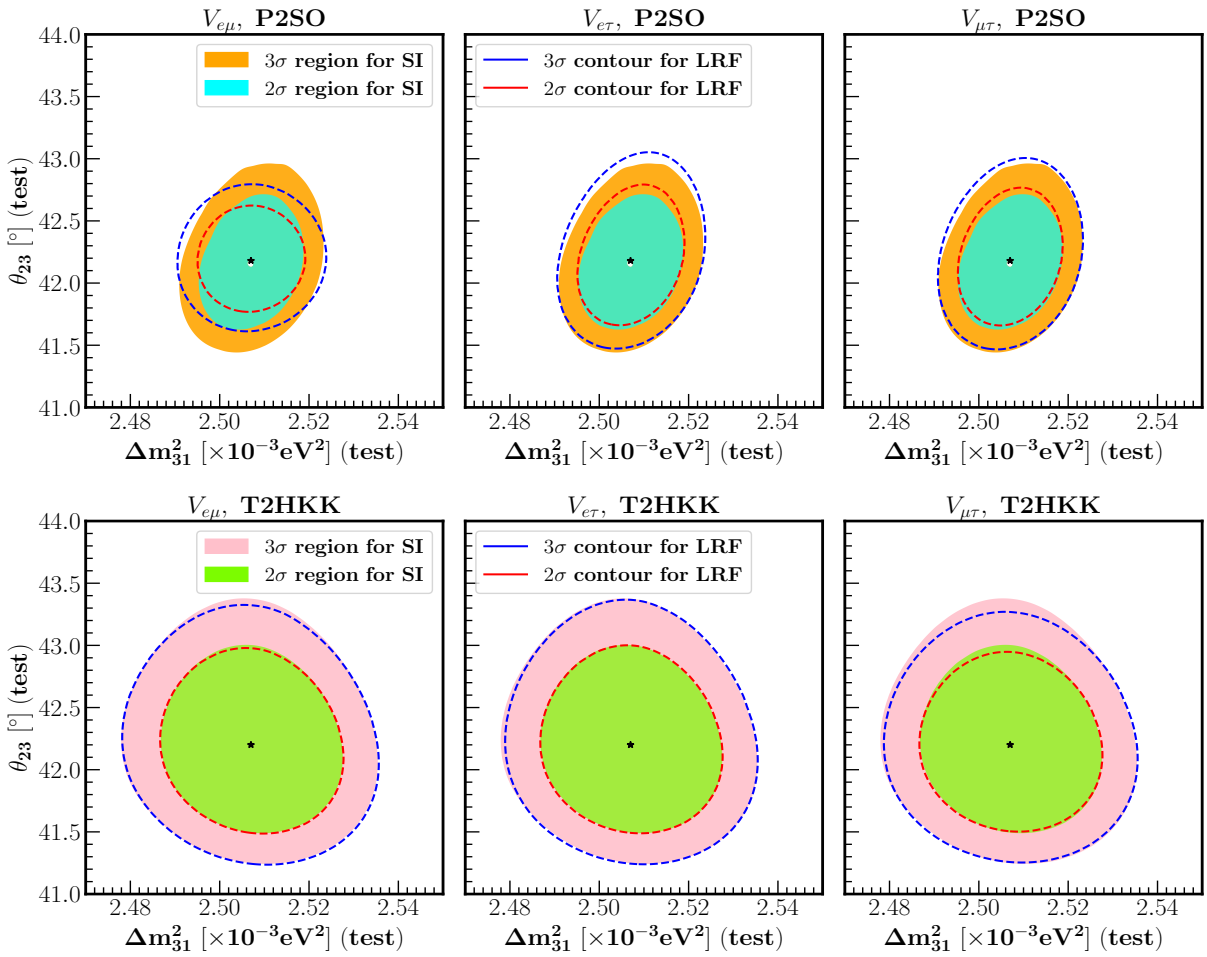



FIG. 10: Allowed parameter space between $\theta_{23}$ (test) $\Delta m^2_{31}$ (test) with $V_{\alpha\beta}$ fixed to their upper bound values at 5 $\sigma$ C.L. for P2SO and T2HKK.

### D. Effect of LRF parameters in the precision of $\delta_{\rm CP}$, $\Delta m^2_{31}$ and $\theta_{23}$

In this section, we study the effect of the LRF parameters in the precision measurements of $\delta_{\rm CP}$, $\Delta m^2_{31}$ and $\theta_{23}$. To show this, we have kept $V_{\alpha\beta}$ fixed in both true and test spectrum of the $\chi^2$ using their $5\sigma$ upper bound values. Then, taking the current true values of $\delta_{\rm CP}$, $\Delta m^2_{31}$ and $\theta_{23}$, we show 2-D allowed regions of these parameters at $2\sigma$ C.L. and $3\sigma$ C.L.

Fig. 9 shows the allowed regions in $\delta_{\rm CP} - \theta_{23}$ plane. The top row is for P2SO and the bottom row is for T2HKK. In each row, left, middle and right panels are for $V_{e\mu}$, $V_{e\tau}$ and $V_{\mu\tau}$ respectively. In each panel, dashed red (blue) curve is $2\sigma$ ($3\sigma$) contours in the presence of LRF and shaded regions are representing the SI in respective experiments. Asterisk symbol in each panel represents the true values of corresponding oscillation parameters. From these panels, we note certain changes in the precision of $\delta_{\rm CP}$ and $\theta_{23}$ in presence of LRF. For T2HKK, the precision of $\theta_{23}$, almost remains same as in the SI case of all $V_{\alpha\beta}$. Whereas for $\delta_{\rm CP}$, the precision improves as compared to SI case for $V_{e\tau}$ and $V_{\mu\tau}$. For P2SO, the precision

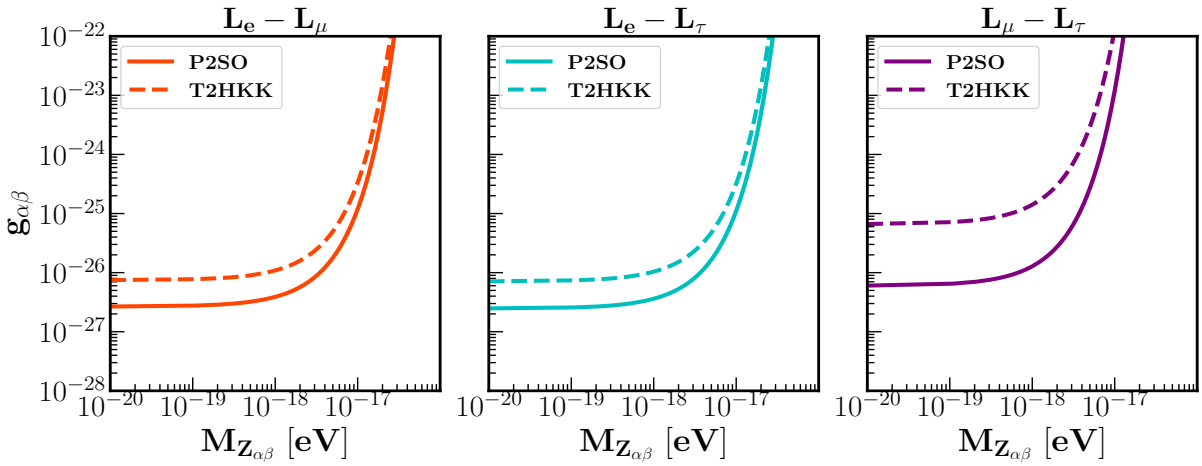



FIG. 11: Plots portraying the allowed range for gauge coupling vs mass of gauge boson for the long range potential in all three cases, is at $2\sigma$ C.L. for both P2SO and T2HKK. For more details please refer to the Sec. VII D

| Model / Experiment | $L_e - L_\mu$ | $L_e - L_\tau$ | $L_\mu - L_\tau$ |
|---|---|---|---|
| P2SO (This work) | $2.66 \times 10^{-27}$ | $2.48 \times 10^{-27}$ | $6.03 \times 10^{-27}$ |
| T2HKK (This work) | $7.47 \times 10^{-27}$ | $7.12 \times 10^{-27}$ | $6.637 \times 10^{-26}$ |
| T2HK [23] | $1.30 \times 10^{-26}$ | $1.24 \times 10^{-26}$ | $4.31 \times 10^{-26}$ |
| DUNE [23] | $8.55 \times 10^{-27}$ | $7.03 \times 10^{-27}$ | $2.59 \times 10^{-26}$ |

TABLE III: Projected upper bound on $g_{\alpha\beta}$ from various long-baseline experiments.

to both $\delta_{\rm CP}$ and $\theta_{23}$ are almost similar to the SI case for $V_{e\tau}$ and $V_{\mu\tau}$. For $V_{e\mu}$, the precision of $\theta_{23}$ is better than the SI and the precision of $\delta_{\rm CP}$ is slightly poor than the SI scenario.

To show the effect of LRF parameter on the precision of $\Delta m^2_{31}$, we adopt the same approach for generating the allowed parameter space in $\Delta m^2_{31} - \theta_{23}$ plane. The results are represented in Fig. 10. From the panels, we see that the precision of $\Delta m^2_{31}$ remains almost unaltered as compared to SI scenario in the presence of LRF.

### E. Sensitivity limits on $M_{Z_{\alpha\beta}}$ and $g_{\alpha\beta}$

Neutrino oscillation experiments also give us opportunity to put bounds on the mass of the new gauge boson $M_{Z_{\alpha\beta}}$ and the new gauge coupling $g_{\alpha\beta}$. In Fig. 11, we present the same at $2\sigma$ C.L., corresponding to a long-range force experienced by the neutrinos at Earth due to the matter density of the Sun, using Eqs. (2) and (3). In each panel, the dashed curve is for T2HKK while the solid curve portrays the allowed region for P2SO. From the

panels, we see that the bounds obtained for P2SO are better than T2HKK and among the three symmetries, the most stringent bound comes for $L_e - L_\tau$. From these plots, we find the upper bound for the coupling of ultra light mediators with mass $M_{Z_{\alpha\beta}} \lesssim 10^{-17}$ eV to be $g_{\alpha\beta} \lesssim 10^{-26}$ for $L_\alpha - L_\beta$ models. Earlier bounds on $g_{\alpha\beta}$ using Superkamiokande data are: $g_{e\mu} < 8.32 \times 10^{-26}$ and $g_{e\tau} < 8.97 \times 10^{-26}$ at 90% confidence level [17]. Using solar and reactor data from KamLAND, the bounds obtained in Ref. [18] are $g_{e\mu} < 2.06 \times 10^{-26}$ and $g_{e\tau} < 1.77 \times 10^{-26}$, at $3\sigma$, assuming $\theta_{13} = 0°$. The projected bounds on $g_{\alpha\beta}$ for the upcoming DUNE and T2HK experiments are obtained in [23]. For comparison, we list the corresponding excluded ranges for the $Z_{\alpha\beta}$ coupling and mass obtained from different long-baseline experiments in Tab. III. From the table we see that the bounds obtained from P2SO are better as compared to other future long-baseline neutrino experiments.

## VIII. SUMMARY AND CONCLUSION

In this paper, we have studied the sensitivity of the two upcoming long-baseline experiments P2SO and T2HKK to the long range force. The proposed P2SO experiment will study the oscillations of the neutrinos originating from protvino and to be detected at the Super-ORCA detector. Whereas the T2HKK experiment, which is an alternative option of the T2HK experiment, will have a dual baseline configuration i.e., one detector will be located in Japan and the another detector will be located in Korea.

When the Standard Model is extended with $L_e - L_\mu$, $L_e - L_\tau$ and $L_\mu - L_\tau$ type $U(1)$ gauge symmetries, it gives rise to a new gauge boson $Z'$ and a new coupling constant $g'$. If the new mediator is inordinately lightweight, it gives rise to a long range potential which can affect the propagation of the neutrinos in the matter. The nature of this new long range potential $V_{\alpha\beta}$ would be different depending on the symmetry i.e., $L_e - L_\mu$, $L_e - L_\tau$ and $L_\mu - L_\tau$.

In the first part of our paper, we have put forward a prescription on how to calculate the neutrino oscillation probabilities in the presence of LRF and studied how this new potential affects the neutrino oscillation probabilities in P2SO and T2HKK. In our study, we find that the effect of the LRF parameters $V_{e\tau}$ and $V_{\mu\tau}$ is more in the appearance probability in neutrino case, whereas in the antineutrino appearance probabilities, only significant change occurs due to $V_{e\mu}$. In the disappearance channel, all the three LRF parameters significantly affect the probabilities in P2SO in both neutrinos and antineutrinos whereas for T2HKK, the effect of $V_{e\mu}$ and $V_{\mu\tau}$ is higher in neutrinos and the effect of $V_{\mu\tau}$ is higher in antineutrinos. We have also noticed some significant differences between the T2HKK and the P2SO probabilities. This is because the effective mixing parameters behave differently at these relevant energies and baselines.

Next, we moved on to study the capability of the above mentioned experiments to constrain the LRF parameters. In our analysis, we find that sensitivity of P2SO is better than T2HKK due to the former's longer baseline and higher statistics. However, the sensitivity

of P2SO suffers from octant degeneracy in the disappearance channel at higher values of $V_{\alpha\beta}$. While comparing our results with the other experiments, we find that P2SO can give the strongest bound on the LRF parameters among the current and future neutrino oscillation experiments. The bounds obtained from T2HKK are better than T2HK but poor than DUNE. The bounds obtained from SK are somewhat poor, and the future atmospheric experiment INO is expected to put stronger bounds on the LRF parameters than T2HKK.

Next, we studied the effect of LRF in the measurement of standard oscillation parameters. Regarding CP violation sensitivity, we find that the sensitivity is a decreasing function of $V_{\alpha\beta}$, whereas for octant sensitivity it is an increasing function of $V_{\alpha\beta}$ for both P2SO and T2HKK. Regarding octant sensitivity, the change of sensitivity with respect to $V_{\alpha\beta}$ is much higher in P2SO as compared to T2HKK. For mass ordering, the sensitivity is an increasing function of $V_{e\mu}$ for T2HKK but it a decreasing function for P2SO. For the other two $V_{\alpha\beta}$, mass ordering sensitivity is an increasing function with respect to $V_{\alpha\beta}$ for both the experiments. Additionally, we noticed that the sensitivity of P2SO is affected by the degeneracy related to $\theta_{23}$ for CP violation measurement and octant measurement and it is affected by large background for the determination of mass ordering. Regarding the precision measurement of $\delta_{\rm CP}$, $\Delta m^2_{31}$ and $\theta_{23}$ we find that, LRF does not affect the precision of $\theta_{23}$ in T2HKK whereas for P2SO, the precision to both $\delta_{\rm CP}$ and $\theta_{23}$ are unaltered except for $V_{e\mu}$. For $V_{e\mu}$, the precision of $\theta_{23}$ is better than the SI and the precision of $\delta_{\rm CP}$ is slightly poor than the SI scenario. Additionally, for T2HKK the precision of $\delta_{\rm CP}$ improves as compared to SI case for $V_{e\tau}$ and $V_{\mu\tau}$. The precision of $\Delta m^2_{31}$ remains unchanged in presence of LRF in both the experiments.

Finally, we have computed the bounds on the mass of the new gauge boson and its coupling strength corresponding to the LRF experienced by the neutrinos at Earth due to the matter density of the Sun. Our results show that the bounds obtained for P2SO are better than T2HKK and among the three symmetries, the most stringent bound comes for $L_e - L_\tau$. While comparing the bounds with the other future long-baseline experiments, we find that the best bound on these parameters comes from P2SO.

## ACKNOWLEDGMENTS

PM would like to thank Prime Minister's Research Fellows (PMRF) scheme for its financial support. This work has been in part funded by the Ministry of Science and Education of the Republic of Croatia grant No. KK.01.1.1.01.0001. SKP and RM acknowledge University of Hyderabad IoE project grant no. RC1-20-012. We gratefully acknowledge the use of the

CMSD HPC facility of Univ. of Hyderabad to carry out the computational work.

---


[1] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, JHEP **09**, 178 (2020), arXiv:2007.14792 [hep-ph].

[2] L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978).

[3] V. D. Barger, K. Whisnant, S. Pakvasa, and R. J. N. Phillips, Phys. Rev. D **22**, 2718 (1980).

[4] J. A. Grifols, E. Masso, and S. Peris, Astropart. Phys. **2**, 161 (1994).

[5] J. A. Grifols, E. Masso, and R. Toldra, Phys. Lett. B **389**, 563 (1996), arXiv:hep-ph/9606377.

[6] J. A. Gonzalez-Garcia:2006vicols and E. Masso, Phys. Lett. B **579**, 123 (2004), arXiv:hep-ph/0311141.

[7] R. Foot, Mod. Phys. Lett. A **6**, 527 (1991).

[8] X. G. He, G. C. Joshi, H. Lew, and R. R. Volkas, Phys. Rev. D **43**, 22 (1991).

[9] X.-G. He, G. C. Joshi, H. Lew, and R. R. Volkas, Phys. Rev. D **44**, 2118 (1991).

[10] R. Foot, X. G. He, H. Lew, and R. R. Volkas, Phys. Rev. D **50**, 4571 (1994), arXiv:hep-ph/9401250.

[11] Y. Farzan and M. Tortola, Front. in Phys. **6**, 10 (2018), arXiv:1710.09360 [hep-ph].

[12] A. Y. Smirnov and X.-J. Xu, JHEP **12**, 046 (2019), arXiv:1909.07505 [hep-ph].

[13] J. G. Williams, X. X. Newhall, and J. O. Dickey, Phys. Rev. D **53**, 6730 (1996).

[14] A. D. Dolgov, Phys. Rept. **320**, 1 (1999).

[15] E. G. Adelberger, B. R. Heckel, and A. E. Nelson, Ann. Rev. Nucl. Part. Sci. **53**, 77 (2003), arXiv:hep-ph/0307284.

[16] J. G. Williams, S. G. Turyshev, and D. H. Boggs, Phys. Rev. Lett. **93**, 261101 (2004), arXiv:gr-qc/0411113.

[17] A. S. Joshipura and S. Mohanty, Phys. Lett. B **584**, 103 (2004), arXiv:hep-ph/0310210.

[18] A. Bandyopadhyay, A. Dighe, and A. S. Joshipura, Phys. Rev. D **75**, 093005 (2007), arXiv:hep-ph/0610263.

[19] M. C. Gonzalez-Garcia, P. C. de Holanda, E. Masso, and R. Zukanovich Funchal, JCAP **01**, 005 (2007), arXiv:hep-ph/0609094.

[20] A. Samanta, JCAP **09**, 010 (2011), arXiv:1001.5344 [hep-ph].

[21] S. S. Chatterjee, A. Dasgupta, and S. K. Agarwalla, JHEP **12**, 167 (2015), arXiv:1509.03517 [hep-ph].

[22] A. Khatun, T. Thakore, and S. Kumar Agarwalla, JHEP **04**, 023 (2018), arXiv:1801.00949 [hep-ph].

[23] M. Singh, M. Bustamante, and S. K. Agarwalla, JHEP **08**, 101 (2023), arXiv:2305.05184 [hep-ph].

[24] S. K. Agarwalla, M. Bustamante, S. Das, and A. Narang, JHEP **08**, 113 (2023), arXiv:2305.03675 [hep-ph].

[25] A. V. Akindinov *et al.*, Eur. Phys. J. C **79**, 758 (2019), arXiv:1902.06083 [physics.ins-det].
[26] K. Abe *et al.* (Hyper-Kamiokande), PTEP **2018**, 063C01 (2018), arXiv:1611.06118 [hep-ex].
[27] K. Abe *et al.* (Hyper-Kamiokande), (2018), arXiv:1805.04163 [physics.ins-det].
[28] B. Abi *et al.* (DUNE), (2020), arXiv:2002.03005 [hep-ex].
[29] A. M. Baldini *et al.* (MEG), Eur. Phys. J. C **76**, 434 (2016), arXiv:1605.05081 [hep-ex].
[30] U. Bellgardt *et al.* (SINDRUM), Nucl. Phys. B **299**, 1 (1988).
[31] A. Abdesselam *et al.* (Belle), JHEP **10**, 19 (2021), arXiv:2103.12994 [hep-ex].
[32] J. N. Bahcall, *Neutrino astrophysics* (Cambridge University Press, 1989).
[33] J. Heeck and W. Rodejohann, J. Phys. G **38**, 085005 (2011), arXiv:1007.2655 [hep-ph].
[34] B. Holdom, Phys. Lett. B **166**, 196 (1986).
[35] S. K. Agarwalla, S. Das, M. Masud, and P. Swain, JHEP **11**, 094 (2021), arXiv:2103.13431 [hep-ph].
[36] S. K. Agarwalla, Y. Kao, and T. Takeuchi, JHEP **04**, 047 (2014), arXiv:1302.6773 [hep-ph].
[37] D. K. Singha, M. Ghosh, R. Majhi, and R. Mohanta, Phys. Rev. D **107**, 075039 (2023), arXiv:2211.01816 [hep-ph].
[38] R. Majhi, D. K. Singha, M. Ghosh, and R. Mohanta, Phys. Rev. D **107**, 075036 (2023), arXiv:2212.07244 [hep-ph].
[39] D. K. Singha, R. Majhi, L. Panda, M. Ghosh, and R. Mohanta, (2023), arXiv:2308.10789 [hep-ph].
[40] P. Huber, M. Lindner, and W. Winter, Comput. Phys. Commun. **167**, 195 (2005), arXiv:hep-ph/0407333.
[41] P. Huber, J. Kopp, M. Lindner, M. Rolinec, and W. Winter, Comput. Phys. Commun. **177**, 432 (2007), arXiv:hep-ph/0701187.
[42] G. L. Fogli, E. Lisi, A. Marrone, D. Montanino, and A. Palazzo, Phys. Rev. D **66**, 053010 (2002), arXiv:hep-ph/0206162.
[43] P. Huber, M. Lindner, and W. Winter, Nucl. Phys. B **645**, 3 (2002), arXiv:hep-ph/0204352.
[44] D. K. Singha, M. Ghosh, R. Majhi, and R. Mohanta, JHEP **05**, 117 (2022), arXiv:2112.04876 [hep-ph].

# Photon-Odderon interference in exclusive $\chi_c$ charmonium production at the Electron-Ion Collider

Sanjin Benić,[1] Adrian Dumitru,[2,3] Abhiram Kaushik,[4] Leszek Motyka,[5] and Tomasz Stebel[5]

[1] *Department of Physics, Faculty of Science, University of Zagreb, Bijenička c. 32, 10000 Zagreb, Croatia*
[2] *Department of Natural Sciences, Baruch College, CUNY, 17 Lexington Avenue, New York, NY 10010, USA*
[3] *The Graduate School and University Center, The City University of New York, 365 Fifth Avenue, New York, NY 10016, USA*
[4] *Centre for Informatics and Computing, Rudjer Bošković Institute, HR-10002 Zagreb, Croatia*
[5] *Jagiellonian University, Institute of Theoretical Physics, Łojasiewicza 11, 30-348 Kraków, Poland*

(Dated: March 1, 2024)

Exclusive $C = +1$ scalar, axial-vector, and tensor quarkonium production in high-energy electron-proton scattering requires a $C$-odd $t$-channel exchange of a photon or a three gluon ladder. We derive the expressions for the corresponding amplitudes. The relative phase of the photon vs. three gluon exchange amplitudes is determined by the sign of the light-front matrix element of the eikonal color current operator $d^{abc}J^{+a}J^{+b}J^{+c}$ at moderate $x$, and is not affected by small-$x$ QCD evolution. Model calculations predict constructive interference, which is particularly strong for momentum transfer $|t| \sim 1$ GeV$^2$ where the cross section for $\chi_{cJ}$ production exceeds that for pure photon exchange by up to a factor of 4. We find that exclusive $\chi_{cJ}$ electroproduction at the Electron-Ion Collider should occur with well measurable rates and measurements of these processes should allow to find an evidence of the perturbative Odderon exchange. We also compute the total electroproduction cross section as a function of energy and provide first estimates of the number of $\chi_{cJ}$ events per month at the Electron-Ion Collider design luminosity.

arXiv:2402.19134v1 [hep-ph] 29 Feb 2024

## CONTENTS

I. Introduction 2

II. The production amplitude of $C$-even quarkonia 3
A. Light-cone wave function of $C$-even quarkonia 5
B. Final expressions for the amplitudes 7
C. The $\gamma^* p \to \mathcal{H} p$ cross section 9
D. Odderon amplitude and its evolution with $x$ 10

III. The Primakoff contribution 11
A. $|t| \to 0$ limit 12
B. Adding the Pauli form factor 13

IV. Boosted Gaussian model for the $\chi_{cJ}$ wave functions 13

V. Numerical results 15

VI. Summary and Conclusions 18

Acknowledgments 19

A. Computation of the light-cone wave functions 19

B. Derivation of the amplitudes for axial and tensor quarkonia 20

C. The Primakoff contribution in specific kinematic limits 22
1. The proof of formula (40) 22
2. The $|t| \to 0$ limit for axial vector quarkonia and its connection to the Landau-Yang theorem 23
3. The NRQCD limit of the Primakoff cross sections 23

D. The relative sign of the eikonal photon and Odderon exchange amplitudes 24

E. BJKP–BLV Odderon 27

References 27

## I. INTRODUCTION

A basic prediction of QCD, related to the non-zero cubic Casimir invariant, is that there exists a contribution to high-energy scattering amplitudes from a $C$-conjugation odd exchange driven mostly by gluons, with a weakly energy dependent cross section. The possible existence of such an exchange, called the Odderon, was predicted 50 years ago from general principles of quantum field theory [1]. If the scattering involves hadrons of size much smaller than the QCD confinement scale, or high momentum transfer, then the composition of the exchanged "state" can be understood from perturbative QCD. At leading order in perturbation theory, it corresponds to an exchange of three gluons which is symmetric under permutations of their colors. At high energies the three gluon exchange is dressed by higher order QCD corrections enhanced by logarithms of energy that have to be resummed into a gluon ladder. It accounts for interactions between the exchanged $t$-channel gluons, and for virtual corrections. The resulting "reggeized" three gluon exchange is the hard Odderon. In the weak field limit the corresponding Bartels-Jaroszewicz-Kwieciński-Praszałowicz (BJKP) linear QCD evolution equation has been obtained long ago [2–4], and it amounts to the reggeization of the exchanged gluons by iteration of the Balitsky-Fadin-Kuraev-Lipatov (BFKL) interaction kernel [5–7] for each pair of gluons in the ladder. The NLO corrections to the BJKP equation were found in Ref. [8]. At asymptotically high energies the leading solution to the BJKP equation corresponds to a configuration of gluons in which two reggeized gluons are combined together [9] – the Bartels–Lipatov–Vacca (BLV) solution. The BLV solution couples to color dipoles so it is particularly important in $C$-even quarkonia (or meson) photo- and electroproduction [9]. The intercept of this solution equals one, in the leading logarithmic approximation [9], and it was argued that it remains at this value to all orders in the perturbative expansion [10].

The non-linear Odderon evolution equation for dipole–proton scattering was established in Refs. [11, 12]. This equation typically produces solutions where the Odderon amplitude decreases with energy [13–16], as will be confirmed also in our present work. The seed for this evolution to small $x$ is provided by the cubic Casimir $d^{abc}J^{+a}J^{+b}J^{+c}$ in the effective action describing color current ($J^{+a}$) fluctuations in the proton at moderately small $x$ [17, 18], which allows the $C$-odd gluon ladder to couple to the proton. The matrix element[1] of this operator could also be evaluated directly from the light-cone wave function of the proton [22, 23]. A key point for the present analysis is that the Odderon evolution equation to small $x$ *does not alter the sign* of the amplitude. Hence, the interference pattern with the electromagnetic amplitude due to single photon exchange is determined by the sign of the matrix element of the $d^{abc}J^{+a}J^{+b}J^{+c}$ operator at the initial, moderately small $x$.

The TOTEM collaboration at the CERN-LHC has measured the differential cross section for $pp$ elastic scattering at $\sqrt{s} = 2.76$ GeV [24]. They observe a significant difference to the cross section for $p\bar{p}$ scattering at $\sqrt{s} = 1.96$ GeV measured by D0 [25]. Assuming that the difference in energy is negligible they conclude that these results provide evidence for the exchange of a color singlet gluonic state with odd C-parity, the Odderon. However, these measurements involve the scattering of hadrons with a size of order the QCD confinement scale as well as low momentum transfers, $|t| \ll 1$ GeV$^2$. Hence, an analysis of the nature of the $t$-channel exchange from perturbative QCD would not appear reliable.

Exclusive production of pseudo-scalar $\eta_c$ quarkonia in DIS has previously been proposed as a process suitable for the discovery of the hard Odderon exchange in QCD [16, 26–32]. In practice, the detection of this process is difficult as the $\eta_c$ has small branching ratios to all relevant decay channels. For example, the radiative $\eta_c \to \gamma\gamma$ mode has a tiny ($\sim 10^{-4}$) branching ratio[2], while the hadronic $\eta_c \to \rho\rho$ mode has a branching ratio of 1.5%. Also, the cross section for exclusive production of $J/\psi$ is expected to be far greater than that for $\eta_c$, and so most $\eta_c$'s would originate from a $J/\psi \to \eta_c\gamma$ decay[3]. The experimental identification of the soft photon ($M_{J/\psi} = 3.097$ GeV, $M_{\eta_c} = 2.984$ GeV) is difficult [34, 35].

Alternative approaches for the discovery of the hard Odderon include Pomeron–Odderon interference (with a background from Pomeron–photon interference) which would manifest in asymmetries in exclusive production of two charged particles in DIS [36–38]; for the $\gamma^* p \to \chi_c p$ process considered here, it is instead the interference of the

[1] We are here concerned with the off-forward matrix element for non-zero momentum transfer $t$. In the $t \to 0$ limit, instead, the "spin dependent Odderon" [19–21] is associated with a spin flip of the proton in DIS.

[2] We take all branching ratios and particle masses from the "Review of Particle Physics" [33].

[3] This issue could be mitigated by considering $\eta_c(2S)$ instead, since it lies between the $J/\psi$ and the $\psi(2S)$. The branching ratio $\psi(2S) \to \eta_c + \gamma$ is favourably small ($(7 \pm 5) \times 10^{-4}$) but the detection of the $\eta_c(2S)$ is again problematic due to the small branching ratios for many channels.

Primakoff-like amplitude with the QCD Odderon exchange amplitude which will be important, i.e. photon–Odderon interference. Another process which involves a $C$-odd $t$-channel gluonic exchange is exclusive production of a $J/\psi$ in double-diffractive proton-proton scattering [39, 40]. At high-energy proton colliders this requires instrumentation over a large range of rapidity, close to the beams.

The GlueX Collaboration at Jefferson Lab has recently reported the observation of $56.5 \pm 8.2$ exclusive $\chi_{c1}(1P)$ and $12.7 \pm 4.5$ $\chi_{c2}(1P)$ production events in DIS near threshold energy [41]. They noticed a "dramatic difference" in the momentum transfer $t$-distribution of events as compared to $J/\psi$ production in that the cross section for $\chi_c$ production appears to drop off much less rapidly with increasing $|t|$. In other words, the probability that the struck proton remains intact at high momentum transfer is much greater in $\gamma^* p \to \chi_c p$ than in $\gamma^* p \to J/\psi p$ events. This remarkable result illustrates the importance of the underlying QCD dynamics as opposed to naive expectations based solely on the mass of the produced quarkonia. A much harder $t$-dependence for exclusive production of $\eta_c$ vs. $J/\psi$ in eikonal dipole–proton scattering is predicted by simple light-front constituent quark models [32]. Inspired by the GlueX measurement of exclusive $\chi_c$ production in DIS at Jefferson Lab, here we consider the same process but at high energies appropriate for the future Electron-Ion Collider (EIC) [42, 43] where dipole model factorization [44, 45] applies, and where the $t$-channel exchange of a $C$-odd color singlet state dominates.

Within the $\chi_{cJ}$ family ($J = 0,\ 1,\ 2$), $\chi_{c1}$ has the largest branching ratio for the radiative $\chi_{c1} \to J/\psi\gamma$ decay (34.3% as opposed to 1.4% for $\chi_{c0}$ and 19.0% for $\chi_{c2}$). Indeed, this was the detection channel used by the GlueX measurement. Therefore, the production of $\chi_{c1}$ axial-vector and $\chi_{c2}$ tensor quarkonia may prove the most promising channels for discovery of the hard Odderon. This is further corroborated by the fact that for $\chi_{c1}$ and $\chi_{c2}$ the Odderon and photon exchange contributions become comparable at lower $|t|$ than for $\chi_{c0}$, within the acceptance of the proton spectrometer of the EIC design detector [43]. However, for the $\chi_{cJ}$ quarkonia there is again a feed-down channel from exclusive production of $\psi(2S)$ with subsequent decay $\psi(2S) \to \chi_{cJ}\gamma$. Hence, the identification and rejection of such feed down is required.

The paper is organized as follows. In Sec. II we start with the computation of the amplitude and the corresponding $\gamma^* p$ cross section for exclusive production of $C$-even scalar, axial vector and tensor quarkonia. We derive the light-cone wave functions of quarkonia, whereas the amplitude is obtained as an overlap with the photon light-cone wave function. We introduce the Odderon exchange amplitude and discuss its evolution to small-$x$. Sec. III is devoted to the computation of the Primakoff contribution, where we pay special attention to the $t \to 0$ limit. In Sec. IV we perform a numerical fit of the $\chi_{cJ}$ wave functions. The main results are shown in Sec. V, where we numerically compute exclusive $\gamma^* p \to \chi_{cJ} p$ and $ep \to \chi_{cJ} ep$ cross sections and the expected number of events at the EIC. Our findings are summarized in the final Sec. VI. Several Appendices follow where we explain the computational steps leading to some of the results from the main text.

## II. THE PRODUCTION AMPLITUDE OF $C$-EVEN QUARKONIA

We consider the process $\gamma^*(q)p(P) \to p(P')\mathcal{H}(\Delta)$ where $\mathcal{H}$ is the $C$-even quarkonium state. In this paper, we specifically have in mind the case when $\mathcal{H}$ are the $P$-wave charmonia $\chi_{cJ}$ states with $J = 0,\ 1,\ 2$. However our derivations are presented in a general form so that they apply also to other $C = +1$ $q\bar{q}$ bound states such as bottomonia, or even mesons with a light flavor content.

Our computation is performed in the dipole frame where $q = (-Q^2/2q^-, q^-, 0, 0)$, $P = (P^+, 0, 0, 0)$ and the proton mass $M_p \ll \sqrt{(q+P)^2}$ is neglected. We use the convention for the components $(v^+, v^-, v^1, v^2)$ of a four vector $v$ where $v^\pm = (v^0 \pm v^3)/\sqrt{2}$ are the $\pm$ light-cone components. The amplitude is computed using light-cone gauge $A^- = A \cdot n = 0$, that is, $n^\mu = \delta^{\mu +}$. Following a similar procedure as in [16, 32] our starting point is the amplitude

$$\langle \mathcal{M}_{\lambda\bar{\lambda}}(\gamma^* p \to \mathcal{H} p)\rangle = 2q^- N_c \int_{\boldsymbol{r}_\perp \boldsymbol{b}_\perp} \mathrm{e}^{-\mathrm{i}\boldsymbol{\Delta}_\perp \cdot \boldsymbol{b}_\perp} \mathrm{i}\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)\mathcal{A}_{\lambda\bar{\lambda}}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp)\,, \tag{1}$$

with $\lambda$ and $\bar{\lambda}$ the helicities of the incoming photon and the outgoing quarkonium, respectively. For brevity, we will often times write the amplitude simply as $\langle \mathcal{M}_{\lambda\bar{\lambda}}\rangle$ (omitting the round brackets). Here $\mathcal{A}_{\lambda\bar{\lambda}}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp)$ is the reduced amplitude, described physically in terms of the photon ($\Psi^\gamma_{\lambda,h\bar{h}}$) and the quarkonia ($\Psi^{\mathcal{H}}_{\lambda,h\bar{h}}$) wave function overlap. We have[4]

$$\mathcal{A}_{\lambda\bar{\lambda}}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) = \int_z \int_{\boldsymbol{l}_\perp \boldsymbol{l}'_\perp} \sum_{h\bar{h}} \Psi^\gamma_{\lambda,h\bar{h}}(\boldsymbol{l}_\perp, z)\Psi^{\mathcal{H}*}_{\lambda,h\bar{h}}(\boldsymbol{l}'_\perp - z\boldsymbol{\Delta}_\perp, z)\mathrm{e}^{\mathrm{i}(\boldsymbol{l}_\perp - \boldsymbol{l}'_\perp + \frac{1}{2}\boldsymbol{\Delta}_\perp)\cdot \boldsymbol{r}_\perp}\,. \tag{2}$$

[4] We have the following abbreviation for the transverse coordinate space integrals: $\int_{\boldsymbol{r}_\perp} \equiv \int \mathrm{d}^2\boldsymbol{r}_\perp$. For transverse and longitudinal momentum space integrals we use: $\int_{\boldsymbol{l}_\perp} \equiv \int \mathrm{d}^2\boldsymbol{l}_\perp/(2\pi)^2$ and $\int_z \equiv \int_0^1 \mathrm{d}z/(4\pi)$.

A particular contribution to the amplitude (1) is shown in Fig. 1, where the upper heavy quark loop represents the reduced amplitude $\mathcal{A}_{\lambda\bar{\lambda}}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp)$, and the three-gluon contribution is the lowest order depiction of the Odderon exchange amplitude $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ for the particular case where the gluons connect to three different quarks in the proton wave function, represented by the green blob.

The photon wave-function in momentum space is given by

$$\Psi^\gamma_{\lambda,h\bar{h}}(\boldsymbol{k}_\perp, z) \equiv \sqrt{z\bar{z}}\,\frac{\bar{u}_h(k) e q_c \not{\epsilon}(\lambda, q) v_{\bar{h}}(q-k)}{\boldsymbol{k}_\perp^2 + \varepsilon^2}\,, \tag{3}$$

where $q_c = +2/3$ is the fractional electric charge of the $c$-quark and $u_h(k)$ ($v_{\bar{h}}(q-k)$) are particle (antiparticle) spinors with $h = \pm 1$ indicating the sign of the helicity $h/2$. For explicit computations we use Lepage-Brodsky (LB) spinors [46] - see App. A for their expressions. $\epsilon(\lambda, q)$ is the photon polarization vector: we have $\epsilon(0, q) = (Q/q^-, 0, 0, 0)$ for longitudinal ($\lambda = 0$) and $\epsilon(\lambda, q) = (0, 0, \boldsymbol{\epsilon}^\lambda_\perp)$ for transverse ($\lambda = \pm 1$) polarization. We are using the LB convention for the 2D polarization vector: $\boldsymbol{\epsilon}^\lambda_\perp \equiv (-\lambda, -\mathrm{i})/\sqrt{2}$, and the common shorthands $\varepsilon \equiv \sqrt{m_c^2 + z\bar{z}Q^2}$ and $\bar{z} \equiv 1 - z$, with $z = k^-/q^-$. We compute the photon wave function as described in App. A, with the result

$$\begin{aligned}
\Psi^\gamma_{+,h\bar{h}}(\boldsymbol{k}_\perp, z) &= -\sqrt{2} e q_c \left[k_\perp \mathrm{e}^{\mathrm{i}\phi_k}\left(z\delta_{h+}\delta_{\bar{h}-} - \bar{z}\delta_{h-}\delta_{\bar{h}+}\right) + m_c \delta_{h+}\delta_{\bar{h}+}\right] \frac{1}{\boldsymbol{k}_\perp^2 + \varepsilon^2}\,, \\
\Psi^\gamma_{-,h\bar{h}}(\boldsymbol{k}_\perp, z) &= -\sqrt{2} e q_c \left[k_\perp \mathrm{e}^{-\mathrm{i}\phi_k}\left(\bar{z}\delta_{h+}\delta_{\bar{h}-} - z\delta_{h-}\delta_{\bar{h}+}\right) + m_c \delta_{h-}\delta_{\bar{h}-}\right] \frac{1}{\boldsymbol{k}_\perp^2 + \varepsilon^2}\,, \\
\Psi^\gamma_{0,h\bar{h}}(\boldsymbol{k}_\perp, z) &= \frac{e q_c 2Q z\bar{z}}{\boldsymbol{k}_\perp^2 + \varepsilon^2}\delta_{h,-\bar{h}}\,,
\end{aligned} \tag{4}$$

where $k_\perp \mathrm{e}^{\pm \mathrm{i}\phi_k} = k^1 \pm \mathrm{i}k^2$. Eq. (4) coincides with the result of Ref. [47] for both $\lambda = 0$ and $\lambda = \pm 1$ after adjusting for an overall factor of $-\sqrt{N_c}$ due to the difference in conventions. Comparing to Ref. [48] the result coincides for $\lambda = 0$ (again up to a factor $-\sqrt{N_c}$) but there is an opposite overall sign for $\lambda = \pm 1$ (the relative factor being $+\sqrt{N_c}$), as also explicitly noted in [48]. While such overall signs play no role for the processes involving overlaps of photons and vector quarkonia, the sign is important for tensor quarkonia, since the tensor wave function can be understood as a linear superposition of the vector wave functions, see Eq. (12) below. The computation described in App. A takes the LB spinors as a starting point from which we proceed to explicitly compute the photon wave function as well as all the subsequent quarkonia wave functions in a systematic and self-contained fashion.

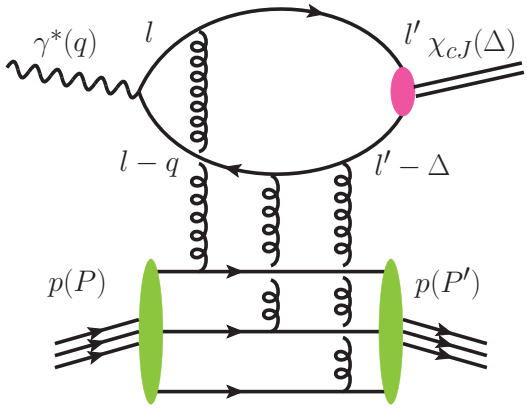



FIG. 1. An example Feynman diagram illustrating the amplitude for exclusive production of $\chi_{cJ}$'s via three gluon (Odderon) exchange from the proton in $\gamma^* p$ collision. The green blobs represent the proton wave function that is shown here in terms of three valence quarks. The magenta blob is the $\chi_{cJ}$ wave function. In the dipole approach the two gluons coupling to the same quark line in the top part of the diagram occupy the same point in transverse position space.

The explicit expressions for the quarkonia wave functions will be given in the following subsection. Here we want to emphasize that thanks to the $C = +1$ parity of the quarkonium wave function (see Eq. (10)) below) and to the $C = -1$ parity of the photon, the amplitude is strictly proportional to the $C = -1$ Odderon amplitude $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$. In the high energy limit of eikonal dipole-proton scattering we have [11, 12]

$$\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp) = -\frac{1}{2\mathrm{i}N_c}\mathrm{tr}\left\langle V(\boldsymbol{x}_\perp)V^\dagger(\boldsymbol{y}_\perp) - V(\boldsymbol{y}_\perp)V^\dagger(\boldsymbol{x}_\perp)\right\rangle\,, \tag{5}$$

where $\boldsymbol{r}_\perp$ is the dipole size and $\boldsymbol{b}_\perp$ the impact parameter

$$\boldsymbol{r}_\perp \equiv \boldsymbol{x}_\perp - \boldsymbol{y}_\perp\,, \qquad \boldsymbol{b}_\perp \equiv \frac{\boldsymbol{x}_\perp + \boldsymbol{y}_\perp}{2}\,, \tag{6}$$

respectively [5]. $V(\boldsymbol{x}_\perp)$ and $V^\dagger(\boldsymbol{x}_\perp)$ represent Wilson lines at transverse coordinate $\boldsymbol{x}_\perp$ which describe the propagation of the quark and anti-quark, respectively, through the color field of the target proton in covariant gauge, and $\langle\cdots\rangle$ denotes an average over the configurations of that field, see App. D for the conventions used in this work. In Fourier space, the Odderon amplitude is given by

$$\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) \equiv \int_{\boldsymbol{b}_\perp} \mathrm{e}^{-\mathrm{i}\boldsymbol{\Delta}_\perp\cdot\boldsymbol{b}_\perp}\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)\,. \tag{7}$$

The $C$-invariance of the full amplitude (1) is now easily verified. We can conventionally start by first exchanging the quark-antiquark coordinates $\boldsymbol{x}_\perp \leftrightarrow \boldsymbol{y}_\perp$ so that $\boldsymbol{r}_\perp \to -\boldsymbol{r}_\perp$ in the Odderon amplitude. Because of the phases in (2) this needs to be followed by an exchange of quark and antiquark transverse momenta $\boldsymbol{l}_\perp \leftrightarrow \boldsymbol{l}'_\perp - z\boldsymbol{\Delta}_\perp$ and $z \leftrightarrow \bar{z}$. By this transformation only the photon wave function picks up a sign which cancels with the sign in the transformation of the Odderon amplitude as $\mathcal{O}(-\boldsymbol{r}_\perp, \boldsymbol{b}_\perp) = -\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$.

### A. Light-cone wave function of $C$-even quarkonia

The light-cone wave function of the $C$-even quarkonia, $\Psi^{\mathcal{H}}_{\bar\lambda,h\bar h}$, is modeled by the following covariant ansatz

$$\Psi^{\mathcal{H}}_{\bar\lambda,h\bar h}(\boldsymbol{k}_\perp, z) \equiv \frac{1}{\sqrt{z\bar z}}\bar u_h(k)\Gamma^{\mathcal{H}}_{\bar\lambda}(k,k')v_{\bar h}(k')\phi_{\mathcal{H}}(\boldsymbol{k}_\perp, z)\,. \tag{8}$$

Here $\bar\lambda$ is the quarkonia helicity and $k$ ($k' = \Delta_0 - k$) is the quark (antiquark) momenta, with $\Delta_0$ being the invariant four-momentum of the $c\bar c$ pair. $\Gamma^{\mathcal{H}}_{\bar\lambda}$ is the appropriate Dirac matrix vertex function for either a scalar ($\mathcal{H} = \mathcal{S}$), axial vector ($\mathcal{A}$) or a tensor ($\mathcal{T}$) quarkonia which we take as[6]

$$\Gamma^{\mathcal{H}}_{\bar\lambda}(k,k') = \begin{cases} 1\,, & \mathcal{H} = \mathcal{S}\,,\\ \mathrm{i}\gamma_5 \not{E}(\bar\lambda, \Delta_0)\,, & \mathcal{H} = \mathcal{A}\,,\\ \frac{1}{2}\left(\gamma_\mu(k_\nu - k'_\nu) + \gamma_\nu(k_\mu - k'_\mu)\right)E^{\mu\nu}(\bar\lambda,\Delta_0)\,, & \mathcal{H} = \mathcal{T}\,. \end{cases} \tag{9}$$

$\phi_{\mathcal{H}}(\boldsymbol{k}_\perp, z)$ is a non-perturbative scalar function that we model later (see Sec. IV).

The structure of (9) is motivated in part to ensure the correct $C$-even property of the quarkonia wave function in (8). Namely, by exchanging the quark and the antiquark momenta ($k \leftrightarrow k'$) and helicity ($h \leftrightarrow \bar h$) we have

$$\bar u_{\bar h}(k')\Gamma^{\mathcal{H}}_{\bar\lambda}(k',k)v_h(k) = +\bar u_h(k)\Gamma^{\mathcal{H}}_{\bar\lambda}(k,k')v_{\bar h}(k')\,, \tag{10}$$

which holds thanks to the relation $C(\Gamma^{\mathcal{H}}_{\bar\lambda}(k',k))^T C^{-1} = +\Gamma^{\mathcal{H}}_{\bar\lambda}(k,k')$ with $C = \mathrm{i}\gamma^0\gamma^2$. Furthermore, the $\Gamma^{\mathcal{T}}_{\bar\lambda}(k,k')$ vertex is modelled as a coupling to the fermionic energy-momentum tensor, see e. g. [49–52].

In (9) $E(\bar\lambda, \Delta_0)$ describe the polarization state of the axial quarkonia. To ensure that the axial vector carries the correct quantum numbers, we require the transversality condition

$$\Delta_0 \cdot E(\bar\lambda, \Delta_0) = 0\,. \tag{11}$$

This way, in the $c\bar c$ rest frame, $E^\mu(\bar\lambda, \Delta_0)$ reduces to its spatial components corresponding to the total spin vector of the $c\bar c$ state. Recall that in the light-cone formulation the "$-$" and $\perp$ components of $\Delta_0$ are exactly equal to the corresponding components of the quarkonia four-momentum $\Delta$, but that the "$+$" components differ: $\Delta^+ = (M^2_{\mathcal{H}} + \boldsymbol{\Delta}^2_\perp)/2\Delta^-$ and $\Delta^+_0 = ({M_0}^2 + \boldsymbol{\Delta}^2_\perp)/2\Delta^-$ with $M^2_0 = \Delta^2_0$ being the invariant mass of the $c\bar c$ pair and $M^2_{\mathcal{H}} = \Delta^2$ the quarkonia mass. This is because the $c\bar c$ pair is allowed to be virtual in the "$+$" component, conjugate to the light-cone time. Explicitly, we have

$$E^\mu(\bar\lambda = \pm1, \Delta_0) = \left(\frac{\boldsymbol{\epsilon}^{\bar\lambda}_\perp\cdot\boldsymbol{\Delta}_\perp}{\Delta^-}, 0, \boldsymbol{\epsilon}^{\bar\lambda}_\perp\right)\,, \quad E^\mu(\bar\lambda = 0, \Delta_0) = \left(\frac{\boldsymbol{\Delta}^2_\perp - M^2_0}{2M_0\Delta^-}, \frac{\Delta^-}{M_0}, \frac{\boldsymbol{\Delta}_\perp}{M_0}\right) = \frac{1}{M_0}\Delta^\mu_0 - \frac{M_0}{\Delta^-}n^\mu\,. \tag{12}$$

---

[5] The impact parameter of the dipole-proton collision is actually $\tilde{\boldsymbol{b}}_\perp = z\boldsymbol{x}_\perp + \bar z\boldsymbol{y}_\perp = \boldsymbol{b}_\perp - (\frac{1}{2} - z)\boldsymbol{r}_\perp$, and $\tilde{\boldsymbol{b}}_\perp$ is the Fourier conjugate of the transverse momentum transfer $\boldsymbol{\Delta}_\perp$. This is the origin of the "off-forward phase" $-i\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp$ in Eq. (18).

[6] Our convention is $\gamma^5 \equiv +\mathrm{i}\gamma^0\gamma^1\gamma^2\gamma^3$ and $\epsilon^{0123} \equiv +1 = -\epsilon^{+-12}$.

Using $\Delta$ instead of $\Delta_0$ has no effect on the $\bar\lambda=\pm1$ components, namely $E^\mu(\bar\lambda=\pm1,\Delta_0)=E^\mu(\bar\lambda=\pm1,\Delta)$. However, the $\bar\lambda=0$ case differs as

$$E(\bar\lambda=0,\Delta)=\frac{1}{M_{\mathcal{A}}}\Delta^\mu-\frac{M_{\mathcal{A}}}{\Delta^-}n^\mu\,. \tag{13}$$

We have checked that the projector $\mathrm{i}\gamma_5\not{E}(\bar\lambda=0,\Delta)$ (instead of $\mathrm{i}\gamma_5\not{E}(\bar\lambda=0,\Delta_0)$) leads to incorrect results for the $c\bar c$ state as it leaves an admixture with spin different from 1. This happens because in the light-cone formulation the rest frames of the quarkonia and of the $c\bar c$ partonic state in general move with different velocities, which is a consequence of the difference in their "+" components of the four-momenta. The necessary condition for the correct projector, $\Delta_0\cdot E(\bar\lambda,\Delta)=0$, is fulfilled if and only if $M_0=M_{\mathcal{A}}$. In other words, we will assume that the amplitude for the transition from $E^\mu(\bar\lambda=0,\Delta_0)$ to $E^\mu(\bar\lambda=0,\Delta)$ is equal to one. A similar approach was used in Ref. [53]. The underlying principle of parton–hadron duality is also the basis for the successful approach to exclusive $\rho^0$ electroproduction [54, 55].

The tensor quarkonia polarizations $E^{\mu\nu}(\bar\lambda,\Delta_0)$ can be obtained in terms of $E^\mu(\bar\lambda,\Delta_0)$, using Clebsch-Gordan coefficients. We have [49]

$$\begin{aligned}
E^{\mu\nu}(\pm2,\Delta_0)&=E^\mu(\pm1,\Delta_0)E^\nu(\pm1,\Delta_0)\,,\\
E^{\mu\nu}(\pm1,\Delta_0)&=\frac{1}{\sqrt2}\big(E^\mu(\pm1,\Delta_0)E^\nu(0,\Delta_0)+E^\mu(0,\Delta_0)E^\nu(\pm1,\Delta_0)\big)\,,\\
E^{\mu\nu}(0,\Delta_0)&=\frac{1}{\sqrt6}\big(E^\mu(+1,\Delta_0)E^\nu(-1,\Delta_0)+E^\mu(-1,\Delta_0)E^\nu(+1,\Delta_0)+2E^\mu(0,\Delta_0)E^\nu(0,\Delta_0)\big)\,.
\end{aligned} \tag{14}$$

These polarization tensors are traceless, $g_{\mu\nu}E^{\mu\nu}(\bar\lambda,\Delta_0)=0$, and symmetric, $E^{\mu\nu}(\bar\lambda,\Delta_0)=E^{\nu\mu}(\bar\lambda,\Delta_0)$. We also have transversality conditions, $\Delta_0^\mu E_{\mu\nu}(\bar\lambda,\Delta_0)=0=\Delta_0^\nu E_{\mu\nu}(\bar\lambda,\Delta_0)$, as a consequence of (11). We have checked that Eqs. (14) are consistent with the polarization tensors written in [51], up to overall signs.

It is instructive to explicitly compute the quarkonia wave functions in this approach. Using LB spinors [46] we find for the scalar quarkonia

$$\Psi^{\mathcal{S}}_{h\bar h}(\boldsymbol{k}_\perp,z)=\frac{1}{z\bar z}\left[hk_\perp\mathrm{e}^{-\mathrm{i}h\phi_k}\delta_{h\bar h}-m_c(z-\bar z)\delta_{h,-\bar h}\right]\phi_{\mathcal{S}}(\boldsymbol{k}_\perp,z)\,. \tag{15}$$

The helicity structure of the wave function agrees completely with the result in [56], see (A.18) there. For the axial quarkonia we have

$$\begin{aligned}
\Psi^{\mathcal{A}}_{+1,h\bar h}(\boldsymbol{k}_\perp,z)&=-\frac{\sqrt2\mathrm{i}}{z\bar z}\left[k_\perp\mathrm{e}^{\mathrm{i}\phi_k}(\bar z\delta_{h+}\delta_{\bar h-}+z\delta_{h-}\delta_{\bar h+})+m_c(z-\bar z)\delta_{h+}\delta_{\bar h+}\right]\phi_{\mathcal{A},T}(\boldsymbol{k}_\perp,z)\,,\\
\Psi^{\mathcal{A}}_{-1,h\bar h}(\boldsymbol{k}_\perp,z)&=-\frac{\sqrt2\mathrm{i}}{z\bar z}\left[k_\perp\mathrm{e}^{-\mathrm{i}\phi_k}(\bar z\delta_{h+}\delta_{\bar h-}+z\delta_{h-}\delta_{\bar h+})-m_c(z-\bar z)\delta_{h-}\delta_{\bar h-}\right]\phi_{\mathcal{A},T}(\boldsymbol{k}_\perp,z)\,,\\
\Psi^{\mathcal{A}}_{0,h\bar h}(\boldsymbol{k}_\perp,z)&=-\frac{\mathrm{i}}{z\bar z}\frac{1}{M_{\mathcal{A}}}\left(2\boldsymbol{k}_\perp^2h\delta_{h,-\bar h}+2m_ck_\perp\mathrm{e}^{-\mathrm{i}h\phi_k}\delta_{h\bar h}\right)\phi_{\mathcal{A},L}(\boldsymbol{k}_\perp,z)\,,
\end{aligned} \tag{16}$$

where $k_\perp\mathrm{e}^{\pm\mathrm{i}\phi_k}\equiv k^1\pm\mathrm{i}k^2$. At this point we have introduced different scalar functions $\phi_{\mathcal{A},T}(r_\perp,z)$ and $\phi_{\mathcal{A},L}(r_\perp,z)$ for transversely and longitudinally polarized quarkonia[7]. The helicity structure of (16) coincides with the axial part of the $Z$-boson wave function from Refs. [57, 58]. In Ref. [59] the axial quarkonia wave function was computed starting from the quarkonia rest frame followed by a Melosh transformation. The resulting wave functions, their Eqs. (A.7) and (A.9), agree up to an overall normalization constant with (16).

For the tensor quarkonia we obtain the following results

$$\begin{aligned}
\Psi^{\mathcal{T}}_{+2,h\bar h}(\boldsymbol{k}_\perp,z)&=-\frac{2}{z\bar z}k_\perp\mathrm{e}^{\mathrm{i}\phi_k}\left[k_\perp\mathrm{e}^{\mathrm{i}\phi_k}(z\delta_{h+}\delta_{\bar h-}-\bar z\delta_{h-}\delta_{\bar h+})+m_c\delta_{h+}\delta_{\bar h+}\right]\phi_{\mathcal{T},T2}(\boldsymbol{k}_\perp,z)\,,\\
\Psi^{\mathcal{T}}_{-2,h\bar h}(\boldsymbol{k}_\perp,z)&=\frac{2}{z\bar z}k_\perp\mathrm{e}^{-\mathrm{i}\phi_k}\left[k_\perp\mathrm{e}^{-\mathrm{i}\phi_k}(\bar z\delta_{h+}\delta_{\bar h-}-z\delta_{h-}\delta_{\bar h+})+m_c\delta_{h-}\delta_{\bar h-}\right]\phi_{\mathcal{T},T2}(\boldsymbol{k}_\perp,z)\,,\\
\Psi^{\mathcal{T}}_{+1,h\bar h}(\boldsymbol{k}_\perp,z)&=\frac{M_{\mathcal{T}}}{z\bar z}\left[-k_\perp\mathrm{e}^{\mathrm{i}\phi_k}\left((3z-4z^2)\delta_{h+}\delta_{\bar h-}+(3\bar z-4\bar z^2)\delta_{h-}\delta_{\bar h+}\right)+m_c(z-\bar z)\delta_{h+}\delta_{\bar h+}\right]\phi_{\mathcal{T},T}(\boldsymbol{k}_\perp,z)\,,\\
\Psi^{\mathcal{T}}_{-1,h\bar h}(\boldsymbol{k}_\perp,z)&=\frac{M_{\mathcal{T}}}{z\bar z}\left[k_\perp\mathrm{e}^{-\mathrm{i}\phi_k}\left((3\bar z-4\bar z^2)\delta_{h+}\delta_{\bar h-}+(3z-4z^2)\delta_{h-}\delta_{\bar h+}\right)+m_c(z-\bar z)\delta_{h-}\delta_{\bar h-}\right]\phi_{\mathcal{T},T}(\boldsymbol{k}_\perp,z)\,,\\
\Psi^{\mathcal{T}}_{0,h\bar h}(\boldsymbol{k}_\perp,z)&=\frac{\sqrt2}{\sqrt3}\frac{1}{z\bar z}\left[(3\boldsymbol{k}_\perp^2+2m_c^2)(z-\bar z)\delta_{h,-\bar h}+m_c(k_\perp\mathrm{e}^{-\mathrm{i}\phi_k}\delta_{h+}\delta_{\bar h+}-k_\perp\mathrm{e}^{\mathrm{i}\phi_k}\delta_{h-}\delta_{\bar h-})\right]\phi_{\mathcal{T},L}(\boldsymbol{k}_\perp,z)\,.
\end{aligned} \tag{17}$$

[7] For the sake of simplicity, in the case of $\bar\lambda=0$, we have redefined $\phi_{\mathcal{A},L}(\boldsymbol{k}_\perp,z)/M_0\to\phi_{\mathcal{A},L}(\boldsymbol{k}_\perp,z)/M_{\mathcal{A}}$, which introduces only subleading effects in the heavy quark limit.

Here we have three different scalar functions $\phi_{\mathcal{T},T2}(\boldsymbol{k}_\perp, z)$, $\phi_{\mathcal{T},T}(\boldsymbol{k}_\perp, z)$ and $\phi_{\mathcal{T},L}(\boldsymbol{k}_\perp, z)$ accounting for the two possible transverse and longitudinal polarizations of the tensor quarkonia[8].

At this point we make a comment on the spin-momentum structure of the obtained wave functions in the rest frame. Since the $C$ parity of a $q\bar{q}$ bound state is $C = (-1)^{L+S}$, the quarkonia with $J = 0$, 1 and 2, are all $P$-waves ($L = 1$), and have $S = 1$. Based on our model wave function (8) we therefore explicitly checked that, after replacing the light-cone LB spinors with Dirac spinors ($u_h(k) \to u_s(k)$ and $v_h(k) \to v_s(k)$), the wave function in the quarkonia rest frame have their expected non-relativistic ${}^3P_J$ spin-momentum structure. The scalar wave function is proportional to $\xi_{\bar{s}}^\dagger(\boldsymbol{\sigma}\cdot\boldsymbol{k})\tilde{\xi}_s$, while for the axial case we find $\xi_{\bar{s}}^\dagger\boldsymbol{\sigma}\cdot(\boldsymbol{k}\times\boldsymbol{E})\tilde{\xi}_s$ [53, 56, 59, 60]. Here $\xi_s$ are the standard two-component Pauli spinors. In case of the tensor wave function we obtain $\xi_{\bar{s}}^\dagger E^{ij}\sigma_i k_j\tilde{\xi}_s$ [61].

### B. Final expressions for the amplitudes

For explicit computations it is convenient to write the reduced amplitude (2) in the following equivalent form, see e.g. Refs. [16, 47, 62]

$$\mathcal{A}_{\lambda\bar{\lambda}}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) = eq_c \int_z \mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp} \int_{\boldsymbol{l}_\perp} \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{l}_\perp\cdot\boldsymbol{r}_\perp}}{\boldsymbol{l}_\perp^2 + \varepsilon^2} \int_{\boldsymbol{l}'_\perp} \mathrm{e}^{-\mathrm{i}(\boldsymbol{l}'_\perp - z\boldsymbol{\Delta}_\perp)\cdot\boldsymbol{r}_\perp} \phi_{\mathcal{H}}(\boldsymbol{l}'_\perp - z\boldsymbol{\Delta}_\perp, z) \frac{1}{z\bar{z}} A_{\lambda\bar{\lambda}}(\boldsymbol{l}_\perp, \boldsymbol{l}'_\perp - z\boldsymbol{\Delta}_\perp, z)\,, \tag{18}$$

where $\boldsymbol{\delta}_\perp = \frac{1}{2}(z-\bar{z})\boldsymbol{\Delta}_\perp = (z-\frac{1}{2})\boldsymbol{\Delta}_\perp$. The helicity sum in (2) is turned into a covariant Dirac trace

$$A_{\lambda\bar{\lambda}}(\boldsymbol{l}_\perp, \boldsymbol{l}'_\perp - z\boldsymbol{\Delta}_\perp, z) = \frac{1}{(2q^-)^2}\mathrm{tr}\left[(\not{l}' + m_c)\not{\epsilon}(\lambda, q)(\not{l}' - \not{q} + m_c)\gamma^-(\not{l}' - \not{\Delta} + m_c)\Gamma^{\mathcal{H}*}_{\bar{\lambda}}(l', \Delta - l')(\not{l}' + m_c)\gamma^-\right], \tag{19}$$

containing the physical information on the polarization dependent part of the photon-quarkonia wave function overlap. Together with Eq. (1), Eqs. (18) and (19) comprise the main formulas that will be used below to write the $\gamma^* p \to \mathcal{H} p$ amplitudes in a form suitable for numerical computations. For explicit evaluation of the traces we have used FeynCalc [63].

We start with the computation of the amplitude for the scalar quarkonia. Inserting the first line of (9) into (19) yields

$$\begin{aligned} A_{\lambda=0}(\boldsymbol{l}_\perp, \boldsymbol{l}_{1\perp}, z) &= -4m_c Q z\bar{z}(z-\bar{z})\,, \\ A_{\lambda=\pm1}(\boldsymbol{l}_\perp, \boldsymbol{l}_{1\perp}, z) &= -2m_c\left[(z-\bar{z})^2(\boldsymbol{\epsilon}_\perp^\lambda\cdot\boldsymbol{l}_\perp) - (\boldsymbol{\epsilon}_\perp^\lambda\cdot\boldsymbol{l}_{1\perp})\right], \end{aligned} \tag{20}$$

where $\boldsymbol{l}_{1\perp} \equiv \boldsymbol{l}'_\perp - z\boldsymbol{\Delta}_\perp$. As a cross check, note that the traces (20) are odd under joint $\boldsymbol{l}_\perp \to -\boldsymbol{l}_\perp$, $\boldsymbol{l}_{1\perp} \to -\boldsymbol{l}_{1\perp}$ and $z \to \bar{z}$ transformation. In other words, the reduced amplitude $\mathcal{A}_\lambda(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp)$ in (18) is odd under $\boldsymbol{r}_\perp \to -\boldsymbol{r}_\perp$. This is consistent with $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp)$ being odd under $\boldsymbol{r}_\perp \to -\boldsymbol{r}_\perp$ in the full amplitude (1).

We now plug (20) into (18) and Fourier transform to coordinate space. After separating the explicit polarization dependence, the reduced amplitudes are found to be

$$\begin{aligned} \mathcal{A}_0(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) &= eq_c \int_z \mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_L(r_\perp)\,, \\ \mathcal{A}_{\lambda=\pm1}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) &= eq_c\lambda\mathrm{e}^{\mathrm{i}\lambda\phi_r}\int_z \mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_T(r_\perp)\,, \end{aligned} \tag{21}$$

where

$$\begin{aligned} \mathcal{A}_L(r_\perp) &\equiv -\frac{2}{\pi}m_c Q(z-\bar{z})K_0(\varepsilon r_\perp)\phi_{\mathcal{S}}(r_\perp, z)\,, \\ \mathcal{A}_T(r_\perp) &\equiv \frac{\mathrm{i}\sqrt{2}}{2\pi}\frac{m_c}{z\bar{z}}\left[(z-\bar{z})^2\varepsilon K_1(\varepsilon r_\perp)\phi_{\mathcal{S}}(r_\perp, z) - K_0(\varepsilon r_\perp)\frac{\partial\phi_{\mathcal{S}}}{\partial r_\perp}\right]. \end{aligned} \tag{22}$$

In the amplitude (1) we are left with the $\boldsymbol{r}_\perp$ integral and the $z$ integral. We use the Jacobi-Anger expansion to compute the $\phi_r$ integral and find

$$\begin{aligned} \langle\mathcal{M}_{\lambda=0}\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=0} \equiv q^-\widetilde{\mathcal{M}}_L\,, \\ \langle\mathcal{M}_{\lambda=\pm1}\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=\pm1} \equiv q^-\lambda\mathrm{e}^{\mathrm{i}\lambda\phi_\Delta}\widetilde{\mathcal{M}}_T\,, \end{aligned} \tag{23}$$

[8] In the case $\bar{\lambda} = \pm1$ we have redefined $M_0\phi_{\mathcal{T},T}(\boldsymbol{k}_\perp, z) \to M_{\mathcal{T}}\phi_{\mathcal{T},T}(\boldsymbol{k}_\perp, z)$.

where in the first equality we have separated the flux factor and in the second equality the explicit polarization dependence (together with the overall phase). The remainder consists of two different scalar functions

$$\begin{aligned}\widetilde{\mathcal{M}}_L &= 8\pi N_c e q_c \sum_{k=0}^{\infty}(-1)^k \int_z \int_0^\infty r_\perp \mathrm{d}r_\perp \mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)\mathcal{A}_L(r_\perp)\,\mathrm{sgn}(z-\bar{z})J_{2k+1}(r_\perp\delta_\perp)\,,\\ \widetilde{\mathcal{M}}_T &= 4\pi \mathrm{i} N_c e q_c \sum_{k=0}^{\infty}(-1)^k \int_z \int_0^\infty r_\perp \mathrm{d}r_\perp \mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)\mathcal{A}_T(r_\perp)\left[J_{2k}(r_\perp\delta_\perp)-J_{2k+2}(r_\perp\delta_\perp)\right]\,,\end{aligned} \tag{24}$$

which are our main results to be used in the numerical computations in Sec. V. The sign function appears due to the definition of $\delta_\perp$ as the modulus of the vector $\boldsymbol{\delta}_\perp$: $\delta_\perp = \frac{1}{2}|z-\bar{z}|\Delta_\perp$.

The quantities $\mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)$ appearing in Eq. (24) are the azimuthal harmonics of the Odderon amplitude,

$$\mathcal{O}_{2k+1}(r_\perp,\Delta_\perp) = 2\pi\mathrm{i}(-1)^{k+1}\int_0^\infty b_\perp \mathrm{d}b_\perp \mathcal{O}_{2k+1}(r_\perp,b_\perp)J_{2k+1}(\Delta_\perp b_\perp)\,, \tag{25}$$

where $\mathcal{O}_{2k+1}(r_\perp,b_\perp)$ is extracted from $\mathcal{O}(\boldsymbol{r}_\perp,\boldsymbol{b}_\perp)$ as its Fourier series coefficients

$$\mathcal{O}(\boldsymbol{r}_\perp,\boldsymbol{b}_\perp) = 2\sum_{k=0}^{\infty}\mathcal{O}_{2k+1}(r_\perp,b_\perp)\cos((2k+1)\phi_{rb})\,,\qquad \mathcal{O}_{2k+1}(r_\perp,b_\perp) = \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\phi_{rb}\mathcal{O}(\boldsymbol{r}_\perp,\boldsymbol{b}_\perp)\cos((2k+1)\phi_{rb})\,, \tag{26}$$

with $\phi_{rb} = \phi_r - \phi_b$.

The calculation for axial vector quarkonia follows the same steps as for scalar case. Detailed derivation of the formulas can be found in App. B. We end up with three different scalar functions:

$$\begin{aligned}\widetilde{\mathcal{M}}_B &= 4\pi\mathrm{i}N_c e q_c \sum_{k=0}^{\infty}(-1)^k\int_z\int_0^\infty r_\perp\mathrm{d}r_\perp\mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)\mathcal{A}_B(r_\perp)\left[J_{2k}(r_\perp\delta_\perp)-J_{2k+2}(r_\perp\delta_\perp)\right]\,, \qquad B=TL,LT\\ \widetilde{\mathcal{M}}_{TT} &= 8\pi N_c e q_c \sum_{k=0}^{\infty}(-1)^k\int_z\int_0^\infty r_\perp\mathrm{d}r_\perp\mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)\mathcal{A}_{TT}(r_\perp)\,\mathrm{sgn}(z-\bar{z})J_{2k+1}(r_\perp\delta_\perp)\,,\end{aligned} \tag{27}$$

with respective reduced amplitudes as

$$\begin{aligned}\mathcal{A}_{LT}(r_\perp) &\equiv \frac{\sqrt{2}}{\pi}QK_0(\varepsilon r_\perp)\frac{\partial\phi_{\mathcal{A},T}}{\partial r_\perp}\,,\\ \mathcal{A}_{TL}(r_\perp) &\equiv \frac{\sqrt{2}}{\pi}\frac{1}{z\bar{z}}\frac{1}{M_{\mathcal{A}}}\left[-m_c^2K_0(\varepsilon r_\perp)\frac{\partial\phi_{\mathcal{A},L}}{\partial r_\perp}+\varepsilon K_1(\varepsilon r_\perp)\boldsymbol{\nabla}_\perp^2\phi_{\mathcal{A},L}\right]\,,\\ \mathcal{A}_{TT}(r_\perp) &\equiv -\frac{\mathrm{i}}{\pi}\frac{z-\bar{z}}{z\bar{z}}\left[\frac{\partial\phi_{\mathcal{A},T}}{\partial r_\perp}\varepsilon K_1(\varepsilon r_\perp)-m_c^2K_0(\varepsilon r_\perp)\phi_{\mathcal{A},T}\right]\,.\end{aligned} \tag{28}$$

Here the notation $\widetilde{\mathcal{M}}_{TT}$ stands for amplitude with transition from transversely polarized photon to transversely polarized axial quarkonia. In principle, both the polarization preserving and polarization flipping transition would be possible. However we find that the polarization flipping transition vanishes, and so $\widetilde{\mathcal{M}}_{TT}$ describes only a polarization preserving transition. Similar notation is used for other contributing amplitudes. We find the amplitude when both the photon and the axial quarkonia would be longitudinal also vanishes, see App. B for more details.

The procedure is again similar for tensor quarkonia, with detailed steps to be found in App. B. The final results

for the amplitudes are

$$\begin{aligned}
\widetilde{\mathcal{M}}_B &= -4\pi N_c e q_c \sum_{k=0}^{\infty}(-1)^k \int_z\int_0^\infty r_\perp \mathrm{d}r_\perp \mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)\mathcal{A}_B(r_\perp) \\
&\qquad \times \mathrm{sgn}(z-\bar{z})\left[J_{2k+3}(r_\perp\delta_\perp)+J_{2k-1}(r_\perp\delta_\perp)\right], \quad B = LT2, TTf, \\
\widetilde{\mathcal{M}}_B &= 4\pi \mathrm{i} N_c e q_c \sum_{k=0}^{\infty}(-1)^k \int_z\int_0^\infty r_\perp \mathrm{d}r_\perp \mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)\mathcal{A}_B(r_\perp) \\
&\qquad \times \left[J_{2k}(r_\perp\delta_\perp)-J_{2k+2}(r_\perp\delta_\perp)\right], \quad B = TT2p, LT, TL, \\
\widetilde{\mathcal{M}}_B &= 8\pi N_c e q_c \sum_{k=0}^{\infty}(-1)^k \int_z\int_0^\infty r_\perp \mathrm{d}r_\perp \mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)\mathcal{A}_B(r_\perp)\mathrm{sgn}(z-\bar{z})J_{2k+1}(r_\perp\delta_\perp), \quad B = TTp, LL, \\
\widetilde{\mathcal{M}}_{TT2f} &= 4\pi \mathrm{i} N_c e q_c \sum_{k=0}^{\infty}(-1)^k \int_z\int_0^\infty r_\perp \mathrm{d}r_\perp \mathcal{O}_{2k+1}(r_\perp,\Delta_\perp)\mathcal{A}_{TT2f}(r_\perp)\left[J_{2k+4}(r_\perp\delta_\perp)-J_{2k-2}(r_\perp\delta_\perp)\right],
\end{aligned} \tag{29}$$

where

$$\begin{aligned}
\mathcal{A}_{LT2}(r_\perp) &\equiv \frac{2}{\pi}(z-\bar{z})QK_0(\varepsilon r_\perp)\left(\frac{\partial^2\phi_{\mathcal{T},T2}}{\partial r_\perp^2}-\frac{1}{r_\perp}\frac{\partial\phi_{\mathcal{T},T2}}{\partial r_\perp}\right), \\
\mathcal{A}_{TT2p}(r_\perp) &\equiv -\frac{\mathrm{i}\sqrt{2}}{\pi}\frac{1}{z\bar{z}}\left((z^2+\bar{z}^2)\varepsilon K_1(\varepsilon r_\perp)\left(\frac{\partial^2\phi_{\mathcal{T},T2}}{\partial r_\perp^2}-\frac{1}{r_\perp}\frac{\partial\phi_{\mathcal{T},T2}}{\partial r_\perp}\right)+m_c^2K_0(\varepsilon r_\perp)\frac{\partial\phi_{\mathcal{T},T2}}{\partial r_\perp}\right), \\
\mathcal{A}_{TT2f}(r_\perp) &\equiv \frac{\mathrm{i}2\sqrt{2}}{\pi}\varepsilon K_1(\varepsilon r_\perp)\left(\frac{\partial^2\phi_{\mathcal{T},T2}}{\partial r_\perp^2}-\frac{1}{r_\perp}\frac{\partial\phi_{\mathcal{T},T2}}{\partial r_\perp}\right), \\
\mathcal{A}_{LT}(r_\perp) &\equiv -\frac{\mathrm{i}}{\pi}QM_{\mathcal{T}}\left(3-4(z^2+\bar{z}^2)\right)K_0(\varepsilon r_\perp)\frac{\partial\phi_{\mathcal{T},T}}{\partial r_\perp}, \\
\mathcal{A}_{TTp}(r_\perp) &= -\frac{M_{\mathcal{T}}}{\sqrt{2}\pi}\frac{z-\bar{z}}{z\bar{z}}\left[m_c^2K_0(\varepsilon r_\perp)\phi_{\mathcal{T},T}(r_\perp,z)-(z-\bar{z})^2\varepsilon K_1(\varepsilon r_\perp)\frac{\partial\phi_{\mathcal{T},T}}{\partial r_\perp}\right], \\
\mathcal{A}_{TTf}(r_\perp) &= -\frac{2\sqrt{2}M_{\mathcal{T}}}{\pi}(z-\bar{z})\varepsilon K_1(\varepsilon r_\perp)\frac{\partial\phi_{\mathcal{T},T}}{\partial r_\perp}, \\
\mathcal{A}_{LL}(r_\perp) &= -\frac{2\sqrt{2}Q}{\sqrt{3}\pi}(z-\bar{z})K_0(\varepsilon r_\perp)\left(3\boldsymbol{\nabla}_\perp^2-2m_c^2\right)\phi_{\mathcal{T},L}(r_\perp,z), \\
\mathcal{A}_{TL}(r_\perp) &= \frac{\mathrm{i}}{\pi\sqrt{3}}\frac{1}{z\bar{z}}\left[\varepsilon K_1(\varepsilon r_\perp)(z-\bar{z})^2\left(3\boldsymbol{\nabla}_\perp^2-2m_c^2\right)\phi_{\mathcal{T},L}-m_c^2K_0(\varepsilon r_\perp)\frac{\partial\phi_{\mathcal{T},L}}{\partial r_\perp}\right].
\end{aligned} \tag{30}$$

In this case the $LL$-type transition is allowed. Also, both the polarization preserving and the polarization flipped contributions are allowed, thus explaining the above used notation: $p$ = preserving, $f$ = flipped.

### C. The $\gamma^* p \to \mathcal{H}p$ cross section

Using the $\widetilde{\mathcal{M}}_{\lambda\bar{\lambda}}$ amplitudes obtained in the previous Sec. II B for the process $\gamma^*(q)p(P) \to \mathcal{H}(\Delta)p(P')$, the cross section for longitudinal ($L$) and transverse ($T$) photons is given by

$$\begin{aligned}
\frac{\mathrm{d}\sigma_L(\gamma^*p\to\mathcal{H}p)}{\mathrm{d}|t|} &= \frac{1}{16\pi}\sum_{\bar{\lambda}=-J}^{J}\left|\widetilde{\mathcal{M}}_{\lambda=0,\bar{\lambda}}(\gamma^*p\to\mathcal{H}p)\right|^2, \\
\frac{\mathrm{d}\sigma_T(\gamma^*p\to\mathcal{H}p)}{\mathrm{d}|t|} &= \frac{1}{16\pi}\frac{1}{2}\sum_{\lambda=\pm1}\sum_{\bar{\lambda}=-J}^{J}\left|\widetilde{\mathcal{M}}_{\lambda\bar{\lambda}}(\gamma^*p\to\mathcal{H}p)\right|^2.
\end{aligned} \tag{31}$$

Inserting the scalar quarkonia amplitudes (23) into (31) gives

$$\begin{aligned}
\frac{\mathrm{d}\sigma_L(\gamma^*p\to\mathcal{S}p)}{\mathrm{d}|t|} &= \frac{1}{16\pi}\left|\widetilde{\mathcal{M}}_L\right|^2, \\
\frac{\mathrm{d}\sigma_T(\gamma^*p\to\mathcal{S}p)}{\mathrm{d}|t|} &= \frac{1}{16\pi}\left|\widetilde{\mathcal{M}}_T\right|^2.
\end{aligned} \tag{32}$$

where $\widetilde{\mathcal{M}}_B$ ($B = L$, $T$) are given in Eqs. (24). For the axial vector quarkonia we find

$$\begin{aligned}\frac{\mathrm{d}\sigma_L(\gamma^* p \to \mathcal{A}p)}{\mathrm{d}|t|} &= \frac{1}{8\pi}\left|\widetilde{\mathcal{M}}_{LT}\right|^2 ,\\ \frac{\mathrm{d}\sigma_T(\gamma^* p \to \mathcal{A}p)}{\mathrm{d}|t|} &= \frac{1}{16\pi}\left(\left|\widetilde{\mathcal{M}}_{TL}\right|^2 + \left|\widetilde{\mathcal{M}}_{TT}\right|^2\right) ,\end{aligned} \tag{33}$$

with $\widetilde{\mathcal{M}}_B$ ($B = LT$, $TL$, $TT$) given in (27). We have used (B3) (see Appendix B) to relate $\widetilde{\mathcal{M}}_{\lambda\bar{\lambda}}$ to $\widetilde{\mathcal{M}}_A$. For the tensor quarkonia, the cross sections are

$$\begin{aligned}\frac{\mathrm{d}\sigma_L(\gamma^* p \to \mathcal{T}p)}{\mathrm{d}|t|} &= \frac{1}{16\pi}\left(\left|\widetilde{\mathcal{M}}_{LL}\right|^2 + 2\left|\widetilde{\mathcal{M}}_{LT}\right|^2 + 2\left|\widetilde{\mathcal{M}}_{LT2}\right|^2\right) ,\\ \frac{\mathrm{d}\sigma_T(\gamma^* p \to \mathcal{T}p)}{\mathrm{d}|t|} &= \frac{1}{16\pi}\left(\left|\widetilde{\mathcal{M}}_{TL}\right|^2 + \left|\widetilde{\mathcal{M}}_{TTp}\right|^2 + \left|\widetilde{\mathcal{M}}_{TTf}\right|^2 + \left|\widetilde{\mathcal{M}}_{TT2p}\right|^2 + \left|\widetilde{\mathcal{M}}_{TT2f}\right|^2\right) ,\end{aligned} \tag{34}$$

with $\widetilde{\mathcal{M}}_B$ ($B = LL$, $LT$, $LT2$, $TL$, $TTp$, $TTf$, $TT2p$, $TT2f$) collected in (29) and the relationship to $\widetilde{\mathcal{M}}_{\lambda\bar{\lambda}}$ in (B7).

### D. Odderon amplitude and its evolution with $x$

Here we discuss the Odderon amplitude $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ used later in Sec. V for numerical predictions of the cross section for exclusive $\chi_{cJ}$ production. The initial $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ at $x_{\mathcal{P}} = x_0 = 0.01$ is approximated by the matrix element of the $C$-odd three gluon operator. This matrix element has been evaluated numerically and tabulated in Ref. [64] using a phenomenological non-perturbative three quark light-cone wave function [65, 66] supplemented by the first correction of perturbative QCD [23]. We refer to these references for further details. Specifically, we use the table for the first azimuthal harmonic $a_1(r_\perp, b_\perp)$ of the $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ defined in Eq. (5) of Ref. [64]. As explained in the App. D the definition of $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ in [64] is the same as in this work[9]. The first azimuthal moment is defined in [64] through $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp) = \cos(\phi_{rb}) a_1(r_\perp, b_\perp) + \ldots$. Thus, from (26) we have the relationship

$$\mathcal{O}_1(r_\perp, b_\perp) = \frac{1}{2} a_1(r_\perp, b_\perp) . \tag{35}$$

To determine $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ for $x_{\mathcal{P}} < x_0$ we solve the impact parameter dependent [67, 68] extension of the Balitsky-Kovchegov (BK) equation (with running coupling kernel – rcBK) to the coupled non-linear evolution equations for the Pomeron and the Odderon derived in Refs. [11, 12] and analyzed in Refs. [13–15]. These equations describe the evolution of the real and imaginary parts of the dipole $S$-matrix, and the initial condition is evolved in rapidity $Y = \ln(x_0/x_{\mathcal{P}})$. For the real part (the Pomeron), we are using the initial condition from [69]. The coupled BK equations for the Pomeron-Odderon system are solved in the "local approximation" [69, 70], where the impact parameter $\boldsymbol{b}_\perp$ becomes an external parameter. This approximation may be less justified for the Odderon than for the Pomeron since the former amplitude peaks at smaller $b_\perp$ [64]. The $t$-dependence of three gluon exchange is then obtained via Fourier transform, Eq. (7). Technical details of the implementation and additional numerical results can be found in Sec. III of Ref. [16].

Fig. 2 shows the first azimuthal harmonic of the odderon amplitude $\mathrm{i}\mathcal{O}_1(r_\perp, \Delta_\perp)$ as a function of $\Delta_\perp^2$ for a dipole of size $r_\perp = 0.3$ fm. We have divided by the parameter $\sigma_0/2 = 16.36$ mb $= 42.027$ GeV$^{-2}$ (area of the proton) [69] to compensate for the dimensionality of $\mathrm{i}\mathcal{O}_1(r_\perp, \Delta_\perp)$. At $x_{\mathcal{P}} = 10^{-2}$ this corresponds to the perturbative exchange of three gluons (with negative $C$-parity) while for smaller $x_{\mathcal{P}}$ the BK resummation has been performed as described above. We note, first of all, the small magnitude of $\mathcal{O}_1(r_\perp, \Delta_\perp) < 10^{-2} \ll 1$, i. e. that scattering of perturbative (small) dipoles via $C$-odd exchanges is predicted to be very weak. However, we also observe that the $C$-odd amplitude peaks at fairly large $\Delta_\perp^2 \simeq 1$ GeV$^2$ and that it decreases by merely a factor of $\sim 2$ as $\Delta_\perp^2$ increases further to $\simeq 3$ GeV$^2$. This is a manifestation of the "Landshoff mechanism" [71], see also [72–74] whereby large-$t$ scattering via three gluon exchange is less likely to break up the proton than scattering via two gluon exchange[10]. In simple terms this is due to the fact that a large momentum transfer can be shared by the exchanged gluons, and so a comparable transverse momentum can be transmitted to up to three partons in the proton, resulting in a smaller increase of their invariant

[9] However, our own numerical investigation and private communication with the authors of Ref. [64] revealed that their plots of $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ and their tabulated values actually correspond to minus the function defined in their Eq. (3) and in our Eq. (D11). We therefore reversed the signs of their table entries for $a_1(r_\perp, b_\perp)$.

[10] The proton model light-cone wave function employed here, however, includes a perturbative correction [23] due to the emission of a gluon from one of the valence quarks, or the internal exchange of a gluon, not considered in the quoted papers by Donnachie and Landshoff.

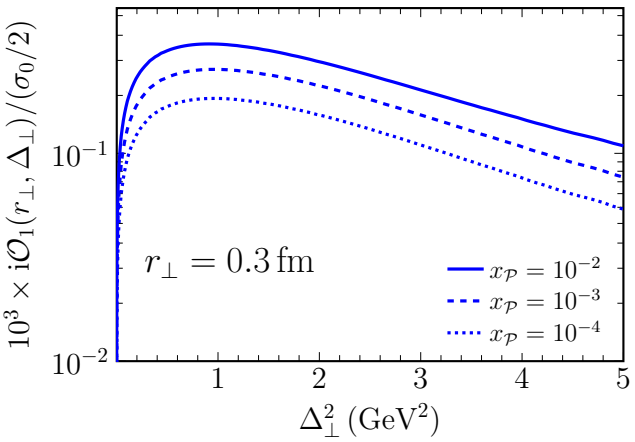



FIG. 2. The first azimuthal harmonic of the Odderon amplitude $\mathrm{i}\mathcal{O}_1(r_\perp, \Delta_\perp)$ as a function of $\Delta_\perp^2$ for $r_\perp = 0.3$ fm and $\alpha_S = 0.25$. Different line styles correspond to different values of $x_{\mathcal{P}}$.

mass than in case of two gluon exchange. Such a contribution has been illustrated on Fig. 1. The light-cone Fock space amplitudes of the proton are strongly suppressed if the invariant mass of the parton system is far from $\sim N_c\Lambda_{\mathrm{QCD}}$. On the other hand, an amplitude which depends weakly on momentum transfer, in impact parameter space would peak at small $b_\perp$. Amplitudes with negative parity must vanish at $b_\perp = 0$, however, and hence are expected to have small magnitude.

Fig. 2 also shows the decrease of the Odderon exchange amplitude with $x_{\mathcal{P}}$ by approximately a factor of 2 from $x_{\mathcal{P}} = 10^{-2}$ to $x_{\mathcal{P}} = 10^{-4}$. The initial decrease of the Odderon amplitude has been understood to originate from gluon Regge trajectory suppressions in the BJKP equation [11, 13, 15]. In App. E we have checked that by omitting the unitarity corrections, the Odderon still decreases but with a smaller rate. Importantly, the sign of the Odderon amplitude is preserved by the evolution equation, which is simply a consequence of the fact that its evolution equation is linear in $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$.

## III. THE PRIMAKOFF CONTRIBUTION

In this section we focus on the Primakoff process, wherein the exclusive $\gamma^* p \to \mathcal{H}p$ production proceeds via photon rather than Odderon exchange. Typically, the cross section for a QCD process far exceeds its QED counterparts, and so the latter can be safely neglected. This is, however, not so in case of the Odderon because its QCD cross section starts at order $\alpha_S^6$ and so the QED contribution becomes a competitive background in Odderon searches, in particular for low momentum transfer.

We treat the Primakoff process in the high-energy approximation of eikonal scattering, replacing the Odderon amplitude $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp)$ in Eq. (1) with [16, 32]

$$\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) \to 8\pi \mathrm{i} q_c \alpha \sin\left(\frac{\boldsymbol{\Delta}_\perp \cdot \boldsymbol{r}_\perp}{2}\right) \frac{F_1(\Delta_\perp)}{\boldsymbol{\Delta}_\perp^2}\,, \tag{36}$$

at leading order in the electromagnetic coupling $\alpha = e^2/4\pi$. The amplitude can be written as

$$\langle \mathcal{M}_{\lambda\bar{\lambda}}(\gamma^* p \to \mathcal{H}p)\rangle \equiv -\frac{eF_1(\ell_\perp)}{\boldsymbol{\ell}_\perp^2} n_\mu \mathcal{M}^\mu_{\lambda\bar{\lambda}}(\gamma^*\gamma^* \to \mathcal{H})\,, \tag{37}$$

where $\mathcal{M}^\mu_{\lambda\bar{\lambda}}(\gamma^*\gamma^* \to \mathcal{H})$ is the $\gamma^*(\lambda, q)\gamma^*(\mu, \ell) \to \mathcal{H}(\bar{\lambda}, \Delta)$ amplitude (in the high energy limit) that is obtained by inserting (36) into (1). In (37), $\ell \equiv P - P'$ is the momentum transfer from the proton, with $\boldsymbol{\ell}_\perp = \boldsymbol{\Delta}_\perp$ in the dipole frame and $F_1(\ell_\perp)$ is the Dirac charge form factor of the proton with $F_1(0) = 1$, see App. D for more detail. Since (36) represents single photon exchange it is more convenient to start from the expression in Eq. (18) for the reduced amplitude and perform the $\boldsymbol{r}_\perp$ integral. This leads to

$$\frac{1}{q^-} n_\mu \mathcal{M}^\mu_{\lambda\bar{\lambda}}(\gamma^*\gamma^* \to \mathcal{H}) = -8\pi \mathrm{i} q_c^2 \alpha N_c \int_z \int_{\boldsymbol{l}_\perp} \left[\frac{A_{\lambda\bar{\lambda}}(\boldsymbol{l}_\perp - \bar{z}\boldsymbol{\Delta}_\perp, \boldsymbol{l}_\perp, z)}{(\boldsymbol{l}_\perp - \bar{z}\boldsymbol{\Delta}_\perp)^2 + \varepsilon^2} - \frac{A_{\lambda\bar{\lambda}}(\boldsymbol{l}_\perp + z\boldsymbol{\Delta}_\perp, \boldsymbol{l}_\perp, z)}{(\boldsymbol{l}_\perp + z\boldsymbol{\Delta}_\perp)^2 + \varepsilon^2}\right] \frac{\phi_{\mathcal{H}}(\boldsymbol{l}_\perp, z)}{z\bar{z}}\,, \tag{38}$$

The term $eF_1(\ell_\perp)/\boldsymbol{\ell}_\perp^2$ in Eq. (37) is understood as a part of the usual Weizsäcker-Williams photon field

$$A^\mu(\ell) = (2\pi)\delta(\ell^-)\frac{1}{\boldsymbol{\ell}_\perp^2} eF_1(\ell_\perp) n^\mu \,, \tag{39}$$

where $n^\mu$ is the light-cone gauge vector.

### A. $|t| \to 0$ limit

In the high energy approximation it is possible to express the $|t| \to 0$ limit of the Primakoff cross section for a $C$-even quarkonia $\mathcal{H}$ with spin $J \neq 1$ in terms of the two-photon decay width $\Gamma(\mathcal{H} \to \gamma\gamma)$

$$\lim_{|t|\to 0} |t| \frac{\mathrm{d}\sigma(\gamma p \to \mathcal{H} p)}{\mathrm{d}|t|} = \frac{8\pi\alpha(2J+1)\Gamma(\mathcal{H}\to\gamma\gamma)}{M_{\mathcal{H}}^3} \,. \tag{40}$$

Eq. (40) is a model independent result and not affected by QCD corrections. Hence it represents a stringent constraint on the model predictions and a useful cross check. The key point for obtaining (40) is the relation of the high energy impact factor for the quarkonia photoproduction, to the quarkonia's two-photon decay amplitude. We provide a derivation of (40) in App. C. For axial vector quarkonia, like $\chi_{c1}$, $\Gamma(\mathcal{A} \to \gamma\gamma)$ vanishes and Eq. (40) implies only that $\lim_{|t|\to 0} |t| \mathrm{d}\sigma(\gamma p \to \mathcal{A} p)/\mathrm{d}|t| = 0$, as is consistent with the Landau-Yang (LY) theorem [75, 76], but does not constrain the value of the cross section at $|t| \to 0$. In App. C we thus provide a separate and general analysis of the $|t| \to 0$ limit for the axial quarkonia.

In practice, we will use (40) to deduce $\Gamma(\mathcal{H} \to \gamma\gamma)$ from the computation of its left hand side. This will be required in Sec. IV where we perform a fit of the quarkonia wave functions. Therefore, we now focus on obtaining explicit expressions of the $\gamma^* p \to \mathcal{H} p$ amplitudes in the $|t| \to 0$ (and $Q^2 \to 0$) limit.

Starting with the scalar quarkonia we use (38) and insert the appropriate expressions for the reduced amplitudes $A_{\lambda,\bar\lambda}(\boldsymbol{l}_\perp, \boldsymbol{l}_{1\perp}, z)$ found in Sec. II B. The relevant reduced amplitude in the $Q^2 \to 0$ limit concerns only the second line of (20). The expression in the square bracket in (38) is conveniently factored as

$$\frac{A_{\lambda=\pm1}(\boldsymbol{l}_\perp - \bar z \boldsymbol{\Delta}_\perp, \boldsymbol{l}_\perp, z)}{(\boldsymbol{l}_\perp - \bar z \boldsymbol{\Delta}_\perp)^2 + \varepsilon^2} - \frac{A_{\lambda=\pm1}(\boldsymbol{l}_\perp + z\boldsymbol{\Delta}_\perp, \boldsymbol{l}_\perp, z)}{(\boldsymbol{l}_\perp + z\boldsymbol{\Delta}_\perp)^2 + \varepsilon^2} \to 2m_c \frac{(\boldsymbol{\epsilon}_\perp^\lambda \cdot \boldsymbol{\Delta}_\perp)}{(\boldsymbol{l}_\perp^2 + m_c^2)^2} \left[\boldsymbol{l}_\perp^2 + (z-\bar z)^2 m_c^2\right] \,, \tag{41}$$

where we have set $Q^2 = 0$, performed an expansion around $\Delta_\perp \to 0$ to linear order, as well as the angular average according to $l_\perp^i l_\perp^j \to \boldsymbol{l}_\perp^2 \delta^{ij}/2$. We now insert (41) into (38) and use second line of (23) to extract $\widetilde{\mathcal{M}}_T$ as

$$\widetilde{\mathcal{M}}_T \to 8\sqrt{2}\pi \mathrm{i} q_c^2 \alpha N_c m_c \frac{eF_1(\Delta_\perp)}{\Delta_\perp} \int_z \int_{\boldsymbol{l}_\perp} \frac{\boldsymbol{l}_\perp^2 + (z-\bar z)^2 m_c^2}{(\boldsymbol{l}_\perp^2 + m_c^2)^2} \frac{\phi_{\mathcal{S}}(\boldsymbol{l}_\perp, z)}{z\bar z} \,. \tag{42}$$

Performing similar steps for axial quarkonia we find that for the square brackets in (38) the $O(\Delta_\perp)$ contributions vanish for $\bar\lambda = 0$ and for $\bar\lambda = \pm 1$. A non-vanishing contribution is found for $\bar\lambda = \pm 1$ at $O(\Delta_\perp^2)$ leading to

$$\widetilde{\mathcal{M}}_{TT} \to -16\pi q_c^2 \alpha N_c e F_1(\Delta_\perp) \int_z \int_{\boldsymbol{l}_\perp} \frac{(z-\bar z)^2 m_c^2}{(\boldsymbol{l}_\perp^2 + m_c^2)^2} \frac{\phi_{\mathcal{A}}(\boldsymbol{l}_\perp, z)}{z\bar z} \,, \tag{43}$$

while the $\bar\lambda = 0$ contribution is $O(\Delta_\perp^3)$. Such a special $\Delta_\perp$ dependence of the axial quarkonia amplitude is explained in App. C 2 from general considerations.

For the calculation of the $|t| \to 0$ limit for the tensor quarkonia we also need

$$l_\perp^i l_\perp^j l_\perp^m l_\perp^n \to \frac{1}{8} \boldsymbol{l}_\perp^4 (\delta^{ij}\delta^{mn} + \delta^{im}\delta^{jn} + \delta^{in}\delta^{jm}) \,. \tag{44}$$

Computing for various tensor polarizations $\bar\lambda$ we find that $\bar\lambda = \pm 1$ does not contribute in the $|t| \to 0$ limit. As for the $\bar\lambda = \pm 2$ and $\bar\lambda = 0$ cases, we extract the amplitudes (29) as

$$\begin{aligned}
&\widetilde{\mathcal{M}}_{TT2,p} \to 16\sqrt{2}\pi \mathrm{i} q_c^2 \alpha N_c \frac{eF_1(\Delta_\perp)}{\Delta_\perp} \int_z \int_{\boldsymbol{l}_\perp} \frac{\boldsymbol{l}_\perp^2}{(\boldsymbol{l}_\perp^2 + m_c^2)^2} \left[(z^2+\bar z^2)\boldsymbol{l}_\perp^2 + m_c^2\right] \frac{\phi_{\mathcal{T},T2}(\boldsymbol{l}_\perp, z)}{z\bar z} \,, \\
&\widetilde{\mathcal{M}}_{TT2,f} \to 0 \,, \\
&\widetilde{\mathcal{M}}_{TL} \to -\frac{16}{\sqrt{3}}\pi \mathrm{i} q_c^2 \alpha N_c \frac{eF_1(\Delta_\perp)}{\Delta_\perp} \int_z \int_{\boldsymbol{l}_\perp} \frac{m_c^2}{(\boldsymbol{l}_\perp^2 + m_c^2)^2} \left[\boldsymbol{l}_\perp^2 - 2(z-\bar z^2)\left(\frac{3}{2}\boldsymbol{l}_\perp^2 + m_c^2\right)\right] \frac{\phi_{\mathcal{T},L}(\boldsymbol{l}_\perp, z)}{z\bar z} \,.
\end{aligned} \tag{45}$$

### B. Adding the Pauli form factor

In general, the $\gamma^* p \to \mathcal{H} p$ amplitude can be related to the $\gamma^*\gamma^* \to \mathcal{H}$ amplitude by separating out the Dirac current with the Dirac ($F_1$) and Pauli ($F_2$) form factors

$$\mathcal{M}_{\lambda\bar{\lambda}}(\gamma^* p \to \mathcal{H} p) = \mathcal{M}^{\mu}_{\lambda\bar{\lambda}}(\gamma^*\gamma^* \to \mathcal{H}) \frac{g_{\mu\nu}}{\ell^2} \bar{u}_{h'}(P') \left[ eF_1(\ell_\perp)\gamma^\nu + \mathrm{i}\sigma^{\nu\rho}\ell_\rho \frac{eF_2(\ell_\perp)}{2m_N} \right] u_h(P)\,, \tag{46}$$

where $u_h(P)$ is a proton spinor with helicity $h$ and $\sigma^{\nu\rho} = \mathrm{i}[\gamma^\nu, \gamma^\rho]/2$. The dominant contribution of the first term in the high energy limit is with $\gamma^+$, while in the second term the dominant contribution is with $\sigma^{+i}$. Thus, we can write in a covariant way

$$\mathcal{M}_{\lambda\bar{\lambda}}(\gamma^* p \to \mathcal{H} p) \approx P_\mu \mathcal{M}^{\mu}_{\lambda\bar{\lambda}}(\gamma^*\gamma^* \to \mathcal{H}) \frac{1}{q\cdot P} \bar{u}_{h'}(P') \left[ \frac{eF_1(\ell_\perp)}{\ell^2} \not{q} + \mathrm{i}\sigma^{\nu\rho} q_\nu \ell_\rho \frac{eF_2(\ell_\perp)}{2m_N \ell^2} \right] u_h(P)\,. \tag{47}$$

The above spinor vertices in the high energy limit become[11]

$$\begin{aligned} &\bar{u}_{h'}(P')\not{q} u_h(P) \approx W^2 \delta_{hh'}\,, \\ &\bar{u}_{h'}(P')\mathrm{i}\sigma^{\nu\rho} q_\nu \ell_\rho u_h(P) \approx \ell_\perp h \mathrm{e}^{\mathrm{i}h\phi_\ell} W^2 \delta_{h,-h'}\,. \end{aligned} \tag{48}$$

Taking into account the connection to the usual notion of the color averaged amplitude $\langle \mathcal{M}_{\lambda\bar{\lambda}} \rangle = q^- \widetilde{\mathcal{M}}_{\lambda\bar{\lambda}} = q^- \mathcal{M}_{\lambda\bar{\lambda}}/W^2$, we have

$$\langle \mathcal{M}_{\lambda\bar{\lambda}}(\gamma^* p \to \mathcal{H} p) \rangle = -n_\mu \mathcal{M}^{\mu}_{\lambda\bar{\lambda}}(\gamma^*\gamma^* \to \mathcal{H}) \left[ \frac{eF_1(\ell_\perp)}{\boldsymbol{\ell}_\perp^2} \delta_{hh'} + \frac{eF_2(\ell_\perp)}{\boldsymbol{\ell}_\perp^2} \frac{\ell_\perp}{2m_N} h \mathrm{e}^{\mathrm{i}h\phi_\ell} \delta_{h,-h'} \right], \tag{49}$$

where the first term reproduces (37). Thanks to the difference in the helicity structure in the high energy limit there is no interference between the $F_1$ and $F_2$ contributions in (49). For the same reason there is no intereference of the $F_2$ term with the Odderon exchange amplitude. The addition of $F_2$ to the Primakoff cross section therefore amounts to a simple replacement

$$F_1^2(\ell_\perp) \to F_1^2(\ell_\perp) + \frac{\boldsymbol{\ell}_\perp^2}{4m_N^2} F_2^2(\ell_\perp)\,. \tag{50}$$

Thanks to the additional $\ell_\perp$ dependence, the $F_2$ contribution is negligible at $|t| = \boldsymbol{\ell}_\perp^2 \to 0$ but in practice becomes up to about 50% correction at finite $|t|$ [77].

## IV. BOOSTED GAUSSIAN MODEL FOR THE $\chi_{cJ}$ WAVE FUNCTIONS

In this section we perform a fit of the scalar part of the $C$-even quarkonia wave function $\phi_{\mathcal{H},i}(\boldsymbol{k}_\perp, z)$. We adopt a Boosted Gaussian ansatz [78] originally written in coordinate space as

$$\phi_{\mathcal{H},B}(r_\perp, z) = \mathcal{N}_{\mathcal{H},B} z\bar{z} \exp\left( -\frac{m_c^2 \mathcal{R}_{\mathcal{H}}^2}{8z\bar{z}} - \frac{2z\bar{z}r_\perp^2}{\mathcal{R}_{\mathcal{H}}^2} + \frac{1}{2} m_c^2 \mathcal{R}_{\mathcal{H}}^2 \right), \tag{51}$$

where $\mathcal{R}_{\mathcal{H}}$ is the radii parameter and $\mathcal{N}_{\mathcal{H},i}$ are normalization parameters for different quarkonia species and polarizations that are to be determined below. For $\mathcal{H} = \mathcal{S}$ we have only one normalization, $\mathcal{N}_{\mathcal{S}}$, for $\mathcal{H} = \mathcal{A}$, we have $B = T$, $L$, denoting transverse and longitudinal polarizations. For $\mathcal{H} = \mathcal{T}$, the index $B$ spans over $B = T2$, $T$, $L$. The wave function is normalized as

$$1 = N_c \sum_{h\bar{h}} \int_z \int_{\boldsymbol{r}_\perp} \left| \Psi^{\mathcal{H}}_{\lambda,h\bar{h}}(\boldsymbol{r}_\perp, z) \right|^2 . \tag{52}$$

---

[11] The calculation steps are very similar to the one performed in App. A and also available in [46].

In practice, we convert (52) to momentum space and express the helicity sum as a Dirac trace as

$$1 = N_c \int_{\boldsymbol{k}_\perp} \int_z \frac{1}{z\bar{z}} \mathrm{tr}\left[(\not{k} - m_c)\gamma^0 \Gamma^{\mathcal{H},\dagger}_{\bar{\lambda}}(k,k')\gamma^0(\not{k}' + m_c)\Gamma^{\mathcal{H}}_{\bar{\lambda}}(k,k')\right] |\phi_{\mathcal{H},B}(\boldsymbol{k}_\perp, z)|^2 \,. \tag{53}$$

To constrain the parameters $\mathcal{R}_\mathcal{H}$ and $\mathcal{N}_{\mathcal{H},i}$, we are using (53) and the $\mathcal{H} \to \gamma\gamma$ decay width for $\chi_{c0}$ and $\chi_{c2}$. The $\chi_{c1} \to \gamma\gamma$ decay is forbidden due to the LY theorem. In this case our assumption is that the $\mathcal{R}$ parameter of $\chi_{c1}$ is equal to that for $\chi_{c2}$.

The decay rate for $\mathcal{H} = \mathcal{S}, \mathcal{T}$ can be obtained from the correspondence in Eq. (40). For scalar quarkonia we first insert the amplitude (42) into (32) to find the $|t| \to 0$ limit of the cross section. We then extract the decay rate via (40) as

$$\begin{aligned} \Gamma(\mathcal{S} \to \gamma\gamma) &= \frac{\pi\alpha^2}{4} M_\mathcal{S}^3 F_\mathcal{S}^2 \,, \\ F_\mathcal{S} &\equiv 4q_c^2 m_c N_c \int_z \int_{\boldsymbol{k}_\perp} \frac{\boldsymbol{k}_\perp^2 + (z-\bar{z})^2 m_c^2}{(\boldsymbol{k}_\perp^2 + m_c^2)^2} \frac{\phi_\mathcal{S}(\boldsymbol{k}_\perp, z)}{z\bar{z}} \,, \end{aligned} \tag{54}$$

which agrees perfectly with Eq. (13) in Ref. [79] after adjusting the conventions. We have additionally confirmed that in the NRQCD limit (54) agrees with the known result in [80] (see App. C).

For the actual fit, the LO result in (54) is supplemented by the NLO QCD corrections [80, 81]. This amounts to a replacement

$$\Gamma(\mathcal{S} \to \gamma\gamma) \to \left(1 + \frac{\alpha_S}{\pi}\left(\frac{\pi^2}{3} - \frac{28}{9}\right)\right)\Gamma(\mathcal{S} \to \gamma\gamma) \approx (1 + 0.06\alpha_S)\Gamma(\mathcal{S} \to \gamma\gamma) \,. \tag{55}$$

For the experimental value of $\chi_{c0} \to \gamma\gamma$ decay width we use the most recent PDG value: $\Gamma(\chi_{c0} \to \gamma\gamma) = 2.203 \times 10^{-6}$ GeV [33]. Using $m_c = 1.4$ GeV and $M_{\chi_{c0}} = 3.414$ GeV and $\alpha_S(2m_c) \approx 0.25$, we determine $\mathcal{N}_{\chi_{c0}} = 1.148$ and $\mathcal{R}_{\chi_{c0}} = 1.539$ GeV$^{-1}$. As expected, the radius parameter $\mathcal{R}$ of the $\chi_{c0}$ is similar to that of the $J/\psi$ [78].

For $\mathcal{T}$ we insert the amplitude (45) into (34) and extract the decay rate via (40) as

$$\begin{aligned} \Gamma(\mathcal{T} \to \gamma\gamma) &= \frac{\pi\alpha^2}{20} M_\mathcal{T}^3 \left(F_{\mathcal{T},L}^2 + F_{\mathcal{T},T2}^2\right) \,, \\ F_{\mathcal{T},L} &= \frac{4\sqrt{2}}{\sqrt{3}} q_c^2 N_c \int_z \int_{\boldsymbol{k}_\perp} \frac{m_c^2\left[\boldsymbol{k}_\perp^2 - 2(z-\bar{z})^2\left(\frac{3}{2}\boldsymbol{k}_\perp^2 + m_c^2\right)\right]}{(\boldsymbol{k}_\perp^2 + m_c^2)^2} \frac{\phi_{\mathcal{T},L}(\boldsymbol{k}_\perp, z)}{z\bar{z}} \,, \\ F_{\mathcal{T},T2} &= 8q_c^2 N_c \int_z \int_{\boldsymbol{k}_\perp} \frac{\boldsymbol{k}_\perp^2\left[(z^2+\bar{z}^2)\boldsymbol{k}_\perp^2 + m_c^2\right]}{(\boldsymbol{k}_\perp^2 + m_c^2)^2} \frac{\phi_{\mathcal{T},T2}(\boldsymbol{k}_\perp, z)}{z\bar{z}} \,, \end{aligned} \tag{56}$$

where $F_{\mathcal{T},L}$ ($F_{\mathcal{T},T2}$) originates from the $\bar{\lambda} = 0$ ($\bar{\lambda} = \pm 2$) polarizations. We find only a partial agreement of our result (56) and Eqs. (22)-(24) of [79]. The third line of (56) can be brought in agreement with Eq. (24) in [79] after appropriate adjustments in conventions and also after a judicious identification of $M_\mathcal{T}$ with the pair invariant mass $M_0$. However, the square brackets of the second line in (56) contain a term $\boldsymbol{k}_\perp^2$ while in (23) of [79] they rather have $-m_c^2$. Using (56) in the NRQCD limit we find the $\bar{\lambda} = 0$ contribution vanishes with the $\bar{\lambda} = \pm 2$ contribution saturating the NRQCD limit completely and in accordance with the known result [80]. Taking the NRQCD limit of (22)-(24) in [79] gives an incorrect result as it leads to a finite contribution from the $\bar{\lambda} = 0$ polarization.

The QCD corrections [80, 81] turn out to be sizeable in this particular channel [80, 81]

$$\Gamma(\mathcal{T} \to \gamma\gamma) \to \left(1 - \frac{16\alpha_S}{3\pi}\right)\Gamma(\mathcal{T} \to \gamma\gamma) \approx (1 - 1.7\alpha_S)\Gamma(\mathcal{T} \to \gamma\gamma) \,. \tag{57}$$

Using $M_{\chi_{c2}} = 3.556$ GeV and the PDG value $\Gamma(\chi_{c2} \to \gamma\gamma) = 5.614 \times 10^{-7}$ GeV [33], we find $\mathcal{N}_{\chi_{c2},T2} = 0.609$ GeV$^{-1}$ $= \mathcal{N}_{\chi_{c2},L}$, $\mathcal{N}_{\chi_{c2},T} = 0.592$ GeV$^{-1}$ and $\mathcal{R}_{\chi_{c2}} = 1.482$ GeV$^{-1}$. In these numerical estimates we used $\alpha_S = 0.25$ in the NLO correction factors.

Finally, for $\chi_{c1}$ we assume $\mathcal{R}_{\chi_{c1}} = \mathcal{R}_{\chi_{c2}}$, allowing us to fix the parameters $\mathcal{N}_{\chi_{c1},T,L}$ via normalization. We find $\mathcal{N}_{\chi_{c1},T} = 1.388$ and $\mathcal{N}_{\chi_{c1},L} = 1.401$.

From the above fits we may compute charge radii of the $\chi_{cJ}$ states via the following formula [82]:

$$\langle r_c^2 \rangle = N_c \sum_{h\bar{h}} \int_z \int_{\boldsymbol{r}_\perp} \left(A_{\bar{\lambda}}^\mathcal{H} \bar{z}^2 \boldsymbol{r}_\perp^2\right) \left|\Psi^\mathcal{H}_{\bar{\lambda},h\bar{h}}(\boldsymbol{r}_\perp, z)\right|^2 \,. \tag{58}$$

$(A_{\bar\lambda}^{\mathcal{H}})^{-1}$ denotes the expectation value of $\boldsymbol{r}_\perp^2/\boldsymbol{r}^2 = \sin^2\theta$ in the state

$$|J,\bar\lambda\rangle = \sum_{m,m_s} C(J,\bar\lambda|L,m;S,m_s)\,|L,m\rangle\,|S,m_s\rangle\ , \tag{59}$$

where $C(J,\bar\lambda|L,m;S,m_s)$ denotes the Clebsch-Gordan coefficient for $L=S=1$. This gives $A^{\chi_{c0}} = 3/2$, $A_0^{\chi_{c1}} = 5/4$, $A_1^{\chi_{c1}} = 5/3$, $A_0^{\chi_{c2}} = 15/8$, $A_1^{\chi_{c2}} = 5/3$, $A_2^{\chi_{c2}} = 5/4$. For $\chi_{c0}$ we then find $\sqrt{\langle r_c^2\rangle} = 0.270$ fm, which is close to the value obtained in Ref. [82] from potential models. For $\chi_{c1}$ we find $\sqrt{\langle r_c^2\rangle} = 0.277$ fm for $\bar\lambda = \pm 1$, while $\sqrt{\langle r_c^2\rangle} = 0.236$ fm was obtained for $\bar\lambda = 0$. For $\chi_{c2}$ we have $\sqrt{\langle r_c^2\rangle} = 0.263$ fm for $\bar\lambda = \pm 2$, $\sqrt{\langle r_c^2\rangle} = 0.262$ fm for $\bar\lambda = \pm 1$ and $\sqrt{\langle r_c^2\rangle} = 0.237$ fm for $\bar\lambda = 0$. Thus, all charge radii are of comparable magnitude and less than 0.3 fm, which appears reasonable.

## V. NUMERICAL RESULTS

In this Section we will firstly show the numerical results for the $\gamma^* p \to \chi_{cJ} p$ cross section based on the formulas in Secs. II B and III and relevant for the EIC kinematics. Using the $\gamma^* p \to \chi_{cJ} p$ cross sections, we will also compute the electro-production cross section $e(k)p(P) \to \chi_{cJ}(\Delta)e(k')p(P')$ through the standard formula [83–86]

$$\frac{\mathrm{d}\sigma_{ep}}{\mathrm{d}x_{\mathcal{P}}\mathrm{d}Q^2\mathrm{d}|t|} = \frac{\alpha}{2\pi Q^2 x_{\mathcal{P}}}\left\{2(1-y)\frac{\mathrm{d}\sigma_L}{\mathrm{d}|t|} + \left(1+(1-y)^2 - 2(1-y)\frac{Q^2_{\min}}{Q^2}\right)\frac{\mathrm{d}\sigma_T}{\mathrm{d}|t|}\right\}\ , \tag{60}$$

where $Q^2_{\min} = m_e^2 y^2/(1-y)$ accounts for projectile mass corrections, with $m_e$ being the electron mass. $y$ is the inelasticity, given by

$$y \equiv \frac{q\cdot P}{k\cdot P} = \frac{W^2+Q^2-m_N^2}{S-m_N^2}\ , \tag{61}$$

where $S \equiv (k+P)^2$ is the $ep$ collision energy, $W^2 \equiv (q+P)^2$ is the $\gamma^* p$ collision energy and $m_N$ is the proton mass. Recall that the Odderon amplitude evolves with Pomeron-$x_{\mathcal{P}}$, which is related to the conventional Bjorken-$x$ variable $x_B = Q^2/(2q\cdot P) = Q^2/(W^2+Q^2-m_N^2)$ via

$$x_{\mathcal{P}} \equiv \frac{x_B}{\beta}\ , \qquad \beta \equiv \frac{Q^2}{2q\cdot\ell} = \frac{Q^2}{Q^2+M_{\mathcal{H}}^2+|t|}\ . \tag{62}$$

In the numerical computations, we augment the Primakoff cross section taking into account the Pauli form factor $F_2$ as explained in Sec. III B. For $\chi_{c0}$ and $\chi_{c2}$ we take into account the available QCD corrections to the Primakoff cross section via (40). The total cross section is based on taking the coherent sum of the Primakoff and the Odderon exchange amplitudes in which case their relative sign becomes crucial. Our careful analysis in App. D reveals that the *Primakoff and Odderon amplitudes are in-phase*, that is, *they interfere constructively* for each value of the Odderon evolution parameter $x_{\mathcal{P}}$, as we have also explained in Sec. II D. The contribution of the $F_2$ does not interfere with the coherent sum of the Odderon and Primakoff amplitudes. The interference between the QCD correction to the Primakoff amplitude (available for $\chi_{c0,2}$) and the Odderon exchange is determined through the relative phase of the Odderon and the Primakoff amplitude at tree level. For the Odderon component we are using the solutions of the rcBK evolution, keeping only the $k=0$ harmonic of the Fourier series (26), as explained in Sec. II D. For the Primakoff component we use the recent fits of $F_1$ and $F_2$ from [77]. Whenever it is not explicitly stated, our results are based on the standard choice for the numerical value of $\alpha_S$: $\alpha_S(2m_c) \approx 0.25$.

We start with the numerical results for the $\gamma^* p \to \chi_{cJ} p$ cross sections shown in Fig. 3 as functions of $|t|$, for different values of $x_{\mathcal{P}}$. We have set $Q^2 = 0.01$ GeV$^2$. Since $Q^2$ is low, we focus on the transverse cross section $\mathrm{d}\sigma_T/\mathrm{d}|t|$. At small $|t|$ the cross section is, of course, dominated by the Primakoff process (photon exchange). However, we note that our predictions for the Primakoff cross sections are substantially lower than those shown in Fig. 4c of Ref. [87]. This emphasizes the importance of the constraint (40) from the two photon decay rate.

The QCD Odderon exchange amplitude reaches a comparable magnitude at $|t| \approx 1$ GeV$^2$, depending on $x_{\mathcal{P}}$ and $J$. The lowest cross-over from the Primakoff dominated to the Odderon dominated regime at $|t| < 1$ GeV$^2$ is seen for $\chi_{c1}$ and $\chi_{c2}$ where the Primakoff-like background is lower than for $\chi_{c0}$. In the regime of $|t|$ where individual Primakoff and Odderon cross sections are of similar magnitude, thanks to their constructive interference, the coherently summed cross section is four times greater than the Primakoff cross section alone. At high momentum transfers Odderon exchange dominates due to its slower fall-off with $|t|$.

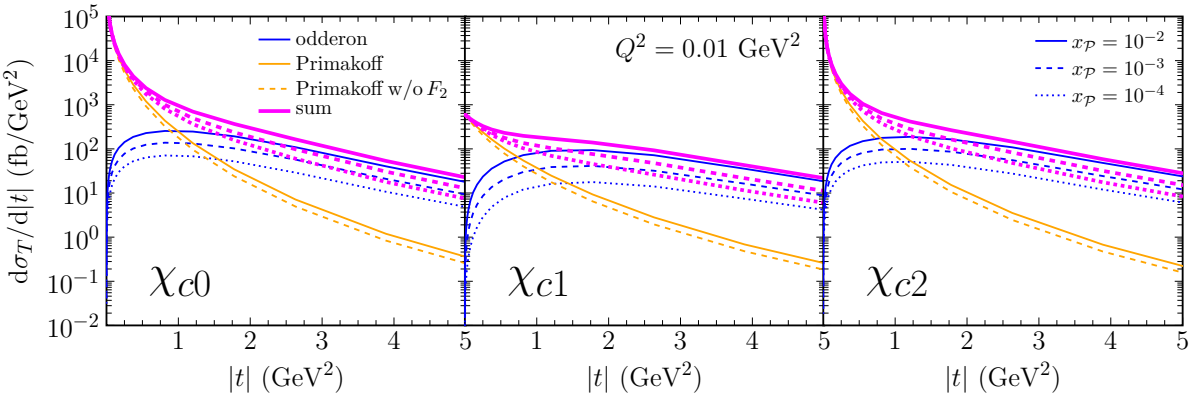



FIG. 3. The $\gamma^* p \to \chi_{cJ} p$ cross sections as functions of $|t|$ for the transverse photon. The label "sum" stands for the coherent sum of the odderon and the Primakoff contributions. We have set $Q^2 = 0.01$ GeV$^2$. Different line styles in the odderon and the summed cross section correspond to different values of $x_{\mathcal{P}}$. The comparison of the Primakoff contribution with the full and dashed orange lines emphasizes the impact of $F_2$.

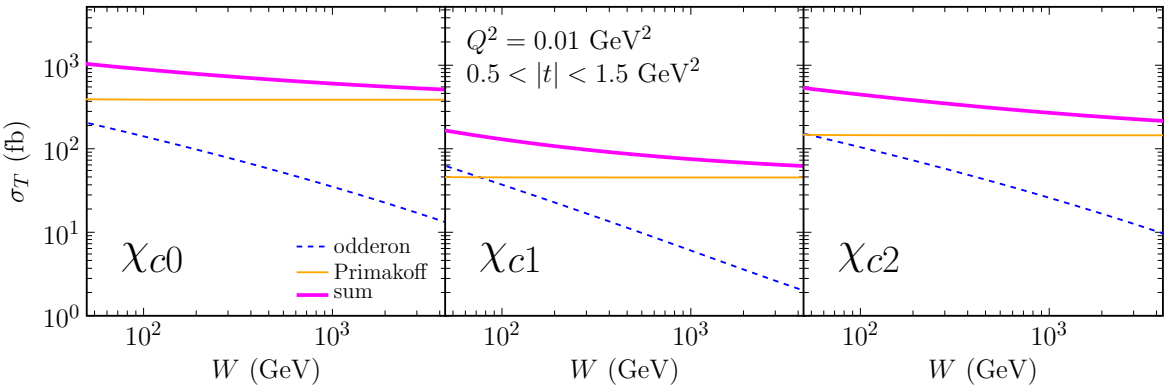



FIG. 4. The $\gamma^* p \to \chi_{cJ} p$ cross sections as functions of $W$ for the transverse photon. The label "sum" stands for the coherent sum of the Odderon and Primakoff contributions. We have set $Q^2 = 0.01$ GeV$^2$ and integrated in $|t|$ over $0.5 < |t| < 1.5$ GeV$^2$.

For Fig. 4 we have integrated $\mathrm{d}\sigma_T/\mathrm{d}|t|$ over the range $0.5 < |t| < 1.5$ GeV$^2$ as appropriate for the EIC design detector [43]. We show the result as a function of of $W$. As a consequence of small-$x$ evolution, the Odderon cross section drops with increasing $W$. With the Primakoff cross section alone being constant in our framework, at the lower end of $W$ the coherent sum is about five times greater than the Primakoff component alone. Interestingly, $\sigma_T$ displays a negative slope due to the decreasing Odderon amplitude towards smaller $x$.

We find that the decrease of the Odderon cross section with $W$ is driven mostly by the nonlinear corrections in the unitarized equation for the Odderon. In App. E (see Fig. 9, right) we have computed $\sigma_T$ based on linear evolution of the BLV Odderon and found a slower decrease with $W$ as one would anticipate from the asymptotics of the BLV solution with the intercept equal to one [9].

We also calculate the differential $ep \to \chi_{cJ} ep$ cross section for top EIC energy, $\sqrt{S} = 140$ GeV, and use the following kinematic cuts: $0.01 < y < 0.95$ [43], $x_{\mathcal{P}} < 0.01$, and $Q^2_{\mathrm{min}} < Q^2 < (2 \times 3.5\,\mathrm{GeV})^2$. The resulting $|t|$-dependence of the cross section is shown in Fig. 5. With these cuts the crossover from photon to Odderon exchange occurs at about $|t| \approx 1$ GeV$^2$ where the total cross section is several times greater than due to the Primakoff process alone.

The total $\chi_{cJ}$ cross sections are shown in Fig. 6. They have been integrated over $|t|$ over the range $0.5$ GeV$^2 < |t| < 1.5$ GeV$^2$ where according to Fig. 5 the Odderon contribution is appreciable. Both the photon and Odderon exchange contributions level off towards top EIC energy since our kinematic cuts are energy independent, and neither exchange involves a positive intercept.

$\chi_{c1}$ has the highest branching ratio $\mathrm{BR}(\chi_{c1} \to J/\psi\gamma) = 34.3\%$ [33]. As an example, lets estimate the total number of $\chi_{c1}$'s per month at the EIC. From Fig. 6 the total cross section is about $\sigma_{ep} \approx 20$ fb at $\sqrt{S} = 140$ GeV. Multiplying by the expected luminosity at the EIC ($L = 10^{34}$ cm$^{-2}$s$^{-1}$ $= 10^{-5}$ fb$^{-1}$ s$^{-1}$) gives about $2 \times 10^{-4}$ events/second or $\approx 518$ events/month. After taking into account $\mathrm{BR}(\chi_{c1} \to J/\psi\gamma)$ results in about 177 events/month. $J/\psi$'s are

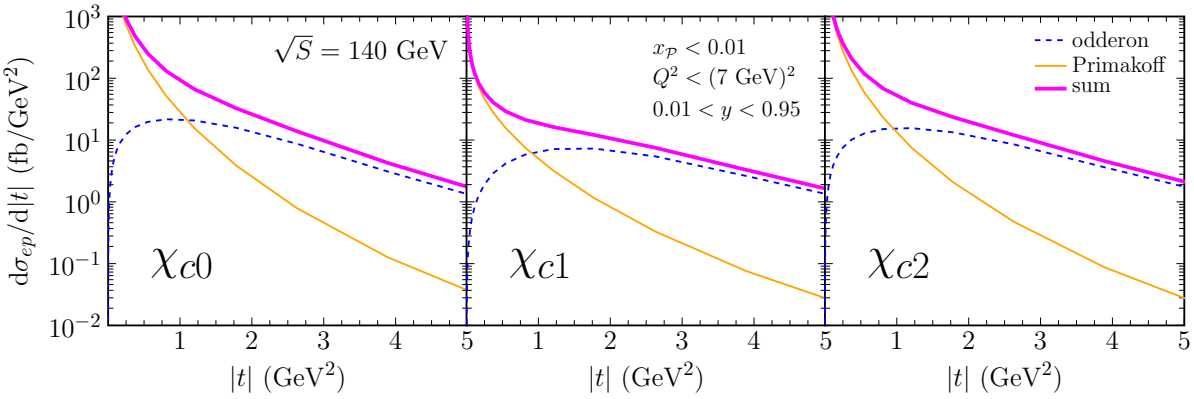



FIG. 5. The $ep \to \chi_{cJ} ep$ ($J = 0,\ 1,\ 2$) cross sections as functions of $|t|$ at the top EIC energy of $\sqrt{S} = 140$ GeV.

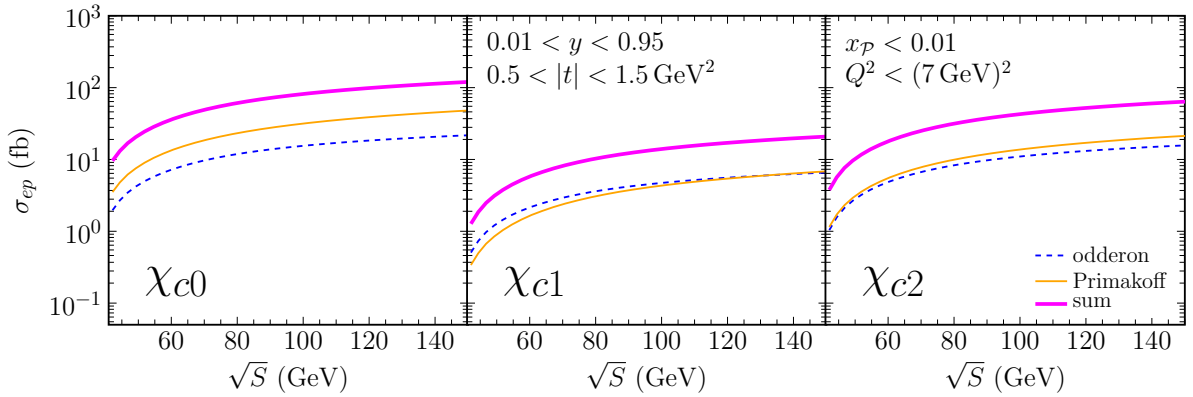



FIG. 6. Total electroproduction cross section of the $\chi_{cJ}$ quarkonia as a function of the $ep$ center of mass energy $\sqrt{S}$. Kinematic cuts are explained in the text.

detected through the $J/\psi \to e^+e^-$ or $\to \mu^+\mu^-$ decays, and the combined corresponding branching ratio is about $\mathrm{BR}(J/\psi \to \ell^+\ell^-) = 12\%$, after which we end up with about 21 events/month.

In the final Fig. 7, we summarize the expected number of exclusive $\chi_{cJ}$ events per month, $N_{\chi_{cJ}}$, at the EIC design luminosity for two ranges of momentum transfer $|t|$. The result is obtained according to:

$$N_{\chi_{cJ}} = L \times \sigma_{ep}(ep \to \chi_{cJ} ep) \times \mathrm{BR}(\chi_{cJ} \to J/\psi\gamma) \times \mathrm{BR}(J/\psi \to \ell^+\ell^-)\,. \tag{63}$$

The most statistics is expected for $\chi_{c1}$ and $\chi_{c2}$ tensor quarkonia with about a factor of 2-3 excess events over the Primakoff process in the interval $0.5\ \mathrm{GeV}^2 < |t| < 1.5\ \mathrm{GeV}^2$. For $\chi_{c0}$, which has the largest cross section from Fig. 6 is compensated by a small branching ratio $\mathrm{BR}(\chi_{c0} \to J/\psi\gamma) = 1.4\%$ [33].

The presented estimates of $\chi_{cJ}$ production cross sections carry some theoretical uncertainties. For the photon exchange contribution they are smaller than for the Odderon exchange. The coupling of the photon to the proton is well constrained experimentally [77], and so the main sources of uncertainty are the $\gamma^*\gamma^* \to \chi_{cJ}$ amplitudes. They depend on the details of the quarkonia wave functions and are sensitive to unknown higher order QCD corrections. For $\chi_{c0}$ and $\chi_{c2}$ the differential cross section $\mathrm{d}\sigma/\mathrm{d}|t|$ for the Primakoff process obeys a stringent constraint at small $|t| \ll 1\ \mathrm{GeV}^2$ imposed by Eq. (40) that holds at all orders in QCD. Hence, for these charmonia the uncertainties mentioned above affect mostly the details of the $t$-shape of $\mathrm{d}\sigma/\mathrm{d}|t|$. They are expected to be small as the $\chi_{cJ}$ wave functions are probed mostly at short distances $\sim 1/m_c$, where they are well constrained by the $\gamma\gamma$ decay width. The value of the coupling $\alpha_S(2m_c) \simeq 0.25$ is not large and the $t \to 0$ limit of the QCD corrections is known, so the uncertainties from QCD corrections at $t \neq 0$ should be small as well. For $\chi_{c1}$ the theoretical uncertainty of the Primakoff cross section is larger than for $\chi_{c0}$ and $\chi_{c2}$, as the constraint (40) cannot be imposed due to the LY theorem and, in consequence, the vanishing $\gamma\gamma$ decay width of $\chi_{c1}$. In this case measurements of the differential cross sections $\mathrm{d}\sigma/\mathrm{d}|t|$ for $|t| \ll 1\ \mathrm{GeV}^2$ where the photon exchange dominates should greatly reduce the uncertainty associated with this contribution at $|t| \sim 1\ \mathrm{GeV}^2$, where we expect to isolate a significant Odderon signal.

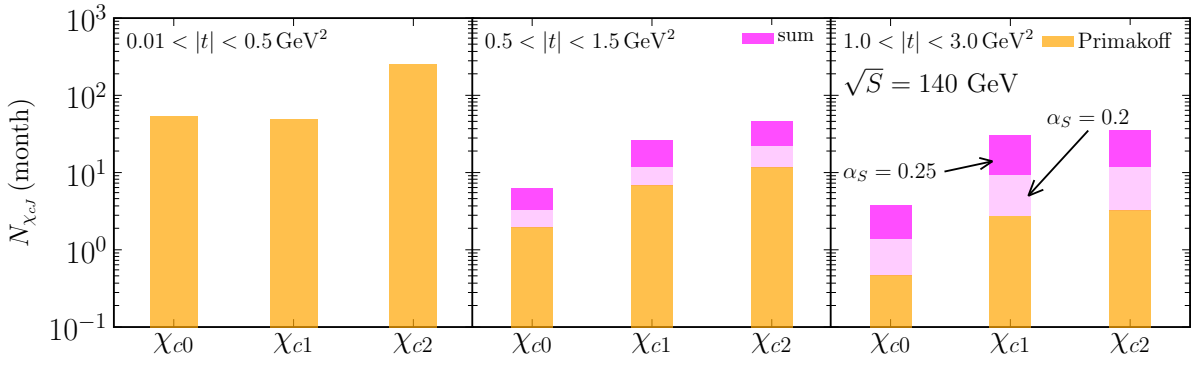



FIG. 7. A summary of expected Primakoff (orange) vs. total (magenta) number of exclusive $\chi_{cJ} \to J/\psi\gamma \to l^+l^-\gamma$ events per month at the top EIC energy of $\sqrt{S} = 140$ GeV, for three different $|t|$ cuts. The remaining kinematic cuts are the same as on Fig. 6. The two different bins for the total number of events correspond to $\alpha_S = 0.25$ and $\alpha_S = 0.2$, respectively.

In the Odderon exchange, in addition to uncertainties from unknown details of the quarkonia wave function and from higher order QCD corrections, there are uncertainties associated with the model for the proton wave function, and from the value of $\alpha_S$. The model of the proton employed here obeys general constraints coming from the measurements of the proton size, exclusive $J/\psi$ production off the proton [32] and open charm electroproduction at HERA [64]. Those constraints, however, originate from measurements in the $C$-even sector, and the emerging $C$-odd correlators have never been probed experimentally. Furthermore, the perturbative Odderon is strongly sensitive to the numerical value of $\alpha_S$ – the amplitude being $O(\alpha_S^3)$ at the lowest order. To quantify the uncertainty in the value of $\alpha_S$ we take an interval $0.2 < \alpha_S < 0.25$. The lower value $\alpha_S = 0.2$ comes from a fit of the proton and $C$-even exchange models to open charm electroproduction at HERA [64]. Therein the average value of the hard scale $\mu = \sqrt{Q^2 + 4m_c^2}$ is about 5 GeV, so the obtained value $\alpha_S = 0.2$ is fully consistent with the running of $\alpha_S(\mu)$. In our case the bulk of exclusive $\chi_{cJ}$ production in $ep$ collisions occurs at small photon virtualities and moderate $|t|$, and so $\alpha_S(2m_c) \approx 0.25$ is more appropriate. The range $0.2 < \alpha_S < 0.25$ results in an uncertainty of about a factor of 2.5 between the minimal and maximal values of the Odderon amplitude, when NLO effects in the proton wave function are included.

For the combined Primakoff and Odderon contributions, the $\alpha_S$-induced uncertainty is negligible for $|t| \ll 1$ GeV$^2$, where the Primakoff process dominates; whereas it about a factor of 2 when Primakoff and Odderon (with $\alpha_S = 0.25$) amplitudes are close to each other, and up to a factor of about 6 when the Odderon strongly dominates over the Primakoff channel. Since the uncertainty coming from $\alpha_S$ is sizeable, we assume that it is the dominant uncertainty for the Odderon exchange and take it as our estimate of the theoretical uncertainty. We consider the choice of $\alpha_S = 0.25$ to be realistic, as the scale $2m_c$ is already fairly high. Another choice is $\mu = m_c$, leading to $\alpha_S \approx 0.35$, as obtained for exclusive $J/\psi$ photoproduction using the same model for the proton at leading order [32]. Thus, we consider the lower value of $\alpha_S = 0.2$ to be a conservative choice.

## VI. SUMMARY AND CONCLUSIONS

In this paper we have derived amplitudes for exclusive scalar, axial-vector and tensor quarkonium production, Eqs. (22, 24), (28, 27), and (30, 29), respectively, in electron-proton scattering. We have provided first estimates of the cross sections for exclusive production of $\chi_{cJ}$ quarkonia with positive $C$-parity at the EIC. This process requires a $C$-odd exchange in the $t$-channel. In the limit of heavy quarks and/or high transverse momentum transfer this could be the exchange of a photon, a Primakoff process, or the exchange of a color-symmetric three gluon ladder, the Odderon. Our estimates suggest that for $|t| \approx 1$ GeV$^2$ the two amplitudes are of similar magnitude and that there are strong interference effects. Importantly, the relative phase is not affected by QCD evolution of the Odderon towards small $x$, and so it is determined by the three gluon exchange amplitude at moderately small $x_0$. In turn, the matrix element of the eikonal color current operator $d^{abc}J^{+a}J^{+b}J^{+c}$ at $x_0$, for transverse momenta of order $\max(R_{\chi_c}^{-1}, |t|^{1/2})$ ($R_{\chi_c}$ being here roughly the size of $\chi_c$), can be computed in a truncated Fock space for the proton that encompasses the states that are relevant in that kinematic regime.

We find that photon and Odderon exchanges interfere constructively, leading to an enhancement of the differential cross section for $\chi_{cJ}$ production around $|t| \approx 1$ GeV$^2$ by up to a factor of 4 over the pure photon exchange contribution. Given that both the normalization and the $t$-dependence of the Primakoff process are reasonably well determined,

this presents a very exciting opportunity to potentially discover at the EIC the hard $C$-odd exchange predicted by QCD. Furthermore, we find that towards top EIC energy, $\sqrt{S} = 140$ GeV, the total electroproduction cross section of $\chi_{cJ}$ quarkonia (with kinematic cuts specified in the previous section) levels off as neither photon nor Odderon exchange involve a positive intercept. Therefore, it would be important to measure the energy dependence [88] of the cross section from $\sqrt{S} = 50$ GeV up to top EIC energy $\sqrt{S} = 140$ GeV, where photon and Odderon amplitudes are similar, and where constructive interference of amplitudes leads to a total cross section for production of $\chi_{c1}$ and $\chi_{c2}$ quarkonia which exceeds the Primakoff component by a factor of $\approx 3 - 4$.

The predictions for the Odderon exchange process of course involve a number of uncertainties such as the matrix element in the proton of $d^{abc}J^{+a}J^{+b}J^{+c}$ or the value of the strong coupling $\alpha_S$. The associated cross section scales approximately like $\alpha_S^6$. Most importantly, the discovery of the hard QCD $C$-odd exchange requires fixing the normalization of the Primakoff background from measurements at low $\sqrt{|t|} < 0.5$ GeV; the relation (40) of the $t \to 0$ limit of the cross section to the $\gamma\gamma$ decay width provides an important constraint on theoretical predictions for this background. The $t$-dependence of the Primakoff cross section is then determined by the Dirac and Pauli electromagnetic form factors of the proton which are well known. Hence, one may then look for excess events above this background, and for a change of slope of the differential cross section $\mathrm{d}\sigma/\mathrm{d}t$, at higher $|t| \sim 1$ GeV$^2$ and beyond.

Our estimates indicate very weak dipole-proton (hard) scattering with $C$-odd exchange; this is largely since these amplitudes are also parity odd and vanish for impact parameter $\boldsymbol{b}_\perp \to 0$, unlike parity even amplitudes, as well as due to the fact that they fall off more rapidly towards large $\boldsymbol{b}_\perp$. The total electro-production cross section of $\chi_{cJ}$ quarkonia at top EIC energy is estimated to have a magnitude $10 - 20$ fb (for $J = 1$) and 60 fb (for $J = 2$). Thus, it was not possible to observe these processes at the HERA accelerator. However, in view of the projected high luminosity of the EIC, data collection over a time span of several months to a year may be sufficient for the discovery of the hard Odderon. A promising alternative would be to allow low-mass excitations of the proton, while requiring a large rapidity gap to the $C = +1$ quarkonia. Such rapidity gap, diffractive processes have greater cross-sections than exclusive ones. Furthermore, they would extend the reach to higher $|t|$ where Odderon exchange would more clearly dominate over photon exchange.

## ACKNOWLEDGMENTS

We thank Y. Hatta, H. Mäntysaari, L. Pentchev, and R. Venugopalan for useful discussions. The work of S. B. and A. K. is supported by the Croatian Science Foundation (HRZZ) no. 5332 (UIP-2019-04). A. D. acknowledges support by the DOE Office of Nuclear Physics through Grant DE-SC0002307, and The City University of New York for PSC-CUNY Research grant 65079-00 53. He also thanks the EIC Theory Institute at Brookhaven National Laboratory for their hospitality during a visit in July 2023 when initial work on this project was done. L. M. acknowledges the support of the Polish National Science Centre (NCN) grant no. 2017/27/B/ST2/02755. T. S. kindly acknowledges the support of the Polish National Science Center (NCN) Grants No. 2019/32/C/ST2/00202 and 2021/43/D/ST2/03375. S. B. thanks for the hospitality at the Jagiellonian University where part of this work was done.

### Appendix A: Computation of the light-cone wave functions

In order to compute the light-cone wave functions, the key element is a vertex contraction between spinors, e. g. $\bar{u}_h(k)\Gamma v_{\bar{h}}(k')$ ($\Gamma$ is some general Dirac vertex). For the spinors we use the LB basis, defined through

$$\begin{aligned} u_h(k) &= \frac{1}{2^{1/4}\sqrt{k^-}}\left(\sqrt{2}k^- + \gamma^0 m + \boldsymbol{\alpha}_\perp \cdot \boldsymbol{k}_\perp\right)\chi_h = \frac{1}{2^{1/4}\sqrt{k^-}}\left(\sqrt{2}k^- + \gamma^0 m + \not{k}_\perp \gamma^0\right)\chi_h\,, \\ v_h(k) &= \frac{1}{2^{1/4}\sqrt{k^-}}\left(\sqrt{2}k^- - \gamma^0 m + \boldsymbol{\alpha}_\perp \cdot \boldsymbol{k}_\perp\right)\chi_{-h} = \frac{1}{2^{1/4}\sqrt{k^-}}\left(\sqrt{2}k^- - \gamma^0 m + \not{k}_\perp \gamma^0\right)\chi_{-h}\,, \end{aligned} \tag{A1}$$

where $\alpha^i = \gamma^0\gamma^i$. In accordance with the conventions used in this work, we have replaced plus and minus light-cone coordinates in the above expression with respect to the original LB convention [46]. Thus the spinors $\chi_h$ are eigenstates of $\gamma^+\gamma^-$, namely

$$\frac{1}{2}\gamma^+\gamma^-\chi_h = \chi_h\,, \qquad \frac{1}{2}\gamma^-\gamma^+\chi_h = 0\,. \tag{A2}$$

Explicitly, we have $\chi_+ = (-1,0,1,0)/\sqrt{2}$, $\chi_- = (0,1,0,1)/\sqrt{2}$. The spinors (A1) can now be written as

$$\begin{aligned} u_h(k) &= \frac{1}{2^{1/4}\sqrt{k^-}}(\not{k} + m)\gamma^0\chi_h\,, \\ v_h(k) &= \frac{1}{2^{1/4}\sqrt{k^-}}(\not{k} - m)\gamma^0\chi_{-h}\,. \end{aligned} \tag{A3}$$

By defining a projection matrix

$$\chi_{h\bar{h}} \equiv \gamma^0\chi_{-\bar{h}}\bar{\chi}_h\gamma^0\,, \tag{A4}$$

one obtains

$$\bar{u}_h(k)\Gamma v_{\bar{h}}(k') = \mathrm{tr}\left[(\not{k} + m)\Gamma(\not{k}' - m)\chi_{h\bar{h}}\right]\,, \tag{A5}$$

and so the computation of the light-cone wave function comes down to the computation of the above Dirac trace. We use this method to compute all the wave functions considered in this work. Using (A4), the explicit forms of the projection matrices are found to be

$$\begin{aligned} \chi_{\pm\pm} &= \pm\frac{1}{2\sqrt{2}}\gamma^-(\gamma^1 \mp \mathrm{i}\gamma^2)\,, \\ \chi_{\pm\mp} &= \frac{1}{2\sqrt{2}}\gamma^-(1 \mp \gamma_5)\,. \end{aligned} \tag{A6}$$

## Appendix B: Derivation of the amplitudes for axial and tensor quarkonia

In this appendix we present the steps of the derivation of the amplitudes from Sec. II B for the axial vector and tensor quarkonia.

In the case of axial quarkonia we plug the second line of (9) into the traces in (19) to obtain

$$\begin{aligned} &A_{\lambda=0,\bar{\lambda}=0}(\boldsymbol{l}_\perp, \boldsymbol{l}_{1\perp}, z) = 0\,, \\ &A_{\lambda=0,\bar{\lambda}=\pm}(\boldsymbol{l}_\perp, \boldsymbol{l}_{1\perp}, z) = 4Qz\bar{z}(\boldsymbol{\epsilon}_\perp^{\bar{\lambda}*} \times \boldsymbol{l}_{1\perp})\,, \\ &A_{\lambda=\pm,\bar{\lambda}=0}(\boldsymbol{l}_\perp, \boldsymbol{l}_{1\perp}, z) = \frac{4}{M_{\mathcal{A}}}\left(m_c^2(\boldsymbol{\epsilon}_\perp^{\lambda} \times \boldsymbol{l}_{1\perp}) + \boldsymbol{l}_{1\perp}^2(\boldsymbol{\epsilon}_\perp^{\lambda} \times \boldsymbol{l}_\perp)\right)\,, \\ &A_{\lambda=\pm,\bar{\lambda}=\pm}(\boldsymbol{l}_\perp, \boldsymbol{l}_{1\perp}, z) = 2(z-\bar{z})\left((\boldsymbol{\epsilon}_\perp^{\lambda} \cdot \boldsymbol{\epsilon}_\perp^{\bar{\lambda}*})(\boldsymbol{l}_\perp \times \boldsymbol{l}_{1\perp}) - (\boldsymbol{l}_\perp \cdot \boldsymbol{l}_{1\perp} + m_c^2)(\boldsymbol{\epsilon}_\perp^{\lambda} \times \boldsymbol{\epsilon}_\perp^{\bar{\lambda}*})\right)\,, \end{aligned} \tag{B1}$$

The last line is understood to contain four different combinations of photon and quarkonia transverse polarizations. We used the following notations for the 2D cross products: $\boldsymbol{x}_\perp \times \boldsymbol{y}_\perp \equiv x^1y^2 - x^2y^1 = -\epsilon^{+-\boldsymbol{x}_\perp\boldsymbol{y}_\perp}$. The non-zero reduced amplitudes are

$$\begin{aligned} \mathcal{A}_{\lambda=0,\bar{\lambda}=\pm}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) &= eq_c\mathrm{e}^{-\mathrm{i}\bar{\lambda}\phi_r}\int_z \mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_{LT}(r_\perp)\,, \\ \mathcal{A}_{\lambda=\pm,\bar{\lambda}=0}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) &= eq_c\mathrm{e}^{\mathrm{i}\lambda\phi_r}\int_z \mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_{TL}(r_\perp)\,, \\ \mathcal{A}_{\lambda=\pm,\bar{\lambda}=\pm}(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) &= eq_c\lambda\delta_{\lambda\bar{\lambda}}\int_z \mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_{TT}(r_\perp)\,, \end{aligned} \tag{B2}$$

where $\mathcal{A}_A$ are given in Eqs. (28). Notice that the first term from the fourth line of (B1) is proportional to $\boldsymbol{r}_\perp \times \boldsymbol{r}_\perp = 0$ in coordinate space and so it does not contribute to $\mathcal{A}_{TT}(r_\perp)$. In the last line of Eq. (B2), out of four possible combinations of photon and axial quarkonia transverse polarizations, only the two polarization preserving transitions survive since this amplitude is proportional to the 2D cross product, $\boldsymbol{\epsilon}_\perp^{\lambda} \times \boldsymbol{\epsilon}_\perp^{\bar{\lambda}*} = (-\mathrm{i}/2)(\lambda + \bar{\lambda}) = -\mathrm{i}\lambda\delta_{\lambda\bar{\lambda}}$ of the polarization vectors of the incoming photon and outgoing axial vector quarkonia. Accordingly, the two polarization flipping transitions vanish, unlike in case of vector quarkonia.

Computing now the $\phi_r$ integral we separate the helicity dependence and obtain the amplitudes

$$\begin{aligned} \left\langle\mathcal{M}_{\lambda=0,\bar{\lambda}=\pm1}\right\rangle &\equiv q^-\widetilde{\mathcal{M}}_{0,\bar{\lambda}=\pm1} \equiv q^-\mathrm{e}^{-\mathrm{i}\bar{\lambda}\phi_\Delta}\widetilde{\mathcal{M}}_{LT}\,, \\ \left\langle\mathcal{M}_{\lambda=\pm1,\lambda=0}\right\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=\pm1,0} \equiv q^-\mathrm{e}^{\mathrm{i}\lambda\phi_\Delta}\widetilde{\mathcal{M}}_{TL}\,, \\ \left\langle\mathcal{M}_{\lambda=\pm1,\bar{\lambda}=\pm1}\right\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=\pm1,\bar{\lambda}=\pm1} \equiv q^-\lambda\delta_{\lambda\bar{\lambda}}\widetilde{\mathcal{M}}_{TT}\,, \end{aligned} \tag{B3}$$

where $\widetilde{\mathcal{M}}_B$ are given in Eqs. (27).

For the tensor quarkonia, we proceed along similar steps as above, only this time there is a total of 15 overlaps that need to be computed. Starting from the traces (19), we plug in the third line of (9) and find

$$\begin{aligned}
A_{\lambda=0,\bar\lambda=\pm2}(\boldsymbol{l}_\perp,\boldsymbol{l}_{1\perp},z) &= -8Qz\bar z(z-\bar z)(\boldsymbol{\epsilon}_\perp^{\pm *}\cdot\boldsymbol{l}_{1\perp})^2\,,\\
A_{\lambda=\pm1,\bar\lambda=\pm2}(\boldsymbol{l}_\perp,\boldsymbol{l}_{1\perp},z) &= -4(\boldsymbol{\epsilon}_\perp^{\pm *}\cdot\boldsymbol{l}_{1\perp})\big[(\boldsymbol{\epsilon}_\perp^{\lambda}\cdot\boldsymbol{\epsilon}_\perp^{\pm *})(\boldsymbol{l}_\perp\cdot\boldsymbol{l}_{1\perp}+m_c^2)\\
&\quad+(z-\bar z)^2(\boldsymbol{\epsilon}_\perp^{\lambda}\cdot\boldsymbol{l}_\perp)(\boldsymbol{\epsilon}_\perp^{\pm *}\cdot\boldsymbol{l}_{1\perp})-(\boldsymbol{\epsilon}_\perp^{\lambda}\cdot\boldsymbol{l}_{1\perp})(\boldsymbol{\epsilon}_\perp^{\pm *}\cdot\boldsymbol{l}_\perp)\big]\,,\\
A_{\lambda=0,\bar\lambda=\pm1}(\boldsymbol{l}_\perp,\boldsymbol{l}_{1\perp},z) &= 2\sqrt{2}QM_{\mathcal{T}}z\bar z(3-4(z^2+\bar z^2))(\boldsymbol{\epsilon}_\perp^{\bar\lambda *}\cdot\boldsymbol{l}_{1\perp})\,,\\
A_{\lambda=\pm1,\bar\lambda=\pm1}(\boldsymbol{l}_\perp,\boldsymbol{l}_{1\perp},z) &= -\sqrt{2}M_{\mathcal{T}}(z-\bar z)\big[(\boldsymbol{\epsilon}_\perp^{\lambda}\cdot\boldsymbol{\epsilon}_\perp^{\bar\lambda *})(\boldsymbol{l}_\perp\cdot\boldsymbol{l}_{1\perp}+m_c^2)\\
&\quad-(3-4(z^2+\bar z^2))(\boldsymbol{\epsilon}_\perp^{\lambda}\cdot\boldsymbol{l}_\perp)(\boldsymbol{\epsilon}_\perp^{\bar\lambda *}\cdot\boldsymbol{l}_{1\perp})-(\boldsymbol{\epsilon}_\perp^{\lambda}\cdot\boldsymbol{l}_{1\perp})(\boldsymbol{\epsilon}_\perp^{\bar\lambda *}\cdot\boldsymbol{l}_\perp)\big]\,,\\
A_{\lambda=0,\bar\lambda=0}(\boldsymbol{l}_\perp,\boldsymbol{l}_{1\perp},z) &= \frac{8\sqrt{2}Q}{\sqrt{3}}z\bar z(z-\bar z)\left(\frac{3}{2}\boldsymbol{l}_{1\perp}^2+m_c^2\right)\,,\\
A_{\lambda=\pm1,\bar\lambda=0}(\boldsymbol{l}_\perp,\boldsymbol{l}_{1\perp},z) &= \frac{2\sqrt{2}}{\sqrt{3}}\left[2(z-\bar z)^2(\boldsymbol{\epsilon}_\perp^{\lambda}\cdot\boldsymbol{l}_\perp)\left(\frac{3}{2}\boldsymbol{l}_{1\perp}^2+m_c^2\right)+(\boldsymbol{\epsilon}_\perp^{\lambda}\cdot\boldsymbol{l}_{1\perp})m_c^2\right].
\end{aligned}\tag{B4}$$

The results for $\bar\lambda=0$ have been simplified using

$$(\boldsymbol{\epsilon}_\perp^{+*}\cdot\boldsymbol{u}_\perp)(\boldsymbol{\epsilon}_\perp^{-*}\cdot\boldsymbol{v}_\perp)+(\boldsymbol{\epsilon}_\perp^{-*}\cdot\boldsymbol{u}_\perp)(\boldsymbol{\epsilon}_\perp^{+*}\cdot\boldsymbol{v}_\perp)=-\boldsymbol{u}_\perp\cdot\boldsymbol{v}_\perp\,,\tag{B5}$$

which holds for general 2D vectors $\boldsymbol{u}_\perp$ and $\boldsymbol{v}_\perp$. Inserting (B4) in (18) we obtain the reduced amplitudes

$$\begin{aligned}
\mathcal{A}_{\lambda=0,\bar\lambda=\pm2}(\boldsymbol{r}_\perp,\boldsymbol{\Delta}_\perp) &= eq_c\mathrm{e}^{-\mathrm{i}\bar\lambda\phi_r}\int_z\mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_{LT2}(r_\perp)\,,\\
\mathcal{A}_{\lambda=\pm1,\bar\lambda=\pm2}(\boldsymbol{r}_\perp,\boldsymbol{\Delta}_\perp) &= eq_c\lambda\mathrm{e}^{\mathrm{i}(\lambda-\bar\lambda)\phi_r}\int_z\mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\left(\delta_{\lambda,\bar\lambda/2}\mathcal{A}_{TT2p}(r_\perp)+\delta_{\lambda,-\bar\lambda/2}\mathcal{A}_{TT2f}(r_\perp)\right)\,,\\
\mathcal{A}_{\lambda=0,\bar\lambda=\pm1}(\boldsymbol{r}_\perp,\boldsymbol{\Delta}_\perp) &= eq_c\bar\lambda\mathrm{e}^{-\mathrm{i}\bar\lambda\phi_r}\int_z\mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_{LT}(r_\perp)\,,\\
\mathcal{A}_{\lambda=\pm1,\bar\lambda=\pm1}(\boldsymbol{r}_\perp,\boldsymbol{\Delta}_\perp) &= eq_c\mathrm{e}^{\mathrm{i}(\lambda-\bar\lambda)\phi_r}\int_z\mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\left(\delta_{\lambda\bar\lambda}\mathcal{A}_{TTp}(r_\perp)-\delta_{\lambda,-\bar\lambda}\mathcal{A}_{TTf}(r_\perp)\right)\,,\\
\mathcal{A}_{\lambda=0,\bar\lambda=0}(\boldsymbol{r}_\perp,\boldsymbol{\Delta}_\perp) &= eq_c\int_z\mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_{LL}(r_\perp)\,,\\
\mathcal{A}_{\lambda=\pm1,\bar\lambda=0}(\boldsymbol{r}_\perp,\boldsymbol{\Delta}_\perp) &= eq_c\lambda\mathrm{e}^{\mathrm{i}\lambda\phi_r}\int_z\mathrm{e}^{-\mathrm{i}\boldsymbol{\delta}_\perp\cdot\boldsymbol{r}_\perp}\mathcal{A}_{TL}(r_\perp)\,.
\end{aligned}\tag{B6}$$

From (B6) we see that for the transverse polarizations of the photon and the tensor quarkonia both the polarization preserving ($\lambda\to\bar\lambda/2$, for the case $\bar\lambda=\pm2$ and $\lambda\to\bar\lambda$ for the case $\bar\lambda=\pm1$) and polarization flipped ($\lambda\to-\bar\lambda/2$, $\lambda\to-\bar\lambda$) transitions are allowed, which explains the notation ($p=$ preserving, $f=$ flipped).

In the final step we compute the amplitudes

$$\begin{aligned}
\langle\mathcal{M}_{\lambda=0,\bar\lambda=\pm2}\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=0,\bar\lambda=\pm2}\equiv q^-\mathrm{e}^{-\mathrm{i}\bar\lambda\phi_\Delta}\widetilde{\mathcal{M}}_{LT2}\,,\\
\langle\mathcal{M}_{\lambda=\pm1,\bar\lambda=\pm2}\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=\pm1,\bar\lambda=\pm2}\equiv q^-\lambda\mathrm{e}^{-\mathrm{i}\lambda\phi_\Delta}\delta_{\lambda,\bar\lambda/2}\widetilde{\mathcal{M}}_{TT2p}+q^-\lambda\mathrm{e}^{3\mathrm{i}\lambda\phi_\Delta}\delta_{\lambda,-\bar\lambda/2}\widetilde{\mathcal{M}}_{TT2f}\,,\\
\langle\mathcal{M}_{\lambda=0,\bar\lambda=\pm1}\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=0,\bar\lambda=\pm1}\equiv q^-\mathrm{e}^{-\mathrm{i}\bar\lambda\phi_\Delta}\widetilde{\mathcal{M}}_{LT}\,,\\
\langle\mathcal{M}_{\lambda=\pm1,\bar\lambda=\pm1}\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=\pm1,\bar\lambda=\pm1}\equiv q^-\delta_{\lambda\bar\lambda}\widetilde{\mathcal{M}}_{TTp}+q^-\delta_{\lambda,-\bar\lambda}\mathrm{e}^{2\mathrm{i}\lambda\phi_\Delta}\widetilde{\mathcal{M}}_{TTf}\,,\\
\langle\mathcal{M}_{\lambda=0,\bar\lambda=0}\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=0,\bar\lambda=0}\equiv q^-\widetilde{\mathcal{M}}_{LL}\,,\\
\langle\mathcal{M}_{\lambda=\pm1,\bar\lambda=0}\rangle &\equiv q^-\widetilde{\mathcal{M}}_{\lambda=\pm1,\bar\lambda=0}\equiv q^-(\delta_{\lambda1}-\delta_{\lambda,-1})\mathrm{e}^{\mathrm{i}\lambda\phi_\Delta}\widetilde{\mathcal{M}}_{TL}\,,
\end{aligned}\tag{B7}$$

where also the explicit helicity depndence was factored out. The scalar functions $\widetilde{\mathcal{M}}_B$ are given by Eqs. (29).

## Appendix C: The Primakoff contribution in specific kinematic limits

### 1. The proof of formula (40)

We begin by considering the amplitude for the decay process $\mathcal{H}(\bar{\lambda},\Delta)\to\gamma(\lambda,q)\gamma(\lambda',\ell)$, where $\mathcal{H}=\mathcal{S},\mathcal{T}$. In what follows, it will be useful to separate out the exchanged photon polarization from the amplitude as

$$\mathcal{M}_{\lambda\lambda'\bar{\lambda}}(\mathcal{H}\to\gamma\gamma)=\epsilon_\mu(\lambda',\ell)\mathcal{M}^\mu_{\lambda\bar{\lambda}}(\mathcal{H}\to\gamma\gamma)\,. \tag{C1}$$

Inserting (C1) into the standard formula for the decay width [33] we get

$$\begin{aligned}\Gamma(\mathcal{H}\to\gamma\gamma)&=\frac{1}{2J+1}\frac{1}{32\pi M_{\mathcal{H}}}\sum_{\bar{\lambda}=-J}^{J}\sum_{\lambda\lambda'}|\mathcal{M}_{\lambda\lambda'\bar{\lambda}}(\mathcal{H}\to\gamma\gamma)|^2\,,\\&=-\frac{1}{2J+1}\frac{1}{32\pi M_{\mathcal{H}}}\sum_{\bar{\lambda}=-J}^{J}\sum_{\lambda}g_{ii'}\mathcal{M}^i_{\lambda\bar{\lambda}}(\mathcal{H}\to\gamma\gamma)\mathcal{M}^{i'*}_{\lambda\bar{\lambda}}(\mathcal{H}\to\gamma\gamma)\,.\end{aligned} \tag{C2}$$

Next, we consider the amplitude for exclusive photoproduction $\mathcal{M}_{\lambda\bar{\lambda}}(\gamma p\to\mathcal{H}p)$, see (46). The cross section reads [33]

$$\frac{\mathrm{d}\sigma_{\mathcal{H}}(\gamma p\to\mathcal{H}p)}{\mathrm{d}|t|}=\frac{1}{2}\sum_{\bar{\lambda}=-J}^{J}\sum_{\lambda}\int\frac{\mathrm{d}\phi_\ell}{2\pi}\frac{1}{16\pi W^4}\left|\mathcal{M}_{\lambda\bar{\lambda}}(\gamma p\to\mathcal{H}p)\right|^2\,. \tag{C3}$$

where $W^2=(q+P)^2\simeq 2q\cdot P\gg M_{\mathcal{H}}^2$, $P^2$, $|\ell^2|$ in the high energy limit[12]. Moreover $\ell\approx(x_{\mathcal{P}}P^+,0,\boldsymbol{\ell}_\perp)$ with $\phi_\ell$ being the respective azimuthal angle. We also have $x_{\mathcal{P}}\approx M_{\mathcal{H}}^2/W^2$ (see Eq. (62) below), and $\ell^2=t\approx-\boldsymbol{\ell}_\perp^2$.

Taking the high energy limit leads to (47). Here we pick up the leading $F_1$ term when $W^2\gg|t|$

$$\mathcal{M}_{\lambda\bar{\lambda}}(\gamma p\to\mathcal{H}p)=P_\mu\,\mathcal{M}^\mu_{\lambda\bar{\lambda}}(\gamma\gamma\to\mathcal{H})\frac{eF_1(\ell_\perp)}{\ell^2}\,\frac{1}{q\cdot P}\bar{u}(P')\not{q}u(P)\,, \tag{C4}$$

To make the connection (40) we rewrite (C4) using QED gauge invariance $\ell_\mu\mathcal{M}^\mu_{\lambda\bar{\lambda}}(\gamma\gamma\to\mathcal{H})=0$ in the high-energy limit (the Collins-Ellis trick [89])

$$P_\mu\mathcal{M}^\mu_{\lambda\bar{\lambda}}(\gamma\gamma\to\mathcal{H})\approx-\frac{1}{x_{\mathcal{P}}}\ell_i\mathcal{M}^i_{\lambda\bar{\lambda}}(\gamma\gamma\to\mathcal{H})\,. \tag{C5}$$

so that

$$\mathcal{M}_{\lambda\bar{\lambda}}(\gamma p\to\mathcal{H}p)=-\frac{2eF_1(\ell_\perp)}{t}\frac{W^2}{M_{\mathcal{H}}^2}\ell_i\mathcal{M}^i_{\lambda\bar{\lambda}}(\gamma\gamma\to\mathcal{H}). \tag{C6}$$

Finally, the result in (C3) is isotropic in $\phi_\ell$ after the polarization sum. We can perform the angular average leading to

$$\begin{aligned}\frac{1}{2}\sum_{\bar{\lambda}=-J}^{J}\sum_{\lambda}\int\frac{\mathrm{d}\phi_\ell}{2\pi}\left|\ell_i\mathcal{M}^i_{\lambda\bar{\lambda}}(\gamma\gamma\to\mathcal{H})\right|^2&=-\frac{1}{4}\boldsymbol{\ell}_\perp^2\sum_{\bar{\lambda}=-J}^{J}\sum_{\lambda}g_{ii'}\mathcal{M}^i_{\lambda\bar{\lambda}}(\gamma\gamma\to\mathcal{H})\mathcal{M}^{i'*}_{\lambda\bar{\lambda}}(\gamma\gamma\to\mathcal{H})\\&=(2J+1)8\pi M_{\mathcal{H}}\boldsymbol{\ell}_\perp^2\Gamma(\mathcal{H}\to\gamma\gamma)\,,\end{aligned} \tag{C7}$$

where in the last line we used (C2). Inserting (C7) into (C3) via (C6) gives Eq. (40), which is the desired result. In the derivation we assumed the scattering off a proton target, but the formula is valid for any charged particle.

---

[12] The amplitudes $\mathcal{M}_{\lambda\bar{\lambda}}$ in this section and the color averaged amplitudes $\widetilde{\mathcal{M}}_{\lambda\bar{\lambda}}$ as obtained through Eqs. (23), (B3) and (B7) are related through $\widetilde{\mathcal{M}}_{\lambda\bar{\lambda}}=\mathcal{M}_{\lambda\bar{\lambda}}/W^2$, compare (C3) with (31).

### 2. The $|t| \to 0$ limit for axial vector quarkonia and its connection to the Landau-Yang theorem

Thanks to the Collins-Ellis trick, the amplitude $\gamma^*\gamma^* \to \mathcal{H}$ for scalar and the tensor quarkonia has a finite $O(\Delta_\perp)$ contribution. The Collins-Ellis trick is a rather general and robust procedure in the high-energy limit and so from this perspective the case of axial quarkonia seems special in that the $O(\Delta_\perp)$ contribution vanishes, with the amplitude scaling as $O(\Delta_\perp^2)$, check (43). In the following we demonstrate that $O(\Delta_\perp^2)$ scaling follows from a general argument and not relying to the choice of axial quarkonia wave function. We take Eq. (37) as a starting point and further factor out the polarization vectors from the helicity amplitude as

$$n_\nu \mathcal{M}^\nu_{\lambda\bar\lambda}(\gamma^*\gamma^* \to \mathcal{A}) = \epsilon_\mu(\lambda, q) n_\nu E_\rho(\bar\lambda, \Delta_0) \mathcal{M}^{\mu\nu\rho}(\gamma^*\gamma^* \to \mathcal{A})\,. \tag{C8}$$

The covariant amplitude $\mathcal{M}^{\mu\nu\rho}(\gamma^*\gamma^* \to \mathcal{A})$ depends only on the vectors $q^\mu$ and $\ell^\nu$. Because $\mathcal{A}$ is an axial vector quarkonia, $\mathcal{M}^{\mu\nu\rho}(\gamma^*\gamma^* \to \mathcal{A})$ has to be proportional to the $\epsilon$-tensor. In general the covariant decomposition is as follows

$$\begin{aligned}\mathcal{M}^{\mu\nu\rho}(\gamma^*\gamma^* \to \mathcal{A}) = &\left(q^\rho - \ell^\rho + \frac{-q^2+\ell^2}{(q+\ell)^2}(q^\rho + \ell^\rho)\right)\epsilon^{\mu\nu q\ell} M_{\mathcal{A}} F_{TT}(q^2, \ell^2) \\ &+ \left(\ell^\mu - \frac{q\cdot\ell}{q^2} q^\mu\right)\epsilon^{\nu\rho q\ell}\sqrt{-q^2} F_{LT}(q^2,\ell^2) + \left(q^\nu - \frac{q\cdot\ell}{\ell^2}\ell^\nu\right)\epsilon^{\mu\rho q\ell}\sqrt{-\ell^2} F_{TL}(q^2,\ell^2)\,,\end{aligned} \tag{C9}$$

see e. g. [59] and references therein. The decomposition (C9) is QED gauge invariant, that is: $q_\mu \mathcal{M}^{\mu\nu\rho}(\gamma^*\gamma^* \to \mathcal{A}) = 0$ and $\ell_\nu \mathcal{M}^{\mu\nu\rho}(\gamma^*\gamma^* \to \mathcal{A}) = 0$, and we also have $(q^\rho + l^\rho)\mathcal{M}_{\mu\nu\rho}(\gamma^*\gamma^* \to \mathcal{A}) = 0$. Here the notation $F_{TT}$, $F_{TL}$ and $F_{LT}$ refers to the $\gamma^*\gamma^*$ polarizations in the center of mass frame. Here only the form-factor $F_{TT}$ is constrained by the LY theorem: $F_{TT}(0,0) = 0$ [75, 76], see also ref. [59]. Due to the Bose symmetry of the full amplitude (C9) we must have $F_{TT}(q^2,\ell^2) = -F_{TT}(\ell^2, q^2)$ and so $F_{TT}(q^2,\ell^2) \propto q^2 - \ell^2$ for small $q^2$ and $\ell^2$. On the other hand, the form-factors $F_{LT}$ and $F_{TL}$ scale as $F_{LT}(q^2,\ell^2) \propto \sqrt{-q^2}$, $F_{TL}(q^2,\ell^2) \propto \sqrt{-\ell^2}$ in order to avoid kinematic singularities [59, 90].

Contracting (C9) with the longitudinal polarization vector $E^\rho(0,\Delta_0)$, only the piece proportional to the gauge vector $n^\rho$ from (12) survives. Further contracting with the transverse photon polarization $\epsilon^\mu(\lambda = \pm 1, q)$ and the gauge vector $n^\nu$ we obtain

$$n_\nu \mathcal{M}^\nu_{\lambda=\pm1,\bar\lambda=0}(\gamma^*\gamma^* \to \mathcal{A}) = -\frac{M_0}{\Delta_0\cdot n}(q\cdot n)\epsilon^{\epsilon(\lambda,q)nq\ell}\left[\left(1 - \frac{q^2-\ell^2}{\Delta_0^2}\right) M_{\mathcal{A}} F_{TT}(q^2,\ell^2) + \sqrt{-\ell^2} F_{TL}(q^2,\ell^2)\right], \tag{C10}$$

where we have used $\ell\cdot n = 0$. The form-factor $F_{LT}(q^2,\ell^2)$ decouples after contracting $\epsilon^{\nu\rho ql}$ with $n_\nu E_\rho(0,\Delta_0)$. Taking into account the scalings of the form-factors the square bracket in (C10) is $O(\Delta_\perp^2)$ while the $\epsilon$-tensor is $O(\Delta_\perp)$ leading to the overall $O(\Delta_\perp^3)$ scaling.

For the transverse polarization we find

$$\begin{aligned}n_\nu \mathcal{M}^\nu_{\lambda=\pm1,\bar\lambda=\pm1}(\gamma^*\gamma^* \to \mathcal{A}) = &(q-l)\cdot E(\bar\lambda,\Delta_0)\epsilon^{\epsilon(\lambda,q)nq\ell} M_{\mathcal{A}} F_{TT}(q^2,\ell^2) + (\epsilon(\lambda,q)\cdot\ell)\epsilon^{nE(\bar\lambda,\Delta_0)q\ell}\sqrt{-q^2} F_{LT}(q,\ell^2) \\ &+ (q\cdot n)\epsilon^{\epsilon(\lambda,q)E(\bar\lambda,\Delta_0)q\ell}\sqrt{-\ell^2} F_{TL}(q^2,\ell^2)\,.\end{aligned} \tag{C11}$$

The first term in (C11) is $O(\Delta_\perp^3)$, the second term decouples in the $q^2 \to 0$ limit and the third term is $O(\Delta_\perp^2)$ which is the leading contribution in the $\Delta_\perp \to 0$ limit.

### 3. The NRQCD limit of the Primakoff cross sections

In this subsection we consider the non-relativisitic QCD (NRQCD) limit of the Primakoff process by taking the heavy quark limit. We also take $Q^2 \to 0$ and focus on transverse photon - proton scattering.

In the NRQCD limit all the momenta in the amplitude are considered to be much smaller than the heavy quark mass. Since the results obtained in the previous Secs. C 1 and C 2 have already been obtained in the $|t| \to 0$ limit, we take these as starting point and further expand to second order in $\xi = z - 1/2$ and $\boldsymbol{l}_\perp$.

For the scalar $\mathcal{S}$ quarkonia using (42) we obtain

$$\widetilde{\mathcal{M}}_T \to 8\sqrt{2}\pi \mathrm{i} q_c^2 \alpha N_c \frac{eF_1(\Delta_\perp)}{\Delta_\perp}\frac{1}{m_c^3}\int_z\int_{\boldsymbol{l}_\perp}(\boldsymbol{l}_\perp^2 + 4\xi^2 m_c^2)\frac{\phi_{\mathcal{S}}(\boldsymbol{l}_\perp, z)}{z\bar z}\,. \tag{C12}$$

In the next step we replace the LCWF $\phi_{\mathcal{S}}(\boldsymbol{l}_\perp, z)$ by the non-relativistic radial wave function $u(l)$. Here $\boldsymbol{l} = (\boldsymbol{l}_\perp, l^3)$ and $l$ is its modulus, with $l^3 = 2\xi m_c$ as appropriate for the NRQCD limit. The correspondence is given by [56]

$$\int_z \int_{\boldsymbol{l}_\perp} \frac{\phi_{\mathcal{S}}(\boldsymbol{l}_\perp, z)}{z\bar{z}} \to \frac{1}{16\pi^2\sqrt{m_c}} \frac{1}{\sqrt{N_c}} \int \mathrm{d}^3\boldsymbol{l} \frac{u(l)}{\boldsymbol{l}^2} \,. \tag{C13}$$

After performing the angular integrations in Eq. (C12) we relate $u(l)$ to the derivative of the radial wave function at the origin, as in Eq. (3.18) of Ref. [56]:

$$\int_0^\infty \mathrm{d}l\, l^2 u(l) = 3\sqrt{\frac{\pi}{2}} R'(0) \,. \tag{C14}$$

This leads to

$$\widetilde{\mathcal{M}}_T \to 12\sqrt{\pi}\mathrm{i} q_c^2 \alpha \sqrt{N_c} \frac{eF_1(\Delta_\perp)}{\Delta_\perp} \frac{1}{m_c^{5/2}} R'(0) \,. \tag{C15}$$

The cross section is obtained from standard formulas, see Eq. (32). We find

$$\frac{\mathrm{d}\sigma(\gamma p \to \mathcal{S} p)}{\mathrm{d}|t|} \to \frac{9\pi q_c^4 \alpha^3 N_c |R'(0)|^2 F_1^2(0)}{m_c^7 |t|} \,. \tag{C16}$$

The result agrees with Eq. (8b) of Jia *et al.* [87].

For the axial vector quarkonia there is an additional factor of $\sqrt{3/2}$ on the r.h.s. of Eq. (C13). Recalling that the leading contribution in the $|t| \to 0$ limit comes from the amplitude $\widetilde{\mathcal{M}}_{TT}$ of (43), and using (33), we find that in the NRQCD limit

$$\frac{\mathrm{d}\sigma(\gamma p \to \mathcal{A} p)}{\mathrm{d}|t|} \to \frac{3\pi q_c^4 \alpha^3 N_c |R'(0)|^2 F_1^2(0)}{m_c^9} \,, \tag{C17}$$

which also agrees with Eq. (8c) in Ref. [87].

For the tensor quarkonia with $\bar{\lambda} = \pm 1, \pm 2$, the r.h.s. of Eq. (C13) is multiplied by a factor of $\sqrt{3}/M_{\mathcal{T}}$. Using (45) and (34) we find that the leading contribution comes from the amplitude $\widetilde{\mathcal{M}}_{TT2,p}$ so that

$$\frac{\mathrm{d}\sigma(\gamma p \to \mathcal{T} p)}{\mathrm{d}|t|} \to \frac{12\pi q_c^4 \alpha^3 N_c |R'(0)|^2 F_1^2(0)}{m_c^7 |t|} \,. \tag{C18}$$

This also agrees with Eq. (8d) in Ref. [87].

## Appendix D: The relative sign of the eikonal photon and Odderon exchange amplitudes

In this appendix we explain in detail our conventions for covariant derivatives, field equations, and Wilson lines. Consistent conventions are important for obtaining the correct relative sign of the Primakoff and Odderon amplitudes.

We write the covariant derivative in the fundamental representation as

$$(D^\mu)_{ij} = \delta_{ij} \left(\partial^\mu + \mathrm{i}eq_c A^\mu\right) + \mathrm{i}g A^{\mu a} (t^a)_{ij} \,. \tag{D1}$$

Here, $i, j = 1, \ldots, N_c$ are fundamental and $a = 1, \ldots, (N_c^2 - 1)$ is an adjoint color index; $A^\mu, A^{\mu a}$ represent the electromagnetic and color fields, respectively, and $t^a$ are the traceless generators of color-SU($N_c$), normalized as $\mathrm{tr}(t^a t^b) = \frac{1}{2}\delta^{ab}$. $q_c$ denotes the fractional electromagnetic charge of the fermion field on which this covariant derivative acts, $q_c = +2/3$ for $c$ quarks.

From the Dirac equation for a massless fermion field, $\psi_j$ $\mathrm{i}\gamma_\mu (D^\mu)_{ij} \psi_j = 0$ in a background with eikonal ("shock-wave") $A^\mu$ and $A^{\mu a}$ fields (where $A^+$ is the only non-vanishing field component and independent of $x^+$), $\psi_j$ is proportional to the Wilson line

$$V(\boldsymbol{x}_\perp) = \mathcal{P} \exp\left\{-\mathrm{i}\int_{-\infty}^{+\infty} \mathrm{d}x^- \left[eq_c A^+(x^-, \boldsymbol{x}_\perp) + gA^{+a}(x^-, \boldsymbol{x}_\perp) t^a\right]\right\} \,. \tag{D2}$$

The Yang-Mills and Maxwell equations for the shockwave fields in covariant gauge are

$$\partial_\mu F_a^{\mu\nu} = J_a^\nu \;\to\; -\boldsymbol{\nabla}_\perp^2 A_a^+ = J_a^+ \;, \tag{D3}$$

$$\partial_\mu F^{\mu\nu} = J^\nu \;\to\; -\boldsymbol{\nabla}_\perp^2 A^+ = J^+ \;, \tag{D4}$$

where $J^{+a}$ and $J^+$ are the plus components of the color and electromagnetic currents, respectively. Note that here $A^+$ denotes the electromagnetic field sourced by the charges in the proton, from which the projectile $c\bar{c}$ dipole scatters.

We now define the $S$-matrix for eikonal scattering of the color singlet dipole, averaged over the colors of the $c, \bar{c}$ quarks

$$S(\boldsymbol{x}_\perp, \boldsymbol{y}_\perp) = \frac{1}{N_c} \operatorname{tr} \left\langle V(\boldsymbol{x}_\perp) V^\dagger(\boldsymbol{y}_\perp) \right\rangle \;. \tag{D5}$$

$\langle \cdots \rangle$ denotes the matrix element of the respective operator between proton states $|P^+, \boldsymbol{P}_\perp\rangle$ and $\langle P^+ + \Delta^+, \boldsymbol{P}_\perp + \boldsymbol{\Delta}_\perp|$, stripped of the $\delta$-functions representing conservation of light-cone and transverse momentum.

We can now define the amplitudes for single photon or three gluon exchange as follows. Setting the QCD coupling $g = 0$ and expanding the Wilson lines to linear order in $A^+$ we have

$$V(\boldsymbol{x}_\perp) = 1 - \mathrm{i} e q_c \int_{-\infty}^{+\infty} \mathrm{d}x^- A^+(x^-, \boldsymbol{x}_\perp) \equiv 1 - \mathrm{i} e q_c \alpha(\boldsymbol{x}_\perp) \,. \tag{D6}$$

Then,

$$\Omega(\boldsymbol{x}_\perp, \boldsymbol{y}_\perp) \equiv -\frac{1}{2\mathrm{i}} \left\langle V(\boldsymbol{x}_\perp) V^\dagger(\boldsymbol{y}_\perp) - V(\boldsymbol{y}_\perp) V^\dagger(\boldsymbol{x}_\perp) \right\rangle = e q_c \left\langle \alpha(\boldsymbol{x}_\perp) - \alpha(\boldsymbol{y}_\perp) \right\rangle \;. \tag{D7}$$

corresponds to the scattering amplitude for single photon exchange. With the field equation (D4) we can write $\langle \alpha(\boldsymbol{x}_\perp) \rangle = -\left\langle \frac{1}{\boldsymbol{\nabla}_\perp^2} \rho(\boldsymbol{x}_\perp) \right\rangle$ where the integrated electric charge density is $\rho(\boldsymbol{x}_\perp) = \int_{-\infty}^{+\infty} \mathrm{d}x^- J^+(x^-, \boldsymbol{x}_\perp)$. Performing a Fourier transform from the transverse coordinate to the transverse momentum space, the matrix element of this operator is simply the Dirac electromagnetic form factor of the proton, $e F_1(q_\perp^2) = \langle \rho(\boldsymbol{q}_\perp) \rangle$. This leads to

$$\Omega(\boldsymbol{r}_\perp, \boldsymbol{\Delta}_\perp) = 8\pi\mathrm{i}\, \alpha q_c \frac{F_1(\Delta_\perp^2)}{\boldsymbol{\Delta}_\perp^2} \sin\left( \frac{\boldsymbol{r}_\perp \cdot \boldsymbol{\Delta}_\perp}{2} \right) \;. \tag{D8}$$

Here we introduced $\boldsymbol{r}_\perp = \boldsymbol{x}_\perp - \boldsymbol{y}_\perp$. This is the expression written in Eq. (36) of the main text. In impact parameter space,

$$\Omega(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp) = 8\pi\alpha q_c \int\limits_{\boldsymbol{\Delta}_\perp} \sin(\boldsymbol{b}_\perp \cdot \boldsymbol{\Delta}_\perp) \frac{-F_1(\Delta_\perp^2)}{\boldsymbol{\Delta}_\perp^2} \sin\left( \frac{\boldsymbol{r}_\perp \cdot \boldsymbol{\Delta}_\perp}{2} \right) \;. \tag{D9}$$

We now proceed to the three gluon exchange amplitude by setting $q_c = 0$ in the Wilson line (D2) followed by an expansion of (5), namely

$$\mathcal{O}(\boldsymbol{x}_\perp, \boldsymbol{y}_\perp) = -\frac{1}{2\mathrm{i} N_c} \operatorname{tr} \left\langle V(\boldsymbol{x}_\perp) V^\dagger(\boldsymbol{y}_\perp) - V(\boldsymbol{y}_\perp) V^\dagger(\boldsymbol{x}_\perp) \right\rangle \;, \tag{D10}$$

to third order in $A^{+a}$:

$$\begin{aligned} \mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp) = \frac{g^3}{4N_c} d^{abc} \int_{\boldsymbol{q}_{1\perp}\boldsymbol{q}_{2\perp}\boldsymbol{q}_{3\perp}} & \frac{\langle \rho^a(\boldsymbol{q}_{1\perp}) \rho^b(\boldsymbol{q}_{2\perp}) \rho^c(\boldsymbol{q}_{3\perp}) \rangle}{\boldsymbol{q}_{1\perp}^2 \boldsymbol{q}_{2\perp}^2 \boldsymbol{q}_{3\perp}^2} \sin\left( \boldsymbol{b}_\perp \cdot (\boldsymbol{q}_{1\perp} + \boldsymbol{q}_{2\perp} + \boldsymbol{q}_{3\perp}) \right) \\ & \times \left[ \sin\left( \frac{1}{2} \boldsymbol{r}_\perp \cdot (\boldsymbol{q}_{1\perp} - \boldsymbol{q}_{2\perp} - \boldsymbol{q}_{3\perp}) \right) + \frac{1}{3} \sin\left( \frac{1}{2} \boldsymbol{r}_\perp \cdot (\boldsymbol{q}_{1\perp} + \boldsymbol{q}_{2\perp} + \boldsymbol{q}_{3\perp}) \right) \right] \;. \end{aligned} \tag{D11}$$

For details see Eq. (77) of Ref. [22], to be multiplied by a factor of i (compare our Eq. (D10) to their Eq. (76)), and their Appendix B. Here $\rho^a(\boldsymbol{x}_\perp) = \int_{-\infty}^{+\infty} \mathrm{d}x^- J^{+a}(x^-, \boldsymbol{x}_\perp)$ denotes the color charge density integrated along the eikonal path. We now write the $C$-conjugation odd part of the color charge correlator as

$$\left\langle \rho^a(\boldsymbol{q}_{1\perp}) \rho^b(\boldsymbol{q}_{2\perp}) \rho^c(\boldsymbol{q}_{3\perp}) \right\rangle = \frac{1}{4} d^{abc} g^3 \, G_3^-(\boldsymbol{q}_{1\perp}, \boldsymbol{q}_{2\perp}, \boldsymbol{q}_{3\perp}) \;. \tag{D12}$$

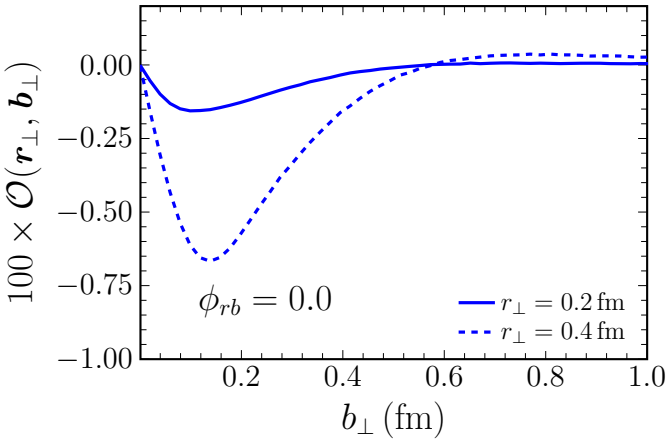



FIG. 8. The Odderon amplitude $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ obtained from a light-front quark model [65, 66], when the relative angle between $\boldsymbol{r}_\perp$ and $\boldsymbol{b}_\perp$ $\phi_{rb} = 0$. At small $b_\perp$ the sign is negative.

In a three-quark model this correlator takes the form [22]

$$\begin{aligned} G_3^-(\boldsymbol{q}_{1\perp}, \boldsymbol{q}_{2\perp}, \boldsymbol{q}_{3\perp}) \;=\; & \int [\mathrm{d}x_i] \int [\mathrm{d}^2 p_i]\, \psi(x_1, \boldsymbol{p}_{1\perp}; x_2, \boldsymbol{p}_{2\perp}; x_3, \boldsymbol{p}_{3\perp}) \\ & \Big[ \psi^*(x_1, \boldsymbol{p}_{1\perp} + (x_1 - 1)\boldsymbol{q}_\perp; x_2, \boldsymbol{p}_{2\perp} + x_2 \boldsymbol{q}_\perp; x_3, \boldsymbol{p}_{3\perp} + x_3 \boldsymbol{q}_\perp) \\ & - \psi^*(x_1, \boldsymbol{p}_{1\perp} - \boldsymbol{q}_{1\perp} + x_1 \boldsymbol{q}_\perp; x_2, \boldsymbol{p}_{2\perp} - \boldsymbol{q}_{2\perp} - \boldsymbol{q}_{3\perp} + x_2 \boldsymbol{q}_\perp; x_3, \boldsymbol{p}_{3\perp} + x_3 \boldsymbol{q}_\perp) \\ & - \psi^*(x_1, \boldsymbol{p}_{1\perp} - \boldsymbol{q}_{1\perp} - \boldsymbol{q}_{3\perp} + x_1 \boldsymbol{q}_\perp; x_2, \boldsymbol{p}_{2\perp} - \boldsymbol{q}_{2\perp} + x_2 \boldsymbol{q}_\perp; x_3, \boldsymbol{p}_{3\perp} + x_3 \boldsymbol{q}_\perp) \\ & - \psi^*(x_1, \boldsymbol{p}_{1\perp} - \boldsymbol{q}_{1\perp} - \boldsymbol{q}_{2\perp} + x_1 \boldsymbol{q}_\perp; x_2, \boldsymbol{p}_{2\perp} - \boldsymbol{q}_{3\perp} + x_2 \boldsymbol{q}_\perp; x_3, \boldsymbol{p}_{3\perp} + x_3 \boldsymbol{q}_\perp) \\ & + 2\, \psi^*(x_1, \boldsymbol{p}_{1\perp} - \boldsymbol{q}_{1\perp} + x_1 \boldsymbol{q}_\perp; x_2, \boldsymbol{p}_{2\perp} - \boldsymbol{q}_{2\perp} + x_2 \boldsymbol{q}_\perp; x_3, \boldsymbol{p}_{3\perp} - \boldsymbol{q}_{3\perp} + x_3 \boldsymbol{q}_\perp) \Big] \;. \end{aligned} \tag{D13}$$

Here, $\boldsymbol{q}_\perp = \boldsymbol{q}_{1\perp} + \boldsymbol{q}_{2\perp} + \boldsymbol{q}_{3\perp}$ and

$$\begin{aligned} [\mathrm{d}x_i] &\equiv \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \delta(1 - x_1 - x_2 - x_3)\,, \\ [\mathrm{d}^2 p_i] &\equiv \frac{1}{(16\pi^3)^2} \mathrm{d}^2 \boldsymbol{p}_{1\perp} \mathrm{d}^2 \boldsymbol{p}_{2\perp} \mathrm{d}^2 \boldsymbol{p}_{3\perp} \delta^{(2)}(\boldsymbol{p}_{1\perp} + \boldsymbol{p}_{2\perp} + \boldsymbol{p}_{3\perp})\,, \end{aligned} \tag{D14}$$

denote integrations over the quark light-cone momentum fractions and transverse momenta. $\psi(x_1, \boldsymbol{p}_{1\perp}; x_2, \boldsymbol{p}_{2\perp}; x_3, \boldsymbol{p}_{3\perp})$ is the light-cone wave function of the model proton. This expression agrees with the $C$-odd three-gluon exchange proton impact factor $E_{3;0}$ by Bartels and Motyka [91] up to a conventional factor of $(-\mathrm{i})^3$. The first term in Eq. (D13) corresponds to the coupling of the three exchanged gluons to the same quark in the proton, and it is equal to the Dirac form factor $F_1(\boldsymbol{q}_\perp^2)$; note the opposite sign as compared to the photon exchange amplitude (D9). This contribution is dominant when all $|\boldsymbol{q}_{\perp i}|$ are much greater than the typical quark transverse momentum while $|\boldsymbol{q}_\perp|$ is on the order of that scale. For high momentum transfer $|t| = \boldsymbol{q}_\perp^2$ on the other hand, contributions due to gluon exchanges with two or three quarks in the proton become dominant. Where this transition occurs is determined by the light-cone wave function $\psi$ which encodes the structure of the proton, specifically the single quark momentum distribution as well as multi-quark momentum correlations.

$G_3^-$ can also be expressed [92] in terms of two-gluon exchanges, $G_2(\boldsymbol{q}_{1\perp}, \boldsymbol{q}_{2\perp}) = 2\delta^{ab} \left\langle \rho^a(\boldsymbol{q}_{1\perp})\, \rho^b(\boldsymbol{q}_{2\perp}) \right\rangle / g^2 (N_c^2 - 1)$, where two of the three gluons are "paired up", plus a genuine 3-body contribution which enforces the Ward identity (vanishing of $G_3^-$) when either one of the transverse momenta vanishes, $\boldsymbol{q}_{\perp i} \to 0$:

$$\begin{aligned} G_3^-(\boldsymbol{q}_{1\perp}, \boldsymbol{q}_{2\perp}, \boldsymbol{q}_{3\perp}) \;=\; & G_2(\boldsymbol{q}_{1\perp} + \boldsymbol{q}_{2\perp}, \boldsymbol{q}_{3\perp}) + G_2(\boldsymbol{q}_{1\perp} + \boldsymbol{q}_{3\perp}, \boldsymbol{q}_{2\perp}) + G_2(\boldsymbol{q}_{2\perp} + \boldsymbol{q}_{3\perp}, \boldsymbol{q}_{1\perp}) \\ & - 2 \int [\mathrm{d}x_i] \int [\mathrm{d}^2 p_i] \, [\psi^*(x_1, \boldsymbol{p}_{1\perp} + (x_1 - 1)\boldsymbol{q}_\perp; x_2, \boldsymbol{p}_{2\perp} + x_2 \boldsymbol{q}_\perp; x_3, \boldsymbol{p}_{3\perp} + x_3 \boldsymbol{q}_\perp) \\ & \qquad - \psi^*(x_1, \boldsymbol{p}_{1\perp} - \boldsymbol{q}_{1\perp} + x_1 \boldsymbol{q}_\perp; x_2, \boldsymbol{p}_{2\perp} - \boldsymbol{q}_{2\perp} + x_2 \boldsymbol{q}_\perp; x_3, \boldsymbol{p}_{3\perp} - \boldsymbol{q}_{3\perp} + x_3 \boldsymbol{q}_\perp)] \times \\ & \qquad \psi(x_1, \boldsymbol{p}_{1\perp}; x_2, \boldsymbol{p}_{2\perp}; x_3, \boldsymbol{p}_{3\perp}) \;. \end{aligned} \tag{D15}$$

Once again, for small $|\boldsymbol{q}_\perp|$ but large $|\boldsymbol{q}_{\perp i}|$ the fourth term on the right-hand-side of this equation becomes $-2F_1(\boldsymbol{q}_\perp^2)$.

A simple light-front quark model wave function from the literature [65, 66] predicts $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp) < 0$ at small impact parameters, see Fig. 8. This corresponds to constructive interference of photon and Odderon exchange amplitudes. Neither the fixed order $\mathcal{O}(\alpha_S)$ correction [64] to the matrix element $\langle \rho^a(\boldsymbol{q}_{1\perp})\rho^b(\boldsymbol{q}_{2\perp})\rho^c(\boldsymbol{q}_{3\perp})\rangle$ nor small-$x$ resummation of $\mathcal{O}(\boldsymbol{r}_\perp, \boldsymbol{b}_\perp)$ change this initial sign; the resulting first harmonic $\mathcal{O}_1(r_\perp, \Delta_\perp)$ at smaller $x$ is shown above in Fig. 2.

## Appendix E: BJKP–BLV Odderon

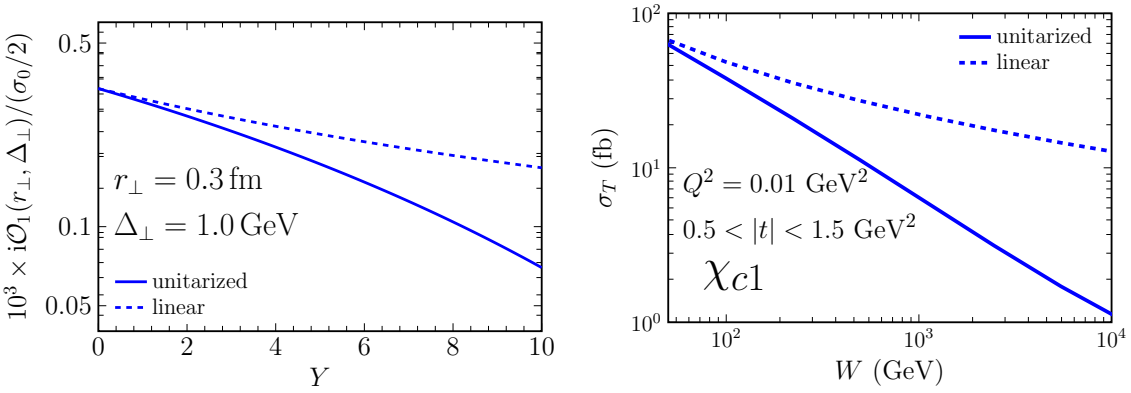



FIG. 9. Left: A comparison of the unitarized vs linear evolution of the Odderon amplitude $\mathcal{O}_1(r_\perp, \Delta_\perp)$ as a function of $Y$. Right: A plot of the resulting $W$ dependence of the Odderon component of the $\gamma^* p \to \chi_{c1} p$ cross section.

In this work the small-$x$ evolution of the Odderon exchange amplitude has been computed using the non-linear evolution of the Odderon coupled to the BK equation [11, 12] – the "unitarized solution". It is interesting to compare to the BLV solution [9] obtained from the BKP equation for the Odderon, that is, without including the unitarity corrections[13] – the "linear solution". Starting from the initial condition at $Y = 0$ obtained in [64] and described in Sec. II D, in Fig. 9 we show $\mathcal{O}_1(r_\perp, \Delta_\perp)$ as a function of $Y = \log(x_0/x_{\mathcal{P}})$ (left) with fixed $r_\perp = 0.3$ fm and $\Delta_\perp = 1$ GeV and up to $Y = 10$. The unitarized (full) and the linear (dashed line) results for $\mathcal{O}_1(r_\perp, \Delta_\perp)$ both drop as a function of $Y$. For the unitarized solution this is more pronounced due to the non-linear corrections suppressing the Odderon exchange amplitude [14, 58, 94].

The impact of the unitarized vs linear evolution on the $W$-dependence of the Odderon component of the cross section is shown on Fig. 9 right. As an example, we consider $\sigma_T(\gamma^* p \to \chi_{c1} p)$ cross section with the kinematic cuts as in Fig. 4 from which we reproduced the unitarized evolution result. The range in $Y$ shown on Fig. 9, left roughly translates to the range in $W$ shown on Fig. 9, right. Though the linear evolution result for $\sigma_T$ leads to a much milder $W$-dependence, the numerical cross section does not seem to fully reach the BLV asymptotics (a constant value) even for $W = 10^4$ GeV. Note however that from the standard saddle point analysis one expects a $1/Y^3$ corrections to the asymptotic behavior of the non-forward BLV exchange cross section – as it was established for an analogous case of the non-forward BFKL cross section [95]. From the perspective of the EIC, taking the top $\sqrt{S} = 140$ GeV collision energy of the $ep$ system, $W$ can reach at most $W \approx 136$ GeV (at $y = 0.95$) resulting in a factor of $2 - 3$ difference in the cross section at the high $y$ end of the photon flux. Due to the $1/y$ leading behavior of the photon flux, however, the exclusive $\chi_{cJ}$ production at the EIC will be dominated by much lower $W$, close to the experimental lower cutoff on $W$ for the exclusive events, say for $2 < Y < 4$, where the unitarity corrections to the Odderon exchange cross section are below 30% – i.e. small in comparison to the theoretical uncertainty.

[1] L. Lukaszuk and B. Nicolescu, A Possible interpretation of p p rising total cross-sections, Lett. Nuovo Cim. **8**, 405 (1973).

[13] Alternatively, one could solve for the BKP Green's function and reconstruct the Odderon amplitude from it [93].

[2] J. Bartels, High-Energy Behavior in a Nonabelian Gauge Theory (II): First Corrections to $T_{n\to m}$ Beyond the Leading $\ln s$ Approximation, Nucl. Phys. B **175**, 365 (1980).

[3] J. Kwiecinski and M. Praszalowicz, Three Gluon Integral Equation and Odd c Singlet Regge Singularities in QCD, Phys. Lett. B **94**, 413 (1980).

[4] T. Jaroszewicz, Infrared Divergences and Regge Behavior in QCD, Acta Phys. Polon. B **11**, 965 (1980).

[5] L. N. Lipatov, Reggeization of the Vector Meson and the Vacuum Singularity in Nonabelian Gauge Theories, Sov. J. Nucl. Phys. **23**, 338 (1976).

[6] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, The Pomeranchuk Singularity in Nonabelian Gauge Theories, Sov. Phys. JETP **45**, 199 (1977).

[7] I. I. Balitsky and L. N. Lipatov, The Pomeranchuk Singularity in Quantum Chromodynamics, Sov. J. Nucl. Phys. **28**, 822 (1978).

[8] J. Bartels, V. S. Fadin, L. N. Lipatov, and G. P. Vacca, NLO Corrections to the kernel of the BKP-equations, Nucl. Phys. B **867**, 827 (2013), arXiv:1210.0797 [hep-ph].

[9] J. Bartels, L. N. Lipatov, and G. P. Vacca, A New odderon solution in perturbative QCD, Phys. Lett. B **477**, 178 (2000), arXiv:hep-ph/9912423.

[10] A. M. Stasto, Small x resummation and the Odderon, Phys. Lett. B **679**, 288 (2009), arXiv:0904.4124 [hep-ph].

[11] Y. V. Kovchegov, L. Szymanowski, and S. Wallon, Perturbative odderon in the dipole model, Phys. Lett. B **586**, 267 (2004), arXiv:hep-ph/0309281.

[12] Y. Hatta, E. Iancu, K. Itakura, and L. McLerran, Odderon in the color glass condensate, Nucl. Phys. A **760**, 172 (2005), arXiv:hep-ph/0501171.

[13] L. Motyka, Nonlinear evolution of pomeron and odderon in momentum space, Phys. Lett. B **637**, 185 (2006), arXiv:hep-ph/0509270.

[14] T. Lappi, A. Ramnath, K. Rummukainen, and H. Weigert, JIMWLK evolution of the odderon, Phys. Rev. D **94**, 054014 (2016), arXiv:1606.00551 [hep-ph].

[15] X. Yao, Y. Hagiwara, and Y. Hatta, Computing the gluon Sivers function at small-$x$, Phys. Lett. B **790**, 361 (2019), arXiv:1812.03959 [hep-ph].

[16] S. Benić, D. Horvatić, A. Kaushik, and E. A. Vivoda, Exclusive $\eta$c production from small-x evolved Odderon at an electron-ion collider, Phys. Rev. D **108**, 074005 (2023), arXiv:2306.10626 [hep-ph].

[17] S. Jeon and R. Venugopalan, A Classical Odderon in QCD at high energies, Phys. Rev. D **71**, 125003 (2005), arXiv:hep-ph/0503219.

[18] S. Jeon and R. Venugopalan, Random walks of partons in SU(N(c)) and classical representations of color charges in QCD at small x, Phys. Rev. D **70**, 105012 (2004), arXiv:hep-ph/0406169.

[19] J. Zhou, Transverse single spin asymmetries at small x and the anomalous magnetic moment, Phys. Rev. D **89**, 074050 (2014), arXiv:1308.5912 [hep-ph].

[20] R. Boussarie, Y. Hatta, L. Szymanowski, and S. Wallon, Probing the Gluon Sivers Function with an Unpolarized Target: GTMD Distributions and the Odderons, Phys. Rev. Lett. **124**, 172501 (2020), arXiv:1912.08182 [hep-ph].

[21] Y. Hagiwara, Y. Hatta, R. Pasechnik, and J. Zhou, Spin-dependent Pomeron and Odderon in elastic proton-proton scattering, Eur. Phys. J. C **80**, 427 (2020), arXiv:2003.03680 [hep-ph].

[22] A. Dumitru, G. A. Miller, and R. Venugopalan, Extracting many-body color charge correlators in the proton from exclusive DIS at large Bjorken x, Phys. Rev. D **98**, 094004 (2018), arXiv:1808.02501 [hep-ph].

[23] A. Dumitru, H. Mäntysaari, and R. Paatelainen, Cubic color charge correlator in a proton made of three quarks and a gluon, Phys. Rev. D **105**, 036007 (2022), arXiv:2106.12623 [hep-ph].

[24] G. Antchev *et al.* (TOTEM), Elastic differential cross-section $\mathrm{d}\sigma/\mathrm{d}t$ at $\sqrt{s} = 2.76$ TeV and implications on the existence of a colourless C-odd three-gluon compound state, Eur. Phys. J. C **80**, 91 (2020), arXiv:1812.08610 [hep-ex].

[25] V. M. Abazov *et al.* (D0), Measurement of the differential cross section $d\sigma/dt$ in elastic $p\bar{p}$ scattering at $\sqrt{s} = 1.96$ TeV, Phys. Rev. D **86**, 012009 (2012), arXiv:1206.0687 [hep-ex].

[26] A. Schafer, L. Mankiewicz, and O. Nachtmann, Diffractive eta(c), eta-prime, J / psi and psi-prime production in electron - proton collisions at HERA energies, in *Workshop on Physics at HERA* (1992).

[27] J. Czyzewski, J. Kwiecinski, L. Motyka, and M. Sadzikowski, Exclusive eta(c) photoproduction and electroproduction at HERA as a possible probe of the odderon singularity in QCD, Phys. Lett. B **398**, 400 (1997), [Erratum: Phys.Lett.B 411, 402 (1997)], arXiv:hep-ph/9611225.

[28] R. Engel, D. Y. Ivanov, R. Kirschner, and L. Szymanowski, Diffractive meson production from virtual photons with odd charge - parity exchange, Eur. Phys. J. C **4**, 93 (1998), arXiv:hep-ph/9707362.

[29] W. Kilian and O. Nachtmann, Single pseudoscalar meson production in diffractive e p scattering, Eur. Phys. J. C **5**, 317 (1998), arXiv:hep-ph/9712371.

[30] E. R. Berger, A. Donnachie, H. G. Dosch, W. Kilian, O. Nachtmann, and M. Rueter, Odderon and photon exchange in electroproduction of pseudoscalar mesons, Eur. Phys. J. C **9**, 491 (1999), arXiv:hep-ph/9901376.

[31] J. Bartels, M. A. Braun, D. Colferai, and G. P. Vacca, Diffractive eta(c) photoproduction and electroproduction with the perturbative QCD odderon, Eur. Phys. J. C **20**, 323 (2001), arXiv:hep-ph/0102221.

[32] A. Dumitru and T. Stebel, Multiquark matrix elements in the proton and three gluon exchange for exclusive $\eta_c$ production in photon-proton diffractive scattering, Phys. Rev. D **99**, 094038 (2019), arXiv:1903.07660 [hep-ph].

[33] R. L. Workman *et al.* (Particle Data Group), Review of Particle Physics, PTEP **2022**, 083C01 (2022).

[34] S. R. Klein, Comment on "$\eta_c$ production in photon-induced interactions at the LHC", Phys. Rev. D **98**, 118501 (2018), arXiv:1808.08253 [hep-ph].

[35] L. A. Harland-Lang, V. A. Khoze, A. D. Martin, and M. G. Ryskin, Searching for the Odderon in Ultraperipheral Proton–Ion Collisions at the LHC, Phys. Rev. D **99**, 034011 (2019), arXiv:1811.12705 [hep-ph].

[36] S. J. Brodsky, J. Rathsman, and C. Merino, Odderon-Pomeron interference, Phys. Lett. B **461**, 114 (1999), arXiv:hep-ph/9904280.

[37] P. Hägler, B. Pire, L. Szymanowski, and O. V. Teryaev, Hunting the QCD-Odderon in hard diffractive electroproduction of two pions, Phys. Lett. B **535**, 117 (2002), [Erratum: Phys.Lett.B 540, 324–325 (2002)], arXiv:hep-ph/0202231.

[38] P. Hagler, B. Pire, L. Szymanowski, and O. V. Teryaev, Pomeron - odderon interference effects in electroproduction of two pions, Eur. Phys. J. C **26**, 261 (2002), arXiv:hep-ph/0207224.

[39] A. Schafer, L. Mankiewicz, and O. Nachtmann, Double diffractive J / psi and phi production as a probe for the odderon, Phys. Lett. B **272**, 419 (1991).

[40] A. Bzdak, L. Motyka, L. Szymanowski, and J. R. Cudell, Exclusive J/psi and Upsilon hadroproduction and the QCD odderon, Phys. Rev. D **75**, 094023 (2007), arXiv:hep-ph/0702134.

[41] L. Pentchev *et al.* (GlueX), Exclusive threshold $J/\psi$ photoproduction with GlueX (2023), presented at DIS2023: XXX International Workshop on Deep-Inelastic Scattering and Related Subjects, Michigan State University, East Lansing, Michigan, USA.

[42] A. Accardi *et al.*, Electron Ion Collider: The Next QCD Frontier: Understanding the glue that binds us all, Eur. Phys. J. A **52**, 268 (2016), arXiv:1212.1701 [nucl-ex].

[43] R. Abdul Khalek *et al.*, Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report, Nucl. Phys. A **1026**, 122447 (2022), arXiv:2103.05419 [physics.ins-det].

[44] N. N. Nikolaev and B. G. Zakharov, Color transparency and scaling properties of nuclear shadowing in deep inelastic scattering, Z. Phys. C **49**, 607 (1991).

[45] A. H. Mueller, Soft gluons in the infinite momentum wave function and the BFKL pomeron, Nucl. Phys. B **415**, 373 (1994).

[46] G. P. Lepage and S. J. Brodsky, Exclusive Processes in Perturbative Quantum Chromodynamics, Phys. Rev. D **22**, 2157 (1980).

[47] H. G. Dosch, T. Gousset, G. Kulzinger, and H. J. Pirner, Vector meson leptoproduction and nonperturbative gluon fluctuations in QCD, Phys. Rev. D **55**, 2602 (1997), arXiv:hep-ph/9608203.

[48] T. Lappi, H. Mäntysaari, and J. Penttala, Relativistic corrections to the vector meson light front wave function, Phys. Rev. D **102**, 054020 (2020), arXiv:2006.02830 [hep-ph].

[49] E. R. Berger, A. Donnachie, H. G. Dosch, and O. Nachtmann, Observing the odderon: Tensor meson photoproduction, Eur. Phys. J. C **14**, 673 (2000), arXiv:hep-ph/0001270.

[50] F. Fillion-Gourdeau and S. Jeon, Tensor Meson Production in Proton-Proton Collisions from the Color Glass Condensate, Phys. Rev. C **77**, 055201 (2008), arXiv:0709.4196 [hep-ph].

[51] R. S. Pasechnik, A. Szczurek, and O. V. Teryaev, Nonperturbative and spin effects in the central exclusive production of tensor $\chi_c(2^+)$ meson, Phys. Rev. D **81**, 034024 (2010), arXiv:0912.4251 [hep-ph].

[52] J. P. Lansberg and T. N. Pham, Effective Lagrangian for Two-photon and Two-gluon Decays of P-wave Heavy Quarkonium chi(c0,2) and chi(b0,2) states, Phys. Rev. D **79**, 094016 (2009), arXiv:0903.1562 [hep-ph].

[53] H.-Y. Cheng, C.-K. Chua, and C.-W. Hwang, Covariant light front approach for s wave and p wave mesons: Its application to decay constants and form-factors, Phys. Rev. D **69**, 074025 (2004), arXiv:hep-ph/0310359.

[54] A. D. Martin, M. G. Ryskin, and T. Teubner, The QCD description of diffractive rho meson electroproduction, Phys. Rev. D **55**, 4329 (1997), arXiv:hep-ph/9609448.

[55] A. D. Martin, M. G. Ryskin, and T. Teubner, Q**2 dependence of diffractive vector meson electroproduction, Phys. Rev. D **62**, 014022 (2000), arXiv:hep-ph/9912551.

[56] I. Babiarz, R. Pasechnik, W. Schäfer, and A. Szczurek, Hadroproduction of scalar $P$-wave quarkonia in the light-front $k_T$ -factorization approach, JHEP **06**, 101, arXiv:2002.09352 [hep-ph].

[57] R. Fiore and V. R. Zoller, Charged currents, color dipoles and xF(3) at small x, JETP Lett. **82**, 385 (2005), arXiv:hep-ph/0508187.

[58] L. Motyka and G. Watt, Exclusive photoproduction at the Tevatron and CERN LHC within the dipole picture, Phys. Rev. D **78**, 014023 (2008), arXiv:0805.2113 [hep-ph].

[59] I. Babiarz, R. Pasechnik, W. Schäfer, and A. Szczurek, Light-front approach to axial-vector quarkonium $\gamma^*\gamma^*$ form factors, JHEP **09**, 170, arXiv:2208.05377 [hep-ph].

[60] C. R. Ji, P. L. Chung, and S. R. Cotanch, Light cone quark model axial vector meson wave function, Phys. Rev. D **45**, 4214 (1992).

[61] S. N. Gupta, J. M. Johnson, and W. W. Repko, Relativistic two photon and two gluon decay rates of heavy quarkonia, Phys. Rev. D **54**, 2075 (1996), arXiv:hep-ph/9606349.

[62] H. Mäntysaari, K. Roy, F. Salazar, and B. Schenke, Gluon imaging using azimuthal correlations in diffractive scattering at the Electron-Ion Collider, Phys. Rev. D **103**, 094026 (2021), arXiv:2011.02464 [hep-ph].

[63] V. Shtabovenko, R. Mertig, and F. Orellana, FeynCalc 9.3: New features and improvements, Comput. Phys. Commun. **256**, 107478 (2020), arXiv:2001.04407 [hep-ph].

[64] A. Dumitru, H. Mäntysaari, and R. Paatelainen, Stronger C-odd color charge correlations in the proton at higher energy, Phys. Rev. D **107**, L011501 (2023), arXiv:2210.05390 [hep-ph].

[65] F. Schlumpf, Relativistic constituent quark model of electroweak properties of baryons, Phys. Rev. D **47**, 4114 (1993), [Erratum: Phys.Rev.D 49, 6246 (1994)], arXiv:hep-ph/9212250.

[66] S. J. Brodsky and F. Schlumpf, Wave function independent relations between the nucleon axial coupling g(A) and the

nucleon magnetic moments, Phys. Lett. B **329**, 111 (1994), arXiv:hep-ph/9402214.
[67] K. J. Golec-Biernat and A. M. Stasto, On solutions of the Balitsky-Kovchegov equation with impact parameter, Nucl. Phys. B **668**, 345 (2003), arXiv:hep-ph/0306279.
[68] J. Berger and A. Stasto, Numerical solution of the nonlinear evolution equation at small x with impact parameter and beyond the LL approximation, Phys. Rev. D **83**, 034015 (2011), arXiv:1010.0671 [hep-ph].
[69] T. Lappi and H. Mäntysaari, Single inclusive particle production at high energy from HERA data to proton-nucleus collisions, Phys. Rev. D **88**, 114020 (2013), arXiv:1309.6963 [hep-ph].
[70] H. Kowalski, T. Lappi, C. Marquet, and R. Venugopalan, Nuclear enhancement and suppression of diffractive structure functions at high energies, Phys. Rev. C **78**, 045201 (2008), arXiv:0805.4071 [hep-ph].
[71] P. V. Landshoff, Model for elastic scattering at wide angle, Phys. Rev. D **10**, 1024 (1974).
[72] A. Donnachie and P. V. Landshoff, Elastic Scattering at Large t, Z. Phys. C **2**, 55 (1979), [Erratum: Z.Phys.C 2, 372 (1979)].
[73] A. Donnachie and P. V. Landshoff, Multi - Gluon Exchange in *pp* Elastic Scattering, Phys. Lett. B **123**, 345 (1983).
[74] A. Donnachie and P. V. Landshoff, p p and anti-p p Elastic Scattering, Nucl. Phys. B **231**, 189 (1984).
[75] L. D. Landau, On the angular momentum of a system of two photons, Dokl. Akad. Nauk SSSR **60**, 207 (1948).
[76] C.-N. Yang, Selection Rules for the Dematerialization of a Particle Into Two Photons, Phys. Rev. **77**, 242 (1950).
[77] Z. Ye, J. Arrington, R. J. Hill, and G. Lee, Proton and Neutron Electromagnetic Form Factors and Uncertainties, Phys. Lett. B **777**, 8 (2018), arXiv:1707.09063 [nucl-ex].
[78] H. Kowalski, L. Motyka, and G. Watt, Exclusive diffractive processes at HERA within the dipole picture, Phys. Rev. D **74**, 074016 (2006), arXiv:hep-ph/0606272.
[79] Y. Li, M. Li, and J. P. Vary, Two-photon transitions of charmonia on the light front, Phys. Rev. D **105**, L071901 (2022), arXiv:2111.14178 [hep-ph].
[80] W. Kwong, P. B. Mackenzie, R. Rosenfeld, and J. L. Rosner, Quarkonium Annihilation Rates, Phys. Rev. D **37**, 3210 (1988).
[81] R. Barbieri, M. Caffo, R. Gatto, and E. Remiddi, Strong QCD Corrections to p Wave Quarkonium Decays, Phys. Lett. B **95**, 93 (1980).
[82] Y. Li, P. Maris, and J. P. Vary, Quarkonium as a relativistic bound state on the light front, Phys. Rev. D **96**, 016022 (2017), arXiv:1704.06968 [hep-ph].
[83] C. F. von Weizsacker, Radiation emitted in collisions of very fast electrons, Z. Phys. **88**, 612 (1934).
[84] E. J. Williams, Correlation of certain collision problems with radiation theory, Kong. Dan. Vid. Sel. Mat. Fys. Med. **13N4**, 1 (1935).
[85] V. M. Budnev, I. F. Ginzburg, G. V. Meledin, and V. G. Serbo, The Two photon particle production mechanism. Physical problems. Applications. Equivalent photon approximation, Phys. Rept. **15**, 181 (1975).
[86] A. Caldwell and H. Kowalski, The J/psi Way to Nuclear Structure, in *13th International Conference on Elastic and Diffractive Scattering (Blois Workshop): Moving Forward into the LHC Era* (2009) pp. 190–192, arXiv:0909.1254 [hep-ph].
[87] Y. Jia, Z. Mo, J. Pan, and J.-Y. Zhang, Photoproduction of C-even quarkonia at the EIC and EicC, Phys. Rev. D **108**, 016015 (2023), arXiv:2207.14171 [hep-ph].
[88] E. C. Aschenauer, S. Fazio, J. H. Lee, H. Mantysaari, B. S. Page, B. Schenke, T. Ullrich, R. Venugopalan, and P. Zurita, The electron–ion collider: assessing the energy dependence of key measurements, Rept. Prog. Phys. **82**, 024301 (2019), arXiv:1708.01527 [nucl-ex].
[89] J. C. Collins and R. K. Ellis, Heavy quark production in very high-energy hadron collisions, Nucl. Phys. B **360**, 3 (1991).
[90] H. Aihara *et al.* (TPC/Two Gamma), Formation of Spin One Mesons by Photon-photon Fusion, Phys. Rev. D **38**, 1 (1988).
[91] J. Bartels and L. Motyka, Baryon scattering at high energies: Wave function, impact factor, and gluon radiation, Eur. Phys. J. C **55**, 65 (2008), arXiv:0711.2196 [hep-ph].
[92] A. Dumitru, V. Skokov, and T. Stebel, Subfemtometer scale color charge correlations in the proton, Phys. Rev. D **101**, 054004 (2020), arXiv:2001.04516 [hep-ph].
[93] G. Chachamis and A. Sabio Vera, Solution of the Bartels-Kwiecinski-Praszalowicz equation via Monte Carlo integration, Phys. Rev. D **94**, 034019 (2016), arXiv:1606.07349 [hep-ph].
[94] Y. Hatta, B.-W. Xiao, and F. Yuan, Gluon Tomography from Deeply Virtual Compton Scattering at Small-x, Phys. Rev. D **95**, 114026 (2017), arXiv:1703.02085 [hep-ph].
[95] J. R. Forshaw and M. G. Ryskin, Diffractive vector meson production at large momentum transfer, Z. Phys. C **68**, 137 (1995), arXiv:hep-ph/9501376.

arXiv:2402.19131v1 [hep-ph] 29 Feb 2024

# Charged Lepton Flavor Violation in the B-L symmetric SSM

Xing-Xing Dong[1,2,3*], Shu-Min Zhao[1,2,3†], Jia-Peng Huo[1,2,3], Tong-Tong Wang[1,2,3‡], Tai-Fu Feng[1,2,3,4§]

[1] *Department of Physics, Hebei University, Baoding, 071002, China*

[2] *Hebei Key Laboratory of High-precision Computation and Application of Quantum Field Theory, 071002, China*

[3] *Hebei Research Center of the Basic Discipline for Computational Physics, Baoding, 071002, China*

[4] *Department of Physics, Chongqing University, Chongqing 401331, China*

## Abstract

Charged lepton flavor violation (CLFV) represents a clear new physics (NP) signal beyond the standard model (SM). In this work, we investigate CLFV processes $l_j^- \to l_i^-\gamma$ utilizing mass insertion approximation(MIA) in the minimal supersymmetric extension of the SM with local B-L gauge symmetry (B-LSSM). The MIA method can provide a set of simple analytic formulae for the form factors and the associated effective vertices, so that the movement of the CLFV decays $l_j^- \to l_i^-\gamma$ with the sensitive parameters will be intuitively analyzed. Considering the SM-like Higgs boson mass and the muon anomalous dipole moment (MDM) within $4\sigma$, $3\sigma$ and $2\sigma$ regions, we discuss the corresponding constraints on the relevant parameter space of the model.

PACS numbers: 12.60.-Jv, 13.35.-r, 13.40.Em

Keywords: Supersymmetric Model, Charged Lepton Flavor Violation, Muon Magnetic Dipole Moment

* dongxx@hbu.edu.cn

† zhaosm@hbu.edu.cn

‡ wtt961018@163.com

§ fengtf@hbu.edu.cn

## I. INTRODUCTION

Although the Standard Model(SM) is considered as a much mature theory, it holds lepton number is conserved, that is, no charged lepton flavor violation (CLFV) occurs[1]. But the CLFV processes can easily occur in new physics (NP) beyond the SM. Therefore, if the CLFV signals are observed in the future experiments, it is obvious evidence of the NP beyond the SM. In TABLE I, we show the latest experimental data for the CLFV processes $l_j^- \to l_i^- \gamma$[2–4], and these processes have been discussed in various theoretical frameworks[5–13]. In this work, we investigate these CLFV processes in the minimal supersymmetric extension of the SM with local B-L gauge symmetry (B-LSSM)[14–18]. We hope to reveal some properties of high energy physics through detailed analyses of these CLFV processes.

TABLE I: The latest experiment limits for the CLFV processes $l_j^- \to l_i^- \gamma$.

| CLFV process | Present limit | confidence level (CL) |
|---|---|---|
| $\mu \to e\gamma$ | $< 4.2 \times 10^{-13}$[2] | 90% |
| $\tau \to e\gamma$ | $< 3.3 \times 10^{-8}$[3] | 90% |
| $\tau \to \mu\gamma$ | $< 4.2 \times 10^{-8}$[4] | 90% |

It is worth noting that we use a novel calculation method called as mass insertion approximation(MIA)[19–24] to study CLFV processes $l_j^- \to l_i^- \gamma$ in the B-LSSM. The CLFV decays $l_j^- \to l_i^- \gamma$ are produced via one-loop contributions, which are influenced by the flavor mixing among the three generations of the B-LSSM sleptons and/or sneutrinos. The MIA works with the sleptons(sneutrinos) in the electroweak interaction eigenstate instead of mass eigenstate. That is to say, the MIA method operates mass insertions inside the propagators of the electroweak interaction sleptons(sneutrinos) eigenstates, instead of performing the exact diagonalization of the mass basis involved in the full one-loop computation. The MIA method has been studied by other LFV works, including the $h, H, A \to \tau\mu$ decays induced from SUSY loops[20], effective LFV $Hl_il_j$ vertex from right-handed neutrinos[21], one-loop effective LFV $Zl_kl_m$ vertex from heavy neutrinos[22], LFV decays $l_j \to l_i\gamma$ in the $U(1)_XSSM$[24] and so on. These works provide references and guidance for our research of CLFV processes $l_j^- \to l_i^- \gamma$ in the B-LSSM.

On the base of the minimum supersymmetric Standard Model (MSSM)[25–28], B-LSSM extends the gauge symmetry group to $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \otimes U(1)_{B-L}$, where $B$ represents the baryon number and $L$ stands for the lepton number. The B-LSSM adds two singlet Higgs superfields $\hat{\eta}$ and $\hat{\bar{\eta}}$ and three generations of right-handed neutrinos superfields $\hat{\nu}_i^c$ to the MSSM. The invariance under $U(1)_{B-L}$ gauge group imposes the R-parity conservation, which is assumed in the MSSM, to avoid proton decay[29]. Besides, through the additional singlet Higgs states and right-handed (s)neutrinos, additional parameter space in the B-LSSM is released from the LEP, Tevatron and LHC constraints to alleviate the hierarchy problem of the MSSM[30, 31]. Furthermore, the B-LSSM can provide much more dark matter (DM) candidates than that in the MSSM[32–35].

Our research of CLFV processes $l_j^- \to l_i^- \gamma$ in the B-LSSM possesses much differences comparing that of $U(1)_X SSM$. Firstly, because the two models contain different fields, and the corresponding quantum numbers are different, the two works are discussed under different models. In the $U(1)_X SSM$, three Higgs singlets and right-handed neutrinos are added to MSSM. This model relieves the so called little hierarchy problem that appears in the MSSM. $\hat{S}$ is the singlet Higgs superfield with a non-zero VEV $v_s/\sqrt{2}$. The terms $\mu \hat{H}_u \hat{H}_d$ and $\lambda_H \hat{S} \hat{H}_u \hat{H}_d$ can produce an effective $\mu_{eff} = \mu + \lambda_H v_s/\sqrt{2}$, which relieves the $\mu$ problem. Comparing with the condition in MSSM, the lightest CP-even Higgs mass at tree level is improved. The second light neutral CP-even Higgs can be at TeV order. Then it easily satisfies the constraints for heavy Higgs from experiments. Secondly, two models contain different parameters, so there are big differences in the analytical calculations, analyses at the analytical level and numerical discussions. In our work, we further discuss the numerical results changing with sensitive parameters within $4\sigma$, $3\sigma$ and $2\sigma$ specifically. We study the two-dimensional distribution of sensitive parameters under experimental constraints, and then the influences of some sensitive parameters on $Br(l_j^- \to l_i^- \gamma)$ are discussed by one-dimensional graphs. There are some differences in numerical discussion methods and ideas comparing with LFV decays $l_j \to l_i \gamma$ in the $U(1)_X SSM$.

Depended on the mass eigenstates of the particles and rotation matrixes, the mass eigenstate method is often not intuitive and clear enough to find the sensitive parameters, which will lead us to pay too much attention on many unimportant parameters. However, the

MIA method provides very simple analytic formulae for the form factors, involved which can be written explicitly in terms of the sensitive parameters after a proper expansion. We can easily find the direct impacts of sensitive parameters on CLFV at the analytic level. Therefore, using MIA method to study CLFV processes will provide a new way to study other CLFV processes in the future.

The paper is organized as follows. In Sec.II, we introduce the B-LSSM briefly including the superpotential and the general soft breaking terms. In Sec.III, we give analytic expressions for muon MDM and the CLFV ratios $Br(l_j^- \to l_i^- \gamma)$ in the B-LSSM. The numerical analyses are given in Sec.IV, and the conclusion is discussed in Sec.V. The tedious formulae are collected in Appendix A. In Appendix B, we discuss chirality flips with two examples, and demonstrate that the chirality flips occurring in the internal gaugino lines may yield dominant contributions comparing in the external lepton lines. In Appendix C, we emphasize the contributions from the incident lepton is dominant by comparing the amplitudes of FIG.2(a1) and (a2).

## II. THE B-LSSM

### A. The B-LSSM

The B-LSSM extends the superfields of the MSSM by introducing $U(1)_{B-L}$ gauge superfield. Therefore, the local gauge group of the B-LSSM is defined as $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \otimes U(1)_{B-L}$. Compared with the MSSM, the B-LSSM adds two singlet Higgs superfields $\hat{\eta}$ and $\hat{\bar{\eta}}$ and three generations of right-handed neutrinos superfields $\hat{\nu}_i^c$. In the TABLE II, we discuss the quantum numbers of gauge symmetry group for the chiral fields in the B-LSSM. Then the B-LSSM superpotential is deduced as

$$\begin{aligned} W_{B-L} = & Y_{u,ij}\hat{Q}_i\hat{H}_u\hat{U}_j^c - Y_{d,ij}\hat{Q}_i\hat{H}_d\hat{D}_j^c + Y_{e,ij}\hat{L}_i\hat{H}_d\hat{E}_j^c + \mu_H\hat{H}_d\hat{H}_u \\ & + Y_{x,ij}\hat{\nu}_i^c\hat{\eta}\hat{\nu}_j^c + Y_{\nu,ij}\hat{L}_i\hat{H}_u\hat{\nu}_j^c - \mu_\eta\hat{\eta}\hat{\bar{\eta}}, \end{aligned} \tag{1}$$

where $i, j$ represent the generation indices, $Y_{u,ij}, Y_{d,ij}, Y_{e,ij}, Y_{x,ij}$ and $Y_{\nu,ij}$ correspond to the Yukawa coupling coefficients. $\mu_H$ and $\mu_\eta$ are both the parameters with mass dimension. $\mu_H$

TABLE II: The chiral superfields and quantum numbers in the B-LSSM.

| Superfield | Spin 0 | Spin $\frac{1}{2}$ | Generations | $U(1)_Y \otimes \mathrm{SU}(2)_L \otimes \mathrm{SU}(3)_C \otimes U(1)_{B-L}$ |
|---|---|---|---|---|
| $\hat{H}_d$ | $H_d$ | $\tilde{H}_d$ | 1 | $(-\frac{1}{2}, \mathbf{2}, \mathbf{1}, 0)$ |
| $\hat{H}_u$ | $H_u$ | $\tilde{H}_u$ | 1 | $(\frac{1}{2}, \mathbf{2}, \mathbf{1}, 0)$ |
| $\hat{Q}_i$ | $\tilde{Q}_i$ | $Q_i$ | 3 | $(\frac{1}{6}, \mathbf{2}, \mathbf{3}, \frac{1}{6})$ |
| $\hat{L}_i$ | $\tilde{L}_i$ | $L_i$ | 3 | $(-\frac{1}{2}, \mathbf{2}, \mathbf{1}, -\frac{1}{2})$ |
| $\hat{D}_i^c$ | $\tilde{D}_i^c$ | $D_i^c$ | 3 | $(\frac{1}{3}, \mathbf{1}, \overline{\mathbf{3}}, -\frac{1}{6})$ |
| $\hat{U}_i^c$ | $\tilde{U}_i^c$ | $U_i^c$ | 3 | $(-\frac{2}{3}, \mathbf{1}, \overline{\mathbf{3}}, -\frac{1}{6})$ |
| $\hat{E}_i^c$ | $\tilde{E}_i^c$ | $E_i^c$ | 3 | $(1, \mathbf{1}, \mathbf{1}, \frac{1}{2})$ |
| $\hat{\nu}_i^c$ | $\tilde{\nu}_i^c$ | $\nu_i^c$ | 3 | $(0, \mathbf{1}, \mathbf{1}, \frac{1}{2})$ |
| $\hat{\eta}$ | $\eta$ | $\tilde{\eta}$ | 1 | $(0, \mathbf{1}, \mathbf{1}, -1)$ |
| $\hat{\bar{\eta}}$ | $\bar{\eta}$ | $\tilde{\bar{\eta}}$ | 1 | $(0, \mathbf{1}, \mathbf{1}, 1)$ |

indicates the supersymmetric mass between $SU(2)_L$ Higgs doublets $\hat{H}_d$ and $\hat{H}_u$, as well as $\mu_\eta$ represents the supersymmetric mass between $U(1)_{B-L}$ Higgs singlets $\hat{\eta}$ and $\hat{\bar{\eta}}$.

In the B-LSSM, the Higgs doublets and Higgs singlets obtain the nonzero vacuum expectation values, then the $SU(2)_L \otimes U(1)_Y \otimes U(1)_{B-L}$ gauge group breaks to $U(1)_{em}$.

$$\begin{aligned} H_d^0 &= \frac{1}{\sqrt{2}}(\phi_d + v_d + i\sigma_d), \quad H_u^0 = \frac{1}{\sqrt{2}}(\phi_u + v_u + i\sigma_u), \\ \eta &= \frac{1}{\sqrt{2}}(\phi_\eta + v_\eta + i\sigma_\eta), \qquad \bar{\eta} = \frac{1}{\sqrt{2}}(\phi_{\bar{\eta}} + v_{\bar{\eta}} + i\sigma_{\bar{\eta}}). \end{aligned} \tag{2}$$

Here, we define $u^2 = v_\eta^2 + v_{\bar{\eta}}^2, v^2 = v_d^2 + v_u^2$ and $\tan\beta' = \frac{v_{\bar{\eta}}}{v_\eta}$ in analogy to the definition $\tan\beta = \frac{v_u}{v_d}$ in the MSSM.

Correspondingly, the soft breaking terms in the B-LSSM are generally written as

$$\begin{aligned} \mathcal{L}_{soft} = & -m_{\tilde{q},ij}^2 \tilde{Q}_i^* \tilde{Q}_j - m_{\tilde{u},ij}^2 \tilde{U}_i^* \tilde{U}_j - m_{\tilde{d},ij}^2 (\tilde{D}_i^c)^* \tilde{D}_j^c - m_{\tilde{L},ij}^2 \tilde{L}_i^* \tilde{L}_j - m_{\tilde{E},ij}^2 (\tilde{E}_i^c)^* \tilde{E}_j^c \\ & - m_{H_d}^2 |H_d|^2 - m_{H_u}^2 |H_u|^2 - m_\eta^2 |\eta|^2 - m_{\bar{\eta}}^2 |\bar{\eta}|^2 - m_{\tilde{\nu},ij}^2 (\tilde{\nu}_i^c)^* \tilde{\nu}_j^c + \Big[ - B_\mu H_d H_u \\ & - B_\eta \eta\bar{\eta} + T_u^{ij} \tilde{Q}_i \tilde{U}_j^c H_u + T_d^{ij} \tilde{Q}_i \tilde{D}_j^c H_d + T_e^{ij} \tilde{L}_i \tilde{E}_j^c H_u + T_\nu^{ij} H_u \tilde{\nu}_i^c \tilde{L}_j + T_x^{ij} \eta \tilde{\nu}_i^c \tilde{\nu}_j^c \\ & - \frac{1}{2}(M_1 \lambda_{\tilde{B}} \lambda_{\tilde{B}} + M_2 \lambda_{\tilde{W}} \lambda_{\tilde{W}} + M_3 \lambda_{\tilde{g}} \lambda_{\tilde{g}} + 2M_{BB'} \lambda_{\tilde{B}'} \lambda_{\tilde{B}} + M_{B'} \lambda_{\tilde{B}'} \lambda_{\tilde{B}'}) + h.c. \Big], \end{aligned} \tag{3}$$

where $\lambda_{\tilde{B}}, \lambda_{\tilde{W}}, \lambda_{\tilde{g}}$ and $\lambda_{\tilde{B}'}$ are the gauginos of $U(1)_Y, SU(2)_L, SU(3)_C$ and $U(1)_{B-L}$ respec-

tively. Besides, the soft breaking terms of the B-LSSM include the mass squared terms of squarks, sleptons, sneutrinos and Higgs bosons, the trilinear scalar coupling terms and the Majorana mass terms.

Comparing with the MSSM or other SUSY models, the two Abelian groups in the B-LSSM produce a new effect called as the gauge kinetic mixing. Although both approaches are equivalent, it is easier to work with noncanonical covariant derivatives instead of off-diagonal field-strength tensors in practice. Hence, the covariant derivatives of the B-LSSM can be considered as

$$D_\mu = \partial_\mu - i\,(Y, B-L) \begin{pmatrix} g_Y, & g'_{YB} \\ g'_{BY}, & g_{B-L} \end{pmatrix} \begin{pmatrix} A'^{Y}_\mu \\ A'^{BL}_\mu \end{pmatrix}, \tag{4}$$

where $Y$ and $B-L$ correspond to the hypercharge and B-L charge, as well as $A'^{Y}_\mu$ and $A'^{BL}_\mu$ denote the gauge fields of $U(1)_Y$ and $U(1)_{B-L}$. With the condition of the two Abelian gauge groups unbroken, choosing matrix R in a proper form, one can write the coupling matrix as

$$\begin{pmatrix} g_Y, & g'_{YB} \\ g'_{BY}, & g_{B-L} \end{pmatrix} R^T = \begin{pmatrix} g_1, & g_{YB} \\ 0, & g_B \end{pmatrix}, \tag{5}$$

here, $g_1$ corresponds to the measured hypercharge coupling which is modified in B-LSSM as given along with $g_B$ and $g_{YB}$ in Refs.[36]. Then, we can redefine the $U(1)$ gauge fields

$$R \begin{pmatrix} A'^{Y}_\mu \\ A'^{BL}_\mu \end{pmatrix} = \begin{pmatrix} A^{Y}_\mu \\ A^{BL}_\mu \end{pmatrix}. \tag{6}$$

The one-loop corrections for the CLFV processes are related with the mass matrices, which can be obtained by SARAH[37, 38]. Besides, the CLFV processes by MIA need to consider the trilinear couplings under the interaction eigenstate, so we show some couplings needed in this work as follows. The lepton-charginos-CP even(odd) sneutrinos are deduced as:

$$\begin{aligned} \mathcal{L}_{\bar{l}_j\chi^-\tilde{\nu}^R} &= \frac{i}{\sqrt{2}}\bar{l}_j\tilde{\nu}^R_L[Y_l^j P_L\tilde{H}^- + g_2 P_R\tilde{W}^-], \\ \mathcal{L}_{\bar{l}_j\chi^-\tilde{\nu}^I} &= \frac{1}{\sqrt{2}}\bar{l}_j\tilde{\nu}^I_L[-Y_l^j P_L\tilde{H}^- + g_2 P_R\tilde{W}^-]. \end{aligned} \tag{7}$$

The lepton-neutralinos-sleptons are deduced as:

$$\mathcal{L}_{\bar{l}_j\chi^0\tilde{L}} = i\bar{l}_j\Big\{ -\Big[\frac{1}{\sqrt{2}}\Big(2g_1P_L\lambda_{\tilde{B}} + (g_B+2g_{YB})P_L\lambda_{\tilde{B}'}\Big)\tilde{L}^R + Y_l^j P_L\tilde{H}_d^0\tilde{L}^L\Big] + \Big[\frac{1}{\sqrt{2}}\Big(g_2P_R\tilde{W}^0 + g_1P_R\lambda_{\tilde{B}} + (g_B+g_{YB})P_R\lambda_{\tilde{B}'}\Big)\tilde{L}^L - Y_l^j P_R\tilde{H}_d^0\tilde{L}^R\Big]\Big\}. \quad (8)$$

### B. Higgs mass in the B-LSSM

Due that the strict constraint from SM-like Higgs boson on the numerical results, we discuss the Higgs boson mass matrix. The $\phi_d, \phi_u, \phi_\eta, \phi_{\bar{\eta}}$ mix together at the tree level. In the base $(\phi_d, \phi_u, \phi_\eta, \phi_{\bar{\eta}})$, the tree-level mass squared matrix for neutral CP-even Higgs boson is deduced as:

$$M_h^2 = u^2\times\begin{pmatrix} \frac{1}{4}\frac{g^2x^2}{1+\tan\beta^2}+n^2\tan\beta & -\frac{1}{4}g^2\frac{x^2\tan\beta}{1+\tan\beta^2}-n^2 & \frac{1}{2}g_Bg_{YB}\frac{x}{T} & -\frac{1}{2}g_Bg_{YB}\frac{x\times\tan\beta'}{T} \\ -\frac{1}{4}g^2\frac{x^2\tan\beta}{1+\tan\beta^2}-n^2 & \frac{1}{4}\frac{g^2x^2\tan\beta^2}{1+\tan\beta^2}+\frac{n^2}{\tan\beta} & \frac{1}{2}g_Bg_{YB}\frac{x\times\tan\beta}{T} & \frac{1}{2}g_Bg_{YB}\frac{x\times\tan\beta\times\tan\beta'}{T} \\ \frac{1}{2}g_Bg_{YB}\frac{x}{T} & \frac{1}{2}g_Bg_{YB}\frac{x\times\tan\beta}{T} & \frac{g_B^2}{1+\tan\beta'^2}+\tan\beta' N^2 & -g_B^2\frac{\tan\beta'}{1+\tan\beta'^2}-N^2 \\ -\frac{1}{2}g_Bg_{YB}\frac{x\times\tan\beta'}{T} & \frac{1}{2}g_Bg_{YB}\frac{x\times\tan\beta\times\tan\beta'}{T} & -g_B^2\frac{\tan\beta'}{1+\tan\beta'^2}-N^2 & g_B^2\frac{\tan\beta'^2}{1+\tan\beta'^2}+\frac{N^2}{\tan\beta'} \end{pmatrix}. \quad (9)$$

Here, $g^2 = g_1^2+g_2^2+g_{YB}^2$, $T = \sqrt{1+\tan\beta^2}\sqrt{1+\tan\beta'^2}$, $x = \frac{v}{u}$, $n^2 = \frac{\mathrm{Re}(B\mu)}{u^2}$ and $N^2 = \frac{\mathrm{Re}(B_\eta)}{u^2}$. The mass of the SM-like Higgs boson can be obtained after considering the leading-log radiative corrections from stop and top particles[39–41].

$$m_h = \sqrt{(m_{h_1}^0)^2 + \Delta m_h^2}, \quad (10)$$

where $m_{h_1}^0$ represents the lightest tree-level Higgs boson mass, and the leading-log radiative corrections $\Delta m_h^2$ can be written as

$$\Delta m_h^2 = \frac{3m_t^4}{4\pi^2v^2}\Big[\Big(\tilde{t}+\frac{1}{2}\tilde{X}_t\Big) + \frac{1}{16\pi^2}\Big(\frac{3m_t^2}{2v^2} - 32\pi\alpha_3\Big)(\tilde{t}^2+\tilde{X}_t\tilde{t})\Big],$$
$$\tilde{t} = \log\frac{M_S^2}{m_t^2}, \quad \tilde{X}_t = \frac{2\tilde{A}_t^2}{M_S^2}(1-\frac{\tilde{A}_t^2}{12M_S^2}). \quad (11)$$

Here, $\alpha_3$ is the strong coupling constant, $M_S = \sqrt{m_{\tilde{t}_1}m_{\tilde{t}_2}}$ with $m_{\tilde{t}_{1,2}}$ are the stop masses, $\tilde{A}_t = A_t - \mu_H\cot\beta$ with $A_t = T_{u,33}$ denotes the trilinear Higgs-stops coupling.

## III. THE $(g-2)_\mu$ AND $l_j^- \to l_i^-\gamma$ IN THE B-LSSM

In this section, we study the muon MDM and the CLFV processes $l_j^- \to l_i^-\gamma$ in the B-LSSM with the MIA method, which shows the factor of the one-loop contributions more clearly. The concrete contents will be discussed as follows.

### A. The one-loop corrections to $(g-2)_\mu$ in the B-LSSM

The effective Lagrangian used here for the muon MDM is given out as follows

$$\mathcal{L}_{\rm MDM} = \frac{e}{4m_\mu}\, a_\mu\, \bar{l}_\mu \sigma^{\alpha\beta} l_\mu\, F_{\alpha\beta}, \tag{12}$$

where $\sigma_{\alpha\beta} = i[\gamma_\alpha, \gamma_\beta]/2$, $l_\mu$ denotes the wave function of muon, $m_\mu$ represents the muon mass, $F_{\alpha\beta}$ is the electromagnetic field strength and $a_\mu$ is the muon MDM.

To obtain the muon MDM, we adopt the effective Lagrangian method[28, 42, 43], which relates with the following dimension-6 operators.

$$\begin{aligned}
\mathcal{O}_1^{\mp} &= \frac{1}{(4\pi)^2}\bar{l}(i\not{\mathcal{D}})^3 P_{L,R} l, & \mathcal{O}_2^{\mp} &= \frac{eQ_f}{(4\pi)^2}\overline{(i\mathcal{D}_\mu l)}\gamma^\mu F\cdot\sigma P_{L,R} l,\\
\mathcal{O}_3^{\mp} &= \frac{eQ_f}{(4\pi)^2}\bar{l}F\cdot\sigma\gamma^\mu P_{L,R}(i\mathcal{D}_\mu l), & \mathcal{O}_4^{\mp} &= \frac{eQ_f}{(4\pi)^2}\bar{l}(\partial^\mu F_{\mu\nu})\gamma^\nu P_{L,R} l,\\
\mathcal{O}_5^{\mp} &= \frac{m_l}{(4\pi)^2}\bar{l}(i\not{\mathcal{D}})^2 P_{L,R} l, & \mathcal{O}_6^{\mp} &= \frac{eQ_f m_l}{(4\pi)^2}\bar{l}F\cdot\sigma P_{L,R} l,
\end{aligned} \tag{13}$$

with $\mathcal{D}_\mu = \partial_\mu + ieA_\mu$. The operators $\mathcal{O}_{2,3,6}^{\mp}$ have relation with muon MDM when adopting on-shell condition for external leptons. Therefore, we only study the Wilson coefficients $C_{2,3,6}^{\mp}$ of the operators $\mathcal{O}_{2,3,6}^{\mp}$ in the effective Lagrangian. Actually, the Wilson coefficients satisfy the relations $C_2^{\mp} = C_3^{\mp *}$ and $C_6^{+} = C_6^{-*}$. Then the muon MDM can be deduced as

$$a_\mu = \frac{4Q_f m_\mu^2}{(4\pi)^2}\Re(C_2^+ + C_2^{-*} + C_6^+). \tag{14}$$

The feynman diagrams of muon MDM by MIA in the B-LSSM are shown in FIG.1. Then, we obtain the concrete forms of the one-loop muon MDM in the B-LSSM adopting the MIA method.

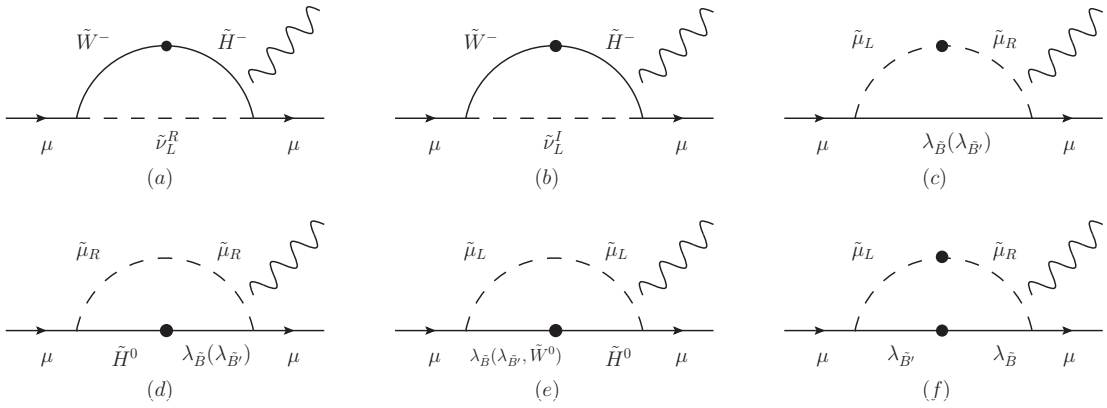



FIG. 1: The feynman diagrams of muon MDM by MIA in the B-LSSM.

1. The one-loop contributions from $\tilde{H}^- - \tilde{W}^- - \tilde{\nu}_L^R(\tilde{\nu}_L^I)$:

$$\begin{aligned}
a_\mu\left(\tilde{\nu}_L^R,\tilde{H}^-,\tilde{W}^-\right) &= \frac{g_2^2}{2}x_\mu\sqrt{x_2 x_{\mu_H}}\tan\beta\Big[2\mathcal{I}\left(x_{\mu_H},x_{\tilde{\nu}_L^R},x_2\right)-\mathcal{J}\left(x_2,x_{\mu_H},x_{\tilde{\nu}_L^R}\right)\\
&\quad+2\mathcal{I}\left(x_2,x_{\tilde{\nu}_L^R},x_{\mu_H}\right)-\mathcal{J}\left(x_{\mu_H},x_2,x_{\tilde{\nu}_L^R}\right)\Big],\\
a_\mu\left(\tilde{\nu}_L^I,\tilde{H}^-,\tilde{W}^-\right) &= \frac{g_2^2}{2}x_\mu\sqrt{x_2 x_{\mu_H}}\tan\beta\Big[2\mathcal{I}\left(x_{\mu_H},x_{\tilde{\nu}_L^I},x_2\right)-\mathcal{J}\left(x_2,x_{\mu_H},x_{\tilde{\nu}_L^I}\right)\\
&\quad+2\mathcal{I}\left(x_2,x_{\tilde{\nu}_L^I},x_{\mu_H}\right)-\mathcal{J}\left(x_{\mu_H},x_2,x_{\tilde{\nu}_L^I}\right)\Big],
\end{aligned}\tag{15}$$

with $x_\mu=\frac{m_\mu^2}{\Lambda^2}$, $x_{\mu_H}=\frac{\mu_H^2}{\Lambda^2}$, and $\Lambda$ is the NP energy scale. The one-loop functions $\mathcal{I}(x,y,z)$, $\mathcal{J}(x,y,z)$ and the following $f(x,y,z,t)$, $g(x,y,z,t)$, $k(x,y,z,t)$ are collected in Appendix.

2. The one-loop contributions from $\lambda_{\tilde{B}}\,(\lambda_{\tilde{B}'})-\tilde{\mu}_L-\tilde{\mu}_R$:

$$\begin{aligned}
a_\mu\left(\tilde{\mu}_R,\tilde{\mu}_L,\lambda_{\tilde{B}}\right) &= g_1^2 x_\mu\sqrt{x_1 x_{\mu_H}}\tan\beta\Big[\mathcal{J}\left(x_1,x_{\tilde{\mu}_L},x_{\tilde{\mu}_R}\right)+\mathcal{J}\left(x_1,x_{\tilde{\mu}_R},x_{\tilde{\mu}_L}\right)\Big],\\
a_\mu\left(\tilde{\mu}_R,\tilde{\mu}_L,\lambda_{\tilde{B}'}\right) &= \frac{(g_B+2g_{YB})(g_B+g_{YB})}{2}x_\mu\sqrt{x_{B'}x_{\mu_H}}\tan\beta\Big[\mathcal{J}\left(x_{B'},x_{\tilde{\mu}_L},x_{\tilde{\mu}_R}\right)\\
&\quad+\mathcal{J}\left(x_{B'},x_{\tilde{\mu}_R},x_{\tilde{\mu}_L}\right)\Big].
\end{aligned}\tag{16}$$

3. The one-loop contributions from $\lambda_{\tilde{B}}\,(\lambda_{\tilde{B}'})-\tilde{H}^0-\tilde{\mu}_R$:

$$\begin{aligned}
a_\mu\left(\tilde{\mu}_R,\lambda_{\tilde{B}},\tilde{H}^0\right) &= -g_1^2 x_\mu\sqrt{x_1 x_{\mu_H}}\tan\beta\Big[\mathcal{J}\left(x_1,x_{\mu_H},x_{\tilde{\mu}_R}\right)+\mathcal{J}\left(x_{\mu_H},x_1,x_{\tilde{\mu}_R}\right)\Big],\\
a_\mu\left(\tilde{\mu}_R,\lambda_{\tilde{B}'},\tilde{H}^0\right) &= -\frac{(g_B+2g_{YB})g_{YB}}{2}x_\mu\sqrt{x_{B'}x_{\mu_H}}\tan\beta\Big[\mathcal{J}\left(x_{B'},x_{\mu_H},x_{\tilde{\mu}_R}\right)\\
&\quad+\mathcal{J}\left(x_{\mu_H},x_{B'},x_{\tilde{\mu}_R}\right)\Big].
\end{aligned}\tag{17}$$

4. The one-loop contributions from $\lambda_{\tilde{B}}\left(\tilde{W}^0,\lambda_{\tilde{B}'}\right)-\tilde{H}^0-\tilde{\mu}_L$:

$$\begin{aligned}
a_\mu\left(\tilde{\mu}_L,\tilde{H}^0,\lambda_{\tilde{B}}\right)&=\frac{1}{2}g_1^2x_\mu\sqrt{x_1x_{\mu_H}}\tan\beta\Big[\mathcal{J}\left(x_1,x_{\mu_H},x_{\tilde{\mu}_L}\right)+\mathcal{J}\left(x_{\mu_H},x_1,x_{\tilde{\mu}_L}\right)\Big],\\
a_\mu\left(\tilde{\mu}_L,\tilde{H}^0,\tilde{W}^0\right)&=-\frac{1}{2}g_2^2x_\mu\sqrt{x_2x_{\mu_H}}\tan\beta\Big[\mathcal{J}\left(x_2,x_{\mu_H},x_{\tilde{\mu}_L}\right)+\mathcal{J}\left(x_{\mu_H},x_2,x_{\tilde{\mu}_L}\right)\Big],\\
a_\mu\left(\tilde{\mu}_L,\tilde{H}^0,\lambda_{\tilde{B}'}\right)&=\frac{g_{YB}(g_B+g_{YB})}{2}x_\mu\sqrt{x_{B'}x_{\mu_H}}\tan\beta\Big[\mathcal{J}\left(x_{B'},x_{\mu_H},x_{\tilde{\mu}_L}\right)\\
&\quad+\mathcal{J}\left(x_{\mu_H},x_{B'},x_{\tilde{\mu}_L}\right)\Big].
\end{aligned}\tag{18}$$

5. The one-loop contributions from $\lambda_{\tilde{B}}-\lambda_{\tilde{B}'}-\tilde{\mu}_R-\tilde{\mu}_L$:

$$\begin{aligned}
a_\mu\left(\tilde{\mu}_R,\tilde{\mu}_L,\lambda_{\tilde{B}},\lambda_{\tilde{B}'}\right)&=g_1\left(4g_{YB}+3g_B\right)x_\mu\sqrt{x_{BB'}x_{\mu_H}}\tan\beta\\
&\quad\times\Big[\sqrt{x_1x_{B'}}f\left(x_{B'},x_1,x_{\tilde{\mu}_L},x_{\tilde{\mu}_R}\right)-g\left(x_{B'},x_1,x_{\tilde{\mu}_L},x_{\tilde{\mu}_R}\right)\Big].
\end{aligned}\tag{19}$$

In Eqs.(15)-(19), one can easily find the factors $x_\mu$ and $\tan\beta$, which possess the same characteristic as these in MSSM[19]. In the B-LSSM, the new gaugino $\lambda_{\tilde{B}'}$ generates the new contributions to muon MDM, which are deduced in Eqs.(16)-(19).

To obtain clearer images of the results, we suppose that all the superparticles masses are almost degenerate. The author[19] gives the one-loop results of the MSSM in the extreme case where the masses for superparticles $M_1,M_2,\mu_H,m_{\tilde{\mu}_L},m_{\tilde{\mu}_R}$ are equal to $\Lambda$

$$a_\mu^{\mathrm{MSSM}}\simeq\frac{1}{192\pi^2}\frac{m_\mu^2}{\Lambda^2}\tan\beta\left(5g_2^2+g_1^2\right).\tag{20}$$

Here, we also use the similar case

$$M_1=M_2=\mu_H=m_{\tilde{\nu}_L^R}=m_{\tilde{\nu}_L^I}=m_{\tilde{\mu}_L}=m_{\tilde{\mu}_R}=|M_{B'}|=|M_{BB'}|=\Lambda.\tag{21}$$

Then, the functions $\mathcal{I}(x,y,z),\mathcal{J}(x,y,z),f(x,y,z,t)$ and $g(x,y,z,t)$ can be simplified as

$$\begin{aligned}
&\mathcal{I}(1,1,1)=\frac{1}{96\pi^2},\qquad &&\mathcal{J}(1,1,1)=\frac{1}{192\pi^2},\\
&f(1,1,1,1)=-\frac{1}{240\pi^2},\quad &&g(1,1,1,1)=-\frac{1}{960\pi^2}.
\end{aligned}\tag{22}$$

In this condition, we obtain the simplified one-loop contributions of muon MDM in the B-LSSM.

$$\begin{aligned}
a_\mu^{B-L}\simeq\frac{1}{960\pi^2}\frac{m_\mu^2}{\Lambda^2}\tan\beta&\Big[5\left(5g_2^2+g_1^2\right)+5\left(g_B^2+3g_{YB}g_B+g_{YB}^2\right)\mathrm{sign}\left[M_{B'}\right]\\
&+g_1\left(3g_B+4g_{YB}\right)\mathrm{sign}\left[M_{BB'}\right]\left(1-4\mathrm{sign}\left[M_{B'}\right]\right)\Big].
\end{aligned}\tag{23}$$

In addition to the one-loop MSSM contributions, the Eq.(23) adds considerable corrections to $a_\mu$ from the new gaugino $\lambda_{\tilde{B}'}$. In the condition $-0.7 < g_{YB} < -0.05$, $0.1 < g_B < 0.85$ and $|g_{YB}| < g_B < \frac{4}{3}|g_{YB}|$ and with the supposition $\text{sign}\,[M_{B'}] = \text{sign}\,[M_{BB'}] = -1$, the corrections beyond MSSM can reach large value.

$$a_\mu^{B-L} \to \frac{1}{192\pi^2}\frac{m_\mu^2}{\Lambda^2}\tan\beta\left[(5g_2^2+g_1^2) - \Big(g_1(3g_B+4g_{YB}) + (g_B^2+3g_{YB}g_B+g_{YB}^2)\Big)\right]. \quad (24)$$

Here, the order analysis shows

$$0 < \frac{-g_1(3g_B+4g_{YB}) - (g_B^2+3g_{YB}g_B+g_{YB}^2)}{5g_2^2+g_1^2} \lesssim 1. \quad (25)$$

Therefore, the B-LSSM contributions beyond MSSM are considerable.

### B. The one-loop corrections to $l_j^- \to l_i^-\gamma$ in the B-LSSM

Generally, the effective amplitude of the CLFV processes $l_j^- \to l_i^-\gamma$ can be written as

$$\mathcal{M} = e\epsilon^\mu \bar{u}_i(p+q)[q^2\gamma_\mu(C_1^L P_L + C_1^R P_R) + m_{l_j} i\sigma_{\mu\nu}q^\nu(C_2^L P_L + C_2^R P_R)]u_j(p). \quad (26)$$

where $p$ ($q$) represents the incident lepton (photon) momentum. $m_{l_j}$ is the j-th generation lepton mass. $\epsilon$ is the photon polarization vector and $u_j(p)$ ($\bar{u}_i(p+q)$) is the incident (outgoing) lepton wave function. In FIG.2, we show the relevant B-LSSM Feynman diagrams of $l_j^- \to l_i^-\gamma$ by MIA. Here, we omit the chirality flips of external lepton in our calculation. In Appendix B, we discuss chirality flips with two examples, and demonstrate that the chirality flips occurring in the internal gaugino lines may yield dominant contributions comparing in the external lepton lines. In Appendix C, we emphasize the contributions from the incident lepton is dominant by comparing the amplitudes of FIG.2(a1) and (a2). The Wilson coefficients $C_\alpha^{iL}$, $C_\alpha^{iR}(\alpha = 1,2,\ i = 1,\ldots,5)$ adopting the MIA method are obtained from the sum of these diagrams' amplitudes.

1. The one-loop contributions from $\tilde{H}^- - \tilde{W}^- - \tilde{\nu}_L^R(\tilde{\nu}_L^I)$:

$$C_2^{1L}\left(\tilde{\nu}_{L_j}^R, \tilde{\nu}_{L_i}^R, \tilde{H}^-, \tilde{W}^-\right) = \frac{g_2^2}{2\Lambda^4}\frac{m_{l_i}}{m_{l_j}}\sqrt{x_2 x_{\mu_H}}\Delta_{ij}^{LL}\tan\beta\Big[k(x_{\mu_H}, x_2, x_{\tilde{\nu}_{L_j}^R}, x_{\tilde{\nu}_{L_i}^R}) + k(x_2, x_{\mu_H}, x_{\tilde{\nu}_{L_j}^R}, x_{\tilde{\nu}_{L_i}^R}) - f(x_2, x_{\mu_H}, x_{\tilde{\nu}_{L_j}^R}, x_{\tilde{\nu}_{L_i}^R})\Big],$$

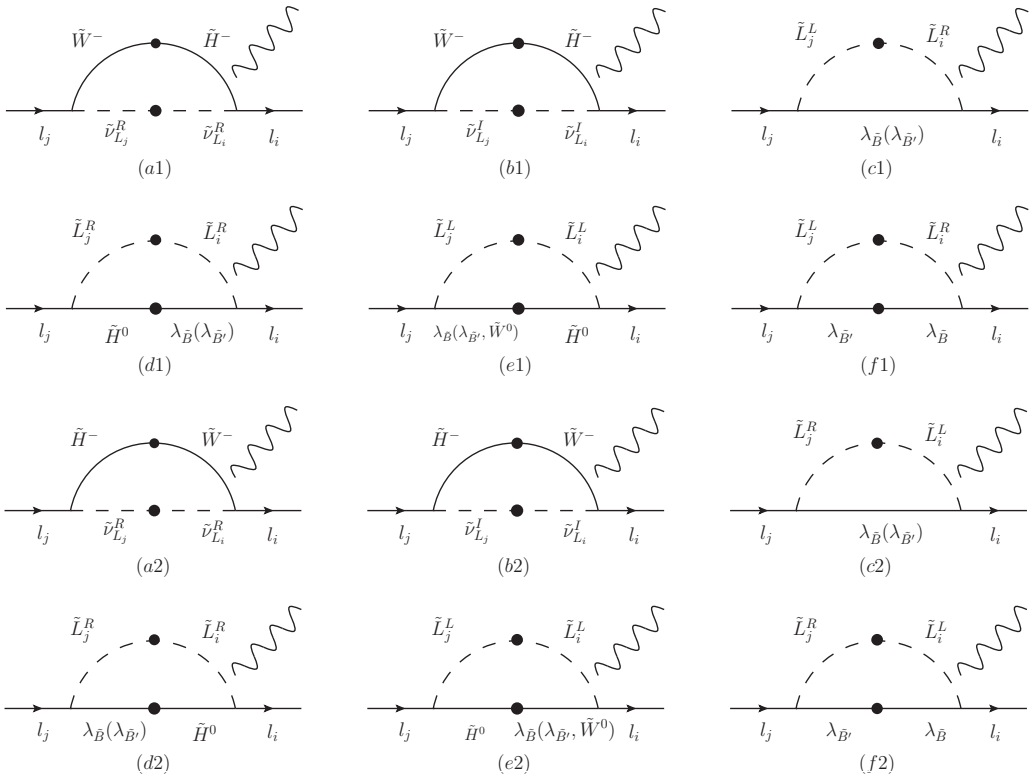



FIG. 2: The feynman diagrams of $l_j^- \to l_i^- \gamma$ by MIA in the B-LSSM.

$$C_2^{1L}(\tilde{\nu}^I_{L_j},\tilde{\nu}^I_{L_i},\tilde{H}^-,\tilde{W}^-) = \frac{g_2^2}{2\Lambda^4}\frac{m_{l_i}}{m_{l_j}}\sqrt{x_2 x_{\mu_H}}\Delta^{LL}_{ij}\tan\beta\Big[k(x_{\mu_H},x_2,x_{\tilde{\nu}^I_{L_j}},x_{\tilde{\nu}^I_{L_i}}) + k(x_2,x_{\mu_H},x_{\tilde{\nu}^I_{L_j}},x_{\tilde{\nu}^I_{L_i}}) - f(x_2,x_{\mu_H},x_{\tilde{\nu}^I_{L_j}},x_{\tilde{\nu}^I_{L_i}})\Big], \tag{27}$$

$$C_2^{1R}\left(\tilde{\nu}^R_{L_j},\tilde{\nu}^R_{L_i},\tilde{H}^-,\tilde{W}^-\right) = \frac{g_2^2}{2\Lambda^4}\sqrt{x_2 x_{\mu_H}}\Delta^{LL}_{ij}\tan\beta\Big[k(x_{\mu_H},x_2,x_{\tilde{\nu}^R_{L_j}},x_{\tilde{\nu}^R_{L_i}}) + k(x_2,x_{\mu_H},x_{\tilde{\nu}^R_{L_j}},x_{\tilde{\nu}^R_{L_i}}) - f(x_2,x_{\mu_H},x_{\tilde{\nu}^R_{L_j}},x_{\tilde{\nu}^R_{L_i}})\Big],$$

$$C_2^{1R}(\tilde{\nu}^I_{L_j},\tilde{\nu}^I_{L_i},\tilde{H}^-,\tilde{W}^-) = \frac{g_2^2}{2\Lambda^4}\sqrt{x_2 x_{\mu_H}}\Delta^{LL}_{ij}\tan\beta\Big[k(x_{\mu_H},x_2,x_{\tilde{\nu}^I_{L_j}},x_{\tilde{\nu}^I_{L_i}}) + k(x_2,x_{\mu_H},x_{\tilde{\nu}^I_{L_j}},x_{\tilde{\nu}^I_{L_i}}) - f(x_2,x_{\mu_H},x_{\tilde{\nu}^I_{L_j}},x_{\tilde{\nu}^I_{L_i}})\Big]. \tag{28}$$

2. The one-loop contributions from $\lambda_{\tilde{B}}\,(\lambda_{\tilde{B}'}) - \tilde{L}^L_j - \tilde{L}^R_i$:

$$C_2^{2L}\left(\tilde{L}^L_j,\tilde{L}^R_i,\lambda_{\tilde{B}}\right) = -\frac{g_1^2}{2\Lambda^3 m_{l_j}}\sqrt{x_1}\Delta^{LR}_{ij}\Big[\mathcal{J}(x_1,x_{\tilde{L}^L_j},x_{\tilde{L}^R_i}) + \mathcal{J}(x_1,x_{\tilde{L}^R_i},x_{\tilde{L}^L_j})\Big],$$

$$C_2^{2L}\left(\tilde{L}^L_j,\tilde{L}^R_i,\lambda_{\tilde{B}'}\right) = -\frac{(g_B+2g_{YB})(g_B+g_{YB})}{4\Lambda^3 m_{l_j}}\sqrt{x_{B'}}\Delta^{LR}_{ij}\Big[\mathcal{J}(x_{B'},x_{\tilde{L}^L_j},x_{\tilde{L}^R_i})$$

$$+\mathcal{J}(x_{B'}, x_{\tilde{L}_i^R}, x_{\tilde{L}_j^L})\Big], \tag{29}$$

$$C_2^{2R}\left(\tilde{L}_j^R, \tilde{L}_i^L, \lambda_{\tilde{B}}\right) = -\frac{g_1^2}{2\Lambda^3 m_{l_j}}\sqrt{x_1}\Delta_{ij}^{LR}\Big[\mathcal{J}(x_1, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^L}) + \mathcal{J}(x_1, x_{\tilde{L}_i^L}, x_{\tilde{L}_j^R})\Big],$$
$$C_2^{2R}\left(\tilde{L}_j^R, \tilde{L}_i^L, \lambda_{\tilde{B}'}\right) = -\frac{(g_B+2g_{YB})(g_B+g_{YB})}{4\Lambda^3 m_{l_j}}\sqrt{x_{B'}}\Delta_{ij}^{LR}\Big[\mathcal{J}(x_{B'}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^L})$$
$$+\mathcal{J}(x_{B'}, x_{\tilde{L}_i^L}, x_{\tilde{L}_j^R})\Big]. \tag{30}$$

3. The one-loop contributions from $\lambda_{\tilde{B}}\,(\lambda_{\tilde{B}'}) - \tilde{H}^0 - \tilde{L}_j^R - \tilde{L}_i^R$:

$$C_2^{3L}\left(\tilde{L}_j^R, \tilde{L}_i^R, \lambda_{\tilde{B}}, \tilde{H}^0\right) = -\frac{g_1^2}{\Lambda^4}\Delta_{ij}^{RR}\Big[\sqrt{x_1 x_{\mu_H}}\tan\beta f(x_1, x_{\mu_H}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^R}) - g(x_1, x_{\mu_H}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^R})\Big],$$
$$C_2^{3L}\left(\tilde{L}_j^R, \tilde{L}_i^R, \lambda_{\tilde{B}'}, \tilde{H}^0\right) = -\frac{(g_B+2g_{YB})g_{YB}}{2\Lambda^4}\Delta_{ij}^{RR}\Big[\sqrt{x_{B'}x_{\mu_H}}\tan\beta f(x_{B'}, x_{\mu_H}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^R})$$
$$-g(x_{B'}, x_{\mu_H}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^R})\Big], \tag{31}$$

$$C_2^{3R}\left(\tilde{L}_j^R, \tilde{L}_i^R, \lambda_{\tilde{B}}, \tilde{H}^0\right) = -\frac{g_1^2}{\Lambda^4}\frac{m_{l_i}}{m_{l_j}}\Delta_{ij}^{RR}\Big[\sqrt{x_1 x_{\mu_H}}\tan\beta f(x_1, x_{\mu_H}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^R})$$
$$-g(x_1, x_{\mu_H}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^R})\Big],$$
$$C_2^{3R}\left(\tilde{L}_j^R, \tilde{L}_i^R, \lambda_{\tilde{B}'}, \tilde{H}^0\right) = -\frac{(g_B+2g_{YB})g_{YB}}{2\Lambda^4}\frac{m_{l_i}}{m_{l_j}}\Delta_{ij}^{RR}\Big[\sqrt{x_{B'}x_{\mu_H}}\tan\beta f(x_{B'}, x_{\mu_H}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^R})$$
$$-g(x_{B'}, x_{\mu_H}, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^R})\Big]. \tag{32}$$

4. The one-loop contributions from $\lambda_{\tilde{B}}\left(\tilde{W}^0, \lambda_{\tilde{B}'}\right) - \tilde{H}^0 - \tilde{L}_j^L - \tilde{L}_i^L$:

$$C_2^{4L}\left(\tilde{L}_j^L, \tilde{L}_i^L, \tilde{H}^0, \lambda_{\tilde{B}}\right) = \frac{g_1^2}{2\Lambda^4}\frac{m_{l_i}}{m_{l_j}}\Delta_{ij}^{LL}\Big[\sqrt{x_1 x_{\mu_H}}\tan\beta f(x_1, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})$$
$$-g(x_1, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})\Big],$$
$$C_2^{4L}\left(\tilde{L}_j^L, \tilde{L}_i^L, \tilde{H}^0, \tilde{W}^0\right) = -\frac{g_2^2}{2\Lambda^4}\frac{m_{l_i}}{m_{l_j}}\Delta_{ij}^{LL}\Big[\sqrt{x_2 x_{\mu_H}}\tan\beta f(x_2, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})$$
$$-g(x_2, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})\Big],$$
$$C_2^{4L}\left(\tilde{L}_j^L, \tilde{L}_i^L, \tilde{H}^0, \lambda_{\tilde{B}'}\right) = \frac{g_{YB}(g_B+g_{YB})}{2\Lambda^4}\frac{m_{l_i}}{m_{l_j}}\Delta_{ij}^{LL}\Big[\sqrt{x_{B'}x_{\mu_H}}\tan\beta f(x_{B'}, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})$$
$$-g(x_{B'}, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})\Big], \tag{33}$$

$$C_2^{4R}\left(\tilde{L}_j^L, \tilde{L}_i^L, \tilde{H}^0, \lambda_{\tilde{B}}\right) = \frac{g_1^2}{2\Lambda^4}\Delta_{ij}^{LL}\Big[\sqrt{x_1 x_{\mu_H}}\tan\beta f(x_1, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})$$

$$-g(x_1, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})\Big],$$

$$C_2^{4R}\left(\tilde{L}_j^L, \tilde{L}_i^L, \tilde{H}^0, \tilde{W}^0\right) = -\frac{g_2^2}{2\Lambda^4}\Delta_{ij}^{LL}\Big[\sqrt{x_2 x_{\mu_H}}\tan\beta f(x_2, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})$$
$$-g(x_2, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})\Big],$$

$$C_2^{4R}\left(\tilde{L}_j^L, \tilde{L}_i^L, \tilde{H}^0, \lambda_{\tilde{B}'}\right) = \frac{g_{YB}(g_B + g_{YB})}{2\Lambda^4}\Delta_{ij}^{LL}\Big[\sqrt{x_{B'} x_{\mu_H}}\tan\beta f(x_{B'}, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})$$
$$-g(x_{B'}, x_{\mu_H}, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^L})\Big]. \tag{34}$$

5. The one-loop contributions from $\lambda_{\tilde{B}} - \lambda_{\tilde{B}'} - \tilde{L}_j^L - \tilde{L}_i^R$:

$$C_2^{5L}\left(\tilde{L}_j^L, \tilde{L}_i^R, \lambda_{\tilde{B}}, \lambda_{\tilde{B}'}\right) = -\frac{g_1(4g_{YB} + 3g_B)}{2\Lambda^3 m_{l_j}}\sqrt{x_{BB'}}\Delta_{ij}^{LR}\Big[\sqrt{x_1 x_{B'}} f(x_{B'}, x_1, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^R})$$
$$-g(x_{B'}, x_1, x_{\tilde{L}_j^L}, x_{\tilde{L}_i^R})\Big], \tag{35}$$

$$C_2^{5R}\left(\tilde{L}_j^R, \tilde{L}_i^L, \lambda_{\tilde{B}}, \lambda_{\tilde{B}'}\right) = -\frac{g_1(4g_{YB} + 3g_B)}{2\Lambda^3 m_{l_j}}\sqrt{x_{BB'}}\Delta_{ij}^{LR}\Big[\sqrt{x_1 x_{B'}} f(x_{B'}, x_1, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^L})$$
$$-g(x_{B'}, x_1, x_{\tilde{L}_j^R}, x_{\tilde{L}_i^L})\Big]. \tag{36}$$

Then the decay width of $l_j^- \to l_i^-\gamma$ can be written as

$$\Gamma(l_j^- \to l_i^-\gamma) = \frac{e^2}{16\pi} m_{l_j}^5 (|C_2^L|^2 + |C_2^R|^2), \tag{37}$$

with the final Wilson coefficient $C_2^{L,R} = \sum_i^{i=1\cdots5} C_2^{iL,iR},\ i = 1, \cdots, 5$. And the branching ratio for $l_j^- \to l_i^-\gamma$ is

$$Br(l_j^- \to l_i^-\gamma) = \Gamma(l_j^- \to l_i^-\gamma)/\Gamma_{l_j}. \tag{38}$$

From the Eqs.(27)-(36), we find that $C_2^{L,R}$ are almost affected by $\tan\beta$ and $\Delta_{ij}^{AB}(A, B = L, R)$. Here, $\Delta_{ij}^{LL} = m_L^2\delta_{ij}^{LL}, \Delta_{ij}^{RR} = m_E^2\delta_{ij}^{RR}$ and $\Delta_{ij}^{LR} = A_e v_d \delta_{ij}^{LR}$, which are related with soft breaking slepton mass squared matrices $m_{\tilde{L},\tilde{E}}^2$ and the trilinear coupling matrix $T_e$, whose off-diagonal terms introduce the slepton flavor mixing.

$$m_{\tilde{L}}^2 = \begin{pmatrix} 1, & \delta_{12}^{LL}, & \delta_{13}^{LL} \\ \delta_{12}^{LL}, & 1, & \delta_{23}^{LL} \\ \delta_{13}^{LL}, & \delta_{23}^{LL}, & 1 \end{pmatrix} m_L^2, \quad m_{\tilde{E}}^2 = \begin{pmatrix} 1, & \delta_{12}^{RR}, & \delta_{13}^{RR} \\ \delta_{12}^{RR}, & 1, & \delta_{23}^{RR} \\ \delta_{13}^{RR}, & \delta_{23}^{RR}, & 1 \end{pmatrix} m_E^2, \quad T_e = \begin{pmatrix} 1, & \delta_{12}^{LR}, & \delta_{13}^{LR} \\ \delta_{12}^{LR}, & 1, & \delta_{23}^{LR} \\ \delta_{13}^{LR}, & \delta_{23}^{LR}, & 1 \end{pmatrix} A_e, \tag{39}$$

In the subsequent numerical analyses, we discuss the branching ratios for CLFV processes $l_j^- \to l_i^- \gamma$ in the B-LSSM depending on the slepton mixing parameters.

In order to more intuitively analyze the factors that affect CLFV processes $l_j^- \to l_i^- \gamma$, we also suppose that the superparticles masses are almost degenerate. In other words, we give the one-loop results in the extreme case where the super particle masses are equal to $\Lambda$.

$$|M_1| = |M_2| = |\mu_H| = m_L = m_E = |M_{B'}| = |M_{BB'}| = \Lambda. \tag{40}$$

Here, the function $k(x, y, z, t)$ is much simplified as $k(1, 1, 1, 1) = -\frac{1}{192\pi^2}$, then the Wilson coefficients from the MSSM can be deduced as

$$\begin{aligned} C_2^{\mathrm{L,MSSM}} \simeq \frac{1}{1920\pi^2\Lambda^4 m_{l_j}}\Big[ & -10\sqrt{2}g_1^2\Lambda A_e v_d \mathrm{sign}\left[M_1\right]\delta_{ij}^{LR} \\ & +2g_1^2(4\tan\beta \mathrm{sign}\left[M_1\mu_H\right]-1)m_{l_j}m_E^2\delta_{ij}^{RR}\Big], \end{aligned} \tag{41}$$

$$\begin{aligned} C_2^{\mathrm{R,MSSM}} \simeq \frac{1}{1920\pi^2\Lambda^4 m_{l_j}}\Big[ & -10\sqrt{2}g_1^2\Lambda A_e v_d \mathrm{sign}\left[M_1\right]\delta_{ij}^{LR} - \Big(g_2^2(8\tan\beta \mathrm{sign}\left[M_2\mu_H\right]+1) \\ & +g_1^2(4\tan\beta \mathrm{sign}\left[M_1\mu_H\right]-1)\Big)m_{l_j}m_L^2\delta_{ij}^{LL}\Big]. \end{aligned} \tag{42}$$

With the supposition $\mathrm{sign}\left[M_1\right] = \mathrm{sign}\left[\mu_H\right] = \mathrm{sign}\left[M_2\right] = 1$ and $4\tan\beta >> 1$, the one-loop corrections to the Wilson coefficient of MSSM can reach

$$C_2^{\mathrm{L,MSSM}} \to \frac{1}{1920\pi^2\Lambda^2 m_{l_j}}\Big[ -10\sqrt{2}g_1^2\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} + 8g_1^2\tan\beta m_{l_j}\delta_{ij}^{RR}\Big], \tag{43}$$

$$C_2^{\mathrm{R,MSSM}} \to \frac{1}{1920\pi^2\Lambda^2 m_{l_j}}\Big[ -10\sqrt{2}g_1^2\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} - (8g_2^2+4g_1^2)\tan\beta m_{l_j}\delta_{ij}^{LL}\Big]. \tag{44}$$

In the same way, the Wilson coefficient from the B-LSSM can be written as

$$\begin{aligned} C_2^{\mathrm{L,B-L}} \simeq -\frac{1}{3840\pi^2\Lambda^4 m_{l_j}}\Big[ & \sqrt{2}\Big(20g_1^2\mathrm{sign}\left[M_1\right] + 10(g_B+g_{YB})(g_B+2g_{YB})\mathrm{sign}\left[M_{B'}\right] \\ & -g_1(3g_B+4g_{YB})\mathrm{sign}\left[M_{BB'}\right](4\mathrm{sign}\left[M_1M_{B'}\right]-1)\Big)\Lambda A_e v_d\delta_{ij}^{LR} - 2\Big(g_1^2(8\tan\beta\mathrm{sign}\left[\mu_H M_1\right] \\ & -2) + g_{YB}(g_B+2g_{YB})(4\tan\beta\mathrm{sign}\left[\mu_H M_{B'}\right]-1)\Big)m_{l_j}m_E^2\delta_{ij}^{RR}\Big], \end{aligned} \tag{45}$$

$$\begin{aligned} C_2^{\mathrm{R,B-L}} \simeq -\frac{1}{3840\pi^2\Lambda^4 m_{l_j}}\Big[ & \sqrt{2}\Big(20g_1^2\mathrm{sign}\left[M_1\right] + 10(g_B+g_{YB})(g_B+2g_{YB})\mathrm{sign}\left[M_{B'}\right] \\ & -g_1(3g_B+4g_{YB})\mathrm{sign}\left[M_{BB'}\right](4\mathrm{sign}\left[M_1M_{B'}\right]-1)\Big)\Lambda A_e v_d\delta_{ij}^{LR} \\ & +2\Big(g_2^2(8\tan\beta\mathrm{sign}\left[\mu_H M_2\right]+1) + g_1^2(4\tan\beta\mathrm{sign}\left[\mu_H M_1\right]-1) \\ & +g_{YB}(g_B+g_{YB})(4\tan\beta\mathrm{sign}\left[\mu_H M_{B'}\right]-1)\Big)m_{l_j}m_L^2\delta_{ij}^{LL}\Big]. \end{aligned} \tag{46}$$

With the supposition $\mathrm{sign}\,[M_1] = \mathrm{sign}\,[M_2] = \mathrm{sign}\,[\mu_H] = 1$, $\mathrm{sign}\,[M_{B'}] = \mathrm{sign}\,[M_{BB'}] = -1$ and $4\tan\beta >> 1$, the Wilson coefficient for the B-LSSM can be

$$\begin{aligned} C_2^{\mathrm{L,B-L}} \to \frac{1}{3840\pi^2\Lambda^2 m_{l_j}}\Big[&-20\sqrt{2}g_1^2\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} + 16g_1^2\tan\beta m_{l_j}\delta_{ij}^{RR} \\ &+\sqrt{2}\Big(10(g_B+g_{YB})(g_B+2g_{YB}) + 5g_1(3g_B+4g_{YB})\Big)\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} \\ &-8g_{YB}(g_B+2g_{YB})\tan\beta m_{l_j}\delta_{ij}^{RR}\Big], \end{aligned} \tag{47}$$

$$\begin{aligned} C_2^{\mathrm{R,B-L}} \to \frac{1}{3840\pi^2\Lambda^2 m_{l_j}}\Big[&-20\sqrt{2}g_1^2\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} - 2(8g_2^2+4g_1^2)\tan\beta m_{l_j}\delta_{ij}^{LL} \\ &+\sqrt{2}\Big(10(g_B+g_{YB})(g_B+2g_{YB}) + 5g_1(3g_B+4g_{YB})\Big)\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} \\ &+8g_{YB}(g_B+g_{YB})\tan\beta m_{l_j}\delta_{ij}^{LL}\Big]. \end{aligned} \tag{48}$$

In the condition $-0.7 < g_{YB} < -0.05$, $0.1 < g_B < 0.85$ and $|g_{YB}| < g_B < 4/3|g_{YB}|$, then $\sqrt{2}\Big(10(g_B+g_{YB})(g_B+2g_{YB}) + 5g_1(3g_B+4g_{YB})\Big) < 0$, $-8g_{YB}(g_B+2g_{YB}) < 0$, $8g_{YB}(g_B+g_{YB}) < 0$. The order analysis shows

$$\begin{aligned} D_2^{L,B-L} = \Big[&\sqrt{2}\Big(10(g_B+g_{YB})(g_B+2g_{YB}) + 5g_1(3g_B+4g_{YB})\Big)\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} \\ &-8g_{YB}(g_B+2g_{YB})\tan\beta m_{l_j}\delta_{ij}^{RR}\Big] < 0, \end{aligned} \tag{49}$$

$$\begin{aligned} D_2^{R,B-L} = \Big[&\sqrt{2}\Big(10(g_B+g_{YB})(g_B+2g_{YB}) + 5g_1(3g_B+4g_{YB})\Big)\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} \\ &+8g_{YB}(g_B+g_{YB})\tan\beta m_{l_j}\delta_{ij}^{LL}\Big] < 0. \end{aligned} \tag{50}$$

As parameters $g_1 \simeq 0.3$, $g_2 \simeq 0.6$, $g_B \simeq 0.3$, $g_{YB} \simeq -0.25$, $\tan\beta m_\mu \simeq \frac{A_e v_d}{\Lambda} \simeq 1$ and $\delta_{12}^{LL} \simeq \delta_{12}^{RR} \simeq \delta_{12}^{LR} \simeq 10^{-5}$, $D_2^{L,B-L}$ and $D_2^{R,B-L}$ are at the order around $-10^{-5}$, which approximately takes the same values as the MSSM contributions from $[-10\sqrt{2}g_1^2\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} + 8g_1^2\tan\beta m_{l_j}\delta_{ij}^{RR}]$ and $[-10\sqrt{2}g_1^2\frac{A_e v_d}{\Lambda}\delta_{ij}^{LR} - (8g_2^2+4g_1^2)\tan\beta m_{l_j}\delta_{ij}^{LL}]$. Therefore, the B-LSSM contributions beyond MSSM are considerable.

## IV. THE NUMERICAL RESULTS

In this section, we research the numerical results of the branching ratios for CLFV processes $l_j^- \to l_i^-\gamma$. We consider some experimental constraints: 1. The updated experimental

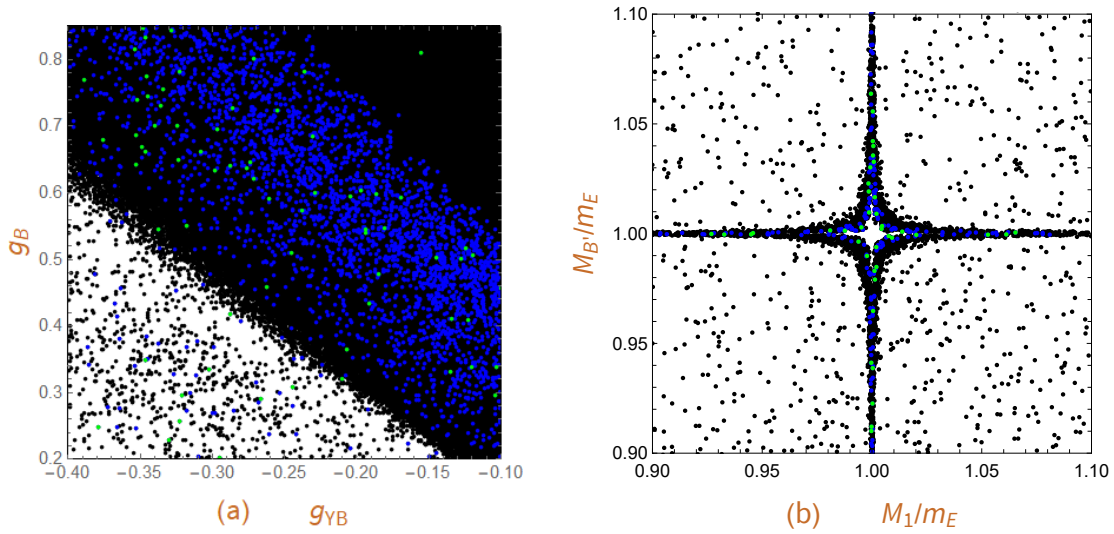



FIG. 3: Under the premise of limit on CLFV process $\mu \to e\gamma$, reasonable parameter space is selected to scatter points. Here, the black, blue and green scatter points correspond to the SM-like Higgs boson mass and muon MDM within $4\sigma, 3\sigma$ and $2\sigma$ regions, respectively.

data indicates that the $Z'$ boson mass satisfies $M'_Z \geq 5.15$ TeV with 95% CL[44]. Refs.[45, 46] give us an upper bound on the ratio between the mass of $Z'$ boson and its gauge coupling at 99% CL: $\frac{M'_Z}{g_B} \geq 6$ TeV, so $g_B$ is restricted in the region of $0 < g_B \leq 0.85$. 2. The large $\tan\beta$ has been excluded by the $\bar{B} \to X_s\gamma$ experiment[47, 48]. Besides, the Yukawa coupling $Y_b$ is defined as $Y_b = \sqrt{2(\tan\beta^2 + 1)}m_b/v$. In general, the value of $Y_b$ is smaller than 1, $m_b \simeq 4.18$ GeV and $v \simeq 246$ GeV, so the parameter $\tan\beta$ should be approximatively smaller than 40. 3. The coupling parameter $g_{YB}$ is taken around $-0.45 \sim -0.05$, which has been discussed in the works[49, 50].

The SM-like Higgs boson mass is $m_h = 125.25 \pm 0.17$ GeV[51–54], which constrains the parameter space strictly. Therefore, we take the suitable parameter space to limit the SM-like Higgs boson mass of the B-LSSM within experimental $4\sigma, 3\sigma$ and $2\sigma$ regions. We also consider the constraints on the NP contribution to the muon MDM in the B-LSSM. The difference between the experimental measurement and SM theoretical prediction of $a_\mu$ is $\Delta a_\mu = a_\mu^{exp} - a_\mu^{SM} = (251 \pm 59) \times 10^{-11}$[51, 55], which possesses $4.2\sigma$ deviation. Therefore, we constrain the NP contribution $\Delta a_\mu^{NP}$ with $4\sigma, 3\sigma$ and $2\sigma$ experimental errors. Then the numerical results of the CLFV process $l_j^- \to l_i^-\gamma$ are discussed detailedly in the follows.

### A. The CLFV process $\mu \to e\gamma$

We all know that the CLFV process $\mu \to e\gamma$ possesses the strict experimental constraint, so we first discuss this process with $\delta_{13}^{AB} = \delta_{23}^{AB} = 0(A, B = L, R), \delta_{12}^{LL} = \delta_{12}^{RR} = \delta_{12}^{LR} = 10^{-5}$. Parameters $\tan\beta, M_1, M_2, \mu_H, M_{BB'}, M_{B'}, A_e, m_L = m_E, g_{YB}$ and $g_B$ are assumed as random variables in the suitable regions. As the SM-like Higgs boson mass and muon MDM are both in $4\sigma, 3\sigma$ and $2\sigma$ regions, and the branching ratio of $\mu \to e\gamma$ satisfies the current experiment constraint, the reasonable parameter space is selected to scatter points, which are shown in FIG.3. Firstly, we discuss the distribution of $g_B$ versus $g_{YB}$ in FIG.3(a). Under the experimental constraint of $4\sigma$ region, the numerical results are mainly concentrated in the upper right half interval of function $g_B = -2g_{YB}$. In the $3\sigma$ interval constraint, the numerical results are mostly located in the region above function $g_B = -2g_{YB}$ and below function $g_B = -2g_{YB} + 0.35$. In the $2\sigma$ constraint, the numerical results are evenly distributed in the lower left half interval of function $g_B = -2g_{YB} + 0.35$. And we find that the region of $g_B$ increases with the decreased $g_{YB}$, even $g_B$ can be anywhere from 0.2 to 0.85 as $g_{YB} \leq -0.25$ with $2\sigma$ limit. So $g_{YB}$ takes $-0.25$ in the following discussion. Besides, in the FIG.3(b), the parameters $M_1$ and $m_E$ are approximately equal to each other as $\frac{M_{B'}}{m_E} \in (0.9, 1.1)$, as well as the parameters $M_{B'}$ and $m_E$ are approximately equal to each other when $\frac{M_1}{m_E} \in (0.9, 1.1)$. As the experimental constraint changes from $4\sigma$ to $2\sigma$, the values of $\frac{M_{B'}}{m_E}$ and $\frac{M_1}{m_E}$ are both closer to 1. Therefore, in the numerical analyses below, we take $M_1 \simeq 0.947$ TeV, $m_E \simeq 0.948$ TeV and $M_{B'} \simeq 0.93$ TeV, which are approximately equal to each other. Besides, we appropriately fix $M_2 = 0.6$ TeV, $\mu_H = 0.5$ TeV and $A_e = 0.4$ TeV to simplify discussion.

The CLFV process $\mu \to e\gamma$ is flavor dependent, which can be influenced by the slepton mixing parameters $\delta_{12}^{AB}(A, B = L, R)$. In order to study the characters of $\delta_{12}^{AB}(A, B = L, R)$ to $Br(\mu \to e\gamma)$, we assume $g_{YB} = -0.25$, $g_B = 0.8$, $M_{BB'} = 0.8$ TeV and $\tan\beta = 25$, then the corresponding analyses are shown in FIG.4. It is obvious that the CLFV rates increase with the enlarged $\delta_{12}^{AB}(A, B = L, R)$ and possess the same changes within $4\sigma$, $3\sigma$ and $2\sigma$ constraints. Besides, the present experimental upper bound of $Br(\mu \to e\gamma)$ constrains $\delta_{12}^{LL} \leq 1.4 \times 10^{-4}$, $\delta_{12}^{RR} \leq 5.1 \times 10^{-4}$ and $\delta_{12}^{LR} \leq 1.1 \times 10^{-5}$.

Then, we study the CLFV rates for $\mu \to e\gamma$ versus $g_B, g_{YB}, M_{BB'}$ and $\tan\beta$ respectively

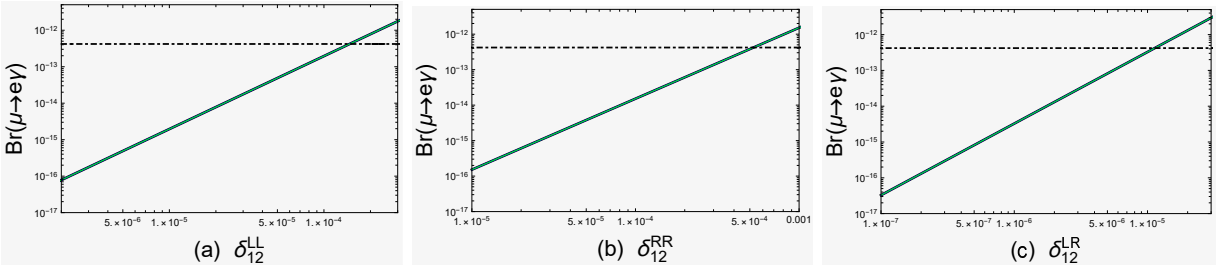



FIG. 4: The CLFV rates for $\mu \to e\gamma$ versus $\delta_{12}^{AB}(A, B = L, R)$, where the dotdashed line denotes the present experimental limit. Here, the black, blue and green lines correspond to $4\sigma$, $3\sigma$ and $2\sigma$ regions, respectively.

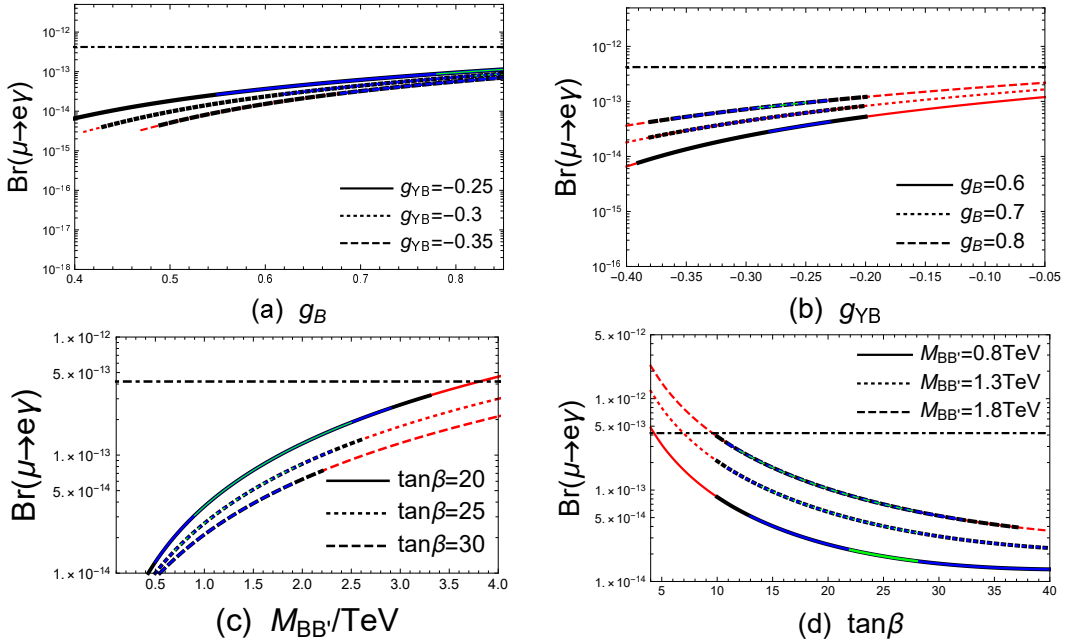



FIG. 5: The $Br(\mu \to e\gamma)$ versus $g_B, g_{YB}, M_{BB'}$ and $\tan\beta$ respectively, where the dotdashed line denotes the present experimental limit. Here, the black, blue and green lines correspond to $4\sigma, 3\sigma$ and $2\sigma$ regions, respectively.

in FIG.5. In general, the numerical results of $Br(\mu \to e\gamma)$ enlarge with the increase of $g_B, g_{YB}, M_{BB'}$, and decrease with the increase of $\tan\beta$. The red line has been excluded due to exceeding $4\sigma$ experimental constraint of SM-like Higgs mass or muon MDM. And the parameter spaces of $g_B, g_{YB}, M_{BB'}$ and $\tan\beta$ all narrow with the increase of muon MDM and SM-like Higgs mass constraints. Obviously, FIG.5(a) and (b) indicate that the larger $g_B$

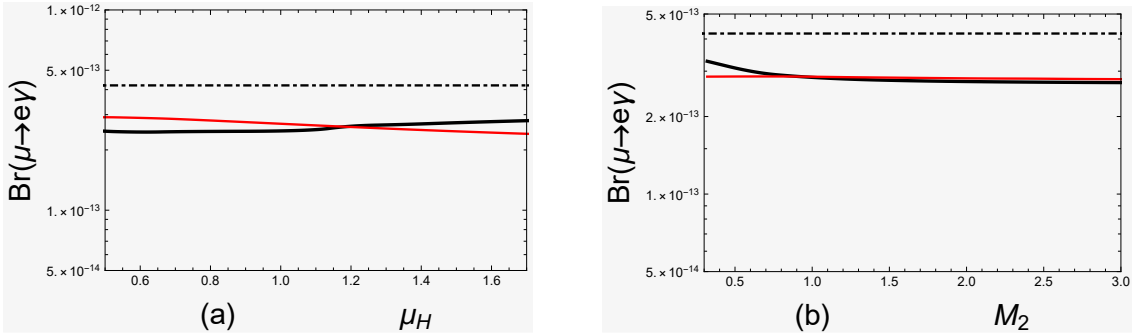



FIG. 6: The CLFV rates for $\mu \to e\gamma$ versus $\mu_H(M_2)$, where the black and red lines correspond to the interaction eigenstate and mass eigenstate respectively.

or $g_{YB}$ is, the more appropriate CLFV rates for $\mu \to e\gamma$ we can obtain. Under the constraint of $2\sigma$, $g_B$ takes the value in the space of $0.78 \sim 0.85$ when $g_{YB} = -0.25$, as well as $g_{YB}$ takes the value in the space of $-0.29 \sim -0.25$ when $g_B = 0.8$. In the FIG.5(c), when the value of $\tan\beta$ is small, the parameter $M_{BB'}$ can obtain a large region, and the upper and lower limits of $M_{BB'}$ also increase. Furthermore, in the $3\sigma$ or $4\sigma$ limit, $\tan\beta$ can obtain a larger region as $M_{BB'} = 1.3$ TeV than that of $M_{BB'} = 0.8$ TeV or $M_{BB'} = 1.8$ TeV. However, in the $2\sigma$ region, the space of $\tan\beta$ decreases with the reduced $M_{BB'}$, and even $\tan\beta$ is confined to a much small interval of $22 \sim 28$ as $M_{BB'} = 0.8$ TeV.

In Refs.[56, 57], the authors show that the MIA results can be obtained if one expands the starting expressions in the mass basis properly. The concrete contents include: 1. The loop particles are much heavier than the external states; 2. The convergent of any one-loop amplitude requires that the moduli of every eigenvalue of the dimensionless mass insertion matrix has to be smaller than one. And our discussion satisfies the basic approximation assumption studied in Refs.[56, 57]. Besides, we further study the numerical results of $Br(\mu \to e\gamma)$ versus $\mu_H(M_2)$ in the interaction eigenstate and mass eigenstate respectively in FIG.6. When parameters are valued in a reasonable space, such as $\delta_{12}^{LL} = 5 \times 10^{-5}$, $\delta_{12}^{RR} = 5 \times 10^{-4}$, $\delta_{12}^{LR} = 9.5 \times 10^{-6}$, $g_{YB} = -0.25$, $g_B = 0.8$ and $\tan\beta = 15$, we find that the $Br(\mu \to e\gamma)$ with the running of $\mu_H(M_2)$ in the interaction eigenstate have the slight deviation comparing with that in mass eigenstate. Even when $\mu_H$ or $M_2$ takes a certain value, the branch ratios in interaction eigenstate and mass eigenstate are exactly the same.

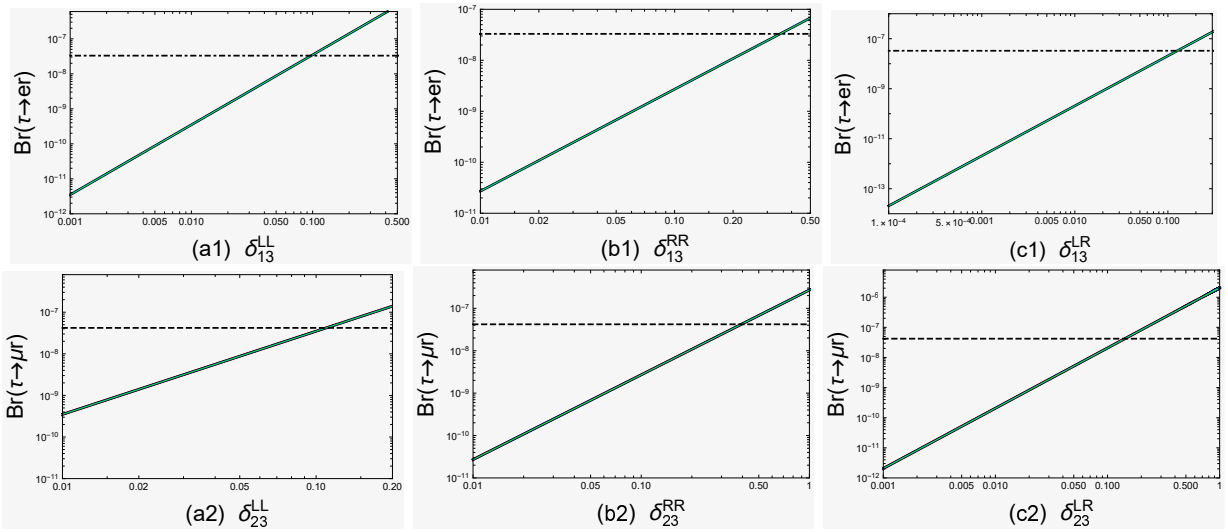



FIG. 7: The CLFV rates for $\tau \to e\gamma(\tau \to \mu\gamma)$ versus $\delta_{13}^{AB}(\delta_{23}^{AB}(A, B = L, R))$, where the dotdashed lines denote the present experimental limits. Here, the black, blue and green lines correspond to $4\sigma, 3\sigma$ and $2\sigma$ regions, respectively.

In this work, the numerical discussion of MIA scheme in the interaction eigenstates is based on satisfying the numerical results of mass eigenstates. Therefore, the compatibility between SARAH and the MIA scheme is guaranteed in this work.

### B. The CLFV processes $\tau \to e\gamma$ and $\tau \to \mu\gamma$

Due that the experimental upper bounds for $Br(\tau \to e\gamma)$ and $Br(\tau \to \mu\gamma)$ do not possess obvious difference, so we research both these processes in this section. In FIG.7, we picture $Br(\tau \to e\gamma)$ $(Br(\tau \to \mu\gamma))$ versus the slepton flavor mixing parameters $\delta_{13}^{AB}(A, B = L, R)$ $(\delta_{23}^{AB}(A, B = L, R))$. Obviously, the CLFV rates of $\tau \to e\gamma$ possess the same trends for $4\sigma$, $3\sigma$ and $2\sigma$ limits, as well as $Br(\tau \to \mu\gamma)$. We can clearly see that all slepton flavor mixing parameters take positive influences on the CLFV rates, even the $Br(\tau \to e\gamma)$ and $Br(\tau \to \mu\gamma)$ are respectively proportional to $\delta_{13}^{AB}$ and $\delta_{23}^{AB}(A, B = L, R)$. Besides, we find that $\delta_{13}^{LL} \leq 0.09$, $\delta_{13}^{RR} \leq 0.35$ and $\delta_{13}^{LR} \leq 0.13$ with the experimental constraint of $Br(\tau \to e\gamma)$. Through the present limit of $Br(\tau \to \mu\gamma)$, $\delta_{23}^{LL}$, $\delta_{23}^{RR}$ and $\delta_{23}^{LR}$ are restricted below 0.11, 0.39 and 0.14 respectively.

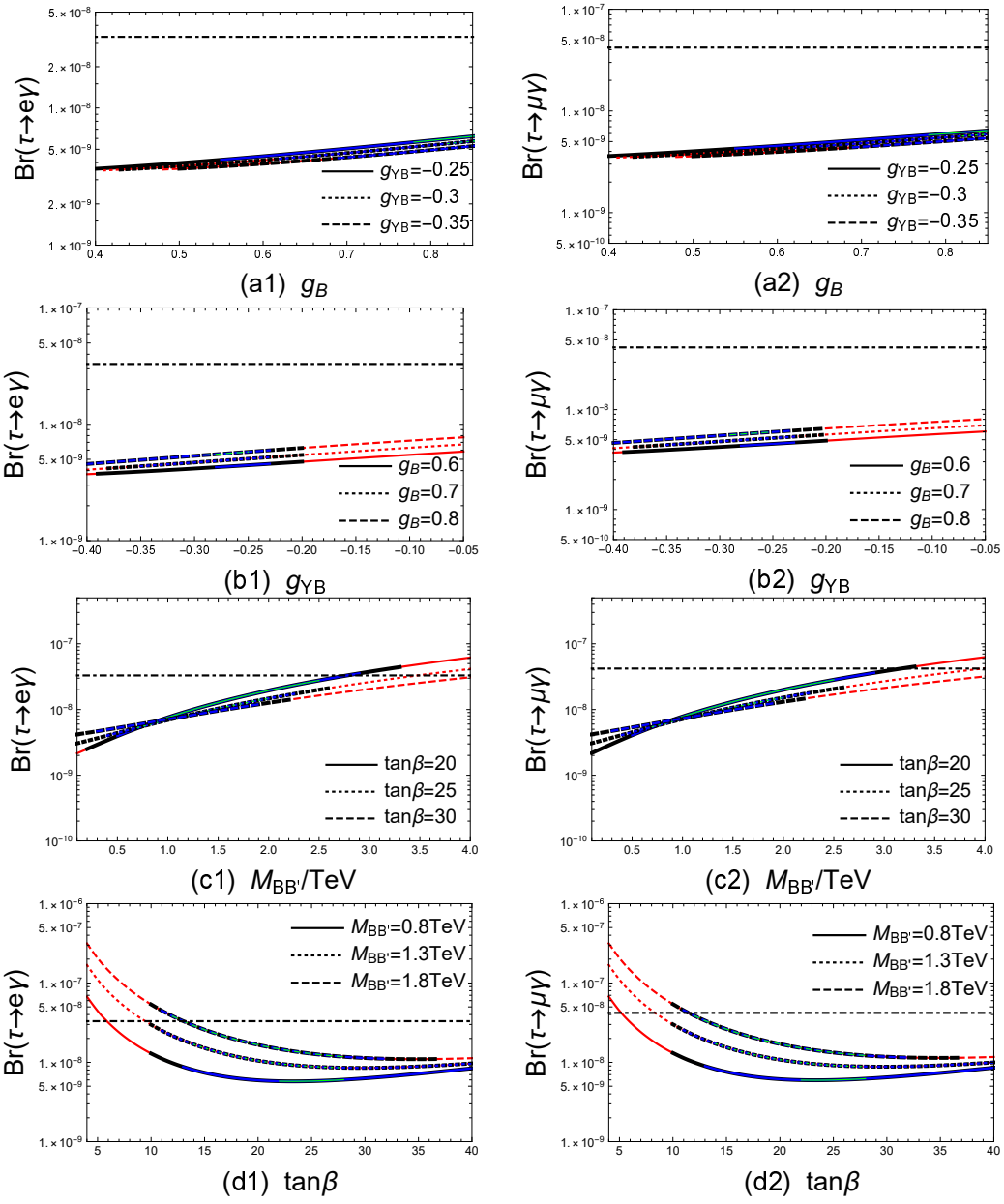



FIG. 8: The CLFV rates change with $g_B, g_{YB}, M_{BB'}$ and $\tan\beta$, where the black, blue and green lines correspond to $4\sigma, 3\sigma$ and $2\sigma$ regions, and the dotdashed line denotes the experimental limit.

In order to discuss the effects from other basic parameters on the numerical results, we set $\delta_{13}^{LL} = \delta_{13}^{RR} = \delta_{13}^{LR} = 0.03$ and $\delta_{23}^{AB}(A, B = L, R) = 0$ for process $\tau \to e\gamma$, as well as $\delta_{23}^{LL} = \delta_{23}^{RR} = \delta_{23}^{LR} = 0.03$ and $\delta_{13}^{AB}(A, B = L, R) = 0$ for process $\tau \to \mu\gamma$. Then, we study the influences on $Br(\tau \to e\gamma)$ and $Br(\tau \to \mu\gamma)$ from parameters $g_B, g_{YB}, M_{BB'}$ and $\tan\beta$.

Overall, the variations of $Br(\tau \to e\gamma)$ with parameters $g_B, g_{YB}, M_{BB'}$ and $\tan\beta$ are all slightly steeper than that of $Br(\tau \to \mu\gamma)$, and the value spaces of $g_B, g_{YB}, M_{BB'}$ and $\tan\beta$ become small as the constraints of SM-like Higgs mass and muon MDM change from $4\sigma$ to $2\sigma$. FIG.8(a1) and (a2) both indicate that the ranges of $g_B$ increase as $g_{YB}$ expands, and when $g_{YB} = -0.25$, $g_B$ can fetch all values from 0.78 to 0.85 under the $2\sigma$ constraint. $g_{YB}$ in FIG.8(b1) and (b2) can both obtain smallest regions when $g_B = 0.7$ with the limits of $4\sigma$. However, within $3\sigma$ limit, the larger $g_B$ is, the larger space for reasonable values $g_{YB}$ is. Under $2\sigma$ constraint, $g_{YB}$ runs in a $-0.29 \sim -0.25$ cell only when $g_B = 0.8$.

In FIG.8(c1) and (c2), we can see that the large regions of $M_{BB'}$ have been eliminated by the experimental upper limits of $Br(\tau \to e\gamma)$ and $Br(\tau \to \mu\gamma)$ both within $4\sigma$ or $3\sigma$ limit, but $M_{BB'}$ can still obtain large region as $\tan\beta = 20$. Besides, the experimental limit of $Br(\tau \to e\gamma)$ is stronger than that of $Br(\tau \to \mu\gamma)$. Under the constraint of $2\sigma$, the value range of $M_{BB'}$ changes from 0.7 TeV to 2 TeV as $\tan\beta = 25$, which is more narrow than that of $\tan\beta = 20$ (here 0.9 TeV$\leq M_{BB'} \leq$2.5 TeV). In addition, considering $4\sigma$ or $3\sigma$ constraint on the FIG.8(d1) and (d2), we find that the larger $M_{BB'}$ is, the more obvious experimental constraints of $Br(\tau \to e\gamma)$ and $Br(\tau \to \mu\gamma)$ on $\tan\beta$ are, and the smaller value space of $\tan\beta$ is. That is to say the suitable parameter space of $\tan\beta$ narrows obviously as $M_{BB'}$ enlarges. When $M_{BB'} = 1.8$ TeV, $\tan\beta$ can be any value between 14 and 28 within $2\sigma$ region, which indicates parameter $\tan\beta$ is more constrained by the experimental limit of $Br(\tau \to e\gamma)$ than $Br(\tau \to \mu\gamma)$.

## V. DISCUSSION AND CONCLUSION

In this work, under the premise that the compatibility between SARAH and the MIA scheme is guaranteed, we focus on CLFV processes $l_j^- \to l_i^-\gamma$ in the B-LSSM using the MIA method. Assuming that all superparticle masses are almost degenerate, we find that the B-LSSM contributions for both muon MDM and CLFV processes beyond MSSM are considerable at the analytical level. In the numerical discussion, we constrain the SM-like Higgs mass and muon MDM both within $4\sigma$, $3\sigma$ and $2\sigma$ regions, which possess strict limits on the CLFV decays. Firstly, the distribution of $g_B$ versus $g_{YB}$ presents different

sensitive parameter regions under the different bounds within $2\sigma$, $3\sigma$ and $4\sigma$. Besides, the parameters $M_1$ and $m_E$ are approximately equal to each other as $\frac{M_{B'}}{m_E} \in (0.9, 1.1)$, as well as the parameters $M_{B'}$ and $m_E$ are approximately equal to each other when $\frac{M_1}{m_E} \in (0.9, 1.1)$. As the experimental constraint changes from $4\sigma$ to $2\sigma$, the values of $\frac{M_{B'}}{m_E}$ and $\frac{M_1}{m_E}$ are both closer to 1. The branching ratios of $l_j^- \to l_i^-\gamma$ depend on the slepton flavor mixing parameters $\delta_{12}^{AB}, \delta_{13}^{AB}$ and $\delta_{23}^{AB}(A, B = L, R)$ obviously, and $Br(l_j^- \to l_i^-\gamma)$ increase with the enlarged $\delta_{12}^{AB}, \delta_{13}^{AB}$ and $\delta_{23}^{AB}$. Considering the latest experimental limits of $Br(l_j^- \to l_i^-\gamma)$, the slepton flavor mixing parameters are restricted as $\delta_{12}^{LL} \leq 1.4 \times 10^{-4}$, $\delta_{12}^{RR} \leq 5.1 \times 10^{-4}$, $\delta_{12}^{LR} \leq 1.1 \times 10^{-5}$, $\delta_{13}^{LL} \leq 0.09$, $\delta_{13}^{RR} \leq 0.35$, $\delta_{13}^{LR} \leq 0.13$, $\delta_{23}^{LL} \leq 0.11$, $\delta_{23}^{RR} \leq 0.39$ and $\delta_{23}^{LR} \leq 0.14$.

The parameter ranges of $g_B$, $g_{YB}$, $M_{BB'}$ and $\tan\beta$ all narrow obviously as the constraints of SM-like Higgs mass and muon MDM change from $4\sigma$ to $2\sigma$. Compared with the MSSM, $g_B$, $g_{YB}$ and $M_{BB'}$ are the new parameters in the B-LSSM, which have an uplift effects on the numerical results. At $3\sigma$ or $2\sigma$ limit, $g_B$ has a large region with the increased $g_{YB}$, similarly, the reasonable parameter space of $g_{YB}$ also widens as the value of $g_B$ enlarges. When $g_{YB} = -0.25(g_B = 0.8)$, $g_B(g_{YB})$ is constrained in the range of $0.78 \sim 0.85(-0.29 \sim -0.25)$, which correspond to the maximum spaces of these parameters under the $2\sigma$ constraint. Under the $2\sigma$ constraint, the parameter $M_{BB'}$ can obtain a large reasonable space $0.9 \sim 2.5$ TeV when $\tan\beta = 20$. The large parameter space of $\tan\beta$ is excluded by the latest experiment limit of $Br(\tau \to e\gamma)$, which possesses the strongest constraint. Therefore, under the $2\sigma$ and $Br(\tau \to e\gamma)$ constraints, $\tan\beta$ can run in a much large region of $14 \sim 28$ when $M_{BB'} = 1.8$ TeV.

**Acknowledgments**

This work is supported by the Major Project of National Natural Science Foundation of China (NNSFC) (No. 12235008), the National Natural Science Foundation of China (NNSFC) (No. 12075074, No. 12075073), the Natural Science Foundation of Hebei province(No. A202201022, No. A2020201002), the Natural Science Foundation of Hebei Education Department(No. QN2022173).

### Appendix A: One-loop functions

The one-loop functions $\mathcal{I}(x,y,z)$, $\mathcal{J}(x,y,z)$, $f(x,y,z,t)$, $g(x,y,z,t)$ and $k(x,y,z,t)$ can be written as

$$
\begin{aligned}
\mathcal{I}(x,y,z) &= \frac{1}{16\pi^2}\Big[\frac{1}{(x-z)(z-y)} + \frac{(z^2-xy)\log z}{(x-z)^2(y-z)^2} - \frac{x\log x}{(x-y)(x-z)^2} \\
&\quad + \frac{y\log y}{(x-y)(y-z)^2}\Big], \\
\mathcal{J}(x,y,z) &= \frac{1}{16\pi^2}\Big[\frac{x(x^2+xz-2yz)\log x}{(x-y)^2(x-z)^3} - \frac{y^2\log y}{(x-y)^2(y-z)^2} + \frac{z[x(z-2y)+z^2]\log z}{(z-x)^3(y-z)^2} \\
&\quad - \frac{x(y-2z)+yz}{(x-y)(x-z)^2(y-z)}\Big], \\
f(x,y,z,t) &= \frac{1}{16\pi^2}\Big[\frac{t\,[t^3-3txy+xy(x+y)]\log t}{(t-x)^3(t-y)^3(t-z)} - \frac{x\,[x^3-3txz+tz(t+z)]\log x}{(t-x)^3(x-y)(x-z)^3} \\
&\quad + \frac{y\,[y^3-3tyz+tz(t+z)]\log y}{(t-y)^3(x-y)(y-z)^3} - \frac{z\,[z^3-3xyz+xy(x+y)]\log z}{(t-z)(z-x)^3(z-y)^3} + \frac{1}{2(x-y)} \\
&\quad \times\left(\frac{t}{(t-x)^2(z-x)} - \frac{2y}{(t-y)(y-z)^2} + \frac{x(2t-3x+z)}{(t-x)^2(x-z)^2} + \frac{t+y}{(t-y)^2(y-z)}\right)\Big], \\
g(x,y,z,t) &= \frac{1}{16\pi^2}\Big\{-\frac{t\,[t^3(x+y)-3t^2xy+x^2y^2]\log t}{(t-x)^3(t-y)^3(t-z)} \\
&\quad + \frac{z\,[x^2y^2+xz^2(z-3y)+yz^3]\log z}{(t-z)(z-x)^3(z-y)^3} + \frac{x^2\,[x^3-3txz+tz(t+z)]\log x}{(t-x)^3(x-y)(x-z)^3} \\
&\quad - \frac{y^2\,[y^3-3tyz+tz(t+z)]\log y}{(t-y)^3(x-y)(y-z)^3} - \frac{x^2(2t-3x+z)}{2(t-x)^2(x-y)(x-z)^2} \\
&\quad + \frac{tx}{2(t-x)^2(x-y)(x-z)} - \frac{y[t(y+z)+y(z-3y)]}{2(t-y)^2(y-x)(y-z)^2}\Big\}, \\
k(x,y,z,t) &= \frac{1}{16\pi^2}\Big\{-\frac{[z\,(x^2-ty)+x^2(t-2x+y)]\log x}{(t-x)^2(x-y)^2(x-z)^2} - \frac{t\log t}{(t-x)^2(t-y)(t-z)} \\
&\quad + \frac{y\log y}{(t-y)(x-y)^2(y-z)} + \frac{1}{(t-x)(x-y)(x-z)} + \frac{z\log z}{(t-z)(z-y)(x-z)^2}\Big\}. \qquad \text{(A1)}
\end{aligned}
$$

### Appendix B: The chirality flips

The chirality flips occurring in the internal gaugino and the external lepton lines were described in detail in the Refs.[58–60]. Since the gaugino mass is much bigger than the lepton mass, these diagrams whose chirality flips occur in the internal gaugino lines may

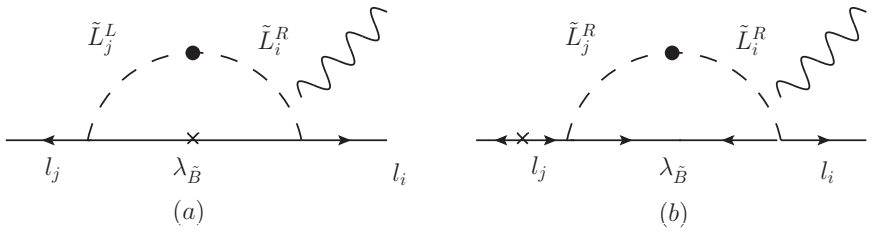



FIG. 9: The feynman diagrams of $l_j^- \to l_i^- \gamma$ by MIA in the B-LSSM, the chirality flips occur in the internal gaugino and external lepton lines, respectively.

yield dominant contributions.

Then, we take two Feynman pictures shown in FIG.9 as the examples, and discuss the chirality flips occurring in the internal gaugino and the external lepton lines as follows. Firstly, we discuss the concrete amplitude of FIG.9(a):

$$i\mathcal{M}_1=\bar{u}_i(p+q)\mu^{4-D}\int\frac{d^Dk}{(2\pi)^D}\Big[\frac{eA_LP_L(\not{k}+M_1)B_LP_L(2p+q-2k)^\mu C}{[k^2-M_1^2][(k-p)^2-m^2_{\tilde{L}_j^L}][(k-(p+q))^2-m^2_{\tilde{L}_j^L}][(k-(p+q))^2-m^2_{\tilde{L}_i^R}]}$$
$$+\frac{eA_LP_L(\not{k}+M_1)B_LP_L(2p+q-2k)^\mu C}{[k^2-M_1^2][(k-p)^2-m^2_{\tilde{L}_j^L}][(k-p)^2-m^2_{\tilde{L}_i^R}][(k-(p+q))^2-m^2_{\tilde{L}_i^R}]}\Big]u_j(p)\epsilon_\mu(q). \quad \text{(B1)}$$

We assume that the internal masses of particles such as gaugino and left(right)-handed slepton are much bigger than the external lepton mass, the function $\frac{1}{(k-p)^2-m^2}\simeq\frac{1}{k^2-m^2}+\frac{2k\cdot p-p^2}{(k^2-m^2)^2}+\frac{4(k\cdot p)^2}{(k^2-m^2)3}$. Then the amplitude can be approximately reduced as

$$i\mathcal{M}_1\simeq\bar{u}_i(p+q)eM_1CA_LB_LP_L\mu^{4-D}\int\frac{d^Dk}{(2\pi)^D}\frac{(2p+q)^\mu}{[k^2-M_1^2][k^2-m^2_{\tilde{L}_j^L}][k^2-m^2_{\tilde{L}_i^R}]}\Big[\frac{1}{k^2-m^2_{\tilde{L}_j^L}}$$
$$+\frac{1}{k^2-m^2_{\tilde{L}_i^R}}-\frac{k^2}{(k^2-m^2_{\tilde{L}_j^L})^2}-\frac{k^2}{(k^2-m^2_{\tilde{L}_i^R})^2}-\frac{k^2}{(k^2-m^2_{\tilde{L}_j^L})(k^2-m^2_{\tilde{L}_i^R})}\Big]u_j(p)\epsilon_\mu(q)$$
$$=-\bar{u}_i(p+q)eM_1CA_LB_LP_Li\sigma^{\mu\nu}q_\nu\frac{i}{2\Lambda^4}\Big[\mathcal{J}(x_1,x_{\tilde{L}_j^L},x_{\tilde{L}_i^R})+\mathcal{J}(x_1,x_{\tilde{L}_i^R},x_{\tilde{L}_j^L})\Big]u_j(p)\epsilon_\mu(q). \quad \text{(B2)}$$

With $A_L=-\sqrt{2}g_1$, $B_L=\frac{g_1}{\sqrt{2}}$ and $C=-\triangle_{ij}^{LR}=-A_ev_d\delta_{ij}^{LR}$, the Wilson coefficient of FIG.9(a) can be written as:

$$C_2\left(\tilde{L}_j^L,\tilde{L}_i^R,\lambda_{\tilde{B}}\right)=-\frac{g_1^2}{2\Lambda^4m_{l_j}}M_1A_ev_d\delta_{ij}^{LR}P_L\Big[\mathcal{J}(x_1,x_{\tilde{L}_j^L},x_{\tilde{L}_i^R})+\mathcal{J}(x_1,x_{\tilde{L}_i^R},x_{\tilde{L}_j^L})\Big]. \quad \text{(B3)}$$

Here, the chirality flip occurs in the internal gaugino line.

Secondly, we take FIG.9(b) as an example, and discuss the chirality flip occurs in the external lepton line. The concrete amplitude is written as:

$$
\begin{aligned}
i\mathcal{M}_2=&\bar{u}_i(p+q)\mu^{4-D}\int\frac{d^Dk}{(2\pi)^D}\Big[\frac{eA_LP_L(\not{k}+M_1)B_RP_R(2p+q-2k)^\mu C'}{[k^2-M_1^2][(k-p)^2-m^2_{\tilde{L}_j^R}][(k-(p+q))^2-m^2_{\tilde{L}_j^R}][(k-(p+q))^2-m^2_{\tilde{L}_i^R}]}\\
&+\frac{eA_LP_L(\not{k}+M_1)B_RP_R(2p+q-2k)^\mu C'}{[k^2-M_1^2][(k-p)^2-m^2_{\tilde{L}_j^R}][(k-p)^2-m^2_{\tilde{L}_i^R}][(k-(p+q))^2-m^2_{\tilde{L}_i^R}]}\Big]u_j(p)\epsilon_\mu(q)\\
&\simeq \bar{u}_i(p+q)eC'A_LB_RP_L\mu^{4-D}\int\frac{d^Dk}{(2\pi)^D}\frac{\not{k}(2p+q)^\mu 2k\cdot(2p+q)}{[k^2-M_1^2][k^2-m^2_{\tilde{L}_j^R}][k^2-m^2_{\tilde{L}_i^R}]}\Big[\frac{1}{(k^2-m^2_{\tilde{L}_j^R})^2}\\
&+\frac{1}{(k^2-m^2_{\tilde{L}_j^R})(k^2-m^2_{\tilde{L}_i^R})}+\frac{1}{(k^2-m^2_{\tilde{L}_i^R})^2}\Big]u_j(p)\epsilon_\mu(q)\\
&\simeq \bar{u}_i(p+q)eC'A_LB_RP_L\mu^{4-D}\int\frac{d^Dk}{(2\pi)^D}\frac{1}{2}\frac{(2p+q)^\mu(2\not{p}+\not{q})}{[k^2-M_1^2][k^2-m^2_{\tilde{L}_j^R}][k^2-m^2_{\tilde{L}_i^R}]}\Big[\frac{k^2}{(k^2-m^2_{\tilde{L}_j^R})^2}\\
&+\frac{k^2}{(k^2-m^2_{\tilde{L}_j^R})(k^2-m^2_{\tilde{L}_i^R})}+\frac{k^2}{(k^2-m^2_{\tilde{L}_i^R})^2}\Big]u_j(p)\epsilon_\mu(q)\\
&=-\bar{u}_i(p+q)eC'A_LB_R(\frac{m_{l_j}}{2}P_L+\frac{m_{l_i}}{2}P_R)i\sigma^{\mu\nu}q_\nu\frac{i}{2\Lambda^4}\Big[2\mathcal{I}(x_{\tilde{L}_j^R},x_1,x_{\tilde{L}_i^R})\\
&+2\mathcal{I}(x_{\tilde{L}_i^R},x_1,x_{\tilde{L}_j^R})-\mathcal{J}(x_{\tilde{L}_j^R},x_{\tilde{L}_i^R},x_1)-\mathcal{J}(x_{\tilde{L}_i^R},x_{\tilde{L}_j^R},x_1)\Big]u_j(p)\epsilon_\mu(q).
\end{aligned}
\tag{B4}
$$

With $A_L=B_R=-\sqrt{2}g_1$, $C'=-\triangle_{ij}^{RR}=-m_E^2\delta_{ij}^{RR}$ and $m_{l_j}\gg m_{l_i}$, the Wilson coefficient of FIG.9(b) can be written as:

$$
\begin{aligned}
C_2'\left(\tilde{L}_j^R,\tilde{L}_i^R,\lambda_{\tilde{B}}\right)\simeq&\frac{g_1^2}{2\Lambda^4}m_E^2\delta_{ij}^{RR}P_L\Big[2\mathcal{I}(x_{\tilde{L}_j^R},x_1,x_{\tilde{L}_i^R})+2\mathcal{I}(x_{\tilde{L}_i^R},x_1,x_{\tilde{L}_j^R})\\
&-\mathcal{J}(x_{\tilde{L}_j^R},x_{\tilde{L}_i^R},x_1)-\mathcal{J}(x_{\tilde{L}_i^R},x_{\tilde{L}_j^R},x_1)\Big].
\end{aligned}
\tag{B5}
$$

If $|M_1|=m_L=m_E=A_e=\Lambda$, then $C_2\left(\tilde{L}_j^L,\tilde{L}_i^R,\lambda_{\tilde{B}}\right)\to-\frac{g_1^2}{192\pi^2\Lambda^4m_{l_j}}M_1A_ev_d\delta_{ij}^{LR}P_L$, and $C_2'\left(\tilde{L}_j^R,\tilde{L}_i^R,\lambda_{\tilde{B}}\right)\to\frac{g_1^2}{64\pi^2\Lambda^4m_{l_j}}m_{l_j}m_E^2\delta_{ij}^{RR}P_L$. If $\delta_{ij}^{RR}$ and $\delta_{ij}^{LR}$ in the same order, $M_1A_ev_d\gg m_{l_j}m_E^2$, then the absolute value of $C_2\left(\tilde{L}_j^L,\tilde{L}_i^R,\lambda_{\tilde{B}}\right)$ is much bigger than $C_2'\left(\tilde{L}_j^R,\tilde{L}_i^R,\lambda_{\tilde{B}}\right)$. Therefore, we omit the chirality flip of incident lepton in our calculation.

**Appendix C: The comparisons from FIG.2(a1) and (a2)**

We will discuss the one-loop contributions from $\tilde{H}^- - \tilde{W}^- - \tilde{\nu}_L^R$ in detail, which derive from FIG.2(a1). The concrete amplitude is

$$i\mathcal{M}_{a1}=\bar{u}_i(p+q)\int\frac{d^4k}{(2\pi)^D}\Big[\frac{eA_LP_L(\not{k}+\not{p}+\not{q}+\mu_H)(B_LP_L+B_RP_R)(\not{k}+\not{p}+\not{q}+M_2)\gamma^\mu(\not{k}+\not{p}+M_2)C_LP_LD}{[(k+p+q)^2-\mu_H^2][(k+p+q)^2-M_2^2][(k+p)^2-M_2^2][k^2-m^2_{\tilde{\nu}^R_{L_j}}][k^2-m^2_{\tilde{\nu}^R_{L_i}}]}$$
$$+\frac{eA_LP_L(\not{k}+\not{p}+\not{q}+\mu_H)\gamma^\mu(\not{k}+\not{p}+\mu_H)(B_LP_L+B_RP_R)(\not{k}+\not{p}+M_2)C_LP_LD}{[(k+p+q)^2-\mu_H^2][(k+p)^2-\mu_H^2][(k+p)^2-M_2^2][k^2-m^2_{\tilde{\nu}^R_{L_j}}][k^2-m^2_{\tilde{\nu}^R_{L_i}}]}\Big]u_j(p)\epsilon_\mu(q). \quad \text{(C1)}$$

We assume that the internal masses of particles are much bigger than the external lepton mass, the function $\frac{1}{(k+p)^2-m^2}\simeq\frac{1}{k^2-m^2}-\frac{2k\cdot p+p^2}{(k^2-m^2)^2}+\frac{4(k\cdot p)^2}{(k^2-m^2)3}$. With $A_L=\frac{Y_{l_j}}{\sqrt{2}}$, $B_L=-\frac{g_2}{\sqrt{2}}v_d\tan\beta$, $B_R=-\frac{g_2}{\sqrt{2}}v_d$ (here $B_L\gg B_R$), $C_L=\frac{g_2}{\sqrt{2}}$ and $D=-\triangle^{LL}_{ij}$, then the amplitude can be approximately reduced as

$$\begin{aligned}
i\mathcal{M}_{a1}&\simeq-\bar{u}_i(p+q)\int\frac{d^4k}{(2\pi)^D}\frac{e\mu_HM_2A_LB_LC_LDP_L(2p+q)^\mu}{[k^2-\mu_H^2][k^2-M_2^2][k^2-m^2_{\tilde{\nu}^R_{L_j}}][k^2-m^2_{\tilde{\nu}^R_{L_i}}]}\\
&\times\quad\Big[\frac{k^2}{(k^2-\mu_H^2)^2}+\frac{k^2}{(k^2-M_2^2)^2}+\frac{k^2}{(k^2-\mu_H^2)(k^2-M_2^2)}\Big]u_j(p)\epsilon_\mu(q)\\
&=i\bar{u}_i(p+q)em_{l_j}i\sigma^{\mu\nu}q_\nu\frac{g_2^2}{2\Lambda^4}\frac{m_{l_i}}{m_{l_j}}\sqrt{x_2x_{\mu_H}}\Delta^{LL}_{ij}\tan\beta\Big[k(x_{\mu_H},x_2,x_{\tilde{\nu}^R_{L_j}},x_{\tilde{\nu}^R_{L_i}})\\
&+k(x_2,x_{\mu_H},x_{\tilde{\nu}^R_{L_j}},x_{\tilde{\nu}^R_{L_i}})-f(x_2,x_{\mu_H},x_{\tilde{\nu}^R_{L_j}},x_{\tilde{\nu}^R_{L_i}})\Big]P_Lu_j(p)\epsilon_\mu(q). \qquad \text{(C2)}
\end{aligned}$$

As the same method with FIG.2(a1), the concrete amplitude of FIG.2(a2) can be written as

$$i\mathcal{M}_{a2}=\bar{u}_i(p+q)\int\frac{d^4k}{(2\pi)^D}\Big[\frac{eA'_RP_R(\not{k}+\not{p}+\not{q}+M_2)(B'_LP_L+B'_RP_R)(\not{k}+\not{p}+\not{q}+\mu_H)\gamma^\mu(\not{k}+\not{p}+\mu_H)C'_RP_RD}{[(k+p+q)^2-M_2^2][(k+p+q)^2-\mu_H^2][(k+p)^2-\mu_H^2][k^2-m^2_{\tilde{\nu}^R_{L_j}}][k^2-m^2_{\tilde{\nu}^R_{L_i}}]}$$
$$+\frac{eA'_RP_R(\not{k}+\not{p}+\not{q}+M_2)\gamma^\mu(\not{k}+\not{p}+M_2)(B'_LP_L+B'_RP_R)(\not{k}+\not{p}+\mu_H)C'_RP_RD}{[(k+p+q)^2-M_2^2][(k+p)^2-M_2^2][(k+p)^2-\mu_H^2][k^2-m^2_{\tilde{\nu}^R_{L_j}}][k^2-m^2_{\tilde{\nu}^R_{L_i}}]}\Big]u_j(p)\epsilon_\mu(q). \quad \text{(C3)}$$

With $A'_R=\frac{g_2}{\sqrt{2}}$, $B'_L=-\frac{g_2}{\sqrt{2}}v_d$, $B'_R=-\frac{g_2}{\sqrt{2}}v_d\tan\beta$ (here $B'_L\ll B'_R$), $C'_R=\frac{Y_{l_j}}{\sqrt{2}}$ and $D=-\triangle^{LL}_{ij}$, then the amplitude can be approximately reduced as

$$i\mathcal{M}_{a2}\simeq-\bar{u}_i(p+q)\int\frac{d^4k}{(2\pi)^D}\frac{e\mu_HM_2A'_RB'_RC'_RDP_R(2p+q)^\mu}{[k^2-\mu_H^2][k^2-M_2^2][k^2-m^2_{\tilde{\nu}^R_{L_j}}][k^2-m^2_{\tilde{\nu}^R_{L_i}}]}$$

$$\times \quad \Big[\frac{k^2}{(k^2-\mu_H^2)^2}+\frac{k^2}{(k^2-M_2^2)^2}+\frac{k^2}{(k^2-\mu_H^2)(k^2-M_2^2)}\Big]u_j(p)\epsilon_\mu(q)$$
$$= i\bar{u}_i(p+q)em_{l_j}i\sigma^{\mu\nu}q_\nu\frac{g_2^2}{2\Lambda^4}\sqrt{x_2x_{\mu_H}}\Delta_{ij}^{LL}\tan\beta\Big[k(x_{\mu_H},x_2,x_{\tilde{\nu}_{L_j}^R},x_{\tilde{\nu}_{L_i}^R})$$
$$+k(x_2,x_{\mu_H},x_{\tilde{\nu}_{L_j}^R},x_{\tilde{\nu}_{L_i}^R})-f(x_2,x_{\mu_H},x_{\tilde{\nu}_{L_j}^R},x_{\tilde{\nu}_{L_i}^R})\Big]P_Ru_j(p)\epsilon_\mu(q). \tag{C4}$$

Therefore, in the one-loop contribution from $\tilde{H}^- - \tilde{W}^- - \tilde{\nu}_L^R$, the term including the incident lepton is dominant.


[1] S. T. Petcov, Sov. J. Nucl. Phys., **25**: 340 (1977) JINR-E2-10176

[2] A. M. Baldini et al., (MEG Collaboration), Eur. Phys. J. C, **76**: 434 (2016)

[3] B. Aubert et al., (BABAR Collaboration), Phys. Rev. Lett., **104**: 021802 (2010)

[4] K. Uno et al., (BELLE Collaboration), J. High Energy Phys., **10**: 019 (2021)

[5] A. Ilakovac and A. Pilaftsis, Nucl. Phys. B, **437**: 491 (1995)

[6] R. Diaz, R. Martinez and J. A. Rodriguez, Phys. Rev. D, **63**: 095007 (2001)

[7] M. Kakizaki, Y. Ogura and F. Shima, Phys. Lett. B, **566**: 210 (2003)

[8] E. Arganda and M. J. Herrero, Phys. Rev. D, **73**: 055003 (2006)

[9] T. Toma and A. Vicente, J. High Energy Phys., **01**: 160 (2014)

[10] H. B. Zhang, T. F. Feng, S. M. Zhao and F. Sun, Int. J. Mod. Phys. A, **29**: 1450123 (2014)

[11] S. M. Zhao, T. F. Feng, H. B. Zhang et al., Phys. Rev. D, **92**: 115016 (2015)

[12] J. L. Yang, T. F. Feng, Y. L. Yan et al., Phys. Rev. D, **99**: 015002 (2019)

[13] T. Nomura, H. Okada, Y. Uesaka, Nucl. Phys. B, **962**: 115236 (2021)

[14] V. Barger, P. F. Perez and S. Spinner, Phys. Rev. Lett., **102**: 181802 (2009)

[15] P. F. Perez and S. Spinner, Phys. Lett. B, **673**: 251 (2009)

[16] M. Ambroso and B. A. Ovrut, Int. J. Mod. Phys. A, **26**: 1569 (2011)

[17] P. F. Perez and S. Spinner, Phys. Rev. D, **83**: 035004 (2011)

[18] J. L. Yang, S. M. Zhao, R. F. Zhu et al., Eur. Phys. J. C, **78**: 714 (2018)

[19] T. Moroi, Phys. Rev. D, **53**: 6565 (1996)

[20] E. Arganda, M. J. Herrero and R. Morales et al., J. High Energy Phys., **03**: 055 (2016)

[21] E. Arganda, M. J. Herrero and X. Marcano et al., Phys. Rev. D, **95**: 095029 (2017)

[22] M. J. Herrero, X. Marcano and R. Morales et al., Eur. Phys. J. C, **78**: 815 (2018)

[23] S. M. Zhao, L. H. Su, X. X. Dong et al., J. High Energy Phys., **03**: 101 (2022)

[24] T. T. Wang, S. M. Zhao, J. F. Zhang et al., Eur. Phys. J. C, **82**: 639 (2022)

[25] H. P. Nilles, Phys. Rept., **110**: 1 (1984)

[26] H. E. Haber and G. L. Kane, Phys. Rept., **117**: 75 (1985)

[27] J. Rosiek, Phys. Rev. D, **41**: 3464 (1990)

[28] T. F. Feng and X. Y. Yang, Nucl. Phys. B, **814**: 101 (2009)

[29] C. S. Aulakh, A. Melfo, A. Rasin and G. Senjanovic, Phys. Lett. B, **459**: 557 (1999)

[30] W. Abdallah, A. Hammad, S. Khalil and S. Moretti, Phys. Rev. D, **95**: 055019 (2017)

[31] J. L. Yang, T. F. Feng and H. B. Zhang, Eur. Phys. J. C, **80**: 210 (2020)

[32] S. Khalil and H. Okada, Phys. Rev. D, **79**: 083510 (2009)

[33] L. Basso, B. O'Leary, W. Porod and F. Staub, J. High Energy Phys., **1209**: 054 (2012)

[34] L. D. Rose, S. Khalil, S. J. D. King et al., Phys. Rev. D, **96**: 055004 (2017)

[35] L. D. Rose, S. Khalil, S. J. D.King et al., J. High Energy Phys., **07**: 100 (2018)

[36] P. H. Chankowski, S. Pokorski and J. Wagner, Eur. Phys. J. C, **47**: 187 (2006)

[37] F. Staub, Adv. High Energy Phys., **2015**: 840780 (2015)

[38] F. Staub, (2008) arxiv/hep-ph: 0806.0538

[39] M. Carena, J. R. Espinosaos, C. E. M. Wagner et al., Phys. Lett. B, **355**: 209 (1995)

[40] M. Carena, M. Quiros and C. E. M. Wagner, Nucl. Phys. B, **461**: 407 (1996)

[41] M. Carena, S. Gori, N. R. Shah et al., J. High Energy Phys., **03**: 014 (2012)

[42] T. F. Feng, L. Sun and X. Y. Yang, Nucl. Phys. B, **800**: 221 (2008)

[43] T. F. Feng, L. Sun and X. Y. Yang, Phys. Rev. D, **77**: 116008 (2008)

[44] G. Aad et al., (ATLAS Collaboration), Phys. Lett. B, **796**: 68-87 (2019)

[45] G. Cacciapaglia, C. Csaki, G. Marandella et al., Phys. Rev. D, **74**: 033011 (2006)

[46] M. Carena, A. Daleo, B. A. Dobrescu et al, Phys. Rev. D, **70**: 093009 (2004)

[47] F. Mamoudi, J. High Energy Phys., **12**: 026 (2007)

[48] K. A. Olive and L. Velasco-Sevilla, J. High Energy Phys., **05**: 052 (2008)

[49] B. O'Leary, W. Porod and F. Staub, J. High Energy Phys., **1205**: 042 (2012)

[50] X. X. Dong, T. F. Feng, H. B. Zhang et al., J. High Energy Phys., **12**: 052 (2021)

[51] R. L. Workman et al., (Particle Data Group), Prog. Theor. Exp. Phys., **2022**: 083C01 (2022)

[52] G. Aad et al., (ATLAS and CMS Collaborations), Phys. Rev. Lett., **114**: 191803 (2015)

[53] M. Aaboud et al., (ATLAS Collaboration), Phys. Lett. B, **784**: 345 (2018)

[54] A. M. Sirunyan et al., (CMS Collaboration), Phys. Lett. B, **805**: 135425 (2020)

[55] B. Abi et al., (Muon g-2 Collaboration), Phys. Rev. Lett., **126**: 14, 141801 (2021)

[56] A. Dedes, M. Paraskevas, J. Rosiek, et al., J. High Energy Phys., **06**: 151 (2015)

[57] J. Rosiek, Comput. Phys. Commun. **201**: 144 (2016)

[58] J. Hisano, T. Moroi, K. Tobe, et al., Phys. Lett. B, **357**: 579-587 (1995)

[59] R. Barbieri, L. Hall and A. Strumia, Nucl. Phys. B, **445**: 219-251 (1995)

[60] J. Hisano, T. Moroi, K. Tobe and M. Yamaguchi, Phys. Rev. D, **53**: 2442-2459 (1996)

arXiv:2402.19115v1 [hep-ph] 29 Feb 2024

# Features of strangeness production in pp and heavy ion collisions

A. Pop and M. Petrovici
*National Institute for Physics and Nuclear Engineering - IFIN-HH*
*Hadron Physics Department*
*Bucharest - Romania*
(Dated: March 1, 2024)

Based on the existing experimental data for A-A collisions starting from the Alternating Gradient Synchrotron up to the CERN Large Hadron Collider energies, various systematics related to strange hadrons and anti-hadrons are presented. As in the case of pions, kaons and protons, the ratio between the average transverse momentum and the square root of the total particle multiplicity per unit rapidity and unit transverse overlap area $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp}$ as a function of collision energy for a given centrality or as a function of centrality for a given collision energy supports the predictions of color glass condensate and percolation based approaches. The dependence on $\sqrt{\langle dN/dy\rangle/S_\perp}$ of the slope and offset, extracted from the $\langle p_T\rangle$ - particle mass correlation and the average transverse expansion velocity and kinetic freeze-out temperature parameters obtained from Boltzmann-Gibbs Blast Wave fits of the $p_T$ spectra, for strange hadrons, is compared to that for pions, kaons and protons, previously studied.

The detailed study of the entropy density ($\langle dN/dy\rangle/S_\perp$) dependence of the ratio of strange hadron yields per unit rapidity to the total particle multiplicity per unit rapidity ($Y^S/\langle dN/dy\rangle$) at different collision energies and centralities reveals the necessity to study separately strange hadrons and anti-hadrons. The correlation between the ratio of the single- and multi- strange anti-hadron yield per unit rapidity to the total particle multiplicity per unit rapidity and the entropy density is presented as a function of the fireball size. A maximum is evidenced in the $Y^S/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_\perp$ correlation for combined and separate species of strange hadrons, at different centralities, in the region where a transition from the baryon-dominated matter to the meson-dominated one takes place. Within the experimental error bars, the position of this maximum does not depend on the mass of the corresponding strange hadron. Comparison with pp experimental data reveals another similarity between pp and Pb-Pb collisions at the CERN Large Hadron Collider energies.

## I. INTRODUCTION

The possibility to produce hot and dense matter in heavy ion collisions [1] such that, based on Quantum Chromo-Dynamics (QCD) asymptotic freedom properties, a transition from the hadronic phase to a high density "quark soup" [2, 3] or "quark-gluon plasma" [4] has motivated an unprecedented international effort in building accelerator facilities and complex experimental devices. The objects produced in such collisions have a size at the fermionic level, are highly inhomogeneous and undergo violent dynamics. Therefore, specific experimental probes have to be studied and theoretical approaches, combining different hypotheses for different stages of the formation and evolution of such systems produced in heavy ion collisions, are required for an unambiguous conclusion. The first estimates of the transition from a gas of free nucleons to hadronic matter and subsequently to deconfined matter as a function of density were done within the percolation approach [5, 6]. Phenomenological models predicted some discontinuities in the behaviour of different observables as a function of collision energy or centrality specific for a phase transition between two thermodynamic states in a closed volume [7–9]. Recently, based on the existing experimental results from the Alternating Gradient Synchrotron (AGS), Super Proton Synchrotron (SPS), BNL Relativistic Heavy Ion Collider (RHIC) and CERN Large Hadron Collider (LHC), it was evidenced such a trend in the dependence of the ratio of the energy density to the entropy density ($\langle dE_T/dy\rangle/\langle dN/dy\rangle$) as a function of entropy density at different collision centralities for A-A collisions [10]. The ratio of the energy density to the entropy density, at a given value of the transverse overlap area, increases with entropy. A tendency towards saturation at values of the entropy density beyond 6-8 $fm^{-2}$, corresponding to the largest collision energies at RHIC, and a steep rise at the LHC energies is evidenced for central collisions. Worth mentioning that for central collisions, a change in the collision energy dependence of the ratio $(1-R_{AA}^{\pi^0})/\langle dN/dy\rangle$ [11] takes place in the same energy range as the one corresponding to the transition from an increase to the saturation in the entropy density dependence of the $\langle dE_T/dy\rangle/\langle dN_{ch}/dy\rangle$. Such trends are in qualitative agreement with theoretical model predictions [8, 9] and [12–14], respectively. 40 years ago, well before the experimental data became available, the enhancement of the strangeness production was advocated as sensitive probe for deconfinement [15]. A series of experiments, from SPS to LHC energies evidenced an enhancement of strange hadron production as a function of centrality relative to the one corresponding to the pp minimum bias collision at the same energy. The influence of the core-corona relative contribution on the centrality dependence of the strangeness production, average transverse momenta, elliptic flow or $p_T$ spectra in heavy-ion collisions at SPS, RHIC and LHC energies was reported in many papers [16–28]. As far as con-

cerns the centrality dependence of the strange hadron production, it was clearly shown that this is the result of the interplay between the corona and core relative contributions. A consistent approach based on string density before hadronization within the EPOS model using some global parameters [29] has shown that the core-corona contribution also explains the similarity between Pb-Pb, p-Pb and pp in terms of strange hadron production as a function of charged particle density at mid-rapidity evidenced at LHC [30]. A core-corona approach based on a new microcanonical hadronic recipe (EPOS4) [31] confirms the previous results. Once the experimental data on strangeness production at LHC energies were available, it has been shown that the ratio of the strange hadron yield to entropy as a function of the entropy density, for the most central collisions [32], qualitatively follows a similar trend as the one observed in the $\langle dE_T/dy\rangle/\langle dN/dy\rangle - \langle dN/dy\rangle/S_\perp$ correlation, indicative for a transition from a hadronic gas to a mixed phase of hadronic and deconfined matter. In the present paper, using the most complete experimental information from AGS up to the LHC energies, a detailed study of the ratio between the single- and multi- strange hadron and anti-hadron yields per unit rapidity and the total particle multiplicity per unit rapidity relative to the entropy density, at different centralities and collision energies, is presented. Similarities between pp and Pb-Pb collisions at LHC energies are evidenced. Results on the entropy density dependence of observables related to the dynamics of the fireball are presented in Section II. Section III is dedicated to a detailed study of the entropy density dependence of the ratio between single- and multi- strange hadron and anti-hadron yields per unit rapidity and the total particle multiplicity per unit rapidity at different centralities and collision energies. In Section IV various aspects of such correlations are presented for different quark/antiquark compositions, fireball sizes and centralities. Similarities in terms of such correlations between A-A and pp collisions at LHC energies are discussed in Section V. Conclusions are presented in Section VI.

## II. COLLISION ENERGY AND CENTRALITY DEPENDENCE OF $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp}$; GEOMETRICAL SCALING

Previous papers [10, 33] confirmed the model prediction based on local parton-hadron duality (LPHD) [34] and dimensionality argument [35, 36] concerning the behaviour of the ratio between the average transverse momentum and the square root of the total particle multiplicity per unit rapidity and unit of the colliding nuclei transverse overlap area ($\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp}$) as a function of centrality and collision energy for pions, kaons and protons. The results of similar studies in the case of strange hadrons $K^0_S$, $\bar{\Lambda}$, $\Xi^-$ are presented in Figure 1 as a function of centrality (number of participating nucleons $\langle N_{part}\rangle$), for different collision energies and in Figure 2 as a function of collision energy, $\sqrt{s_{NN}}$, for different centralities. Similar with the trends shown for pions, kaons and protons [10], a decrease of the $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp}$ with centrality at all collision energies and as a function of collision energy at a given centrality is evidenced. $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp}$ decreases from $\sqrt{s_{NN}} = 7.7$ GeV up

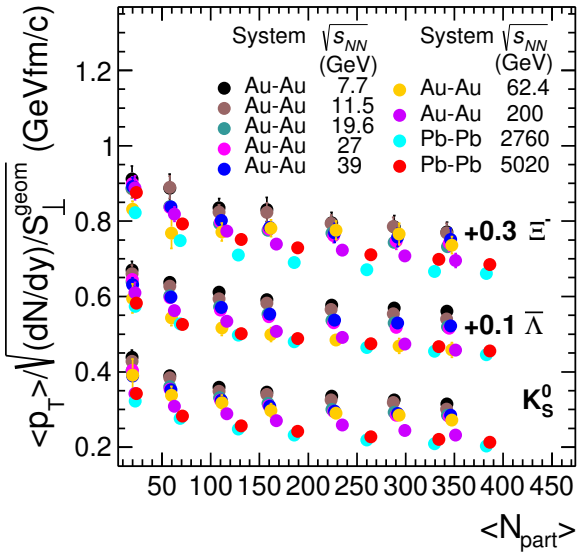



FIG. 1. $\langle N_{part}\rangle$ dependence of $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp^{geom}}$ for $K^0_S$, $\bar{\Lambda}$, $\Xi^-$ at different collision energies.

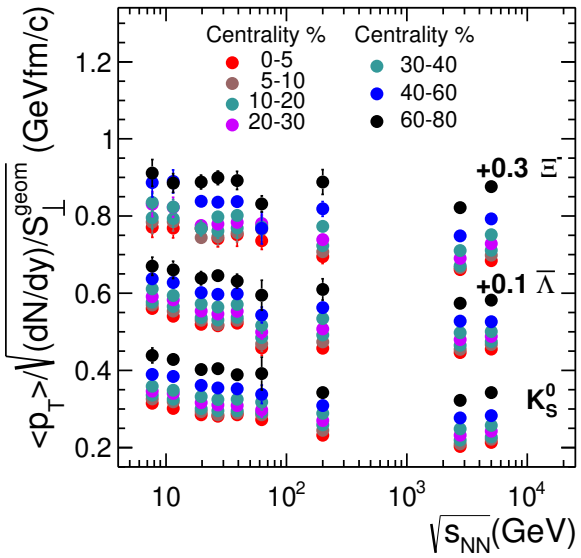



FIG. 2. Collision energy, $\sqrt{s_{NN}}$, dependence of $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp^{geom}}$ for $K^0_S$, $\bar{\Lambda}$, $\Xi^-$ at different centralities.

to 200 GeV. The values corresponding to LHC energies are slightly smaller than the ones corresponding to the top RHIC energy. The geometrical transverse overlap area ($S_{\perp}^{geom}$) of the two colliding nuclei for a given incident energy and centrality was estimated on the basis of a Glauber Monte-Carlo (MC) approach [37–40] as explained in [33] where the calculated values were compiled. These values are used in the present paper. The total final state particle multiplicity per unit rapidity, $dN/dy$, a measure of the entropy produced in the collision, was estimated according to:

$$\frac{dN}{dy} \simeq \frac{3}{2}\frac{dN}{dy}^{(\pi^+ + \pi^-)} + 2\frac{dN}{dy}^{(p+\bar{p},\Xi^- + \bar{\Xi}^+)} + \frac{dN}{dy}^{(K^+ + K^-, \Lambda + \bar{\Lambda}, \Omega^- + \bar{\Omega}^+)} + 2\frac{dN}{dy}^{K^0_S} \quad (1)$$

The experimental data in terms of transverse momentum spectra, yields and average transverse momenta were taken from: [41, 42] (AGS and SPS), [43, 44] (Au-Au, Beam Energy Scan (BES)), [45, 46] (Au-Au at $\sqrt{s_{NN}}$ = 62.4 GeV), [45, 47, 48] (Au-Au at $\sqrt{s_{NN}}$ = 200 GeV), [49–51] (Pb-Pb at $\sqrt{s_{NN}}$ = 2.76 TeV) and [52–56] (Pb-Pb at $\sqrt{s_{NN}}$ = 5.02 TeV). Eq. 1 was applied by taking into consideration all the measured particles and antiparticles depending, from system to system and case to case, on the most complete published experimental information.

Results of geometrical scaling, in terms of estimated slopes and offsets of the linear $\langle p_T \rangle$-hadron mass dependence and average transverse flow velocity ($\langle \beta_T \rangle$) and kinetic freeze-out temperature ($T_{kin}$) parameters, from the Boltzmann-Gibbs Blast Wave (BGBW) fits of $p_T$ spectra in relation to the geometrical variable $\sqrt{\langle dN/dy \rangle / S_{\perp}}$ and comparison with pp collisions at LHC energies for pions, kaons and protons were reported in [33] and for $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$ in [57]. The similarities and differences in terms of these correlations between $\pi^+$, $K^+$, p and $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$ can be followed in Figures 3 - 6. Qualitatively, there are similarities between the two categories of hadrons, i.e. a linear dependence on $\sqrt{\langle dN/dy \rangle / S_{\perp}}$ of the slopes of $\langle p_T \rangle$ as a function of hadron mass and a constant value of the corresponding offsets, separately for RHIC and LHC energies. Quantitatively, there are differences, i.e. the slopes corresponding to $\pi^+$, $K^+$ and p are larger than the ones corresponding to $K^0_S, \Lambda, \Xi^-$ and $\Omega^-$ while the offsets are larger for strange hadrons relative to $\pi^+$, $K^+$ and p. This is in agreement with the dynamical freeze out scenario, namely, $K^0_S, \Lambda, \Xi^-$ and $\Omega^-$, suffering negligible interaction with the hadron environment, conserve the momentum distribution corresponding to their formation time, while $\pi^+$, $K^+$ and p could build up extra expansion via elastic interaction with the hadronic medium, resulting in an increase in the slope of $\langle p_T \rangle$ as a function of mass. The offsets, within the error bars, are constant as a function of $\sqrt{\langle dN/dy \rangle / S_{\perp}}$, the values at LHC energies being larger than at RHIC energies. In both energy ranges, the offsets associated to strange hadrons are larger than in the case of $\pi^+$, $K^+$ and p. Within the string percolation scenario, inspired by color glass condensate (CGC) [58, 59] such a difference could

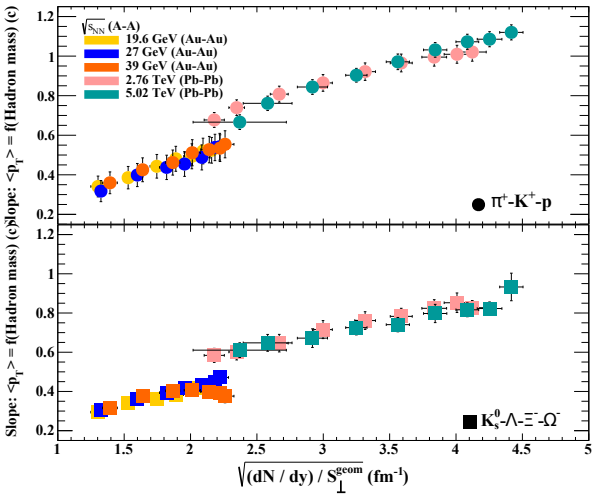



FIG. 3. $\sqrt{\langle dN/dy \rangle / S_{\perp}}$ dependence of the slope obtained from the linear fit of the $\langle p_T \rangle$-hadron mass dependence for $\pi^+$, $K^+$ and p - top and for $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$ - bottom.

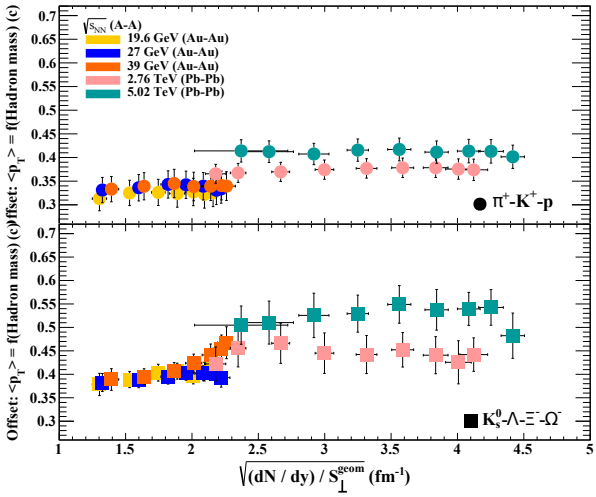



FIG. 4. $\sqrt{\langle dN/dy \rangle / S_{\perp}}$ dependence of the offset obtained from the linear fit of the $\langle p_T \rangle$ - hadron mass dependence for $\pi^+$, $K^+$ and p - top and for $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$ - bottom.

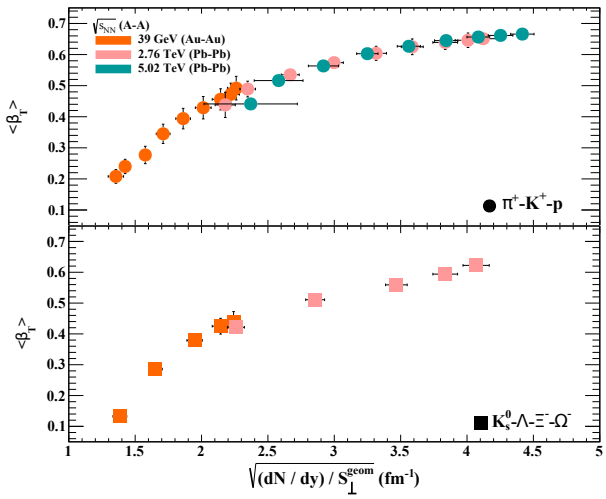



FIG. 5. $\sqrt{\langle dN/dy\rangle/S_\perp}$ dependence of $\langle\beta_T\rangle$ obtained from the BGBW simultaneous $p_T$ spectra fits of $\pi^+$, $K^+$ and p - top and of $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$ - bottom.

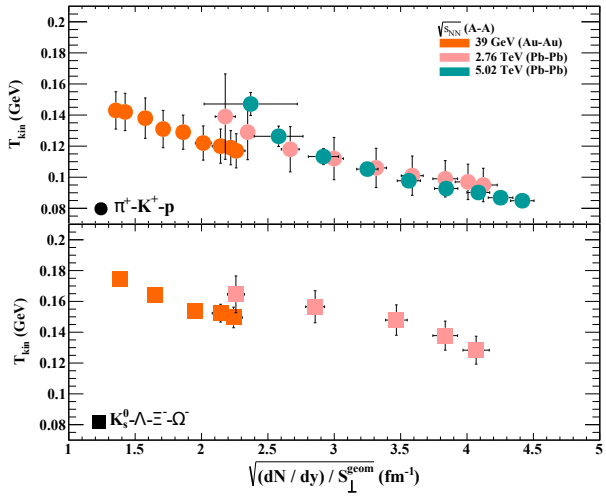



FIG. 6. $\sqrt{\langle dN/dy\rangle/S_\perp}$ dependence of $T_{kin}$ obtained from the BGBW simultaneous $p_T$ spectra fits of $\pi^+$, $K^+$ and p - top and of $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$ - bottom.

be explained by a higher percolation probability at LHC energies relative to the RHIC ones, at LHC energies the gluon density per unit of rapidity and unit overlap area being by a factor 3-4 larger than at the top RHIC energy [60]. Simultaneous fits of the $p_T$ spectra of $\pi^+$, $K^+$ and p and separately of $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$ using the BGBW expression inspired by hydrodynamical models [61] were performed and the values of the fit parameters namely, the average transverse expansion velocity ($\langle\beta_T\rangle$) and the kinetic freeze-out temperature ($T_{kin}$), are presented in Figures 5 and 6, respectively. $\langle\beta_T\rangle$ scales as a function of $\sqrt{\langle dN/dy\rangle/S_\perp}$, separately for $\pi^+$, $K^+$ and p and $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$ with lower values for strange hadrons at a given $\sqrt{\langle dN/dy\rangle/S_\perp}$, in agreement with the trends observed in Figure 3, having the same explanation. The kinetic freeze-out temperature decreases with increasing $\sqrt{\langle dN/dy\rangle/S_\perp}$, the values for $\pi^+$, $K^+$ and p being lower than the ones coresponding to $K^0_S$, $\Lambda$, $\Xi^-$ and $\Omega^-$. For both hadron families, the extracted values at LHC energies are larger than the corresponding ones at RHIC energies for a given value of $\sqrt{\langle dN/dy\rangle/S_\perp}$.

## III. ENTROPY DENSITY DEPENDENCE OF THE RATIO OF STRANGE HADRON YIELDS PER UNIT RAPIDITY TO THE TOTAL PARTICLE MULTIPLICITY PER UNIT RAPIDITY AS A FUNCTION OF QUARK/ANTIQUARK COMPOSITION, COLLISION ENERGY AND CENTRALITY

Since the entropy and strangeness are conserved during the confinement process, the ratio between the strange hadron yield and entropy has been proposed as diagnostic tool for deconfinement [62, 63]. The entropy density per unit transverse overlap area, assuming an isentropic expansion, is proportional to the total particle multiplicity per unit rapidity and unit transverse overlap area. In order to compare different collision systems Au-Au at RHIC energies, Pb-Pb and pp at LHC energies, based on the present experimental information, the correlated dependence of the ratio between the strange hadron yield per unit rapidity and the total particle multiplicity per unit rapidity on the entropy density ($\langle dN/dy\rangle/S_\perp$) was scrutinized. Such correlations are investigated separately for single- and multi-strange hadrons. The total single- ($Y^{S-1s}$) and multi- ($Y^{S-ms}$) strange hadron yields per unit rapidity, are estimated from the existing experimental data with:

$$Y^{S-1s} = \frac{dN}{dy}^{S-1s} = \frac{dN}{dy}^{(K^+ + K^-)} + 2\frac{dN}{dy}^{K^0_S} + \frac{dN}{dy}^{(\Lambda+\bar{\Lambda})} \quad (2)$$

$$Y^{S-ms} = \frac{dN}{dy}^{S-ms} = 2\frac{dN}{dy}^{(\Xi^- + \bar{\Xi}^+)} + \frac{dN}{dy}^{(\Omega^- + \bar{\Omega}^+)} \quad (3)$$

By examining Figure 7, where the ratio of the total single- strange hadron yield per unit rapidity (Eq. 2) to the

total particle multiplicity per unit rapidity as a function of entropy density is presented, one can conclude that below $\sqrt{s_{NN}}$ = 62.4 GeV a scaling is evidenced, $Y^{S-1s}/\langle dN/dy\rangle$ increasing with the entropy density. A similar trend is observed at $\sqrt{s_{NN}}$ = 62.4 GeV with a systematic lower value of $Y^{S-1s}/\langle dN/dy\rangle$ for a given entropy density. Within the error bars, the results at $\sqrt{s_{NN}}$ = 200 GeV seem to be in line with the LHC values, $Y^{S-1s}/\langle dN/dy\rangle$ being almost independent on the entropy density. Early theoretical studies of strangeness in relativistic heavy ion collisions [64, 65] have concluded that especially strange anti-hadrons are sensitive probes for deconfinment. Therefore, the same type of representations, separately for single-strange hadrons and anti-hadrons are presented in Figures 8 and 9, respectively. Figure 8 clearly shows that $Y^{S-1s-part}/\langle dN/dy\rangle$ for single-strange hadrons increases with entropy density for a given collision energy and decreases with the collision energy as a consequence of in-medium effects in the presence of baryonic matter in the mid-rapidity region, decreasing with the collision energy as the net baryonic matter at mid-rapidity approaches a negligible value at the top RHIC energy, the collision becoming almost transparent [43]. The trend for the single-strange anti-hadrons presented in Figure 9 is completely different, i.e. $Y^{S-1s-antipart}/\langle dN/dy\rangle$ increases with the collision energy up to $\sqrt{s_{NN}}$ = 200 GeV and for a given collision energy increases with entropy density bending towards larger entropy density. The collision energy and entropy density dependence of single-strange anti-hadrons is indicative of their production once the creation of pairs of quarks and antiquarks becomes possible. The tendency towards a levelling off at large entropy density is the consequence of the surrounding hadronic matter which increases towards central collisions or larger $\langle dN/dy\rangle/S_{\perp}$. At LHC energies where the mechanism of creation of pairs of quarks and antiquarks becomes dominant and most of the fireball reaches a deconfined phase, there is no difference between the behaviour of single-strange hadrons and single-strange anti-hadrons and the dependence of $Y^{S}/\langle dN/dy\rangle$ on the entropy density becomes negligible. The results from $\sqrt{s_{NN}}$ = 200 GeV are in-line with the LHC ones at the same entropy density and no dependence on the entropy density can be observed within the experimental errors. The $Y^{S}/\langle dN/dy\rangle$ values corresponding to the Pb-Pb collision at $\sqrt{s_{NN}}$ = 5.02 TeV are systematically a bit larger than the ones corresponding to the Pb-Pb collision at $\sqrt{s_{NN}}$ = 2.76 TeV. The multi-strange hadron yields are more sensitive to the production mechanism originating in the fireball, therefore are expected to be more sensitive to the fireball composition. Similar to the previous cases, it is worth to look to $Y^{S-ms}/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_{\perp}$ for combined multi-strange hadrons and their anti-partners. The result is presented in Figure 10. A rather good scaling can be observed with a steep linear rise up to $\langle dN/dy\rangle/S_{\perp} \approx 3\ fm^{-2}$ at the top RHIC energy followed by a transition to the regions corresponding to LHC energies. However,

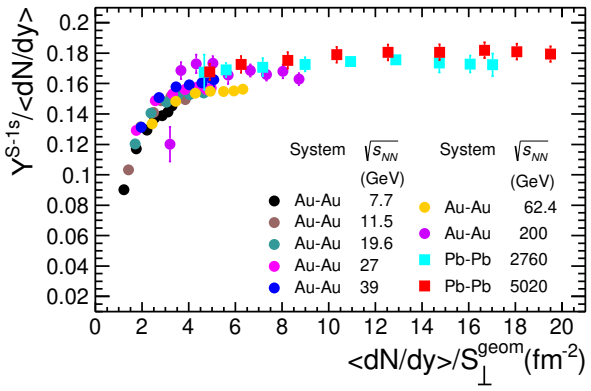



FIG. 7. The ratio of the single-strange hadron+anti-hadron yield per unit rapidity (Eq. 2) to the total particle multiplicity per unit rapidity as a function of entropy density from $\sqrt{s_{NN}}$ = 7.7 GeV up to 5.02 TeV.

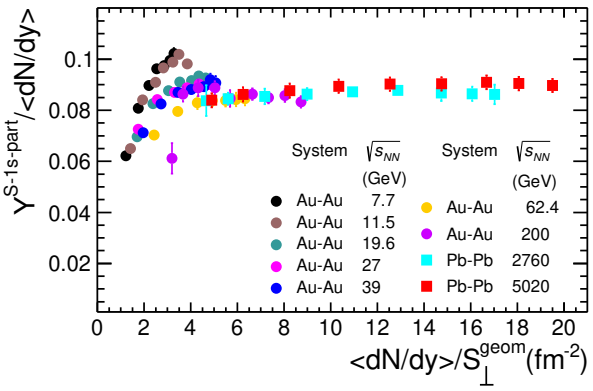



FIG. 8. The ratio of the single-strange hadron yield per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density from $\sqrt{s_{NN}}$ = 7.7 GeV up to 5.02 TeV.

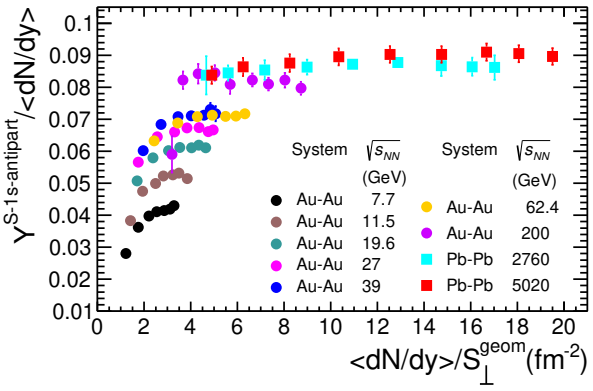



FIG. 9. The ratio of the single-strange anti-hadron yield per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density from $\sqrt{s_{NN}}$ = 7.7 GeV up to 5.02 TeV.

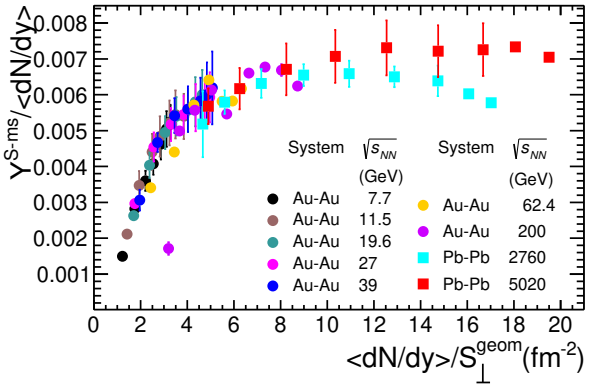



FIG. 10. The ratio of the multi-strange hadron+anti-hadron yield per unit rapidity (Eq. 3) to the total particle multiplicity per unit rapidity as a function of entropy density from $\sqrt{s_{NN}}$ = 7.7 GeV up to 5.02 TeV.

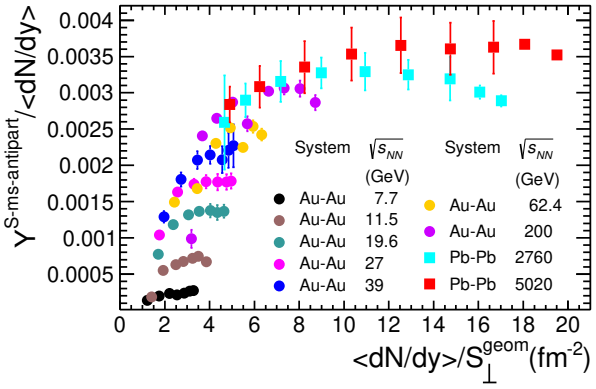



FIG. 12. The ratio of the multi-strange anti-hadron yield per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density from $\sqrt{s_{NN}}$ = 7.7 GeV up to 5.02 TeV.

this result is nothing else than a combination of two very different dependences of $Y^S/\langle dN/dy\rangle$ as a function of $\langle dN/dy\rangle/S_{\perp}$ for multi-strange hadrons and multi-strange anti-hadrons as can be followed in Figures 11 and 12.

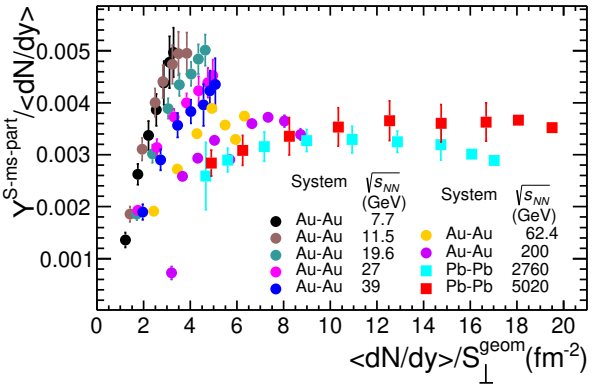



FIG. 11. The ratio of the multi-strange hadron yield per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density from $\sqrt{s_{NN}}$ = 7.7 GeV up to 5.02 TeV.

For multi-strange hadrons, the considerations made relative to the trends evidenced in Figure 8 are more evident. The $Y^{S-ms-part}/\langle dN/dy\rangle$ for multi-strange hadrons has a steep linear increase with the entropy density, the slope decreasing going from $\sqrt{s_{NN}}$ = 7.7 GeV up to 39 GeV, at 200 GeV converging towards the values corresponding to $\sqrt{s_{NN}}$ = 5.02 TeV. At the LHC energy of $\sqrt{s_{NN}}$ = 5.02 TeV, a very small increase of $Y^{S-ms}/\langle dN/dy\rangle$ as a function of entropy density is observed. The entropy density dependence of $Y^{S-ms-antipart}/\langle dN/dy\rangle$ for the multi-strange anti-hadrons presented in Figure 12 is completely different. $Y^{S-ms-antipart}/\langle dN/dy\rangle$ increases with collision energy and for a given collision energy increases with the entropy density with a tendency towards a plateau at larger entropy density/larger centrality. These results show that an integral representation as those presented in Figures 7 and 10, in which strange hadrons and anti-hadrons are summed together, hides all the trends revealed in separate representations, Figures 8, 9, 11 and 12, and makes it difficult to take it as testing ground for theoretical models. Moreover, because of the much larger yield of single-strange hadrons relative to the multi-strange hadrons one, the behaviour of the established correlations in the second case is obscured in an integral representation for the total strange hadron yield. The trends evidenced in Figures 9 and 12 are very similar with those observed in the $\langle dE_T/dy\rangle/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_{\perp}$ correlation for different collision energies and different centralities [10]. This shows that there is a strong correlation between the strange anti-hadron production and energy density. A similar conclusion was reached by studying the degree of strangeness suppression in hadronic and nuclear collisions [66].

## IV. VARIOUS ASPECTS RELATED TO THE $Y^S/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_{\perp}$ CORRELATION FOR DIFFERENT QUARK/ANTIQUARK COMPOSITIONS AT DIFFERENT CENTRALITIES

Four different values of the transverse overlap area, namely, $S_{\perp}$ = 149, 108, 64, 20 $fm^2$ have been chosen to inspect the fireball size dependence of the strangeness/entropy-entropy density correlation. The corresponding values of the ratio of the strange anti-hadron yields per unit rapidity to the total particle multiplicity per unit rapidity were obtained with the interpolation of the experimental data measured at different centralities. The results are presented in Figures 13 and 14 for single- and multi- strange anti-hadrons, respectively.

A clear dependence on the size of the transverse over-

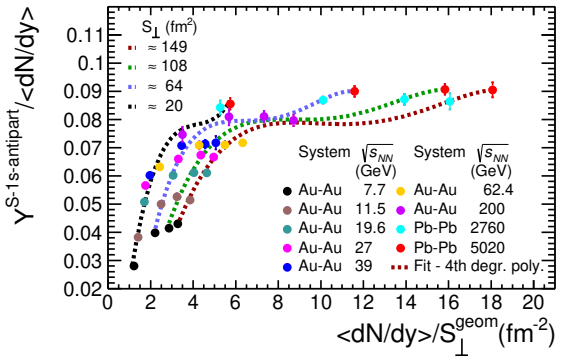



FIG. 13. The ratio of the single-strange anti-hadron yield per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density for four values of the transverse overlap area. The dashed lines are the result of a fourth degree polynomial fit used to guide the eye.

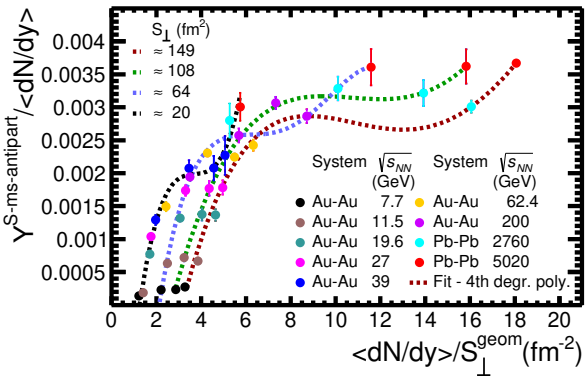



FIG. 14. The ratio of the multi-strange anti-hadron yield per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density for four values of the transverse overlap area. The dashed lines are the result of a fourth degree polynomial fit used to guide the eye.

lap area is observed. The rise in $Y^{S-1s-antipart}/\langle dN/dy\rangle$ and $Y^{S-ms-antipart}/\langle dN/dy\rangle$ at RHIC energies becomes steeper and the range in $\langle dN/dy\rangle/S_\perp$, corresponding to a close to a plateau, decreases from central to peripheral collisions. The trends resemble those observed in Fig. 5 from [10] for the $\langle dE_T/dy\rangle/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_\perp$ correlation. The dashed lines represent the result of the fit of the corresponding points with a fourth degree polynomial. At each inflection point of the fit lines, the corresponding values of the ratio of the yield per unit rapidity to the total particle multiplicity per unit rapidity was estimated and represented in Figures 15 and 16 for single- and multi-strange anti-hadrons, respectively.

While the inflection point values are constant in Figure 15, Figure 16 looks very much like Figure 6 in [10], a slight dependence on the overlap area being evidenced.

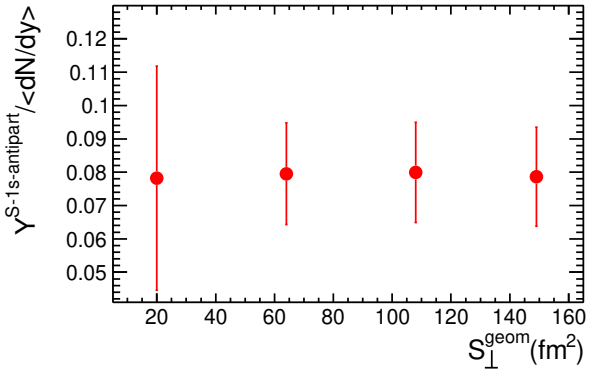



FIG. 15. $Y^{S-1s-antipart}/\langle dN/dy\rangle$ calculated at the inflection point in the $Y^{S-1s-antipart}/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_\perp$ correlation at different collision energies for four values of $S_\perp^{geom}$.

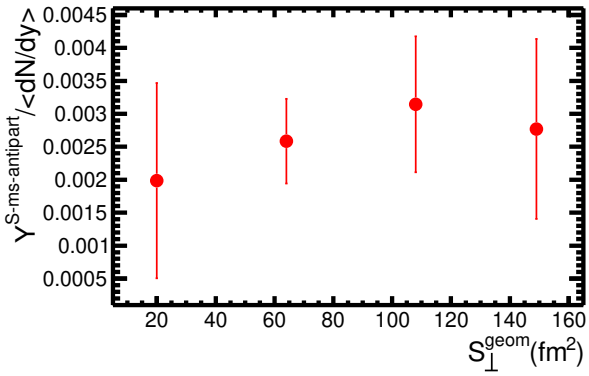



FIG. 16. $Y^{S-ms-antipart}/\langle dN/dy\rangle$ calculated at the inflection point in the $Y^{S-ms-antipart}/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_\perp$ correlation at different collision energies for four values of $S_\perp^{geom}$.

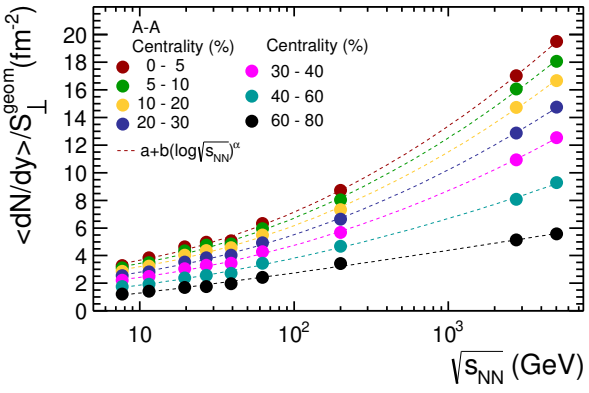



FIG. 17. Collision energy dependence of $(dN/dy)/S_\perp$ for different centralities.

These considerations and the following ones can be better understood based on the collision energy dependence of $(dN/dy)/S_\perp$ for different centralities, presented in Figure 17. The range in $< dN/dy > /S_\perp$ as a function of collision energy is increasing from peripheral to central collisions. The dashed lines are the result of the fit with the formula $< dN/dy > /S_\perp = a + b(\log\sqrt{s_{NN}})^\alpha$ and the fit coefficients corresponding to the different centralities are listed in Table I. In Figures 18 and 19 are represented the ratios of single- and multi-strange hadron (full symbols) and single- and multi-strange anti-hadron (open symbols), respectively, yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density, for four values of centrality. A steep increase at lower collision energies is followed by a leveling off and another increase is observed at the largest LHC energy, more pronounced in the case of multi-strange hadrons. The ratio for single- and multi-strange hadrons relative to anti-hadrons, at lower entropy density, decreases from central to peripheral collisions, becoming similar towards the plateau region. This shows that decreasing the number of participants the single collisions become the main source of strange hadron production.

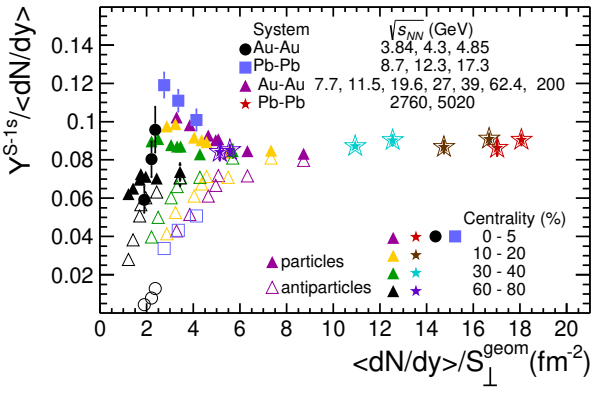



FIG. 18. The ratio of single-strange hadron (full symbols) and single-strange anti-hadron (open symbols) yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density for four values of centrality.

The data for multi-strange hadrons exist at AGS energies only for 0-5 % centrality. Nevertheless, a maximum in $Y^{S-1s}/\langle dN/dy\rangle$ and $Y^{S-ms}/\langle dN/dy\rangle$ in the case of particles as a function of entropy density starts to be evidenced going from central to peripheral collisions. While the entropy density corresponding to the maximum decreases from central to peripheral collisions, the corresponding value of the collision energy is slightly increasing, as can be seen in Figure 17, where the $\langle dN/dy\rangle$ - $\sqrt{s_{NN}}$ correlation for different collision centralities is represented. Such a behaviour is in a qualitative agrement with the results of the parton-hadron string dynamics (PHSD) transport model [67] which predicts a slight increase in the collision energy at which the maximum in the $K^+/\pi^+$ versus the collision energy representation appears going from Au-Au to C-C central collisions. However, one should mention that the model does not evidence a similar trend in $(\Lambda+\Sigma^0)$ as a function of collision energy for the three symmetric colliding systems, i.e. Au-Au, Ca-Ca and C-C. The decrease of the multi-strange hadron yield per unit rapidity to the total particle multiplicity per unit rapidity for peripheral collisions supports the conclusion of the same approach that the main contribution for the observed increase at lower entropy density/collision energy is due to chiral symmetry restoration which should decrease towards peripheral collisions, where the transparency is larger and the baryon density of the fireball is lower. Focusing only on the more complete data for central collisions a kink is clearly evidenced in the correlation for single- and multi- strange hadrons around the entropy density of $\approx 2.5$ - 3 $fm^2$. Beyond this value, $Y^S/\langle dN/dy\rangle$ decreases and converges to the value of single- and multi-strange anti-hadrons at $\sqrt{s_{NN}} = 200$ GeV. The $Y^S/\langle dN/dy\rangle$ for single- and multi-strange anti-hadrons increases with $\langle dN/dy\rangle/S_\perp$ up to the largest RHIC energies, the value corresponding to $\sqrt{s_{NN}} = 200$ GeV being similar to that of the single- and multi-strange hadrons, respectively. The same type of representation for kaons can be followed in Figure 20. Apart the absolute values, the trend in $\langle dN/dy\rangle/S_\perp$ dependence of $K^+$, $K^-$ yields per unit rapidity to the total particle multiplicity per unit rapidity are similar with those corresponding to single- and multi-strange hadrons represented in Figures 18 and 19, the kink location in $\langle dN/dy\rangle/S_\perp$ being the same. $K^0_S$ follows more or less the same trend as $K^-$ as its formation requires the presence of strange quark-antiquark pairs. The same type of representation for the $\Lambda$ and $\bar{\Lambda}$ ratios of yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density, presented in Figure 21, shows that the kink position is the same as in the case of Figures 18, 19 and 20. Figure 22 is the same type of representation for

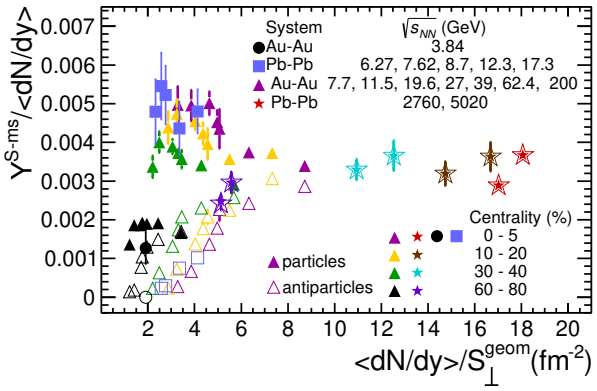



FIG. 19. The ratio of multi-strange hadron (full symbols) and multi-strange anti-hadron (open symbols) yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density for four values of centrality.

TABLE I. Coefficients of the fit with the formula from Figure 17

| Centrality (%) | a | b | $\alpha$ |
|---|---|---|---|
| 0 - 5 | 2.6399 ± 0.2284 | 1.0239 ± 0.1365 | 1.0700 ± 0.0475 |
| 5 - 10 | 2.4497 ± 0.2338 | 0.9829 ± 0.1413 | 1.0590 ± 0.0511 |
| 10 - 20 | 2.3246 ± 0.1970 | 0.8712 ± 0.1176 | 1.0718 ± 0.0481 |
| 20 - 30 | 1.9471 ± 0.1528 | 0.8518 ± 0.0944 | 1.0341 ± 0.0392 |
| 30 - 40 | 1.7053 ± 0.1438 | 0.7293 ± 0.0893 | 1.0293 ± 0.0433 |
| 40 - 60 | 1.0456 ± 0.1861 | 0.8217 ± 0.1314 | 0.8767 ± 0.0544 |
| 60 - 80 | 0.1873 ± 0.3645 | 1.1015 ± 0.3083 | 0.6086 ± 0.0844 |

the ratio of the $\Xi^-$, $\bar{\Xi}^+$ yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density. The difference between the ratio of

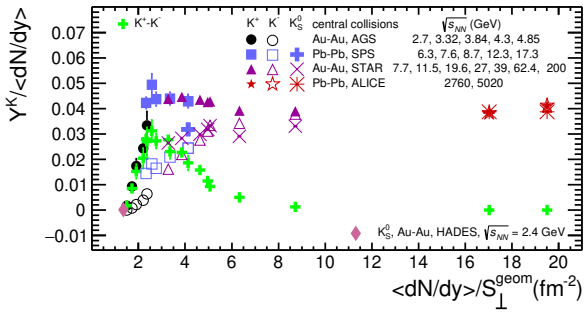



FIG. 20. The ratio of the yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density for $K^+$ (full symbols), $K^-$ (open symbols), their difference (green crosses) and $K^0_S$, from AGS to LHC energies, for central collisions.

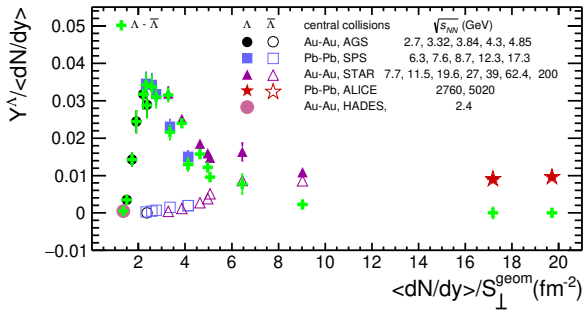



FIG. 21. The ratio of the yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density for $\Lambda$ (full symbols), $\bar{\Lambda}$ (open symbols), their difference (green crosses) from AGS to LHC energies, for central collisions.

the yields to the total particle multiplicity corresponding to hadrons and anti-hadrons is represented by the green crosses in Figures 20 - 22. In such representation the kink is better evidenced, its location in $\langle dN/dy\rangle/S_\perp$ remaining unchanged at around the entropy density of $\approx$ 2.5 - 3 $fm^2$. All these representations also highlight

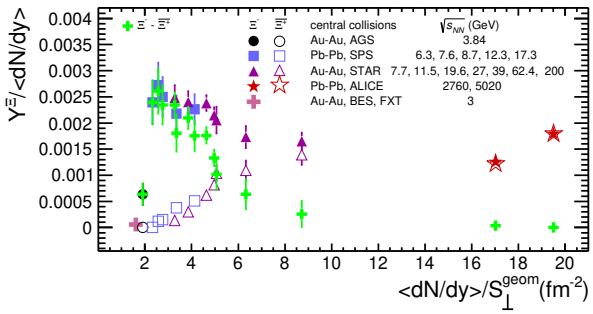



FIG. 22. The ratio of the yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density for $\Xi^-$ (full symbols), $\bar{\Xi}^+$ (open symbols),their difference (green crosses) from AGS to LHC energies, for central collisions.

the fact that the new data from HADES [68] and BES fixed-target (FXT) [69] experiments fit well into the general systematics. Such a sharp peak, evidenced for the first time at SPS based on data from AGS and SPS in the $K^+/\pi^+$ ratio as a function of $\sqrt{s_{NN}}$ for central collisions, coined "horn" was considered to be a signature of a phase transition from a hadron gas to QGP [70, 71]. Within the framework of a thermal model [72] it was shown that such a kink type behaviour is related to a transition from a baryon dominated hadronic gas to a meson dominated one. Based on the PHSD approach [73] the enhancement in $K^+/\pi^+$ and $(\Lambda + \Sigma^0)/\pi^-$ evidenced at $\sqrt{s_{NN}} \approx 6$ GeV is caused by chiral symmetry restoration while deconfinement is essential in producing the maximum. In Figure 23 is represented the dependence on the entropy density of the ratio of the yield per unit rapidity to the total particle multiplicity per unit rapidity for pions and protons for central collisions. The ratio of the proton yield per unit rapidity to the total particle multiplicity per unit rapidity decreases up to the top RHIC energy, becoming negligible at the LHC energies. The ratio of the pion yield per unit rapidity to the total particle multiplicity per unit rapidity increases up to the top RHIC energy and remains constant up to the LHC energies. They cross each other at the entropy density of $\approx$ 2.5 - 3 $fm^{-2}$ where a transition from a baryon-dominated to a meson-dominated matter takes place [72]. This value is in the region corresponding to the kinks positions ev-

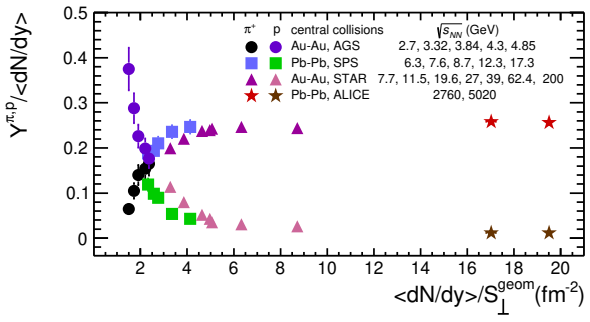



FIG. 23. Entropy density dependence of the ratio of the yield per unit rapidity to the total particle multiplicity per unit rapidity of pions and protons for central collisions.

idenced in Figures 18 - 22 and corresponds to $\sqrt{s_{NN}} \approx$ 6 GeV. It is well below the energy where a plateau is reached in the $\langle dE_T/dy\rangle/\langle dN/dy\rangle - \langle dN/dy\rangle/S_\perp$ correlation or in the collision energy dependence of the slope of the Bjorken energy density times the interaction time as a function of $\langle dN/dy\rangle/S_\perp$ [10], thus indicating a transition from a hadronic phase to a coexistence of hadronic and deconfined matter. In view of the results obtained using the PHSD transport approach [73], the strangeness enhancement up to the maximum is due to chiral symmetry restoration at high baryon density. Therefore, the position of the maximum could be indicative for a maximum of the baryon density of the fireball [72, 74]. Contrary to the expectations based on the statistical model [72] the kink position stays at the same value of entropy density/collision energy for $K$, $\Lambda$ and $\Xi^-$, no mass dependence being observed. The centrality dependence of the kink position evidenced in the $Y^S/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_\perp$ correlation for particles excludes the interpretation as signature for critical point [75].

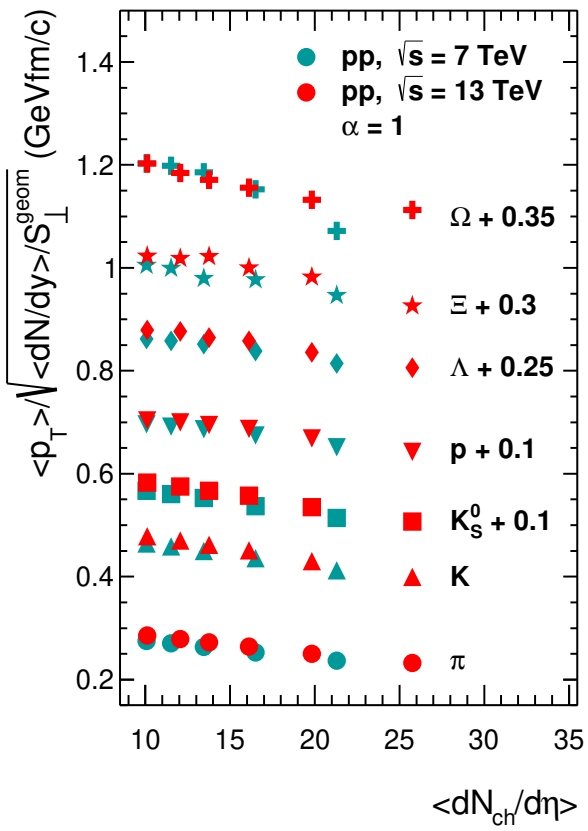



FIG. 24. The $\langle dN_{ch}/d\eta\rangle$ dependence of $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp}$ for $\pi$, K, $K^0_S$, p, $\Lambda$, $\Xi$ and $\Omega$ measured in pp collisions at $\sqrt{s}$ =7 TeV and $\sqrt{s}$ = 13 TeV.

## V. COMPARISON WITH THE RESULTS FROM *pp* COLLISIONS AT LHC ENERGIES

Measurements in pp collisions up to very high charged particle multiplicities at LHC energies, revealed similarities between pp and Pb-Pb in many respects, i.e. near-side long range pseudorapidity correlations [76], the correlation among the parameters extracted from the $p_T$ distributions fitted with a BGBW formula ($\langle\beta_T\rangle - T^{fo}_{kin}$) as a function of charged particle multiplicity [77], azimuthal angular correlations [78], geometrical scaling [33, 57, 79], etc. Therefore, in order to have an unambiguous conclusion on the origin of such similarities, the studies of the types of correlations presented in the previous chapters for A-A collisions are extended to pp. Experimental data on transverse momentum spectra for light flavors as a function of the average charged particle multiplicity at LHC energies were published in [80–82]. In order to calculate the entropy density, an estimate of the transverse

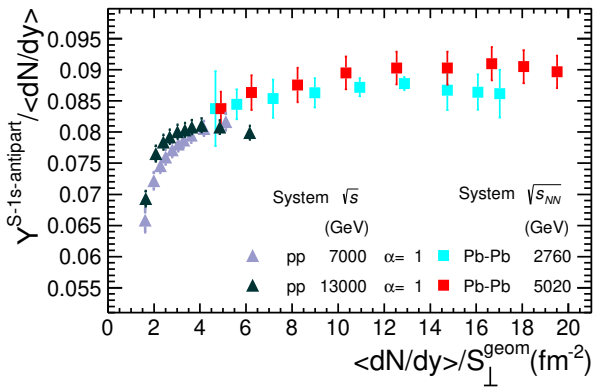



FIG. 25. The ratio of the yield of single-strange anti-hadrons to the total particle multiplicity per unit rapidity as a function of entropy density for Pb-Pb and pp collisions at the LHC energies. Up-triangles correspond to $\alpha = 1$ for estimating the transverse overlap area in pp collisions (see text).

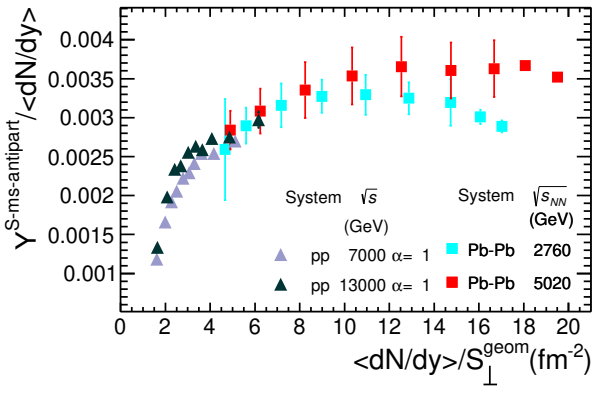



FIG. 26. The ratio of the yield of multi-strange anti-hadrons to the total particle multiplicity per unit rapidity as a function of entropy density for Pb-Pb and pp collisions at the LHC energies. Up-triangles correspond to $\alpha = 1$ for estimating the transverse overlap area in pp collisions (see text).

overlap area in the case of a pp collision as a function of charged particle multiplicity per unit rapidity is required. As in the previous papers [33, 57, 79], the transverse overlap area for pp collisions, $S_{\perp}^{pp}=\pi\ r_{max}^2$ as a function of charged particle multiplicity was estimated using the result of [83], computed in the IP-Glasma model. In this approach $r_{max}$ is the maximal radius for which the energy density of the Yang-Mills fields is above $\varepsilon = \alpha\Lambda_{QCD}^4$ ($\alpha \in [1, 10]$). The model results were fitted for $\alpha$ =1, with:

$$f_{pp} = \begin{cases} 0.387 + 0.0335x + 0.274x^2 - 0.0542x^3 & \text{if } x < 3.4 \\ 1.538 & \text{if } x \geq 3.4 \end{cases} \quad (4)$$

where x=$(dN/dy)^{1/3}$ and $r_{max}$=1fm$\cdot f_{pp}(x)$. In Figure 24 is presented the dependence of $\langle pT\rangle/\sqrt{\langle dN/dy\rangle/S_{\perp}}$ on the charged particle multiplicity per unit pseudorapidity ($\langle dN_{ch}/d\eta\rangle$) for $\pi$, K, $K_S^0$, p, $\bar{\Lambda}$, $\Xi$ and $\Omega$ for pp collisions at $\sqrt{s}$ =7 TeV and $\sqrt{s}$ = 13 TeV. A clear decrease as a function of $\langle dN_{ch}/d\eta\rangle$ is observed. The transverse overlap area was considered to be the same for all species, corresponding to $\alpha = 1$ ( Eq. 4). Qualitatively the trend is similar with the one corresponding to A-A collisions (see Figure 1 of the present paper and Figure 1 of Ref. [10]), supporting the model predictions [35, 36].

As far as in the case of pp collisions at LHC energies the production of pairs of quarks and antiquarks is the main process, the difference in the representation of the $Y^S/\langle dN/dy\rangle$ - $\langle dN/dy\rangle/S_{\perp}$ correlations between strange hadrons and strange anti-hadrons is minor. Plotting the results for $Y^{S-1s}/\langle dN/dy\rangle$ and $Y^{S-ms}/\langle dN/dy\rangle$ for anti-hadrons for pp collisions along with the corresponding data for Pb-Pb at LHC energies, Figures 25 and 26 are obtained. The results corresponding to $\alpha = 1$ are represented by up triangles. As could be seen in Figures 25 and 26, for both single- and multi-strange anti-hadrons the pp data using $\alpha = 1$ for estimating $S_{\perp}$ are laying on top of the extrapolation of Pb-Pb results towards smaller values of the entropy density.

## VI. CONCLUSIONS

The present paper is focused on strange hadron production in heavy ion collisions and similarities between pp and Pb-Pb at LHC energies. This study is based on the impressive amount of experimental information obtained in A-A collisions starting from AGS up to LHC energies. For strange hadrons, the observed trends in the ratio between the average transverse momentum and the square root of the total particle multiplicity per unit rapidity and unit transverse overlap area $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_{\perp}}$ as a function of collision energy for a given centrality or as a function of centrality for a given collision energy are similar with the results for $\pi$ , K and p presented in a previous study, supporting the predictions of CGC and percolation based approaches. The slope and offset of the $\langle p_T\rangle$ dependence on the particle mass and the parameters of BGBW simultaneous fits of the $p_T$ spectra for strange and multi-strange hadrons namely, the average transverse expansion velocity, $\langle\beta_T\rangle$ and the kinetic freeze-out temperature, $T_{kin}$, are extracted and compared to the the results of the same analysis for $\pi$ , K and p. Represented as a function of $\sqrt{\langle dN/dy\rangle/S_{\perp}}$, the larger slope and $\langle\beta_T\rangle$ values for $\pi$, K and p relative to the ones corresponding to strange- and multi-strange hadrons and the smaller offset and $T_{kin}$ values for $\pi$, K and p relative to those corresponding to strange and multi-strange hadrons, support the dynamical freeze out scenario. $K_S^0$, $\Lambda$, $\Xi^-$ and $\Omega^-$ suffer negligible interaction with the hadron environment conserving the momentum distribution corresponding to the their formation time, while the $\pi^+$ , $K^+$ and p build up extra expansion via elastic interaction with the hadronic medium which cools down due to expansion.

The detailed study of the entropy density dependence of the ratio of strange hadron yields per unit rapidity to the total particle multiplicity per unit rapidity at different collision energies and centralities reveals the necessity to study separately single-, multi- strange hadrons and anti-hadrons, in the multi-strange case the sensitivity to the composition of the fireball being higher. The ratio of the strange anti-hadron yields per unit rapidity to the total particle multiplicity per unit rapidity as a function of entropy density studied as a function of the fireball size shows a similarity with the correlation between the ratio of the energy density to the entropy density as a function of entropy density [10]. A kink type behaviour in the ratio of the single-, multi- strange hadrons, $K^+$, $\Lambda$ and $\Xi^-$ yields per unit rapidity to the total particle multiplicity per unit rapidity and in the difference between particles and anti-particles, showing up in the region where a transition from a baryon-dominated gas to a meson-dominated one takes place, is evidenced in agreement with the expectations based on a thermal

model [72]. In this energy range the baryon transparency sets-in and a hot mesonic fireball starts to be formed by particle production in collisions as the precursor of deconfinement. The maximum in $Y^{S-1s-part}/\langle dN/dy\rangle$ and $Y^{S-ms-part}/\langle dN/dy\rangle$ as a function of entropy density starts to be evidenced going from central to peripheral collisions. While the entropy density corresponding to the maximum decreases from central to peripheral collisions, the corresponding value of the collision energy is slightly increasing. Within the experimental error bars, the position of this maximum does not depend on the mass of the corresponding strange hadrons. Similar with A-A collisions, a clear decrease of $\langle p_T\rangle/\sqrt{\langle dN/dy\rangle/S_\perp}$ with $\langle dN_{ch}/d\eta\rangle$ for $\pi$, K, $K^0_S$, p, $\Lambda$, $\Xi$ and $\Omega$ for pp collisions at $\sqrt{s}$ = 7 TeV and $\sqrt{s}$ = 13 TeV is evidenced supporting the expectations based on CGC and percolation models. In the comparison between Pb-Pb and pp for single- and multi- strange anti-hadrons, the pp results nicely follow the extrapolation of Pb-Pb data at lower entropy density.

## VII. ACKNOWLEDGEMENTS

The work was carried out under the contracts sponsored by the Ministry of Research, Innovation and Digitization: RONIPALICE-07/03.01.2022 (via IFA Coordinating Agency) and PN-23 21 01 03.

[1] G. F. Chapline and A. K. Kerman, MIT report CTP-695 (1978).
[2] J. C. Collins and M. J. Perry, Phys. Rev. Lett. **34**, 1353 (1975).
[3] N. Cabibbo and G. Parisi, Phys. Lett. B **59**, 67 (1975).
[4] E. V. Shuryak, Phys. Lett. B **78**, 150 (1978).
[5] G. Baym, Physica A **96**, 131 (1979).
[6] T. Çelik, F. Karsch, and H. Satz, Phys. Lett. B **97**, 128 (1980).
[7] L. Van Hove, Phys. Lett. B **118**, 138 (1982).
[8] J. P. Blaizot and J. Y. Ollitrault, Nucl. Phys. A **458**, 745 (1986).
[9] J. P. Blaizot and J. Y. Ollitrault, Phys. Lett. B **191**, 21 (1987).
[10] M. Petrovici and A. Pop, Phys. Rev. C **107**, 034913 (2023).
[11] M. Petrovici, A. Lindner, and A. Pop, Phys. Rev. C **103**, 034903 (2021).
[12] J. Liao and E. V. Shuryak, Phys. Rev. Lett. **102**, 202302 (2009).
[13] K. M. Burke, A. Buzzatti, N. Chang, C. Gale, M. Gyulassy, U. Heinz, S. Jeon, A. Majumder, B. Müller, G. Qin, *et al.* (JET Collaboration), Phys. Rev. C **90**, 014909 (2014).
[14] S. Shi, J. Liao, and M. Gyulassy, Chinese Phys. C **43**, 044101 (2019).
[15] J. Rafelski and B. Müller, Phys. Rev. Lett. **48**, 1066 (1982).
[16] F. Becattini, M. Gaździcki, A. Keränen, J. Manninen, and R. Stock, Phys. Rev. C **69**, 024905 (2004).
[17] P. Bożek, Acta Phys. Pol. B **36**, 3071 (2005).
[18] K. Werner, Phys. Rev. Lett. **98**, 152301 (2007).
[19] F. Becattini and J. Manninen, Jour. Phys. G: Nucl. Part. Phys. **35**, 104013 (2008).
[20] F. Becattini and J. Manninen, Phys. Lett. B **673**, 19 (2009).
[21] J. Aichelin and K. Werner, Phys. Rev. C **79**, 064907 (2009).
[22] P. Bożek, Phys. Rev. C **79**, 054901 (2009).
[23] J. Aichelin and K. Werner, Phys. Rev. C **81**, 029902(E) (2010).
[24] J. Aichelin and K. Werner, J. Phys. G **37**, 094006 (2010).
[25] J. Steinheimer and M. Bleicher, Phys. Rev. C **84**, 024905 (2011).
[26] C. Schreiber, K. Werner, and J. Aichelin, *Exploring Fundamental Issues in Nuclear Physics* (World Scientific Publishing, 2012) pp. 252–261.
[27] M. Gemard and J. Aichelin, Astronom. Nachr. **335**, 660 (2014).
[28] M. Petrovici, I. Berceanu, A. Pop, M. Tarzila, and C. Andrei, Phys. Rev. C **96**, 014908 (2017).
[29] K. Werner, EPJ Web Conf **171**, 09002 (2018).
[30] J. Adam *et al.* (ALICE Collaboration), Nature **13**, 535 (2017).
[31] K. Werner, Phys. Rev. C **109**, 014910 (2024).
[32] M. Petrovici and A. Pop, EuNPC2022, October 24-28, 2022, University of Santiago de Compostela, Spain, https://indico.cern.ch/event/1104299 /contributions/5055299/attachments/2536779/4366087/EuNPC talk mp.pdf.
[33] M. Petrovici, A. Lindner, and A. Pop, Phys. Rev. C **98**, 024904 (2018).
[34] Y. L. Dokshitzer, V. A. Khoze, and S. I. Troyan, J. Phys. G: Nucl. Part. Phys. **17**, 1585 (1991).
[35] E. Levin and A. H. Rezaeian, Phys. Rev. D **83**, 114001 (2011).
[36] T. Lappi, Eur. Phys. J. C **71**, 1699 (2011).
[37] R. J. Glauber, Phys. Rev. **100**, 242 (1955).
[38] V. Franco and R. J. Glauber, Phys. Rev. **142**, 1195 (1966).
[39] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, Ann. Rev. Nucl. Part. Sci. **57**, 205 (2007).
[40] M. Rybczyński, G. Stefanek, W. Broniowski, and P. Bożek, Comput. Phys. Commun. **185**, 1759 (2014).
[41] S. Chatterjee, S. Das, L. Kumar, D. Mishra, B. Mohanty, R. Sahoo, and N. Sharma, Advances in High Energy Physics **2015**, 349013 (2015).
[42] P. Chung, N. N. Ajitanand, J. Alexander, M. Anderson, D. Best, F. P. Brady, T. Case, W. Caskey, D. Cebra, J. L. Chance, *et al.*, Phys. Rev. Lett. **91**, 202301 (2003).
[43] L. Adamczyk *et al.* (STAR Collaboration), Phys. Rev. C **96**, 044904 (2017).
[44] J. Adam *et al.* (STAR Collaboration), Phys. Rev. C **102**, 034909 (2020).
[45] B. I. Abelev *et al.* (STAR Collaboration), Phys. Rev. C **77**, 044908 (2008).
[46] M. M. Aggarwal *et al.* (STAR Collaboration), Phys. Rev. C **83**, 024901 (2011).

[47] J. Adams *et al.* (STAR Collaboration), Phys. Rev. Lett. **98**, 062301 (2007).
[48] G. Agakishiev *et al.* (STAR Collaboration), Phys. Rev. Lett. **108**, 072301 (2012).
[49] B. Abelev *et al.* (ALICE Collaboration), Phys. Rev. C **88**, 044910 (2013).
[50] B. Abelev *et al.* (ALICE Collaboration), Phys. Rev. Lett. **111**, 222301 (2013).
[51] B. Abelev *et al.* (ALICE Collaboration), Phys. Lett. B **728**, 216 (2014); **734**, 409 (2014).
[52] S. Acharya *et al.* (ALICE Collaboration), Phys. Rev. C **101**, 044907 (2020).
[53] D. S. de Albuquerque, Ph.D. thesis (2019), CERN-THESIS-2019-135.
[54] P. Kalinak for the ALICE Collaboration, European Physical Society Conference on High Energy Physics, 5-12 July 2017, Venice, Italy, PoS(EPS-HEP2017)168 (2017), https://pos.sissa.it/314/168/pdf.
[55] D. S. de Albuquerque for the ALICE Collaboration, Nucl. Phys. A **982**, 823 (2019), XXVIIth International Conference on Ultrarelativistic Nucleus-Nucleus Collisions (Quark Matter 2018).
[56] A. Balbino, Ph.D. thesis, Politecnico di Torino (2023), https://iris.polito.it/handle/11583/2979899.
[57] A. Lindner, M. Petrovici, and A. Pop, Proceedings of Science (PoS), **380**, 197 (2021), Particles and Nuclei International Conference (PANIC2021), https://pos.sissa.it/380/197/.
[58] T. S. Biro, H. B. Nielsen, and J. Knoll, Nucl. Phys. B **245**, 449 (1984).
[59] J. Dias de Deus and C. Pajares, Phys. Lett. B **695**, 211 (2011).
[60] A. H. Mueller, Nucl. Phys. A **715**, 20c (2003).
[61] J. S. E, Schnedermann and U. Heinz, Phys. Rev. C **48**, 2462 (1993).
[62] J. Rafelski, I. Kuznetsova, and J. Letessier, Jour. Phys. G: Nucl. Part. Phys. **35**, 044011 (2008).
[63] J. Rafelski and M. Petran, Acta Phys. Polonica Supp. **7**, 35 (2017).
[64] P. Koch and J. Rafelski, South Afr. J. Phys. **9**, 8 (1986).
[65] P. Koch, B. Müller, and J. Rafelski, Phys. Rep. **142**, 167 (1986).
[66] P. Castorina, S. Plumari, and H. Satz, Int. J. Mod. Phys. **E25**, 1650058 (2016).
[67] E. I. Bratkovskaya, W. Cassing, A. Palmese, and P. Moreau, Acta Phys. Pol. B **10**, 507 (2017).
[68] J.Adamczewski-Musch *et al.* (HADES Collaboration), Phys. Lett. B **793**, 457 (2019).
[69] M. Abdallah *et al.* (STAR Collaboration), Phys. Lett. B **831**, 137152 (2022).
[70] M. Gaździcki (NA49 Collaboration), Jour. Phys. G Nucl. Part. Phys. **30**, S701 (2004).
[71] C. Alt *et al.* (NA49 Collaboration), Phys. Rev. C **77**, 024903 (2008).
[72] J. Cleymans, H. Oeschler, K. Redlich, and S. Wheaton, Eur. Phys. J A **29**, 119 (2006).
[73] W. Cassing, A. Palmese, P. Moreau, and E. L. Bratkovskaya, Phys. Rev. C **93**, 014902 (2016).
[74] W. Busza and A. Goldhaber, Phys. Lett. B **139**, 235 (1984).
[75] R. Stock, arXiv:hep-ph , 0404125 (2004).
[76] CMS Collaboration, J. High Energ. Phys. **9**, 91 (2010).
[77] C. Andrei, ALICE Collaboration, Nucl. Phys. A **931**, c888 (2014).
[78] CMS Collaboration, Phys. Lett. B **765**, 193 (2017).
[79] M. Petrovici, A. Lindner, and A. Pop, AIP Conference Proceedings, **2076**, 040001 (2019), EXOTIC NUCLEI AND NUCLEAR/PARTICLE ASTROPHYSICS (VII). PHYSICS WITH SMALL ACCELERATORS: Proceedings of the Carpathian Summer School of Physics 2018 (CSSP18).
[80] S. Acharya *et al.* (ALICE Collaboration), Phys. Rev. C **99**, 024906 (2019).
[81] ALICE Collaboration, Eur. Phys. J C **80**, 167 (2020).
[82] ALICE Collaboration, Eur. Phys. J C **80**, 693 (2020).
[83] A. Bzdak, B. Schenke, P. Tribedy, and R. Venugopalan, Phys. Rev. C **87**, 064906 (2013).

# Cosmic Inflation, Dark Energy and Gravitational Waves

K. Dimopoulos[1, *]

[1] *Consortium for Fundamental Physics, Physics Department, Lancaster University, Lancaster LA1 4YB, UK*

(Received xx.xx.2023; Revised xx.xx.2023; Accepted xx.xx.2023)

We briefly discuss cosmic inflation, which is the dominant paradigm for the generation of the large scale structure in the Universe and also for arranging for the initial conditions of the hot Big Bang. We then present quintessential inflation, which also accounts of the observed dark energy. We discuss how quintessential inflation can be successfully modelled in modified gravity in the Palatini formalism. Finally, we focus on the generation of primordial gravitational waves by inflation and how their spectrum can be enhanced when the early Universe goes through periods of stiff equation of state. This results in gravitational waves with a characteristic spectrum, which may well be observed in the near future, providing insights for the background theory.

PACS numbers: Suggested PACS
Keywords: Suggested keywords

arXiv:2402.19086v1 [hep-ph] 29 Feb 2024

## 1. COSMIC INFLATION

The history of the Universe requires special initial conditions, which are arranged by cosmic inflation [1, 2]. In a nutshell, cosmic inflation can be defined as a period of accelerated expansion in the Early Universe [3, 4]. Inflation produces a Universe which is large, uniform and spatially flat according to observations. Typically, inflation is realised via the inflationary paradigm, which states that the Universe inflates when dominated by the potential energy density of a scalar field, called the inflaton field.

The Klein-Gordon equation of motion of a homogeneous scalar field $\phi$ is

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0\,, \tag{1}$$

where $H$ is the rate of the Universe expansion (Hubble parameter), the dot denotes derivative with respect to the cosmic time $t$ and the prime denotes derivative with respect to the field: $' \equiv \partial/\partial\phi$. The above is of the same form as the equation of motion of a ball sliding down a potential under friction determined by $H$ (see Fig. 1). Potential domination, therefore, suggests that the kinetic energy density is subdominant to the potential energy density $V$, and the field slowly rolls (slowly varies in field space) down a potential plateau, called the inflationary plateau. Inflation ends at a characteristic value $\phi_{\rm end}$ when the potential becomes steep and curved. After the end of inflation, the inflaton field oscillates around its vacuum expectation value (VEV). These coherent oscillations amount to inflaton particles, which decay into the primordial plasma, through a process called reheating.

Inflation however, should not make the Universe perfectly uniform, because in order for galaxies to form, initial perturbations in the density of the Universe are needed. Indeed, inflation makes the Universe largely uniform but also introduces minor deviations from uniformity which give rise to the Primordial Density Perturbations (PDPs), which in turn become the seeds for the formation of structures such as galaxies [4]. Inflation does this through the particle productions process which roughly operates as follows:

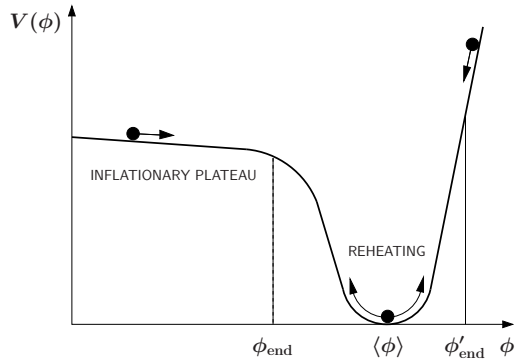



Figure 1. Sketch of the typical inflationary potential. Due to the form of the equation of motion (1), we can envisage the system as a ball rolling down a flat region of the potential, called the inflationary plateau. At some critical value of the inflaton field $\phi_{\rm end}$, the potential becomes steep and curved such that inflation is terminated. Afterwards, the field oscillates around its vacuum expectation value $\langle\phi\rangle$. The figure also depicts the possibility that the inflaton field slow-rolls down a steep potential under excessive friction (and inflation ends when this friction is not enough for inflation at $\phi'_{\rm end}$) but this possibility is not favoured by the observations.

Accelerated expansion of space is superluminal. This superluminal expansion during inflation amplifies the quantum fluctuations of the inflaton field, which become classical perturbations of the field through quantum decoherence. Consequently, inflation continues a little bit more in some locations than in others. Thus, at the end of inflation space expands in a different way in neighbouring locations, which introduces the PDPs (see Fig. 2).

The PDP reflects itself onto the Cosmic Microwave Background radiation (CMB) through the Sachs-

* E-mail: k.dimopoulos1@lancaster.ac.uk

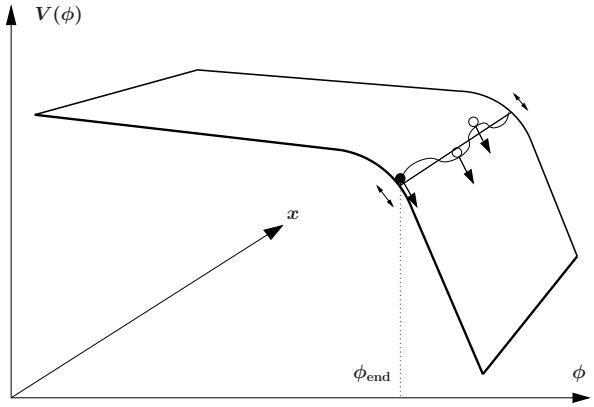



Figure 2. Sketch of the edge of the inflationary plateau, where inflation is terminated at the critical value $\phi_{\rm end}$. The spatial direction $x$ is also depicted (assumed one-dimensional for illustrative purposes). Perturbations of the inflaton field $\phi$ imply that, while it is rolling down its potential, it reaches the critical value $\phi_{\rm end}$ at different times at different locations (values of $x$). This means that inflation continues a little bit more in some locations than in others, which leads to the generation of the primordial density perturbations.

Wolfe effect [5]. Indeed, precise CMB observations have revealed the existence of the PDP at the level of $\sim 10^{-5}$, with the characteristics suggested by inflation (acoustic peaks). The agreement with the observations (see Fig. 3) is spectacular [6].

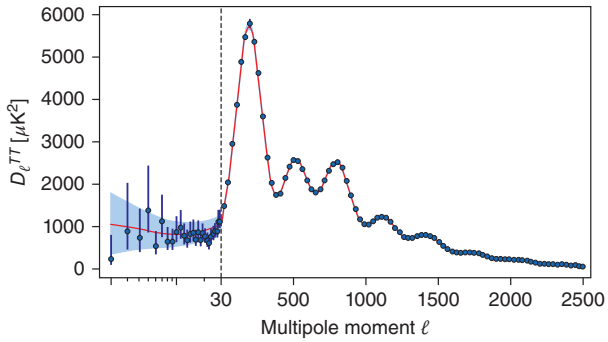



Figure 3. CMB temperature perturbation spectrum. The graph is a plot of $\mathcal{D}_\ell = \ell(\ell+1)C_\ell/2\pi$, where $C_\ell$ is the CMB temperature anisotropy, as a function of the multipole moment $\ell$ of spherical harmonics. For $\ell < 30$ the scaling of the horizontal axis changes. The shaded area depicts the region under the influence of cosmic variance, which suggests that deviations from the theoretical curve (depicted by the solid line) within the shaded area are due to poor statistics and do not have a physical meaning. The effect of cosmic variance diminishes (virtually disappears) for large $\ell$. For $\ell \geq 30$ we have the clear depiction of seven acoustic peaks. The line between the binned data points is not the line which connects the dots; it is the theoretical line predicted by inflation. The agreement with the data is impressive. Figure taken from Ref. [6].

The PDPs are predominately adiabatic, Gaussian and scale invariant [4]. Adiabaticity suggests that they are the product of a single degree of freedom, such as the inflaton field. Gaussianity reflects the randomness of the original quantum fluctuations. Approximate scale invariance suggests that inflation is of quasi-de Sitter type, when the density $\rho$ is roughly constant during inflation. The barotropic (equation of state) parameter of a homogeneous scalar field is

$$w \equiv \frac{p}{\rho} = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}\,, \tag{2}$$

where $p$ is the pressure. According to the inflationary paradigm, the kinetic energy density $\rho_{\rm kin}$ is subdominant to the potential during inflation, which means $\rho_{\rm kin} \equiv \frac{1}{2}\dot{\phi}^2 \ll V$. As a result, during quasi-de Sitter inflation we have $w \approx -1$.

There is a lot of emphasis put on the PDP generated by inflation, because the perturbations generated can discriminate between different inflation models. In particular, two observables are of prime interest: the scalar spectral index $n_s$ and the tensor-to-scalar ratio $r$. For the scalar perturbations, the spectrum can be written as $\mathcal{P}_\zeta(k) \propto k^{n_s(k)-1}$, where $k$ is the momentum scale. A perfectly scale invariant spectrum would correspond to $n_s = 1$, in which case all $k$-dependence of $\mathcal{P}_\zeta$ would disappear. Indeed, observations suggest that, for negligible tensor perturbations, the spectral index is very close to unity $n_s = 0.965 \pm 0.004$ [6]. Crucially, the spectral index is not exactly equal to unity because the end of inflation is near and the inflationary plateau is not exactly flat. Therefore, a slightly red spectrum is expected, exactly as observed. The tensor-to-scalar ratio is $r \equiv \mathcal{P}_\zeta/\mathcal{P}_t$, where $\mathcal{P}_t$ is the spectrum of the tensor perturbations (primordial gravitational waves). Observations suggests that there is a stringent upper limit $r < 0.036$ [7], which can be even tighter under certain conditions. Precision observations of $n_s$ and $r$ have already resulted in the exclusion of many, otherwise well motivated inflation models (see Fig. 4).

### 1.1. $R^2$ inflation

The seminal paper of Alexei A. Starobinsky, which appeared in 1980, even before the name 'cosmic inflation' was coined, introduced the very first and the most successful to date inflation model [2]. It is a modified gravity theory with a simple Lagrangian density

$$\mathcal{L} = \frac{1}{2}m_P^2 R + \alpha R^2\,, \tag{3}$$

where $m_P$ is the reduced Planck mass and $R$ is the scalar curvature (Ricci scalar). The first term in the above corresponds to the standard Einstein-Hilbert action. However, the second term, whose importance is parametrised by the non-perturbative coefficient $\alpha$, corresponds to modified gravity. This quadratic gravity term introduces an additional degree of freedom. The latter can be flushed out if we switch from the modified gravity frame (called the Jordan

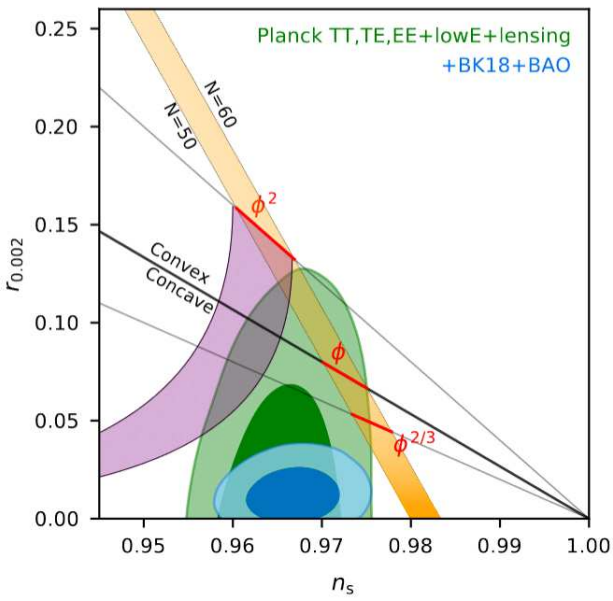



Figure 4. Observations of BICEP/Keck Collaboration plot the tensor to scalar ratio $r$ with respect to the scalar spectral index $n_s$. The allowed parameter space is shown in blue (contours depict the 95% c.l. [light] and 67% c.l. [dark]). Observations exclude families of otherwise motivated inflation models. For example, the clearly excluded purple band and orange band correspond to natural and chaotic monomial inflation models respectively. The figure was taken from Ref. [7].

frame) to the frame of Einstein gravity (called the Einstein frame) through a conformal transformation of the spacetime metric, of the form $g_{\mu\nu} \to \Omega^2 g_{\mu\nu}$, where the conformal factor for this theory is $\Omega^2 = 1 + \frac{4\alpha}{m_P^2} R$. In the Einstein frame the Lagrangian density becomes

$$\mathcal{L} = \frac{1}{2} m_P^2 R + \frac{1}{2}(\partial\phi)^2 - V(\phi)\,, \qquad (4)$$

where $R$ is now calculated using the new metric and, apart from the Einstein-Hilbert term, $\mathcal{L}$ features a minimally coupled (to gravity), canonically normalised scalar field $\phi$, where $(\partial\phi)^2 \equiv -\partial_\mu\phi\,\partial^\mu\phi$. This scalar field corresponds to the new degree of freedom introduced by the original quadratic gravity term, and it is called the scalaron field with $\Omega^2 = \exp\left(\sqrt{\frac{2}{3}}\,\phi/m_P\right)$.

For this theory, in the Einstein frame, the scalar potential is

$$V(\phi) = \frac{m_P^4}{16\alpha}\left(1 - e^{-\sqrt{\frac{2}{3}}\,\phi/m_P}\right)^2 . \qquad (5)$$

The form of the above potential is shown in Fig. 5. The model predicts

$$n_s = 1 - \frac{2}{N} \quad \text{and} \quad r = \frac{12}{N^2}\,, \qquad (6)$$

where $N$ is the number of e-folds (exponential expansions) of remaining inflation when the cosmological scales exit the horizon during inflation (they are pushed out by the superluminal expansion). Typically, $N = 50 - 60$ depending on the details of reheating (prompt reheating results in $N = 60$). The above suggest that the Starobinsky inflation predictions are in the sweet spot of the observations. For the amplitude of the PDP to match the CMB observations we require $\alpha = 5.5225 \times 10^8$.

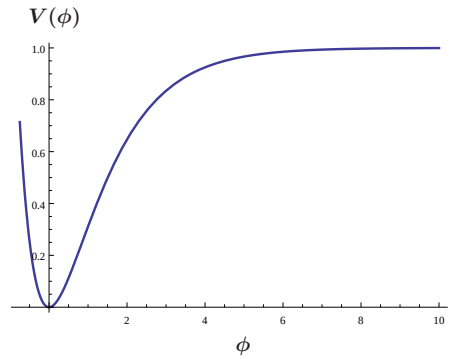



Figure 5. Form of the Starobinsky potential, which applies to both the cases of $R^2$-inflation and Higgs inflation. The plot clearly depicts the inflationary plateau at large values of the inflaton field. At small values, the potential becomes steep and curved and the inflaton oscillates around its VEV, which is zero in this case. Fiducial units are used.

### 1.2. Higgs inflation

Another seminal inflation model was put forward by Fedor L.Bezrukov and Mikhail Shaposhnikov in 2008 [8]. The model considers the electroweak Higgs field $h$ (first observed in CERN in 2012) as the inflaton. It is another modified gravity theory with Lagrangian density

$$\mathcal{L} = \frac{1}{2} m_P^2 R + \frac{1}{2}\xi R h^2 + \frac{1}{2}(\partial h)^2 - U(h)\,, \qquad (7)$$

where $\xi$ parametrises the strength of the non-minimal coupling to gravity of the Higgs field, which is naturally introduced by quantum corrections in curved spacetime. The potential is $U(h) = \frac{1}{4}\lambda_h(h^2 - v^2)^2$, where $v = 246\,\text{GeV}$ is the Higgs VEV and $\lambda_h$ is its self-coupling, which at low energies is $\lambda_h = 0.129$.

We switch to Einstein gravity using the conformal transformation $\Omega^2 = 1 + \frac{\xi h^2}{m_P^2}$. In the Einstein frame the Lagrangian density is given by Eq. (4) with potential

$$V(\phi) = \frac{\lambda_h m_P^4}{4\xi^2}\left(1 - e^{-\sqrt{\frac{2}{3}}\,\phi/m_P}\right)^2 . \qquad (8)$$

The above is of exactly the same form as the potential in Eq. (5) if we identify $\xi^2 = 4\alpha\lambda_h$. Thus, the inflationary predictions are equally successful. In fact, because the branching rations of the Higgs decays into the standard model particles are known, reheating is unambiguously defined and $N \simeq 57$. As a result, Eq. (6)

suggests $n_s = 0.965$ and $r = 0.0037$, which are in excellent agreement with observations. To have the correct amplitude of the PDP we require $\xi \simeq 47000/\sqrt{\lambda_h}$.

## 2. DARK ENERGY

Observations suggest that the late Universe is also undergoing accelerated expansion. This is attributed to an exotic substance called dark energy, which makes up almost 70% of the Universe content at present [9]. Dark energy could correspond to a positive cosmological constant $\Lambda$ but this requires phenomenal fine-tuning (of order $10^{-120}$) which has been called "the worse fine-tuning in physics" (Lawrence Krauss). As a result, other proposals have been put forward. A prominent such suggestion is quintessence [10], which amounts to the fifth element after baryonic (normal) matter, dark matter, photons (mainly CMB) and neutrinos. Quintessence is a scalar field like the inflaton, typically slow-rolling down a runaway potential; a feature called the quintessential tail. Thus, the dark energy observations suggest that the Universe is undergoing a late inflation period, driven by quintessence. This late inflation is also quasi-de Sitter, with barotropic parameter $-1 \leq w \leq -0.95$ [6].

Being a dynamical degree of freedom, quintessence introduces an additional requirement (compared with just $\Lambda$), that of its initial conditions. Indeed, in thawing quintessence the field is initially frozen only to unfreeze near the present time to become dominant. But the initial (frozen) value of the field must be such that the potential energy density of the field today is comparable ($\sim 70\%$) to the density of the Universe at present. This is called the coincidence requirement [9].

## 3. QUINTESSENTIAL INFLATION

A promising way to overcome the coincidence requirement is the idea of quintessential inflation, proposed first by P. James E. Peebles and Alexander Vilenkin in 1999 [11]. In a nutshell, quintessential inflation is the identification of quintessence with the inflaton field.

Quintessential inflation is a natural idea, as quintessence and the inflaton are both scalar fields. It is also economic because one employs only a single degree of freedom. Furthermore, it allows to treat the physics of inflation in the early Universe and dark energy in the late Universe in a common theoretical framework, valid over a vast range of energies. Quintessential inflation has to satisfy the observations both cosmic inflation and of dark energy, which is very difficult but not impossible. Finally, the initial conditions of quintessence (the coincidence requirement) are fixed by the inflationary attractor.

The potential of quintessential inflation models typically features two flat regions, the inflationary plateau and the quintessential tail [3] (see Fig. 6). In contrast to the traditional inflationary paradigm, the inflaton field does not decay after the end of inflation, because it has to survive until the present time and become quintessence. Consequently, reheating has to occur without the decay of the inflaton field. Fortunately, there are many motivated mechanisms that achieve this.

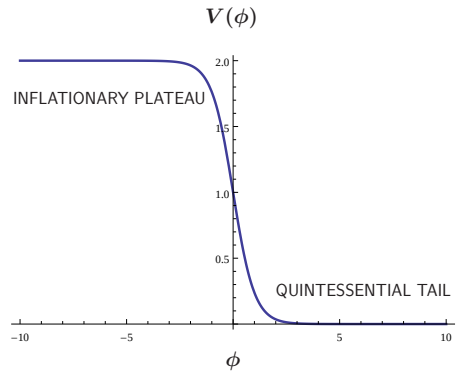



Figure 6. The typical potential of quintessential inflation. It depicts a runaway direction with VEV displaced at infinity. There are two flat regions, the inflationary plateau and the quintessential tail, chosen (without loss of generality) to be at small and large values of the scalar field.

### 3.1. Quintessential inflation in Palatini gravity

A promising framework for the construction of quintessential inflation models is modified gravity on the Palatini formalism [12]. The Palatini formalism considers both the spacetime metric and the connection to be independent degrees of freedom, in contrast to the metric formalism, where the connection is the Levi-Civita. With the Einstein-Hilbert action, both formalisms result in general relativity. However, modified gravity actions result in different theories when assuming the metric or the Palatini formalism. Here we will argue that the Palatini formalism gives rise naturally to a potential suitable for quintessential inflation.

We consider the Lagrangian density in the Jordan frame

$$\mathcal{L} = \frac{1}{2} m_P^2 R + \alpha R^2 + \frac{1}{2}(\partial\varphi)^2 - V(\varphi)\,, \qquad (9)$$

where we considered a quadratic gravity term, as in Starobinsky inflation discussed in Sec. 3, and we have explicitly introduced a canonically normalised scalar field $\varphi$. We did this because, in the Palatini formalism, the quadratic gravity term $R^2$ does not introduce an additional degree of freedom (there is no scalaron).

Using a suitable conformal transformation, we switch to the Einstein frame, where the Lagrangian density is

$$\mathcal{L} = \frac{1}{2} m_P^2 R + \frac{\frac{1}{2}(\partial\varphi)^2}{1 + 16\alpha\frac{V(\varphi)}{m_P^4}} - \frac{V(\varphi)}{1 + 16\alpha\frac{V(\varphi)}{m_P^4}}\,. \qquad (10)$$

We see that in the Einstein frame the scalar field has a non-canonical kinetic term. Notice however, that, when $V$ is very large, the unity in the denominator is negligible and the potential density approaches the constant value: $m_P^4/16\alpha$, This corresponds to the inflationary plateau, as the potential energy density does not depend of the value of the inflaton field, i.e. it is the same for a range of $\varphi$ values.

In the opposite limit, when $V$ is very small then the denominator becomes unity and the field becomes canonically normalised. Thus, we only need a suitable runaway potential to arrange for the quintessential tail, while the Palatini setup generates the inflationary plateau by "flattening" the potential at high energies.

We apply this finding in the following setup. Now the Lagrangian density in the Jordan frame is [13]

$$\mathcal{L} = \frac{1}{2}m_P^2 R + \alpha R^2 + \frac{1}{2}\xi R\varphi^2 + \frac{1}{2}(\partial\varphi)^2 - V(\varphi)\,, \quad (11)$$

where we also consider a non-minimal coupling of $\varphi$ to gravity, which is expected naturally by quantum corrections in curved spacetime, as we mention in Sec. 1.2.

After performing a conformal transformation we bring the theory in the Einstein frame. We also redefine a canonically normalised inflaton field $\phi$ by using the relation

$$\left(\frac{\mathrm{d}\phi}{\mathrm{d}\varphi}\right)^2 = \frac{1+\xi\varphi^2/m_P^2}{(1+\xi\varphi^2/m_P^2)^2 + \frac{16\alpha}{m_P^4}V(\varphi)}\,. \quad (12)$$

Then, the Lagrangian density becomes

$$\mathcal{L} = \frac{1}{2}m_P^2 R + \frac{1}{2}(\partial\phi)^2 - U(\phi)\,, \quad (13)$$

where the potential is

$$U(\phi) = \frac{V(\varphi(\phi))}{[1+\xi\varphi^2(\phi)/m_P^2]^2 + \frac{16\alpha}{m_P^4}V(\varphi(\phi))}\,. \quad (14)$$

We see that, again, if $V$ is very large, we approach the inflationary plateau with $U \simeq m_P^4/16\alpha$. We only need a runaway potential. We choose a simple exponential potential [13], which is amply motivated by fundamental theory:

$$V(\varphi) = V_0\, e^{-\kappa\varphi/m_P}\,, \quad (15)$$

where $\kappa$ is the strength of the exponential. As we approach the quintessential limit the potential of the quintessential tail approaches the form [13]

$$U(\phi) \simeq V_0 \frac{\exp\left[-\frac{\kappa}{\sqrt{\xi}}\sinh\left(\sqrt{\xi}\frac{\phi}{m_P}\right)\right]}{\cosh^4\left(\sqrt{\xi}\frac{\phi}{m_P}\right)}\,, \quad (16)$$

where we used the fact that, at late times, the quadratic gravity term is negligible, which means that Eq. (12) can be solved exactly to give $\sqrt{\xi}\varphi = m_P \sinh\left(\sqrt{\xi}\phi/m_P\right)$. In the limit $\sqrt{\xi}\varphi \ll m_P$, we recover the classic exponential quintessential tail $U(\phi) \simeq V_0\, e^{-\kappa\phi/m_P}$.

## 4. KINATION

The Universe history in quintessential inflation typically includes a period, just after inflation, where the Universe is dominated by the kinetic energy density of the rolling scalar field $\rho_{\rm kin} = \frac{1}{2}\dot\phi^2$. This period is called kination [14] and usually follows the end of inflation until reheating, when the radiation era of the hot Big Bang begins.

Indeed, as the inflaton rolls off the inflationary plateau, it plunges down the potential cliff (see Fig. 6) and becomes kinetically dominated, i.e. $V \ll \rho_{\rm kin}$. Because the inflaton continues to dominate the Universe, according to Eq. (2), the barotropic barotropic parameter of the Universe is $w \approx 1$. The Klein-Gordon in Eq. (1) becomes oblivious of the potential: $\ddot\phi + 3H\dot\phi \simeq 0$, whose solution suggests that $\rho = \rho_{\rm kin} \propto a^{-6}$, where $a(t)$ is the scale factor of the Universe. During kination the field rolls down to the quintessential tail of its potential.

Because the energy density of radiation is diluted less efficiently with the Universe expansion, $\rho_r \propto a^{-4}$, if some mechanism generates some (initially subdominant) radiation at the end of inflation (or afterwards), this radiation eventually catches up with the kinetic energy density of the rolling scalar field and the Universe becomes radiation dominated. This is the moment of reheating $\rho_{\rm kin}^{\rm reh} = \rho_r^{\rm reh}$, the onset of the hot Big Bang.

After the radiation era begins, it can be shown that the field's roll is impeded and soon the field freezes in some value $\phi_F$ (see Fig. 7). Because the roll of the scalar field is oblivious of the potential during kination and afterwards until it freezes, it can be studied in a model independent way. In particular, the total excursion of the scalar field in field space from the end of inflation and until it freezes is [15]

$$\Delta\phi_F = \sqrt{\frac{2}{3}}\left(2 - \frac{3}{2}\ln\Omega_r^{\rm end}\right) m_P\,, \quad (17)$$

where $\Omega_r^{\rm end} = (\rho_r/\rho)_{\rm end}$ is the density parameter of radiation at the end of inflation, sometimes called reheating efficiency.[1] Thus, we see that the smaller the reheating efficiency $\Omega_r^{\rm end}$ is the largest the roll of the scalar field after inflation and until it freezes and the longer the corresponding kination period is. The significance of this is discussed next.

## 5. STIFF PERIOD AND GRAVITATIONAL WAVES

In the inflationary paradigm, inflation generates an almost scale invariant superhorizon spectrum of tensor perturbations (gravitational waves) [16]. This is

[1] Here, we assume that the mechanism which generates the radiation bath operates at the end of inflation, as it is typically the case.

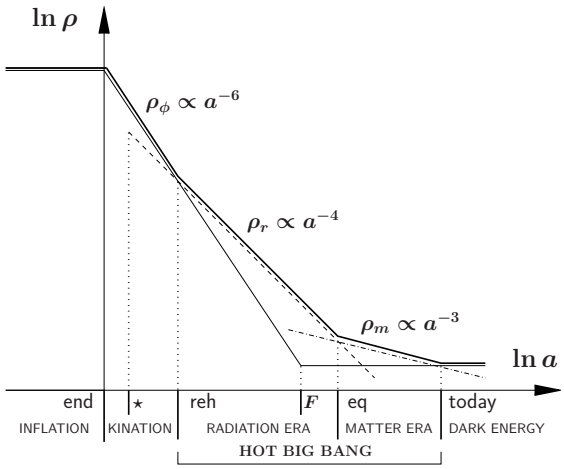



Figure 7. Schematic log-log plot of the density of the Universe and its individual components with respect to the scale factor according to quintessential inflation. The thin solid line depicts the density of the scalar field, which is roughly constant during inflation (i.e. until inflation ends, denoted by 'end') and after freezing (denoted by '$F$') but scales as $\rho_\phi \propto a^{-6}$ in between, regardless of whether it is dominant or not. The dashed line depicts the density of radiation (formed at the moment denoted by '$\star$', but it is identified with 'end' in most cases), which scales as $\rho_r \propto a^{-4}$. Radiation dominates at reheating (denoted by 'reh'), which marks the onset of the hot Big Bang. The dot-dashed line depicts the density of matter, which scales as $\rho_m \propto a^{-3}$ and takes over from radiation at the time of equality (denoted by 'eq'). Today is the end of the hot Big Bang, when the density of the scalar field takes over again accounting for dark energy. The thick solid line depicts the total density.

so only for the scales which re-enter the horizon during the radiation era of the hot Big Bang, because the density ratio $\rho_{\rm GW}/\rho$ is constant, as the energy density of the gravitational waves (GWs) scales as radiation $\rho_{\rm GW} \propto a^{-4}$ and $\rho = \rho_r \propto a^{-4}$ during the radiation era. However, as we have seen, during kination $\rho = \rho_{\rm kin} \propto a^{-6}$, which means that the ratio $\rho_{\rm GW}/\rho$ is not constant any more. In general, this would be true for any scales which re-enter the horizon during a period with stiff equation of state, such that $\frac{1}{3} < w \leq 1$, as $\rho \propto a^{-3(1+w)}$.

The density parameter of GWs per logarithmic frequency interval is $\Omega_{\rm GW}(f) \equiv \frac{{\rm d}\Omega_{\rm GW}}{{\rm d}\ln f}$. Then, it can be shown that the GW spectrum is of the form [17]

$$\Omega_{\rm GW}(f) \propto f^{2(\frac{3w-1}{3w+1})}, \qquad (18)$$

where $w$ is the barotropic parameter at the time when the scales in question re-enter the horizon. For the scales, which do so during the radiation era where $w = \frac{1}{3}$ we obtain $\Omega_{\rm GW}(f) = $ constant, which results in a scale invariant spectrum (independent of $f$). If there is a period of kination with $w = 1$, then the above suggests $\Omega_{\rm GW} \propto f$ and we have a peak in the spectrum, at the time when kination begins (usually, the end of inflation; see Fig. 7). The longer kination is the larger the peak in the GW spectrum. This peak cannot be arbitrarily large however, because too much GWs can disturb the process of Big Bang Nucleosynthesis (BBN), as we explain below.

### 5.1. Kination and GWs

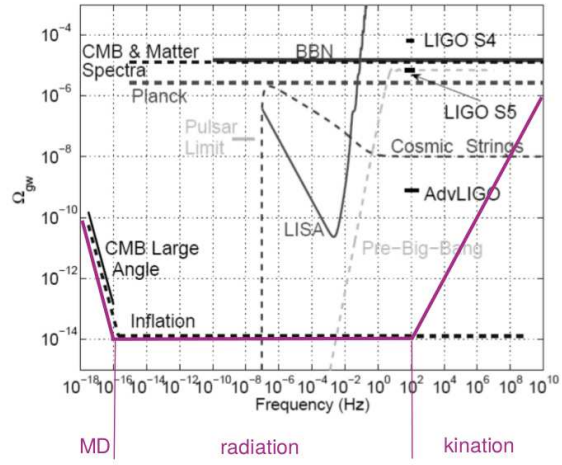



Figure 8. Plot of the GW spectrum per logarithmic frequency interval $\Omega_{\rm GW}$ with respect to frequency, up to the largest possible frequency which corresponds to the inflation energy scale ($f_{\rm end} \sim 10^{10}$ Hz) assumed at the scale of grand unification $\rho_{\rm end}^{1/4} \sim 10^{15-16}$ GeV, as is typically the case. The solid purple line depicts the spectrum of primordial gravitational waves in the case of kination following right after inflation, with barotropic parameter $w = 1$ and reheating temperature $T_{\rm reh} \sim 10^2$ MeV. The peak in the GW spectrum almost saturates the BBN bound $\Omega_{\rm GW} < 10^{-6}$, depicted by the horizontal solid black line. In the figure, the expected observational capability of Advanced LIGO [22] and LISA [21] are shown. It is evident that this scenario produces unobservable (by LISA and Advanced LIGO) primordial GW. Observational bounds have been taken from Ref. [18].

The GW energy density has to be at most 1% the the time of BBN, i.e. $\Omega_{\rm GW}^{\rm BBN} < 10^{-2}$ [19]. Now, as we have said, the energy density of GWs and that of radiation stay at a constant ratio. This means that there is an upper bound on the GW energy density today, which can be computed as

$$\Omega_{\rm GW}^0 = \frac{\rho_{\rm GW}^0}{\rho^0} = \left.\frac{\rho_{\rm GW}}{\rho_r}\right|_0 \left.\frac{\rho_r}{\rho}\right|_0 = \left.\frac{\rho_{\rm GW}}{\rho_r}\right|_{\rm BBN} \Omega_r^0 = \Omega_{\rm GW}^{\rm BBN}\,\Omega_r^0 < 10^{-6}, \qquad (19)$$

where we used that $\Omega_r^0 = (\rho_r/\rho)_0 \simeq 10^{-4}$ and that $\rho_{\rm GW}/\rho_r = $ constant.

This bound means that kination cannot be arbitrary extended to late times (by considering a small

reheating efficiency) because the peak in the GW spectrum would become too large and violate the bound in Eq. (19). Consequently, kination can occur only at very early times, which translates to high frequencies. A peak in the GW spectrum cannot be extended to lower frequencies and come in contact with future observations, e.g. of the LISA space interferometer. This can be understood better in Fig. 8.

### 5.2. GWs from a stiff period with $w \simeq \frac{1}{2}$

One way to avoid violating the bound in Eq. (19) and still extend the GW peak down to observable frequencies is to consider that, the stiff period after inflation is not as stiff as kination proper with $w = 1$. Indeed, in Ref. [20] it was shown that contact with the LISA observations [21] can be achieved if $0.46 \lesssim w \lesssim 0.56$ with low reheating temperature in the range $1\ \text{MeV} < T_{\rm reh} \lesssim 150\ \text{MeV}$. A realisation of this possibility was recently discussed in Ref. [23].

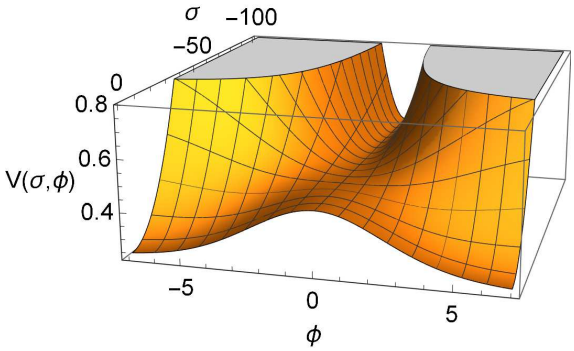



Figure 9. Qualitative form of the scalar potential in Eq. (22) in terms of the canonical scalar fields $\sigma$ (primordial inflaton) and $\phi$ (waterfall). The axes are in fiducial units. The system origitally finds itself gradually rolling down $V(\sigma)$ inside the steep valey at $\phi = 0$ when $\sigma$ is large. When the inflaton is reduced below the critical value $\sigma_c$, a phase transition sends the waterfall field away from the origin. The minima along the canonical waterfall direction are displaced at infinity. As a result, after the phase transition, there is an initial period of fast-roll hilltop inflation along the waterfall direction, followed by a stiff period. Figure taken by Ref. [23].

Consider two flat directions meeting at an enhanced symmetry point (ESP). The scalar potential can be written as

$$V(\varphi,\sigma) = \frac{1}{2}g^2\sigma^2\varphi^2 + \frac{1}{4}\lambda(\varphi^2 - M^2)^2 + V(\sigma)\,, \quad (20)$$

where the strength of the interaction is parametrised by the perturbative coupling constant $g < 1$ (we assume that the ESP is at the origin), $\lambda$ is the self-coupling of the $\varphi$-field, whose VEV is $M$ and $V(\sigma)$ is the potential along the $\sigma$-direction. As we discuss below, $M \sim m_P$, so $\varphi$ is a flat direction lifted by Planck-suppressed operators. The above potential in Eq. (20) is the standard perturbative potential in hybrid inflation [24], where $\sigma$ plays the role of the primordial inflaton field. As in standard hybrid inflation, during primordial inflation the interaction term sends the waterfall field $\varphi$ to zero, while $V(\sigma)$ provides a gentle slope which allows the inflaton to slow-roll towards the origin. Primordial inflation is terminated by a phase transition, when $\sigma = \sigma_c \equiv \sqrt{\lambda}\,M/g$ and the waterfall field is released from the origin towards its VEV.

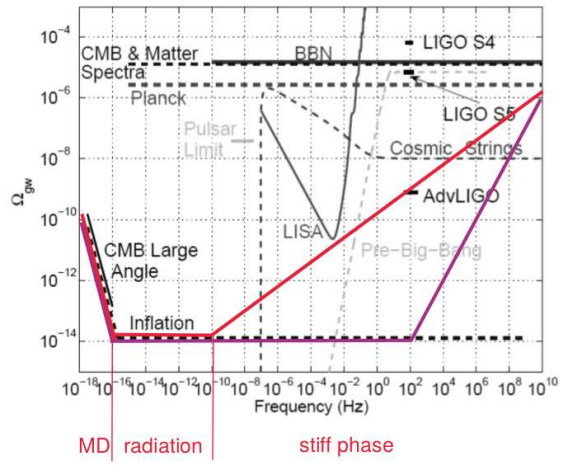



Figure 10. Plot of the GW spectrum per logarithmic frequency interval $\Omega_{\rm GW}$ with respect to frequency, up to the largest possible frequency which corresponds to the inflation energy scale ($f_{\rm end} \sim 10^{10}$ Hz) assumed at the scale of grand unification $\rho_{\rm end}^{1/4} \sim 10^{15-16}$ GeV, as is typically the case. The solid purple line depicts the spectrum of primordial GWs in the case of kination following right after inflation, with barotropic parameter $w = 1$, while the solid red line depicts the spectrum of primordial GWs in the stiff barotropic parameter $w = \frac{1}{2}$ scenario, with reheating temperature $T_{\rm reh} \sim 10^2$ MeV. The peak in the GW spectrum almost saturates the BBN bound $\Omega_{\rm GW} < 10^{-6}$, depicted by the horizontal solid black line. In the figure, the expected observational capability of Advanced LIGO and LISA are shown. It is evident that this scenario produces marginally observable (by LISA and Advanced LIGO) primordial GWs. Observational bounds have been taken from Ref. [18].

This is when things become different, because we assume there is a kinetic pole at the VEV of the waterfall field. Such a pole can be due to some non-trivial Kähler geometry as with $\alpha$-attractors [25]. The Lagrangian density is

$$\mathcal{L} = \frac{1}{2}(\partial\sigma)^2 + \frac{\frac{1}{2}(\partial\varphi)^2}{(1-\varphi^2/M^2)^2} - V(\varphi,\sigma)\,. \quad (21)$$

To assist our intuition we redefine the waterfall field so that it also is characterised by a canonical kinetic term. The redefinition is $\varphi = M\tanh(\phi/M)$, and $\phi$ is now canonically normalised. The scalar potential now

becomes

$$V(\phi,\sigma)=\frac{1}{2}g^2M^2\sigma^2\tanh^2(\phi/M)+\frac{\frac{1}{4}\lambda M^4}{\cosh^4(\phi/M)}+V(\sigma)\,. \tag{22}$$

The form of the above potential is shown in Fig. 9.

After the phase transition which ends primordial inflation, $\sigma$ quickly goes to zero and the system rolls along the runaway waterfall direction, since the VEV of $\phi$ is at infinity. Now, exactly because $M$ is Planckian, we have a further boot of fast-roll inflation [26] as the waterfall field leaves the origin, which is now a potential hill

$$V(\phi\ll M)\simeq\frac{\frac{1}{4}\lambda M^4}{[1+\frac{1}{2}(\phi/M)^2]^4}\simeq\frac{1}{4}\lambda M^4\left[1-2\left(\frac{\phi}{M}\right)^2\right]. \tag{23}$$

Based on the value of $M$, this gives about 13.5 e-folds of hilltop fast-roll inflation, following the primordial inflation. Inflation ends when $\phi_{\rm end}\simeq M/\sqrt{2}$. Afterwards, the canonical waterfall rolls along the potential tail with

$$V(\phi\gg M)\simeq\frac{\frac{1}{4}\lambda M^4}{[\frac{1}{2}\exp(\phi/M)]^4}\simeq 4\lambda M e^{-4\phi/M}. \tag{24}$$

Since the potential tail is of exponential form, the field soon follows the power-law inflation attractor [27] which is characterised by the barotropic parameter $w=-1+\frac{16}{3}\left(\frac{m_P}{M}\right)^2$. Requiring $w\approx\frac{1}{2}$ suggests that $M/m_P\approx\frac{4\sqrt{2}}{3}=1.88$. This is why we considered that $M\sim m_P$ in the first place.

As can be seen in Fig. 10, the peak in the GW spectrum is much more mild and it can come to contact with the future LISA observations.

### 5.3. Hyperkination and GWs

Another possibility to boost the primordial GWs generated during inflation at observable frequencies is by considering the model of Sec. 3.1, which is characterised by the Lagrangian density in Eq. (11). Switching to Einstein frame through a suitable conformal transformation, the Lagrangian density becomes

$$\mathcal{L}=\frac{1}{2}m_P^2R+\frac{1}{2}(\partial\phi)^2+\alpha\frac{h^2+16\alpha V}{h^2m_P^4}(\partial\phi)^4-\frac{Vm_P^4}{h^2+16\alpha V}\,, \tag{25}$$

where $h^2\equiv m_P^2+\xi\varphi^2$, with $\varphi$ being the non-canonical field, which appears in Eq. (11). Then, the Klein-Gordon equation of motion of the field is

$$\left[1+12\alpha\left(1+\frac{16\alpha V}{h^2}\right)\right]\ddot{\phi}+\left[1+4\alpha\left(1+\frac{16\alpha V}{h^2}\right)\frac{\dot{\phi}^2}{m_P^4}\right]3H\dot{\phi}+48\alpha^2\frac{\dot{\phi}^4}{m_P^4}\frac{\rm d}{{\rm d}\phi}\left(\frac{V}{h^2}\right)+\frac{\rm d}{{\rm d}\phi}\frac{Vm_P^4}{h^2+16\alpha V}=0\,. \tag{26}$$

From the energy-momentum tensor, it is straightforward to obtain the energy density and the pressure of the scalar field, which respectively read

$$\rho_\phi=\frac{1}{2}\left[1+6\alpha\left(1+\frac{16\alpha V}{h^2}\right)\frac{\dot{\phi}^2}{m_P^4}\right]\dot{\phi}^2+\frac{Vm_P^4}{h^2+16\alpha V}\,, \tag{27}$$

$$p_\phi=\frac{1}{2}\left[1+2\alpha\left(1+\frac{16\alpha V}{h^2}\right)\frac{\dot{\phi}^2}{m_P^4}\right]\dot{\phi}^2-\frac{Vm_P^4}{h^2+16\alpha V}\,.$$

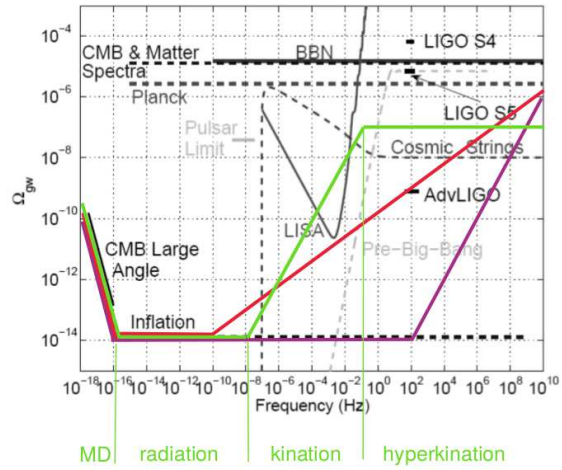



Figure 11. Plot of the GW spectrum per logarithmic frequency interval $\Omega_{\rm GW}$ with respect to frequency, up to the largest possible frequency which corresponds to the inflation energy scale ($f_{\rm end}\sim10^{10}$ Hz) assumed at the scale of grand unification $\rho_{\rm end}^{1/4}\sim10^{15-16}$ GeV, as is typically the case. The solid purple line depicts the spectrum of primordial GWs in the case of kination following right after inflation, with barotropic parameter $w=1$, while the solid red line depicts the spectrum of primordial GWs in the stiff barotropic parameter $w=\frac{1}{2}$ scenario and the solid green line depicts the spectrum of primordial GWs in the case when inflation is followed first by a period of hyperkination and then by a period of regular kination, with reheating temperature $T_{\rm reh}\sim10^2$ MeV. This time, the spectrum of the GWs is nowhere near the BBN bound $\Omega_{\rm GW}<10^{-6}$, depicted by the horizontal solid black line. In the figure, the expected observational capability of Advanced LIGO and LISA are shown. It is evident that this scenario produces a distinct, characteristic spectrum of observable (by LISA and Advanced LIGO) primordial GWs. Observational bounds have been taken from Ref. [18].

The above complicated expressions are much simplified if, after the end of inflation, the scalar field becomes kinetically dominated. In this case, we can set $V\to0$ so that Eqs. (26) and (27) are respectively reduced to

$$\left(1+12\alpha\frac{\dot{\phi}^2}{m_P^4}\right)\ddot{\phi}+\left(1+4\alpha\frac{\dot{\phi}^2}{m_P^4}\right)3H\dot{\phi}\simeq0 \tag{28}$$

and

$$\rho_\phi \simeq \frac{1}{2}\left(1+6\alpha\frac{\dot\phi^2}{m_P^4}\right)\dot\phi^2, \qquad p_\phi \simeq \frac{1}{2}\left(1+2\alpha\frac{\dot\phi^2}{m_P^4}\right)\dot\phi^2. \tag{29}$$

Now, if the standard quadratic kinetic term is dominant in Eq. (25), then this is equivalent to setting $\alpha \to 0$. In this case, Eq. (29) suggests that $w_\phi = p_\phi/\rho_\phi = 1$ and we have regular kination, which results in $\Omega_{\rm GW}(f) \propto f$ as we have discussed. If however, the quartic kinetic term is dominant in Eq. (25), then this is equivalent to considering only the terms proportional to $\alpha$. In this case, Eq. (29) suggests that $w_\phi = p_\phi/\rho_\phi = \frac{1}{3}$, similarly to radiation, for which $\Omega_{\rm GW}(f) =$ constant. We call this period hyperkination [28].

Because, after inflation, the quartic kinetic term dominates before the quadratic one takes over (and not the other way around), after inflation we have a period of hyperkination followed by a period of regular kination until reheating. This results in a GW spectrum, which features a truncated peak such that boosting the primordial GWs can occur at lower observable frequencies without violating the BBN bound in Eq. (19). The situation is depicted in Fig. 11. The distinct GW spectrum, if observed, could provide information on the theoretical background, such as the energy scale of inflation and the value of the $\alpha$ parameter, which characterises the contribution of the quadratic gravity, as shown in Eq. (11).

## 6. CONCLUSIONS

Cosmic Inflation determines the initial conditions of the Universe history and leads to a large and uniform Universe. Inflation also generates the primordial density perturbations which seed galaxy formation and are reflected on the CMB primordial anisotropy. Agreement between the CMB observations and the predictions of inflation is spectacular. In addition, the Universe today engages into a late inflationary period, which may be due to quintessence, a form of dark energy.

Quintessential inflation employs a common theoretical framework for the early and late Universe and leads to a surge in primordial gravitational waves. Typically, quintessential inflation is modelled considering a flat direction with a runaway scalar potential, which has minimum at infinity and features two flat regions: the inflationary plateau and the quintessential tail. Palatini gravity is a natural framework for model-building quintessential inflation because it "flattens" a runaway potential to generate the inflationary plateau.

Quintessential inflation is typically followed by a stiff period called kination, which generated a peak of primordial gravitational waves (GWs). However, the kination GW peak corresponds to unobservable frequencies. One way to overcome this is by considering that the peak is milder, which can be achieved in the stiff period after inflation is not as stiff as kination proper. A model realisation of this possibility considers two flat directions intersecting at an enhanced symmetry point in field space, giving rise to the hybrid mechanism, with Planckian waterfall field VEV, which is also a kinetic pole of the waterfall field, following the $\alpha$-attractors proposal.

Another interesting possibility to obtain a boost of primordial GWs down at observable frequencies in by considering higher order kinetic terms (as in k-essence [29]). This is possible to realise in Palatini modified gravity. Indeed, considering $R + R^2$ gravity and a non-minimally coupled scalar field gives rise to additional quartic kinetic terms. When these dominate, this leads to hyperkination which is followed by regular kination when the field becomes canonical. the resulting characteristic truncated GW peak can be extended to observable frequencies without disturbing BBN.

Forthcoming observations of Advanced LIGO, LISA, DesiGO or BBO may well detect the primordial GWs generated by inflation. A distinct GW spectrum will provide insights to the background theory. Finally, it should be pointed out that the detection of primordial GWs will not only confirm another prediction of cosmic inflation but also offer tantalising evidence for the quantum nature of gravity itself.


[1] A. H. Guth, Phys. Rev. D **23** (1981), 347-356.
[2] A. A. Starobinsky, Phys. Lett. B **91** (1980), 99-102.
[3] K. Dimopoulos, *Introduction to Cosmic Inflation and Dark Energy*, CRC Press, 2022, ISBN 978-0-367-61104-0
[4] D. H. Lyth and A. R. Liddle, *The primordial density perturbation: Cosmology, inflation and the origin of structure*, CUP, 2009, ISBN 978-0521828499
[5] R. K. Sachs and A. M. Wolfe, Astrophys. J. **147** (1967), 73-90.
[6] P. A. R. Ade *et al.* [Planck], Astron. Astrophys. **594** (2016), A20.
[7] P. A. R. Ade *et al.* [BICEP and Keck], Phys. Rev. Lett. **127** (2021) no.15, 151301.
[8] F. Bezrukov, A. Magnin, M. Shaposhnikov and S. Sibiryakov, JHEP **01** (2011), 016.
[9] A. G. Riess *et al.* [Supernova Search Team], Astron. J. **116** (1998), 1009-1038; S. Perlmutter *et al.* [Supernova Cosmology Project], Astrophys. J. **517** (1999), 565-586.
[10] R. R. Caldwell, R. Dave and P. J. Steinhardt, Phys. Rev. Lett. **80** (1998), 1582-1585; S. M. Carroll, Phys. Rev. Lett. **81** (1998), 3067-3070.
[11] P. J. E. Peebles and A. Vilenkin, Phys. Rev. D **59**

(1999), 063505.
[12] K. Dimopoulos and S. Sánchez López, Phys. Rev. D **103** (2021) no.4, 043533.
[13] K. Dimopoulos, A. Karam, S. Sánchez López and E. Tomberg, Galaxies **10** (2022) no.2, 57; JCAP **10** (2022), 076.
[14] M. Joyce and T. Prokopec, Phys. Rev. D **57** (1998), 6022-6049.
[15] K. Dimopoulos and C. Owen, JCAP **06** (2017), 027.
[16] M. S. Turner, Phys. Rev. D **55** (1997), R435-R439.
[17] Y. Gouttenoire, G. Servant and P. Simakachorn, [arXiv:2111.01150 [hep-ph]].
[18] E. Thrane and J. D. Romano, Phys. Rev. D **88** (2013) no.12, 124032.
[19] R. H. Cyburt, B. D. Fields, K. A. Olive and E. Skillman, Astropart. Phys. **23** (2005), 313-323.
[20] D. G. Figueroa and E. H. Tanin, JCAP **08** (2019), 011.
[21] P. Amaro-Seoane *et al.* [LISA], [arXiv:1702.00786 [astro-ph.IM]].
[22] J. Aasi *et al.* [LIGO Scientific], Class. Quant. Grav. **32** (2015), 074001.
[23] K. Dimopoulos, JCAP **10** (2022), 027.
[24] A. D. Linde, Phys. Rev. D **49** (1994), 748-754.
[25] R. Kallosh, A. Linde and D. Roest, JHEP **11** (2013), 198.
[26] A. D. Linde, JHEP **11** (2001), 052.
[27] F. Lucchin and S. Matarrese, Phys. Rev. D **32** (1985), 1316.
[28] S. Sánchez López, K. Dimopoulos, A. Karam and E. Tomberg, Eur. Phys. J. C **83** (2023) no.12, 1152.
[29] C. Armendariz-Picon, V. F. Mukhanov and P. J. Steinhardt, Phys. Rev. Lett. **85** (2000), 4438-4441; Phys. Rev. D **63** (2001), 103510.

# Slicing Pomerons in ultraperipheral collisions using forward neutrons from nuclear breakup

M. Alvioli,[1,2] V. Guzey,[3] and M. Strikman[4]

[1] *Consiglio Nazionale delle Ricerche, Istituto di Ricerca per la Protezione Idrogeologica, via Madonna Alta 126, I-06128, Perugia, Italy*
[2] *Istituto Nazionale di Fisica Nucleare, Sezione di Perugia, via Pascoli 23c, I-06123, Perugia, Italy*
[3] *University of Jyvaskyla, Department of Physics, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland and Helsinki Institute of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland*
[4] *Pennsylvania State University, University Park, PA, 16802, USA*

(Dated: March 1, 2024)

We argue that measurements of forward neutrons from nuclear breakup in inclusive high energy photon-nucleus ($\gamma A$) scattering provide a novel complementary way to study small-$x$ dynamics of QCD in heavy-ion ultraperipheral collisions (UPCs). Using the leading twist approximation to nuclear shadowing, we calculate the distribution over the number of evaporation neutrons produced in $\gamma Pb$ collisions at the LHC. We demonstrate that it allows one to determine the distribution over the number of wounded nucleons (inelastic collisions), which constrains the mechanism of nuclear shadowing of nuclear parton distributions.

PACS numbers:
Keywords: Heavy-ion scattering, ultraperipheral collisions, nuclear shadowing

arXiv:2402.19060v1 [hep-ph] 29 Feb 2024

## I. INTRODUCTION AND MOTIVATION

Understanding of the QCD dynamics of hard high energy interactions and the structure of nuclei and nucleons is one of the main directions of theoretical and experimental studies at the Large Hadron Collider (LHC) and the Relativistic Heavy Ion Collider (RHIC). Of particular interest is the limit of very small momentum fractions $x$, when the linear Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) approximation is expected to break down [1, 2] and a regime close to the black disk limit (BDL) [3] may set in. Its observation is one of the prime objectives of the planned Electron-Ion Collider (EIC) at Brookhaven National Laboratory [4, 5], which will ultimately reach $x \sim 10^{-3}$ for momentum transfers of a few GeV. At the same time, it was pointed out some time ago that ultraperipheral collisions (UPCs) of two ions at the LHC, where a photon emitted by one of the nuclei interacts with the other nucleus, allow one to probe down to $x \sim 10^{-5} - 10^{-4}$, depending on a particular reaction channel and the detector geometry [6, 7].

Electron-nucleus collisions at the EIC and UPCs of heavy ions at the LHC present two options for studying small-$x$ dynamics, which are largely complementary. At the LHC practically all data are collected for one heavy nucleus and one cannot directly access the dependence of cross sections on the virtuality of the probe. At the same time, one can reach very small $x$, which is provided by a wide rapidity coverage of the LHC detectors and the large invariant photon-nucleus collision energies exceeding by a factor of 100 the design energies at the EIC.

Over the last decade the data taken in the LHC and RHIC kinematics discovered a significant nuclear suppression of coherent $J/\psi$ photoproduction in Pb-Pb and Au-Au UPCs compared to the impulse approximation prediction [8–17]. When interpreted in the leading twist approximation (LTA) [18], it amounts to strong gluon nuclear shadowing [19, 20]:

$$R_A^g(x,Q^2) = \frac{g_A(x,Q^2)}{A g_N(x,Q^2)} < 1\,, \tag{1}$$

where $g_A(x,Q^2)$ and $g_N(x,Q^2)$ are the nucleus and nucleon gluon densities, respectively. Typical numbers reported by the LHC experiments [8–16] correspond to

$$\begin{aligned} R_{\rm Pb}^g(x=10^{-3}, Q^2_{\rm eff}=3\,{\rm GeV}^2) &\approx 0.6\,,\\ R_{\rm Pb}^g(x=10^{-4}, Q^2_{\rm eff}=3\,{\rm GeV}^2) &\approx 0.5\,, \end{aligned} \tag{2}$$

with a similar suppression extending down to $x \sim 10^{-5}$. Here $Q_{\rm eff}$ is the effective resolution scale determined by the charm quark mass. These values of $R_{\rm Pb}^g$ agree very well with the LTA predictions for nuclear shadowing made more than 10 years ago [18]. Note that this interpretation of the $J/\psi$ UPC data is complicated at the next-to-leading order (NLO) of the perturbative expansion in powers of $\log Q^2$ (perturbative QCD) due to large cancellations between the leading-order (LO) and NLO gluon terms, which leaves a numerically important quark contribution [21, 22]. A way to

stabilize the perturbation series and restore the gluon dominance in this process on the proton target was suggested in [23, 24].

Other hard UPC processes considered in the literature are inclusive [6, 25] and diffractive [26] dijet photoproduction, which are still to produce final results [27, 28], timelike Compton scattering [29–31], and heavy quark photoproduction [32–34]. Overall these findings confirm the conclusion of [7] that UPCs provide a very effective tool to access the small-$x$ dynamics of the strong interactions and the nuclear structure in hard, semi-hard and soft regimes.

So far the experimental studies of UPCs have mainly been focusing on coherent and incoherent production of light and heavy vector mesons. In this paper, we would like to outline several possible directions of future UPC studies, which were not discussed in the review [7] due to the lack of experimental confirmations of large nuclear shadowing. We explore for the first time possibilities of testing the small-$x$ nuclear shadowing dynamics by measuring the rates of forward neutron production from nuclear breakup in the zero degree calorimeters (ZDCs) at the LHC. Our numerical studies demonstrate that the number of produced neutrons is correlated with the number of wounded nucleons (inelastic photon-nucleon interactions), which presents a complementary way to study the mechanism of nuclear shadowing of nuclear parton distributions. In particular, this allows one to effectively probe the centrality (impact parameter) dependence of nuclear shadowing by experimentally controlling the number of unitarity cuts of diffractive exchanges (Pomerons), which build up the leading twist nuclear shadowing.

This paper is organized as follows. In Sec. II we briefly summarize expectations based on applications of the Abramovski-Gribov-Kancheli (AGK) theorem to photon-nucleus scattering, model the distribution over the number of wounded nucleons and estimate its average value in the current UPC kinematics. Section III presents our predictions for the distributions over the number of emitted forward neutrons from nuclear breakup in inelastic photon-nucleus scattering and its connection to parameters of the leading twist approximation of nuclear shadowing. Our conclusions and outlook are given Sec. IV.

## II. ABRAMOVSKI-GRIBOV-KANCHELI (AGK) CUTTING RULES, NUCLEAR SHADOWING AND THE NUMBER OF WOUNDED NUCLEONS IN $\gamma A$ SCATTERING

It was demonstrated by Abramovski, Gribov and Kancheli in 1973 [35] that different unitary cuts of the diagrams corresponding to multi-Pomeron (color singlet) exchanges result in different multiplicities of produced particles in the central rapidity region and that the absorptive part of the amplitude can be expressed in terms of a small number of cut diagrams, which are related by combinatorial factors; these are the so-called Abramovski-Gribov-Kancheli (AGK) cutting rules or the AGK cancellation. For the interpretation of the AGK rules in QCD and other effective field theories, see Refs. [36–40].

The application of the AGK cutting rules to photon-nucleus scattering allows one to express the nuclear shadowing correction to the total nuclear cross section $\sigma_{\rm tot}^{\gamma A}$ in terms of the diffractive cross section on individual nucleons [41, 42]. Indeed, the nuclear amplitude can be constructed as a series corresponding to the interactions with $k = 1, 2, \ldots, A$ nucleons of the target, which are mediated by Pomeron exchanges. The interaction with one nucleon gives the impulse approximation (IA),

$$\sigma_{\rm tot}^{\gamma A({\rm IA})} \;=\; A\,\sigma_{\rm tot}^{\gamma N}\,, \tag{3}$$

where $\sigma_{\rm tot}^{\gamma N}$ is the total photon-nucleon cross section. For the interaction with two nucleons, the nuclear cross section is given by a sum of cut two-Reggeon amplitudes corresponding to both Reggeons uncut ($\sigma_0$), one cut Reggeon ($\sigma_1$), and two cut Reggeons ($\sigma_2$),

$$\sigma_{\rm tot}^{\gamma A(k=2)} = \sigma_0 + \sigma_1 + \sigma_2\,. \tag{4}$$

Figure 1 shows the diagrams corresponding to the three terms in Eq. (4). The AGK rules relate these cross sections as follows

$$\begin{aligned} \sigma_1 \;&=\; -4\,\sigma_0\,, \\ \sigma_2 \;&=\; 2\,\sigma_0\,, \end{aligned} \tag{5}$$

where $\sigma_1$ is the absorptive (screening) correction to the IA cross section, and $\sigma_2$ corresponds to the inelastic interaction with both nucleons. In Eq. (5), we neglected the small real part of the amplitude for simplicity.

Substituting Eq. (5) in Eq. (4), one finds for the nuclear shadowing correction $\Delta\sigma_{\rm tot}^{\gamma A}$ in the approximation, where one keeps the interaction with only one and two nucleons,

$$\Delta\sigma_{\rm tot}^{\gamma A} = \sigma_{\rm tot}^{\gamma A} - \sigma_{\rm tot}^{\gamma A({\rm IA})} = \sigma_{\rm tot}^{\gamma A(k=2)} = -\sigma_0\,. \tag{6}$$

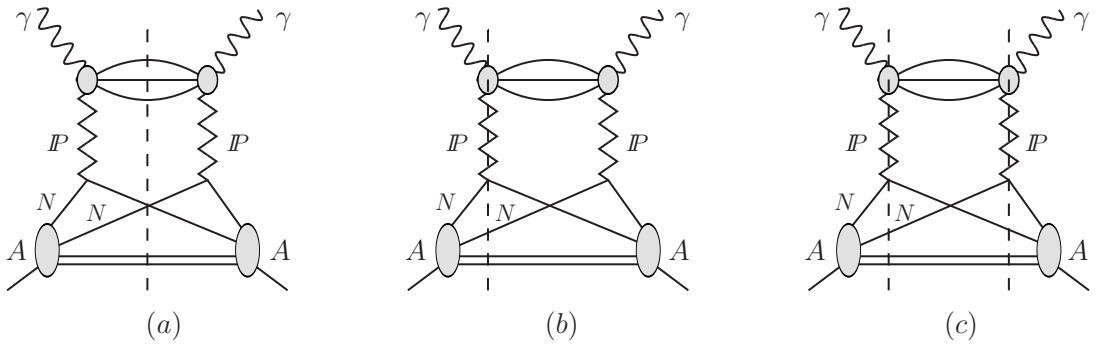



FIG. 1: Unitary cuts of the two-Reggeon forward amplitude, which gives rise to the nuclear shadowing correction due to the interaction with $k = 2$ nucleons of the nuclear target: $(a)$ the diffractive cut with both Reggeons uncut and the cross section $\sigma_0$, $(b)$ one cut Reggeon corresponding to a single multiplicity and the cross section $\sigma_1$, and $(c)$ two cut Reggeons corresponding to a double multiplicity and the cross section $\sigma_2$.

This relation constitutes the foundation of the Gribov model of nuclear shadowing for the total cross section of pion-deuteron scattering, which relates the nuclear shadowing correction to the pion-nucleon diffractive cross section [43]. Note that the AGK rules were derived later and were not explicitly used in [43]. For the generalization of this approach for the calculation of total cross sections of photon scattering on heavier nuclei, see [44–47].

Equation (6) is general and does not require an expansion over twists or the partonic interpretation. However, in lepton-proton deep inelastic scattering (DIS), $\sigma_0$ is a leading twist cross section, which by the virtue of the QCD collinear factorization theorem for hard diffraction [48] can be interpreted in terms of diffractive parton distribution functions (PDFs). As a result, it allows one to formulate the leading twist approximation for small-$x$ nuclear PDFs of individual flavors (valence and sea quarks and gluons) [18, 42, 49].

Another important application of the AGK cutting rules involves the total hadron-nucleus inelastic cross section defined as the difference between the total and total elastic (coherent plus incoherent) cross sections, which is obtained using the unitary form of the Glauber theory [50]. Generalizing this result to photon-nucleus scattering by taking into account that the photon takes part in the strong interactions by means of its hadronic fluctuations, the photon-nucleus total inelastic cross section can be presented in the following form [51]

$$\sigma_{\rm inel}^{\gamma A} = \sum_{\nu=1}^{A} \sigma_\nu \,, \tag{7}$$

where

$$\sigma_\nu = \frac{A!}{(A-\nu)!\nu!} \int d^2\vec{b} \int d\sigma P_\gamma(\sigma) (\sigma_{\rm inel} T_A(\vec{b}))^\nu (1 - \sigma_{\rm inel} T_A(\vec{b}))^{A-\nu} \,. \tag{8}$$

In Eq. (8), $\vec{b}$ is the impact parameter (transverse coordinate) of the interacting nucleon, $T_A(\vec{b}) = \int dz \rho_A(\vec{b}, z)$, where $\rho_A(\vec{b}, z)$ is the nuclear density normalized to unity, and $\sigma_{\rm inel} = 0.85\,\sigma$ is the inelastic cross section for the interaction of a hadronic fluctuation of the photon with a target nucleon, which is based on the estimate that the $\rho$ meson-nucleon elastic cross section constitutes approximately 15% of the total one. Unlike the sign-alternating Glauber series for the nuclear total cross section [52], the cross sections $\sigma_\nu$ in Eqs. (7) and (8) are positive and represent the physical process, where $\nu$ nucleons undergo inelastic scattering, while the remaining $A - \nu$ nucleons provide absorption. In the literature, one uses the term "wounded nucleons" [53] and the notation $\nu = N_{\rm coll}$.

The distribution $P_\gamma(\sigma)$ gives the probability density for hadronic fluctuations of the real photon to interact with nucleons with the cross section $\sigma$ [51, 54]. While the shape of $P_\gamma(\sigma)$ cannot be calculated from the first principles, one can reliably model it using constraints on its first moments and the small-$\sigma$ and large-$\sigma$ limits. Indeed, the total photon-proton cross section $\sigma_{\gamma p}(W)$ and the cross section of photon diffractive dissociation on the proton $d\sigma_{\gamma p \to Xp}(W, t = 0)/dt$ constrain the first two moments of $P_\gamma(\sigma)$ as follows

$$\begin{aligned} \sigma_{\gamma p}(W) &= \int d\sigma P_\gamma(\sigma) \sigma \,, \\ \frac{d\sigma_{\gamma p}(W, t=0)}{dt} &= \frac{1}{16\pi} \int d\sigma P_\gamma(\sigma) \sigma^2 \,, \end{aligned} \tag{9}$$

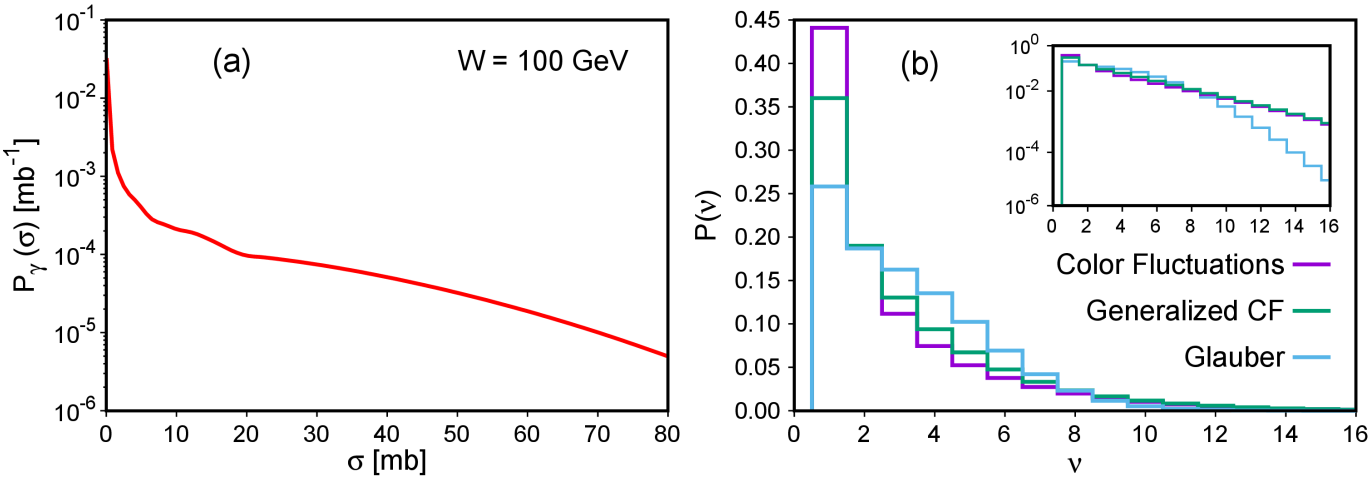



FIG. 2: (a) The $P_\gamma(\sigma)$ probability density for hadronic fluctuations of the real photon to interact with nucleons with the cross section $\sigma$ at $W = 100$ GeV. (b) The $P(\nu)$ distribution over the number of wounded nucleons (inelastic interactions) $\nu$ in inelastic photon-nucleus (Pb) scattering. The three curves correspond to the three models for $P_\gamma(\sigma)$ parametrizing hadronic fluctuations of the real photon, see text for details. The insert emphasizes the region of large $\nu$. The figures are adopted from [51].

where $W$ is the invariant photon-nucleon center-of-mass-energy. Further, in the small-$\sigma$ limit, one can express $P_\gamma(\sigma)$ in terms of the quark-antiquark component of the photon light-cone wave function and the color dipole cross section, which results in $P_\gamma(\sigma) \propto 1/\sigma$. In the opposite limit of large $\sigma$, the photon behaves as a superposition of the $\rho$, $\omega$ and $\phi$ vector mesons in the spirit of the vector meson dominance model and, hence, $P_\gamma(\sigma)$ can be modeled using hadronic (cross section) fluctuations in $\rho$ mesons, which in turn are related to those for pions. Finally, the small-$\sigma$ and large-$\sigma$ regimes can be smoothly interpolated. Note that this matching is achieved best, when the light quark masses $m_q$ are taken to be those of the constituent quarks, $m_q \sim 300$ MeV. For details, see [51, 54].

The left panel of Fig. 2 presents the resulting distribution $P_\gamma(\sigma)$ as a function of $\sigma$ at $W = 100$ GeV. Its shape and normalization are constrained by the procedure outlined above, the parametrization of $\sigma_{\gamma p}(W)$ [55], and the data on $d\sigma_{\gamma p \to X p}(W, t = 0)/dt$ [56]. Since the $W$ dependence of $P_\gamma(\sigma)$ is weak, the presented distribution is applicable to a wide range of energies probed in heavy-ion UPCs at the LHC. Note that the distribution $P_\gamma(\sigma)$ parametrizes the so-called resolved photon contribution to photon-induced scattering and does contain the direct photon contribution.

Using Eqs. (7) and (8), one can readily define the $P(\nu)$ probability distribution for the number of wounded nucleons $\nu$ in inelastic photon-nucleus scattering as follows [51],

$$P(\nu) = \frac{\sigma_\nu}{\sum_{\nu=1}^{A} \sigma_\nu}\,, \tag{10}$$

where $\sigma_\nu$ is given by Eq. (8). To calculate it, we use the Monte Carlo generator for nucleon configurations in complex nuclei [57], which also includes nucleon-nucleon correlations in the nucleus wave function [58, 59], and the Gribov-Glauber model for photon-nucleus scattering, where the hadronic structure of the photon is described by the distribution $P_\gamma(\sigma)$ [51]. The resulting distribution $P(\nu)$ as a function of $\nu$ for lead (Pb) is shown by the curve labeled "Color Fluctuations" in the right panel of Fig. 2.

Our modeling of small-$\sigma$ hadronic fluctuations of the photon is based on the quark-antiquark component of the photon wave function and corresponds to a weak nuclear shadowing in disagreement with the observed strong gluon nuclear shadowing, see Eqs. (1) and (2). To take the latter into account, we model the nuclear suppression of the dipoles with $\sigma \le \sigma_0 = 20$ mb by the factor of $R_A^g$, which leads to the modified expression for the distribution $\tilde{P}_\gamma(\sigma)$,

$$\tilde{P}_\gamma(\sigma) = \left[R_A^g(x, Q^2_{\rm eff})\theta(\sigma_0 - \sigma) + \theta(\sigma - \sigma_0)\right] P_\gamma(\sigma)\,, \tag{11}$$

where $x = Q^2_{\rm eff}/W^2$ and $P_\gamma(\sigma)$ is shown in Fig. 2 (left panel). The distribution $P(\nu)$ corresponding to $\sigma_\nu$ calculated using $\tilde{P}_\gamma(\sigma)$ is given by the curve "Generalized CF" in the right panel of Fig. 2.

Finally, to test the importance of cross section (color) fluctuations in the real photon, we calculated $P(\nu)$ neglecting these fluctuations and using

$$P_\gamma(\sigma) = \delta(\sigma - 25\,{\rm mb}) \tag{12}$$

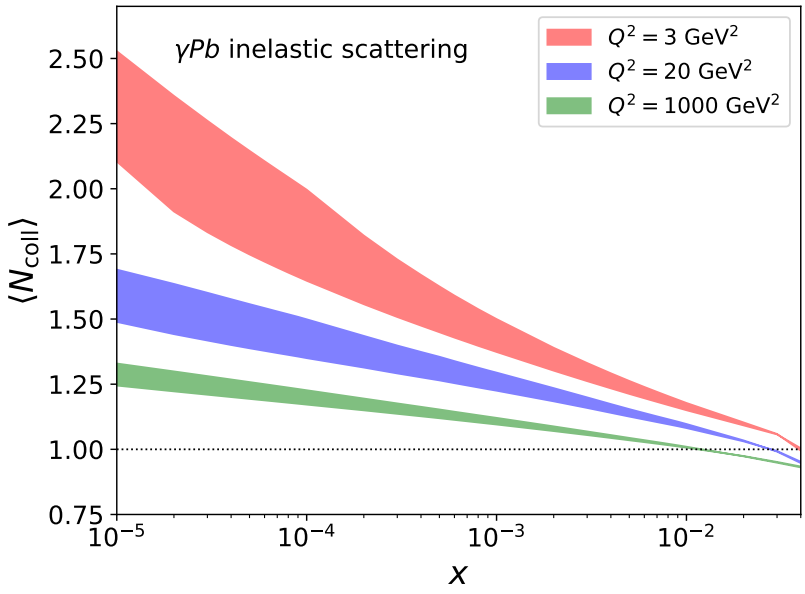



FIG. 3: The LTA predictions for the average number of wounded nucleons $\langle N_{\rm coll}\rangle$ in inelastic photon-nucleus (Pb) scattering as a function of $x$ at $Q^2 = 3$, 20, and 1000 GeV$^2$.

in Eq. (8). The result is given by the curve "Glauber" in Fig. 2. One can see from this figure that color fluctuations significantly increase the distribution $P(\nu)$ at small and especially large $\sigma$; the latter is emphasized in the insert.

In the total inelastic photon-nucleus cross section, the AGK cancellations manifest themselves as the observation that the average number of wounded nucleons $\langle N_{\rm coll}\rangle$ is inversely proportional to the nuclear shadowing factor. Generalizing the result of [50] for hadron-nucleus scattering to the case of photon-induced scattering, one obtains

$$\langle N_{\rm coll}\rangle \equiv \sum_{\nu=1}^{A} P(\nu)\nu = \frac{\sum_{\nu=1}^{A}\nu\sigma_\nu}{\sum_{\nu=1}^{A}\sigma_\nu} = \frac{A\sigma_{\rm inel}^{\gamma N}}{\sigma_{\rm inel}^{\gamma A}}\,, \tag{13}$$

where $\sigma_{\rm inel}^{\gamma N}$ is the photon-nucleon inelastic cross section. Considering a particular hard process in inelastic photon-nucleus scattering that probes the nuclear gluon distribution, *e.g.*, inclusive charmonium (bottomonium) production $\gamma + A \to J/\psi(\Upsilon) + X$ or inclusive heavy-quark dijet production $\gamma + A \to Q\bar{Q} + X$, one obtains using Eq. (13):

$$\langle N_{\rm coll}\rangle \approx \frac{1}{R_{\rm Pb}^{g}(x,Q^2)} \lesssim 2\,. \tag{14}$$

In this estimate we used the following considerations. In general, the effect of nuclear shadowing for the nuclear inelastic cross section is somewhat larger than that for the total cross section. However, the theoretical uncertainties of the leading twist approximation [18] largely mask the differences and make the shadowing effects approximately equal (within uncertainties) in the two cases. Finally, in the last step, we used the results of Eq. (2).

Figure 3 shows predictions of the leading twist approximation for the average number of wounded nucleons $\langle N_{\rm coll}\rangle = 1/R_{\rm Pb}^{g}(x,Q^2)$ in inelastic photon-nucleus (Pb) scattering as a function of $x$ at $Q^2 = 3$, 20, and 1000 GeV$^2$. These values of $Q^2$ correspond to photoproduction $J/\psi$, $\Upsilon$, and high-$p_T$ dijets, respectively. One can see from the figure that in the discussed kinematics, the average number of wounded nucleons is modest and the series in Eq. (13) converges rather rapidly. In particular, we have checked that it is saturated by first six terms with a 5% precision. Note, however, that the convergence slows down in the limit of small $x$.

Measurements of $\langle N_{\rm coll}\rangle$ present a new method to study nuclear shadowing in inelastic photon-nucleus scattering. Unlike the observables used so far, the constraint of Eq. (14) indicates that one can perform a "Pomeron surgery" of nuclear shadowing by cutting a small number of Pomeron exchanges controlling the number of inelastic interactions with target nucleons. As a result, it gives an opportunity for an experimental determination of a small number of parameters quantifying nuclear shadowing, which leads to a systematic improvement of its theoretical description. We elaborate on it in the following section.

## III. THE DISTRIBUTIONS OVER THE NUMBER OF WOUNDED NUCLEON AND FORWARD NEUTRONS FROM NUCLEAR BREAKUP

While the average number of inelastic interactions $\langle N_{\rm coll}\rangle$ encodes information on the energy and scale dependence of nuclear shadowing, its dependence on the impact parameter of the collision is averaged over. To obtain a more microscopic description of nuclear shadowing and restore its impact parameter dependence, one needs to determine not only $\langle N_{\rm coll}\rangle$, but also the entire distribution over the number of wounded nucleons. This can be done using experimental data on the neutron emission resulting from nucleus fragmentation in a given UPC process, *e.g.*, in inclusive quarkonium or dijet photoproduction in heavy-ion UPCs with an additional condition of $Xn$ neutrons in the zero degree calorimeter (ZDC) on the nuclear target side [27, 28].

Very little is known about the dynamics of neutron emission in high energy scattering off heavy nuclei. The ALICE collaboration measured the distribution over the number of collisions $\langle N_{\rm coll}\rangle$ in proton-nucleus scattering as determined by the energy release ($E_T$) at central rapidities and the neutron multiplicity as a function of $E_T$. It was observed that $\langle N_{\rm coll}(E_T)\rangle$ is linearly proportional to number of evaporation neutrons $\langle M_n(E_T)\rangle$ for the same $E_T$ bins at least up to $\langle N_{\rm coll}\rangle \sim 10$, for details, see [60]. Note that in our case, $\langle N_{\rm coll}\rangle$ is much lower, see Eq. (14).

Another important observation made by E665 experiment at Fermilab is that in muon-nucleus deep inelastic scattering (DIS) in coincidence with detection of slow neutrons, $\mu^- + A \to n + X$, the average neutron multiplicity $\langle M_n\rangle$ for the lead target is [61]

$$\langle M_n\rangle \approx 5\,. \tag{15}$$

This result has been understood in the framework of cascade models of nuclear DIS [62, 63], where soft neutrons are produced either directly in DIS on a bound nucleon or through statistical decay (de-excitation) of the excited residual nucleus, leading to neutron evaporation[1]. The latter mechanism depends strongly on the hadron formation time: to describe the energy spectrum of emitted neutrons, one has to assume that only nucleons and pions with momenta $\leq 1$ GeV/c could be involved in final state interactions. A similar conclusion was reached by M. Baker (private communication, 2022) using the BeAGLE Monte Carlo generator [65].

This suggests the following space-time picture of forward neutron production in high energy photon-nucleus scattering. Well before the target, the incoming photon fluctuates into long-lived hadronic components, which pass through the nucleus and interact inelastically with several nucleons. This leads to the creation of holes in the nucleus (particle-hole excitations in terminology of a nuclear shell model), which de-excite and cool the nucleus by evaporating neutrons. It also produces a number of soft particles with the momenta less than 1 GeV/c, which in turn generate more neutrons.

The nucleon fragmentation weakly depends on the incident energy due to Feynman scaling and, hence, it is not significantly affected by the energy conservation constraint, which is important in the case, when one uses hadron multiplicities at central rapidities [66]. In this case, the energy transferred to the rest of the nucleus, which heats the residual nuclear system, is proportional to $\langle N_{\rm coll}\rangle$. Since the Fermilab data [61] corresponds to the average momentum fraction $\langle x\rangle = 0.015$, where the nuclear shadowing effect is small, one finds that $\langle N_{\rm coll}\rangle \approx 1$, see Eq. (13). Thus, every inelastic photon-nucleon interaction results on average in 5 forward neutrons.

To test this hypothesis, we perform two numerical studies. First, we consider a simple model, which assumes that the probability density of neutron emission is given by the Poisson distribution and that each hole created in the target nucleus generates independently on average $\langle M_n\rangle$ neutrons. Therefore, the neutron probability distribution for $\nu = \langle N_{\rm coll}\rangle$ wounded nucleons is

$$P_{\rm Poisson}(N;\lambda=\nu\langle M_n\rangle) = \frac{(\nu\langle M_n\rangle)^N e^{-\nu\langle M_n\rangle}}{N!}\,, \tag{16}$$

where $N$ is the number of produced neutrons (neutron multiplicity). The resulting probability density (frequency) for Pb as a function of $N$ is shown in the left panel of Fig. 4. In this estimate, we used $\langle M_n\rangle = 5$, see Eq. (15) and Ref. [61], and $\nu = \langle N_{\rm coll}\rangle = 1$, 2 or 3 independent neutron emissions (wounded nucleons). One can see from the figure that since the distributions for different $\nu$ are peaked at different values of $N$ and do not significantly overlap, the measurement of the forward neutron multiplicity can be used to reliably separate contributions of different numbers of wounded nucleons $\langle N_{\rm coll}\rangle$. Our analysis also shows that this separation becomes even cleaner with an increase of $\langle M_n\rangle$ (not shown here).

---

[1] The geometry of the heating is very different from the case of $AA$ collisions, where in each collision a large postion of nucleons in each nucleus interacts and de-excitation of the spectators only occurs close to the interaction surface [64].

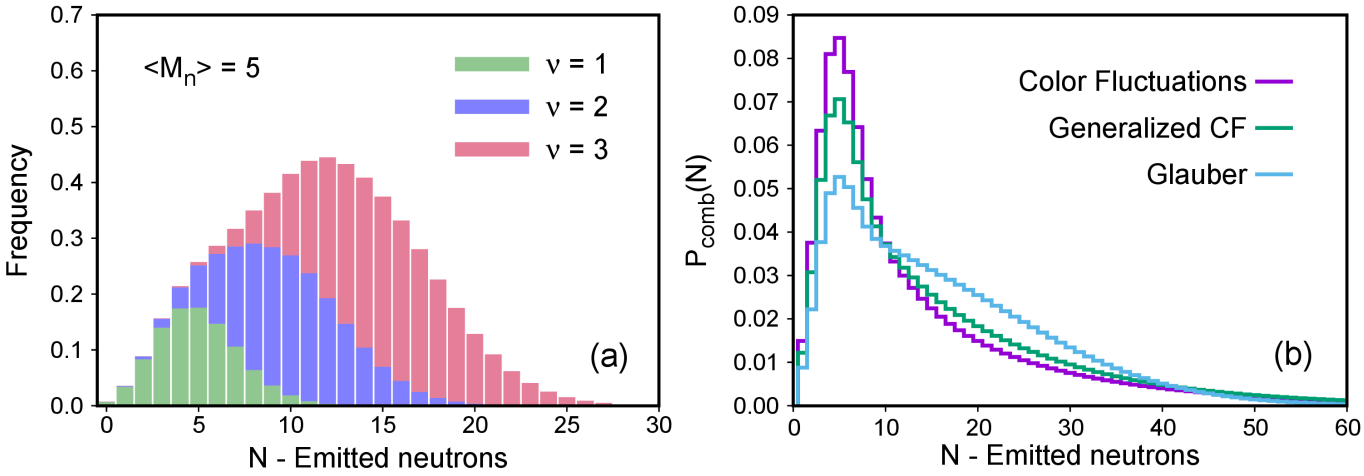



FIG. 4: (a) The probability distribution (frequency) for Pb to produce $N$ forward neutrons assuming their Poisson distribution with $\nu = \langle N_{\rm coll}\rangle$=1, 2, and 3 independent emissions (wounded nucleons) and the $\langle M_n\rangle = 5$ average multiplicity for a single inelastic scattering. (b) The probability distribution of forward neutron emission $P_{\rm comb}(N)$ (17) in the model combining the distribution over the number of wounded nucleons, which is calculated using three models for hadronic fluctuations in the photon (see Fig. 2), with the Poisson distribution of produced neutrons.

In the second study, we combine the distribution over the number of wounded nucleons $P(\nu)$ that we discussed in Sec. II, see Eq. (10) and Fig. 2 (right panel), with the Poisson distribution of produced neutrons, see Eq. (16) and Fig. 4 (left panel). The resulting probability distribution of forward neutrons is given by the following expression,

$$P_{\rm comb}(N) = \sum_{\nu=1}^{A} P(\nu) P_{\rm Poisson}(N; \nu\langle M_n\rangle)\,. \tag{17}$$

The right panel of Fig. 4 presents $P_{\rm comb}(N)$ as a function of forward neutrons $N$ for $\langle M_n\rangle = 5$. The three curves correspond to the three models for the hadronic (color) fluctuations of the real photon, see the right panel of Fig. 2. One can see from the figure that cross section fluctuations in the real photon noticeably affect the shape of the neutron distribution: its maximum is shifted toward smaller $N$ compared to the "Glauber" result and is also somewhat suppressed by the leading twist shadowing in the "Generalized CF" case.

Our numerical studies suggest that one can examine details of the theoretical description of small-$x$ nuclear shadowing in QCD using the distribution of forward neutrons from nuclear breakup emitted in a given hard UPC process, which is directly correlated with the number of inelastic interactions (wounded nucleons). By studying the dispersion of this distribution, one can single out the individual contributions of $\nu = 1, 2, 3, \ldots$ wounded nucleons. In particular, using Eq. (8) in the approximation that the series in Eq. (13) is saturated by first few terms, which corresponds to the limit of small-to-modest nuclear shadowing, one obtains [51, 67]

$$\sigma_2(x,Q^2) \equiv \frac{\langle \sigma^2_{\rm inel}\rangle}{\langle \sigma_{\rm inel}\rangle} = \frac{\langle N_{\rm coll}-1\rangle}{\frac{(A-1)}{2}\int d^2\vec{b}\, T_A^2(\vec{b})}\,, \tag{18}$$

and

$$\sigma_3(x,Q^2) \equiv \frac{\langle \sigma^3_{\rm inel}\rangle}{\langle \sigma^2_{\rm inel}\rangle} = \frac{\langle \sigma^3_{\rm inel}\rangle}{\langle \sigma_{\rm inel}\rangle}\frac{\langle \sigma^2_{\rm inel}\rangle}{\langle \sigma^2_{\rm inel}\rangle} = \frac{(\langle N_{\rm coll}-2\rangle)(\langle N_{\rm coll}-1\rangle)}{\langle N_{\rm coll}-1\rangle}\frac{\frac{(A-1)}{2}\int d^2\vec{b}\, T_A^2(\vec{b})}{\frac{(A-2)(A-1)}{2}\int d^2\vec{b}\, T_A^3(\vec{b})}\,, \tag{19}$$

where $\langle \sigma^n_{\rm inel}\rangle = \int d\sigma P_\gamma(\sigma)\sigma^n_{\rm inel}$. The cross sections $\sigma_2(x,Q^2)$ and $\sigma_3(x,Q^2)$ are essential ingredients of the LTA approach [18]. While $\sigma_2(x,Q^2)$ is determined using the HERA data on inclusive diffraction in lepton-proton DIS, $\sigma_3(x,Q^2)$ is model-dependent and leads to theoretical uncertainties of this approach. Thus, an independent determination $\sigma_2(x,Q^2)$ and $\sigma_3(x,Q^2)$ using UPCs with forward neutrons will supply new constraints on these quantities.

Note that to reach a high accuracy in such an analysis, one needs to calibrate the theoretical description against the kinematics, where only one target nucleon is struck, *e.g.*, $\gamma + A \to 2$ jets $+ X$ or quasi-elastic $J/\psi$ production for $x_A \geq 0.01$, where the effect of nuclear shadowing is small, and $\langle N_{\rm coll}\rangle \approx 1$.

## IV. CONCLUSIONS AND OUTLOOK

In this paper, we advertise measurements of forward neutrons from nuclear breakup in inclusive high energy photon-nucleus scattering in heavy-ion UPCs, e.g., charmonium (bottomonium) production $\gamma + A \to J/\psi(\Upsilon) + X$ or heavy-quark dijet production $\gamma + A \to Q\bar{Q} + X$, as a novel way to study the QCD dynamics at small $x$. The key quantity is the number of inelastic photon-nucleon interactions (the number of wounded nucleons): its average value $\langle N_{\rm coll}\rangle$ is proportional to inverse of the gluon nuclear shadowing and its first moments are sensitive to details of the leading twist mechanism of nuclear shadowing. Our numerical analysis has demonstrated that the number of forward neutrons from nuclear breakup detected in the ZDC on the nuclear target side is rather unambiguously proportional to the number of wounded nucleons, which provides a practical opportunity for novel studies of nuclear shadowing.

On top of providing stringent tests of the dynamics of leading twist shadowing of gluon PDFs, it would be possible to explore effects related to proximity to the black disk limit of the strong interaction. For example, one can study fragmentation of leading hadrons in $\gamma A$ scattering and look for suppression of their multiplicity as a function of Feynman $x_F$ and $W$ as well as for broadening of their transverse momentum distribution [3]. These effects should be more pronounced for central collisions characterized by an enhanced activity in the ZDC. It should be possible to construct from the data an analog of the $R_{CP}$ ratio, which would probe the density dependence of fragmentation. It would also be useful to construct similar quantities for low $p_T$ charm production.

Another interesting application is for multiparton interactions in proton-nucleus ($pA$) scattering. It was argued in [68] that the single and double scattering can be separated using their dependence on the impact parameter: the former is proportional to $A$, while the latter $\propto A^{4/3}$. However, since both hard interactions are typically detected in a limited range of rapidities $|y| \leq 3-4$, centrality is difficult to determine from the transverse energy $E_T$ signal because multiparton interactions also contribute to $E_T$. The use of forward neutrons in ZDCs would alleviate this problem.

One should point out that the neutrons detected in ZDCs can be a promising complementary way to determine centrality of various photon-nucleus and proton-nucleus inelastic collisions expanding the use of ZDCs beyond their current use in vector meson diffractive production and for determining of centrality of the heavy-ion collisions. The main advantage of using forward neutrons rather than the transverse energy $E_T$ for the determination of centrality is a much larger distance in rapidity between the rapidity of the hard process and that of the process used for determination of the centrality.

The methods presented in this paper can be readily generalized to the case of virtual photons and allow one to predict the distribution over the number of forward neutrons in inelastic photon-nucleus scattering at the EIC.

### Acknowledgments

The research of V.G. was funded by the Academy of Finland project 330448, the Center of Excellence in Quark Matter of the Academy of Finland (projects 346325 and 346326), and the European Research Council project ERC-2018-ADG-835105 YoctoLHC. The research of M.S. was supported by the US Department of Energy Office of Science, Office of Nuclear Physics under Award No. DE- FG02-93ER40771.

[1] Y. L. Dokshitzer, D. Diakonov and S. I. Troian, Phys. Rept. **58**, 269-395 (1980)
[2] L. V. Gribov, E. M. Levin and M. G. Ryskin, Phys. Rept. **100**, 1-150 (1983)
[3] L. Frankfurt, V. Guzey, M. McDermott and M. Strikman, Phys. Rev. Lett. **87**, 192301 (2001) [arXiv:hep-ph/0104154 [hep-ph]].
[4] A. Accardi, J. L. Albacete, M. Anselmino, N. Armesto, E. C. Aschenauer, A. Bacchetta, D. Boer, W. K. Brooks, T. Burton and N. B. Chang, *et al.* Eur. Phys. J. A **52**, no.9, 268 (2016) [arXiv:1212.1701 [nucl-ex]].
[5] R. Abdul Khalek, A. Accardi, J. Adam, D. Adamiak, W. Akers, M. Albaladejo, A. Al-bataineh, M. G. Alexeev, F. Ameli and P. Antonioli, *et al.* Nucl. Phys. A **1026**, 122447 (2022) [arXiv:2103.05419 [physics.ins-det]].
[6] M. Strikman, R. Vogt and S. N. White, Phys. Rev. Lett. **96**, 082001 (2006) [arXiv:hep-ph/0508296 [hep-ph]].
[7] A. J. Baltz, G. Baur, D. d'Enterria, L. Frankfurt, F. Gelis, V. Guzey, K. Hencken, Y. Kharlov, M. Klasen and S. R. Klein, *et al.* Phys. Rept. **458**, 1-171 (2008)
[8] E. Abbas *et al.* [ALICE], Eur. Phys. J. C **73**, no.11, 2617 (2013) [arXiv:1305.1467 [nucl-ex]].
[9] B. Abelev *et al.* [ALICE], Phys. Lett. B **718**, 1273-1283 (2013) [arXiv:1209.3715 [nucl-ex]].
[10] V. Khachatryan *et al.* [CMS], Phys. Lett. B **772**, 489-511 (2017) [arXiv:1605.06966 [nucl-ex]].
[11] S. Acharya *et al.* [ALICE], Eur. Phys. J. C **81**, no.8, 712 (2021) [arXiv:2101.04577 [nucl-ex]].
[12] S. Acharya *et al.* [ALICE], Phys. Lett. B **798**, 134926 (2019) [arXiv:1904.06272 [nucl-ex]].
[13] R. Aaij *et al.* [LHCb], JHEP **07**, 117 (2022) [arXiv:2107.03223 [hep-ex]].

[14] R. Aaij *et al.* [LHCb], JHEP **06**, 146 (2023) [arXiv:2206.08221 [hep-ex]].
[15] A. Tumasyan *et al.* [CMS], Phys. Rev. Lett. **131**, no.26, 262301 (2023) [arXiv:2303.16984 [nucl-ex]].
[16] S. Acharya *et al.* [ALICE], JHEP **10**, 119 (2023) [arXiv:2305.19060 [nucl-ex]].
[17] [STAR], [arXiv:2311.13632 [nucl-ex]].
[18] L. Frankfurt, V. Guzey and M. Strikman, Phys. Rept. **512**, 255-393 (2012) [arXiv:1106.2091 [hep-ph]].
[19] V. Guzey, E. Kryshen, M. Strikman and M. Zhalov, Phys. Lett. B **726**, 290-295 (2013) [arXiv:1305.1724 [hep-ph]].
[20] V. Guzey and M. Zhalov, JHEP **10**, 207 (2013) [arXiv:1307.4526 [hep-ph]].
[21] K. J. Eskola, C. A. Flett, V. Guzey, T. Löytäinen and H. Paukkunen, Phys. Rev. C **106**, no.3, 035202 (2022) [arXiv:2203.11613 [hep-ph]].
[22] K. J. Eskola, C. A. Flett, V. Guzey, T. Löytäinen and H. Paukkunen, Phys. Rev. C **107**, no.4, 044912 (2023) [arXiv:2210.16048 [hep-ph]].
[23] S. P. Jones, A. D. Martin, M. G. Ryskin and T. Teubner, J. Phys. G **43**, no.3, 035002 (2016) [arXiv:1507.06942 [hep-ph]].
[24] S. P. Jones, A. D. Martin, M. G. Ryskin and T. Teubner, Eur. Phys. J. C **76**, no.11, 633 (2016) [arXiv:1610.02272 [hep-ph]].
[25] V. Guzey and M. Klasen, Phys. Rev. C **99**, no.6, 065202 (2019) [arXiv:1811.10236 [hep-ph]].
[26] V. Guzey and M. Klasen, JHEP **04**, 158 (2016) [arXiv:1603.06055 [hep-ph]].
[27] [ATLAS], "Photo-nuclear dijet production in ultra-peripheral Pb+Pb collisions," ATLAS-CONF-2017-011.
[28] [ATLAS], "Photo-nuclear jet production in ultra-peripheral Pb+Pb collisions at $\sqrt{s}_{\rm NN} = 5.02$ TeV with the ATLAS detector," ATLAS-CONF-2022-021.
[29] B. Pire, L. Szymanowski and J. Wagner, Phys. Rev. D **79**, 014010 (2009) [arXiv:0811.0321 [hep-ph]].
[30] W. Schafer, G. Slipek and A. Szczurek, Phys. Lett. B **688**, 185-191 (2010) [arXiv:1003.0610 [hep-ph]].
[31] G. M. Peccini, L. S. Moriggi and M. V. T. Machado, Phys. Rev. D **103**, no.5, 054009 (2021) [arXiv:2101.08338 [hep-ph]].
[32] S. R. Klein, J. Nystrand and R. Vogt, Phys. Rev. C **66**, 044906 (2002) [arXiv:hep-ph/0206220 [hep-ph]].
[33] V. P. Goncalves, M. V. T. Machado and A. R. Meneses, Phys. Rev. D **80**, 034021 (2009) [arXiv:0905.2067 [hep-ph]].
[34] V. P. Gonçalves, G. Sampaio dos Santos and C. R. Sena, Nucl. Phys. A **976**, 33-45 (2018) [arXiv:1711.04497 [hep-ph]].
[35] V. A. Abramovsky, V. N. Gribov and O. V. Kancheli, Yad. Fiz. **18**, 595-616 (1973), Sov. J. Nucl. Phys. 18, 308-317 (1974).
[36] D. Treleani, Int. J. Mod. Phys. A **11**, 613-654 (1996)
[37] J. Jalilian-Marian and Y. V. Kovchegov, Phys. Rev. D **70**, 114017 (2004) [erratum: Phys. Rev. D **71**, 079901 (2005)] [arXiv:hep-ph/0405266 [hep-ph]].
[38] F. Gelis and R. Venugopalan, Nucl. Phys. A **776**, 135-171 (2006) [arXiv:hep-ph/0601209 [hep-ph]].
[39] A. Kovner and M. Lublinsky, JHEP **11**, 083 (2006) [arXiv:hep-ph/0609227 [hep-ph]].
[40] N. N. Nikolaev and W. Schafer, Phys. Rev. D **74**, 074021 (2006) [arXiv:hep-ph/0607307 [hep-ph]].
[41] L. L. Frankfurt and M. I. Strikman, Phys. Lett. B **382**, 6-12 (1996)
[42] L. Frankfurt and M. Strikman, Eur. Phys. J. A **5**, 293-306 (1999) [arXiv:hep-ph/9812322 [hep-ph]].
[43] V. N. Gribov, Sov. Phys. JETP **29**, 483-487 (1969) [Zh. Eksp. Teor. Fiz. **56**, 892-901 (1969)]
[44] V. N. Gribov, Sov. Phys. JETP **30**, 709-717 (1970) [Zh. Eksp. Teor. Fiz. **57**, 1306-1323 (1969)]
[45] G. Piller and W. Weise, Phys. Rept. **330**, 1-94 (2000) [arXiv:hep-ph/9908230 [hep-ph]].
[46] A. Adeluyi and G. Fai, Phys. Rev. C **74**, 054904 (2006) [arXiv:hep-ph/0610214 [hep-ph]].
[47] A. Adeluyi and T. Nguyen, Phys. Rev. C **75**, 054911 (2007)
[48] J. C. Collins, Phys. Rev. D **57**, 3051-3056 (1998) [erratum: Phys. Rev. D **61**, 019902 (2000)] [arXiv:hep-ph/9709499 [hep-ph]].
[49] L. Frankfurt, V. Guzey and M. Strikman, Phys. Rev. D **71**, 054001 (2005) [arXiv:hep-ph/0303022 [hep-ph]].
[50] L. Bertocchi and D. Treleani, J. Phys. G **3**, 147 (1977)
[51] M. Alvioli, L. Frankfurt, V. Guzey, M. Strikman and M. Zhalov, Phys. Lett. B **767**, 450-457 (2017) [arXiv:1605.06606 [hep-ph]].
[52] V. Franco and R. J. Glauber, Phys. Rev. **142**, 1195-1214 (1966)
[53] A. Bialas, M. Bleszynski and W. Czyz, Nucl. Phys. B **111**, 461-476 (1976)
[54] L. Frankfurt, V. Guzey, A. Stasto and M. Strikman, Rept. Prog. Phys. **85**, no.12, 126301 (2022) [arXiv:2203.12289 [hep-ph]].
[55] K. A. Olive *et al.* [Particle Data Group], Chin. Phys. C **38**, 090001 (2014)
[56] T. J. Chapin, R. L. Cool, K. A. Goulianos, K. A. Jenkins, J. P. Silverman, G. R. Snow, H. Sticker, S. N. White and Y. H. Chou, Phys. Rev. D **31**, 17-30 (1985)
[57] M. Alvioli, H. J. Drescher and M. Strikman, Phys. Lett. B **680**, 225-230 (2009) [arXiv:0905.2670 [nucl-th]].
[58] M. Alvioli, C. Ciofi degli Atti, I. Marchino, V. Palli and H. Morita, Phys. Rev. C **78**, 031601 (2008) [arXiv:0807.0873 [nucl-th]].
[59] M. Alvioli, C. Ciofi degli Atti and H. Morita, Phys. Rev. Lett. **100**, 162503 (2008)
[60] S. Acharya *et al.* [ALICE], JHEP **08**, 086 (2022) [arXiv:2107.10757 [nucl-ex]].
[61] M. R. Adams *et al.* [E665], Phys. Rev. Lett. **74**, 5198-5201 (1995) [erratum: Phys. Rev. Lett. **80**, 2020-2021 (1998)]
[62] M. Strikman, M. G. Tverskoii and M. B. Zhalov, Phys. Lett. B **459**, 37-42 (1999) [arXiv:nucl-th/9806099 [nucl-th]].
[63] A. B. Larionov and M. Strikman, Phys. Rev. C **101**, no.1, 014617 (2020) [arXiv:1812.08231 [hep-ph]].
[64] M. Alvioli and M. Strikman, Phys. Rev. C **83**, 044905 (2011) [arXiv:1008.2328 [nucl-th]].
[65] W. Chang, E. C. Aschenauer, M. D. Baker, A. Jentsch, J. H. Lee, Z. Tu, Z. Yin and L. Zheng, Phys. Rev. D **106**, no.1, 012007 (2022) [arXiv:2204.11998 [physics.comp-ph]].
[66] G. Aad *et al.* [ATLAS], Eur. Phys. J. C **76**, no.4, 199 (2016) [arXiv:1508.00848 [hep-ex]].
[67] M. Alvioli, L. Frankfurt, V. Guzey and M. Strikman, Phys. Rev. C **90**, 034914 (2014) [arXiv:1402.2868 [hep-ph]].
[68] M. Alvioli, M. Azarkin, B. Blok and M. Strikman, Eur. Phys. J. C **79**, no.6, 482 (2019) [arXiv:1901.11266 [hep-ph]].

arXiv:2402.19048v1 [hep-ph] 29 Feb 2024

# Nuclear medium meson structures from the Schwinger proper-time Nambu–Jona-Lasinio model

Geoffry Gifari,[1,*] Parada T. P. Hutauruk ,[2,†] and Terry Mart [1,‡]

[1] *Departemen Fisika, FMIPA, Universitas Indonesia, Depok 16424, Indonesia*
[2] *Department of Physics, Pukyong National University (PKNU), Busan 48513, Korea*

(Dated: March 1, 2024)

In this paper, we study the nuclear medium modifications of the meson electromagnetic form factors in the framework of the Nambu–Jona-Lasinio (NJL) model with the help of the Schwinger proper-time regularization scheme to tame loop divergence and simulate the effect of QCD confinement. In our current approach, the meson structure and nuclear medium are constructed in the same NJL model at the quark level. We examine the free space and in-medium charge radii for the kaon and pion, in addition to the spacelike elastic electromagnetic form factors of the kaon and pion as well as their quark-sector form factors, which reflect their internal structure. By comparing to the experimental data, we found that the free space elastic electromagnetic form factors for the mesons are consistent with the data, while the in-medium elastic electromagnetic form factors of the mesons are found to decrease as the nuclear matter density increases, leading to increase of meson charge radius, which is consistent with the prediction of other theory calculations. We also predict the axial nucleon coupling constant $g_A$ in nuclear medium computed via the Goldberger-Treiman relation (GTR), which is crucial for the search for the neutrinoless double beta decay ($0\nu\beta\beta$).

## I. INTRODUCTION

Kaons and pions play an important role in the description of the low-energy dynamics and properties of the strong interactions nonperturbative QCD [1] such as color QCD confinement and dynamical chiral symmetry breaking (D$\chi$SB), which are related to the so-called hadron mass generation or emergent hadron mass (EHM) [2]. In the nuclear medium, it is expected the partial restoration of the chiral symmetry breaking to occur at higher nuclear matter density. However, the question of how the restoration mechanism happens in the nuclear medium and how it will affect the structure of hadron remains unresolved and is still poorly understood. It is widely known that not only the weak properties of hadrons are modified in a nuclear medium, but also the structure of hadrons is expected to change, as observed for the first time by the Stanford Linear Accelerator Center (SLAC) experiment, which is then so-called as the European Muon Collaboration (EMC) effect [3, 4].

In the past many studies on kaon and pion properties in nuclear medium have been made within the various theoretical models, such as the hybrid light-front–quark-meson coupling (LF-QMC) model [5, 6], the QCD sum rules [7], the hybrid light-front constituent quark–quark-meson coupling (LFCQM-QMC) model [8], and the hybrid Nambu–Jona-Lasino–quark meson coupling (NJL-QMC) model [9, 10]. In the literature, most of those studies and attempts have made use of the hybrid model, meaning in their calculation they combined two different models in calculating the nuclear matter and the structure of hadron. For instance, the authors of Ref. [5] studied the pion structure using the LF-QMC model, where the LF model is used to calculate the pion structure in a relativistic manner, on one hand, and the QMC model is used to calculate the quark mass properties in the nuclear medium, on the other hand. It is worth mentioning that in the QMC model, the light quark is coupled to the strong scalar and vector meson mean fields to yield the medium quark masses. However, the dynamical spontaneous chiral symmetry breaking that dynamically generates the constituent quark mass is not clearly described in both LF and QMC models. The authors in Ref. [11] have tried to capture this quark condensate in the QMC model through the NJL model, where the dynamical symmetry breaking is well explained in the model. In the present study, we consistently calculate the meson elastic electromagnetic form factors and nuclear matter in the NJL model, where, in the NJL model, dynamical spontaneous chiral symmetry breaking and its partial restoration in the nuclear medium are captured in the chiral quark condensate as the order parameter quantity.

In this paper, we evaluate the kaon and pion properties and their medium modifications of the electromagnetic form factors using the NJL model with the help of the Schwinger proper-time regularization scheme. In the calculation, besides the meson structure, the nuclear medium is also computed using the NJL model. The NJL model is an uncomplicated model and very useful tool. The NJL model has been widely and successfully used to describe number of physics phenomena such as kaon and pion parton distribution functions [12], gluon distribution functions for the kaon and pion [13, 14], gluon EMC effects in nuclear matter [15, 16], color superconducting in neutron star [17–19], charge symmetry breaking effect in the pion and kaon structures [20], transverse momentum dependent (TMD) quark distribution function [21], transverse momentum dependent fragmenta-

*Electronic address: geoffrygifari.work@gmail.com
†Electronic address: phutauruk@pknu.ac.kr
‡Electronic address: terry.mart@sci.ui.ac.id

tion functions [22], pion cloud effect on the nucleon form factors [23]. Very recently, the NJL model was used to study the charge symmetry-breaking effect on the parton distribution functions [24] and pseudoscalar properties at finite nuclear matter density and temperature [25].

In this work, we found that the in-medium modifications of the electromagnetic form factors for kaon in nuclear matter decrease as the nuclear matter density increases, as expected. Oppositely, we found that the charge radii of the pion and kaon increase as the nuclear matter density increases, which is consistent with other predictions [6, 7, 9, 10]. At nuclear matter density $\rho_B = \rho_0$, it is estimated that the enhancements of the charge radii for the pion, positively charged, and neutral kaon are approximately 10.33%, 8.87%, and 21.69%, respectively, in comparison to those in free space. Similarly for the elastic electromagnetic form factors, the decay constant and mass of the pion decrease as the nuclear density increases. On the contrary, the kaon decay constant and mass in the nuclear medium are found to increase as the nuclear matter density increases.

In the experiment sector, the results of our current study provide very useful information for future experiments that plan to measure the properties of the quark and meson in the higher nuclear matter density (nuclear chemical potential) at Compressed Baryonic Matter Experiment (CBM) at Antiproton and Ion Research (FAIR) in Germany, the Nuclotron Ion Collider Facility (NICA) at the JINR in Rusia, and at (Japan Proton Accelerator Research Complex (J-PARC) in Japan. In addition, the results for the weak axial-vector coupling $g_A$ in the nuclear medium are very important for the search for neutrinoless double beta decay ($0\nu\beta\beta$) and weak interaction nuclear processes [26–28].

This paper is organized as follows. In Sec. II, we briefly describe the NJL model in free space with the help of the Schwinger proper-time regularization scheme to simulate QCD confinement. In Sec. III, we present the in-medium modifications of the pion and kaon properties that are computed within the NJL model. Section IV presents the symmetric nuclear matter properties in the NJL model. In Sec. V, we present the in-medium modifications of the kaon and pion form factors that are calculated in the NJL model. In Sec. VI, we present the numerical results for the kaon and pion form factors in nuclear medium and free space. Summary and conclusion are devoted in Sec. VII.

## II. FREE SPACE MESON PROPERTIES IN THE NJL MODEL

In this section, the generic expression of the three flavors NJL model Lagrangian in free space is presented, in addition to the free space quark and meson properties. All these quantities are needed to compute before computing the medium modifications of meson form factors.

### A. NJL Lagrangian

The three flavors NJL model Lagrangian in terms of the four-fermion interactions, where the gluon interaction was integrated out, is given by

$$\begin{aligned}\mathcal{L}_{\rm NJL} &= \bar{\psi}_q\left(i\partial\!\!\!/ - \hat{m}\right)\psi_q \\ &+ G_\pi\Big[\left(\bar{\psi}_q\lambda_a\psi_q\right)^2 - \left(\bar{\psi}_q\lambda_a\gamma_5\psi_q\right)^2\Big] \\ &+ G_V\Big[\left(\bar{\psi}_q\lambda_a\gamma^\mu\psi\right)^2 + \left(\hat{\psi}_q\lambda_a\gamma^\mu\gamma_5\psi_q\right)^2\Big], \quad (1)\end{aligned}$$

where $\psi_q$ represents the quark field with flavor $q = (u, d, s)$ and $\hat{m} = \mathrm{diag}(m_u, m_d, m_s)$ is the current quark mass matrix. $\lambda_1, \cdots, \lambda_8$ are the Gell-Mann matrices in flavor space with $\lambda_0 \equiv \sqrt{\frac{1}{3}}\mathbf{1}$. It is worth noting that the four-fermion interaction term is proportional to $G_\pi$ coupling constant, which shows the direct terms of the antiquark-quark interaction in the scalar and pseudoscalar channels and has responsibility for the dynamical chiral symmetry breaking (dressed quark mass generation). $G_V$ is the vector coupling constant. Using the mean-field approach, the dressed quark masses are determined through the quark self-energy interaction. The gap equation can be defined in the Schwinger proper-time regularization scheme and it gives

$$\begin{aligned}M_q &= m_q - 4G_\pi\langle\bar{\psi}_q\psi_q\rangle, \\ &= m_q + \frac{3G_\pi M_q}{\pi^2}\int_{\tau_{\rm UV}}^{\tau_{\rm IR}}\frac{d\tau}{\tau^2}e^{-\tau M_q^2}, \quad (2)\end{aligned}$$

where $\langle\bar{\psi}_q\psi_q\rangle$ is the chiral quark condensate, which is the order parameter of chiral symmetry breaking. $\tau_{\rm IR} = 1/(\Lambda_{\rm IR})^2$ and $\tau_{\rm UV} = 1/(\Lambda_{\rm UV})^2$ stand for the upper and lower limits of integration, respectively, where $\Lambda_{\rm IR}$ and $\Lambda_{\rm UV}$ are the infrared and ultraviolet cutoffs.

### B. Kaon and pion properties

The kaon and pion $t$-matrices can be described by the dressed quark and dressed antiquark interaction in the kaon and pion channels in the random phase approximation (RPA), equivalently to the ladder approximation. Summation of the bubble diagram from infinite interaction can be simply written by

$$\tau_\pi = \gamma_5\lambda_a\left(\frac{-2iG_\pi}{1+2G_\pi\Pi_\pi(p^2)}\right)\gamma_5\lambda_a, \quad (3)$$

$$\tau_K = \gamma_5\lambda_a\left(\frac{-2iG_\pi}{1+2G_\pi\Pi_K(p^2)}\right)\gamma_5\lambda_a, \quad (4)$$

where, for the pion, the sum over $a = 1, 2, 3$ and for the kaon $a = 4, 5, 6, 7$. The polarization insertion propaga-

tors for the pion and kaon are respectively given by

$$\begin{aligned}\Pi_\pi(p^2)\delta_{ab} &= i\int\frac{d^4k}{(2\pi)^4}\mathrm{Tr}\Big[\gamma_5\lambda_a S_l(p+k)\gamma_5\lambda_b S_l(k)\Big],\\ \Pi_K(p^2)\delta_{ab} &= i\int\frac{d^4k}{(2\pi)^4}\mathrm{Tr}\Big[\gamma_5\lambda_a S_l(p+k)\gamma_5\lambda_b S_s(k)\Big],\end{aligned} \tag{5}$$

where the trace is taken for the Dirac, flavor, and color space.

### C. Kaon and pion masses

The kaon and pion masses can be straightforwardly determined by the pole position of the corresponding $T$ matrix in Eqs. (3) and (4) and it gives

$$1+2G_\pi\Pi_\pi(p^2=m_\pi^2) = 0, \tag{6}$$
$$1+2G_\pi\Pi_K(p^2=m_K^2) = 0, \tag{7}$$

where near the bound state pole position of the $T$ matrix, it is given by

$$T_{\pi,K}\simeq\gamma_5\lambda_a\left(\frac{ig_{mq}^2}{p^2-m_k^2+i\epsilon}\right)\gamma_5\lambda_a. \tag{8}$$

With $g_{mq}$ represents the meson-quark coupling constant. To determine the meson-quark coupling constant, we expand the $t-$matrix in Eq. (3)at around the pole $p^2=m_{\pi,K}^2$ and it gives

$$\begin{aligned}\Pi_{\pi,K}(p^2) &= \Pi_{\pi K}(m_{\pi K}^2)+\left.\frac{\partial\Pi_{\pi K}(p^2)}{\partial p^2}\right|_{p^2=m_{\pi K}^2}\\ &\times\ (p^2-m_{\pi K}^2)+\cdots,\end{aligned} \tag{9}$$

we then can define the meson-quark coupling constant by

$$Z_K^{-1}=\Big[g_{\pi qq}\Big]^{-2} = -\left.\frac{\partial\Pi_\pi(p^2)}{\partial p^2}\right|_{p^2=m_\pi^2}, \tag{10}$$

$$Z_\pi^{-1}=\Big[g_{Kqq}\Big]^{-2} = -\left.\frac{\partial\Pi_K(p^2)}{\partial p^2}\right|_{p^2=m_K^2}. \tag{11}$$

### D. Meson weak-decay constants

The kaon and pion weak-decay constants are defined by the matrix element of the meson to hadronic in a vacuum, which is given by $\langle 0|j_a^\mu(0)|\pi(K)_b(p)\rangle$ with $j_a^\mu$ stands for the axial-vector current operator for flavor $a$. Thus, the kaon and pion weak-decay constants are respectively given by

$$\langle 0|j_a^\mu(0)|\pi_b(p)\rangle = ip^\mu f_\pi\delta_{ab}, \tag{12}$$
$$\langle 0|j_a^\mu(0)|K_b(p)\rangle = ip^\mu f_K\delta_{ab}, \tag{13}$$

where in the NJL model, it can be defined by

$$\begin{aligned}ip^\mu f_\pi\delta_{ab} &= -\frac{g_{\pi qq}}{2}\int\frac{d^4k}{(2\pi)^4}\\ &\times\ \mathrm{Tr}\Big[\gamma^\mu\gamma_5\lambda_a S_l(k+p)\gamma_5\lambda_b S_l(k)\Big],\end{aligned} \tag{14}$$

$$\begin{aligned}ip^\mu f_K\delta_{ab} &= -\frac{g_{Kqq}}{2}\int\frac{d^4k}{(2\pi)^4}\\ &\times\ \mathrm{Tr}\Big[\gamma^\mu\gamma_5\lambda_a S_l(k+p)\gamma_5\lambda_b S_s(k)\Big].\end{aligned} \tag{15}$$

We then apply the Feynman parameterization, and Wick rotation, and introduce the Schwinger proper-time regularization scheme. The final expression of the kaon and pion decay constants are respectively given by

$$\begin{aligned}f_K &= \frac{3g_{Kqq}}{4\pi^2}\int_0^1 dx\int_{\tau_{\rm UV}}^{\tau_{\rm IR}}\frac{d\tau}{\tau}\Big[M_s+x(M_l-M_s)\Big]\\ &\times\ e^{-\tau\Big(M_s^2-x(M_s^2-M_l^2)-x(1-x)m_K^2\Big)},\end{aligned} \tag{16}$$

$$f_\pi = \frac{3g_{\pi qq}M_l}{4\pi^2}\int_0^1 dx\int_{\tau_{\rm UV}}^{\tau_{\rm IR}}\frac{d\tau}{\tau}e^{-\tau\Big(M_l^2-x(1-x)m_\pi^2\Big)}. \tag{17}$$

Next, we will extend these free space NJL expressions into the nuclear medium in Sec. III.

## III. KAON AND PION PROPERTIES IN NUCLEAR MEDIUM

Here, we describe the in-medium properties of the quark and kaon, which is an extension of the kaon and pion properties in Sec. II. The NJL dynamical quark masses in the nuclear medium in the Schwinger proper-time regularization scheme are given by

$$M_q^* = m_q^*+\frac{3G_\pi M_q^*}{\pi^2}\int_{\tau_{\rm UV}}^\infty d\tau\frac{e^{\tau M_q^{*2}}}{\tau^2}, \tag{18}$$

where $M_q^*$, and $m_q^*$ represent the effective, and current quark masses, respectively. The coupling constant of $G_\pi$ is the four-fermion coupling constant, where the values are the same as in the free space. It is worth noting that the $\tau_{\rm IR}$ is taken to infinity in the nuclear medium. The in-medium dressed up and strange quark propagators in Eq. (18) are respectively given by

$$S_l^*(k^*) = \frac{\not{k}^*+M_l^*}{k^{*2}-M_l^{*2}+i\epsilon}, \tag{19}$$

$$S_s^*(k^*) = S_s(k)=\frac{\not{k}+M_s}{k^2-M_s^2+i\epsilon}, \tag{20}$$

where the asterisk sign in Eq. (19) denotes the nuclear medium effect, which enters the light quark momentum $k^\mu$ by $k^{*\mu}=k^\mu+V^\mu$. This is due to the vector meson

mean field $V^\mu = \left(V^0, \mathbf{0}\right)$. Note that the medium modifications of the space component of the light quark momentum of $k^{*\mu}$ can be ignored since the size contribution is insignificant [29]. Substituting the in-medium modifications of the quark propagator into the constituent quark mass formula in Eq. (18), we then apply the shift variable of the integral to eliminate the vector potential that enters the light quark momentum. For the strange quark, it is widely known that the strange quark is weakly coupled to the scalar and vector mean fields in the NJL nuclear matter, so it is reasonably assumed that the in-medium modifications of the strange quark propagator are similar to those in the free space, implying that the in-medium constituent quark mass is the same as that in free space ($M_s^* = M_s$).

Now we turn to present the description of the bound state of the kaon and pion, which is obtained by elucidating the Bethe-Salpeter equation (BSE) in random phase approximation (RPA). The solution of the BSE in the pion and kaon channels is determined by the two-body scattering $t$-matrix. The reduced $t$-matrices in the pion and kaon mesons and vector mesons are respectively obtained by

$$
\begin{aligned}
t_K^*(p^*) &= \frac{-2iG_\pi}{1+2G_\pi\Pi_K^*(p^{*2})}, && (21)\\
t_\pi^*(p^*) &= \frac{-2iG_\pi}{1+2G_\pi\Pi_\pi^*(p^{*2})}, && (22)\\
t_V^{*\mu\nu}(p^*) &= \frac{-2iG_\rho}{1+2G_\rho\Pi_V^*(p^{*2}} \\
&\times \left[g^{\mu\nu}+2G_\rho\Pi_V^*(p^{*2})\frac{p^{*\mu}p^{*\nu}}{p^2}\right], && (23)
\end{aligned}
$$

where $G_\rho$ is the four-fermion coupling constant for the vector meson channels and the polarization insertions, which are the so-called bubble diagrams in the nuclear medium can be introduced by

$$
\begin{aligned}
\Pi_K^*(p^{*2}) &= 6i\int\frac{d^4k}{(2\pi)^4}\mathrm{Tr}\Big[\gamma_5 S_l^*(k^*)\gamma_5 S_s^*(k^*+p^*\Big], \\
\Pi_\pi^*(p^{*2}) &= 6i\int\frac{d^4k}{(2\pi)^4}\mathrm{Tr}\Big[\gamma_5 S_l^*(k^*)\gamma_5 S_l^*(k^*+p^*\Big], \\
\Pi_V^{ll}(p^{*2})P_T^{\mu\nu} &= 6i\int\frac{d^4k}{(2\pi)^4}\mathrm{Tr}\Big[\gamma^\mu S_l^*(k^*)\gamma^\nu S_l^*(k^*+p^*)\Big], \qquad (24)
\end{aligned}
$$

where $P_T^{\mu\nu} = g^{\mu\nu} - q^\mu q^\nu/q^2$, Tr are only included for the Dirac indices. It is worth noting that the polarization insertion for the light quarks, $\Pi_\rho^* = \Pi_\omega^* = \Pi_V^{*ll}$, while for the strange quarks, $\Pi_\phi^* = \Pi_V^{*ss}$.

After extracting the polarization insertion for the pion and kaon, we can also straightforwardly determine the in-medium meson masses for the kaon and pion from the pole position of the corresponding $t$-matrix, which are given by

$$
\begin{aligned}
1+2G_\pi\Pi_K^*(p^{*2}=m_K^{*2}) &= 0, && (25)\\
1+2G_\pi\Pi_\pi^*(p^{*2}=m_K^{*2}) &= 0. && (26)
\end{aligned}
$$

Next, using the polarization insertions of the kaon and pion, other properties, such as the kaon-quark coupling constant in medium and pion-quark coupling constant in nuclear medium, can be easily computed by taking the first derivative of the polarization insertion of the kaon and pion, respectively, with $p^{*2} = m_K^{*2}$. Mathematically, it can be formulated by

$$
\begin{aligned}
[Z_K^*]^{-1} = \Big[g_{Kqq}^*\Big]^{-2} &= -\frac{\partial\Pi_K^*(p^{*2})}{\partial p^{*2}}\bigg|_{p^{*2}=m_K^{*2}}, && (27)\\
[Z_\pi^*]^{-1} = \Big[g_{\pi qq}^*\Big]^{-2} &= -\frac{\partial\Pi_\pi^*(p^{*2})}{\partial p^{*2}}\bigg|_{p^{*2}=m_\pi^{*2}}. && (28)
\end{aligned}
$$

## IV. NUCLEAR MATTER IN THE NJL MODEL

In this section, we present the symmetric nuclear matter NJL model built at the quark level, which is adapted from Ref. [18]. After applying the quark bilinear into NJL Lagrangian in Eq. (1), the NJL effective Lagrangian in the symmetric nuclear matter (NM) is given by

$$
\begin{aligned}
\mathcal{L}_{\mathrm{NM-NJL}} &= \bar\psi\left(i\partial\!\!\!/ - M - V\!\!\!/\right)\psi - \frac{(M-m)^2}{4G_\pi} \\
&+ \frac{V_\mu V^\mu}{2G_\omega} + \mathcal{L}_I, && (29)
\end{aligned}
$$

where $\mathcal{L}_I$ represents the interaction Lagrangian. The isoscalar-vector mean field and the effective constituent quark mass can be respectively defined by

$$
\begin{aligned}
V^\mu &= 2G_\omega\langle\rho|\bar\psi_q\gamma^\mu\psi_q|\rho\rangle = 2\delta^{0\mu}G_\omega\langle\psi_q^\dagger\psi_q\rangle, && (30)\\
M &= m - 2G_\pi\langle\rho|\bar\psi_q\psi_q|\rho\rangle, && (31)
\end{aligned}
$$

where the vector potential is defined by $V = (V_0, \mathbf{0})$. Using the standard hadronization technique, the expression of the effective potential for symmetric nuclear matter can be determined from the NJL Lagrangian. In the mean-field approximation, the effective potential for NM is given by

$$
\mathcal{E} = \mathcal{E}_V - \frac{V_0^2}{4G_\omega} + 4\int\frac{d^3p}{(2\pi)^3}\Theta(P_F - |\mathbf{p}|)\epsilon_N, \qquad (32)
$$

where $\epsilon_N = \sqrt{M_N^{*2}+\mathbf{p}^2}+3V_0 \equiv E_N+3V_0$. The last term is the Fermi motion of nucleons in the scalar and vector mean fields. The vacuum contribution of the quark loop can be defined by

$$
\begin{aligned}
\mathcal{E}_V &= 12i\int\frac{d^4k}{(2\pi)^4}\ln\left(\frac{k^2-M^2+i\epsilon}{k^2-M_0^2+i\epsilon}\right) \\
&+ \frac{(M-m)^2}{4G_\pi} - \frac{(M_0-m)^2}{4G_\pi}. && (33)
\end{aligned}
$$

Next, using the minimum condition $\partial\mathcal{E}/\partial V_0 = 0$, the value of $V_0$ is obtained as

$$
V_0 = 6G_\omega\rho_B, \qquad (34)
$$

where $\rho_B = 2p_F^3/3\pi^2$ represents the baryon density with $p_F$ is the Fermi momentum of the baryon. It is worth noting that the constituent quark mass for fixed baryon density must satisfy the minimum condition $\partial\mathcal{E}/\partial M_q = 0$ to give a similar expression of the in-medium constituent quark mass in Eq. (30).

Using the expression of effective potential, we determine the binding energy per nucleon for symmetric nuclear matter, which is given by

$$\frac{E_B}{A} = \frac{\mathcal{E}}{\rho_B} - M_{\rm N0}, \tag{35}$$

where $M_{\rm N0}$ stands for the free space nucleon mass. Our symmetry nuclear matter NJL model reproduces well the binding energy per nucleon $E_B/A = -15.7$ MeV at saturation density $\rho_0 = 0.16$ fm$^{-3}$. Using the quark properties calculated in the nuclear matter NJL model, we obtain the quantities such as dressed quark masses, kaon and pion masses, pion and kaon quark coupling constants, and the quark condensates in nuclear medium. A summary of these in-medium modification quantities is tabulated in Table. I. Next, these quantities are needed to compute the kaon and pion electromagnetic form factors in nuclear medium.

## V. NUCLEAR MEDIUM FORM FACTORS

Here we present the electromagnetic form factor for the kaon and pion in the nuclear medium. The electromagnetic form factors can be extracted from the matrix element of the electromagnetic current, which is given by

$$\mathcal{J}^\mu(p^{'*}, p^*) = \left(p^{'*\mu} + p^{*\mu}\right) F_K^*(Q^2), \tag{36}$$

where $F_K^*(Q^2)$ is the medium modifications of the kaon form factor and it can be written in terms of the quark sector form factors,

$$F_K^*(Q^2) = e_u F_K^{*u}(Q^2) + e_d F_K^{*d}(Q^2) + e_s F_K^{*s}(Q^2) + \cdots. \tag{37}$$

Analogously, it can be applied for the pion case by replacing the strange quark with the light quark.

Following the free space calculation [12], the in-medium modifications kaon and pion form factors can be computed with the sum of two Feynman diagrams, as used in Ref. [12]. It is worth noting that we consider the in-medium dressed quark-photon vertex $\Lambda^{*\mu}_{\gamma q} = \gamma^\mu F^*_{1Q}(Q^2)$, rather than in medium bare quark-photon vertex $\Lambda^{*\mu}_{\gamma q} = \hat{Q}\gamma^\mu$. This form satisfies the Ward-Takahashi identity (WTI) defined by

$$q_\mu \Lambda^{*\mu}_{\gamma Q}(p^{'*}, p^*) = e_Q\Big[S^{*-1}_Q(p^{'*}) - S^{*-1}_Q(p^*)\Big]. \tag{38}$$

The dressed quarks for up and strange quarks are respectively defined by

$$F^*_{1U}(Q^2) = \frac{e_u}{1 + 2G_\rho \Pi^{*ll}_V(Q^2)}, \tag{39}$$

$$F^*_{1D}(Q^2) = \frac{e_d}{1 + 2G_\rho \Pi^{*ll}_V(Q^2)}, \tag{40}$$

$$F^*_{1S}(Q^2) = \frac{e_s}{1 + 2G_\rho \Pi^{*ss}_V(Q^2)}, \tag{41}$$

where the in-medium bubble diagrams are the same as defined in Eq. (24).

After performing the calculation such as Feynman parameterization and applying the Schwinger proper-time regularization scheme, the final expression for the in-medium kaon and pion form factors are respectively given by

$$\begin{aligned} F^*_K(Q^2) &= F^*_{1U}(Q^2) f^{*ls}_K(Q^2) - F^*_{1S}(Q^2) f^{*sl}_K(Q^2), \\ F^*_\pi(Q^2) &= F^*_{1U}(Q^2) f^{*ll}_\pi(Q^2) - F^*_{1D}(Q^2) f^{*ll}_\pi(Q^2), \end{aligned} \tag{42}$$

where the $f^{*ls}_K(Q^2)$ and $f^{*ll}_\pi(Q^2)$ are respectively the bare kaon and pion form factors in the nuclear medium.

## VI. NUMERICAL RESULT

In this section, we present the numerical results for the kaon and pion properties and their elastic and quark sector form factors in the nuclear medium. The NJL parameter used in this present work is $G_\pi$, $G_\omega$, $G_\rho$, $G_a$, and $G_s$, as used in Refs. [12, 13, 17, 18]. In this work, the regularization parameters of the $\Lambda_{\rm IR}$ in the Schwinger proper-time regularization scheme are fixed to $\Lambda_{\rm IR} = 240$ MeV, which is set the same as the order of $\Lambda_{\rm QCD}$ and the constituent quark mass $M_l = 400$ MeV. The remaining parameters are fitted to the pion physical mass $m_\pi = 140$ MeV, kaon physical mass $m_K = 495$ MeV, nucleon mass $M_N = M_{N0} = 940$ MeV, and $\rho$-meson mass $m_\rho = 770$ MeV together with the pion weak-decay constant $f_\pi = 93$ MeV. The fitted results give ultraviolet cutoff $\Lambda_{\rm UV} = 645$ MeV, pion coupling constant $G_\pi = 19.04 \times 10^{-6}$ MeV$^{-2}$, axial-vector diquark coupling constant $G_a = 2.8 \times 10^{-6}$ MeV$^{-2}$, scalar diquark coupling constant $G_s = 7.49 \times 10^{-6}$ MeV$^{-2}$, and strange constituent quark mass $M_s = 611$ MeV. The $\phi$ meson mass is $m_\phi = 1001$ MeV, scalar diquark mass $M_s = 687$ MeV, and axial diquark mass $M_a = 1027$ MeV. It is worth noting that the $G_s$ and $G_a$ are determined using the Faddeev equation to produce the vacuum nucleon mass $M_N$ and axial coupling constant of nucleon $g_A = 1.267$, as used in Ref. [17, 18]. Symmetric nuclear matter parameters are determined by fitting the binding energy per nucleon $E_B/A = -15.7$ MeV at saturation baryon density $\rho_0 = 0.16$ fm$^{-3}$, it gives $G_\omega = 6.03 \times 10^{-6}$ MeV$^{-2}$. The current quark mass for the up quark is given by $m_u = 16$ MeV and the current quark mass for the strange quark is obtained $m_s = 356$ MeV, as used in Ref. [12, 13, 17, 18].

TABLE I: The in-medium modifications of the constituent quark mass, kaon mass, kaon decay constant, kaon-quark coupling constants, the constituent quark mass, pion mass, pion decay constant, pion-quark coupling constants, quark condensate, and effective nucleon mass are calculated in the NJL model. The units are in MeV. Note that $\langle \bar{s}s\rangle^{*1/3} = 0.150$ GeV and unit of $\langle \bar{u}u\rangle^{*1/3}$ is in GeV.

| $\rho_B/\rho_0$ | $M_q^*$ | $M_s^*$ | $m_{\rm K}^*$ | $f_K^*$ | $g_{\rm Kqq}^*$ | $m_\pi^*$ | $f_\pi^*$ | $g_{\pi{\rm qq}}^*$ | $-\langle \bar{u}u\rangle^{*1/3}$ | $M_N^*$ | $g_A^*/g_A$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.00 | 400 | 611 | 495 | 97 | 4.572 | 140 | 93 | 4.225 | 0.171 | 934.36 | 1.00 |
| 0.25 | 379 | 611 | 495 | 99 | 4.467 | 137 | 92 | 4.040 | 0.168 | 887.13 | 0.991 |
| 0.50 | 360 | 611 | 496 | 100 | 4.375 | 136 | 91 | 3.882 | 0.165 | 842.54 | 0.992 |
| 0.75 | 343 | 611 | 499 | 101 | 4.296 | 134 | 90 | 3.750 | 0.162 | 806.43 | 0.995 |
| 1.00 | 328 | 611 | 501 | 102 | 4.224 | 133 | 89 | 3.639 | 0.160 | 778.03 | 0.998 |
| 1.25 | 316 | 611 | 504 | 103 | 4.170 | 132 | 87 | 3.555 | 0.158 | 756.01 | 0.989 |
| 1.50 | 306 | 611 | 506 | 104 | 4.123 | 132 | 86 | 3.488 | 0.156 | 738.98 | 0.991 |
| 1.75 | 297 | 611 | 509 | 104 | 4.084 | 132 | 85 | 3.429 | 0.154 | 725.73 | 0.992 |
| 2.00 | 289 | 611 | 511 | 105 | 4.047 | 132 | 84 | 3.379 | 0.153 | 715.34 | 0.993 |

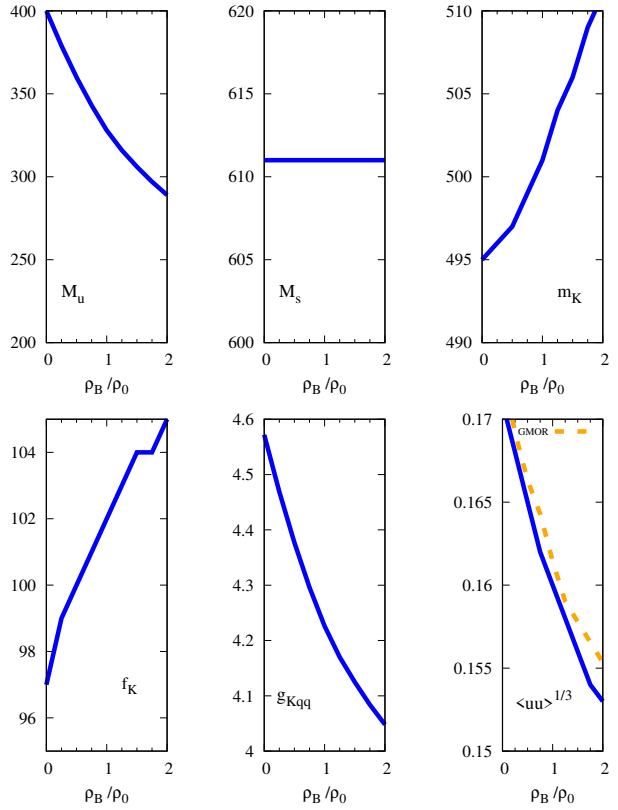



FIG. 1: Kaon and quark properties as a function of $\rho_B/\rho_0$. Top panel (left to right): constituent quark mass of the up quark, constituent quark mass of the strange quark, and kaon mass. Lower panel (left to right): kaon decay constant, kaon-quark coupling constant, and quark condensate in comparison with that calculated via GMOR (dashed line).

### A. Meson properties in the nuclear medium

Figure 1 shows the in-medium modifications of the constituent mass of the up-quark, constituent quark mass of strange-quark, kaon mass, kaon weak-decay constant, kaon-quark coupling constant, and the quark condensate, respectively. The top-left panel of Fig. 1 indicates that the in-medium modifications of the constituent quark mass decrease as the density increases. Relative to the vacuum values, the decreasing rate is approximately 18% at $\rho_B = \rho_0$. In the top-middle panel of Fig. 1, it is shown that the constituent mass of the strange quark does not change as the density increases. This is due to the fact that in the present work we take $M_s^* = M_s$ and the strange quark decoupled to the meson mean fields in the nuclear matter. The parameterization for the in-medium modifications of the constituent mass of the up-quark is given by

$$\frac{M_u^*}{M_u} = 1 - 0.22\left(\frac{\rho_B}{\rho_0}\right) + 0.04\left(\frac{\rho_B}{\rho_0}\right)^2. \qquad (43)$$

The top-right panel of Fig. 1 shows the results for the in-medium modifications of the kaon mass. It indicates that the effective kaon mass slowly increases as the density increases. The increased value of the kaon mass in the nuclear medium is relatively small, which is about 1.2% at $\rho_B = \rho_0$ relative to that in the vacuum. Our prediction for the effective positively charged kaon mass in the nuclear medium is consistent with the theoretical result obtained in Ref. [30]. The parameterization form for the kaon mass in the nuclear medium can be defined by

$$\frac{m_K^*}{m_K} = 1 + 0.009\left(\frac{\rho_B}{\rho_0}\right) + 0.04\left(\frac{\rho_B}{\rho_0}\right)^2. \qquad (44)$$

Results for the kaon weak-decay constant in the nuclear medium are shown in the lower-left panel of Fig. 1. Similarly with the in-medium modifications of the kaon mass, the in-medium modifications of the kaon weak-decay constant increase with increasing densities. The increasing rate at $\rho_B = \rho_0$ is approximately about 5.2% relative to that in a vacuum. The parameterization for the in-medium modifications of the kaon weak-decay constant

is given by

$$\frac{f_K^*}{f_K} = 1 + 0.06\left(\frac{\rho_B}{\rho_0}\right) - 0.01\left(\frac{\rho_B}{\rho_0}\right)^2. \quad (45)$$

For the in-medium modifications of the kaon-quark coupling constant, it is found that the coupling constant decreases as the density increases, as shown in the lower-middle panel of Fig. 1. The decreasing rate of the in-medium kaon-quark coupling constant at $\rho_B = \rho_0$ is about 7.6% relative to that in vacuum. The parameterization for the in-medium modifications of the kaon-quark coupling constant is given by

$$\frac{g_{Kqq}^*}{g_{Kqq}} = 1 - 0.09\left(\frac{\rho_B}{\rho_0}\right) + 0.02\left(\frac{\rho_B}{\rho_0}\right)^2. \quad (46)$$

Results for $\langle\bar{u}u\rangle^{1/3}$ as a function of $\rho_B/\rho_0$ are depicted in the lower-right panel of Fig. 1. The chiral condensate in the nuclear medium decreases as the density decreases as expected. The in-medium modifications of the chiral condensate decrease about 6.4% relative to that in vacuum at $\rho_B = \rho_0$. The parameterization for the in-medium modifications of the kaon-quark coupling constant is given by

$$\frac{\langle\bar{u}u\rangle^*}{\langle\bar{u}u\rangle} = 1 - 0.08\left(\frac{\rho_B}{\rho_0}\right) + 0.01\left(\frac{\rho_B}{\rho_0}\right)^2. \quad (47)$$

In the lower-right panel of Fig. 1, the results of the quark condensate (dashed line) obtained through the Gell-Mann-Oakes-Renner relation (GMOR), which is formulated by

$$f_\pi^2 m_\pi^2 = -\frac{1}{2}(m_u + m_d)\langle\bar{u}u + \bar{d}d\rangle, \quad (48)$$

by plugging the values of the $f_\pi$, $m_\pi$ in Table I, and $m_u = m_d = 16.4$ MeV to Eq. (48) and then the quark condensate for various density can be obtained. The lower-right panel of Fig. 1 clearly shows that the results for the quark condensate from the GMOR in Eq. (48) have a similar tendency compared to that calculated directly in the NJL model. This explicitly indicates that the NJL model with the Schwinger proper-time regularization scheme is consistent with the chiral symmetry in QCD.

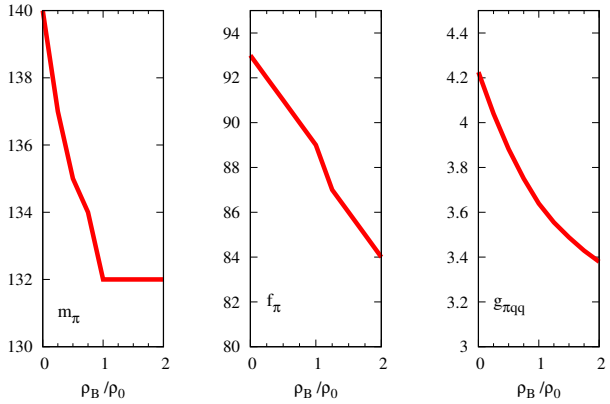



FIG. 2: Pion properties as a function of $\rho_B/\rho_0$. Left to right: pion mass, pion weak-decay constant, and pion-quark coupling constant.

Results for the in-medium modifications of the pion mass are shown in the right panel of Fig. 2. It clearly shows that the pion mass decreases as density increases up to $\rho_B = 0.75\rho_0$ and at a large density of $\rho_B > 0.75\rho_0$, the values of the pion mass are the same. The decreasing rate for this regime is approximately about 5.7% relative to that in a vacuum. The parameterization form for the in-medium modifications of the pion mass is given by

$$\frac{m_\pi^*}{m_\pi} = 1 - 0.08\left(\frac{\rho_B}{\rho_0}\right) + 0.02\left(\frac{\rho_B}{\rho_0}\right)^2. \quad (49)$$

Results for the in-medium pion weak-decay constant are shown in the middle panel of Fig. 2. The in-medium modifications of the pion weak-decay constant decrease as the baryon density increases, which is rather different from the results for the in-medium modifications of the kaon weak-decay constant, as shown in the bottom-left panel of Fig. 1. The suppression of the pion weak-decay constant at $\rho_B = \rho_0$ is approximated at about 4.3% relative to that in the vacuum. The parameterization form for the in-medium modifications of the pion mass is given by

$$\frac{f_\pi^*}{f_\pi} = 1 - 0.05\left(\frac{\rho_B}{\rho_0}\right) - 0.002\left(\frac{\rho_B}{\rho_0}\right)^2. \quad (50)$$

Moreover, we also numerically compute the pion-quark coupling constant in the nuclear medium. Our results for the in-medium modifications of the pion-quark coupling constant are shown in the right panel of Fig. 2. It clearly shows that the in-medium modifications of the pion-quark coupling constant decrease with increasing baryon densities. The decreasing rate is approximately about 13.9% at $\rho_B = \rho_0$ relative to that in the vacuum. The parameterization form for the in-medium modifications of the pion-quark coupling constant is given by

$$\frac{g_{\pi qq}^*}{g_{\pi qq}} = 1 - 0.17\left(\frac{\rho_B}{\rho_0}\right) + 0.04\left(\frac{\rho_B}{\rho_0}\right)^2. \quad (51)$$

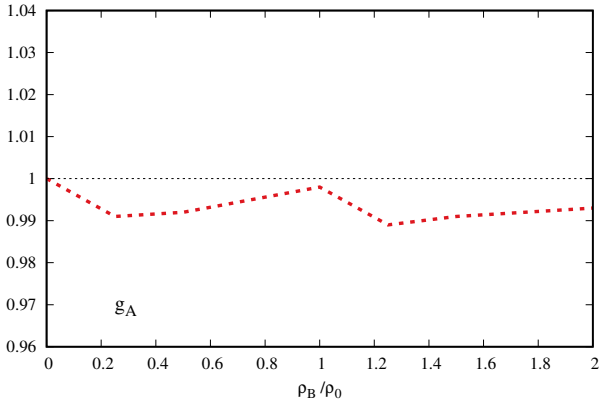



FIG. 3: Quenching factor of the axial-vector coupling constant as a function of $\rho_B/\rho_0$.

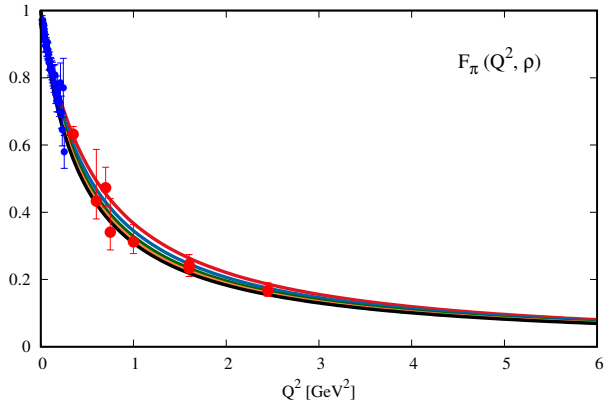



FIG. 4: Pion form factor as a function of $Q^2$ for various nuclear matter densities. The red, blue, green, orange, and black lines represent $\rho_B/\rho_0 = 0.0$ (free space), $\rho_B/\rho_0 = 0.50$, $\rho_B/\rho_0 = 1.00$, $\rho_B/\rho_0 = 1.50$, and $\rho_B/\rho_0 = 2.00$, respectively. Experimental data are taken from Refs. [34, 35].

Using the results for the pion properties in the nuclear medium, we can straightforwardly calculate the quenching of the nuclear weak axial-vector coupling constant through the Goldberger-Treiman relation (GTR). In the nuclear medium, the GTR at the quark level can be simply written as

$$\frac{g_A^*}{g_A} = \left(\frac{g_{\pi qq}^*}{g_{\pi qq}}\right)\left(\frac{f_\pi^*}{f_\pi}\right)\left(\frac{M_q}{M_q^*}\right), \tag{52}$$

where $M_N^*$ and $M_N$ are respectively effective nucleon mass and vacuum nucleon mass. The values of $M_N^*$ calculated in the NJL model for various densities are provided in Table I. Parameterization of the in-medium modifications of the $g_A$ is given by

$$\frac{g_A^*}{g_A} = 1 - 0.007\left(\frac{\rho_B}{\rho_0}\right) + 0.002\left(\frac{\rho_B}{\rho_0}\right)^2. \tag{53}$$

Results for the in-medium modifications of the axial-vector coupling constant for various densities are shown in Fig. 3. It clearly shows that the quenching factor $q = g_A^*/g_A$ monotonically decreases as the nuclear matter density increases. By taking an experimental value of $g_A = 1.267$ in free space, at baryon density $\rho_B = \rho_0$, the reduction of $g_A$ is approximately about 0.2% relative to that in free space, indicating very small quenching factor. This result is consistent with the results obtained in Refs. [31–33]. The values of $g_A^*/g_A$ are also given in Table. I.

### B. Meson structure in the nuclear medium

Here, numerical results for the in-medium modifications of the elastic form factors and charge radius of the pion and kaon in free space and nuclear medium calculated in the NJL model are presented.

#### 1. *Elastic form factors*

Results for the total pion and kaon electromagnetic form factors and their quark sector form factors in the free space and nuclear medium are depicted in Figs. 4-10. Figure 4 shows total pion electromagnetic form factors for different nuclear matter densities. The red solid line represents the free space total pion electromagnetic form factor. The result for the total pion electromagnetic form factor in free space is in excellent agreement with the experimental data given in Refs. [34, 35] and other theoretical results [12]. In Fig. 4, the pion electromagnetic form factors for nuclear matter densities $\rho/\rho_0$ =0.50, 1.00, 1.50, and 2.00, which are represented by blue, green, orange, and black solid lines, respectively are shown. It reveals that, as the nuclear matter density increases, the pion electromagnetic form factors decrease, as expected.

To see more clearly the nuclear medium effect in the pion electromagnetic form factors, we compute the pion electromagnetic form factors multiplied by $Q^2$ as shown in Fig. 5. Also, the result for the up quark sector electromagnetic form factor of the pion multiplied by $e_u Q^2$ is shown in Fig. 6, where $e_u$ is the elementary electric charge for the up quark. The significant contribution of the up quark sector form factor to the total pion electromagnetic form factor can be also seen in Fig. 6. Following the total pion electromagnetic form factor, the up quark sector form factor decreases with the increase of the nuclear matter density. It is worth noting that the pion electromagnetic form factors in free space and nuclear medium reproduce $F_\pi(Q^2,\rho) \approx 1/\mathrm{Q}^2$ for $Q^2 \to \infty$. At higher $Q^2$, the $Q^2F_\pi(Q^2,\rho_B=0)$ begins to show plateau with $Q^2F_\pi(Q^2,\rho_B) \simeq 0.49$ for free space. For $\rho_B/\rho_0 =$ 2.0, $Q^2F_\pi(Q^2,\rho_B) \simeq 0.405$.

Besides the pion electromagnetic form factor in free space and nuclear medium, we now turn to compute the

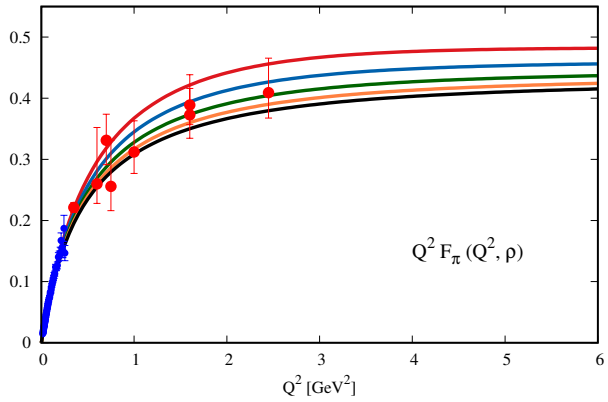



FIG. 5: Same as in Fig. 4, but for $F_\pi(Q^2)$ multiplied by $Q^2$.

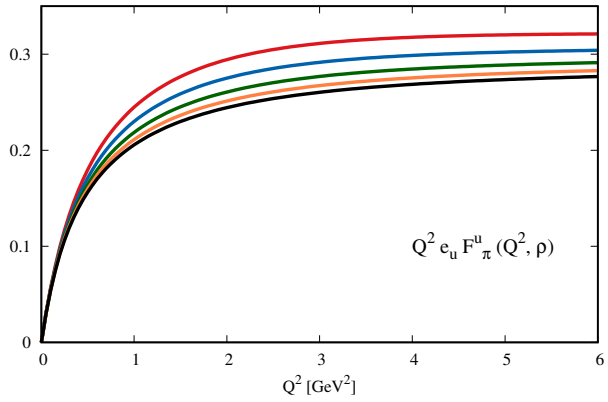



FIG. 6: Same as in Fig. 4, but for $e_u F_\pi^u(Q^2)$ multiplied by $Q^2$.

kaon electromagnetic form factor in the free space and nuclear medium, and results are shown in Fig. 7. In comparison with the low $Q^2$ existing data ($0 < Q^2 < 0.1$ GeV$^2$) [36], the kaon electromagnetic form factor in free space is in excellent agreement with the data [36]. As for the pion case, we also compute the kaon electromagnetic form factor at $\rho_B/\rho_0 = 0.50$ (blue solid line), 1.00 (green solid line). 1.50 (orange solid line), and 2.00 (black solid line), respectively. As the nuclear matter density increases, it clearly shows that the kaon electromagnetic form factor decreases. The tendency of the kaon electromagnetic form factor is almost similar to the case of the pion. However, they differ in the quark sector form factors. For the up quark sector form factors, kaon and pion indicate the same feature, where they increase as the nuclear matter density increases as can be seen in Fig. 9. However, they have different magnitudes. Similar to the pion case, in Fig. 8, we also show our results for the kaon electromagnetic form factors for different densities multiplied by $Q^2$. It also indicates that the kaon electromagnetic form factor at larger $Q^2 \to \infty$ (asymptotic feature) satisfies $F_K(Q^2, \rho_B) \approx 1/\mathrm{Q}^2$ with $Q^2 F_K(Q^2, \rho_B = 0) \simeq 0.55$.

In addition, results for the in-medium up and strange quark sector form factors of the kaon multiplied by $Q^2$ for different densities are shown in Figs. 9 and 10, respectively. Similar to the up quark sector form factor of the pion in free space and nuclear matter density in Fig. 6, the up quark sector form factor of the kaon in nuclear medium decreases as the nuclear matter density increases. However, the up quark sector form factor of the kaon decreases faster in comparison with that of the pion. Also, we observe the up quark sector form factor for the pion is larger than that of the kaon for the corresponding nuclear matter density as clearly shown in Figs. 6 and 9, respectively.

Results for the in-medium modifications of the strange sector form factor for the kaon multiplied by $Q^2$ are depicted in Fig. 10. Interestingly, the results for the strange quark sector form factor are different from the up-quark sector form factor of the kaon. It clearly shows that the strange quark form factors increase as the nuclear matter density increases. The increasing rate is rather slow for different densities. Also, we observe, in free space, the strange quark sector form factor begins to contribute dominantly at $Q^2 \geq 1.6$ GeV$^2$ to the total electromagnetic form factor of the kaon, which is consistent with that found in Ref. [12]. It is worth noting that by comparing the in-medium modifications kaon electromagnetic form factor calculated by using the hybrid NJL-QMC model [9] and LF-QMC model [6] to the present result of this work, we found that the nuclear medium kaon form factor is not too sensitive to the medium modifications kaon mass. In Refs. [6, 9], the medium modifications kaon masses decrease as the nuclear matter density increases, which is in contrast to our present result, where the medium modifications kaon mass increases as the nuclear matter density increases as can be seen in Fig. 1. The decreasing kaon mass in Refs. [6, 9] originates from the fact that the kaon potential is less repulsive. However, this can be improved by rescaling the repulsive vector potential to some extent (e.g., by rescaling with a factor of $(1.4)^2$) as in Refs. [37, 38].

Next, we are in a position to compute the charge radii of the pion and kaon as well as their sector form factors in free space and nuclear medium, which are calculated from the derivative of the form factors with respect to $Q^2$, as formulated in Eq. (54).

### 2. *Charge radius*

Using the expression for the in-medium modifications of the kaon and pion electromagnetic form factors in Eq. (42), respectively, the root-mean-square charge radius of the pion and kaon for various densities can be obtained from

$$\langle r^2_{\pi(K)} \rangle = -6 \frac{dF_{\pi(K)}(Q^2, \rho_B)}{dQ^2}\bigg|_{Q^2=0} . \qquad (54)$$

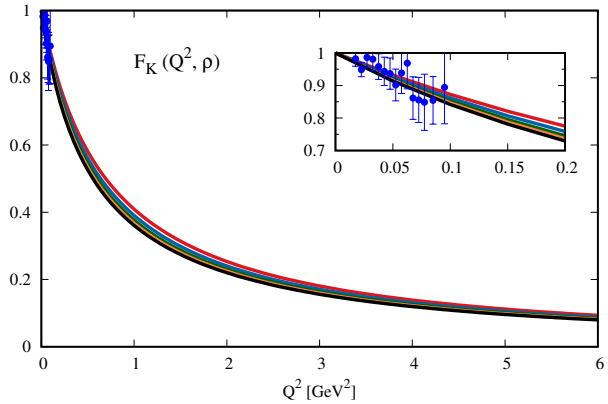



FIG. 7: Same as in Fig. 4, but for the kaon. Here, the experimental data is from Ref. [36].

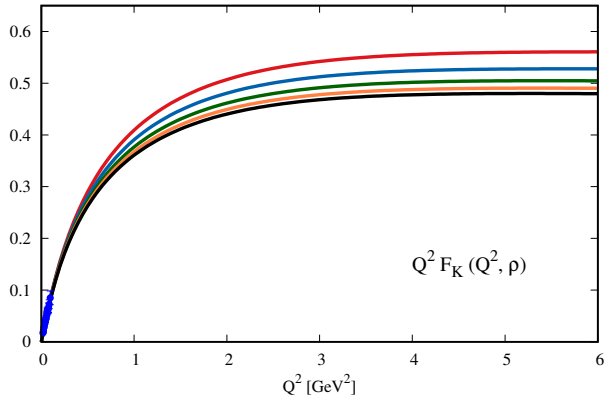



FIG. 8: Same as in Fig. 7, but for $F_K(Q^2, \rho_B)$ multiplied by $Q^2$.

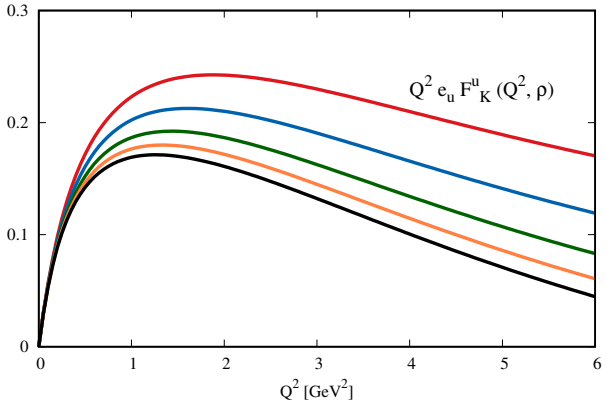



FIG. 9: Same as in Fig. 7, but for $e_u F_K^u(Q^2, \rho_B)$ multiplied by $Q^2$.

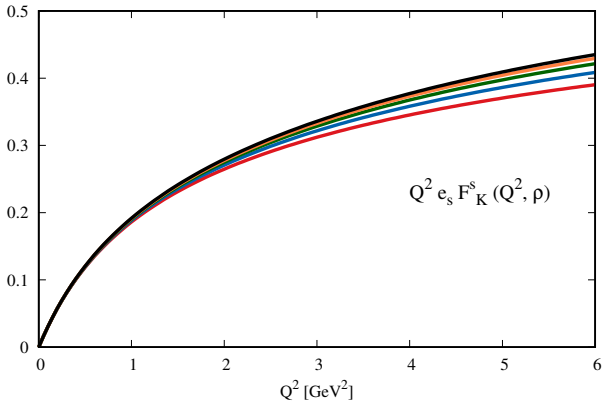



FIG. 10: Same as in Fig. 7, but for $e_s F_K^s(Q^2, \rho_B)$ multiplied by $Q^2$.

Results for the kaon and pion charge radii in free space and nuclear medium are shown in Figs. 11 and 12. Figure 11 shows the kaon and its quark sector charge radii. In free space, the kaon charge radius is in good agreement with the experimental data [40]. The kaon charge radius (solid line) in nuclear medium increases as the nuclear matter density increases. Our results are also consistent with other calculations obtained by using different approaches [9].

In Fig. 11, it is shown that the contribution of the up-quark charge radius (dotted line) increases as the nuclear matter density increases. The numerical values of the up-quark charge radius contributions can be also seen in Table. II. Also, we show the strange-quark charge radius (dashed line) contribution to the total kaon charge radius. It reveals that the strange quark charge radius contribution does not significantly vary in the nuclear medium, as expected. This is because the strange quark mass does not significantly change in the nuclear medium and the strange quark weakly interacts with vector and scalar mean fields in the nuclear medium.

Besides the positively charged kaon radius result, we also provide the result for the charge radius for the neutral kaon (dash-dotted line) in free space and nuclear medium. It shows that the charge radius for the neutral kaon decreases as the nuclear matter density increases as shown in Fig. 11. For the positively charged and neutral kaons the contributions of strange and up quarks charge radii are similar, as indicated in Table. II. Only their total charges are different.

Finally, we show our result for the pion charge radius in free space and nuclear medium in Fig.12. The free space charge radius for the pion reproduces the experimental data [39] nicely, while the nuclear medium charge radius increases as the nuclear matter density increases. This result is consistent with other calculations available in the literature [5, 10].

TABLE II: Kaon and pion charge radius in free space and nuclear medium and their quark sector form factors. All radii are in units of fm.

| $\pi^+$ | $\rho_B/\rho_0$ | $r_\pi$ | $r_u$ | $r_d$ | $r_{\text{expt.}}$ [39] |
|---|---|---|---|---|---|
| | 0.0 | 0.629 | 0.629 | −0.629 | $0.672 \pm 0.008$ |
| | 0.5 | 0.664 | 0.664 | −0.664 | |
| | 1.0 | 0.694 | 0.694 | −0.694 | |
| | 1.5 | 0.714 | 0.714 | −0.714 | |
| | 2.0 | 0.730 | 0.730 | −0.730 | |
| $K^+$ | $\rho_B/\rho_0$ | $r_K$ | $r_u$ | $r_s$ | $r_{\text{expt.}}$ [40] |
| | 0.0 | 0.586 | 0.646 | −0.441 | $0.560 \pm 0.031$ |
| | 0.5 | 0.614 | 0.686 | −0.434 | |
| | 1.0 | 0.638 | 0.719 | −0.434 | |
| | 1.5 | 0.655 | 0.742 | −0.433 | |
| | 2.0 | 0.668 | 0.759 | −0.432 | |
| $K^0$ | $\rho_B/\rho_0$ | $r_K$ | $r_d$ | $r_s$ | $r_{\text{expt.}}$ [41] |
| | 0.0 | −0.272 | 0.646 | −0.441 | $-0.277 \pm 0.018$ |
| | 0.5 | −0.307 | 0.686 | −0.434 | |
| | 1.0 | −0.331 | 0.719 | −0.434 | |
| | 1.5 | −0.348 | 0.742 | −0.433 | |
| | 2.0 | −0.361 | 0.759 | −0.432 | |

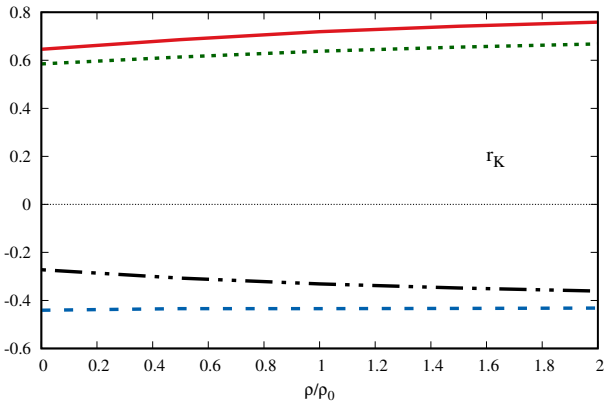



FIG. 11: Kaon radius and its quark sector charge radius contributions in the nuclear medium. The solid, dotted, dashed, and dash-dotted lines represent respectively $r_u$, $r_{K^+}$, $r_s$, and $r_{K^0}$.

## VII. SUMMARY

In the present work, we have investigated the in-medium modifications of the kaon and pion properties and structures as well as their charge radii in the framework of the NJL model with the help of the Schwinger proper-time regularization scheme to simulate the effect of QCD confinement. The nuclear medium effect is also computed with the same NJL model consistently, where the chiral symmetry-breaking partial restoration is captured in the model *via* the chiral quark condensate as an order parameter. In this work, we compute the free space pion and kaon properties and structures as well as in the nuclear medium. We then compare our free space results with the free space experimental data and other theoretical calculations. We found that the free space kaon and pion electromagnetic form factors are in good agreement with the experimental data, which is followed by the kaon and pion charge radii, which result in good agreement with data.

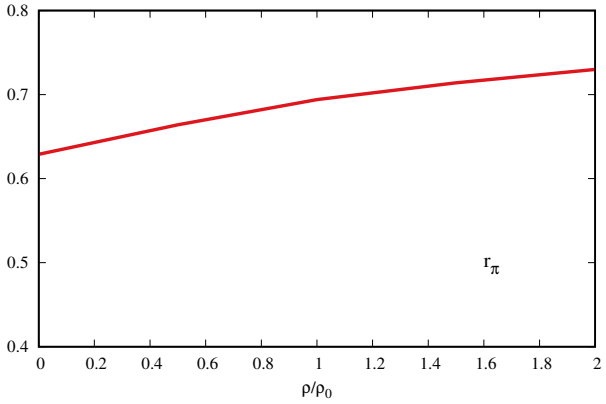



FIG. 12: Pion charge radius in free space and the nuclear matter.

With these good free space results, we predict the kaon and pion electromagnetic form factors in the nuclear medium. It is found that the kaon electromagnetic form factor decreases as the nuclear matter density increases. Similarly, the pion electromagnetic form factor in the nuclear medium also decreases as the nuclear matter density increases. Their quark sector form factor contributions for the pion and kaon are also investigated.

Also, we predict the kaon and pion charge radii in the nuclear medium. It is found that the kaon charge radius in the nuclear medium increases as the nuclear matter density increases. A similar trend is also shown by the pion charge radius.

Another interesting result of this work is the in-medium modifications of the axial-vector $g_A^*$ coupling that is calculated via GMOR. We found that $g_A^*$ does not significantly change in the nuclear medium, where at nuclear matter density $\rho_B = \rho_0$, the reduction of $g_A$ is estimated about 0.2% relative to that in free space, indicating a very small quenching factor.

The present studies of the pion and kaon properties in free space and nuclear medium are very useful to provide valuable and relevant information to the deeply bound pionic and kaonic atom experiments that are planned to be measured at J-PARC in Japan [42]. Also, the results of this study may provide useful and helpful guidance for the lattice QCD. Next, for future work, we plan to apply this consistent NJL model in calculating the electromagnetic form factors of spin-1 vector mesons with equal and unequal quark masses in the nuclear medium. Related works are still in progress and will appear elsewhere.

**Acknowledgements**

This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korean government (MSIT) Nos. 2018R1A5A1025563, 2022R1A2C1003964, and 2022K2A9A1A0609176. The work of TM was supported by the PUTI Q2 Grant from University of Indonesia under contract No. NKB-663/UN2.RST/HKP.05.00/2022.

[1] B. L. Ioffe, Prog. Part. Nucl. Phys. **56**, 232-277 (2006) doi:10.1016/j.ppnp.2005.05.001 [arXiv:hep-ph/0502148 [hep-ph]].

[2] A. Accardi, P. Achenbach, D. Adhikari, A. Afanasev, C. S. Akondi, N. Akopov, M. Albaladejo, H. Albataineh, M. Albrecht and B. Almeida-Zamora, *et al.* [arXiv:2306.09360 [nucl-ex]].

[3] J. J. Aubert *et al.* [European Muon], Phys. Lett. B **123**, 275-278 (1983) doi:10.1016/0370-2693(83)90437-9

[4] D. F. Geesaman, K. Saito and A. W. Thomas, Ann. Rev. Nucl. Part. Sci. **45**, 337-390 (1995) doi:10.1146/annurev.ns.45.120195.002005

[5] J. P. B. C. de Melo, K. Tsushima, B. El-Bennich, E. Rojas and T. Frederico, Phys. Rev. C **90**, no.3, 035201 (2014) doi:10.1103/PhysRevC.90.035201 [arXiv:1404.5873 [hep-ph]].

[6] G. H. S. O. Yabusaki, J. P. B. C. de Melo, K. Tsushima, T. Frederico and W. de Paula, [arXiv:2310.11262 [hep-ph]].

[7] N. Er and K. Azizi, Eur. Phys. J. C **82**, no.5, 397 (2022) doi:10.1140/epjc/s10052-022-10333-w [arXiv:2202.01504 [hep-ph]].

[8] A. J. Arifi, P. T. P. Hutauruk and K. Tsushima, Phys. Rev. D **107**, no.11, 114010 (2023) doi:10.1103/PhysRevD.107.114010 [arXiv:2302.12382 [hep-ph]].

[9] P. T. P. Hutauruk and K. Tsushima, [arXiv:1910.08133 [nucl-th]].

[10] P. T. P. Hutauruk, Y. Oh and K. Tsushima, Phys. Rev. C **99**, no.1, 015202 (2019) doi:10.1103/PhysRevC.99.015202 [arXiv:1810.08874 [nucl-th]].

[11] D. L. Whittenbury, M. E. Carrillo-Serrano and A. W. Thomas, Phys. Lett. B **762**, 467-472 (2016) doi:10.1016/j.physletb.2016.09.057 [arXiv:1606.03158 [nucl-th]].

[12] P. T. P. Hutauruk, I. C. Cloet and A. W. Thomas, Phys. Rev. C **94**, no.3, 035201 (2016) doi:10.1103/PhysRevC.94.035201 [arXiv:1604.02853 [nucl-th]].

[13] P. T. P. Hutauruk and S. i. Nam, Phys. Rev. D **105**, no.3, 034021 (2022) doi:10.1103/PhysRevD.105.034021 [arXiv:2112.05435 [hep-ph]].

[14] P. T. P. Hutauruk and S. i. Nam, [arXiv:2302.05566 [hep-ph]].

[15] X. G. Wang, W. Bentz, I. C. Cloët and A. W. Thomas, J. Phys. G **49**, no.3, 03LT01 (2022) doi:10.1088/1361-6471/ac4c90 [arXiv:2109.03591 [hep-ph]].

[16] P. T. P. Hutauruk and S. i. Nam, Makara J. Sci. **26**, no.4, 296-303 (2022) doi:10.7454/mss.v26i4.1402 [arXiv:2303.01652 [hep-ph]].

[17] T. Tanimoto, W. Bentz and I. C. Cloët, Phys. Rev. C **101**, no.5, 055204 (2020) doi:10.1103/PhysRevC.101.055204 [arXiv:1903.06851 [nucl-th]].

[18] W. Bentz and A. W. Thomas, Nucl. Phys. A **696**, 138-172 (2001) doi:10.1016/S0375-9474(01)01119-8 [arXiv:nucl-th/0105022 [nucl-th]].

[19] P. T. P. Hutauruk, H. Gil, S. i. Nam and C. H. Hyun, [arXiv:2307.09038 [nucl-th]].

[20] P. T. P. Hutauruk, W. Bentz, I. C. Cloët and A. W. Thomas, Phys. Rev. C **97**, no.5, 055210 (2018) doi:10.1103/PhysRevC.97.055210 [arXiv:1802.05511 [nucl-th]].

[21] Y. Ninomiya, W. Bentz and I. C. Cloët, Phys. Rev. C **96**, no.4, 045206 (2017) doi:10.1103/PhysRevC.96.045206 [arXiv:1707.03787 [nucl-th]].

[22] W. Bentz, A. Kotzinian, H. H. Matevosyan, Y. Ninomiya, A. W. Thomas and K. Yazaki, Phys. Rev. D **94**, no.3, 034004 (2016) doi:10.1103/PhysRevD.94.034004 [arXiv:1603.08333 [nucl-th]].

[23] I. C. Cloët, W. Bentz and A. W. Thomas, Phys. Rev. C **90**, 045202 (2014) doi:10.1103/PhysRevC.90.045202 [arXiv:1405.5542 [nucl-th]].

[24] P. T. P. Hutauruk, Mod. Phys. Lett. A **38**, no.28n29, 2350132 (2023) doi:10.1142/S0217732323501328 [arXiv:2204.11520 [hep-ph]].

[25] P. T. P. Hutauruk and S. i. Nam, Mod. Phys. Lett. A **37**, no.14, 2250087 (2022) doi:10.1142/S0217732322500870 [arXiv:2103.09426 [hep-ph]].

[26] J. Suhonen and O. Civitarese, Phys. Rept. **300**, 123-214 (1998) doi:10.1016/S0370-1573(97)00087-2

[27] V. Cirigliano, Z. Davoudi, J. Engel, R. J. Furnstahl, G. Hagen, U. Heinz, H. Hergert, M. Horoi, C. W. Johnson and A. Lovato, *et al.* J. Phys. G **49**, no.12, 120502 (2022) doi:10.1088/1361-6471/aca03e [arXiv:2207.01085 [nucl-th]].

[28] P. Gysbers, G. Hagen, J. D. Holt, G. R. Jansen, T. D. Morris, P. Navrátil, T. Papenbrock, S. Quaglioni, A. Schwenk and S. R. Stroberg, *et al.* Nature Phys. **15**, no.5, 428-431 (2019) doi:10.1038/s41567-019-0450-7 [arXiv:1903.00047 [nucl-th]].

[29] G. Krein, A. W. Thomas and K. Tsushima, Nucl. Phys. A **650**, 313-325 (1999) doi:10.1016/S0375-9474(99)00117-7 [arXiv:nucl-th/9810023 [nucl-th]].

[30] T. Waas, N. Kaiser and W. Weise, Phys. Lett. B **379**, 34-38 (1996) doi:10.1016/0370-2693(96)00472-8

[31] Y. L. Ma and M. Rho, Phys. Rev. Lett. **125**, no.14, 142501 (2020) doi:10.1103/PhysRevLett.125.142501 [arXiv:2002.03310 [nucl-th]].

[32] J. T. Suhonen, Front. in Phys. **5**, 55 (2017) doi:10.3389/fphy.2017.00055 [arXiv:1712.01565 [nucl-th]].

[33] C. A. Dominguez, M. Loewe, C. Villavicencio and R. Zamora, Phys. Rev. D **108**, no.7, 074024 (2023) doi:10.1103/PhysRevD.108.074024 [arXiv:2308.05663 [hep-ph]].

[34] S. R. Amendolia, B. Badelek, G. Batignani, G. A. Beck,

F. Bedeschi, E. H. Bellamy, E. Bertolucci, D. Bettoni, H. Bilokon and G. Bologna, *et al.* Phys. Lett. B **146**, 116-120 (1984) doi:10.1016/0370-2693(84)90655-5.

[35] G. M. Huber *et al.* [Jefferson Lab], Phys. Rev. C **78**, 045203 (2008) doi:10.1103/PhysRevC.78.045203 [arXiv:0809.3052 [nucl-ex]].

[36] S. R. Amendolia, G. Batignani, G. A. Beck, E. H. Bellamy, E. Bertolucci, G. Bologna, L. Bosisio, C. Bradaschia, M. Budinich and M. Dell'orso, *et al.* Phys. Lett. B **178**, 435-440 (1986) doi:10.1016/0370-2693(86)91407-3

[37] K. Tsushima, K. Saito, A. W. Thomas and S. V. Wright, Phys. Lett. B **429**, 239-246 (1998) [erratum: Phys. Lett. B **436**, 453-453 (1998)] doi:10.1016/S0370-2693(98)00488-2 [arXiv:nucl-th/9712044 [nucl-th]].

[38] K. Tsushima, A. Sibirtsev and A. W. Thomas, J. Phys. G **27**, 349-354 (2001) doi:10.1088/0954-3899/27/3/312 [arXiv:nucl-th/0008039 [nucl-th]].

[39] C. Patrignani *et al.* [Particle Data Group], Chin. Phys. C **40**, no.10, 100001 (2016) doi:10.1088/1674-1137/40/10/100001

[40] R. L. Workman *et al.* [Particle Data Group], PTEP **2022**, 083C01 (2022) doi:10.1093/ptep/ptac097

[41] M. Tanabashi *et al.* [Particle Data Group], Phys. Rev. D **98**, no.3, 030001 (2018) doi:10.1103/PhysRevD.98.030001

[42] K. Aoki, H. Fujioka, T. Gogami, Y. Hidaka, E. Hiyama, R. Honda, A. Hosaka, Y. Ichikawa, M. Ieiri and M. Isaka, *et al.* [arXiv:2110.04462 [nucl-ex]].

arXiv:2402.19006v1 [hep-ph] 29 Feb 2024

# Light quarkonium hybrid mesons

B. Barsbay,[1] K. Azizi,[2,3,4,*] and H. Sundu[5]

[1]*Division of Optometry, School of Medical Services and Techniques, Doğuş University, 34775 Istanbul, Türkiye*
[2]*Department of Physics, University of Tehran, North Karegar Ave. Tehran 14395-547, Iran*
[3]*Department of Physics, Doğuş University, Dudullu-Ümraniye, 34775 Istanbul, Türkiye*
[4]*School of Particles and Accelerators, Institute for Research in Fundamental Sciences (IPM) P.O. Box 19395-5531, Tehran, Iran*
[5]*Department of Physics Engineering, Istanbul Medeniyet University, 34700 Istanbul, Türkiye*

(Dated: March 1, 2024)

We investigate the light quarkonium hybrid mesons of various spin-parities in QCD. Considering different interpolating currents made of the valence light quarks and single gluon, we calculate the mass and current coupling of the strange and nonstrange members of light hybrid mesons by including into computations the nonperturbative quark and gluon condensates up to ten dimensions in order to increase the accuracy of the results. The obtained results may be useful for future experimental searches of these hypothetical states. They can also be used in the calculations of different parameters related to the decays/interactions of light hybrid mesons to/with other states.

## I. INTRODUCTION

As successful theory of strong interaction, the quantum chromodynamics (QCD) together with the powerful quark model have predicted the existence of exotic hadrons beyond the standard mesons and baryons. The most known categories for exotic hadrons are tetraquarks, pentaquarks, hexaquarks, quark-gluon hybrids and glueballs. Starting from 2003, many tetraquark and pentaquark states have been discovered in the experiment. Concerning the hexaquarks, WASA-at-COSY collaboration has reported observation of a six-quark candidate $d^*(2380)$ with $J^P = 3^+$ [1], starting a wide research on the properties of hexaquarks and dibaryons as interesting objects: a hypothetical SU(3) flavor-singlet, highly symmetric, deeply bound neutral particle termed the scalar hexaquark $S = uuddss$ has been introduced as a potential candidate for dark matter [2]. Although some resonances have been introduced as the potential condidates for the hybrids and glueballs, there is no discovered particles with high confidence level in these categories yet. It is time to investigate the hybrid states and glueballs with a fresh breath, though they are searched for a long time. Hybrid states was predicted in 1976 [3]. There are some states with quantum numbers $J^{PC} = 0^{--}$, $0^{+-}$, $1^{-+}$, $2^{-+}$, which cannot be explained by $q\bar{q}$ picture and are considered as potential hybrid meson candidates [4]. Among them are $\pi_1(1400)$ [5], $\pi_1(1600)$ [6], $\pi_1(2015)$ [7], and $\eta_1(1855)$ [8] with exotic quantum numbers $J^{PC} = 1^{-+}$ evidenced by some experiments. They are possible single-gluon hybrid candidates of a quark-antiquark pair together with a valence gluon. Designed to search for the hybrid mesons as its primary goal, the GlueX experiment at Jefferson Lab is expected to give crucial insights into the existence and structure of the exotic hybrid mesons.

Investigation of the light and heavy hybrid mesons is of great importance not only for determination of their nature and quark-gluon organization but for gaining useful information about the nonperturbative nature of QCD. Light hybrid states, which are subjects of the present study, have been intensively investigated in the framework of different theoretical methods such as lattice QCD [9, 10], the Schwinger-Dyson formalism [11–13],the flux tube model [14, 15], the MIT bag model [16, 17] and QCD Laplace sum-rules(LSRs) [18–34]. In particular, Ref. [21] contains a comprehensive LSRs analysis of the light hybrids for $J = 0$ and 1 with all the possible combinations for the parity and charge quantum numbers. Analyses show that the $0^{++}$, $0^{--}$, $1^{++}$, and $1^{--}$ states are mainly stable, while $0^{+-}$, $0^{-+}$, $1^{+-}$, and $1^{-+}$ quantum numbers lead to unstable and controversial results. Expected to be the lightest hybrid mesons, the ones with $1^{-+}$ have been the subject of much additional studies. In the QCD sum rules, predictions on the $1^{-+}$ light hybrid mesons are inconsistent among different works. For instance, I. I. Balitsky *et al*'s prediction of mass for $1^{-+}$ state is in the range, 1.0–1.3 GeV [18, 22], J . I. Latorre *et al* predicted the related mass to be around 2.1 GeV [24] and the obtained mass for $1^{-+}$ state is around 2.5 GeV in Refs. [19, 35, 36]. The mass of $1^{-+}$ was re-analyzed by including the quark-gluon condensates up to 8 dimensions [34] and it was found that the mass value increases to be in the range 1.72–2.60 GeV. The obtained mass range did not favor the $\pi_1(1400)$ and the $\pi_1(1600)$ to be pure hybrid states and suggests the $\pi_1(2015)$, observed by E852, to have much of a hybrid constituent. The masses of the light hybrid mesons with $J^{PC} = 1^{-+}$ quantum numbers have also been calculated using different theoretical methods other than QCD sum rules [37–40].

Concerning other quantum numbers, the masses of the $0^{++}$ and $0^{-+}$ light hybrid mesons were calculated using the QCD sum rule method by considering two kinds of the interpolated currents with the same quantum numbers. While the approximately equal mass was predicted for the $0^{-+}$ hybrid states from the two different currents, different masses were obtained for the $0^{++}$ hybrid

*Corresponding Author

states from the two considered different currents. The masses of light (nonstrange and strange) quarkonium hybrid mesons with $J^{PC} = 0^{+-}$ were investigated using Gaussian sum rules accompanied by the Hölder inequality [41]. The obtained mass predictions of the nonstrange and strange states were 2.60 GeV and 3.57 GeV, respectively. Unfortunately, the predictions for the masses of the light hybrid mesons with various spin and parity are inconsistent with each other. Thus, further theoretical studies for the light hybrid states with all possible quantum numbers are necessary.

In this article, inspired by this situation, we will compute the mass and current coupling of the light quarkonium hybrid mesons with different quantum numbers. We will utilize the Borel QCD sum rule method to carry out the calculations. This method is recognized as a powerful and predictive nonperturbative approach in the field of hadron physics [42–45]. It has demonstrated good success not only in analyzing the properties of conventional hadrons but also in examining the exotic particles (see for instance Refs. [46–60]). The results obtained through this approach have effectively confirmed the existing experimental data. We construct different currents to interpolate the light quarkonium hybrid mesons with various possible spin-parities. The aim is to explore the spectroscopic characteristics of various types of light hybrid mesons. Note that, the scalar and vector states; and pseudoscalar and axial-vector states, couple simultaneously to the same currents. We isolate the contributions of different states by choosing appropriate Lorentz structures entering the calculations.

In sec. II, the sum rules for the mass and current coupling of the light strange and nonstrange hybrid mesons of different spin-parities are derived. In sec. III, the obtained sum rules are numerically analyzed. The last section is dedicated to the summary of the calculations and conclusions. We move the expressions for the contributions of different perturbative and nonperturbative operators to the Appendix.

## II. FORMALISM

The sum rules for the mass and current coupling of the light hybrid mesons can be extracted from analysis of the following two-point correlation function:

$$\Pi_{\mu\nu}(q) = i\int d^4x e^{iqx}\langle 0|\mathcal{T}\{J_\mu(x)J_\nu^\dagger(0)\}|0\rangle = \frac{q_\mu q_\nu}{q^2}\Pi_{S(PS)}(q^2) + \left(\frac{q_\mu q_\nu}{q^2} - g_{\mu\nu}\right)\Pi_{V(AV)}(q^2), \tag{1}$$

where $J_\mu(x)$ is the current representing the vector ($V$) and axial-vector ($AV$) light hybrid states coupling to the scalar ($S$) and pseudoscalar ($PS$) states as well. Possible interpolating currents with different valence quark-gluon contents and quantum numbers to be considered in the present study are given in Table I. Here $g_s$ is the strong coupling constant, $a, b = 1, 2, 3$ are color indices, $\lambda^n$ with $n = 1, 2, \cdots, 8$ are the Gell-Mann matrices, $\tilde{G}^n_{\mu\theta}(x) = \epsilon_{\mu\theta\alpha\beta}G^n_{\alpha\beta}(x)/2$ is the dual field strength of $G^n_{\mu\theta}(x)$, and $q = u, d, s$ is the light quark fields. In Table I one can see the states with different quantum numbers that couple to each current, simultaneously. This situation causes some pollution that should be removed to calculate the physical quantities of the desired states.

TABLE I: The currents of the light hybrid states.

| Currents | $J^{PC}$ |
|---|---|
| $J^1_\mu = g_s\overline{q}_a\gamma_\theta\gamma_5\frac{\lambda^n_{ab}}{2}\tilde{G}^n_{\mu\theta}q_b$ | $0^{+-}, 1^{--}$ |
| $J^2_\mu = g_s\overline{q}_a\gamma_\theta\gamma_5\frac{\lambda^n_{ab}}{2}G^n_{\mu\theta}q_b$ | $0^{--}, 1^{+-}$ |
| $J^3_\mu = g_s\overline{q}_a\gamma_\theta\frac{\lambda^n_{ab}}{2}\tilde{G}^n_{\mu\theta}q_b$ | $0^{-+}, 1^{++}$ |
| $J^4_\mu = g_s\overline{q}_a(x)\gamma_\theta\frac{\lambda^n_{ab}}{2}G^n_{\mu\theta}q_b$ | $0^{++}, 1^{-+}$ |

To obtain the sum rules for the mass and current coupling of different states we need to relate these hadronic quantities to the fundamental QCD parameters like the quark masses, quark-gluon condensates of different nonperturbative mass dimensions, strong coupling constants and some auxiliary parameters entering the calculations at different stages based on the standard prescriptions of the method. Therefore, we need to calculate the aforesaid correlation function in two hadronic and QCD languages. Matching the coefficients of different Lorentz structures from the both representations will give us the aiming sum rules. In technique language, the hadronic representation at time-like region is found by inserting complete sets of hadronic states with the same quantum properties as the interpolating currents between the two creating and annihilating currents in coordinate space. As an example, for the current $J^1_\mu$ coupling simultaneously to the scalar $0^{+-}$ and vector $1^{--}$ states, after performing the four integrals ove four-$x$, we get

$$\Pi^{\rm Phys}_{\mu\nu}(q) = \frac{\langle 0|J^1_\mu|H_S(q)\rangle\langle H_S(q)|J^{1\dagger}_\nu|0\rangle}{m^2_{H_S} - q^2} + \frac{\langle 0|J^1_\mu|H_V(q)\rangle\langle H_V(q)|J^{1\dagger}_\nu|0\rangle}{m^2_{H_V} - q^2} + \ldots, \tag{2}$$

where $\Pi^{\rm Phys}_{\mu\nu}(q)$ stands for the physical or hadronic side; and $m_{H_S}$ and $m_{H_V}$ are the masses of the light $S$ and $V$ hybrid states, respectively. To proceed, we need to define the following matrix elements:

$$\langle 0|J_\mu|H_S(q)\rangle = q_\mu f_{H_S} \tag{3}$$

and

$$\langle 0|J_\mu|H_V(q)\rangle = m_{H_V} f_{H_V}\varepsilon_\mu, \tag{4}$$

where $\varepsilon_\mu$ is the polarization vector of the $V$ state; and $f_{H_S}$ and $f_{H_V}$ represent the corresponding current couplings. Up to here, the function $\Pi_{\mu\nu}^{\rm Phys}(q)$ takes the form,

$$\Pi_{\mu\nu}^{\rm Phys}(q) = \frac{f_{H_S}^2}{m_{H_S}^2-q^2}q_\mu q_\nu + \frac{m_{H_V}^2 f_{H_V}^2}{m_{H_V}^2-q^2} \times \left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{q^2}\right)+\cdots . \quad (5)$$

As it is clear, the function $\Pi_{\mu\nu}^{\rm Phys}(q)$ contains the $S$ and $V$ hybrid contributions, simultaneously. To isolate the $S$ contribution alone, it is enough to multiply Eq. (5) by $q_\mu q_\nu/q^2$ which results in,

$$\frac{q_\mu q_\nu}{q^2}\Pi_{\mu\nu}^{\rm Phys}(q) = -f_{H_S}^2 + \frac{m_{H_S}^2 f_{H_S}^2}{m_{H_S}^2-q^2}+\cdots . \quad (6)$$

The contribution of vector state is found by choosing the structure $g_{\mu\nu}$ which is free of the $S$ hybrid state and contains only the vector contribution:

$$\Pi_{\mu\nu}^{\prime\rm Phys}(q) = \frac{m_{H_V}^2 f_{H_V}^2}{m_{H_V}^2-q^2}\times(-g_{\mu\nu})+\cdots . \quad (7)$$

To suppress the contributions of the higher states and continuum at each channel, we apply Borel transformation with respect to $q^2$. We find,

$$\mathcal{B}_{q^2}\Pi_{\mu\nu}^{\prime\rm Phys}(q) = m_{H_V}^2 f_{H_V}^2 e^{-m_{H_V}^2/M^2}\,(-g_{\mu\nu})+\cdots , \quad (8)$$

and

$$\mathcal{B}_{q^2}\frac{q_\mu q_\nu}{q^2}\Pi_{\mu\nu}^{\rm Phys}(q) = m_{H_S}^2 f_{H_S}^2 e^{-m_{H_S}^2/M^2}\ +\cdots . \quad (9)$$

for the final representations of the physical contributions of the $V$ and $S$ states in Borel scheme, respectively. In these equations $M^2$ is the Borel parameter to be fixed in next section.

On the QCD side, the correlation function is obtained in deep Euclidean space-like region, $q^2\to-\infty$, where the time ordering operator of two currents are expanded in terms of different perturbative and nonperturbative contributions using operator product expansion (OPE). We take into account the nonperturbative contributions up to ten dimensions. In fact, to do this, we insert the explicit forms of the currents into the correlation function and perform all the possible contractions among the quark fields in vacuum. The result is obtained in terms of the participating quarks propagators and vacuum expectation value of the two valence gluons. As an example, for the current $J_\mu^1$ the function $\Pi_{\mu\nu}^{\rm OPE}(q)$ takes the form

$$\Pi_{\mu\nu}^{\rm OPE}(q) = \frac{g_s^2\epsilon_{\mu\theta\alpha\beta}\epsilon_{\nu\delta\alpha'\beta'}}{4}i\int d^4x e^{iqx} \times\langle 0|G_{\alpha\beta}^n(x)G_{\alpha'\beta'}^m(0)|0\rangle {\rm t}^n[a,b]{\rm t}^m[a',b'] \times{\rm Tr}\left[S_q^{a'a}(-x)\gamma_\theta\gamma_5 S_q^{bb'}(x)\gamma_5\gamma_\delta\right], \quad (10)$$

where $t^n=\lambda^n/2$ and $S_q^{ab}(x)$ is the light-quark propagator, whose expression in coordinate space is given by

$$S_q^{ab}(x) = i\frac{\not{x}\delta_{ab}}{2\pi^2x^4} - \frac{m_q\delta_{ab}}{4\pi^2x^2} - \frac{\langle\overline{q}q\rangle}{12}\left(1-i\frac{m_q}{4}\not{x}\right)\delta_{ab} -\frac{x^2}{192}\langle\overline{q}g_s\sigma Gq\rangle\left(1-i\frac{m_q}{6}\not{x}\right)\delta_{ab} - \frac{\not{x}x^2g_s^2}{7776}\langle\overline{q}q\rangle^2\delta_{ab} -\frac{ig_sG_{ab}^{\mu\nu}}{32\pi^2x^2}[\not{x}\sigma_{\mu\nu}+\sigma_{\mu\nu}\not{x}] - \frac{x^4\langle\overline{q}q\rangle\langle g_s^2G^2\rangle}{27648}\delta_{ab}+\cdots . \quad (11)$$

When replacing this propagator into Eq. (10), besides the different perturbative and nonperturbative contributions of different dimensions resulting from the multiplications of two propagators, there appears another term from the multiplications of $-\frac{ig_sG_{ab}^{\mu\nu}}{32\pi^2x^2}[\not{x}\sigma_{\mu\nu}+\sigma_{\mu\nu}\not{x}]$ from both the propagators. This term makes another two-gluon condensate in the presence of vacuum. We apply the following short-hand notations:

$$G_{ab}^{\alpha\beta} = G_A^{\alpha\beta}\lambda_{ab}^A/2, \quad G^2 = G_{\alpha\beta}^A G_{\alpha\beta}^A, \quad (12)$$

in the calculations.

Now, we proceed to discuss the matrix element $\langle 0|G_{\alpha\beta}^n(x)G_{\alpha'\beta'}^m(0)|0\rangle$ in Eqs. (10). It is considered by two different ways: In the first approach, it is considered as the full gluon propagator in coordinate space between two points 0 and $x$, whose expresion is given by

$$\langle 0|G_{\alpha\beta}^n(x)G_{\alpha'\beta'}^m(0)||0\rangle = \frac{\delta^{mn}}{2\pi^2x^4}[g_{\beta\beta'}(g_{\alpha\alpha'}-\frac{4x_\alpha x_{\alpha'}}{x^2}) +(\beta,\beta')\leftrightarrow(\alpha,\alpha')-\beta\leftrightarrow\alpha-\beta'\leftrightarrow\alpha']. \quad (13)$$

This is multiplied to the two light quark propagators and all the calculations are done. This contribution is equal to the diagrams at which the two quarks and the valence-gluon all are full propagators. In the second approach, the matrix element $\langle 0|G_{\alpha\beta}^n(x)G_{\alpha'\beta'}^m(0)|0\rangle$ is considered as the two-gluon condensate. To this end the gluon field at point $x$ is expanded around $x=0$ and the first term is taken into account. Hence, one has,

$$\langle 0|G_{\alpha\beta}^n(0)G_{\alpha'\beta'}^m(0)|0\rangle = \frac{\langle G^2\rangle}{96}\delta^{mn}[g_{\alpha\alpha'}g_{\beta\beta'} -g_{\alpha\beta'}g_{\alpha'\beta}]. \quad (14)$$

The contribution coming from this approach is equivalent to the diagrams with full quark propagators and the gluon interacting with the QCD vacuum. By using Eq. (13) or (14) into Eqs. (10), we make use of the following relation:

$${\rm t}^n[a,b]{\rm t}^n[a',b'] = \frac{1}{2}\left(\delta^{ab'}\delta^{a'b}-\frac{1}{3}\delta^{ab}\delta^{a'b'}\right). \quad (15)$$

After putting all the elements explained above together, we find expressions in OPE side in terms of fundamental QCD parameters like the quark masses, quark and gluon condensates, strong coupling constant etc. in coordinate

space. We apply Fourier transformation to transfer the calculations to the momentum space. To this end, we make use of the formula

$$\frac{1}{(x^2)^m} = \int \frac{d^Dk}{(2\pi)^D} e^{-ik\cdot x} i(-1)^{m+1} 2^{D-2m} \pi^{D/2} \times \frac{\Gamma[D/2-m]}{\Gamma[m]} \Big(-\frac{1}{k^2}\Big)^{D/2-m}, \quad (16)$$

then perform integral over four-$x$. To perform the integrals over other four-parameters, we use the Feynman parametrization and the formula,

$$\int d^4\ell \frac{(\ell^2)^m}{(\ell^2+\Delta)^n} = \frac{i\pi^2(-1)^{m-n}\Gamma[m+2]\Gamma[n-m-2]}{\Gamma[2]\Gamma[n](-\Delta)^{n-m-2}}, \quad (17)$$

where, $\Delta$ depends on the parameters of the problem except $\ell$. Finally, we use

$$\Gamma\Big[\frac{D}{2}-n\Big]\Big(-\frac{1}{\Delta}\Big)^{D/2-n} = \frac{(-1)^{n-1}}{(n-2)!}(-\Delta)^{n-2} ln[-\Delta]. \quad (18)$$

to find the imaginary parts of the resultant expressions for the perturbative and nonperturbative contributions up to six dimensions, which gives the spectral densities in dispersion representation. Note that, for the contributions equal or grater than the six dimensions, we directly find the contributions of the nonperturbative operators based on the standard prescriptions of the method. The next step is applying the Borel transformation to the OPE side the same as the physical side. The final procedure, is to apply continuum subtraction to further suppress the contributions of higher states and continuum. When performing this, a threshold $s_0$ is set to use the quark-hadron duality assumption. The continuum threshod depends on the energy of the first excited state and will be fixed in next section as well. As a result, for the OPE side we get

$$\Pi(M^2, s_0) = \int_{4m_q^2}^{s_0} ds \rho^{\rm OPE}(s) e^{-s/M^2} + \Gamma(M^2), \quad (19)$$

where $\rho^{\rm OPE}(s) = \frac{1}{\pi} Im[\Pi(q^2)]$ is the two-point spectral density calculated up to dimension six nonperturbative contributions. The invariant amplitude $\Gamma(M^2)$ contains nonperturbative contributions calculated directly from the OPE side and for the nonperturbative contributions from six to ten mass dimensions. Note that the six dimension appears both in the $\rho^{\rm OPE}(s)$ and $\Gamma(M^2)$. The explicit expressions of these functions are presented in the Appendix, as samples, for the strange hybrid meson with the quantum numbers $J^{PC} = 1^{--}$.

The last task in this section is to choose appropriate Lorentz structures from both sides and match the invariant coefficients. We shall say that, for our case, these structures are $g_{\mu\nu}$ and $I$, which typify contributions of $V(AV)$ and $S(PC)$ particles, respectively. For the mass $m_H$ and current coupling $f_H$ of the hybrid states associated to the current $J^1_\mu$, we get

$$m_H^2 = \frac{d/dx[\Pi(M^2, s_0)]}{\Pi(M^2, s_0)}, \quad (20)$$

and

$$f_H^2 = \frac{e^{m_H^2/M^2}}{m_H^2} \Pi(M^2, s_0), \quad (21)$$

where $x = -1/M^2$. Similar computations are done for the currents $J^2_\mu$, $J^3_\mu$ and $J^4_\mu$ to get sum rules for the physical quantities of the corresponding hybrid states.

## III. NUMERICAL RESULTS

In this section, we perform numerical analysis to get the values of the masses and current couplings. The quark masses are taken from PDG, while for the quark-gluon condensates, we use:

$$\begin{aligned} &\langle\overline{q}q\rangle = -(0.24 \pm 0.01)^3 \text{ GeV}^3, \\ &\langle\overline{s}s\rangle = 0.8 \ \langle\overline{q}q\rangle, \\ &\langle\overline{q}g_s\sigma Gq\rangle = m_0^2\langle\overline{q}q\rangle, \ m_0^2 = (0.8 \pm 0.1) \text{ GeV}^2, \\ &\langle\frac{\alpha_s G^2}{\pi}\rangle = (0.012 \pm 0.004) \text{ GeV}^4. \end{aligned} \quad (22)$$

The higher dimensional operators up to ten are obtained in terms of the condensates presented above (see the Appendix).

Besides the above-mentioned input parameters, the obtained sum rules in the previous section depend on two auxiliary parameters: The Borel parameter $M^2$ and the continuum threshold $s_0$. Their working windows are found based on the standard requirements of the method. These are the relatively weak dependence of the physical quantities, dominance of pole contribution at each channel and convergence of the OPE. For the convergence of OPE series we need that the perturbative part exceeds the total nonperturbative contribution and the higher the dimension of the nonperturbative operator the lower its contribution. In technique language, we define the following formula for the pole contribution ($PC$):

$$\text{PC} = \frac{\Pi(M^2, \ s_0)}{\Pi(M^2, \ \infty)}. \quad (23)$$

The condition $PC \geq 0.5$ sets the higher limit of the Borel parameter $M^2$. Its lower limit is fixed by the convergence of the OPE series. For this, we employ the condition

$$R(M^2_{\rm min}) = \frac{\Pi^{\rm DimN}(M^2_{\rm min}, \ s_0)}{\Pi(M^2_{\rm min}, \ s_0)} \leq 0.05, \quad (24)$$

where $\Pi^{\rm DimN}(M^2_{\rm min}, s_0)$ with $\text{DimN} = \text{Dim}(8+9+10)$ denotes the contributions of the last three nonperturbative operators. The continuum threshold $s_0$ depends on the energy of the first excited states at each channel

and it is fixed such that the above conditions are satisfied and we witness the relatively weak dependence of the mass and current coupling with respect to its variations in the working window. The residual dependence arising from the determination of the working intervals of the $M^2$ and $s_0$ appear as the uncertainties of the results. The uncertainties coming from the auxiliary parameters constitute the main errors in the presented values. The intervals for the $M^2$ and the values of $PC$ at lower and higher values of $M^2$ for all the considered channels at average values of the continuum threshold are given in tables II and III for the strange and nonstrange hybrid mesons, respectively. The presented values for $PC$ in these tables are acceptable for the exotic states. We will present the exact intervals of the continuum thresholds for each channel in the following tables. As an example, the perturbative and different nonperturbative contributions (in units of $GeV^8$) for the strange hybrid meson of the $J^{PC} = 1^{--}$ quantum numbers as functions of $M^2$ and at average $s_0$ is depicted in Fig. 1 . As it is seen, the OPE convergence condition is nicely satisfied.

TABLE II: Values of $PC$ and intervals of $M^2$ for the ground state light strange hybrid states at average value of the continuum threshold.

| $J^{PC}$ | $s_0(average)$ $(GeV^2)$ | $[M^2_{\rm min}, M^2_{\rm max}]$ $(GeV^2)$ | PC |
|---|---|---|---|
| $0^{++}$ | 19.0 | [5.0, 6.0] | [0.27, 0.22] |
| $0^{+-}$ | 17.0 | [4.5, 5.5] | [0.64, 0.48] |
| $0^{--}$ | 19.0 | [5.5, 6.5] | [0.22, 0.17] |
| $0^{-+}$ | 18.0 | [4.5, 5.5] | [0.68, 0.52] |
| $1^{++}$ | 10.0 | [4.5, 5.5] | [0.08, 0.05] |
| $1^{+-}$ | 9.0 | [3.0, 4.0] | [0.48, 0.27] |
| $1^{--}$ | 9.0 | [2.5, 3.5] | [0.46, 0.25] |
| $1^{-+}$ | 9.0 | [2.5, 3.5] | [0.62, 0.36] |

TABLE III: Values of $PC$ and intervals of $M^2$ for the ground state light nonstrange hybrid states at average value of the continuum threshold.

| $J^{PC}$ | $s_0(average)$ $(GeV^2)$ | $[M^2_{\rm min}, M^2_{\rm max}]$ $(GeV^2)$ | PC |
|---|---|---|---|
| $0^{++}$ | 19.0 | [5.5, 6.5] | [0.23, 0.18] |
| $0^{+-}$ | 17.0 | [4.5, 5.5] | [0.65, 0.48] |
| $0^{--}$ | 19.0 | [5.5, 6.5] | [0.23, 0.18] |
| $0^{-+}$ | 17.0 | [4.5, 5.5] | [0.65, 0.48] |
| $1^{++}$ | 10.0 | [4.5, 5.5] | [0.08, 0.05] |
| $1^{+-}$ | 9.0 | [2.5, 3.5] | [0.63, 0.36] |
| $1^{--}$ | 9.0 | [2.5, 3.5] | [0.48, 0.26] |
| $1^{-+}$ | 9.0 | [2.5, 3.5] | [0.63, 0.36] |

Having calculated the working intervals of the auxiliary parameters, we depict the variations of the mass $m_H$ and current coupling $f_H$ as functions of the Borel parameter and continuum threshold in Figs. 2 and 3. We observe that the physical quantities demonstrate good stability with respect to the variations of the parameters $M^2$ and $s_0$ in their working intervals. This is the case for other members and quantum numbers as well.

Now, we proceed to present the values of the masses and current couplings of the light hybrid mesons with the quantum numbers $0^{++}$, $0^{+-}$, $0^{--}$, $0^{-+}$, $1^{++}$, $1^{+-}$, $1^{--}$ and $1^{-+}$ for the strange and nonstrange members in tables IV-V, respectively. For completeness, we also depict the intervals of the Borel parameter $M^2$ and continuum threshold $s_0$ for all the channels in these tables. The uncertainties present in the values of the masses and current couplings are due to the uncertainties of the auxiliary parameters and errors of other input values. The order of

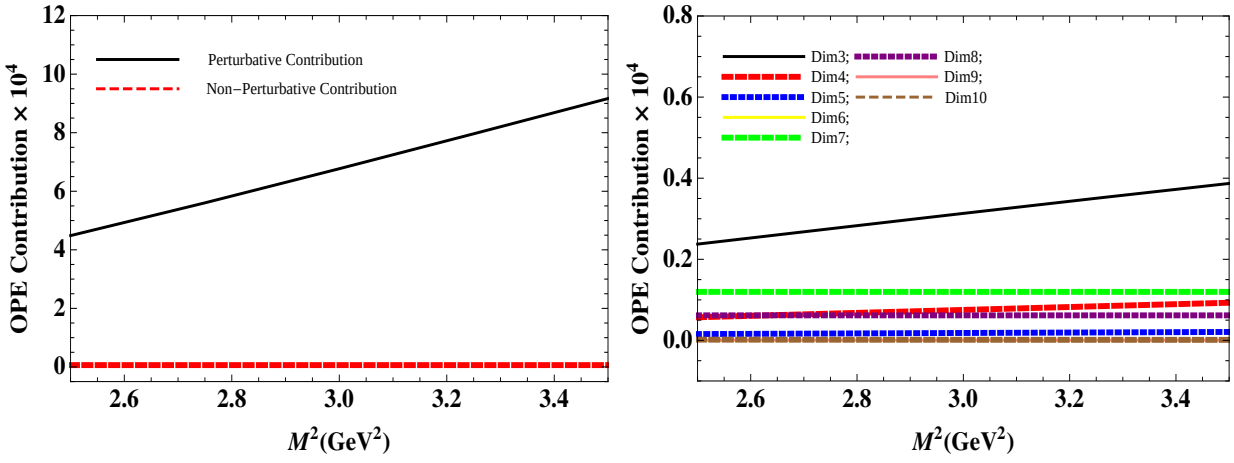



FIG. 1: The perturbative and different nonperturbative contributions (in units of $GeV^8$) for the strange hybrid meson of the $J^{PC} = 1^{--}$ quantum numbers as functions of $M^2$ and at average $s_0$.

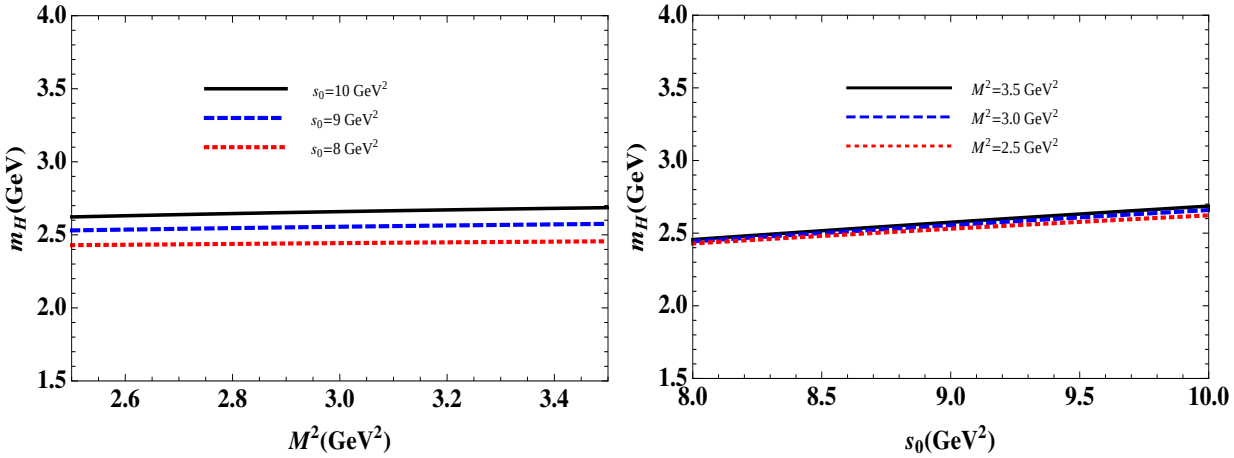



FIG. 2: The mass of the strange hybrid meson with $J^{PC} = 1^{--}$ as functions of $M^2$ and $s_0$.

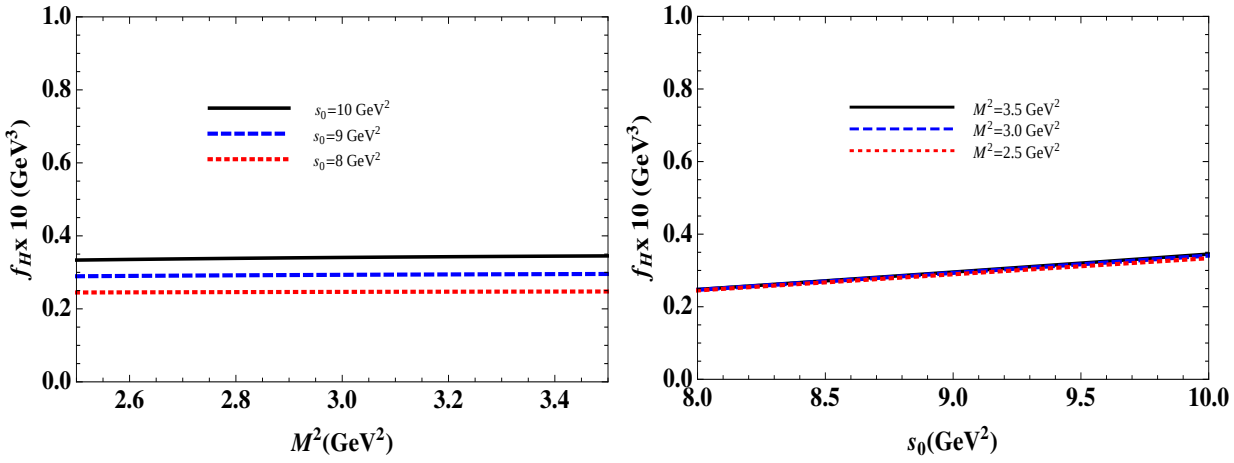



FIG. 3: The current coupling of the strange hybrid meson with $J^{PC} = 1^{--}$ as functions of $M^2$ and $s_0$.

uncertainties in the values of the masses are lower compared to those of the current coupling constants. This

TABLE IV: Values of the mass and current coupling for different quantum numbers of the ground state strange hybrid mesons.

| $J^{PC}$ | $M^2 \pm \delta M^2\ (GeV^2)$ | $s_0 \pm \delta s_0\ (GeV^2)$ | $m_H \pm \delta m_H\ (GeV)$ | $(f_H \pm \delta f_H) \times 10\ (GeV^3)$ |
|---|---|---|---|---|
| $0^{++}$ | $5.5 \pm 0.5$ | $19.0 \pm 1.0$ | $4.58^{+0.29}_{-0.17}$ | $0.63^{+0.11}_{-0.11}$ |
| $0^{+-}$ | $5.0 \pm 0.5$ | $17.0 \pm 1.0$ | $3.14^{+0.14}_{-0.13}$ | $0.65^{+0.05}_{-0.04}$ |
| $0^{--}$ | $6.0 \pm 0.5$ | $19.0 \pm 1.0$ | $4.64^{+0.29}_{-0.17}$ | $0.59^{+0.09}_{-0.10}$ |
| $0^{-+}$ | $5.0 \pm 0.5$ | $18.0 \pm 1.0$ | $3.20^{+0.14}_{-0.13}$ | $0.69^{+0.05}_{-0.05}$ |
| $1^{++}$ | $5.0 \pm 0.5$ | $10.0 \pm 1.0$ | $3.22^{+0.13}_{-0.07}$ | $0.26^{+0.06}_{-0.06}$ |
| $1^{+-}$ | $3.5 \pm 0.5$ | $9.0 \pm 1.0$ | $2.35^{+0.17}_{-0.16}$ | $0.38^{+0.05}_{-0.05}$ |
| $1^{--}$ | $3.0 \pm 0.5$ | $9.0 \pm 1.0$ | $2.55^{+0.14}_{-0.12}$ | $0.29^{+0.05}_{-0.04}$ |
| $1^{-+}$ | $3.0 \pm 0.5$ | $9.0 \pm 1.0$ | $2.31^{+0.18}_{-0.16}$ | $0.38^{+0.05}_{-0.04}$ |

TABLE V: Values of the mass and current coupling for different quantum numbers of the ground state nonstrange hybrid mesons.

| $J^{PC}$ | $M^2 \pm \delta M^2\ (GeV^2)$ | $s_0 \pm \delta s_0\ (GeV^2)$ | $m_H \pm \delta m_H\ (GeV)$ | $(f_H \pm \delta f_H) \times 10\ (GeV^3)$ |
|---|---|---|---|---|
| $0^{++}$ | $6.0 \pm 0.5$ | $19.0 \pm 1.0$ | $4.53^{+0.21}_{-0.13}$ | $0.58^{+0.08}_{-0.09}$ |
| $0^{+-}$ | $5.0 \pm 0.5$ | $17.0 \pm 1.0$ | $3.13^{+0.14}_{-0.13}$ | $0.65^{+0.05}_{-0.04}$ |
| $0^{--}$ | $6.0 \pm 0.5$ | $19.0 \pm 1.0$ | $4.52^{+0.21}_{-0.13}$ | $0.58^{+0.08}_{-0.08}$ |
| $0^{-+}$ | $5.0 \pm 0.5$ | $17.0 \pm 1.0$ | $3.14^{+0.14}_{-0.13}$ | $0.65^{+0.05}_{-0.04}$ |
| $1^{++}$ | $5.0 \pm 0.5$ | $10.0 \pm 1.0$ | $3.19^{+0.11}_{-0.07}$ | $0.26^{+0.06}_{-0.06}$ |
| $1^{+-}$ | $3.0 \pm 0.5$ | $9.0 \pm 1.0$ | $2.30^{+0.18}_{-0.17}$ | $0.38^{+0.06}_{-0.03}$ |
| $1^{--}$ | $3.0 \pm 0.5$ | $9.0 \pm 1.0$ | $2.51^{+0.15}_{-0.14}$ | $0.30^{+0.05}_{-0.04}$ |
| $1^{-+}$ | $3.0 \pm 0.5$ | $9.0 \pm 1.0$ | $2.30^{+0.18}_{-0.17}$ | $0.38^{+0.05}_{-0.04}$ |

is due the fact that the mass is ratio of two sum rules, whose uncertainties partly kill each other, while the current coupling is found from one sum rule and receives relatively more errors. The masses of the ground state light hybrid mesons for the strange and nonstrange members are available from different sources, as well. However, they have been presented either without uncertainties or are not available for all the members and quantum numbers. In this view, our results calculated with a higher accuracy can be a useful guide for the related experimental groups at different hadron colliders in their search for these yet unseen interesting particles. The comparison of our results for the strange and nonstrange members with other existing predictions in the literature are made in tables VI and VII, respectively. As it is seen from the tables, there is a consistency among the presented results in some channels within the uncertainties, but we see some serious differences of the presented values of different sources for some other channels. The previous studies done by QCD sum rules method are up to five, six, or eight mass dimensions of nonperturbative operators and do not include the uncertainties. This is the case for other frameworks as well. An exact comparison can be made, when the predictions from other studies/approaches are available for all the strange and nonstrange light hybrid mesons with different quantum numbers with their uncertainties.

## IV. SUMMARY AND CONCLUDING NOTES

The QCD as quantum field theory of strong interaction and the quark model as the most powerful model for the categorizing the particles made of quarks and gluons allow the existence of exotic states out of the usual mesons and baryons/antibaryons. The existence of exotic states of different quark-gluon configurations was predicted by Jaffe in the middle of 1970s. However, the first observation of exotic states was made in 2003 by discovery of the tetraquark $X(3872)$ by Belle collaboration. After that many tetraquark and pentaquark states of charmed-anticharmed content were observed by different collaborations. By the progresses made in the experiments on the identification of doubly heavy baryons, the discovery of the first doubly heavy charmed tetraquak $T_{cc}^{+}(3875)$ was reported by the LHCb collaboration in 2021 that was another revolution in discovery of the exotic states. Considering these developments, it is not far from waiting to see other categories of the exotic states like hybrid mesons, glueballs, etc.

Inspired by the experimental and theoretical progresses on the exotic states, we investigated the light strange and nonstrange hybrid mesons with quantum numbers $0^{++}$, $0^{+-}$, $0^{--}$, $0^{-+}$, $1^{++}$, $1^{+-}$, $1^{--}$ and $1^{-+}$ in the present study. We calculated the mass and current coupling of all the members up to a high accuracy, nonperturbative operator of ten dimension. We had the vector-scalar and axialvecor-pseudoscalar mixing as they couple to the same currents: They are separated by the tech-

TABLE VI: The masses of the ground state strange hybrid mesons in GeV and comparison of the results with other predictions.

| $J^{PC}$ | This work | Ref. [21] | Ref. [22] | Ref. [24] | Ref. [29] | Ref. [41] |
|---|---|---|---|---|---|---|
| $0^{++}$ | $4.58^{+0.29}_{-0.17}$ | 3.5 | – | – | – | - |
| $0^{+-}$ | $3.14^{+0.14}_{-0.13}$ | 3.4 | – | – | – | 3.57 |
| $0^{--}$ | $4.64^{+0.29}_{-0.17}$ | 3.7 | – | 3.8 | – | - |
| $0^{-+}$ | $3.20^{+0.14}_{-0.13}$ | 3.1 | – | – | – | - |
| $1^{++}$ | $3.22^{+0.13}_{-0.07}$ | 2.8 | – | – | – | - |
| $1^{+-}$ | $2.35^{+0.17}_{-0.16}$ | 2.8 | – | – | – | - |
| $1^{--}$ | $2.55^{+0.14}_{-0.12}$ | 2.9 | – | – | $2.54-2.62$ | - |
| $1^{-+}$ | $2.31^{+0.18}_{-0.16}$ | 2.5 | $\sim 1.75$ | $\sim 1.6-2.1$ | – | - |

TABLE VII: The masses of the ground state nonstrange hybrid mesons in GeV and comparison of the results with other predictions.

| $J^{PC}$ | This work | Ref. [20] | Ref. [22] | Ref. [23] | Ref. [25] | Ref. [27] | Ref. [29] | Ref. [30] | Ref. [31] | Ref. [32] | Ref. [41] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $0^{++}$ | $4.53^{+0.21}_{-0.13}$ | – | – | – | $2.35-3.4$ | $1.75-2.00$ | – | – | $>4$ | – | - |
| $0^{+-}$ | $3.13^{+0.14}_{-0.13}$ | – | – | $\sim 2.1-2.5$ | – | – | – | – | – | | 2.60 |
| $0^{--}$ | $4.52^{+0.21}_{-0.13}$ | $3.1\pm0.2$ | – | – | – | – | – | – | – | – | - |
| $0^{-+}$ | $3.14^{+0.14}_{-0.13}$ | – | – | – | 2.3 | – | – | – | – | – | - |
| $1^{++}$ | $3.19^{+0.11}_{-0.07}$ | – | – | – | – | – | – | – | – | – | - |
| $1^{+-}$ | $2.30^{+0.18}_{-0.17}$ | – | – | – | – | – | – | – | – | – | - |
| $1^{--}$ | $2.51^{+0.15}_{-0.14}$ | – | – | – | – | – | $2.33-2.43$ | – | – | – | - |
| $1^{-+}$ | $2.30^{+0.18}_{-0.17}$ | $1.7\pm0.1$ | $\sim 1.5$ | – | – | 1.55 | – | $1.86-2.2$ | $1.71\pm0.22$ | $1.72-2.60$ | - |

niques mentioned in section II. We found the working intervals of the auxiliary parameters based on the standard requirements of the method to make the physical observables like mass and current coupling possibly independent of these helping parameters. We obtained the stable sum rules and presented the numerical results for the mass and current coupling of all the strange and nonstrange members of $0^{++}$, $0^{+-}$, $0^{--}$, $0^{-+}$, $1^{++}$, $1^{+-}$, $1^{--}$ and $1^{-+}$ quantum numbers. We also presented and compared the previously existing predictions in the literature. We hope that, our results obtained with a high accuracy will help the experimental groups in their search for the light hybrid mesons. Comparison of any future experimental data on the considered states with our predictions will shed light on the nature, quark-gluon configuration and inner structure of these interesting particles.

## ACKNOWLEDGMENTS

The work of H. S. and B. B. was supported in part by the TUBITAK via the grant No: 123F197. K. A. thanks Iran national science foundation (INSF) for the partial financial support provided under the elites Grant No. 4025036. We dedicate this paper to Durmuş Ali Demir for his important contributions to developments of physics in the country: He left us too early.

## Appendix: Some expressions in the OPE side of the calculations for the strange hybrid meson of $J^{PC}=1^{--}$

In this appendix, as examples, we present the explicit expressions of the invariant functions $\rho^{\rm OPE}(s)$ and $\Gamma(M^2)$ for the light strange hybrid meson with quantum numbers $J^{PC}=1^{--}$. We can write,

$$\rho^{\rm OPE}(s)=\rho^{\rm pert.}(s)+\sum_{N=3}^{6}\rho^{\rm DimN}(s),\quad \Gamma(M^2)=\sum_{N=6}^{10}\Gamma^{\rm DimN}(M^2), \tag{A.1}$$

where,

$$\rho^{\rm DimN}(s)=\int_0^1 d\alpha\rho^{\rm DimN}(s), \tag{A.2}$$

and

$$\Gamma^{\rm DimN}(M^2)=\int_0^1 d\alpha\Gamma^{\rm DimN}(M^2,\alpha). \tag{A.3}$$

The perturbative and nonperturbative components of the spectral density are found as

$$\rho^{\rm pert.}(s)=\frac{g_s^2s^2(-5m_s^2+2s)}{2^9\cdot3\cdot5\pi^4}, \tag{A.4}$$

$$\rho^{\rm Dim3}(s)=\frac{g_s^2m_s\langle\overline{s}s\rangle s}{2^2\cdot3^2\pi^2}, \tag{A.5}$$

$$\rho^{\rm Dim4}(s)=\langle\alpha_sG^2/\pi\rangle\left[\frac{g_s^2s}{2^6\cdot3^3\pi^2}-\frac{(-9m_s^2+2s)}{2^5\cdot3^2}\right], \tag{A.6}$$

$$\rho^{\rm Dim5}(s)=-\frac{g_s^2m_o^2m_s\langle\overline{s}s\rangle}{2^6\cdot3\pi^2}, \tag{A.7}$$

and

$$\rho^{\rm Dim6}(s)=\frac{g_s^4\langle\overline{s}s\rangle^2}{2^4\cdot3^4\pi^2}. \tag{A.8}$$

Components of the function $\Gamma(M^2)$ are:

$$\Gamma^{\rm Dim6}(M^2)=-\frac{g_s^2m_s^2\langle\overline{s}s\rangle^2}{2^4\cdot3}, \tag{A.9}$$

$$\Gamma^{\rm Dim7}(M^2)=\frac{m_s\langle\alpha_sG^2/\pi\rangle\langle\overline{s}s\rangle}{2^3\cdot3^2}\left[\frac{-g_s^2}{2^2\cdot3}+7\pi^2\right], \tag{A.10}$$

$$\Gamma^{\rm Dim8}(M^2)=-\frac{g_s^2m_o^2\langle\overline{s}s\rangle^2}{2^3\cdot3^2}\left[1+\frac{m_s^2}{2M^2}\right]-5\frac{\langle\alpha_sG^2/\pi\rangle^2\pi^2}{2^6\cdot3^3}, \tag{A.11}$$

$$\Gamma^{\rm Dim9}(M^2)=\frac{m_s\langle\overline{s}s\rangle}{2^2\cdot3^3M^2}\left[\frac{g_s^4\langle\overline{s}s\rangle^2}{3^2}-\frac{m_o^2\pi^2\langle\alpha_sG^2/\pi\rangle}{2^2}\right], \tag{A.12}$$

$$\Gamma^{\rm Dim10}(M^2)=\frac{g_s^2\langle\overline{s}s\rangle^2}{2^2\cdot3^2M^2}\left[-\frac{m_o^4}{2^3}-\langle\alpha_sG^2/\pi\rangle\pi^2(\frac{1}{3^2}-\frac{1}{3^4})\right]. \tag{A.13}$$


[1] P. Adlarson *et al.* [WASA-at-COSY], "Neutron-proton scattering in the context of the $d^*(2380)$ resonance," Phys. Rev. C **90**, no.3, 035204 (2014) [arXiv:1408.4928 [nucl-ex]].

[2] K. Azizi, S. S. Agaev and H. Sundu, "The Scalar Hexaquark $uuddss$: a Candidate to Dark Matter?," J. Phys. G **47**, no.9, 095001 (2020) [arXiv:1904.09913 [hep-ph]].

[3] R. L. Jaffe and K. Johnson, "Unconventional States of Confined Quarks and Gluons,"Phys. Lett. B **60**, 201-204 (1976).

[4] R. L. Workman *et al.* [Particle Data Group], "Review of Particle Physics,"PTEP **2022**, 083C01 (2022).

[5] D. Alde *et al.* [IHEP-Brussels-Los Alamos-

Annecy(LAPP)], "Evidence for a $1^{-+}$ Exotic Meson,"Phys. Lett. B **205**, 397 (1988).
[6] G. S. Adams *et al.* [E852], "Observation of a new J(PC) = 1-+ exotic state in the reaction $\pi^- p \to \pi^+\pi^-\pi^- p$ at 18 GeV/c," Phys. Rev. Lett. **81**, 5760-5763 (1998).
[7] J. Kuhn *et al.* [E852], "Exotic meson production in the f(1)(1285) pi- system observed in the reaction $\pi^- p \to \eta\pi^+\pi^-\pi^- p$ at 18 GeV/c,' 'Phys. Lett. B **595**, 109-117 (2004), [arXiv:0401004 [hep-ex]].
[8] M. Ablikim *et al.* [BESIII], "Observation of an Isoscalar Resonance with Exotic $J^{PC} = 1^{-+}$ Quantum Numbers in $J/\psi \to \gamma\eta\eta'$," Phys. Rev. Lett. **129**, no.19, 192002 (2022), [arXiv:2202.00621 [hep-ex]].
[9] J. J. Dudek, "The lightest hybrid meson supermultiplet in QCD,"Phys. Rev. D **84**, 074023 (2011), [arXiv:1106.5515[hep-ph]].
[10] J. J. Dudek *et al.* [Hadron Spectrum], "Toward the excited isoscalar meson spectrum from lattice QCD,"Phys. Rev. D **88**, no.9, 094505 (2013), [arXiv:1309.2608[hep-lat]]. Copy to ClipboardDownload
[11] C. J. Burden, L. Qian, C. D. Roberts, P. C. Tandy and M. J. Thomson, "Ground state spectrum of light quark mesons,"Phys. Rev. C **55**, 2649-2664 (1997), [arXiv:9605027[nucl-th]].
[12] C. J. Burden and M. A. Pichowsky, "$J^{PC}$ exotic mesons from the Bethe-Salpeter equation,"Few Body Syst. **32**, 119-126 (2002), [arXiv:0206161[hep-ph]].
[13] T. Hilger, M. Gomez-Rocha and A. Krassnigg, "Light-quarkonium spectra and orbital-angular-momentum decomposition in a Bethe–Salpeter-equation approach,"Eur. Phys. J. C **77**, no.9, 625 (2017), [arXiv:1508.07183[hep-ph]].
[14] F. E. Close and P. R. Page, "The Production and decay of hybrid mesons by flux tube breaking,"Nucl. Phys. B **443**, 233-254 (1995), [arXiv:9411301[hep-ph]].
[15] T. Barnes, F. E. Close and E. S. Swanson, "Hybrid and conventional mesons in the flux tube model: Numerical studies and their phenomenological implications,"Phys. Rev. D **52**, 5242-5256 (1995), [arXiv:9501405[hep-ph]].
[16] T. Barnes and F. E. Close, "A LIGHT EXOTIC q anti-q g HERMAPHRODITE MESON?,"Phys. Lett. B **116**, 365-368 (1982).
[17] M. S. Chanowitz and S. R. Sharpe, "Hybrids: Mixed States of Quarks and Gluons,"Nucl. Phys. B **222**, 211-244 (1983).
[18] I. I. Balitsky, D. Diakonov and A. V. Yung, "EXOTIC MESONS WITH $J^{PC} = 1^{-+}$ FROM QCD SUM RULES,"Phys. Lett. B **112**, 71-75 (1982).
[19] J. Govaerts, F. de Viron, D. Gusbin and J. Weyers, "QCD Sum Rules and Hybrid Mesons,"Nucl. Phys. B **248**, 1-18 (1984).
[20] J. I. Latorre, S. Narison, P. Pascual and R. Tarrach, "Hermaphrodite Mesons and QCD Sum Rules,"Phys. Lett. B **147**, 169-174 (1984).
[21] J. Govaerts, L. J. Reinders and J. Weyers, "Radial Excitations and Exotic Mesons via QCD Sum Rules,"Nucl. Phys. B **262**, 575-592 (1985).
[22] I. I. Balitsky, D. Diakonov and A. V. Yung, "Exotic Mesons With $J^{PC} = 1^{-+}$, Strange and Non-strange,"Z. Phys. C **33**, 265-273 (1986).
[23] V. M. Braun and A. V. Kolesnichenko, "Exotic Scalar $J^{PC} = 0^{+-}$ Meson From QCD Sum Rules,"Phys. Lett. B **175**, 485-491 (1986).
[24] J. I. Latorre, P. Pascual and S. Narison, "Spectra and Hadronic Couplings of Light Hermaphrodite Mesons,"Z. Phys. C **34**, 347 (1987).
[25] T. Huang, H. Y. Jin and A. L. Zhang, "The $0^{++}$ and $0^{-+}$ Mass of Light Quark Hybrid in QCD Sum Rules,"Eur. Phys. J. C **8**, 465-471 (1999), [arXiv:9809331[hep-ph]].
[26] K. G. Chetyrkin and S. Narison, "Light hybrid mesons in QCD,"Phys. Lett. B **485**, 145-150 (2000), [arXiv:0003151[hep-ph]].
[27] H. Y. Jin and J. G. Korner, "Radiative corrections to the correlator of ($0^{++}$, $1^{-+}$) light hybrid currents,"Phys. Rev. D **64**, 074002 (2001), [arXiv:0003202[hep-ph]].
[28] H. Y. Jin, J. G. Korner and T. G. Steele, "Improved determination of the mass of the $1^{-+}$ light hybrid meson from QCD sum rules,"Phys. Rev. D **67**, 014025 (2003), [arXiv:0211304 [hep-ph]].
[29] F. K. Guo, P. N. Shen, Z. G. Wang, W. H. Liang and L. S. Kisslinger, "Light vector hybrid states via QCD sum rules,"[arXiv:hep-ph/0703062 [hep-ph]].
[30] S. Narison, "$1^{-+}$ light exotic mesons in QCD,"Phys. Lett. B **675**, 319-325 (2009), [arXiv:0903.2266[hep-ph]].
[31] Z. F Zhang, H. Y. Jin and T. G. Steele, "Revisiting $1^{-+}$ and $0^{++}$ light hybrids from Monte-Carlo based QCD sum rules,"Chin. Phys. Lett. **31**, 051201 (2014), [arXiv:1312.5432[hep-ph]].
[32] Z. R. Huang, H. Y. Jin and Z. F. Zhang, "New predictions on the mass of the $1^{-+}$ light hybrid meson from QCD sum rules,"JHEP **04**, 004 (2015), [arXiv:1411.2224[hep-ph]].
[33] Z. R. Huang, H. Y. Jin, T. G. Steele and Z. F. Zhang, "Revisiting the $b_1\pi$ and $\rho\pi$ decay modes of the $1^{-+}$ light hybrid state with light-cone QCD sum rules," Phys. Rev. D **94**, no.5, 054037 (2016), [arXiv:1608.03028[hep-ph]].
[34] Z. R. Huang, H. Y. Jin, T. G. Steele and Z. F. Zhang, "The Mass and Decay Properties of the $1^{-+}$ Light Hybrid Meson,"Nucl. Part. Phys. Proc. **294-296**, 113-118 (2018), [arXiv:1709.10233[hep-ph]].
[35] J. Govaerts, L. J. Reinders, H. R. Rubinstein and J. Weyers, "Hybrid quarkonia from QCD sum rules,"Nucl. Phys. B **258**, 215-229 (1985).
[36] J. Govaerts, L. J. Reinders, P. Francken, X. Gonze and J. Weyers, "Coupled QCD Sum Rules for Hybrid Mesons,"Nucl. Phys. B **284**, 674 (1987).
[37] F. J. Llanes-Estrada and S. R. Cotanch, "Coulomb gauge hybrid meson calculation,"Nucl. Phys. B Proc. Suppl. **90**, 117-119 (2000).
[38] L. Bellantuono, P. Colangelo and F. Giannuzzi, "Exotic $J^{PC} = 1^{-+}$ mesons in a holographic model of QCD,"Eur. Phys. J. C **74**, no.4, 2830 (2014), [arXiv:1402.5308[hep-ph]].
[39] J. N. Hedditch, W. Kamleh, B. G. Lasscock, D. B. Leinweber, A. G. Williams and J. M. Zanotti, "$1^{-+}$ exotic meson at light quark masses,"Phys. Rev. D **72**, 114507 (2005), [arXiv:0509106[hep-lat]].
[40] C. McNeile *et al.* [UKQCD], "Decay

width of light quark hybrid meson from the lattice," Phys. Rev. D **73**, 074506 (2006), [arXiv:0603007[hep-lat]].

[41] J. Ho, R. Berg, W. Chen, D. Harnett and T. G. Steele, "Mass Calculations of Light Quarkonium, Exotic $J^{PC} = 0^{+-}$ Hybrid Mesons from Gaussian Sum-Rules," Phys. Rev. D **98**, no.9, 096020 (2018), [arXiv:1806.02465[hep-ph]].

[42] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, "QCD and Resonance Physics. Theoretical Foundations," Nucl. Phys. B **147**, 385-447 (1979).

[43] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, "QCD and Resonance Physics: Applications," Nucl. Phys. B **147**, 448-518 (1979).

[44] L. J. Reinders, H. Rubinstein and S. Yazaki, "Hadron Properties from QCD Sum Rules," Phys. Rept. **127**, 1 (1985).

[45] S. Narison, "QCD spectral sum rules," World Sci. Lect. Notes Phys. **26**, 1-527 (1989).

[46] Z. G. Wang, "Analysis of the vector tetraquark states with P-waves between the diquarks and antidiquarks via the QCD sum rules," Eur. Phys. J. C **79**, no.1, 29 (2019), [arXiv:1811.02726 [hep-ph]].

[47] J. Wu, X. Liu, Y. R. Liu and S. L. Zhu, "Systematic studies of charmonium-, bottomonium-, and $B_c$-like tetraquark states," Phys. Rev. D **99**, no.1, 014037 (2019), [arXiv:1810.06886 [hep-ph]].

[48] M. B. Voloshin, "$Z_c(4100)$ and $Z_c(4200)$ as hadrocharmonium," Phys. Rev. D **98**, no.9, 094028 (2018), [arXiv:1810.08146 [hep-ph]].

[49] X. Cao and J. P. Dai, "Spin parity of $Z_c^-(4100)$, $Z_1^+(4050)$ and $Z_2^+(4250)$," Phys. Rev. D **100**, no.5, 054004 (2019), [arXiv:1811.06434 [hep-ph]].

[50] S. Agaev, K. Azizi and H. Sundu, "Four-quark exotic mesons," Turk. J. Phys. **44**, no.2, 95-173 (2020), [arXiv:2004.12079 [hep-ph]].

[51] S. S. Agaev, K. Azizi and H. Sundu, "Resonance X(3960) as a hidden charm-strange scalar tetraquark," Phys. Rev. D **107**, no.5, 054017 (2023), [arXiv:2211.14129 [hep-ph]].

[52] S. S. Agaev, K. Azizi, B. Barsbay and H. Sundu, "Exploring fully heavy scalar tetraquarks $QQ\overline{QQ}$," Phys. Lett. B **844**, 138089 (2023), [arXiv:2304.03244 [hep-ph]].

[53] B. Barsbay, K. Azizi and H. Sundu, "Heavy-light hybrid mesons with different spin-parities," Eur. Phys. J. C **82**, no.12, 1086 (2022), [arXiv:2205.14597 [hep-ph]].

[54] S. S. Agaev, K. Azizi and H. Sundu, "Newly observed exotic doubly charmed meson $T_{cc}^+$," Nucl. Phys. B **975**, 115650 (2022), [arXiv:2108.00188 [hep-ph]].

[55] Z. G. Wang and Q. Xin, "Analysis of hidden-charm pentaquark molecular states with and without strangeness via the QCD sum rules ," Chin. Phys. C **45**, no.12, 123105 (2021), [arXiv:2103.08239 [hep-ph]].

[56] K. Azizi, Y. Sarac and H. Sundu, "Investigation of the strange pentaquark candidate $P_{\psi s}^{\Lambda}(4338)^0$ recently observed by LHCb," Phys. Rev. D **108**, no.7, 074010 (2023), [arXiv:2304.00604 [hep-ph]].

[57] Z. G. Wang, H. J. Wang and Q. Xin, "The hadronic coupling constants of the lowest hidden-charm pentaquark state with the QCD sum rules in rigorous quark-hadron duality," Chin. Phys. C **45**, 063104 (2021), [arXiv:2005.00535 [hep-ph]].

[58] Z. G. Wang, "Analysis of the $\bar{D}\Sigma_c$, $\bar{D}\Sigma_c^*$, $\bar{D}^*\Sigma_c$ and $\bar{D}^*\Sigma_c^*$ pentaquark molecular states with QCD sum rules," Int. J. Mod. Phys. A **34**, no.19, 1950097 (2019), [arXiv:1806.10384 [hep-ph]].

[59] Z. G. Wang, "Strong decays of the $Y(4660)$ as a vector tetraquark state in solid quark-hadron duality," Eur. Phys. J. C **79**, no.3, 184 (2019), [arXiv:1901.02177 [hep-ph]].

[60] Z. G. Wang and X. Wang, "Analysis of the strong decays of the $P_c(4312)$ as a pentaquark molecular state with QCD sum rules," Chin. Phys. C **44**, 103102 (2020), [arXiv:1907.04582 [hep-ph]].

arXiv:2402.18906v1 [hep-ph] 29 Feb 2024

# Neutrino Mixing and Resonant Leptogenesis in Inverse Seesaw and $\Delta(54)$ Flavor Symmetry

Hrishi Bora[1*], Ng. K. Francis[1†], Bikash Thapa[1‡] and Shawan Kumar Jha[2§]

[1] *Department of Physics, Tezpur University, Tezpur - 784028, India*

[2] *Department of Physics, Indian Institute of Technology Guwahati, Guwahati 781039, India*

## Abstract

The current work involves augmenting the $\Delta(54)$ discrete flavor model by incorporating two Standard Model Higgs particles into the Inverse Seesaw mechanism. We introduced Weyl fermions and Vector like fermions, which are gauge singlets in the Standard Model and produces Majorana mass terms in our lagrangian. The resulting mass matrix deviates from the tribimaximal neutrino mixing pattern producing a non-zero reactor angle ($\theta_{13}$) . We have determined the effective Majorana neutrino mass, which is the parameter of relevance in neutrinoless double beta decay investigations, using the model's limited six-dimensional parameter space. We additionally investigate the possibility of baryogenesis in the proposed framework via resonant leptogenesis. We have the non-zero value for resonantly enhanced CP asymmetry originating from the decay of right-handed neutrinos at the TeV scale, accounting for flavor effects. The evolution of lepton asymmetry is systematically analyzed by numerically solving a set of Boltzmann equations, leading to the determination of the baryon asymmetry with a magnitude of $|\eta_B| \approx 6 \times 10^{-10}$. This outcome is achieved by selecting specific values for the right-handed neutrino mass $M_1 = 10$ TeV and mass splitting, $d \approx 10^{-8}$.

***Keywords***— Majorana neutrinos, Double inverse seesaw, Jarlskog invariant, Tribimaximal neutrino mixing, Neutrinoless double-beta decay

PACS numbers: 12.60.-i, 14.60.Pq, 14.60.St

*hrishi@tezu.ernet.in (Corresponding author)

†francis@tezu.ernet.in

‡bikash2@tezu.ernet.in

§shawankumar@iitg.ac.in

# 1. Introduction

The neutrino masses along with their flavor mixing as observed in neutrino oscillations, leads to a question about where these tiny masses come from [1–5]. Since the standard model do not include right-handed neutrinos unlike other fermions, it is unlikely that neutrino masses work the same way as the masses of charged fermions. The origin of neutrino masses can be explained by various frameworks beyond the standard model (BSM), including the seesaw mechanism [6–8], radiative seesaw mechanism [9], extra-dimensional models [10, 11] and others. The extension of the standard model with Inverse Seesaw mechanism can explain the observed Baryon Asymmetry of the Universe (BAU) through leptogenesis [12]. These references includes many current reviews on neutrino physics [13–29].

In leptogenesis, the asymmetry in leptons, obtained from the CP-violating decay of heavy right-handed neutrinos, is transformed into an asymmetry in baryons through sphaleron processes [30]. According to Ref. [31], a mass scale of around $\mathcal{O}(10^9)$ for the right-handed neutrino is necessary to explain the observed BAU. However, this requirement can be lessened if the masses of right-handed neutrinos are nearly the same. In such cases, the effects that violate CP symmetry become significantly amplified, and with relatively low masses (TeV scale), sufficient asymmetry in leptons can be generated to account for the Baryon Asymmetry of the Universe (BAU). This condition is termed resonant leptogenesis. It is important to mention that recent research, utilizing the $SU(5) \times \mathcal{T}_{13}$ model [32], has shown the possibility of resonant leptogenesis at the GeV–TeV scale within the type-I seesaw model, considering active sterile mixing within the sensitivity range of DUNE. Additionally, considerable attention has been devoted over time to investigating the origin of neutrino flavor mixing. Among the available explanations, Tri-bimaximal mixing (TBM) appears to be the most probable. However, experimental results from Daya Bay, RENO, and Double Chooz suggest that TBM needs to be adjusted to incorporate a non-zero value for $\theta_{13}$. In this study, we introduce a model constructed within the minimum seesaw model framework, employing $\Delta(54)$ discrete symmetry. By employing resonant leptogenesis, the resulting mass matrix can potentially explain the Baryon Asymmetry of the Universe (BAU) concurrently. To achieve successful resonant leptogenesis, we introduce a higher-order term. We specifically choose the Majorana mass matrix for right-handed neutrinos, $M_R$, so that these neutrinos have degenerate masses at the dimension five-level. In essence, our work expands upon the model proposed in [33], making it suitable for investigating resonant leptogenesis in scenarios involving the minimum seesaw model.

Similar study on resonant leptogenesis, utilizing $S_4$ symmetry within the minimum seesaw model has been carried out in Ref. [34]. However, in contrast to our current research, the models discussed in Ref. [34] achieve resonant leptogenesis differently. They achieve this by creating mass differences among the heavy right-handed neutrinos through minimal seesaw model based on $S_4$ discrete flavor symmetry that leads to $\mathrm{TM}_1$ mixing. We explore the investigation of resonant leptogenesis within the inverse seesaw model based on $\Delta(54)$ discrete symmetry, while considering the discovery of a non-zero $\theta_{13}$.

This paper is structured as follows: In Section 2, we introduce the $\Delta(54)$ discrete symmetry with inverse seesaw mechanism and discuss the characteristics of the flavor group relevant to constructing the model. Section 3 outlines the allowable range for model parameters based on the constraints imposed by the $3\sigma$ range of neutrino oscillation data. The numerical solution of the Boltzmann equations, which govern the evolution of lepton number density and the baryon asymmetry parameter, is presented in Section 4, along with the framework for resonant leptogenesis.

In Section 5, we conclude our study and provide numerical results regarding neutrinoless double beta decay within the model.

## 2. Framework of the Model

Extending the fermion sector within the Standard Model framework is necessary to achieve the implementation of the Inverse Seesaw mechanism. Here, we have introduced Vector-like (VL) fermions, $N_1$ and $N_2$, which have the property of being gauge singlets inside the Standard Model framework. After symmetry breaking, this new piece generates a Majorana mass ($M_N$) term which is negligible under the seesaw hierarchy $M_N, M_S << m_{\nu N} << M_{NS}$. We also introduced a Weyl fermion denoted as $S_1$. In fact, the $\phi$ VEV induces a Majorana mass term for the $S_1$ fermion. The fields associated with right-handedness and left-handedness are indicated by subscripts 1 and 2 respectively. The $\Delta(54)$ group includes irreducible representations $1_1$, $1_2$, $2_1$, $2_2$, $2_3$, $2_4$, $3_{1(1)}$, $3_{1(2)}$, $3_{2(1)}$ and $3_{2(2)}$. The products of $3_{2(1)} \otimes 3_{2(2)}$, $3_{1(1)} \otimes 3_{1(2)}$ , $3_{1(2)} \otimes 3_{2(1)}$ and $3_{1(1)} \otimes 3_{2(2)}$ lead to the trivial singlet.

The rules for multiplication are as follow [35]:

$$3_{1(1)} \otimes 3_{1(1)} = 3_{1(2)} \oplus 3_{1(2)} \oplus 3_{2(2)}$$

$$3_{1(2)} \otimes 3_{1(2)} = 3_{1(1)} \oplus 3_{1(1)} \oplus 3_{2(1)}$$

$$3_{2(1)} \otimes 3_{2(1)} = 3_{1(2)} \oplus 3_{1(2)} \oplus 3_{2(2)}$$

$$3_{2(2)} \otimes 3_{2(2)} = 3_{1(1)} \oplus 3_{1(1)} \oplus 3_{2(1)}$$

$$3_{1(1)} \otimes 3_{1(2)} = 1_1 \oplus 2_1 \oplus 2_2 \oplus 2_3 \oplus 2_4$$

$$3_{1(2)} \otimes 3_{2(1)} = 1_2 \oplus 2_1 \oplus 2_2 \oplus 2_3 \oplus 2_4$$

$$3_{2(1)} \otimes 3_{2(2)} = 1_1 \oplus 2_1 \oplus 2_2 \oplus 2_3 \oplus 2_4$$

$$3_{1(1)} \otimes 3_{2(2)} = 1_2 \oplus 2_1 \oplus 2_2 \oplus 2_3 \oplus 2_4$$

| Field | L | $l$ | $H_1$ | $H_2$ | $N_1$ | $N_2$ | $S_1$ | $\chi$ | $\chi'$ | $\zeta$ | $\zeta'$ | $\xi$ | $\Phi_S$ | $\phi$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Delta(54)$ | $3_{1(1)}$ | $3_{2(2)}$ | $1_1$ | $1_2$ | $3_{1(1)}$ | $3_{2(1)}$ | $3_{2(2)}$ | $1_2$ | $2_1$ | $1_2$ | $1_1$ | $3_{2(1)}$ | $3_{1(1)}$ | $3_{1(2)}$ |
| $Z_2$ | 1 | -1 | 1 | 1 | -1 | 1 | 1 | -1 | -1 | -1 | 1 | -1 | -1 | 1 |
| $Z_3$ | $\omega$ | $\omega$ | 1 | 1 | 1 | $\omega$ | 1 | 1 | 1 | 1 | $\omega$ | $\omega$ | $\omega$ | 1 |
| $Z_4$ | 1 | -1 | 1 | 1 | 1 | -1 | 1 | -1 | -1 | 1 | -1 | 1 | 1 | 1 |
| U(1) | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 1: Particle content of our model

We have developed a model based on the $\Delta(54)$ model, which incorporates the inclusion of additional flavons, namely $\chi$, $\chi'$, $\xi$, $\zeta$, $\zeta'$, $\Phi_S$, and $\phi$. In order to avoid undesired terms, we added additional symmetry $Z_2 \otimes Z_3 \otimes Z_4$. Details on the particle composition and associated charge assignment according to the symmetry group are given in Table 1. The left-handed leptons doublets and the right-handed charged lepton are assigned using the triplet representation of $\Delta(54)$.

The Lagrangian is as follows :

$$\begin{aligned}\mathcal{L} =&\frac{y_1}{\Lambda}(l\bar{L})\chi H_1 + \frac{y_2}{\Lambda}(l\bar{L})\chi' H_1 + \frac{\bar{L}\tilde{H}_2 N_1}{\Lambda}y_\xi \xi + \frac{\bar{L}\tilde{H}_1 N_1}{\Lambda}y_s \Phi_s + \frac{\bar{L}\tilde{H}_2 N_1}{\Lambda}y_a \Phi_s \\ &+ y_{NS}\bar{N}_1 S_1 \zeta + y'_{NS}\bar{N}_2 S_1 \zeta' + \frac{y_{s_1}}{\Lambda}\bar{S}_1 S_1 \phi + \frac{y_{s_2}}{\Lambda^2}\bar{S}_1 S_1 \phi \end{aligned} \tag{1}$$

The vacuum expectation values are considered naturally as,

$$\langle\chi\rangle = (v_\chi) \qquad \langle\chi'\rangle = (v_{\chi'}, v_{\chi'}) \qquad \langle\xi\rangle = (v_\xi, v_\xi, v_\xi) \qquad \langle\phi\rangle = (v_\phi, v_\phi, v_\phi)$$
$$\langle\Phi_S\rangle = (v_s, v_s, v_s) \qquad \langle\zeta\rangle = (v_\zeta) \qquad \langle\zeta'\rangle = (v'_\zeta)$$

The charged lepton mass matrix is given as [36]

$$M_l = \frac{y_1 v}{\Lambda}\begin{pmatrix} v_\chi & 0 & 0 \\ 0 & v_\chi & 0 \\ 0 & 0 & v_\chi \end{pmatrix} + \frac{y_2 v}{\Lambda}\begin{pmatrix} -\omega v_{\chi'} + v_{\chi'} & 0 & 0 \\ 0 & -\omega^2 v_{\chi'} + \omega^2 v_{\chi'} & 0 \\ 0 & 0 & -v_{\chi'} + \omega v_{\chi'} \end{pmatrix}$$

where $y_1$ and $y_2$ are coupling constants.

## 2.1 Effective neutrino mass matrix

After applying $\Delta(54)$ and electroweak symmetry breaking, the mass matrices related to the neutrino sector may be derived using the above-mentioned Lagrangian. The fundamental assumption of the ISS theory is the small $M_S$ scale, which guarantees small neutrino masses. The $M_S$ scale must be at the KeV level in order to decrease the right-handed neutrino masses to the TeV scale. The inverse seesaw model is a TeV-scale seesaw model that maintains compatibility with light neutrino masses in the sub-eV range while allowing heavy neutrinos to remain as light as a TeV and Dirac masses to be as large as those of charged leptons.

$$M_{NS} = y_{_{NS}}\begin{pmatrix} v_\zeta & 0 & 0 \\ 0 & v_\zeta & 0 \\ 0 & 0 & v_\zeta \end{pmatrix} \tag{2}$$

$$M_S = \frac{y_{s_1}}{\Lambda^2}\begin{pmatrix} v_\phi & 0 & 0 \\ 0 & v_\phi & 0 \\ 0 & 0 & v_\phi \end{pmatrix} \tag{3}$$

$$M'_{NS} = y'_{_{NS}}\begin{pmatrix} v'_\zeta & 0 & 0 \\ 0 & v'_\zeta & 0 \\ 0 & 0 & v'_\zeta \end{pmatrix} \tag{4}$$

$$M_{\nu N} = \frac{v}{\Lambda}\begin{pmatrix} y_\xi v_\xi & y_s v_s + y_a v_a & y_s v_s - y_a v_a \\ y_s v_s - y_a v_a & y_\xi v_\xi & y_s v_s + y_a v_a \\ y_s v_s + y_a v_a & y_s v_s - y_a v_a & y_\xi v_\xi \end{pmatrix} \tag{5}$$

In the inverse seesaw framework, the effective neutrino mass matrix can be written as

$$m_\nu = M_{\nu N}(M'_{NS})^{-1} M_S M_{NS}^{-1} M'_{\nu N} \tag{6}$$

$$m_\nu = M_{\nu N} M_{mid}^{-1} M_{\nu N}^T \quad \text{with} \quad M_{mid} = M'_{NS} M_S^{-1} M_{NS} \tag{7}$$

$$M_{mid} = \begin{pmatrix} M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & M \end{pmatrix} \tag{8}$$

$$M_{\nu N} = \begin{pmatrix} x & c+a & c-a \\ c-a & x & c+a \\ c+a & c-a & x \end{pmatrix} \tag{9}$$

where $x = y_\xi v_\xi$, $c = y_s v_s$, $a = y_a v_s$, $M = \frac{\Lambda^2 y_{NS} {y'}_{NS} v_\varsigma v'_\varsigma}{v_\phi y_{s_1}}$.

The resultant mass matrix from Eq. (6)

$$m_\nu = \frac{v^2}{M\Lambda^2} \begin{pmatrix} 2a^2+2c^2+x^2 & -a^2+c^2+2cx & -a^2+c^2+2cx \\ -a^2+c^2+2cx & 2a^2+2c^2+x^2 & -a^2+c^2+2cx \\ -a^2+c^2+2cx & -a^2+c^2+2cx & 2a^2+2c^2+x^2 \end{pmatrix} \tag{10}$$

$$m_\nu = \frac{1}{M} \begin{pmatrix} 2a'^2+2c'^2+x'^2 & -a'^2+c'^2+2c'x' & -a'^2+c'^2+2c'x' \\ -a'^2+c'^2+2c'x' & 2a'^2+2c'^2+x'^2 & -a'^2+c'^2+2c'x' \\ -a'^2+c'^2+2c'x' & -a'^2+c'^2+2c'x' & 2a'^2+2c'^2+x'^2 \end{pmatrix} \tag{11}$$

where $a' = \frac{av}{\Lambda}$, $c' = \frac{cv}{\Lambda}$, $x' = \frac{xv}{\Lambda}$. The dimension of the problem is absorbed by the term $M$ and the components of the matrix are unaffected . Phase redefinitions of the charged lepton fields can absorb the phase of $M$, allowing it to be considered as a real parameter without losing generality [37].

In the following sections, we have presented the numerical approaches and discussed baryogenesis via resonant leptogenesis, within the context of our model.

## 3. Numerical Analysis

The mass matrix in Eq. (11) gives the effective neutrino mass matrix in terms of the model complex parameters $a'$, $c'$, and $x'$. We find the values of the model parameters by fitting the model to the current neutrino oscillation data. We use the $3\sigma$ interval for the neutrino oscillation parameters ( $\theta_{12}, \theta_{23}, \theta_{13}, \Delta m^2_{21}, \Delta m^2_{31}$ ) as presented in Table 2. A further constraint on the model parameters was applied on the sum of absolute neutrino masses from the cosmological bound $\sum_i m_i < 0.12eV$. In our analysis, the three complex parameters of the model are treated as free parameters and are allowed to run over the following ranges: $|a'| \in [0, 0.3]eV$, $\phi_a \in [-\pi, \pi]$ ; $|c'| \in [0, 10^{-1}]eV$ ; $\phi_c \in [-\pi, \pi]$ ; $|x'| \in [0, 10^{-3}]eV$ , $\phi_x \in [-\pi, \pi]$

where $\phi_a$ , $\phi_c$ and $\phi_x$ are the phases.

| Parameters | NH ($3\sigma$) | IH ($3\sigma$) |
|---|---|---|
| $\Delta m^2_{21}[10^{-5}eV^2]$ | $6.82 \to 8.03$ | $6.82 \to 8.03$ |
| $\Delta m^2_{31}[10^{-3}eV^2]$ | $2.428 \to 2.597$ | $-2.581 \to -2.408$ |
| $\sin^2\theta_{12}$ | $0.270 \to 0.341$ | $0.270 \to 0.341$ |
| $\sin^2\theta_{13}$ | $0.02029 \to 0.02391$ | $0.02047 \to 0.02396$ |
| $\sin^2\theta_{23}$ | $0.406 \to 0.620$ | $0.410 \to 0.623$ |
| $\delta_{CP}$ | $108 \to 404$ | $192 \to 360$ |

Table 2: The neutrino oscillation parameters from NuFIT 5.2 (2022)

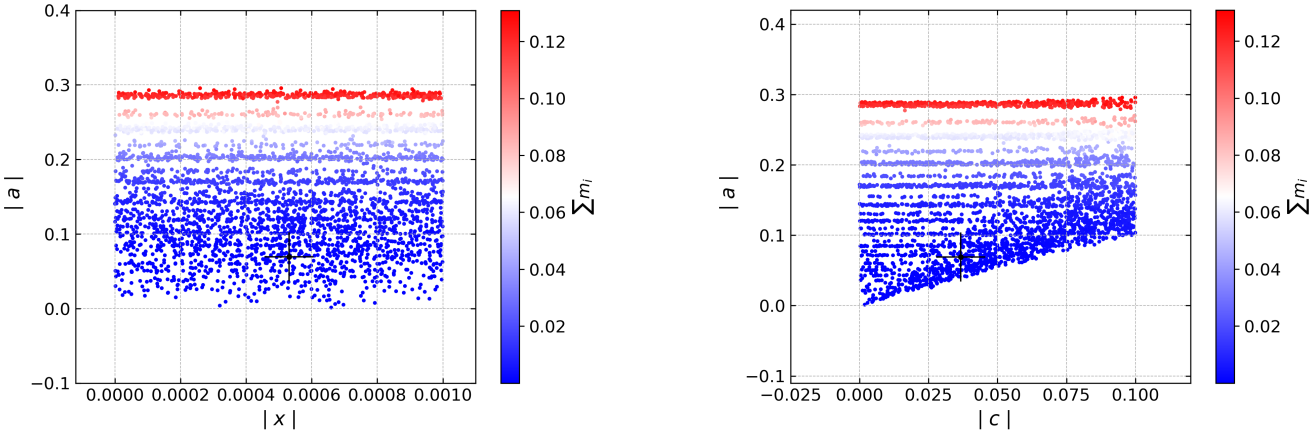



Figure 1: Left and right panel shows the correlation of model parameters along with the variation of $\sum m_i$. The black marker indicate the best-fit values.

The neutrino mass matrix $m_\nu$ is diagonalized by the PMNS matrix $U$ as follow [37]:

$$U^\dagger m_\nu^{(i)} U^* = \text{diag}(m_1, m_2, m_3) \tag{12}$$

We numerically calculated $U$ using the relation $U^\dagger M_\nu U = \text{diag}(m_1^2, m_2^2, m_3^2)$, where $M_\nu = m_\nu m_\nu^\dagger$. The neutrino oscillation parameters $\theta_{12}$, $\theta_{13}$, $\theta_{23}$ and $\delta$ can be obtained from $U$ as

$$s_{12}^2 = \frac{|U_{12}|^2}{1-|U_{13}|^2}, \qquad s_{13}^2 = |U_{13}|^2, \qquad s_{23}^2 = \frac{|U_{23}|^2}{1-|U_{13}|^2} \tag{13}$$

and $\delta$ may be given by

$$\delta = \sin^{-1}\left(\frac{8\,\text{Im}(h_{12}h_{23}h_{31})}{P}\right) \tag{14}$$

with

$$P = (m_2^2 - m_1^2)(m_3^2 - m_2^2)(m_3^2 - m_1^2)\sin 2\theta_{12}\sin 2\theta_{23}\sin 2\theta_{13}\cos\theta_{13} \tag{15}$$

We adjusted the modified $\Delta(54)$ model to suit the experimental data by minimizing the ensuing $\chi^2$ function in order to evaluate how the neutrino mixing parameters contrast with the most current experimental data:

$$\chi^2 = \sum_i \left(\frac{\lambda_i^{model} - \lambda_i^{expt}}{\Delta\lambda_i}\right)^2, \tag{16}$$

where $\lambda_i^{model}$ is the $i^{th}$ observable predicted by the model, $\lambda_i^{expt}$ stands for $i^{th}$ experimental best-fit value and

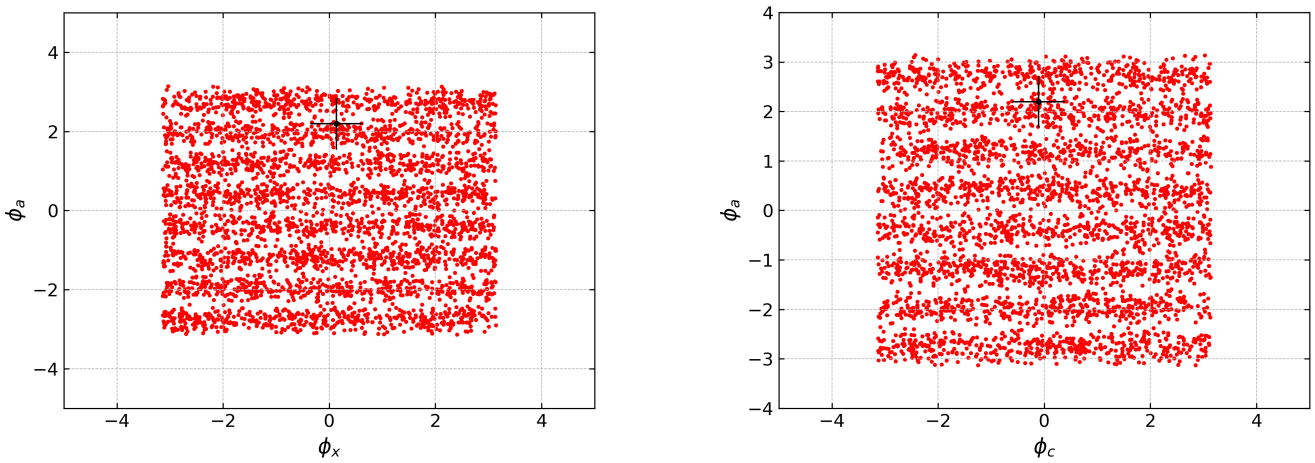



Figure 2: Correlation between the phases $\phi_x$ , $\phi_c$ and $\phi_a$. The black marker indicate the best-fit values corresponding to $\chi^2$ min.

$\Delta\lambda_i$ is the $1\sigma$ range of the observable.

The best-fit values for $|x'|$, $|c'|$, $|a'|$, $\phi_x$, $\phi_c$ and $\phi_a$ obtained are (0.00054, 0.03667, 0.06914, 0.13258$\pi$, -0.10508$\pi$, 2.23488$\pi$).

Correspondingly, the best-fit values for the neutrino oscillation parameters are: $sin^2\theta_{12} = 0.31940$, $\sin^2\theta_{13} = 0.02394$, $\sin^2\theta_{23} = 0.51231$, $\sin\delta_{CP} = 0.094$. The best-fit values for other parameters, such as $\Delta m^2_{21}/\Delta m^2_{31}$ is 0.029, are correspond to the $\chi^2$-minimum.

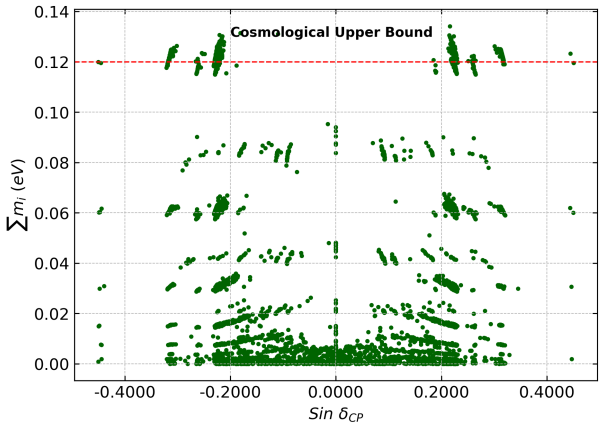



Figure 3: Correlation between $\sum m_i$ with sin $\delta_{CP}$. The red dotted horizontal line gives the upper limit of absolute neutrino mass.

## 4. Resonant Leptogenesis

Fukugita and Yanagida originally proposed the leptogenesis mechanism, which is one of the most commonly accepted explanations for the Baryon Asymmetry of the Universe(BAU). The mass of the lightest right-handed neutrino, $M_1 = 10^9 GeV$, has a lower bound in the simplest case of thermal leptogenesis with a hierarchical mass spectrum

of right-handed neutrinos [31]. Although one can lower this limit if their masses are nearly degenerate. This scenario is popularly known as resonant leptogenesis [38,39]. In this scenario, the resonant enhancement amplifies the one-loop self-energy contribution, leading to the flavor-dependent asymmetry resulting from the decay of a right-handed neutrino into a lepton and Higgs.

$$\epsilon_{i\alpha} = \frac{\Gamma(N_i \to l_\alpha + H) - \Gamma(N_i \to \bar{l}_\alpha + \bar{H})}{\sum_\alpha \left(\Gamma(N_i \to l_\alpha + H) + \Gamma(N_i \to \bar{l}_\alpha + \bar{H})\right)} \tag{17}$$

$$= \sum_{i \neq j} \frac{\mathrm{Im}\left\{(Y_\nu^*)_{\alpha i}(Y_\nu)_{\alpha j}\left[(Y_\nu^\dagger Y_\nu)_{ij} + \xi_{ij}(Y_\nu^\dagger Y_\nu)_{ji}\right]\right\}}{(Y_\nu^\dagger Y_\nu)_{ii}(Y_\nu^\dagger Y_\nu)_{jj}} \times \frac{\xi_{ij}\zeta_j(\xi_{ij}^2 - 1)}{(\xi_{ij}\zeta_j)^2 + (\xi_{ij}^2 - 1)^2} \tag{18}$$

where $\xi_{ij} = M_i/M_j$ and we took $M_1 = 10\ TeV$ and $d = (M_3 - M_1)/M_1 = 10^{-8}$

In our model, we have three right-handed neutrinos with exactly degenerate masses, $M_1 = M_2 = M_3 = M$. However, a tinymass separation between any two right-handed neutrinos is necessary for successful leptogenesis, and this is included to our model by including a higher dimension term in Eq.(1). Such term leads to a minor correction in the Majorana mass matrix of Eq.(8), and the resultant structure of the mass matrix may be written as

$$M_{mid} = \begin{pmatrix} M & e & e \\ e & M & e \\ e & e & M \end{pmatrix} \tag{19}$$

where $e = \frac{y_{s_2} v_\phi}{\Lambda^2}$ is a parameter that quantifies the tiny difference between masses required for leptogenesis. The mass matrix in Eq.(19) is diagonalized using a (3×3) matrix of the form

$$D = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

with real eigenvalues $M_1 = M - e$ and $M_2 = M - e$ and $M_3 = M - 2e$ . In the basis where the charged-lepton and Majorana mass matrix are diagonal, the dirac mass matrix Eq.(9) takes the form

$$M'_{\nu N} = M_{\nu N}.D = \begin{pmatrix} \frac{v(a+s-x)}{\Lambda} & -\frac{v(a-s+x)}{\Lambda} & \frac{v(2s+x)}{\Lambda} \\ \frac{v(a-s+x)}{\Lambda} & \frac{v(a-v+x)}{\Lambda} & \frac{v(2s+x)}{\Lambda} \\ -\frac{v(2a)}{\Lambda} & -\frac{v(a+s-x)}{\Lambda} & \frac{v(2s+x)}{\Lambda} \end{pmatrix}$$

From this point onward, we will take $Y_{\nu N} = M'_{\nu N}/v$, which is relevant for calculating CP asymmetry that arises during the decay of right-handed neutrinos in out-of-equilibrium way.

The CP-violating asymmetries $\varepsilon_{i\alpha}$ in the flavored resonant leptogenesis scenario under study are related to the baryon-to-photon ratio $\eta_B$ as follows [40]:

$$\eta_B \simeq -9.6 \times 10^{-3} \sum_\alpha (\varepsilon_{1\alpha} K_{1\alpha} + \varepsilon_{2\alpha} K_{2\alpha}) \tag{20}$$

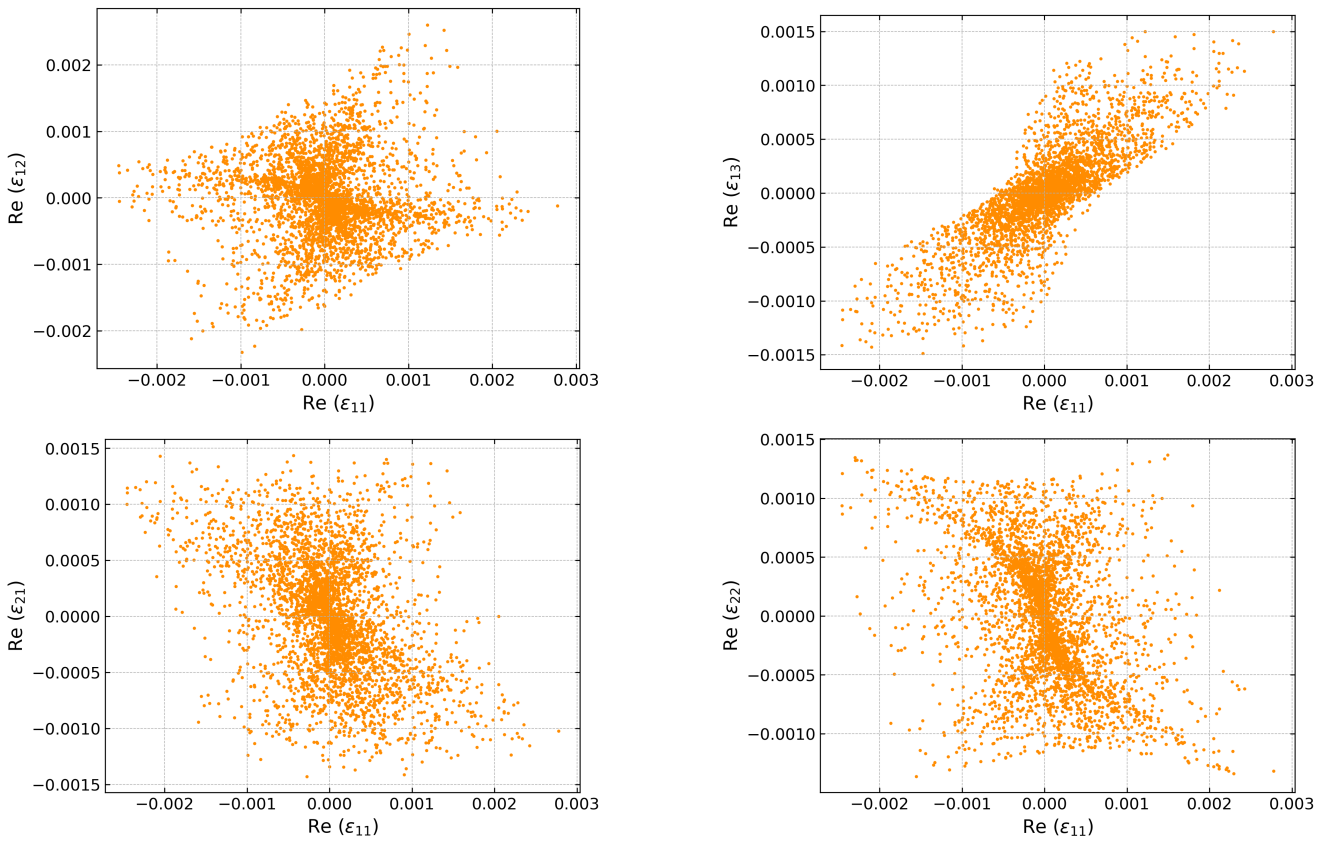



Figure 4: Correlations between the flavor-dependent CP-violating asymmetries $\varepsilon_{11}$ with $\varepsilon_{12}$, $\varepsilon_{21}$ and $\varepsilon_{22}$ respectively.

where $K_{1\alpha}$ and $K_{2\alpha}$ are the conversion efficiency factors. The region in which the lepton flavor takes effect determines the sum over the flavor index $\alpha$. . To evaluate the sizes of $K_{i\alpha}$, let us first of all figure out the effective light neutrino masses.

$$m_{i\alpha} \simeq \frac{v^2 \left|(Y_v)_{\alpha i}\right|^2}{M_i} \tag{21}$$

The decay parameters $K_{i\alpha} \equiv \tilde{m}_{i\alpha}/m_*$ can be calculated, where $m_* = 8\pi v^2 H(M_1)/M_1^2 \simeq 1.08 \times 10^{-3} eV$ gives the equilibrium neutrino mass and $H(M_1)$ is called the Hubble expansion parameter of the Universe.

Now we define a dimensionless parameter $d \equiv (M_2 - M_1)/M_1 = \xi_{21} - 1$ to calculate the level of degeneracy for two of the three heavy Majorana neutrinos. Allowing for $d << 1$, we have $\kappa_{1\alpha} \simeq \kappa_{2\alpha} \equiv \kappa(K_\alpha)$ with $K_\alpha \equiv K_{1\alpha} + K_{2\alpha}$. Given the initial thermal abundance of heavy Majorana neutrinos, the efficiency factor $\kappa(K_\alpha)$ can be expressed as:

$$\kappa(K_\alpha) \simeq \frac{2}{K_\alpha z_B(K_\alpha)} \left[1 - exp\left(\frac{-1}{2} K_\alpha z_B(K_\alpha)\right)\right] \tag{22}$$

where $z_B(K_\alpha) \simeq 2 + 4K_\alpha^{0.13} exp(-2.5/K_\alpha)$.

We illustrated that our resonant lepto-genesis scenario works well. We have the correlation of Baryon Asymmetry of the Universe with the flavor dependent CP-violating asymmetries in Fig. 5.

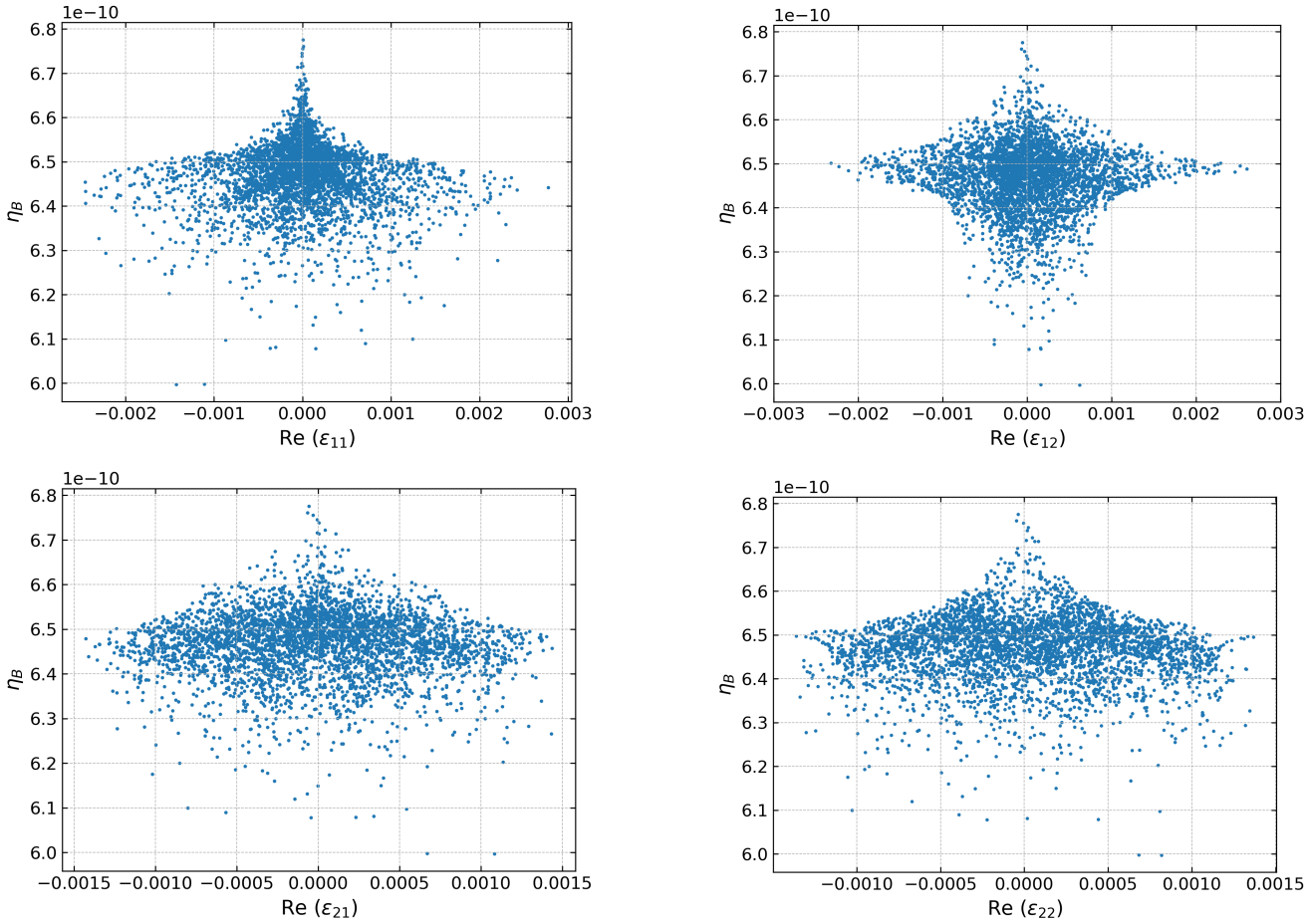



Figure 5: Correlations between the Baryon Asymmetry ($\eta_B$) with flavor-dependent CP-violating asymmetries $\varepsilon_{11}$ , $\varepsilon_{12}$, $\varepsilon_{21}$ and $\varepsilon_{22}$ respectively.

## 5. Conclusion

We showed the $\Delta(54)$ flavor symmetry with SM Higgs boson and $Z_2 \otimes Z_3 \otimes Z_4$ symmetry which generates a neutrino mass matrix. Our model incorporates the ISS mechanism in order to provide a flavor-symmetric approach. This includes accounting for the solar mixing angle , upper octant of atmospheric mixing angle , non-zero reactor angle and CP violation ($\delta_{CP}$). The values predicted for the absolute neutrino masses are within the range of the cosmological bound $\sum_i m_i < 0.12$ eV. Neutrino oscillation parameter predictions based on the resulting mass matrix agree with the best-fit values obtained via $\chi^2$ analysis. However, the Inverted Hierarchy (IH) scenario's predicted mixing angles, mass-squared disparities, and CP violation phase disagree with experimental findings.

Furthermore, we investigated baryogenesis via flavoured resonant leptogenesis. The right-handed neutrinos are degenerate at the dimension 5 level, introducing a higher dimension term resulted in a tiny splitting. We have taken the splitting parameter, $d \approx 10^{-8}$ and thus, obtained a nonzero, resonantly enhanced CP asymmetry ($\epsilon_{i\alpha}$) from the out-of equilibrium decay of right-handed Majorana neutrinos. We determined the baryon-to-photon ratio ($\eta_B$) in the flavored resonant leptogenesis scenario using the values of the CP asymmetries. It was found that the model can explain the observed value of BAU with particular choice of RHN mass scale, $M_1 = 10TeV$ and mass

splitting, $d \approx 10^{-8}$.

Future work is reserved for examining the model to explore phenomena such as Asymmetric Dark Matter.

## Acknowledgements

HB acknowledges Tezpur University, India for Institutional Research Fellowship. The research of NKF is funded by DST-SERB, India under Grant no. EMR/2015/001683.

## References

[1] Aker, M. *et al.* Improved upper limit on the neutrino mass from a direct kinematic method by katrin. *Physical review letters* **123**, 221802 (2019).

[2] Faessler, A. Status of the determination of the electron–neutrino mass. *Progress in Particle and Nuclear Physics* **113**, 103789 (2020).

[3] Araki, T. *et al.* Measurement of neutrino oscillation with kamland: Evidence of spectral distortion. *Physical review letters* **94**, 081801 (2005).

[4] Cao, S. *et al.* Physics potential of the combined sensitivity of T2K-II, NO $\nu$ A extension, and JUNO. *Physical Review D* **103**, 112010 (2021).

[5] Nath, A. & Francis, N. K. Detection techniques and investigation of different neutrino experiments. *International Journal of Modern Physics A* **36**, 2130008 (2021).

[6] Minkowski, P. $\mu \to e\gamma$ at a rate of one out of 109 muon decays? *Physics Letters B* **67**, 421–428 (1977).

[7] King, S. & Yanagida, T. Testing the see-saw mechanism at collider energies. *Progress of theoretical physics* **114**, 1035–1043 (2005).

[8] Mohapatra, R. N. Mechanism for understanding small neutrino mass in superstring theories. *Physical review letters* **56**, 561 (1986).

[9] Ma, E. Radiative inverse seesaw mechanism for nonzero neutrino mass. *Physical Review D* **80**, 013013 (2009).

[10] Mohapatra, R. N., Nandi, S. & Perez-Lorenzana, A. Neutrino masses and oscillations in models with large extra dimensions. *Physics Letters B* **466**, 115–121 (1999).

[11] Arkani-Hamed, N., Dimopoulos, S., Dvali, G. & March-Russell, J. Neutrino masses from large extra dimensions. *Physical Review D* **65**, 024032 (2001).

[12] Fukugita, M. & Yanagida, T. Barygenesis without grand unification. *Physics Letters B* **174**, 45–47 (1986).

[13] Nguyen, T. P., Le, T. T., Hong, T. T. & Hue, L. T. Decay of standard model-like Higgs boson $h \to \mu\tau$ in a 3-3-1 model with inverse seesaw neutrino masses. *Phys. Rev. D* **97**, 073003 (2018). 1802.00429.

[14] King, S. F., Merle, A., Morisi, S., Shimizu, Y. & Tanimoto, M. Neutrino Mass and Mixing: from Theory to Experiment. *New J. Phys.* **16**, 045018 (2014). 1402.4271.

[15] King, S. F. Neutrino mass models. *Rept. Prog. Phys.* **67**, 107–158 (2004). hep-ph/0310204.

[16] Cao, S. *et al.* Physics potential of the combined sensitivity of T2K-II, NO$\nu$A extension, and JUNO. *Phys. Rev. D* **103**, 112010 (2021). 2009.08585.

[17] Ahn, Y. H. & Gondolo, P. Towards a realistic model of quarks and leptons, leptonic $CP$ violation, and neutrinoless $\beta\beta$-decay. *Phys. Rev. D* **91**, 013007 (2015). 1402.0150.

[18] McDonald, A. B. Nobel lecture: the sudbury neutrino observatory: observation of flavor change for solar neutrinos. *Reviews of Modern Physics* **88**, 030502 (2016).

[19] Nguyen, T. P., Thuc, T. T., Si, D. T., Hong, T. T. & Hue, L. T. Low-energy phenomena of the lepton sector in an A4 symmetry model with heavy inverse seesaw neutrinos. *PTEP* **2022**, 023B01 (2022). 2011.12181.

[20] Hong, T. T., Nha, N. H. T., Nguyen, T. P., Phuong, L. T. T. & Hue, L. T. Decays h $\to$ eaeb, eb $\to$ ea$\gamma$, and (g $-$ 2)e,$\mu$ in a 3-3-1 model with inverse seesaw neutrinos. *PTEP* **2022**, 093B05 (2022). 2206.08028.

[21] Kajita, T. Nobel lecture: Discovery of atmospheric neutrino oscillations. *Reviews of Modern Physics* **88**, 030501 (2016).

[22] de Salas, P. F. *et al.* 2020 global reassessment of the neutrino oscillation picture. *Journal of High Energy Physics* **2021**, 1–36 (2021).

[23] Okada, H. & Tanimoto, M. Spontaneous CP violation by modulus $\tau$ in A4 model of lepton flavors. *Journal of High Energy Physics* **2021**, 1–27 (2021).

[24] Ahn, Y. H., Kang, S. K. & Lee, H. M. Toward a model of quarks and leptons. *Phys. Rev. D* **106**, 075029 (2022). 2112.13392.

[25] Phong Nguyen, T., Hue, L. T., Si, D. T. & Thuc, T. T. CP violations in a predictive $A_4$ symmetry model. *PTEP* **2020**, 033B04 (2020). 1711.05588.

[26] Buravov, L. I. Confining potential and mass of elementary particles. *J. Mod. Phys.* **7**, 129–133 (2017). 1502.00958.

[27] Buravov, L. Elementary muon, pion, and kaon particles as resonators for neutrino quanta. Calculations of mass ratios for e, $\mu$, $\pi$ 0, $\pi\pm$, K 0, K$\pm$, and $\nu$ e. *Russian Physics Journal* **52**, 25–32 (2009).

[28] Barman, A., Francis, N. K. & Bora, H. Neutrino Mixing Phenomenology: $A_4$ Discrete Flavor Symmetry with Type-I Seesaw Mechanism. *arXiv preprint arXiv:2306.11461* (2023).

[29] Bora, H., Francis, N. K., Barman, A. & Thapa, B. Majorana neutrinos in Inverse Seesaw and $\Delta(54)$ Flavor Models. *arXiv e-prints* arXiv–2311 (2023).

[30] Kuzmin, V. A., Rubakov, V. A. & Shaposhnikov, M. E. On anomalous electroweak baryon-number non-conservation in the early universe. *Physics Letters B* **155**, 36–42 (1985).

[31] Davidson, S. & Ibarra, A. A lower bound on the right-handed neutrino mass from leptogenesis. *Physics Letters B* **535**, 25–32 (2002).

[32] Fong, C. S., Rahat, M. H. & Saad, S. Low-scale resonant leptogenesis in SU(5) GUT with T13 family symmetry. *Phys. Rev. D* **104**, 095028 (2021). 2103.14691.

[33] Bora, H., Francis, N. K., Barman, A. & Thapa, B. Neutrino mass model in the context of $\Delta(54) \otimes Z_2 \otimes Z_3 \otimes Z_4$ flavor symmetries with Inverse Seesaw mechanism. *Physics Letters B* **848**, 138329 (2024).

[34] Thapa, B. & Francis, N. K. Resonant leptogenesis and $TM_1$ mixing in minimal type-I seesaw model with $S_4$ symmetry. *The European Physical Journal C* **81**, 1–8 (2021).

[35] Ishimori, H. *et al.* Non-abelian discrete symmetries in particle physics. *Progress of Theoretical Physics Supplement* **183**, 1–163 (2010).

[36] Ishimori, H., Kobayashi, T., Okada, H., Shimizu, Y. & Tanimoto, M. Lepton flavor model from $\Delta$ (54) symmetry. *Journal of High Energy Physics* **2009**, 011 (2009).

[37] Lei, M. & Wells, J. D. Minimally modified $A_4$ Altarelli-Feruglio model for neutrino masses and mixings and its experimental consequences. *Physical Review D* **102**, 016023 (2020).

[38] Pilaftsis, A. & Underwood, T. E. Resonant leptogenesis. *Nuclear Physics B* **692**, 303–345 (2004).

[39] Pilaftsis, A. Cp violation and baryogenesis due to heavy majorana neutrinos. *Physical Review D* **56**, 5431 (1997).

[40] Xing, Z.-z. & Zhang, D. Bridging resonant leptogenesis and low-energy cp violation with an rge-modified seesaw relation. *Physics Letters B* **804**, 135397 (2020).

# Revealing the nature of $\Omega_c$-like states from pentaquark perspective

Ulaş Özdem[1, *]

[1] *Health Services Vocational School of Higher Education, Istanbul Aydin University, Sefakoy-Kucukcekmece, 34295 Istanbul, Türkiye*

(Dated: March 1, 2024)

We systematically study the electromagnetic properties of controversial states whose internal structure is not elucidated and we try to offer a different point of view to unravel the internal structure of these states. Inspired by the $\Omega_c$ states observed by the LHCb Collaboration, we study the electromagnetic properties of the $\Omega_c$-like states as the compact diquark-diquark-antiquark pentaquarks with both $J^P = \frac{1}{2}^-$ and $J^P = \frac{3}{2}^-$ in the context of the QCD light-cone sum rule model. From the obtained numerical results, we conclude that the magnetic dipole moments of the $\Omega_c$-like states can reflect their inner structures, which can be used to distinguish their spin-parity quantum numbers. Measuring the magnetic moment of the $\Omega_c$-like states in future experimental facilities can be very helpful for understanding the internal organization and identifying the quantum numbers of these states.

Keywords: Electromagnetic form factors, diquark-diquark-antiquark picture, QCD light-cone sum rules, magnetic dipole moments

arXiv:2402.18901v1 [hep-ph] 29 Feb 2024

## I. MOTIVATION

Many heavy baryon states have been discovered in recent years by experimental collaborations. One of the main challenges in non-perturbative QCD is to understand and shed light on the precise nature of these states. Through further theoretical explorations involving the study of hadrons comprising a single heavy quark, it offers an exquisite basis to probe the dynamics of a light diquark in a heavy quark background, to enhance the understanding of the non-perturbative nature of QCD, and to test the predictions of different phenomenological models. In recent decades, there have been significant experimental advancements in the field of singly-charm/bottom baryons, resulting in a dramatic increase in the number of particles [1–30]. As the data for the presence of some of these states is scarce and their internal structure, as well as quantum numbers, are not well defined, additional experimental exploration is therefore required. Therefore, researchers in hadron physics continue to study these topics through theoretical and experimental research, as they have not yet been fully understood.

In 2017, the LHCb collaboration studied the $\Xi_c^+ K^-$ mass spectrum and observed five new narrow excited $\Omega_c$ states, $\Omega_c(3000)$, $\Omega_c(3050)$, $\Omega_c(3066)$, $\Omega_c(3090)$, $\Omega_c(3119)$ [22]. The parameters that have been measured are as follows

$$\begin{aligned}
M_{\Omega_c(3000)} &= 3000.4 \pm 0.2 \pm 0.1 \text{ MeV}, \quad \Gamma_{\Omega_c(3000)} = 4.5 \pm 0.6 \pm 0.3 \text{ MeV}, \\
M_{\Omega_c(3050)} &= 3050.2 \pm 0.1 \pm 0.1 \text{ MeV}, \quad \Gamma_{\Omega_c(3050)} = 0.8 \pm 0.2 \pm 0.1 \text{ MeV}, \\
M_{\Omega_c(3066)} &= 3065.6 \pm 0.1 \pm 0.3 \text{ MeV}, \quad \Gamma_{\Omega_c(3066)} = 3.5 \pm 0.4 \pm 0.2 \text{ MeV}, \\
M_{\Omega_c(3090)} &= 3090.2 \pm 0.3 \pm 0.5 \text{ MeV}, \quad \Gamma_{\Omega_c(3090)} = 8.7 \pm 1.0 \pm 0.8 \text{ MeV}, \\
M_{\Omega_c(3119)} &= 3119.1 \pm 0.3 \pm 0.9 \text{ MeV}, \quad \Gamma_{\Omega_c(3119)} = 1.1 \pm 0.8 \pm 0.4 \text{ MeV}.
\end{aligned} \tag{1}$$

Later, the former four states $\Omega_c(3000)$, $\Omega_c(3050)$, $\Omega_c(3066)$ and $\Omega_c(3090)$ were confirmed by the LHCb and Belle Collaborations [31, 32]. The discovery of the LHCb collaboration has led to a new experimental situation, which requires a more detailed study of heavy baryons and their properties. The discovery of these states, although their quantum numbers have not been determined, could provide new insights into QCD and its complex behavior. This could lead to a deeper understanding of the underlying properties of QCD. However, our information about their properties is still not sufficient and further suggestions for experimental exploration of $\Omega_c$-like states should be addressed. The experimental discoveries were followed by various theoretical studies, investigating them in the conventional baryon, molecular, and compact pentaquark states to shed light on their exact nature and quantum numbers (for details see the Refs. [33–37]).

The literature review indicates that the majority of research has concentrated on computing the spectroscopic and decay parameters of these states. However, it is evident that relying exclusively on spectroscopic and decay

* ulasozdem@aydin.edu.tr

parameters is insufficient for resolving the disputed nature of these states. Therefore, to determine the precise internal configurations of these states, additional studies such as obtaining electromagnetic multipole moments, radiative decays, and weak decays are required. To unveil the nature and internal structures of unconventional states, the physical quantities associated with the electromagnetic properties of hadrons are valuable parameters. The charge and magnetism distributions inside the hadrons can be studied using the electromagnetic form factors (FFs) and the resulting multipole moments. This information can be used to determine the distribution of quark-antiquark pairs (both valence and sea-quarks) and gluons within the hadron volume. It also provides valuable insights into the geometric forms, electric radii, and magnetic radii of the hadrons. With this motivation, in this work, we will derive the electromagnetic features of $\Omega_c$-like states using the method of QCD light-cone sum rules [38–40]. Throughout the analysis, these states are considered in the compact $[ss][qc][\bar{q}]$, $[sq][sc][\bar{q}]$ and $[qs][sc][\bar{q}]$ diquark-diquark-antiquark configurations, and are assumed to have both $J^P = \frac{1}{2}^-$ and $J^P = \frac{3}{2}^-$ quantum numbers. There are several studies in the literature that have extracted the electromagnetic features of the singly-heavy nonconventional hadron states [41–48].

The article is structured as follows: we derive the QCD light-cone sum rules for the electromagnetic features of the $\Omega_c$-like states in Sect. II. The numerical results of the magnetic dipole moments and discussions are presented in Sect. III. Sect. IV is reserved for our summary and concluding remarks.

## II. ELECTROMAGNETIC FORM FACTORS FROM QCD LIGHT-CONE SUM RULES

To study the electromagnetic properties of hadrons in the low energy regime, reliable and efficient non-perturbative methods are needed. The QCD light-cone sum rules method is one of these methods and has been very successful in predicting the static and dynamical properties of both conventional and nonconventional hadrons. In this method, the correlation function, the key ingredient of the model, is calculated in terms of both hadronic parameters and QCD degrees of freedom. The Borel transform and the continuum subtraction procedures are then applied to suppress the contribution of the effects caused by the continuum and higher states. Together with these procedures, the QCD light-cone sum rules for the physical parameter of interest are obtained via dispersion integrals and by using the quark-hadron duality assumption.

The study of the electromagnetic form factors of the $\Omega_c$-like states ($\frac{1}{2}^- \to \Omega_c$, $\frac{3}{2}^- \to \Omega_c^*$) in QCD light cone sum rules begins with the writing of the correlation function in the weak electromagnetic background field ($\gamma$) as follows

$$\Pi(p,q) = i \int d^4x e^{ip\cdot x} \langle 0|T\left\{J^{\Omega_c}(x)\bar{J}^{\Omega_c}(0)\right\}|0\rangle_\gamma \,, \tag{2}$$

$$\Pi_{\mu\nu}(p,q) = i \int d^4x e^{ip\cdot x} \langle 0|T\left\{J_\mu^{\Omega_c^*}(x)\bar{J}_\nu^{\Omega_c^*}(0)\right\}|0\rangle_\gamma \,, \tag{3}$$

where $T$ is the time ordered product; and the $J^{\Omega_c}(x)$ and $J_\mu^{\Omega_c^*}(x)$ are the interpolating currents for the $\Omega_c$-like states. The explicit form of the $J^{\Omega_c}(x)$ and $J_\mu^{\Omega_c^*}(x)$ are written as follows

$$J^{\Omega_c}(x) = \varepsilon^{abc}\varepsilon^{ade}\varepsilon^{bfg}\big[q_1^{d^T}(x)C\gamma_\mu q_2^e(x)\, q_3^{f^T}(x)C\gamma^\mu c^g(x)\, C\bar{q}^{c^T}(x)\big], \tag{4}$$

$$J_\mu^{\Omega_c^*}(x) = \varepsilon^{abc}\varepsilon^{ade}\varepsilon^{bfg}\big[q_1^{d^T}(x)C\gamma_\mu q_2^e(x)\, q_3^{f^T}(x)C\gamma_5 c^g(x)\, C\bar{q}^{c^T}(x)\big], \tag{5}$$

where $q_1$, $q_2$, $q_3$ are u-, d- and/or s-quark, $q$ is u- or d-quark, the $C$ is the charge conjugation operator and $a$, $b$... are color indices. It should be noted here that the mass values obtained with these interpolating currents are compatible with the mass values of the $\Omega_c(3050)$, $\Omega_c(3066)$, $\Omega_c(3090)$, and $\Omega_c(3119)$ states [49, 50].

After this brief introduction to the methodology, we begin the analysis of electromagnetic form factors in terms of hadronic parameters, following the above procedure. To evaluate the hadronic side, a complete set of hadronic states with the same quantum numbers is inserted into the correlation functions with the considered interpolating currents. Integration over x is then performed, yielding

$$\Pi^{Had}(p,q) = \frac{\langle 0 \mid J^{\Omega_c}(x) \mid \Omega_c(p,s)\rangle}{[p^2 - m_{\Omega_c}^2]}\langle \Omega_c(p,s) \mid \Omega_c(p+q,s)\rangle_\gamma \frac{\langle \Omega_c(p+q,s) \mid \bar{J}^{\Omega_c}(0) \mid 0\rangle}{[(p+q)^2 - m_{\Omega_c}^2]} + \cdots \,, \tag{6}$$

$$\Pi_{\mu\nu}^{Had}(p,q) = \frac{\langle 0 \mid J_\mu^{\Omega_c^*}(x) \mid \Omega_c^*(p,s)\rangle}{[p^2 - m_{\Omega_c^*}^2]}\langle \Omega_c^*(p,s) \mid \Omega_c^*(p+q,s)\rangle_\gamma \frac{\langle \Omega_c^*(p+q,s) \mid \bar{J}_\nu^{\Omega_c^*}(0) \mid 0\rangle}{[(p+q)^2 - m_{\Omega_c^*}^2]} + \cdots \,, \tag{7}$$

where $\cdots$ stands for the continuum and higher states. The Eqs. (6) and (7) contain matrix elements containing hadronic parameters like residues ($\lambda$), form factors ($F_i$), spinors ($u_{(\mu)}$) and so on. The explicit expressions for these

matrix elements are as follows

$$\langle 0 \mid J^{\Omega_c}(x) \mid \Omega_c(p,s)\rangle = \lambda_{\Omega_c}\gamma_5\, u(p,s), \tag{8}$$

$$\langle \Omega_c(p+q,s) \mid \bar{J}^{\Omega_c}(0) \mid 0\rangle = \lambda_{\Omega_c}\gamma_5\, \bar{u}(p+q,s), \tag{9}$$

$$\langle \Omega_c(p,s) \mid \Omega_c(p+q,s)\rangle_\gamma = \varepsilon^\mu\, \bar{u}(p,s)\Big[\big[f_1(q^2)+f_2(q^2)\big]\gamma_\mu + f_2(q^2)\frac{(2p+q)_\mu}{2m_{\Omega_c}}\Big]\, u(p+q,s), \tag{10}$$

$$\langle 0 \mid J^{\Omega_c^*}_\mu(x) \mid \Omega_c^*(p,s)\rangle = \lambda_{\Omega_c^*} u_\mu(p,s), \tag{11}$$

$$\langle \Omega_c^*(p+q,s) \mid \bar{J}^{\Omega_c^*}_\nu(0) \mid 0\rangle = \lambda_{\Omega_c^*} \bar{u}_\nu(p+q,s), \tag{12}$$

$$\begin{aligned}\langle \Omega_c^*(p,s) \mid \Omega_c^*(p+q,s)\rangle_\gamma = -e\bar{u}_\mu(p,s)\Big[&F_1(q^2)g_{\mu\nu}\not{\epsilon} - \frac{1}{2m_{\Omega_c^*}}\Big[F_2(q^2)g_{\mu\nu} + F_4(q^2)\frac{q_\mu q_\nu}{(2m_{\Omega_c^*})^2}\Big]\not{\epsilon}\not{q} \\ &+ F_3(q^2)\frac{1}{(2m_{\Omega_c^*})^2}q_\mu q_\nu \not{\epsilon}\Big] u_\nu(p+q,s).\end{aligned} \tag{13}$$

With the help of the above formulas and the application of some mathematical simplifications, the following result is derived for the hadronic side of the correlation functions,

$$\Pi^{Had}(p,q) = \lambda^2_{\Omega_c}\gamma_5\frac{\big(\not{p}+m_{\Omega_c}\big)}{[p^2-m^2_{\Omega_c}]}\varepsilon^\mu\Big[\big[f_1(q^2)+f_2(q^2)\big]\gamma_\mu + f_2(q^2)\,\frac{(2p+q)_\mu}{2m_{\Omega_c}}\Big]\gamma_5\frac{\big(\not{p}+\not{q}+m_{\Omega_c}\big)}{[(p+q)^2-m^2_{\Omega_c}]}, \tag{14}$$

$$\begin{aligned}\Pi^{Had}_{\mu\nu}(p,q) = &-\frac{\lambda^2_{\Omega_c^*}\big(\not{p}+\not{q}+m_{\Omega_c^*}\big)}{[(p+q)^2-m^2_{\Omega_c^*}][p^2-m^2_{\Omega_c^*}]}\Big[g_{\mu\nu}-\frac{1}{3}\gamma_\mu\gamma_\nu-\frac{2\,p_{1\mu}p_{1\nu}}{3\,m^2_{\Omega_c^*}}+\frac{(p+q)_\mu\gamma_\nu-(p+q)_\nu\gamma_\mu}{3\,m_{\Omega_c^*}}\Big] \\ &\times\Big\{F_1(q^2)g_{\mu\nu}\not{\epsilon}-\frac{1}{2m_{\Omega_c^*}}\Big[F_2(q^2)g_{\mu\nu}+F_4(q^2)\frac{q_\mu q_\nu}{(2m_{\Omega_c^*})^2}\Big]\not{\epsilon}\not{q}+\frac{F_3(q^2)}{(2m_{\Omega_c^*})^2}q_\mu q_\nu\not{\epsilon}\Big\} \\ &\times\big(\not{p}+m_{\Omega_c^*}\big)\Big[g_{\mu\nu}-\frac{1}{3}\gamma_\mu\gamma_\nu-\frac{2\,p_\mu p_\nu}{3\,m^2_{\Omega_c^*}}+\frac{p_\mu\gamma_\nu-p_\nu\gamma_\mu}{3\,m_{\Omega_c^*}}\Big].\end{aligned} \tag{15}$$

The magnetic dipole moments ($\mu$) are determined from magnetic form factors at zero momentum transfer,

$$\mu_{\Omega_c} = \frac{e}{2m_{\Omega_c}}F_M(0), \tag{16}$$

$$\mu_{\Omega_c^*} = \frac{e}{2m_{\Omega_c^*}}G_M(0), \tag{17}$$

where $F_M(0)$ and $G_M(0)$ stand for magnetic form factor of the spin-$\frac{1}{2}^-$ and spin-$\frac{3}{2}^-$ $\Omega_c$-like states, respectively. Explicit forms of these terms are given as,

$$F_M(0) = f_1(0)+f_2(0), \tag{18}$$

$$G_M(0) = F_1(0)+F_2(0). \tag{19}$$

Now that we have finished calculating the analysis regarding hadronic parameters, it is time to calculate in terms of QCD parameters. The magnetic dipole moments are analyzed in terms of QCD parameters using the interpolating current in the correlation function. Applying Wick's theorem, we contract all the light and heavy quark fields to get the desired outcomes:

$$\begin{aligned}\Pi^{\rm QCD}(p,q) = &-i\,\varepsilon^{abc}\varepsilon^{a'b'c'}\varepsilon^{ade}\varepsilon^{a'd'e'}\varepsilon^{bfg}\varepsilon^{b'f'g'}\int d^4x e^{ip\cdot x}\langle 0|\Big\{\mathrm{Tr}\Big[\gamma^\mu \mathrm{S}_{\rm c}^{\rm gg'}(\mathrm{x})\gamma^\nu\widetilde{\mathrm{S}}_{\rm q_3}^{\rm ff'}(\mathrm{x})\Big]\mathrm{Tr}\Big[\gamma_\mu \mathrm{S}_{\rm q_2}^{\rm ee'}(\mathrm{x})\gamma_\nu\widetilde{\mathrm{S}}_{\rm q_1}^{\rm dd'}(\mathrm{x})\Big] \\ &-\mathrm{Tr}\Big[\gamma^\mu \mathrm{S}_{\rm c}^{\rm gg'}(\mathrm{x})\gamma^\nu\widetilde{\mathrm{S}}_{\rm q_3}^{\rm ff'}(\mathrm{x})\Big]\mathrm{Tr}\Big[\gamma_\mu \mathrm{S}_{\rm q_2q_1}^{\rm ed'}(\mathrm{x})\gamma_\nu\widetilde{\mathrm{S}}_{\rm q_1q_2}^{\rm de'}(\mathrm{x})\Big]-\mathrm{Tr}\Big[\gamma^\mu \mathrm{S}_{\rm c}^{\rm gg'}(\mathrm{x})\gamma^\nu\widetilde{\mathrm{S}}_{\rm q_2q_3}^{\rm ef'}(\mathrm{x})\gamma_\mu \mathrm{S}_{\rm q_1}^{\rm dd'}(\mathrm{x})\gamma_\nu\widetilde{\mathrm{S}}_{\rm q_3q_2}^{\rm fe'}(\mathrm{x})\Big] \\ &-\mathrm{Tr}\Big[\gamma^\mu \mathrm{S}_{\rm c}^{\rm gg'}(\mathrm{x})\gamma^\nu\widetilde{\mathrm{S}}_{\rm q_1q_3}^{\rm df'}(\mathrm{x})\gamma_\mu\widetilde{\mathrm{S}}_{\rm q_2}^{\rm ee'}(\mathrm{x})\gamma_\nu\widetilde{\mathrm{S}}_{\rm q_3q_1}^{\rm fd'}(\mathrm{x})\Big]+\mathrm{Tr}\Big[\gamma^\mu \mathrm{S}_{\rm c}^{\rm gg'}(\mathrm{x})\gamma^\nu\widetilde{\mathrm{S}}_{\rm q_1q_3}^{\rm df'}(\mathrm{x})\gamma_\mu \mathrm{S}_{\rm q_2q_1}^{\rm ed'}(\mathrm{x})\gamma_\nu\widetilde{\mathrm{S}}_{\rm q_3q_2}^{\rm fe'}(\mathrm{x})\Big] \\ &+\mathrm{Tr}\Big[\gamma^\mu \mathrm{S}_{\rm c}^{\rm gg'}(\mathrm{x})\gamma^\nu\widetilde{\mathrm{S}}_{\rm q_2q_3}^{\rm ef'}(\mathrm{x})\gamma_\mu \mathrm{S}_{\rm q_1q_2}^{\rm de'}(\mathrm{x})\gamma_\mu\widetilde{\mathrm{S}}_{\rm q_3q_1}^{\rm fd'}(\mathrm{x})\Big]\Big\}\widetilde{\mathrm{S}}_{\rm q}^{\rm c'c}(-\mathrm{x})|0\rangle_\gamma,\end{aligned} \tag{20}$$

$$\Pi^{\rm QCD}_{\mu\nu}(p,q) = -\,i\,\varepsilon^{abc}\varepsilon^{a'b'c'}\varepsilon^{ade}\varepsilon^{a'd'e'}\varepsilon^{bfg}\varepsilon^{b'f'g'}\int d^4x e^{ip\cdot x}\langle 0|\Big\{{\rm Tr}\Big[\gamma_5 {\rm S}_{\rm c}^{\rm gg'}({\rm x})\gamma_5\widetilde{\rm S}_{\rm q_3}^{\rm ff'}({\rm x})\Big]{\rm Tr}\Big[\gamma_\mu {\rm S}_{\rm q_2}^{\rm ee'}({\rm x})\gamma_\nu\widetilde{\rm S}_{\rm q_1}^{\rm dd'}({\rm x})\Big]$$
$$-\,{\rm Tr}\Big[\gamma_5 {\rm S}_{\rm c}^{\rm gg'}({\rm x})\gamma_5\widetilde{\rm S}_{\rm q_3}^{\rm ff'}({\rm x})\Big]{\rm Tr}\Big[\gamma_\mu {\rm S}_{\rm q_2q_1}^{\rm ed'}({\rm x})\gamma_\nu\widetilde{\rm S}_{\rm q_1q_2}^{\rm de'}({\rm x})\Big] - {\rm Tr}\Big[\gamma_5 {\rm S}_{\rm c}^{\rm gg'}({\rm x})\gamma_5\widetilde{\rm S}_{\rm q_2q_3}^{\rm ef'}({\rm x})\gamma_\mu {\rm S}_{\rm q_1}^{\rm dd'}({\rm x})\gamma_\nu\widetilde{\rm S}_{\rm q_3q_2}^{\rm fe'}({\rm x})\Big]$$
$$-\,{\rm Tr}\Big[\gamma_5 {\rm S}_{\rm c}^{\rm gg'}({\rm x})\gamma_5\widetilde{\rm S}_{\rm q_1q_3}^{\rm df'}({\rm x})\gamma_\mu\widetilde{\rm S}_{\rm q_2}^{\rm ee'}({\rm x})\gamma_\nu\widetilde{\rm S}_{\rm q_3q_1}^{\rm fd'}({\rm x})\Big] + {\rm Tr}\Big[\gamma_5 {\rm S}_{\rm c}^{\rm gg'}({\rm x})\gamma_5\widetilde{\rm S}_{\rm q_1q_3}^{\rm df'}({\rm x})\gamma_\mu {\rm S}_{\rm q_2q_1}^{\rm ed'}({\rm x})\gamma_\nu\widetilde{\rm S}_{\rm q_3q_2}^{\rm fe'}({\rm x})\Big]$$
$$+\,{\rm Tr}\Big[\gamma_5 {\rm S}_{\rm c}^{\rm gg'}({\rm x})\gamma_5\widetilde{\rm S}_{\rm q_2q_3}^{\rm ef'}({\rm x})\gamma_\mu {\rm S}_{\rm q_1q_2}^{\rm de'}({\rm x})\gamma_\mu\widetilde{\rm S}_{\rm q_3q_1}^{\rm fd'}({\rm x})\Big]\Big\}\widetilde{\rm S}_{\rm q}^{\rm c'c}(-{\rm x})|0\rangle_\gamma, \tag{21}$$

where $\widetilde{S}^{ij}_{c(q)}(x) = CS^{ij\rm T}_{c(q)}(x)C$. $S_{q_iq_j}$ exists when $q_i = q_j$ but it vanishes when $q_i \neq q_j$. The $S_c(x)$ and $S_q(x)$ are the heavy and light-quark propagators whose expressions in the coordinate space are given as [51, 52],

$$S_q(x) = S_q^{free}(x) - \frac{\langle\bar{q}q\rangle}{12}\Big(1 - i\frac{m_q\not{x}}{4}\Big) - \frac{\langle\bar{q}q\rangle}{192}m_0^2x^2\Big(1 - i\frac{m_q\not{x}}{6}\Big) - \frac{ig_s}{32\pi^2x^2}\,G^{\mu\nu}(x)\Big[\not{x}\sigma_{\mu\nu} + \sigma_{\mu\nu}\not{x}\Big], \tag{22}$$

$$S_c(x) = S_c^{free}(x) - \frac{g_sm_c}{16\pi^2}\int_0^1 dv\,G^{\mu\nu}(vx)\Big[(\sigma_{\mu\nu}\not{x} + \not{x}\sigma_{\mu\nu})\frac{K_1\big(m_c\sqrt{-x^2}\big)}{\sqrt{-x^2}} + 2\sigma_{\mu\nu}K_0\big(m_c\sqrt{-x^2}\big)\Big], \tag{23}$$

with

$$S_q^{free}(x) = \frac{1}{2\pi x^2}\Big(i\frac{\not{x}}{x^2} - \frac{m_q}{2}\Big), \tag{24}$$

$$S_c^{free}(x) = \frac{m_c^2}{4\pi^2}\left[\frac{K_1\big(m_c\sqrt{-x^2}\big)}{\sqrt{-x^2}} + i\frac{\not{x}\;K_2\big(m_c\sqrt{-x^2}\big)}{(\sqrt{-x^2})^2}\right], \tag{25}$$

where $m_q$ and $\langle\bar{q}q\rangle$ stand for the light-quark mass and condensates, respectively, $G^{\mu\nu}$ being the gluon field-strength tensor, $v$ for the line variable, and $K_0$, $K_1$, and $K_2$ for the modified second type Bessel functions.

The correlation function in Eqs. (20) and (21) would both receive perturbative and non-perturbative contributions. The interactions between photon and light or heavy quarks at short distances are referred to as perturbative contributions, whilst interactions between photon and quarks at large distances are referred to as non-perturbative contributions. To ensure the completeness and reliability of the magnetic dipole moment results obtained from the QCD light-cone sum rule method, it is necessary to include contributions from both regions. For the evaluation of the perturbative contributions we use the following formula:

$$S^{free}(x) \to \int d^4y\,S^{free}(x-z)\not{A}(z)\,S^{free}(z)\,, \tag{26}$$

where one of the propagators for either the light or heavy quark interacts with the photon at a short distance, and the surviving four propagators are assumed to be free.

In the case of the non-perturbative contributions, the following equation is to be used

$$S^{ab}_{\alpha\beta}(x) \to -\frac{1}{4}\big[\bar{q}^a(x)\Gamma_iq^b(0)\big]\big(\Gamma_i\big)_{\alpha\beta}, \tag{27}$$

where it is assumed that the four surviving propagators are full, with one of the quarks interacting with the photon at a large distance. Here $\Gamma_i = \{\mathbf{1},\,\gamma_5,\,\gamma_\mu,\,i\gamma_5\gamma_\mu,\,\sigma_{\mu\nu}/2\}$. Substituting Eq. (27) into Eqs. (20) and (21) yields matrix elements $\langle\gamma(q)\,|\bar{q}(x)\Gamma_iG_{\alpha\beta}q(0)|\,0\rangle$ and $\langle\gamma(q)\,|\bar{q}(x)\Gamma_iq(0)|\,0\rangle$, which are defined in terms of the photon distribution amplitudes. These matrix elements, expressed by photon wave functions, play a crucial role in the evaluation of non-perturbative contributions (See Ref. [53] for details on the distribution amplitudes (DAs) of the photon). It should be noted that the photon DAs used in this study only take into account the contributions of the light quarks. However, in principle, photons can be emitted at a long distance from the c-quark. The matrix elements of nonlocal operators are proportional to the quark condensates, the product of DAs, and some constants. In the case of c-quark, the contribution of non-perturbative constants is negligible and can be neglected. The c-quark condensates are proportional to $1/m_c$. The condensates for the c-quark are to a large extent suppressed because of their large mass [54]. In our calculations, we will therefore not use DAs with c-quark. We have only taken into account the short-distance photon emission from the c-quark. By following the technical schemes mentioned above, we arrive at the QCD representation of the analysis of the magnetic dipole moments. Details of the procedure used to obtain the expression of the perturbative and non-perturbative contributions can be found in Refs. [55, 56] for those who are interested.

After completing the procedures described above, we have finished analytical calculations of the magnetic dipole moments for the relevant states. The results for the magnetic dipole moments of $\Omega_c$-like states are presented in the following section through numerical analysis.

## III. NUMERICAL RESULTS AND DISCUSSION

In this section, we numerically study the magnetic dipole moments of $\Omega_c$-like states. The physical parameters to be calculated are analyzed using the following parameters: $m_u = 2.16^{+0.49}_{-0.26}$ MeV, $m_d = 4.67^{+0.48}_{-0.17}$ MeV, $m_s = 93.4^{+8.6}_{-3.4}$ MeV, $m_c = 1.27 \pm 0.02$ GeV [57], $\langle\bar{q}q\rangle$=$(-0.24 \pm 0.01)^3$ GeV$^3$, $\langle\bar{s}s\rangle$= $0.8\langle\bar{q}q\rangle$ [58], $m_0^2 = 0.8 \pm 0.1$ GeV$^2$, and $\langle g_s^2 G^2\rangle = 0.48 \pm 0.14$ GeV$^4$ [59].The mass and residue values for these states have been calculated using the two-point sum rules of QCD, which have been used in our analysis [49, 50]. The photon distribution amplitudes and the parameters utilized in these distribution amplitudes are borrowed from Ref. [53].

Beyond the parameters listed above, the QCD light-cone sum rules method includes two other parameters known as the Borel mass parameter $M^2$ and the continuum threshold $s_0$, and in the ideal scenario, these two parameters are expected to have a very small impact on the results. To do this, it is necessary to find the working regions of these parameters, where the variation of the results with respect to these parameters should be minimal. The $s_0$ is the point at which the correlation function begins to include contributions from both excited states and the continuum. To determine the working region of this parameter, the assumption $s_0 = (M_H + 0.4^{+0.1}_{-0.1})$ GeV$^2$ is usually adopted. The results are then analyzed for their dependence on slight variations of this parameter. We search for the ideal Borel parameters $M^2$ according to the two criteria: pole contribution (PC) and OPE convergence (CVG). These two conditions are both satisfied in the region 11.4 GeV$^2 \leq s_0 \leq$ 12.8 GeV$^2$ and 3.0 GeV$^2 \leq M^2 \leq$ 4.2 GeV$^2$ for the $\Omega_c$ states; and 11.9 GeV$^2 \leq s_0 \leq$ 13.3 GeV$^2$ and 3.2 GeV$^2 \leq M^2 \leq$ 4.6 GeV$^2$ for the $\Omega_c^*$ states. Our numerical evaluations show that considering these working intervals for the $M^2$, the magnetic dipole moments of the $\Omega_c$-like states PC vary within the interval $34\% \leq$ PC $\leq 61\%$ and $31\% \leq$ PC $\leq 57\%$ for $\Omega_c$ and $\Omega_c^*$ states, respectively. Analyzing the CVG, one sees that the contribution of the higher twist and higher dimensional terms in the OPE is $< 3\%$ of the total and the series shows good convergence for both $\Omega_c$ and $\Omega_c^*$ states. As an example, for fixed values of the continuum threshold $s_0$, the variation of the extracted magnetic dipole moments of the spin-$\frac{1}{2}^-$ $\Omega_c$-like states with Borel mass $M^2$ is shown in Fig.1. As it is seen from this figure, the magnetic dipole moments of these states exhibit a slight dependence on the parameter $M^2$. While the magnetic dipole moments of the $\Omega_c$-like states exhibit some dependence on $s_0$, this is within the limits authorized by the approach and constitutes the main part of the uncertainties.

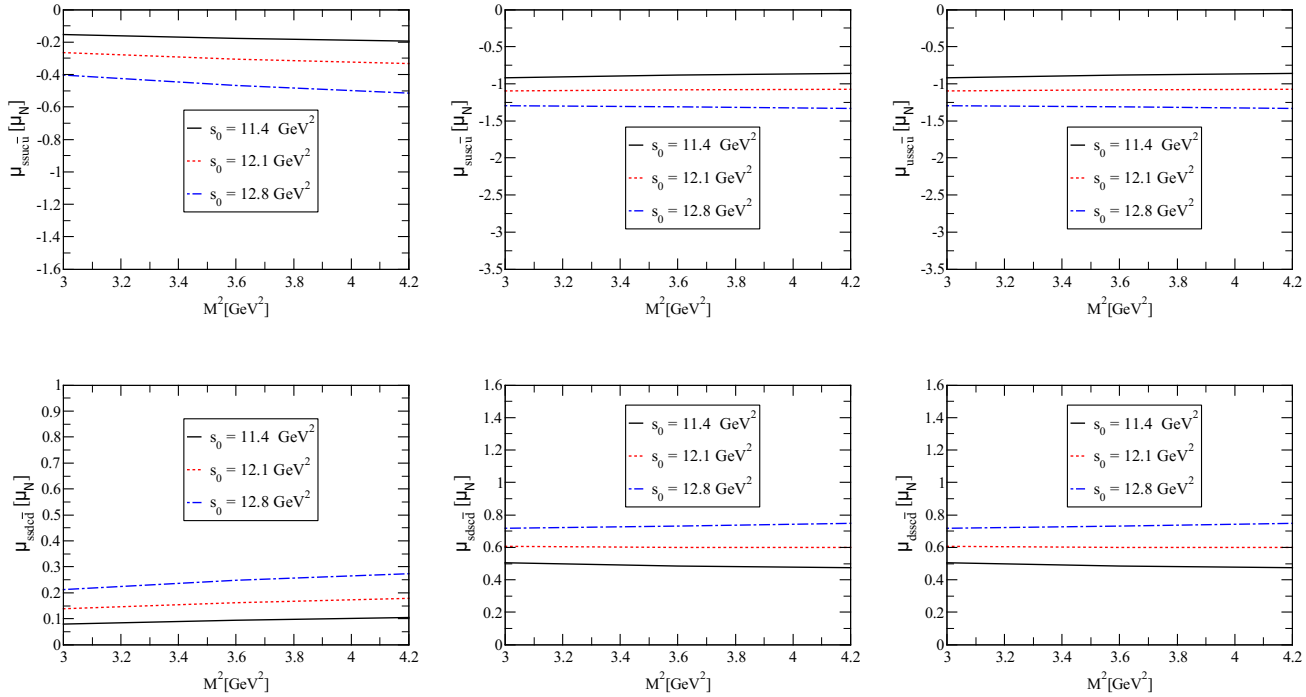



FIG. 1. Dependencies of the magnetic dipole moments of spin-$\frac{1}{2}^-$ $\Omega_c$ states on $M^2$ at three values of $s_0$.

After setting all the relevant parameters, we are finally ready to estimate the numerical values of the corresponding magnetic dipole moments of $\Omega_c$-like states. The predicted results for the magnetic dipole moments are shown in

Table I. The errors of the results, caused by errors in the input parameters and uncertainties in the estimation of the working intervals of the auxiliary parameters, are also shown in Table I.

TABLE I. Results of the magnetic dipole moments of the $\Omega_c$-like states.

| Parameters | $J^P$ | $ssuc\bar{u}$ | $susc\bar{u}$ | $usuc\bar{u}$ | $ssdc\bar{d}$ | $sdsc\bar{d}$ | $dssc\bar{d}$ |
|---|---|---|---|---|---|---|---|
| $\mu[\mu_N]$ | $\frac{1}{2}^-$ | $-0.31^{+0.20}_{-0.16}$ | $-1.10^{+0.23}_{-0.25}$ | $-1.09^{+0.24}_{-0.23}$ | $0.17^{+0.11}_{-0.08}$ | $0.62^{+0.14}_{-0.14}$ | $0.61^{+0.15}_{-0.13}$ |
| $\mu[\mu_N]$ | $\frac{3}{2}^-$ | $0.28^{+0.17}_{-0.13}$ | $0.88^{+0.15}_{-0.18}$ | $0.87^{+0.16}_{-0.16}$ | $-0.058^{+0.032}_{-0.031}$ | $-0.37^{+0.05}_{-0.07}$ | $-0.37^{+0.05}_{-0.07}$ |

The order of magnetic dipole moments can yield insights into their experimental measurability. The analysis of the obtained results indicates that the extracted magnetic dipole moments for both spin-1/2 and spin-3/2 $\Omega_c$-like states can be experimentally measured. Our analysis shows that although hadrons have the same quark content, different diquark components can significantly affect the magnetic dipole moment results. This indicates that the magnetic dipole moment is a crucial parameter in determining the internal structure of hadrons. As previously stated, the interpolating currents used for both spin-1/2 and spin-3/2 $\Omega_c$-like states predicted compatible with mass values. However, the magnetic dipole moments obtained for these states have different magnitudes and signs. This information may help us understand their quantum numbers and internal structure in experimental measurements. Comparing the magnetic dipole moment results obtained in this study with those extracted using different theoretical approaches can provide insight into the consistency of our predictions. The measurement of the magnetic dipole moments of unstable hadrons is a challenging experimental task. Nevertheless, using models approximating key features of QCD can give clues to the relevant dynamical mechanisms behind the observed structure. Together with spectroscopic/decay parameters and electromagnetic properties, it will also be interesting to characterize the branching ratios of the different decay modes and decay channels of $\Omega_c$-like states.

Finally, this section briefly discusses the measurement of the magnetic dipole moment of the $\Omega_c$-like states in the experiment. Due to their short lifetimes, measuring the magnetic dipole moments of $\Omega_c$-like states by spin precession experiments is a challenging task. Instead, the magnetic dipole moment of these unstable states can only be measured indirectly through a three-step process. The particle is first produced, followed by the emission of a low-energy photon that acts as an external magnetic field. As the final step, the particle decays. There is a procedure for the indirect determination of the magnetic dipole moments of these states, which is based on the soft photon emission of hadrons proposed in Ref. [61]. The basic concept underlying the procedure is that the photon carries information about the magnetic dipole and other higher multipole moments of the emitted particle. From this point of view, we can determine the magnetic dipole moment of the particles under consideration through measuring the cross-section or decay width of the radiative process. The magnetic dipole moment of states under investigation can affect the cross sections, whether total or differential. Determining the dipole moment of states under investigation requires comparing theoretical predictions with measured cross-sections. The magnetic dipole moment of $\Delta(1232)$ resonance has been achieved by using the experimental data obtained in the $\gamma N \to \Delta(1232) \to \Delta(1232)\gamma \to \pi N\gamma$ process employing this procedure [62–70].

## IV. SUMMARY AND OUTLOOK

Electromagnetic form factors, which describe the response of composite particles to electromagnetic surveys, provide an important tool for understanding the structure of bound states in QCD. The magnetic dipole moment is the leading-order response of a bound system to a weak external magnetic field. It therefore offers an exquisite laboratory for studying the internal structure of composite particles, which is governed by the quark-gluon dynamics of QCD. Therefore, we systematically study the magnetic dipole moment of controversial states whose internal structure is not elucidated and we try to offer a different point of view to unravel the internal structure of these states. Inspired by the $\Omega_c$ states observed by the LHCb Collaboration, we focus on the scenario of the diquark-diquark-antiquark pentaquark interpretation of the $\Omega_c$-like states with both $J^P = \frac{1}{2}^-$ and $J^P = \frac{3}{2}^-$ quantum numbers, and study the electromagnetic properties of these states with the QCD light-cone sum rules by employing the distribution amplitudes of the photon. From the obtained numerical results, we conclude that the magnetic dipole moments of the $\Omega_c$-like states can reflect their inner structures, which can be used to distinguish their spin-parity quantum numbers. The order of the magnetic dipole moments obtained for these states implies that they are accessible by experiments.

Measuring the magnetic dipole moments of the $\Omega_c$-like states in future experimental facilities can be very helpful for understanding the internal organization and identifying the quantum numbers of these states. We hope that in the future we will have more experimental data on a variety of physical quantities of these states.


[1] H. Albrecht, et al., Observation of a new charmed baryon, Phys. Lett. B 317 (1993) 227–232. doi:10.1016/0370-2693(93)91598-H.

[2] K. W. Edwards, et al., Observation of excited baryon states decaying to Lambda(c)+ pi+ pi-, Phys. Rev. Lett. 74 (1995) 3331–3335. doi:10.1103/PhysRevLett.74.3331.

[3] H. Albrecht, et al., Evidence for Lambda(c)+ (2593) production, Phys. Lett. B 402 (1997) 207–212. doi:10.1016/S0370-2693(97)00503-0.

[4] P. L. Frabetti, et al., An Observation of an Excited State of the $\Lambda_c^+$ Baryon, Phys. Rev. Lett. 72 (1994) 961–964. doi:10.1103/PhysRevLett.72.961.

[5] M. Artuso, et al., Observation of new states decaying into Lambda+(c) pi- pi+, Phys. Rev. Lett. 86 (2001) 4479–4482. arXiv:hep-ex/0010080, doi:10.1103/PhysRevLett.86.4479.

[6] B. Aubert, et al., Observation of a charmed baryon decaying to D0p at a mass near 2.94-GeV/c**2, Phys. Rev. Lett. 98 (2007) 012001. arXiv:hep-ex/0603052, doi:10.1103/PhysRevLett.98.012001.

[7] K. Abe, et al., Experimental constraints on the possible J**P quantum numbers of the Lambda(c)(2880)+, Phys. Rev. Lett. 98 (2007) 262001. arXiv:hep-ex/0608043, doi:10.1103/PhysRevLett.98.262001.

[8] R. Aaij, et al., Study of the $D^0p$ amplitude in $\Lambda_b^0 \to D^0p\pi^-$ decays, JHEP 05 (2017) 030. arXiv:1701.07873, doi:10.1007/JHEP05(2017)030.

[9] V. V. Ammosov, et al., Observation of the production of charmed Sigma(c)*++ baryons in neutrino interactions at the SKAT bubble chamber, JETP Lett. 58 (1993) 247–251.

[10] G. Brandenburg, et al., Observation of two excited charmed baryons decaying into Lambda(c)+ pi+-, Phys. Rev. Lett. 78 (1997) 2304–2308. doi:10.1103/PhysRevLett.78.2304.

[11] R. Mizuk, et al., Observation of an isotriplet of excited charmed baryons decaying to Lambda+(c) pi, Phys. Rev. Lett. 94 (2005) 122002. arXiv:hep-ex/0412069, doi:10.1103/PhysRevLett.94.122002.

[12] P. Avery, et al., Observation of a narrow state decaying into Xi(c)+ pi-, Phys. Rev. Lett. 75 (1995) 4364–4368. arXiv:hep-ex/9508010, doi:10.1103/PhysRevLett.75.4364.

[13] L. Gibbons, et al., Observation of an excited charmed baryon decaying into Xi(c)0 pi+, Phys. Rev. Lett. 77 (1996) 810–813. doi:10.1103/PhysRevLett.77.810.

[14] P. L. Frabetti, et al., Observation of a narrow state decaying into xi0(c) pi+, Phys. Lett. B 426 (1998) 403–410. doi:10.1016/S0370-2693(98)00348-7.

[15] S. E. Csorna, et al., Evidence of new states decaying into Xi-prime(c) pi, Phys. Rev. Lett. 86 (2001) 4243–4246. arXiv:hep-ex/0012020, doi:10.1103/PhysRevLett.86.4243.

[16] J. P. Alexander, et al., Evidence of new states decaying into Xi(c)* pi, Phys. Rev. Lett. 83 (1999) 3390–3393. arXiv:hep-ex/9906013, doi:10.1103/PhysRevLett.83.3390.

[17] R. Aaij, et al., Observation of New $\Xi_c^0$ Baryons Decaying to $\Lambda_c^+K^-$, Phys. Rev. Lett. 124 (22) (2020) 222001. arXiv:2003.13649, doi:10.1103/PhysRevLett.124.222001.

[18] B. Aubert, et al., A Study of anti-B —> Xi(c) anti-Lambda-(c) and anti-B —> anti-Lambda+(c) anti-Lambda-(c) anti-K decays at BABAR, Phys. Rev. D 77 (2008) 031101. arXiv:0710.5775, doi:10.1103/PhysRevD.77.031101.

[19] R. Chistov, et al., Observation of new states decaying into Lambda(c)+ K- pi+ and Lambda(c)+ K0(S) pi-, Phys. Rev. Lett. 97 (2006) 162001. arXiv:hep-ex/0606051, doi:10.1103/PhysRevLett.97.162001.

[20] B. Aubert, et al., A Study of Excited Charm-Strange Baryons with Evidence for new Baryons Xi(c)(3055)+ and Xi(c)(3123)+, Phys. Rev. D 77 (2008) 012002. arXiv:0710.5763, doi:10.1103/PhysRevD.77.012002.

[21] B. Aubert, et al., Observation of an excited charm baryon Omega(C)* decaying to Omega(C)0 gamma, Phys. Rev. Lett. 97 (2006) 232001. arXiv:hep-ex/0608055, doi:10.1103/PhysRevLett.97.232001.

[22] R. Aaij, et al., Observation of five new narrow $\Omega_c^0$ states decaying to $\Xi_c^+K^-$, Phys. Rev. Lett. 118 (18) (2017) 182001. arXiv:1703.04639, doi:10.1103/PhysRevLett.118.182001.

[23] R. Aaij, et al., Observation of excited $\Lambda_b^0$ baryons, Phys. Rev. Lett. 109 (2012) 172003. arXiv:1205.3452, doi:10.1103/PhysRevLett.109.172003.

[24] R. Aaij, et al., Observation of a new baryon state in the $\Lambda_b^0\pi^+\pi^-$ mass spectrum, JHEP 06 (2020) 136. arXiv:2002.05112, doi:10.1007/JHEP06(2020)136.

[25] R. Aaij, et al., Observation of New Resonances in the $\Lambda_b^0\pi^+\pi^-$ System, Phys. Rev. Lett. 123 (15) (2019) 152001. arXiv:1907.13598, doi:10.1103/PhysRevLett.123.152001.

[26] R. Aaij, et al., Observation of Two Resonances in the $\Lambda_b^0\pi^\pm$ Systems and Precise Measurement of $\Sigma_b^\pm$ and $\Sigma_b^{*\pm}$ properties, Phys. Rev. Lett. 122 (1) (2019) 012001. arXiv:1809.07752, doi:10.1103/PhysRevLett.122.012001.

[27] A. M. Sirunyan, et al., Observation of a New Excited Beauty Strange Baryon Decaying to $\Xi_b^-\pi^+\pi^-$, Phys. Rev. Lett. 126 (25) (2021) 252003. arXiv:2102.04524, doi:10.1103/PhysRevLett.126.252003.

[28] R. Aaij, et al., Observation of a new $\Xi_b^-$ resonance, Phys. Rev. Lett. 121 (7) (2018) 072002. arXiv:1805.09418, doi:10.1103/PhysRevLett.121.072002.

[29] R. Aaij, et al., Observation of a new $\Xi_b^0$ state, Phys. Rev. D 103 (1) (2021) 012004. arXiv:2010.14485, doi:10.1103/PhysRevD.103.012004.

[30] R. Aaij, et al., First observation of excited $\Omega_b^-$ states, Phys. Rev. Lett. 124 (8) (2020) 082002. arXiv:2001.00851, doi:10.1103/PhysRevLett.124.082002.

[31] J. Yelton, et al., Observation of Excited $\Omega_c$ Charmed Baryons in $e^+e^-$ Collisions, Phys. Rev. D 97 (5) (2018) 051102. arXiv:1711.07927, doi:10.1103/PhysRevD.97.051102.

[32] R. Aaij, et al., Observation of New Ωc0 States Decaying to the Ξc+K- Final State, Phys. Rev. Lett. 131 (13) (2023) 131902. arXiv:2302.04733, doi:10.1103/PhysRevLett.131.131902.

[33] F.-K. Guo, C. Hanhart, U.-G. Meißner, Q. Wang, Q. Zhao, B.-S. Zou, Hadronic molecules, Rev. Mod. Phys. 90 (1) (2018) 015004, [Erratum: Rev.Mod.Phys. 94, 029901 (2022)]. arXiv:1705.00141, doi:10.1103/RevModPhys.90.015004.

[34] S. L. Olsen, T. Skwarnicki, D. Zieminska, Nonstandard heavy mesons and baryons: Experimental evidence, Rev. Mod. Phys. 90 (1) (2018) 015003. arXiv:1708.04012, doi:10.1103/RevModPhys.90.015003.

[35] N. Brambilla, S. Eidelman, C. Hanhart, A. Nefediev, C.-P. Shen, C. E. Thomas, A. Vairo, C.-Z. Yuan, The $XYZ$ states: experimental and theoretical status and perspectives, Phys. Rept. 873 (2020) 1–154. arXiv:1907.07583, doi:10.1016/j.physrep.2020.05.001.

[36] H.-Y. Cheng, Charmed baryon physics circa 2021, Chin. J. Phys. 78 (2022) 324–362. arXiv:2109.01216, doi:10.1016/j.cjph.2022.06.021.

[37] H.-X. Chen, W. Chen, X. Liu, Y.-R. Liu, S.-L. Zhu, An updated review of the new hadron states, Rept. Prog. Phys. 86 (2) (2023) 026201. arXiv:2204.02649, doi:10.1088/1361-6633/aca3b6.

[38] V. L. Chernyak, I. R. Zhitnitsky, B meson exclusive decays into baryons, Nucl. Phys. B 345 (1990) 137–172. doi:10.1016/0550-3213(90)90612-H.

[39] V. M. Braun, I. E. Filyanov, QCD Sum Rules in Exclusive Kinematics and Pion Wave Function, Z. Phys. C 44 (1989) 157. doi:10.1007/BF01548594.

[40] I. I. Balitsky, V. M. Braun, A. V. Kolesnichenko, Radiative Decay Sigma+ —> p gamma in Quantum Chromodynamics, Nucl. Phys. B 312 (1989) 509–550. doi:10.1016/0550-3213(89)90570-1.

[41] U. Özdem, Analysis of the $X_{AV}$ state through its electromagnetic properties (12 2023). arXiv:2401.00481.

[42] U. Özdem, Magnetic dipole moments of the Ωc(3185)0 and Ωc(3327)0 states from molecular perspective, Phys. Lett. B 849 (2024) 138432. arXiv:2311.02925, doi:10.1016/j.physletb.2023.138432.

[43] U. Ozdem, Electromagnetic properties of the $\Sigma_c(2800)^+$ and $\Lambda_c(2940)^+$ states via light-cone QCD, Eur. Phys. J. C 83 (11) (2023) 1077. arXiv:2309.00959, doi:10.1140/epjc/s10052-023-12251-x.

[44] U. Özdem, K. Azizi, Magnetic moment of the $X_1(2900)$ state in the diquark–antidiquark picture, Eur. Phys. J. A 58 (9) (2022) 171. arXiv:2202.11466, doi:10.1140/epja/s10050-022-00815-6.

[45] U. Özdem, Magnetic moment of the $\Xi_b(6227)$ as a molecular pentaquark state, Eur. Phys. J. Plus 137 (1) (2022) 103. arXiv:2109.09313, doi:10.1140/epjp/s13360-022-02339-w.

[46] K. Azizi, U. Özdem, Magnetic dipole moments of the Tcc+ and ZV++ tetraquark states, Phys. Rev. D 104 (11) (2021) 114002. arXiv:2109.02390, doi:10.1103/PhysRevD.104.114002.

[47] K. Azizi, U. t. Özdem, The electromagnetic multipole moments of the possible charm-strange pentaquarks in light-cone QCD, Eur. Phys. J. C 78 (9) (2018) 698. arXiv:1807.06503, doi:10.1140/epjc/s10052-018-6187-0.

[48] K. Azizi, U. Özdem, The electromagnetic multipole moments of the charged open-flavor $Z_{\bar{c}q}$ states, J. Phys. G 45 (5) (2018) 055003. arXiv:1802.07711, doi:10.1088/1361-6471/aab56b.

[49] H.-J. Wang, Z.-Y. Di, Z.-G. Wang, Analysis of the excited Ω c states as the ± pentaquark states with QCD sum rules, Commun. Theor. Phys. 73 (3) (2021) 035201. doi:10.1088/1572-9494/abc7b1.

[50] Z.-G. Wang, J.-X. Zhang, Possible pentaquark candidates: new excited $\Omega_c$ states, Eur. Phys. J. C 78 (6) (2018) 503. arXiv:1804.06195, doi:10.1140/epjc/s10052-018-5989-4.

[51] K.-C. Yang, W. Y. P. Hwang, E. M. Henley, L. S. Kisslinger, QCD sum rules and neutron proton mass difference, Phys. Rev. D 47 (1993) 3001–3012. doi:10.1103/PhysRevD.47.3001.

[52] V. M. Belyaev, B. Y. Blok, CHARMED BARYONS IN QUANTUM CHROMODYNAMICS, Z. Phys. C 30 (1986) 151. doi:10.1007/BF01560689.

[53] P. Ball, V. M. Braun, N. Kivel, Photon distribution amplitudes in QCD, Nucl. Phys. B 649 (2003) 263–296. arXiv:hep-ph/0207307, doi:10.1016/S0550-3213(02)01017-9.

[54] D. Antonov, J. E. F. T. Ribeiro, Quark condensate for various heavy flavors, Eur. Phys. J. C 72 (2012) 2179. arXiv:1209.0408, doi:10.1140/epjc/s10052-012-2179-7.

[55] U. Özdem, Electromagnetic form factors of the Bc-like tetraquarks: Molecular and diquark-antidiquark pictures, Phys. Lett. B 838 (2023) 137750. arXiv:2211.10169, doi:10.1016/j.physletb.2023.137750.

[56] U. Özdem, Electromagnetic properties of doubly heavy pentaquark states, Eur. Phys. J. Plus 137 (2022) 936. arXiv:2201.00979, doi:10.1140/epjp/s13360-022-03125-4.

[57] R. L. Workman, et al., Review of Particle Physics, PTEP 2022 (2022) 083C01. doi:10.1093/ptep/ptac097.

[58] B. L. Ioffe, QCD at low energies, Prog. Part. Nucl. Phys. 56 (2006) 232–277. arXiv:hep-ph/0502148, doi:10.1016/j.ppnp.2005.05.001.

[59] S. Narison, $\overline{m}_{c,b}$, $< \alpha_s G^2 >$ and $\alpha_s$ from Heavy Quarkonia, Nucl. Part. Phys. Proc. 300-302 (2018) 153–164. doi:10.1016/j.nuclphysbps.2018.12.026.

[60] J. Rohrwild, Determination of the magnetic susceptibility of the quark condensate using radiative heavy meson decays, JHEP 09 (2007) 073. arXiv:0708.1405, doi:10.1088/1126-6708/2007/09/073.

[61] V. I. Zakharov, L. A. Kondratyuk, L. A. Ponomarev, Bremsstrahlung and determination of electromagnetic parameters of particles, Yad. Fiz. 8 (1968) 783–792.

[62] V. Pascalutsa, M. Vanderhaeghen, Magnetic moment of the Delta(1232)-resonance in chiral effective field theory, Phys. Rev. Lett. 94 (2005) 102003. arXiv:nucl-th/0412113, doi:10.1103/PhysRevLett.94.102003.

[63] V. Pascalutsa, M. Vanderhaeghen, Chiral effective-field theory in the Delta(1232) region: I. Pion electroproduction on the nucleon, Phys. Rev. D 73 (2006) 034003. arXiv:hep-ph/0512244, doi:10.1103/PhysRevD.73.034003.

[64] V. Pascalutsa, M. Vanderhaeghen, Chiral effective-field theory in the Delta(1232) region. II. Radiative pion photoproduction, Phys. Rev. D 77 (2008) 014027. arXiv:0709.4583, doi:10.1103/PhysRevD.77.014027.

[65] M. Kotulla, et al., The Reaction gamma p —> pi zero gamma-prime p and the magnetic dipole moment of the delta+(1232) resonance, Phys. Rev. Lett. 89 (2002) 272001. arXiv:nucl-ex/0210040, doi:10.1103/PhysRevLett.89.272001.

[66] D. Drechsel, M. Vanderhaeghen, Magnetic dipole moment of the Delta+ (1232) from the gamma p —> gamma pi0 pi reaction, Phys. Rev. C 64 (2001) 065202. arXiv:hep-ph/0105060, doi:10.1103/PhysRevC.64.065202.

[67] A. I. Machavariani, A. Faessler, A. J. Buchmann, Field-theoretical description of electromagnetic Delta resonance production and determination of the magnetic moment of the $\Delta^+(1232)$ resonance by the $ep \to e'N'\pi'\gamma'$ and $\gamma p \to N'\pi'\gamma'$ reactions, Nucl. Phys. A 646 (1999) 231–257, [Erratum: Nucl.Phys.A 686, 601–603 (2001)]. doi:10.1016/S0375-9474(98)00611-3.

[68] D. Drechsel, M. Vanderhaeghen, M. M. Giannini, E. Santopinto, Inelastic photon scattering and the magnetic moment of the Delta (1232) resonance, Phys. Lett. B 484 (2000) 236–242. arXiv:nucl-th/0003035, doi:10.1016/S0370-2693(00)00654-7.

[69] W.-T. Chiang, M. Vanderhaeghen, S. N. Yang, D. Drechsel, Unitary model for the gamma p —> gamma pi0 p reaction and the magnetic dipole moment of the Delta+ (1232), Phys. Rev. C 71 (2005) 015204. arXiv:hep-ph/0409078, doi:10.1103/PhysRevC.71.015204.

[70] A. I. Machavariani, A. Faessler, Propagator of the Delta resonance and determination of the magnetic moment of the Delta+ from the gamma p —> gamma pi0 p reaction, Phys. Rev. C 72 (2005) 024002. doi:10.1103/PhysRevC.72.024002.

# Beauty-charm Meson Family with Coupled Channel Effects and Their Strong Decays

Wei Hao[1] and Ruilin Zhu[2, 3, *]

[1] School of Physics, Nankai University, Tianjin 300071, China

[2] Department of Physics and Institute of Theoretical Physics, Nanjing Normal University, Nanjing, Jiangsu 210023, China

[3] Peng Huanwu Innovation Research Center, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China

arXiv:2402.18898v1 [hep-ph] 29 Feb 2024

We systematically calculate the spectrum and hadronic decays of the beauty-charm system in a coupled channel framework. The unquenched effects are induced by the $^3P_0$ model. Our results can good explain the observed $B_c$ meson spectrum. For the coupled channel components, we predicted the $1S$ is about 4%, the $2S$, $1P$, $2P$, $1D$, and $2D$ states are about 14%, 10%, 33%, and 17% respectively. For the $3S$, $2P$ and $2D$ states, the strong decay is allowed, The hadronic decay widths of the $3^1S_0$, $3^3S_1$, $2^3P_2$ states are about 110 MeV, 69 MeV, and 3 MeV, respectively. While the decay widths of the $2^3D_1$, $2D$, $2D'$, and $2^3D_2$ states are 60 MeV, 149 MeV, 65 MeV, and 72 MeV, respectively.

## I. INTRODUCTION

The understanding of hadron structures and their transitions at Fermi scale is fundamental issue from both the theoretical and experimental aspects in Particle Physics. For the conventional meson spectrum composed of a quark and an antiquark, the beauty-charm meson family is relatively incomplete. Up to now, only three beauty-charm mesons have been observed in experiments, i.e, the $B_c(1S)$, the $B_c(2S)$ and $B_c^*(2S)$.

The ground state $B_c(1^1S_0)$ of beauty-charm family was first discovered in 1998 by CDF at Fermilab [1]. The latest average mass for this state is $6274.47 \pm 0.27 \pm 0.17$ MeV [2]. One of the radial excited beauty-charm states was first discovered in $B_c^+\pi^+\pi^-$ invariant mass spectrum with subprocess $B_c^+ \to J/\psi\pi^+$ using the sample corresponding to 4.9 fb$^{-1}$ of 7 TeV and 19.2 fb$^{-1}$ of 8 TeV $pp$ collision data collected by the ATLAS experiment at the LHC in 2014 [3]. After that, two excited beauty-charm states $B_c(2^1S_0)$ and $B_c^*(2^3S_1)$ instead of one peak are discovered in $B_c^+\pi^+\pi^-$ invariant mass spectrum from both the CMS and LHCb experiments [4, 5]. The combined average mass for the $B_c(2^1S_0)$ is determined as $6871.2 \pm 0.1$ MeV [2]. For the vector excited state, $B_c^*(2^3S_1)$ first decays via hadronic transition $B_c^{*+}(2^3S_1) \to B_c^{*+}(1^3S_1)\pi^+\pi^-$, and then the vector $B_c^{*+}(1^3S_1)$ decays via electromagnetic transition $B_c^{*+}(1^3S_1) \to B_c^+(1^1S_0) + \gamma$. However, the radiated photon is soft with energy around 60MeV, which is not reconstructed in both the CMS and LHCb experiments. Thus the value of the mass of $B_c^{*+}(2^3S_1)$ relies on the precise information of $B_c^{*+}(1^3S_1)$. On the other hand, other beauty-charm meson states including orbitally excited states are not observed at experiments yet.

In theoretical aspects, the mass spectra of beauty-charm mesons have been studied by many groups. For example, the quark potential models [6–15], QCD sum rule [16–20], the heavy quark effective theory [21], and the Dyson-Schwinger equation approach of QCD [22, 23]. Besides, the properties of the low-lying $B_c$ mesons are also investigated in the lattice QCD based on the first principles [24–26].

The quark potential models in Refs. [6–12, 14] are usually called quenched quark models. In quenched quark models, the convention mesons are constituted by a quark and an untiquark, and the mass spectrum comes from the interactions between constituent quarks. Therein the Godfrey-Isgur relativistic quark model [14] is usu-

*Electronic address: rlzhu@njnu.edu.cn

ally thought to provide a good description for most of the meson spectra. However, the quenched quark models sometimes explain poorly for higher excited states beyond the two-body threshold because they miss the generation of the light quark-antiquark pairs which enlarge the Fock space of the initial state [27]. These multiquark components will change the Hamiltonian of the quark potential models, and then lead to mass shift and mixing among states with the same quantum numbers. If the initial state is above two-body threshold, the open channel strong decay will be allowed [27]. In other words, the unquenched quark model includes virtual hadronic loops. The hadronic loop has turned out to be highly nontrivial and can give rise to mass shifts to the bare hadron states and contribute continuum components to the physical hadron states [28].

The coupled-channel model as one of the unquenched quark model, which is usually neglected, will manifest as a coupling to meson-meson (meson-baryon) channels and lead to mass shifts. The effects of quark-antiquark pairs are introduced explicitly into the constituent quark model via a QCD-inspired $^3P_0$ pair-creation mechanism. The approach is based on a constituent quark model to which the quark-antiquark pairs with vacuum quantum numbers are added as a perturbation. The pair-creation mechanism is inserted at the quark level and the one-loop diagrams are calculated by summing over the possible intermediate states [29]. It has been shown that the coupled-channel effects play an important role for describing the mesons spectra, such as charmonium [30–33], bottomonium [27–29, 34], and charmed-strange mesons [35–45]. In this paper, we will use this kind of unquenched quark model to study the $B_c$ mesons. we will investigate the mass spectrum of the beauty-charm mesons within the nonrelativistic quark model by taking into account the mass shifts from the coupled-channel effect.

This paper is organized as follows. In Section II, we explain the calculation framework of coupled channel effects, the details of $^3P_0$ model and the quenched model. Section III is devoted to the analysis of the results of our coupled channel calculation with various detailed comparisons and elucidation of $B_c$ mesons. Finally, we give a short summary of this work in Section IV.

## II. THEORETICAL FORMALISM

### A. Non-relativistic quark model

The quenched part of the Hamiltonian $H_A$ is taken from the potential model. This model was proposed by Lakhina and Swanson [46], and has been used to study the bottom mesons [47] and open charm mesons [48]. The Hamiltonian $H_A$ can be split into a term $H_0$ which could be solved non-perturbatively and a spin-orbital and tensor term $H_{sd}$ which we solve in perturbation theory.

$$H_A = H_0 + H_{sd}, \tag{1}$$

$$H_0 = m_q + m_{\bar{q}} + \frac{\nabla^2}{2M_r} - \frac{4}{3}\frac{\alpha_s}{r} + br + C_{q\bar{q}} + \frac{32\alpha_s\sigma^3 e^{-\sigma^2 r^2}}{9\sqrt{\pi}m_q m_{\bar{q}}}\boldsymbol{S}_q\cdot\boldsymbol{S}_{\bar{q}}, \tag{2}$$

where $M_r$ is the reduced mass of which equals to $M_r = m_q m_{\bar{q}}/(m_q + m_{\bar{q}})$. When combined together with the subscript, $\boldsymbol{S}$ and $m$ stand for the spin and mass of quark or antiquark, respectively. where $m_q, m_{\bar{q}}, \alpha_s, b, C_{q\bar{q}}$ and $\sigma$ are the free parameters we need to refit in this quenched model.

The $H_{sd}$ can be compressed as

$$H_{sd} = \left(\frac{\boldsymbol{S}_q}{2m_q^2} + \frac{\boldsymbol{S}_{\bar{q}}}{2m_{\bar{q}}^2}\right)\cdot\boldsymbol{L}\left(\frac{1}{r}\frac{dV_c}{dr} + \frac{2}{r}\frac{dV_1}{dr}\right) + \frac{\boldsymbol{S}_+\cdot\boldsymbol{L}}{m_q m_{\bar{q}}}\left(\frac{1}{r}\frac{dV_2}{r}\right) + \frac{3\boldsymbol{S}_q\cdot\hat{\boldsymbol{r}}\boldsymbol{S}_{\bar{q}}\cdot\hat{\boldsymbol{r}} - \boldsymbol{S}_q\cdot\boldsymbol{S}_{\bar{q}}}{3m_q m_{\bar{q}}}V_3 + \left[\left(\frac{\boldsymbol{S}_q}{m_q^2} - \frac{\boldsymbol{S}_{\bar{q}}}{m_{\bar{q}}^2}\right) + \frac{\boldsymbol{S}_-}{m_q m_{\bar{q}}}\right]\cdot\boldsymbol{L}V_4, \tag{3}$$

where

$$\begin{aligned} V_c &= -\frac{4}{3}\frac{\alpha_s}{r} + br, \\ V_1 &= -br - \frac{2}{9\pi}\frac{\alpha_s^2}{r}[9\ln(\sqrt{m_q m_{\bar{q}}}r) + 9\gamma_E - 4], \\ V_2 &= -\frac{4}{3}\frac{\alpha_s}{r} - \frac{1}{9\pi}\frac{\alpha_s^2}{r}[-18\ln(\sqrt{m_q m_{\bar{q}}}r) + 54\ln(\mu r) \\ &\quad +36\gamma_E + 29], \\ V_3 &= -\frac{4\alpha_s}{r^3} - \frac{1}{3\pi}\frac{\alpha_s^2}{r^3}[-36\ln(\sqrt{m_q m_{\bar{q}}}r) + 54\ln(\mu r) \\ &\quad +18\gamma_E + 31], \\ V_4 &= \frac{1}{\pi}\frac{\alpha_s^2}{r^3}\ln\left(\frac{m_{\bar{q}}}{m_q}\right), \end{aligned} \tag{4}$$

where $\boldsymbol{S}_\pm = \boldsymbol{S}_q \pm \boldsymbol{S}_{\bar{q}}$, $\boldsymbol{L}$ is the orbital angular momentum of the $q\bar{q}$ system. $\gamma_E$ is Euler constant, $C_F = 4/3$ and $C_A = 3$ are SU(3) gauge group factors and $\mu$ is renormalization scale. We have set the same value for the renormalization scale $\mu = 1$ GeV as in Ref. [46].

The spin-orbit term in the $H_{sd}$ can be further rewritten into symmetric part $H_{sym}$ and antisymmetric part $H_{anti}$,

$$\begin{aligned} H_{sym} &= \frac{\boldsymbol{S}_+ \cdot \boldsymbol{L}}{2}\left[\left(\frac{1}{2m_q^2} + \frac{1}{2m_{\bar{q}}^2}\right)\left(\frac{1}{r}\frac{dV_c}{dr} + \frac{2}{r}\frac{dV_1}{dr}\right)\right. \\ &\quad \left. + \frac{2}{m_q m_{\bar{q}}}\left(\frac{1}{r}\frac{dV_2}{r}\right) + \left(\frac{1}{m_q^2} - \frac{1}{m_{\bar{q}}^2}\right)V_4\right], \end{aligned} \tag{5}$$

$$\begin{aligned} H_{anti} &= \frac{\boldsymbol{S}_- \cdot \boldsymbol{L}}{2}\left[\left(\frac{1}{2m_q^2} - \frac{1}{2m_{\bar{q}}^2}\right)\left(\frac{1}{r}\frac{dV_c}{dr} + \frac{2}{r}\frac{dV_1}{dr}\right)\right. \\ &\quad \left. + \left(\frac{1}{m_q^2} + \frac{1}{m_{\bar{q}}^2} + \frac{2}{m_q m_{\bar{q}}}\right)V_4\right]. \end{aligned} \tag{6}$$

The antisymmetric part $H_{anti}$ gives rise to the the spin-orbit mixing of the mesons, such as $B_c(n^3L_L) - B_c(n^1L_L)$. This mixing can be parametrized by a mixing angle $\theta_{nL}$ by the following formula [14, 49],

$$\begin{pmatrix} B_{cL}(nL) \\ B'_{cL}(nL) \end{pmatrix} = \begin{pmatrix} \cos\theta_{nL} & \sin\theta_{nL} \\ -\sin\theta_{nL} & \cos\theta_{nL} \end{pmatrix} \begin{pmatrix} B_c(n^1L_L) \\ B_c(n^3L_L) \end{pmatrix}, \tag{7}$$

where $B_{cL}(nL)$ and $B'_{cL}(nL)$ represent the physical observed states.

### B. $^3P_0$ Model and Coupled Channel Effects

In the coupled channel model, the Hamiltonian is defined as

$$H = H_A + H_{BC} + H_I, \tag{8}$$

where $H_A$ is the Hamiltonian of the quark-antiquark component. In this work, we use a nonrelativistic potential model in Eq. (1) which has been introduced in above section to calculate the quenched part. $H_{BC}$ is the Hamiltonian between the meson pairs component, which is expressed as

$$H_{BC} = E_{BC} = \sqrt{m_B^2 + p^2} + \sqrt{m_C^2 + p^2}. \tag{9}$$

$H_I$ is the term which induces the mixing between $q\bar{q}$ bare state and BC meson pair system. In this work, we use the widely used $^3P_0$ model or called the quark pair creation model [50–52] to analyse the mixing of the bare states and mesons pairs system. In the $^3P_0$ model, the generated quark-antiquark pairs have quantum numbers with $J^{PC} = 0^{++}$ identical to that of vacuum. And the pairs can mix with initial quark-antiquark mesons to create two final mesons.

In the $^3P_0$ model, the quark-antiquark pair-creation operator $T^\dagger$ is expressed as [29, 32, 34]

$$\begin{aligned} T^\dagger &= -3\,\gamma_0^{eff}\int d\vec{p}_3\, d\vec{p}_4\, \delta(\vec{p}_3 + \vec{p}_4)\, C_{34}\, F_{34}\, e^{-r_q^2(\vec{p}_3 - \vec{p}_4)^2/6} \\ &\quad [\chi_{34} \times \mathcal{Y}_1(\vec{p}_3 - \vec{p}_4)]_0^{(0)}\; b_3^\dagger(\vec{p}_3)\, d_4^\dagger(\vec{p}_4)\;, \end{aligned} \tag{10}$$

where $C_{34}$, $F_{34}$ and $\chi_{34}$ are the color singlet wave function, flavor singlet wave function and spin triplet wave function of the $q\bar{q}$ respectively. $b_3^\dagger(\vec{p}_3)$ and $d_4^\dagger(\vec{p}_4)$ are the creation operators for a quark and an antiquark with momenta $\vec{p}_3$ and $\vec{p}_4$, respectively. $\gamma_0^{eff} = \frac{m_n}{m_i}\gamma_0$ is the pair-creation strength, whose value is a typical value $\gamma_0 = 0.4$ that gives a reasonably accurate description of the overall scale of decay widths of both light and heavy mesons [7]. $m_n$ refers to the light quark mass $m_u$, and $m_i$ refers to the quark mass $m_u, m_d$ or $m_s$. In the $^3P_0$ model, the operator $T^\dagger$ creates a pair of constituent quarks with an actual size, where the pair-creation point has to be smeared out by a Gaussian factor, whose width $r_q$ was determined from meson decays to be in the range 0.25 to 0.35 fm [53–56]. In our calculation, we take the value $r_q = 0.3$ fm.

The eigenfunction of the full Hamiltonian $H$ can be expressed as,

$$|\psi\rangle = c_0|\psi_0\rangle + \sum_{BC}\int d^3p\, c_{BC}(p)|BC;p\rangle, \tag{11}$$

TABLE I: Parameters refitted in this work.

| Parameters | Values |
|---|---|
| $m_n$ | 0.45 GeV |
| $m_s$ | 0.55 GeV |
| $m_c$ | 1.43 GeV |
| $m_b$ | 4.5 GeV |
| $\alpha_s$ | 0.51881 |
| $b$ | 0.16164 GeV$^2$ |
| $\sigma$ | 1.3381 GeV |
| $C_{bc}$ | 0.449 GeV |
| $\gamma_0$ | 0.4 |

where $c_0$ is the normalization constant before the $q\bar{q}$ bare state. $c_{BC}(p)$ is the normalization constant with specific momentum $p$ for $BC$ molecular components.

Meanwhile, the eigenvalue $M$ of $H$ in Eq. 8 is the theoretical prediction of the coupled channel model which can be decomposed into two terms [30],

$$M = M_0 + \Delta M \tag{12}$$

$$\Delta M = \sum_{BC} \int_0^\infty p^2 dp \frac{\left|\langle BC; p| T^\dagger |\psi_0\rangle\right|^2}{M - E_{BC} + i\epsilon} \tag{13}$$

where $M_0$ is the eigenvalue of the quenched Hamiltonian $H_A$ and $\Delta M$ is the mass shift which signifies the deviation between $H$ and $H_A$. If the initial state $A$ is above the threshold of $B$ and $C$, a strong decay process $A \to BC$ will happen and $\Delta M$ will pick up a imaginary part which equals to the one half of the decay width.

$$\Gamma_{BC} = 2\pi p_0 \frac{E_B(p_0) E_C(p_0)}{m_A} \left|\langle BC; p_0| T^\dagger |\psi_0\rangle\right|^2 \tag{14}$$

For states below the $BC$ threshold, it is reasonable to normalize the physical state $|\psi\rangle$, and the probability of quenched quark pairs can be calculated as,

$$\begin{aligned} P_{q\bar{q}} &\equiv |c_0|^2 \\ &= \left(1 + \sum_{BC} \int_0^\infty p^2 dp \frac{\left|\langle BC; p\ell J| T^\dagger |\psi_0\rangle\right|^2}{(M - E_{BC})^2}\right)^{-1}, \end{aligned} \tag{15}$$

and the probability of the molecular component are naturally expressed as $P_{BC} = 1 - P_{q\bar{q}}$. We also use $P_{\text{molecule}}$ to represent the sum of all the coupled channel probabilities.

TABLE II: The mass spectrum (in MeV) of the $b\bar{c}$ mesons. Columns from 3 to 5 stand for spectrum from the potential model, the mass shift, the spectrum with coupled channel effects. The last Column is the experimental values taken from RPP [2]. The mixing angle of the $1P$, $2P$, $1D$ and $2D$ are $-35.0^\circ$, $-36.4^\circ$, $-42.8^\circ$ and $-44.1^\circ$, respectively.

| $n^{2S+1}L_J$ | State | $M_0$ | $\Delta M$ | $M$ | RPP [2] |
|---|---|---|---|---|---|
| $1^1S_0$ | $B_c^\pm$ | 6332 | $-60$ | 6272 | $6274.47 \pm 0.32$ |
| $1^3S_1$ | – | 6399 | $-64$ | 6335 | – |
| $2^1S_0$ | $B_c^\pm(2S)$ | 6983 | $-108$ | 6875 | $6871.2 \pm 1.0$ |
| $2^3S_1$ | – | 7007 | $-110$ | 6897 | – |
| $3^1S_0$ | – | 7388 | $-137$ | 7250 | – |
| $3^3S_1$ | – | 7404 | $-124$ | 7280 | – |
| $1^3P_0$ | – | 6798 | $-91$ | 6707 | – |
| $1P$ | – | 6846 | $-96$ | 6751 | – |
| $1P'$ | – | 6881 | $-95$ | 6786 | – |
| $1^3P_2$ | – | 6901 | $-99$ | 6802 | – |
| $2^3P_0$ | – | 7220 | $-122$ | 7099 | – |
| $2P$ | – | 7258 | $-131$ | 7127 | – |
| $2P'$ | – | 7294 | $-129$ | 7165 | – |
| $2^3P_2$ | – | 7310 | $-138$ | 7172 | – |
| $1^3D_1$ | – | 7142 | $-117$ | 7025 | – |
| $1D$ | – | 7148 | $-116$ | 7033 | – |
| $1D'$ | – | 7157 | $-116$ | 7041 | – |
| $1^3D_3$ | – | 7157 | $-115$ | 7042 | – |
| $2^3D_1$ | – | 7489 | $-129$ | 7360 | – |
| $2D$ | – | 7499 | $-146$ | 7354 | – |
| $2D'$ | – | 7503 | $-127$ | 7376 | – |
| $2^3D_3$ | – | 7508 | $-170$ | 7337 | – |

## III. RESULTS AND DISCUSSIONS

After solving the Schrödinger equation with Hamiltonian $H_0$, the wave functions are used to determine the leading order correction of $H_{sd}$ in the quenched quark model and the $^3P_0$ model to the spectrum and the strong decays.

The refitted parameters are listed in Table I. Because only two $B_c$ mesons have experimental information, it is difficult to fit all parameters. So We adopt a strategy similar to that in Ref. [7]. In our calculation, the parameters $\alpha_s$, $b$ and $\sigma$ are determined by fitting the

TABLE III: Mass shift $\Delta M$ (in MeV) of each coupled channel.

| State | $BD$ | $BD^*$ | $B^*D$ | $B^*D^*$ | $B_sD_s$ | $B_sD_s^*$ | $B_s^*D_s$ | $B_s^*D_s^*$ | Total |
|---|---|---|---|---|---|---|---|---|---|
| $1^1S_0$ | 0 | −11 | −12 | −22 | 0 | −4 | −4 | −8 | −60 |
| $1^3S_1$ | −4 | −8 | −9 | −28 | −1 | −3 | −3 | −9 | −64 |
| $2^1S_0$ | 0 | −20 | −24 | −41 | 0 | −5 | −6 | −11 | −108 |
| $2^3S_1$ | −8 | −14 | −16 | −49 | −2 | −4 | −4 | −13 | −110 |
| $3^1S_0$ | 0 | −34 | −25 | −57 | 0 | −5 | −6 | −10 | −136 |
| $3^3S_1$ | 3 | −30 | 2 | −75 | −2 | −3 | −4 | −12 | −122 |
| $1^3P_0$ | −12 | 0 | 0 | −58 | −3 | 0 | 0 | 18 | −91 |
| $1P$ | 0 | −18 | −20 | −36 | 0 | −5 | −6 | −11 | −95 |
| $1P'$ | 0 | −14 | −16 | −43 | 0 | −4 | −5 | −13 | −95 |
| $1^3P_2$ | −10 | −13 | −15 | −39 | −3 | −4 | −4 | −11 | −99 |
| $2^3P_0$ | −33 | 0 | 0 | −68 | −4 | 0 | 0 | −17 | −122 |
| $2P$ | 0 | −25 | −35 | −48 | 0 | −5 | −6 | −11 | −131 |
| $2P'$ | 0 | −23 | −34 | −51 | 0 | −5 | −5 | −12 | −129 |
| $2^3P_2$ | −19 | −17 | −22 | −58 | −3 | −4 | −4 | −12 | −138 |
| $1^3D_1$ | −13 | −4 | −6 | −70 | −3 | −1 | −1 | −19 | −117 |
| $1D$ | 0 | −22 | −26 | −44 | 0 | −6 | −6 | −12 | −116 |
| $1D'$ | 0 | −19 | −23 | −50 | 0 | −5 | −5 | −14 | −116 |
| $1^3D_3$ | −13 | −16 | −18 | −44 | −3 | −4 | −5 | −11 | −115 |
| $2^3D_1$ | −6 | 1 | 1 | −99 | −6 | −1 | −2 | −16 | −129 |
| $2D$ | 0 | −26 | −18 | −80 | 0 | −5 | −6 | −10 | −145 |
| $2D'$ | 0 | −11 | −12 | −76 | 0 | −5 | −7 | −12 | −123 |
| $2^3D_3$ | −10 | −21 | −20 | −98 | −3 | −3 | −4 | −10 | −169 |

TABLE IV: Probabilities (in %) of the coupled channels considered in this work.

| State | $BD$ | $BD^*$ | $B^*D$ | $B^*D^*$ | $B_sD_s$ | $B_sD_s^*$ | $B_s^*D_s$ | $B_s^*D_s^*$ | $P_{molecule}$ | $P_{b\bar{c}}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $1^1S_0$ | 0 | 0.7 | 0.8 | 1.4 | 0 | 0.2 | 0.2 | 0.4 | 3.7 | 96.3 |
| $1^3S_1$ | 0.3 | 0.5 | 0.6 | 1.8 | 0.1 | 0.1 | 0.2 | 0.5 | 4.3 | 95.7 |
| $2^1S_0$ | 0 | 2.8 | 3.9 | 5.1 | 0 | 0.5 | 0.7 | 1.1 | 14.0 | 86.0 |
| $2^3S_1$ | 1.6 | 1.9 | 2.7 | 6.2 | 0.2 | 0.4 | 0.5 | 1.3 | 14.8 | 85.2 |
| $1^3P_0$ | 1.8 | 0 | 0 | 5.2 | 0.4 | 0 | 0 | 1.3 | 8.6 | 91.4 |
| $1^1P_1$ | 0 | 1.8 | 2.4 | 3.5 | 0 | 0.4 | 0.5 | 0.9 | 9.6 | 90.4 |
| $1^3P_1$ | 0 | 1.6 | 2.1 | 4.0 | 0 | 0.4 | 0.4 | 1.0 | 9.5 | 90.5 |
| $1^3P_2$ | 1.2 | 1.3 | 1.7 | 4.1 | 0.3 | 0.3 | 0.4 | 1.0 | 10.2 | 89.8 |
| $2^3P_0$ | 22.6 | 0 | 0 | 9.0 | 0.7 | 0 | 0 | 1.4 | 33.7 | 66.3 |
| $2^1P_1$ | 0 | 5.3 | 16.8 | 8.4 | 0 | 0.6 | 0.8 | 1.1 | 32.9 | 67.1 |
| $2^3P_1$ | 0 | 5.3 | 18.3 | 7.5 | 0 | 0.5 | 0.7 | 1.1 | 33.4 | 66.6 |
| $1^3D_1$ | 4.3 | 0.9 | 1.5 | 8.1 | 0.4 | 0.1 | 0.2 | 1.7 | 17.3 | 82.7 |
| $1^1D_2$ | 0 | 3.2 | 4.7 | 5.8 | 0 | 0.6 | 0.7 | 1.1 | 16.2 | 83.8 |
| $1^3D_2$ | 0 | 3.1 | 4.8 | 6.2 | 0 | 0.5 | 0.7 | 1.2 | 16.5 | 83.5 |
| $1^3D_3$ | 2.2 | 2.0 | 2.6 | 6.6 | 0.4 | 0.4 | 0.5 | 1.2 | 15.9 | 84.1 |

mass spectrum of the $B_c$, $B_c^*$ and $B_c^{(*)}(2S)$, and the other mesons come from previous papers [7, 46, 48]. For the $B_c(1^3S_1)$ and $B_c(2^3S_1)$, there are no experimental data, but their values can be estimate to be around 6334 MeV and 6900 MeV, respectively. Because the hyperfine mass splitting in bottomonium family is measured as $\Delta M_{b\bar{b}(1S)} = 62.3 \pm 3.2$ MeV and $\Delta M_{b\bar{b}(2S)} = 24 \pm 4$ MeV [2], the hyperfine mass splitting in beauty-charm meson family is believed to be smaller as $m_{B_c^*} - m_{B_c} \geq \Delta M_{b\bar{b}(1S)} = 62.3 \pm 3.2$ MeV and $m_{B_c^*(2S)} - m_{B_c(2S)} \geq \Delta M_{b\bar{b}(2S)} = 24 \pm 4$ MeV.

With the parameters in Table I, the mass spectrum and mass shifts of the $B_c$ mesons can be estimated. The results are shown in Fig. 1, with numbers listed in Table II. The spectrum can good describe the experimental masses of the $B_c^+$ and $B_c(2S)^+$. Besides, the mixing angle of the $1P$, $2P$, $1D$ and $2D$ states can also be calculated, which are $-35.0°$,$-36.4°$,$-42.8°$ and $-44.1°$, respectively. The mixing angles are close to the $B$ mesons $\theta_{1P} = 34.6$, $\theta_{2P} = -36.1°$, $\theta_{1D} = -39.6°$, $\theta_{2D} = -39.7°$ and $B_s$ mesons $\theta_{1P} = 34.9$, $\theta_{2P} = -36.1°$, $\theta_{1D} = -39.8°$, $\theta_{2D} = -39.8°$ [47].

The mass spectrum is listed in Table II. We predicted the masses of the $B_c(1^3S_0)$ and the $B_c(2^1S_0)$ are 6272 MeV and 6875 MeV, which are close to the experimental values 6274.47±0.32 MeV and 6871.2±1.0 MeV. Besides, we also show the theoretical results of the other exotic states. We expect more experimental information can be found to support our results.

For states below $BD$ threshold, the probabilities of each coupled channel can be estimated, The probabilities are listed in Table IV. As we see, all the states have coupled channel components. Especially when comparing with the $1S$, $2S$, $1P$, and the $1D$ states, the $2P$ states have larger non-$b\bar{c}$ components. For the $1S$-wave states, we predicted they have 96% $b\bar{c}$ components, that means the coupled channel components is just 4%. For other sates, we predicted the $b\bar{c}$ component probabilities for the $2S$, $1P$, $2P$, $1D$, and $2D$ states are about 86%, 90%, 67%, and 83%, respectively.

For states have large masses, the strong decay channels will be open, the strong decay width are shown in Table V. As we see, for the $1S$-wave, $2S$-wave, $1P$-wave, $2P$-wave and $1D$-wave states, their masses are two small to strong decay final states. So we just discuss the strong decay of the $B_c(2^3P_2)$, $3S$-wave and the $2D$-wave $B_c$ mesons here. For the $B_c(2^3P_2)$, it can just decay to $BD$ final states with decay width 3 MeV. For the $3S$-wave states, the $B_c(3^1S_0)$ can strong decay to $B^*D$ with predicted width 110 MeV, the $B_c(3^3S_1)$ can strong decay to $BD$ and $B^*D$ with predicted width 10 MeV and 59 MeV, the total width for the state is 69 MeV. For the $2D$-wave states, The total decay widths of the $B_c(2^3D_1)$, $B_c(2D)$, $B_c(2D')$ and $2^3D_3$ states are 60 MeV, 149 MeV, 65 MeV and 72 MeV. The $B_c(2^3D_1)$ dominantly decay into $BD^*$ and $B^*D^*$ with predicted widths 28 MeV and 22 MeV. The $B_c(2D)$ can mainly decay into $B^*D$ and $BD^*$ with predicted widths 92 MeV and 54 MeV. And the $B_c(2D')$ can mainly decay into $BD^*$and $B^*D^*$ with predicted widths 18 MeV and 44 MeV. The mixing angle of the two states is $-44.1°$. The $B_c(2^3D_3)$ have three dominant decay channels $BD$, $B^*D$ and $B^*D^*$, the decay width for the four channels are 27 MeV, 30 MeV and 11 MeV. These difference will be helpful to distinguish these states.

TABLE V: Hadronic decay widths (in MeV) of the beauty-charm mesons.

| State | $BD$ | $BD^*$ | $B^*D$ | $B^*D^*$ | $B_sD_s$ | $B_sD_s^*$ | $B_s^*D_s^{(*)}$ | Total |
|---|---|---|---|---|---|---|---|---|
| $3^1S_0$ | 0 | 0 | 110 | 0 | 0 | 0 | 0 | 110 |
| $3^3S_1$ | 10 | 0 | 59 | 0 | 0 | 0 | 0 | 69 |
| $2^3P_2$ | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| $2^3D_1$ | 2 | 28 | 1 | 22 | 7 | 0 | 0 | 60 |
| $2D$ | 0 | 92 | 54 | 3 | 0 | 0 | 0 | 149 |
| $2D'$ | 0 | 18 | 4 | 44 | 0 | 0 | 0 | 65 |
| $2^3D_3$ | 27 | 4 | 30 | 11 | 0 | 0 | 0 | 72 |

## IV. SUMMARY

We made a coupled channel calculation of the beauty-charm mesons, where the spectrum is coherently calculated. We also calculated the strong decay properties by using the $^3P_0$ model. The wave functions in our calcu-

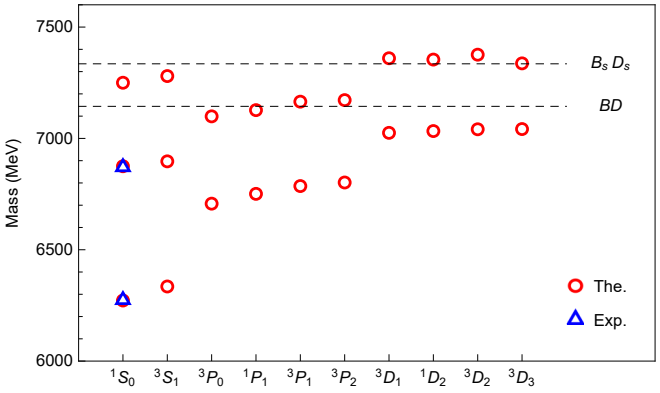



FIG. 1: The spectrum of the beauty-charm meson family. "Exp." denotes the current experimental values from the latest RPP [2] and our theoretical results are depicted as "The.". The dashed lines represent the threshold positions of $BD$ and $B_sD_s$, respectively

lations are obtained by solving the Hamiltonian of the potential model with Gaussian Expansion Method.

Our results support all of the $B_c$ states have coupled channel components. The components are different for various states. Generally, the coupled channel effects are smaller for bound states than exotic states. The $1S$ states are about $3 \sim 5\%$, the $1P$ states are abort $8 \sim 11\%$, The $1D$ states are about $15 \sim 18\%$, the $2S$ states are about $14 \sim 15\%$. And the four $2P$ states have larger couple channel components which is about $32 \sim 34\%$.

For the strong decays, the $3^1S_0$ mainly strong decay to $B^*D$ channel and the $3^3S_1$ mainly strong decay to $BD$ and $B^*D$, the $B_c(2^3P_2)$ mainly decay to $BD$. In addition, the $B_c(3S)$ and $B_c^*(3S)$ can also be detected in $B_c(3S) \to B_c(1S/2S) + \pi^+ + \pi^-$ and $B_c^*(3S) \to B_c(1S/2S)+\pi^+ +\pi^- +\gamma$ processes practically. For the four $D$-wave states, the $2D$ dominantly decay to $BD^*$ and $B^*D$ final states with total width 149 MeV. Comparing with the states, the $B_c(2^3D_1)$, $B_c(2D')$, and $B_c(2^3D_3)$ have smaller decay widths which are about 60 $\sim$ 72 MeV. For the electromagnetical and weak decays of beauty-charm mesons with polarization analysis, one can also refer to Refs. [57–61].

As we know, for the $B_c$ mesons, only a few states of have been observed experimentally, the theoretical studies will be useful to reveal their nature, and promote future experimental findings.

## V. ACKNOWLEDGEMENTS

This work is supported by NSFC under grant No. 12322503, No. 12047503, and No. 12075124, and by Natural Science Foundation of Jiangsu under Grant No. BK20211267.

[1] F. Abe et al. (CDF), Phys. Rev. Lett. **81**, 2432 (1998), hep-ex/9805034.
[2] R. L. Workman et al. (Particle Data Group), PTEP **2022**, 083C01 (2022).
[3] G. Aad et al. (ATLAS), Phys. Rev. Lett. **113**, 212004 (2014), 1407.1032.
[4] A. M. Sirunyan et al. (CMS), Phys. Rev. Lett. **122**, 132001 (2019), 1902.00571.
[5] R. Aaij et al. (LHCb), Phys. Rev. Lett. **122**, 232001 (2019), 1904.00081.
[6] N. Akbar, Phys. Atom. Nucl. **83**, 634 (2020), 1911.02078.
[7] Q. Li, M.-S. Liu, L.-S. Lu, Q.-F. Lü, L.-C. Gui, and X.-H. Zhong, Phys. Rev. D **99**, 096020 (2019), 1903.11927.
[8] I. Asghar, F. Akram, B. Masud, and M. A. Sultan, Phys. Rev. D **100**, 096002 (2019), 1910.02680.
[9] N. Akbar, F. Akram, B. Masud, and M. Atif Sultan, Eur. Phys. J. A **55**, 82 (2019), 1811.07552.
[10] A. P. Monteiro, M. Bhat, and K. B. Vijaya Kumar, Phys. Rev. D **95**, 054016 (2017), 1608.05782.
[11] T.-y. Li, L. Tang, Z.-y. Fang, C.-h. Wang, C.-q. Pang, and X. Liu, Phys. Rev. D **108**, 034019 (2023), 2204.14258.
[12] X.-J. Li, Y.-S. Li, F.-L. Wang, and X. Liu, Eur. Phys. J. C **83**, 1080 (2023), 2308.07206.
[13] Z.-b. Gao, Y.-y. Fan, H. Chen, and C.-q. Pang (2024), 2402.10629.
[14] S. Godfrey and N. Isgur, Phys. Rev. D **32**, 189 (1985).
[15] L. Chang, M. Chen, X.-q. Li, Y.-x. Liu, and K. Raya, Few Body Syst. **62**, 4 (2021), 1912.08339.
[16] C. A. Dominguez, K. Schilcher, and Y. L. Wu, Phys. Lett. B **298**, 190 (1993).

[17] S. S. Gershtein, V. V. Kiselev, A. K. Likhoded, and A. V. Tkabladze, Phys. Rev. D **51**, 3613 (1995), hep-ph/9406339.

[18] E. Bagan, H. G. Dosch, P. Gosdzinsky, S. Narison, and J. M. Richard, Z. Phys. C **64**, 57 (1994), hep-ph/9403208.

[19] W. Chen, T. G. Steele, and S.-L. Zhu, J. Phys. G **41**, 025003 (2014), 1306.3486.

[20] Z.-G. Wang, Eur. Phys. J. A **49**, 131 (2013), 1203.6252.

[21] J. Zeng, J. W. Van Orden, and W. Roberts, Phys. Rev. D **52**, 5229 (1995), hep-ph/9412269.

[22] L. Chang, M. Chen, and Y.-x. Liu, Phys. Rev. D **102**, 074010 (2020), 1904.00399.

[23] M. Chen, L. Chang, and Y.-x. Liu, Phys. Rev. D **101**, 056002 (2020), 2001.00161.

[24] C. T. H. Davies, K. Hornbostel, G. P. Lepage, A. J. Lidsey, J. Shigemitsu, and J. H. Sloan, Phys. Lett. B **382**, 131 (1996), hep-lat/9602020.

[25] G. M. de Divitiis, M. Guagnelli, R. Petronzio, N. Tantalo, and F. Palombi, Nucl. Phys. B **675**, 309 (2003), hep-lat/0305018.

[26] I. F. Allison, C. T. H. Davies, A. Gray, A. S. Kronfeld, P. B. Mackenzie, and J. N. Simone (HPQCD, FNAL Lattice, UKQCD), Nucl. Phys. B Proc. Suppl. **140**, 440 (2005), hep-lat/0409090.

[27] Y. Lu, M. N. Anwar, and B.-S. Zou, Phys. Rev. D **94**, 034021 (2016), 1606.06927.

[28] J.-F. Liu and G.-J. Ding, Eur. Phys. J. C **72**, 1981 (2012), 1105.0855.

[29] J. Ferretti, G. Galata, E. Santopinto, and A. Vassallo, Phys. Rev. C **86**, 015204 (2012).

[30] Y. S. Kalashnikova, Phys. Rev. D **72**, 034010 (2005), hep-ph/0506270.

[31] B.-Q. Li, C. Meng, and K.-T. Chao, Phys. Rev. D **80**, 014012 (2009), 0904.4068.

[32] J. Ferretti, G. Galatà, and E. Santopinto, Phys. Rev. C **88**, 015207 (2013), 1302.6857.

[33] Q. Deng, R.-H. Ni, Q. Li, and X.-H. Zhong (2023), 2312.10296.

[34] J. Ferretti and E. Santopinto, Phys. Rev. D **90**, 094022 (2014), 1306.2874.

[35] E. van Beveren and G. Rupp, Eur. Phys. J. C **32**, 493 (2004), hep-ph/0306051.

[36] E. van Beveren and G. Rupp, Phys. Rev. Lett. **91**, 012003 (2003), hep-ph/0305035.

[37] S. Coito, G. Rupp, and E. van Beveren, Phys. Rev. D **84**, 094020 (2011), 1106.2760.

[38] D. S. Hwang and D.-W. Kim, Phys. Lett. B **601**, 137 (2004), hep-ph/0408154.

[39] Y. A. Simonov and J. A. Tjon, Phys. Rev. D **70**, 114013 (2004), hep-ph/0409361.

[40] I. W. Lee, T. Lee, D. P. Min, and B.-Y. Park, Eur. Phys. J. C **49**, 737 (2007), hep-ph/0412210.

[41] F.-K. Guo, S. Krewald, and U.-G. Meissner, Phys. Lett. B **665**, 157 (2008), 0712.2953.

[42] Z.-Y. Zhou and Z. Xiao, Phys. Rev. D **84**, 034023 (2011), 1105.6025.

[43] A. M. Badalian, Y. A. Simonov, and M. A. Trusov, Phys. Rev. D **77**, 074017 (2008), 0712.3943.

[44] Y.-B. Dai, X.-Q. Li, S.-L. Zhu, and Y.-B. Zuo, Eur. Phys. J. C **55**, 249 (2008), hep-ph/0610327.

[45] W. Hao, Y. Lu, and B.-S. Zou, Phys. Rev. D **106**, 074014 (2022), 2208.10915.

[46] O. Lakhina and E. S. Swanson, Phys. Lett. B **650**, 159 (2007), hep-ph/0608011.

[47] Q.-F. Lü, T.-T. Pan, Y.-Y. Wang, E. Wang, and D.-M. Li, Phys. Rev. D **94**, 074012 (2016), 1607.02812.

[48] D.-M. Li, P.-F. Ji, and B. Ma, Eur. Phys. J. C **71**, 1582 (2011), 1011.1548.

[49] S. Godfrey and R. Kokoski, Phys. Rev. D **43**, 1679 (1991).

[50] L. Micu, Nucl. Phys. B **10**, 521 (1969).

[51] A. Le Yaouanc, L. Oliver, O. Pene, and J. C. Raynal, Phys. Rev. D **8**, 2223 (1973).

[52] A. Le Yaouanc, L. Oliver, O. Pene, and J. C. Raynal, Phys. Rev. D **9**, 1415 (1974).

[53] B. Silvestre-Brac and C. Gignoux, Phys. Rev. D **43**, 3699 (1991).

[54] P. Geiger and N. Isgur, Phys. Rev. D **44**, 799 (1991).

[55] P. Geiger and N. Isgur, Phys. Rev. Lett. **67**, 1066 (1991).

[56] P. Geiger and N. Isgur, Phys. Rev. D **55**, 299 (1997), hep-ph/9610445.

[57] Y. Geng, M. Cao, and R. Zhu (2023), 2310.03425.

[58] W. Tao, R. Zhu, and Z.-J. Xiao, Phys. Rev. D **106**, 114037 (2022), 2209.15521.

[59] W. Wang and R. Zhu, Int. J. Mod. Phys. A **34**, 1950195 (2019), 1808.10830.

[60] R. Zhu, Nucl. Phys. B **931**, 359 (2018), 1710.07011.

[61] C.-F. Qiao, P. Sun, D. Yang, and R.-L. Zhu, Phys. Rev. D **89**, 034008 (2014), 1209.5859.

arXiv:2402.18870v1 [hep-ph] 29 Feb 2024

# B-mesons as essential probes of hot QCD matter

Vinod Chandra[1*] and Santosh K. Das[2]

[1*]Department of Physics, Indian Institute of Technology Gandhinagar, Gandhinagar, 382355, Gujarat, India.
[2]School of Physical Sciences, Indian Institute of Technology Goa, Ponda, Goa, 403401, Goa, India.

*Corresponding author(s). E-mail(s): vchandra@iitgn.ac.in;
Contributing authors: santosh@iitgoa.ac.in;

**Abstract**

This article elucidates the pivotal role of b-mesons and bottomonium states in exploring the existence and properties of hot QCD matter (commonly known as quark-gluon-plasma (QGP) produced within the crucible heavy-ion collision experiments). Owing to the complex and confounding nature of strong interaction force the direct detection of probing the hot QCD matter is not feasible. In light of this, investigating the dynamics of b-quarks and anti-quarks within the hot QCD medium emerges as an invaluable indirect probe. The impact of b-quarks and the mesons spans a spectrum of interesting domains regarding the physics of QCD at finite temperature, encompassing the QCD phase transition, color screening, quarkonia dissociation, heavy quark energy loss and collective flow, anisotropic aspects, and strongly coupled nature of hot QCD medium. These aspects underscore the indispensable nature of B-mesons in the quest to create and explore the complex nature of strong interaction force through the QGP/hot QCD matter. In this context, we mainly focus on works related to transport studies of b-mesons in hot QCD medium, lattice QCD, and effective field theory studies on bottomonium states, and finally, open quantum system frameworks to quarkonia to explore the properties of hot QCD medium in relativistic heavy-ion collision experiments.

**Keywords:** B-mesons, heavy-quark dynamics, hot QCD medium, lattice QCD

# 1 Introduction

The study of B-mesons in heavy-ion collisions is a critical component of our quest to understand the fundamental properties of nuclear matter under extreme temperatures and energy densities. Heavy quarks Beraudo et al. (2018); Cao et al. (2019); Aarts et al. (2017), such as charm (c) and bottom (b), are essential for investigating and understanding the properties of hot QCD matter in relativistic heavy-ion collisions (RHIC), commonly termed as quark-gluon plasma (QGP)Gyulassy and McLerran (2005); Shuryak (2017). Due to their larger masses, the thermalization time of heavy quarks is greater than that of the light quarks and gluons that constitute the bulk of the hot QCD matter. Thus, they traverse the hot QCD medium as non-equilibrium degrees of freedom and sense the entire spacetime evolution of the medium. Therefore, their study in heavy-ion collisions is a critical component of our quest to understand the fundamental properties of nuclear matter under extreme temperatures and energy densities, addressing the grand quest of understanding the complex nature of the strong interaction force.

The bound states of charm and bottom quarks, known as quarkonia, are integral to the exploration of the properties of the QGP medium Djordjevic et al. (2005). Quarkonia exhibit distinct behavior in the QGP, undergoing phenomena such as sequential thermal broadening of various states and suppression, which are crucial for understanding the strongly coupled nature of the hot QCD medium. There have been extensive investigations into the in-medium properties of charmonium and bottomonium states using theoretical frameworks such as lattice QCD, incorporating the non-relativistic QCD (NRQCD) approach. By employing lattice QCD techniques, the fate of quarkonia under extreme conditions of the QGP has been explored Ratti (2018), providing insights into their dynamics and interactions. One key aspect of studying quarkonia in the QGP is their suppression. Recent studies have been conducted regarding bottomonium suppression within an open quantum framework, where the QGP medium effectively acts as a heat bath, influencing the behavior and stability of bottomonium states. The most crucial observable in these studies has been the nuclear suppression factor, $R_{AA}$, for bottomonium, which turned out to be consistent with data from RHIC (Au-Au collisions) and LHC (Pb-Pb Collisions). For recent reviews on the usage of open quantum framework to quarkonia as a probe of the QGP, we refer the reader to Akamatsu (2022); Yao (2021) and for the phenomenological aspects of quarkonia production, we refer to Lansberg (2020).

In this article, we primarily concentrate on the bottom quarks and B-mesons and their pivotal role in investigating the bulk and transport properties of hot QCD matter. They provide invaluable information regarding the properties and behavior of the QGP, the phase transition from hadronic matter to the QGP, and the dynamics of heavy quarks in extreme conditions. We will now proceed into a detailed discussions of key methodologies for studying b-mesons and bottomonia within the context of a hot QCD medium, as outlined in Sections 2-4. The conclusive insights and future prospects will be presented in Section 5.

## 2 Dynamics B-mesons and b-quarks in hot QCD medium

Bottom quarks are considered as a noble probe of the hot QCD matter produced in high-energy heavy-ion collisions. As the relaxation time for the bottom quarks of mass $M$ at a temperature $T$ are larger than the corresponding quantities for light partons by a factor of $M/T(>1)$ Moore and Teaney (2005) *i.e.* the light quarks and the gluons get thermalized faster than the bottom quarks. Hence, the propagation of bottom quarks within a GQP can be conceptualized as the interplay between equilibrium and non-equilibrium degrees of freedom. The Fokker-Planck (FP) equation offers a suitable framework for describing such phenomena. To derive the Fokker-Planck equation, let us begin with the Boltzmann transport equation.

The Boltzmann transport equation that describes a non-equilibrium statistical system is given by:

$$\left[\frac{\partial}{\partial t}+\frac{\mathbf{p}}{E}.\nabla_{\mathbf{x}}+\mathbf{F}.\nabla_{\mathbf{p}}\right]f(x,p,t)=\left[\frac{\partial f}{\partial t}\right]_{col} \tag{1}$$

where $p$ and $E$ denote momentum and energy, $\nabla_{\mathbf{x}}$ ($\nabla_{\mathbf{p}}$) are spatial (momentum space) gradient and $f(x,p,t)$ is the phase space distribution (in the present case $f$ stands for bottom quark distribution). The assumption of plasma uniformity and the absence of external forces leads to

$$\frac{\partial f}{\partial t}=\left[\frac{\partial f}{\partial t}\right]_{col} \tag{2}$$

If we define $\omega(p,k)$, the rate of collisions which change the momentum of the bottom quark from $p$ to $p-k$, then we have Svetitsky (1988)

$$\left[\frac{\partial f}{\partial t}\right]_{col}=\int d^3k\left[\omega(p+k,k)f(p+k)-\omega(p,k)f(p)\right] \tag{3}$$

The first term in the integrand represents a gain of probability through collisions that knock the bottom quark into the volume element of momentum space at $p$, and the second the term denotes loss out of that volume element. $\omega$ is the sum of contributions arising from the scattering of gluons, light quarks, and antiquarks.

If we expand $\omega(p+k,k)f(p+k)$ around $k$,

$$\omega(p+k,k)f(p+k)\approx\omega(p,k)f(p)+k\frac{\partial}{\partial p}(\omega p)+\frac{1}{2}k_ik_j\frac{\partial^2}{\partial p_i\partial_j}(\omega p) \tag{4}$$

and substitute in Eq. 3, we get:

$$\left[\frac{\partial f}{\partial t}\right]_{col}=\frac{\partial}{\partial p_i}\left[A_i(p)f+\frac{\partial}{\partial p_i}[B_{ij}(p)f]\right] \tag{5}$$

where we have defined the kernels

$$A_i = \int d^3k\omega(p,k)k_i$$
$$B_{ij} = \int d^3k\omega(p,k)k_ik_j. \tag{6}$$

for $\mid \mathbf{p} \mid \to \mathbf{0}$, $A_i \to \gamma p_i$ and $B_{ij} \to D\delta_{ij}$ where $\gamma$ and $D$ stand for drag and diffusion coefficients respectively. The function $\omega(p,k)$ is given by

$$\omega(p,k) = g\int \frac{d^3q}{(2\pi)^3} f'(q)v\sigma_{p,q\to p-k,q+k} \tag{7}$$

where $f'$ is the phase space distribution, representing light quarks, anti-quarks, and gluons in the current context. $v$ represents the relative velocity between the two collision partners, $\sigma$ denotes the cross-section and $g$ is the statistical degeneracy. The coefficients in the first two terms of the expansion in Equation 5 exhibit comparable magnitudes due to greater cancellation in the averaging of $k_i$ compared to the averaging of the quadratic term $k_ik_j$. The higher powers of $k_i$'s are assumed to be relatively smaller.

With these approximations, the Boltzmann equation simplifies into a non-linear integro-differential equation, commonly referred to as the Landau kinetic equation:

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial p_i}\left[A_i(p)f + \frac{\partial}{\partial p_i}[B_{ij}(p)f]\right] \tag{8}$$

The nonlinearity arises from the presence of $f'$ in $A_i$ and $B_{ij}$ through $w(p,k)$. It emerges from the simple fact that we are investigating collision process which involves two particles - it should, therefore, depend on the states of the two participating particles and hence on the product of the two distribution functions. Significant simplification can be attained by substituting the distribution functions of one of the collision partners by their equilibrium distribution either Fermi-Dirac or Bose-Einstein (depending on the particle type) in the expressions of $A_i$ and $B_{ij}$. Then Eq. 8 reduces to a linear partial differential equation - usually referred to as the Fokker-Planck equation. This equation describes the interaction of a particle that is out of thermal equilibrium with the particles in a thermal bath. As mentioned earlier, the quantities $A_i$ and $B_{ij}$ are related to the usual drag and diffusion coefficients, and we denote them by $\gamma$ and $D_{ij}$ respectively (*i.e.* these quantities can be obtained from the expressions for $A_i$ and $B_{ij}$ by replacing the distribution functions by their thermal counterparts):

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial p_i}\left[\gamma_i(p)f + \frac{\partial}{\partial p_i}[D_{ij}(p)f]\right] \tag{9}$$

One needs the drag and diffusion coefficient as inputs to solve the Fokker-Planck equation along with the initial distribution. The drag coefficient can be defined as the thermal average of the momentum transfer weighted by the square of the invariant

transition amplitude. For the elastic two-body collisional process $HQ(P) + l(Q) \rightarrow HQ(P^{'}) + l(Q^{'})$, where $l$ represents constituent particles in the thermal medium (quarks, antiquarks, and gluons) and $P, Q$ are the four-momentum of the bottom quark and constituent particle before the collision, the heavy quark drag and momentum diffusion can be described as,

$$\begin{aligned} A_i =& \frac{1}{2E_p}\int \frac{d^3\mathbf{q}}{(2\pi)^3 2E_q}\int \frac{d^3\mathbf{q}'}{(2\pi)^3 2E_{q'}}\int \frac{d^3\mathbf{p}'}{(2\pi)^3 2E_{p'}}\frac{1}{\gamma_{HQ}} \\ & \times \sum |\mathcal{M}_{2\rightarrow 2}|^2 (2\pi)^4 \delta^4(P+Q-P'-Q') f_k(\mathbf{q}) \\ & \times \Big(1 + a_k f_k(\mathbf{q}')\Big)\Big[(\mathbf{p}-\mathbf{p}')_i\Big] \\ & = \langle\langle (\mathbf{p}-\mathbf{p}')_i \rangle\rangle, \end{aligned} \tag{10}$$

and

$$\begin{aligned} B_{ij} =& \frac{1}{2E_p}\int \frac{d^3\mathbf{q}}{(2\pi)^3 2E_q}\int \frac{d^3\mathbf{q}'}{(2\pi)^3 2E_{q'}}\int \frac{d^3\mathbf{p}'}{(2\pi)^3 2E_{p'}}\frac{1}{\gamma_{HQ}} \\ & \times \sum |\mathcal{M}_{2\rightarrow 2}|^2 (2\pi)^4 \delta^4(P+Q-P'-Q') f_k(\mathbf{q}) \\ & \times \Big(1 + a_k f_k(\mathbf{q}')\Big)\frac{1}{2}\Big[(\mathbf{p}-\mathbf{p}')_i(\mathbf{p}-\mathbf{p}')_j\Big] \\ & = \frac{1}{2}\langle\langle (\mathbf{p}-\mathbf{p}')_i(\mathbf{p}-\mathbf{p}')_j \rangle\rangle, \end{aligned} \tag{11}$$

respectively. It is important to note that the delta function guarantees energy-momentum conservation in the system. and $f_k$ is the equilibrium phase space distribution for the light quarks, antiquarks, and gluons. Here, $\gamma_{HQ}$ denotes the statistical degeneracy factor of the heavy quark, and $|\mathcal{M}_{2\rightarrow 2}|$ represents the matrix element for the two-body elastic collisions of the bottom with light quarks, antiquarks, and gluons Svetitsky (1988). It is crucial to highlight that the drag coefficient for the bottom quark represents the thermal average of the momentum transfer, while the momentum diffusion captures the square of the momentum transfer resulting from the interaction. As $A_i$ depends on the momentum, we have the following decomposition of the bottom quark drag,

$$A_i = p_i A(p^2), \qquad\qquad A = \langle\langle 1 \rangle\rangle - \frac{\langle\langle \mathbf{p}.\mathbf{p}' \rangle\rangle}{p^2}, \tag{12}$$

where $p^2 = |\mathbf{p}|^2$ and $A$ is the drag coefficient of the heavy quark. Similarly, the momentum diffusion $B_{ij}$ can be decomposed in terms of longitudinal and transverse components as follows,

$$B_{i,j} = \left(\delta_{ij} - \frac{p_i p_j}{p^2}\right) B_0(p^2) + \frac{p_i p_j}{p^2} B_1(p^2), \tag{13}$$

with the transverse and longitudinal diffusion coefficients, they respectively assume the following forms:

$$B_0 = \frac{1}{4}\left[\langle\langle p'^2\rangle\rangle - \frac{\langle\langle(\mathbf{p}'.\mathbf{p})^2\rangle\rangle}{p^2}\right], \tag{14}$$

$$B_1 = \frac{1}{2}\left[\frac{\langle\langle(\mathbf{p}'.\mathbf{p})^2\rangle\rangle}{p^2} - 2\langle\langle(\mathbf{p}'.\mathbf{p})\rangle\rangle + p^2\langle\langle 1\rangle\rangle\right]. \tag{15}$$

The standard approach to HQ dynamics in the QGP is to follow their evolution using a Fokker-Plank equation solved stochastically by the Langevin equations. The relativistic Langevin equations of motion governing the evolution of the position and momentum of the heavy quarks can be expressed in the following form: Das et al. (2015):

$$\begin{aligned} dx_i &= \frac{p_i}{E}dt, \\ dp_i &= -Ap_i dt + C_{ij}\rho_j\sqrt{dt} \end{aligned} \tag{16}$$

where $dx_i$ and $dp_i$ are the changes of the coordinate and momentum in each time step $dt$. $A$ and $C_{ij}$ are the drag coefficient and the covariance matrix in terms of independent Gaussian-normal distributed random variables $\rho$,$P(\rho) = (2\pi)^{-3/2}e^{-\rho^2/2}$, which obey the relations $< \rho_i\rho_j >= \delta_{ij}$ and $< \rho_i >= 0$, respectively. The covariance matrix is related to the diffusion tensor,

$$C_{ij} = \sqrt{2B_0}P_{ij}^{\perp} + \sqrt{2B_1}P_{ij}^{\parallel}, \tag{17}$$

where $P_{ij}^{\perp} = \delta_{ij} - p_ip_j/p^2$ and $P_{ij}^{\parallel} = p_ip_j/p^2$ represent the transverse and longitudinal projector operators, respectively. With the assumption, $B_0 = B_1 = D$, Eq (17) becomes $C_{ij} = \sqrt{2D(p)}\delta_{ij}$. This assumption is strictly valid only for $p \to 0$; however, it is commonly employed at finite $p$ when applied to the dynamics of bottom (also charm quark) quarks in hot QCD matter.

The Fokker-Planck or Langevin equation results can be utilized to calculate experimental observables, such as the nuclear suppression factor, $R_{AA}$, and elliptic flow, $v_2$. The nuclear suppression factor, $R_{AA}$, can be evaluated using the initial charm and bottom quark distributions at initial time $t = \tau_i$ and final time $t = \tau_f$ at the freeze-out temperature as $R_{AA}(p) = \frac{f(p,\tau_f)}{f(p,\tau_i)}$. Along with $R_{AA}$, the anisotropic momentum distribution induced by the spatial anisotropy of the bulk medium can be calculated and defined as

$$v_2 = \left\langle \frac{p_x^2 - p_y^2}{p_x^2 + p_y^2} \right\rangle, \tag{18}$$

which measures the momentum space anisotropy.

## 2.1 Boltzmann vs Fokker-Planck dynamics

The propagation of heavy quarks, mainly charm and bottom, in QGP has been quite often treated within the framework of the Fokker-Planck equation. This is primarily because their motion can be likened to Brownian motion, attributed to their perturbative interaction and large mass. This characteristic typically results in collisions that are sufficiently forward-peaked and/or entail small momentum transfers. Under these constraints, it is recognized that the Boltzmann transport equation also simplifies to Fokker-Planck dynamics. This reduction represents a substantial simplification of in-medium dynamics. This approach has been extensively utilized in the literature for the computation of experimentally observed nuclear suppression factor ($R_{AA}$) and elliptic flow ($v_2$).

Indeed, the initial motivation for adopting a Fokker-Planck approach was primarily linked to the assumption that heavy quarks, both charm and bottom quarks, experience relatively small momentum transfers. However, the experimentally observed suppression factors ($R_{AA}$) and elliptic flows ($v_2$) for light and charmed hadrons in the QGP indicate that the momentum transfer for heavy quarks may not be sufficiently small. Hence, investigating the effects of the approximations inherent in the Fokker-Planck equation through a direct comparison with the full collisional integral within the Boltzmann transport equation framework would be valuable. This analysis could help determine whether heavy quarks indeed exhibit Brownian motion in the QGP.

The momentum evolution of heavy quarks is explored using both the Langevin and Boltzmann approaches, with the initial distribution of charm and bottom quarks assumed to be a delta distribution at p=10 GeV. The Debye mass is taken to be $m_D = gT$ to regularize the divergence arising from the t-channel process. The heavy quark drag and diffusion coefficients with the gluons in the bulk have been computed within the framework of perturbative Quantum Chromodynamics (pQCD) to serve as input in the Langevin equation. Numerical solutions for the Boltzmann equation are obtained by discretizing space into a three-dimensional lattice and employing the test particle method to sample distribution functions. The collision integral is addressed through a stochastic implementation of the collision probability, where the heavy quark-gluon cross-section is treated as an input, computed within the framework of pQCD. For further details, please refer to reference Das et al. (2014).

The objective is to compare the time evolution of charm and bottom quarks, originating from the same initial momentum distribution within a bulk at a temperature of T=0.4 GeV, using both the Langevin and Boltzmann equations. The momentum evolution of the charm quarks is displayed in Fig.1 (left panel) obtained within the Langevin dynamics. Fig.2 (left panel) obtained within the Langevin dynamics displays the bottom quark momentum evolution. The observation reveals that both charm and bottom quarks exhibit Gaussian distributions, as anticipated by the construction of the model. This present calculation indicates that the equilibrium distribution can be achieved at the end of the evolution by implementing the Fluctuation-dissipation theorem (FDT). As is well-known, the Langevin dynamics consists of a shift of the average momenta accompanied by fluctuations around this average, allowing the heavy

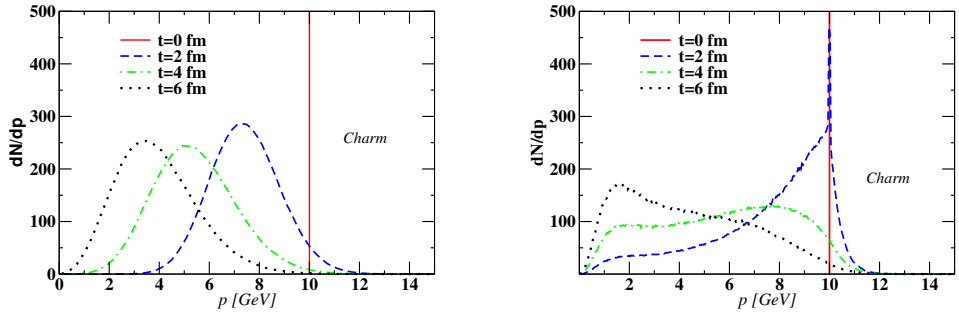



**Fig. 1** Left panel: Charm quark momentum evolution obtained within Langevin dynamical considering the initial momentum distribution of the charm quarks as a delta distribution at p=10 GeV. Right panel: Charm quark momentum evolution obtained within the Boltzmann equation considering the initial momentum distribution of the charm quarks as a delta distribution at p=10 GeV. Figures adapted from Ref. Das et al. (2014).

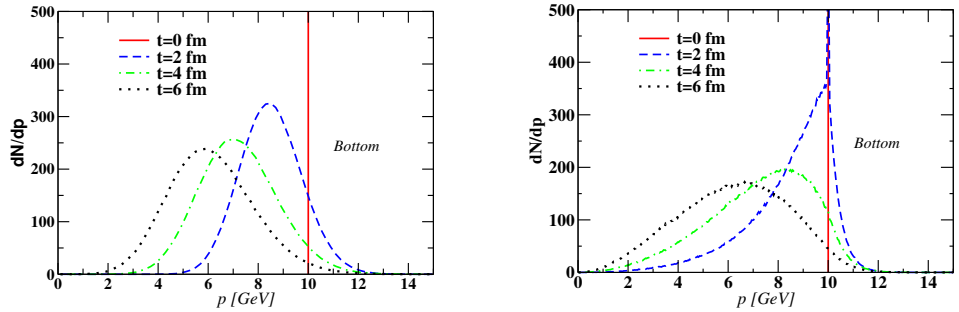



**Fig. 2** Left panel: Bottom quark momentum evolution obtained within Langevin equation considering the initial momentum distribution of the bottom quarks as a delta distribution at p=10 GeV. Right panel: Bottom quark momentum evolution obtained within the Boltzmann equation considering the initial momentum distribution of the bottom quarks as a delta distribution at p=10 GeV. Figures adapted from Ref. Das et al. (2014).

quark the possibility of gaining energy. This is evident from the tail of the momentum distribution that overshoots the initial momentum p = 10GeV at t = 2 fm/c, the black solid line in Fig.1 (left panel) and Fig.2 (left panel).

In Fig. 1 (right panel), the momentum distribution for charm quarks within the Boltzmann equation is depicted. There is a distinctly different evolution of the particle momentum, characterized by a non-Gaussian shape. By t=2vfm/c, the distribution displays a broader spread in momentum, with a notable contribution from processes where charm quarks can gain energy. Additionally, there is a long tail at low momenta, indicating a probability of losing a significant amount of energy. Overall, the global shape deviates significantly from a Gaussian form. This observation suggests that the evolution of the charm quark at a temperature T=0.4 GeV is not Brownian in nature. In Fig.2 (right panel) we present the momentum distribution for the bottom quark obtained within the Boltzmann equation. For the bottom quarks, shown in Fig.2 (right panel), the momentum evolution gives a much better agreement between the Boltzmann and the Langevin evolution because $M_b/T \simeq 10$. This suggests that the

Langevin approach serves as a good approximation for the bottom quark, while for the charm quark, there can be significant deviations Das et al. (2014).

To compute the $R_{AA}$ and $v_2$ of B-meson, bottom quarks are distributed in momentum space according to the Fixed Order + Next-to-Leading Log (FONLL) Cacciari et al. (2005) calculation, which describes the B-mesons spectra in proton-proton collisions. The solution of the Langevin equation can be convoluted with the fragmentation functions of the bottom quarks at the quark-hadron transition temperature $T_c$ to obtain the momentum distribution of the B mesons. One of the prime goals of all phenomenological studies of open heavy-flavor observables is to extract the heavy quark spatial diffusion coefficient, $D_x$, which quantifies the interaction of heavy quarks with the bulk medium. $D_x$ is also directly related to the heavy quark thermalization time and can be evaluated using lattice QCD (lQCD). Transport models are quite successful in describing the experimentally measured observables, the nuclear modification factor, $R_{AA}$, and elliptic flow, $v_2$, of the B-meson. Drag and momentum diffusion coefficients are inputs in the transport models (cross sections are the inputs in the Boltzmann equation), which contain the microscopic details mentioning how the heavy quarks interact with the hot QGP medium. Bottom quark drag and diffusion coefficients can be computed within the framework of pQCD. However, pQCD does not work to describe the measured heavy quark $R_{AA}$ and $v_2$ at RHIC and LHC Das et al. (2015). At low $p_T$ and low T, the non-perturbative effect plays a crucial role in simultaneously describing both the $R_{AA}$ and $v_2$ of heavy quark. Non-perturbative effects can be taken into account through the quasi-particle model (QPM) Das et al. (2015). Then, they compare their results with the experimental data. The value of the transport coefficients with which they can describe the experimental data, they relate it with the spatial diffusion coefficient, $D_x$. The spatial diffusion coefficient, $D_x$, can be calculated in the static limit ($p \to 0$) from the drag coefficient $D_x = T/MF$, where $T$ is the temperature of the thermal bath, $M$ is the mass of the heavy quark and $F$ is the drag coefficient. The standard quantification of the space diffusion coefficient is done in terms of a dimensionless quantity $2\pi TD_x$ He et al. (2013) which is independent of mass. One can estimate the charm quark thermalization time as $\tau_{th} \equiv \Gamma^{-1}(p \to 0)$ Scardina et al. (2017):

$$\tau_{th} = \frac{MD_x}{T} = \frac{M}{2\pi T^2}(2\pi TD_x) \cong 1.8\,\frac{2\pi TD_x}{(T/T_c)^2}\ \text{fm/c}, \tag{19}$$

where $T_c$ is the quark-hadron transition temperature. The $D_x$ can also computed from the diffusion coefficient ($\Gamma$) through the fluctuation-dissipation theorem (FDT), $\Gamma = MFT$.

In Fig.3, the variation of $D_x$ is shown as a function of temperature for b quarks (high temperature), $\bar{B}$ and $\Lambda_b$ (low temperature). We notice that the $\bar{B}$ and $\Lambda_b$ diffusion coefficient also support a continuous evolution with a minimum around $T_c$ like the D meson case. This indicates the strength of the interaction is the maximum near the quark-hadron transition temperature. The diffusion coefficient of the $\bar{B}$ and $\Lambda_b$ in the hadronic matter is shown in Fig.3 computed within the effective field theory framework. For further details, please refer to reference Tolos et al. (2016).

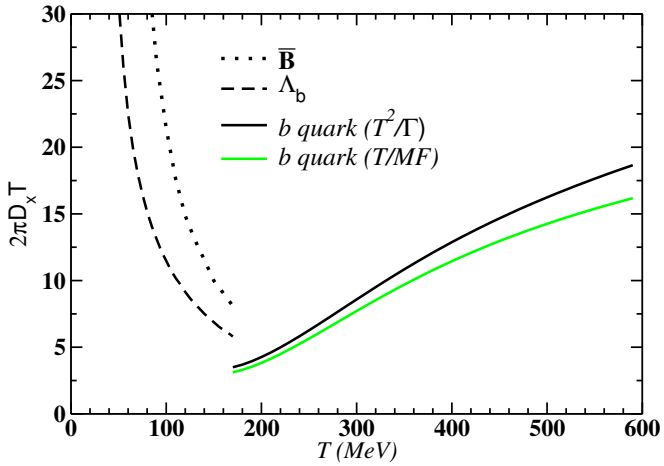



**Fig. 3** : Spatial diffusion coefficients $D_x$ (in units of the thermal wavelength, $1/(2\pi \mathrm{T})$) as a function of temperature (T) for b quark, $\bar{B}$ and $\Lambda_b$. Figure adapted from Ref. Das et al. (2016).

# 3 B-mesons at finite temperature in the lattice

The in-medium characteristics of bottomonium have been explored through lattice QCD investigations, as documented in previous studies Aarts et al. (2011, 2013, 2014); Kim et al. (2015, 2018). These investigations primarily concentrate on discerning the modifications occurring in the ground states of S- and P-wave quarkonia within the medium. However, a notable challenge in these lattice studies arises from the use of point meson operators, which exhibit limited overlap with quarkonia wavefunctions, particularly those associated with excited states. Consequently, the correlators tend to be predominantly influenced by the vacuum continuum portions of the spectral function, posing a significant obstacle in accurately computing in-medium bottomonium properties Burnier et al. (2015); Mocsy and Petreczky (2008).

Efforts have been made to address this limitation by employing extended meson operators, known for their improved compatibility with S- and P-state bottomonia Larsen et al. (2019). However, further refinement is needed to delve into the in-medium properties of excited bottomonium states, such as 3S and 2P-states. A step in this direction has been taken by introducing extended meson operators within the lattice NRQCD framework Larsen et al. (2020), revealing evidence of sequential in-medium modifications aligned with the sizes of the bottomonium states.

The investigation into the thermal broadening of quarkonia states within a lattice framework remains an active research area. More nuanced and refined results in this domain are anticipated, especially considering the temperature regimes accessible at RHIC and LHC energies. This ongoing work is crucial for advancing our understanding of the behavior of bottomonium in a hot and dense medium.

# 4 Bottomonium suppression within open quantum system framework

Quankonia suppression was proposed as a potential indicator of the QGP format long ago by Matsui and Satz Matsui and Satz (1986). Since then quakonia suppression has been an integral part of heavy-ion physics investigations in experiments at RHIC and LHC. The idea of Matsui was primarily based on the screening of chromoelectric fields at a distance inverse of the Debye mass ($m_D \sim g(T)T$, where $g$ is QCD coupling constant and $T$ is the temperature of the system. In this picture, whenever the binding energy of a given quakonia is overcome by the temperature the said quakonia dissociates. This mechanism has been revised by the first principle QCD computations at finite temperature that computes the heavy quark-antiquark static potential in the QGP medium Laine et al. (2007). The potential turned out to the complex. The imaginary part leads to the thermal width and it is this part that dominates the in-medium heavy quarkonia suppression rather than the screening effects. Employing this potential in non-relativistic QCD effective theory Brambilla et al. (2013, 2018) framework, it is shown that the in-medium quantum evolution of heavy quarkonium depends on its momentum diffusion coefficient and its dispersive counterpart. Both the transport coefficients could be defined in terms of nonperturbative correlators of chromoelectric fields.

On the other hand, Non-relativistic Effective Field Theories (NREFTs) serve the purpose of appropriately defining the heavy quark potential to enable a way for the systematic calculation of quarkonia properties in the hot QCD medium. This framework relies on the separation of various scales that are characteristic of NR bond states (both at zero as well as finite temperatures). Heavy quark mass, $m$, the inverse of the Bohr radius, $1/r_0 \sim mv$, and the binding energy, $e \sim mv^2$ ($v << 1$, being the heavy quark velocity) are the energy scales at zero temperature. At finite temperature, more scales enter in the picture, for instance, the temperature T and, at weak coupling, the Debye mass $m_D$. Both the cases Nevertheless, at leading order in $v$ the equation of motion is still a Schrödinger equation, which describes the real-time evolution of the $Q\bar{Q}$ pair in the medium. The effective field theory (EFT) framework that is obtained by integrating out degrees of freedom associated with the scale $m$ is Non-Relativistic QCD (NRQCD) Caswell and Lepage (1986); Bodwin et al. (1995). The EFT that integrates out gluons with momentum or energy scaling like the inverse of the Bohr radius is the potential NRQCD (pNRQCD) Pineda and Soto (1998); Brambilla et al. (2000, 2005). At the leading order in heavy quark velocity, $v$, the equation of motion of pNRQCD is the Schro"dinger equation for an NR bound state. Note that pNRQCD provides an unambiguous field theoretical definition of the potential (encoding the contributions coming from modes with energy and momentum above the scale of the binding energy). This framework includes systematically higher order corrections to leading of Schro'dinger equation. The first corrections in the weak coupling regime are carried out by chromoelectric dipole terms.

The essential processes regarding quarkoina physics in the QGP, viz., Quarkonium scattering in the medium, its dissociation into an unbound color octet $Q\bar{Q}$ pair, and the inverse processes of $Q\bar{Q}$ pair generation need a framework that describes the

quarkonium non-equilibrium evolution in the said medium. This calls for the open quantum system framework (OQS) where the background QGP medium plays the role of the environment with temperature scale T. The OQS framework in this context was proposed in a seminal work by Akamatsu Akamatsu (2015).

In the context of bottomonium, the suppression has been investigated in an OQS framework in Brambilla et al. (2021). In this work, the Lindblad equation has been derived in the framework of pNRQCD accounting for the quantum and non-Abelian nature of the quarkonia system. The bottomonium nuclear suppression factor computed within the framework taking lattice QCD inputs was shown to agree well with the LHC data for 5.02 TeV Pb-Pb collisions. In a very recent work Strickland and Thapa (2023), NLO correction has been considered to above mentioned OQS framework within pNRQCD, and predictions on bottomonium suppression were found to be in good agreement with LHC data but some disagreements were found with RHIC data.

These investigations have gained momentum and we hope for a more refined understanding regarding bottomonium production and suppression in the hot QCD medium shortly. It is to be noted that the usage of the OQS framework in high-energy physics extends beyond the confines of the heavy-quarkonia transport. The OQS framework has been employed in diverse areas such as dark matter formation Binder et al. (2020), deeply inelastic reactions Braaten et al. (2016), and studies related to inflation Boyanovsky (2015, 2016) etc., demonstrating its applicability across a spectrum of scientific investigations.

# 5 Conclusions and outlook

The study of the properties of hot QCD matter is a field of contemporary interest and the heavy quarks (HQ) namely, charm and bottom quarks play crucial roles in this endeavor. Heavy flavor particles are not anticipated to undergo complete thermalization. As non-equilibrium probes, they retain a memory of their interaction history, providing a gauge of their interaction strength with the medium. Unlike charm quarks, bottom quarks exhibit behavior akin to Brownian motion, and the Langevin approach provides a suitable approximation for their dynamics. However, the momentum evolution of charm quarks is not Brownian in nature at the temperatures anticipated in experiments conducted at the Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC). Additionally, it has been suggested that the heavy quark diffusion with an infinite-mass limit may not be a suitable approximation for the charm quark Pandey et al. (2023). However, for the dynamics of the bottom quark in QCD matter, the infinite-mass limit remains a reasonable approximation. This emphasizes the significance of bottom quarks as a probe of the hot QCD matter. They possess the potential to establish a connection between the phenomenology constrained by experimental data and lQCD, offering insights into the transport properties of the hot QCD matter.

The presence of the hadronic phase is unavoidable in high-energy heavy-ion collisions. Following the QGP phase, the bottom quark undergoes hadronization, resulting

in the formation of B-mesons. Subsequently, these B-mesons interact with the surrounding hadronic matter, which comprises particles such as pions, kaons, etas, and nucleons. Rescattering within the hadronic medium can have a significant impact on observables related to B-mesons, particularly their elliptic flow. The study of these effects constitutes a vibrant and active field of research Das et al. (2012); Torres-Rincon et al. (2014). Furthermore, the ratio of heavy baryons to mesons ($\Lambda_c/D$ and $\Lambda_b/B$) is of fundamental importance in gaining insights into in-medium hadronization processes Das et al. (2016). The ratio $\Lambda_b/B$ can serve as a valuable tool for disentangling different hadronization mechanisms. This is one of the primary focuses of the ongoing experimental efforts at both RHIC and LHC colliding energies, aimed at understanding the mechanisms of hadronization. The B-meson, acting as a non-equilibrium probe and being produced early in the collision process, offers an excellent avenue for probing the early stages of heavy-ion collisions. This is a rapidly emerging field of research, with ongoing investigations aimed at gaining a deeper understanding of the dynamics and properties of the early stages of these high-energy collisions Das (2023).

The OQS framework, incorporating inputs from effective field theory and lattice QCD (lQCD) within the context of bottomonium, proves to be a promising approach for examining its behavior in the hot QCD medium. The findings related to quarkonia suppression align consistently with data obtained from RHIC and LHC experiments. Anticipated to undergo essential refinements, this methodology is poised to yield more reliable results in the near future.

## Acknowledgement

We would like to acknowledge the Science and Engineering Research Board (SERB) for Core Research Grant (CRG) No. CRG/2020/002320.

**Data availability statement**: No data associated with the manuscript.

## References

Beraudo, A., *et al.*: Extraction of Heavy-Flavor Transport Coefficients in QCD Matter. Nucl. Phys. A **979**, 21–86 (2018) https://doi.org/10.1016/j.nuclphysa.2018.09.002 arXiv:1803.03824 [nucl-th]

Cao, S., *et al.*: Toward the determination of heavy-quark transport coefficients in quark-gluon plasma. Phys. Rev. C **99**(5), 054907 (2019) https://doi.org/10.1103/PhysRevC.99.054907 arXiv:1809.07894 [nucl-th]

Aarts, G., *et al.*: Heavy-flavor production and medium properties in high-energy nuclear collisions - What next? Eur. Phys. J. A **53**(5), 93 (2017) https://doi.org/10.1140/epja/i2017-12282-9 arXiv:1612.08032 [nucl-th]

Gyulassy, M., McLerran, L.: New forms of QCD matter discovered at RHIC. Nucl. Phys. A **750**, 30–63 (2005) https://doi.org/10.1016/j.nuclphysa.2004.10.034

arXiv:nucl-th/0405013

Shuryak, E.: Strongly coupled quark-gluon plasma in heavy ion collisions. Rev. Mod. Phys. **89**, 035001 (2017) https://doi.org/10.1103/RevModPhys.89.035001 arXiv:1412.8393 [hep-ph]

Djordjevic, M., Gyulassy, M., Wicks, S.: The Charm and beauty of RHIC and LHC. Phys. Rev. Lett. **94**, 112301 (2005) https://doi.org/10.1103/PhysRevLett.94.112301 arXiv:hep-ph/0410372

Ratti, C.: Lattice QCD and heavy ion collisions: a review of recent progress. Rept. Prog. Phys. **81**(8), 084301 (2018) https://doi.org/10.1088/1361-6633/aabb97 arXiv:1804.07810 [hep-lat]

Akamatsu, Y.: Quarkonium in quark–gluon plasma: Open quantum system approaches re-examined. Prog. Part. Nucl. Phys. **123**, 103932 (2022) https://doi.org/10.1016/j.ppnp.2021.103932 arXiv:2009.10559 [nucl-th]

Yao, X.: Open quantum systems for quarkonia. Int. J. Mod. Phys. A **36**(20), 2130010 (2021) https://doi.org/10.1142/S0217751X21300106 arXiv:2102.01736 [hep-ph]

Lansberg, J.-P.: New Observables in Inclusive Production of Quarkonia. Phys. Rept. **889**, 1–106 (2020) https://doi.org/10.1016/j.physrep.2020.08.007 arXiv:1903.09185 [hep-ph]

Moore, G.D., Teaney, D.: How much do heavy quarks thermalize in a heavy ion collision? Phys. Rev. C **71**, 064904 (2005) https://doi.org/10.1103/PhysRevC.71.064904 arXiv:hep-ph/0412346

Svetitsky, B.: Diffusion of charmed quarks in the quark-gluon plasma. Phys. Rev. D **37**, 2484–2491 (1988) https://doi.org/10.1103/PhysRevD.37.2484

Das, S.K., Scardina, F., Plumari, S., Greco, V.: Toward a solution to the $R_{AA}$ and $v_2$ puzzle for heavy quarks. Phys. Lett. B **747**, 260–264 (2015) https://doi.org/10.1016/j.physletb.2015.06.003 arXiv:1502.03757 [nucl-th]

Das, S.K., Scardina, F., Plumari, S., Greco, V.: Heavy-flavor in-medium momentum evolution: Langevin versus Boltzmann approach. Phys. Rev. C **90**, 044901 (2014) https://doi.org/10.1103/PhysRevC.90.044901 arXiv:1312.6857 [nucl-th]

Cacciari, M., Nason, P., Vogt, R.: QCD predictions for charm and bottom production at RHIC. Phys. Rev. Lett. **95**, 122001 (2005) https://doi.org/10.1103/PhysRevLett.95.122001 arXiv:hep-ph/0502203

He, M., Fries, R.J., Rapp, R.: $\mathbf{D_s}$-Meson as Quantitative Probe of Diffusion and Hadronization in Nuclear Collisions. Phys. Rev. Lett. **110**(11), 112301 (2013) https://doi.org/10.1103/PhysRevLett.110.112301 arXiv:1204.4442 [nucl-th]

Scardina, F., Das, S.K., Minissale, V., Plumari, S., Greco, V.: Estimating the charm quark diffusion coefficient and thermalization time from D meson spectra at energies available at the BNL Relativistic Heavy Ion Collider and the CERN Large Hadron Collider. Phys. Rev. C **96**(4), 044905 (2017) https://doi.org/10.1103/PhysRevC.96.044905 arXiv:1707.05452 [nucl-th]

Das, S.K., Torres-Rincon, J.M., Tolos, L., Minissale, V., Scardina, F., Greco, V.: Propagation of heavy baryons in heavy-ion collisions. Phys. Rev. D **94**(11), 114039 (2016) https://doi.org/10.1103/PhysRevD.94.114039 arXiv:1604.05666 [nucl-th]

Tolos, L., Torres-Rincon, J.M., Das, S.K.: Transport coefficients of heavy baryons. Phys. Rev. D **94**(3), 034018 (2016) https://doi.org/10.1103/PhysRevD.94.034018 arXiv:1601.03743 [hep-ph]

Aarts, G., Kim, S., Lombardo, M.P., Oktay, M.B., Ryan, S.M., Sinclair, D.K., Skullerud, J.-I.: Bottomonium above deconfinement in lattice nonrelativistic QCD. Phys. Rev. Lett. **106**, 061602 (2011) https://doi.org/10.1103/PhysRevLett.106.061602 arXiv:1010.3725 [hep-lat]

Aarts, G., Allton, C., Kim, S., Lombardo, M.P., Ryan, S.M., Skullerud, J.-I.: Melting of P wave bottomonium states in the quark-gluon plasma from lattice NRQCD. JHEP **12**, 064 (2013) https://doi.org/10.1007/JHEP12(2013)064 arXiv:1310.5467 [hep-lat]

Aarts, G., Allton, C., Harris, T., Kim, S., Lombardo, M.P., Ryan, S.M., Skullerud, J.-I.: The bottomonium spectrum at finite temperature from $N_f = 2 + 1$ lattice QCD. JHEP **07**, 097 (2014) https://doi.org/10.1007/JHEP07(2014)097 arXiv:1402.6210 [hep-lat]

Kim, S., Petreczky, P., Rothkopf, A.: Lattice NRQCD study of S- and P-wave bottomonium states in a thermal medium with $N_f = 2 + 1$ light flavors. Phys. Rev. D **91**, 054511 (2015) https://doi.org/10.1103/PhysRevD.91.054511 arXiv:1409.3630 [hep-lat]

Kim, S., Petreczky, P., Rothkopf, A.: Quarkonium in-medium properties from realistic lattice NRQCD. JHEP **11**, 088 (2018) https://doi.org/10.1007/JHEP11(2018)088 arXiv:1808.08781 [hep-lat]

Burnier, Y., Kaczmarek, O., Rothkopf, A.: Quarkonium at finite temperature: Towards realistic phenomenology from first principles. JHEP **12**, 101 (2015) https://doi.org/10.1007/JHEP12(2015)101 arXiv:1509.07366 [hep-ph]

Mocsy, A., Petreczky, P.: Can quarkonia survive deconfinement? Phys. Rev. D **77**, 014501 (2008) https://doi.org/10.1103/PhysRevD.77.014501 arXiv:0705.2559 [hep-ph]

Larsen, R., Meinel, S., Mukherjee, S., Petreczky, P.: Thermal broadening of bottomonia: Lattice nonrelativistic QCD with extended operators. Phys. Rev. D **100**(7), 074506 (2019) https://doi.org/10.1103/PhysRevD.100.074506 arXiv:1908.08437 [hep-lat]

Larsen, R., Meinel, S., Mukherjee, S., Petreczky, P.: Excited bottomonia in quark-gluon plasma from lattice QCD. Phys. Lett. B **800**, 135119 (2020) https://doi.org/10.1016/j.physletb.2019.135119 arXiv:1910.07374 [hep-lat]

Matsui, T., Satz, H.: $J/\psi$ Suppression by Quark-Gluon Plasma Formation. Phys. Lett. B **178**, 416–422 (1986) https://doi.org/10.1016/0370-2693(86)91404-8

Laine, M., Philipsen, O., Romatschke, P., Tassler, M.: Real-time static potential in hot QCD. JHEP **03**, 054 (2007) https://doi.org/10.1088/1126-6708/2007/03/054 arXiv:hep-ph/0611300

Brambilla, N., Escobedo, M.A., Ghiglieri, J., Vairo, A.: Thermal width and quarkonium dissociation by inelastic parton scattering. JHEP **05**, 130 (2013) https://doi.org/10.1007/JHEP05(2013)130 arXiv:1303.6097 [hep-ph]

Brambilla, N., Escobedo, M.A., Soto, J., Vairo, A.: Heavy quarkonium suppression in a fireball. Phys. Rev. D **97**(7), 074009 (2018) https://doi.org/10.1103/PhysRevD.97.074009 arXiv:1711.04515 [hep-ph]

Caswell, W.E., Lepage, G.P.: Effective Lagrangians for Bound State Problems in QED, QCD, and Other Field Theories. Phys. Lett. B **167**, 437–442 (1986) https://doi.org/10.1016/0370-2693(86)91297-9

Bodwin, G.T., Braaten, E., Lepage, G.P.: Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium. Phys. Rev. D **51**, 1125–1171 (1995) https://doi.org/10.1103/PhysRevD.55.5853 arXiv:hep-ph/9407339. [Erratum: Phys.Rev.D 55, 5853 (1997)]

Pineda, A., Soto, J.: Effective field theory for ultrasoft momenta in NRQCD and NRQED. Nucl. Phys. B Proc. Suppl. **64**, 428–432 (1998) https://doi.org/10.1016/S0920-5632(97)01102-X arXiv:hep-ph/9707481

Brambilla, N., Pineda, A., Soto, J., Vairo, A.: Potential NRQCD: An Effective theory for heavy quarkonium. Nucl. Phys. B **566**, 275 (2000) https://doi.org/10.1016/S0550-3213(99)00693-8 arXiv:hep-ph/9907240

Brambilla, N., Pineda, A., Soto, J., Vairo, A.: Effective Field Theories for Heavy Quarkonium. Rev. Mod. Phys. **77**, 1423 (2005) https://doi.org/10.1103/RevModPhys.77.1423 arXiv:hep-ph/0410047

Akamatsu, Y.: Heavy quark master equations in the Lindblad form at high temperatures. Phys. Rev. D **91**(5), 056002 (2015) https://doi.org/10.1103/PhysRevD.91.

056002 arXiv:1403.5783 [hep-ph]

Brambilla, N., Escobedo, M.A., Strickland, M., Vairo, A., Vander Griend, P., Weber, J.H.: Bottomonium suppression in an open quantum system using the quantum trajectories method. JHEP **05**, 136 (2021) https://doi.org/10.1007/JHEP05(2021)136 arXiv:2012.01240 [hep-ph]

Strickland, M., Thapa, S.: Bottomonium suppression at RHIC and LHC in an open quantum system approach. Phys. Rev. D **108**(1), 014031 (2023) https://doi.org/10.1103/PhysRevD.108.014031 arXiv:2305.17841 [hep-ph]

Binder, T., Blobel, B., Harz, J., Mukaida, K.: Dark matter bound-state formation at higher order: a non-equilibrium quantum field theory approach. JHEP **09**, 086 (2020) https://doi.org/10.1007/JHEP09(2020)086 arXiv:2002.07145 [hep-ph]

Braaten, E., Hammer, H.-W., Lepage, G.P.: Open Effective Field Theories from Deeply Inelastic Reactions. Phys. Rev. D **94**(5), 056006 (2016) https://doi.org/10.1103/PhysRevD.94.056006 arXiv:1607.02939 [hep-ph]

Boyanovsky, D.: Effective field theory during inflation: Reduced density matrix and its quantum master equation. Phys. Rev. D **92**(2), 023527 (2015) https://doi.org/10.1103/PhysRevD.92.023527 arXiv:1506.07395 [astro-ph.CO]

Boyanovsky, D.: Effective field theory during inflation. II. Stochastic dynamics and power spectrum suppression. Phys. Rev. D **93**, 043501 (2016) https://doi.org/10.1103/PhysRevD.93.043501 arXiv:1511.06649 [astro-ph.CO]

Pandey, H., Schlichting, S., Sharma, S.: Heavy quark momentum broadening in a non-Abelian plasma away from thermal equilibrium (2023) arXiv:2312.12280 [hep-lat]

Das, S.K., Ghosh, S., Sarkar, S., Alam, J.-e.: Drag and diffusion coefficients of $B$ mesons in hot hadronic matter. Phys. Rev. D **85**, 074017 (2012) https://doi.org/10.1103/PhysRevD.85.074017 arXiv:1109.3359 [hep-ph]

Torres-Rincon, J.M., Tolos, L., Romanets, O.: Open bottom states and the $\bar{B}$-meson propagation in hadronic matter. Phys. Rev. D **89**(7), 074042 (2014) https://doi.org/10.1103/PhysRevD.89.074042 arXiv:1403.1371 [hep-ph]

Das, S.K.: Open heavy flavors: Theory. In: 11th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions: Hard Probes 2023 (2023)

arXiv:2402.18738v1 [hep-ph] 28 Feb 2024

# QCD at Finite Temperature and Density - Equation of State

*Jamie* Karthein[1*]

[1]Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

**Abstract.** As an important set of thermodynamic quantities, knowledge of the equation of state over a broad range of temperatures and chemical potentials in the QCD phase diagram is crucial for our understanding of strongly-interacting matter. There is a good understanding from first-principles results in lattice QCD, perturbative QCD and chiral effective field theory about the equation of state. However, these approaches are valid in different regimes of the phase diagram, and therefore, a method of providing an equation of state that covers a full range of the phase diagram involves matching together these results with appropriate models in order to fill in the gaps between these regions. Furthermore, with such equations of state, important questions about QCD phase structure can begin to be addressed, such as whether there is a critical point in the QCD phase diagram. In this contribution to the proceedings, equations of state from first-principles and effective theories will be discussed in order to understand how QCD thermodynamics is affected by the presence of a critical point.

## 1 Introduction

The phase structure of strongly-interacting matter has been a focus of efforts for both the theoretical and experimental communities for decades. Upon exploration of the QCD phase diagram, the question being probed is how does strongly-interacting matter respond to heat, i.e. increasing temperature, and compression, i.e. increasing density. Therefore, the QCD equation of state (EoS) is the object of study lending from the fact that these are thermodynamic properties. First-priniciples lattice QCD calculations have given insight into the zero and low-to-moderate density equation of state [1–5]. Also from the lattice with good precision, the transition temperature at vanishing chemical potential has been calculated as well as the curvature of the transition line and where it lies up to $\mu_B \simeq 2.5T$ [6, 7]. In the low-temperature region of the QCD phase diagram, the system is rather well-described by an ideal gas of hadrons and their resonances [8–11]. However, as the density, or equivalently chemical potential, increases, nuclear matter exhibits a liquid-gas phase transition in the low temperature regime that is described by an interacting van der Waals equation of state [9]. Furthermore, chiral effective field theory is the appropriate low energy theory that covers the low-temperature and low-to-moderate-density regime [12]. On the other hand, at asymptotically high temperatures and densities, perturbative QCD (pQCD) results yield access to the asymptotic regime of the QCD phase diagram [13–15]. From the perspective of connecting to the observable universe, zero temperature calculations from chiral effective field theory and pQCD can be combined with astrophysical constraints to learn about the QCD equation

*e-mail: jmkar@mit.edu

of state [16, 17]. On the other hand in terms of terrestrial experiments, along the $T-$axis of the QCD phase diagram are high energy heavy-ion collisions, such as those conducted at the Large Hadron Collider (LHC) at CERN [18, 19]. Furthermore, by decreasing collider energy, the phase diagram at non-zero baryon chemical potential can be explored, which is a main focus of the Beam Energy Scan (BES) program at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Lab [20].

In light of these varied sources across theory, experiment, and astrophysical observation, there is a plethora of information about the phase structure of QCD within the regimes of validity of each of them [21]. These, therefore, yield the ground truth knowledge of the QCD EoS, which, however, does not cover the entire reach of the phase diagram. Thus, many questions remain about the phases of QCD. For example, does the crossover phase transitions proven by lattice calculations [22] become a proper first order phase transition at larger chemical potentials? This would imply that there is a critical point somewhere in the phase diagram where the change in transition type occurs. The search for this critical point has been pursued in part by studying critical fluctuations related to the diverging correlation length at the critical point [23–25]. On the other hand, where does the transition line lie at high densities which are beyond the current reach of lattice calculations? This question in essence probes whether the densities found in neutron stars are large enough to create deconfined quark matter. There are a number of further profound questions about the nature of strongly-interacting matter beyond the reaches of current experimental and theoretical methods. Thus, in order to explore these questions from the theory side, a method of piecing the current first-principles knowledge together with models of QCD is required in order to fill in the gaps in the phase diagram. The goal of the MUSES collaboration is to provide equations of state along these lines. The forthcoming publicly available software will allow for user-specified coverage of the phase diagram and user-defined conditions in addition to the choice of model in the region of applicability.

As mentioned previously there are many open questions about QCD phase structure, here I will focus on the regime covered by high energy heavy-ion collisions due to the brevity required for proceedings. In this contribution, I will address the question of how QCD thermodynamics are affected by critical features in the phase diagram.

## 2 Equation of State at Finite Temperature and Density

### 2.1 Utilizing a critical equation of state to estimate equilibrium proton fluctuations

In order to study the effect of a critical point that could potentially be observed in the BES-II program at RHIC, the 3D Ising model is utilized to map such critical behavior onto the phase diagram of QCD. The 3D Ising model was chosen for this approach because it exhibits the same scaling features in the vicinity of a critical point as QCD, in other words, they belong to the same universality class [26, 27]. The mapping of the critical behavior onto the QCD phase diagram is, however, not universal, which is to say that there are no strict mapping parameters that exist *a priori* between the Ising phase diagram and the one for QCD. Such a map can be constructed in the following way as in the BEST equation of state [28, 29]:

$$\frac{T - T_c}{T_c} = \omega(\rho r \sin\alpha_1 + h \sin\alpha_2) \tag{1}$$

$$\frac{\mu_B - \mu_{B,c}}{T_c} = \omega(-\rho r \cos\alpha_1 - h \cos\alpha_2) \tag{2}$$

Utilizing this mapping, we study the critical fluctuations in the Ising model in order to calculate the proton fluctuations near a critical point. We provide an update to the approach

laid out in Ref. [30] for calculating such particle fluctuations. This update includes utilizing the critical correlation length from universality. We, additionally, further rely on universality in order to extract the higher order couplings, $\lambda$, of the fluctuations:

$$\kappa_3 = 2\lambda_3 VT^2\xi^6 \qquad \kappa_4 = 6VT^3[2(\lambda_3\xi)^2 - \lambda_4]\xi^8 \tag{3}$$

Following the work of Pradeep, Rajagopal, Stephanov, and Yin on Gaussian fluctuations [31], we explore the higher order non-Gaussian fluctuations in- and out-of-equilibrium. As pertains to this contribution centered on the equation of state, an equilibrium thermodynamic property, we will focus on the equilibrium results related to that work.

The first part of updating the estimates for proton fluctuations involves the input for the correlation length from universality. The scaling form of the correlation length in the Ising model, which follows Widom's scaling form in terms of Ising model variables, can be written as [32]:

$$\xi^2(r, M) = M^{-2\nu/\beta} f_\xi(r/M^{1/\beta}), \tag{4}$$

where $\nu$ is the correlation length critical exponent, $f_\xi$ is the scaling function and the scaling parameter is $x = \frac{r}{M^{1/\beta}}$.

In the $\epsilon$-expansion, the function $f_\xi$ is given to $O(\epsilon^2)$ as [33]:

$$\begin{aligned} f_\xi(x) = (f_1^+)^2 6^{-2\nu} z\Big\{1 - \frac{\epsilon}{36}[(5 + 6\ln 3)z - 6(1+z)\ln z] + \epsilon^2\Big[\frac{1+2z^2}{72}\ln^2 z + \Big(\frac{z}{18}\Big(z - \frac{1}{2}\Big)(1 - \ln 3) \\ - \frac{1}{216}\Big(16z^2 - \frac{47}{3}z - \frac{56}{3}\Big)\Big)\ln z + \frac{1}{216}\Big(\frac{101}{6} + \frac{2}{3}I + 6\ln^2 3 + 4\ln 3 - 10\Big)z^2 \\ - \frac{1}{216}\Big(6\ln^2 3 + \frac{44}{3}\ln 3 + \frac{137}{9} + \frac{8}{3}I\Big)z\Big]\Big\} \end{aligned} \tag{5}$$

where $f_1^+$ is a non-universal critical amplitude, $z \equiv \frac{2}{1+x/3}$, $I \equiv \int_0^1 \frac{\ln[x(1-x)]}{1-x(1-x)}dx \sim -2.3439$, and $\epsilon = 4 - d$, where $d$ is the spatial dimension.

One may also write the parametric form of the correlation length in the Ising model in terms of the parametric variables $(R, \theta)$ as [32]:

$$\xi^2(M, t) = R^{-2\nu} g_\xi(\theta), \quad g_\xi(\theta) = g_\xi(0)(1 - \frac{5}{18}\epsilon\theta^2 + O(\epsilon^2)). \tag{6}$$

In this case the parametric equation of state is defined as:

$$M = M_0 R^\beta \theta \qquad h = h_0 R^{\beta\delta}\tilde{h}(\theta) \qquad r = x_0 R(1 - \theta^2) \tag{7}$$

where $x_0 = b^{1/\beta}/(b^2 - 1)$, $h(\theta) = h_0\theta(b^2 - \theta^2)(1 + c\theta^2) + O(\epsilon^4)$, $b^2 = \frac{3}{2}(1 - \frac{\epsilon^2}{12})$, $c = -\frac{\epsilon^3}{18}(\zeta(3) + \frac{I-1}{4})$. With these calculations, we further expand the correlation length to $O(\epsilon^2)$. Thus, the $\theta$-dependence of the correlation length to $O(\epsilon^2)$ is then given by:

$$\tilde{g}_\xi = \tilde{g}_\xi(0)\left(1 - \frac{5}{18}\epsilon\theta^2 + \Big[\frac{1}{972}(24I - 25)\theta^2 + \frac{1}{324}(4I + 41)\theta^4\Big]\epsilon^2\right). \tag{8}$$

We utilize this updated correlation length to $O(\epsilon^2)$ in order to be consistent with the equation of state which is valid to the same order in the $\epsilon$-expansion.

On the other hand, we also take into account the $\mu_B$-dependence of the couplings $\lambda_3$ and $\lambda_4$, as shown in Eq. 3. Since we know that the 4th order moment is expected to exhibit

a sign change [25], these couplings will not simply be constants, which provides another modernization of the work in Ref. [30]. We extract these higher order couplings via the knowledge of the correlation length and fluctuations in the Ising model. In order to study this $\mu_B$-dependence of the critical correlation, we need to choose the non-universal parameters of the Ising-QCD mapping. We choose parameters for our EoS in order to be consistent with recent results from STAR [34]. Furthermore, our choice also obeys constraints from universality studied by Pradeep and Stephanov in Ref. [35] as well as those given by causality and stability of the BEST EoS as shown by Mroczek et al in Ref. [36]. As such, we move away from the choice of orthogonal Ising axes as given by the canonical choice in the BEST EoS in Refs. [28, 29] in order to agree with constraints from the universal features of this mapping shown by Pradeep and Stephanov. Consequently, in order to have both the constraints properly taken into account and a strength of divergence that will reasonably agree with results from STAR net-proton energy dependence, we must modify the choice of scaling parameters, $w, \rho$. However, as shown in the analysis of the BEST EoS parameter space by Mroczek et al in Ref. [36], decreasing the angle difference, i.e. moving away from orthogonality, leads to a restriction of the allowed values for $w, \rho$. Thus, we arrive at the following choice of mapping parameters.

$$\mu_{B,c} = 420 \text{ MeV} \quad \Delta\alpha = \alpha_1 - \alpha_2 = 10^\circ \quad w = 8 \quad \rho = 0.2$$

Finally, with inputs from universality for both the correlation length and the $\mu_B$-dependent couplings, we calculate the 3rd and 4th order proton fluctuations, similar to Ref. [30],

$$\omega_{3p,\sigma} = \frac{2\tilde{\lambda}_3}{T^{3/2}\, n_p} \xi^{9/2} \left( d_p\, g_p \int_k \frac{v_k^2}{\gamma_k} \right)^3 \qquad \omega_{4p,\sigma} = \frac{6(2\tilde{\lambda}_3^2 - \tilde{\lambda}_4)}{T^2\, n_p} \xi^7 \left( d_p\, g_p \int_k \frac{v_k^2}{\gamma_k} \right)^4 \tag{9}$$

where the dimensionless $\lambda$'s are given by $\tilde{\lambda}_3 = \lambda_3(T^{1/2}\xi^{3/2})$, $\tilde{\lambda}_4 = \lambda_4(T\xi)$, $n_p$ is the proton density, and the kinematic integrals involve the degeneracy, $d_p$, and the proton coupling to the sigma field $g_p$. We consider a coupling of $g_p \sim 7$ as in Ref. [30].

Figure 1 shows the 4th order proton fluctuations in the phase diagram (left) and along several freeze-out trajectories (right). The freeze-out curves are taken to be 3, 5, and 7 MeV below the critical point. The freeze-out curve at $\Delta T = 5$ MeV is from the fit to experimental results performed in Ref. [37], while the other two are simply a shift from this established freeze-out curve to show a range of possibilities within errors. The order of magnitude of the fluctuations is very sensitive to the path of the freeze-out curve through the phase diagram. Taking the experimentally fitted curve at $\Delta T = 5$ MeV, the fluctuations are $O(10^2)$. However, these are still equilibrium quantities which are not to be compared with experimental results. A treatment such as the one in Ref. [31] for these higher order moments is first required. Therefore, given these large equilibrium results and what is known about the effects of freezing out critical fluctuations from Ref. [31], we expect the experimental results to be reduced. Furthermore, here we have made one choice of parameters for the Ising-QCD mapping that is motivated by utilizing the available universal inputs. However, this remains a choice which also effects the size and shape of the fluctuations. In a forthcoming publication, we will explore the effect of the parameters in detail.

## 3 Conclusions

Many open questions remain about the phase structure of QCD at finite temperatures and densities which are outside the current range of first-principle approaches. I focused here on the effects of a potential critical point on the equilibrium particle fluctuations. By utilizing universal properties of the equation of state, we have made estimates for the equilibrium

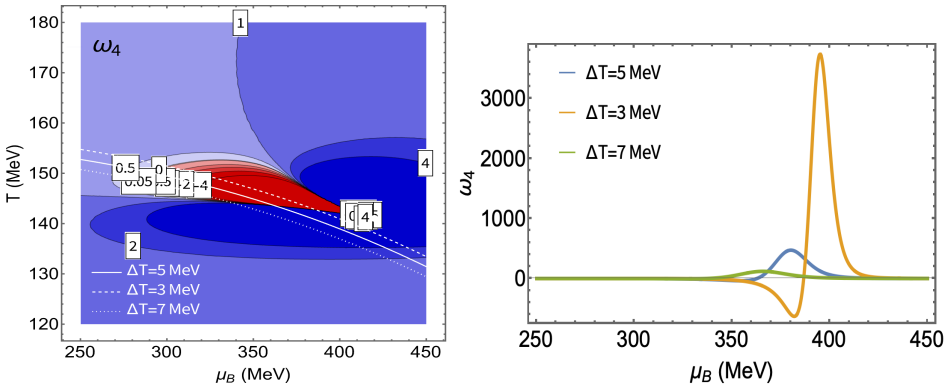



**Figure 1.** Fourth order proton fluctuations calculated as shown in Eq. 9 in the phase diagram (left) and along freeze-out trajectories (right).

critical fluctuations of protons. Importantly, these equilibrium quantities are ready to in turn be used as input for estimating out-of-equilibrium fluctuations in a similar manner as was previously shown for Gaussian fluctuations in Ref. [31] such that they can more readily be compared to experimental data. Furthermore, given these tools to make such estimates, once the experimental data becomes available we will be able to constrain the non-universal features of this approach in order to yield the proper mapping between the QCD phase diagram and its analogous critical features from the 3D Ising model.

## 4 Acknowledgements

The author is grateful to the conference organizers for the opportunity to present this and other work at Quark Matter 2023. The author is supported by an Ascending Postdoctoral Scholar Fellowship from the National Science Foundation under Award No. 2138063.

## References

[1] S. Borsanyi, G. Endrodi, Z. Fodor, A. Jakovac, S.D. Katz, S. Krieg, C. Ratti, K.K. Szabo, JHEP **11**, 077 (2010), 1007.2580

[2] A. Bazavov, T. Bhattacharya, C. DeTar, H. Ding, S. Gottlieb, R. Gupta et al. (HotQCD), Phys. Rev. **D90**, 094503 (2014), 1407.6387

[3] A. Bazavov et al., Phys. Rev. D **95**, 054504 (2017), 1701.04325

[4] S. Borsányi, Z. Fodor, J.N. Guenther, R. Kara, S.D. Katz, P. Parotto, A. Pásztor, C. Ratti, K.K. Szabó, Phys. Rev. Lett. **126**, 232001 (2021), 2102.06660

[5] S. Mondal, S. Mukherjee, P. Hegde (arXiv:2106.03165, 2021), 2106.03165

[6] S. Borsanyi, Z. Fodor, J.N. Guenther, R. Kara, S.D. Katz, P. Parotto, A. Pasztor, C. Ratti, K.K. Szabo, Phys. Rev. Lett. **125**, 052001 (2020), 2002.02821

[7] G. Aarts et al., Prog. Part. Nucl. Phys. **133**, 104070 (2023), 2301.04382

[8] S. Borsanyi, Z. Fodor, C. Hoelbling, S.D. Katz, S. Krieg, C. Ratti, K.K. Szabo (Wuppertal-Budapest), JHEP **09**, 073 (2010), 1005.3508

[9] V. Vovchenko, M.I. Gorenstein, H. Stoecker, Phys. Rev. Lett. **118**, 182301 (2017), 1609.03975

[10] P. Alba, V.M. Sarti, J. Noronha-Hostler, P. Parotto, I. Portillo-Vazquez, C. Ratti, J.M. Stafford, Phys. Rev. C **101**, 054905 (2020), 2002.12395
[11] J.S. San Martin, R. Hirayama, J. Hammelmann, J.M. Karthein, P. Parotto, J. Noronha-Hostler, C. Ratti, H. Elfner (2023), 2309.01737
[12] C. Drischler, J.W. Holt, C. Wellenhofer, Ann. Rev. Nucl. Part. Sci. **71**, 403 (2021), 2101.01709
[13] J.O. Andersen, M. Strickland, Phys. Rev. D **66**, 105001 (2002), hep-ph/0206196
[14] A. Kurkela, P. Romatschke, A. Vuorinen, Phys. Rev. D **81**, 105021 (2010), 0912.1856
[15] N. Haque, J.O. Andersen, M.G. Mustafa, M. Strickland, N. Su, Phys. Rev. D **89**, 061701 (2014), 1309.3968
[16] E. Annala, T. Gorda, A. Kurkela, J. Nättilä, A. Vuorinen, Nature Phys. **16**, 907 (2020), 1903.09121
[17] D. Mroczek, M.C. Miller, J. Noronha-Hostler, N. Yunes (2023), 2309.02345
[18] (2022), 2211.04384
[19] W. Busza, K. Rajagopal, W. van der Schee, Ann. Rev. Nucl. Part. Sci. **68**, 339 (2018), 1802.04801
[20] A. Bzdak, S. Esumi, V. Koch, J. Liao, M. Stephanov, N. Xu, Phys. Rept. **853**, 1 (2020), 1906.00936
[21] R. Kumar et al. (MUSES) (2023), 2303.17021
[22] Y. Aoki, G. Endrodi, Z. Fodor, S.D. Katz, K.K. Szabo, Nature **443**, 675 (2006), hep-lat/0611014
[23] M.A. Stephanov, K. Rajagopal, E.V. Shuryak, Phys. Rev. Lett. **81**, 4816 (1998), hep-ph/9806219
[24] M.A. Stephanov, K. Rajagopal, E.V. Shuryak, Phys. Rev. D **60**, 114028 (1999), hep-ph/9903292
[25] M. Stephanov, Phys. Rev. Lett. **107**, 052301 (2011), 1104.1627
[26] R.D. Pisarski, F. Wilczek, Phys. Rev. D **29**, 338 (1984)
[27] K. Rajagopal, F. Wilczek, Nucl. Phys. B **399**, 395 (1993), hep-ph/9210253
[28] P. Parotto, M. Bluhm, D. Mroczek, M. Nahrgang, J. Noronha-Hostler, K. Rajagopal, C. Ratti, T. Schäfer, M. Stephanov, Phys. Rev. C **101**, 034901 (2020), 1805.05249
[29] J.M. Karthein, D. Mroczek, A.R. Nava Acuna, J. Noronha-Hostler, P. Parotto, D.R.P. Price, C. Ratti, Eur. Phys. J. Plus **136**, 621 (2021), 2103.08146
[30] C. Athanasiou, K. Rajagopal, M. Stephanov, Phys. Rev. D **82**, 074008 (2010), 1006.4636
[31] M. Pradeep, K. Rajagopal, M. Stephanov, Y. Yin, Phys. Rev. D **106**, 036017 (2022), 2204.00639
[32] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Oxford University Press, 2002), ISBN 9780198509233, https://doi.org/10.1093/acprof:oso/9780198509233.001.0001
[33] E. Brezin, J.C. Le Guillou, J. Zinn-Justin, Physics Letters A **47**, 285 (1974)
[34] J. Adam et al. (STAR), Phys. Rev. Lett. **126**, 092301 (2021), 2001.02852
[35] M.S. Pradeep, M. Stephanov, Phys. Rev. D **100**, 056003 (2019), 1905.13247
[36] D. Mroczek, M. Hjorth-Jensen, J. Noronha-Hostler, P. Parotto, C. Ratti, R. Vilalta (2022), 2203.13876
[37] A. Andronic, P. Braun-Munzinger, K. Redlich, J. Stachel, Nature **561**, 321 (2018), 1710.09425

Nikhef-2023-020
TIF-UNIMI-2023-23
Edinburgh 2023/29

# The Path to N$^3$LO Parton Distributions

**The NNPDF Collaboration**:
Richard D. Ball[1], Andrea Barontini[2], Alessandro Candido[2,3], Stefano Carrazza[2], Juan Cruz-Martinez[3], Luigi Del Debbio[1], Stefano Forte[2], Tommaso Giani[4,5], Felix Hekhorn[2,6,7], Zahari Kassabov[8], Niccolò Laurenti,[2] Giacomo Magni[4,5], Emanuele R. Nocera[9], Tanjona R. Rabemananjara[4,5], Juan Rojo[4,5], Christopher Schwan[10], Roy Stegeman[1], and Maria Ubiali[8]

[1] *The Higgs Centre for Theoretical Physics, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3JZ, Scotland*
[2] *Tif Lab, Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy*
[3] *CERN, Theoretical Physics Department, CH-1211 Geneva 23, Switzerland*
[4] *Department of Physics and Astronomy, Vrije Universiteit, NL-1081 HV Amsterdam*
[5] *Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands*
[6] *University of Jyvaskyla, Department of Physics, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland*
[7] *Helsinki Institute of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland*
[8] *DAMTP, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom*
[9] *Dipartimento di Fisica, Università degli Studi di Torino and INFN, Sezione di Torino, Via Pietro Giuria 1, I-10125 Torino, Italy*
[10] *Universität Würzburg, Institut für Theoretische Physik und Astrophysik, 97074 Würzburg, Germany*

*This paper is dedicated to the memory of Stefano Catani, Grand Master of QCD, great scientist and human being*

arXiv:2402.18635v1 [hep-ph] 28 Feb 2024

## Abstract

We extend the existing leading (LO), next-to-leading (NLO), and next-to-next-to-leading order (NNLO) NNPDF4.0 sets of parton distribution functions (PDFs) to approximate next-to-next-to-next-to-leading order (aN$^3$LO). We construct an approximation to the N$^3$LO splitting functions that includes all available partial information from both fixed-order computations and from small and large $x$ resummation, and estimate the uncertainty on this approximation by varying the set of basis functions used to construct the approximation. We include known N$^3$LO corrections to deep-inelastic scattering structure functions and extend the FONLL general-mass scheme to $\mathcal{O}\left(\alpha_s^3\right)$ accuracy. We determine a set of aN$^3$LO PDFs by accounting both for the uncertainty on splitting functions due to the incomplete knowledge of N$^3$LO terms, and to the uncertainty related to missing higher corrections (MHOU), estimated by scale variation, through a theory covariance matrix formalism. We assess the perturbative stability of the resulting PDFs, we study the impact of MHOUs on them, and we compare our results to the aN$^3$LO PDFs from the MSHT group. We examine the phenomenological impact of aN$^3$LO corrections on parton luminosities at the LHC, and give a first assessment of the impact of aN$^3$LO PDFs on the Higgs and Drell-Yan total production cross-sections. We find that the aN$^3$LO NNPDF4.0 PDFs are consistent within uncertainties with their NNLO counterparts, that they improve the description of the global dataset and the perturbative convergence of Higgs and Drell-Yan cross-sections, and that MHOUs on PDFs decrease substantially with the increase of perturbative order.

# Contents


**1 Introduction** **2**

**2 Approximate N$^3$LO evolution** **3**
2.1 Construction of the approximation . . . 3
2.2 The approximate anomalous dimension matrix and its uncertainty . . . 5
2.3 aN$^3$LO anomalous dimensions: the nonsinglet sector . . . 6
2.4 aN$^3$LO anomalous dimensions: singlet sector . . . 8
2.5 Results: aN$^3$LO splitting functions . . . 11
2.6 Results: aN$^3$LO evolution . . . 14
2.7 Comparison to other groups . . . 15

**3 N$^3$LO partonic cross-sections** **17**
3.1 N$^3$LO corrections to DIS structure functions . . . 17
3.2 A general-mass variable flavor number scheme at N$^3$LO . . . 20
3.3 N$^3$LO corrections to hadronic processes . . . 23

**4 NNPDF4.0 at aN$^3$LO** **25**
4.1 Fit quality . . . 26
4.2 Parton distributions . . . 28
4.3 PDF uncertainties . . . 35
4.4 Implications for intrinsic charm . . . 35
4.5 Dependence on the treatment of aN$^3$LO corrections. . . . 40
4.6 Comparison with MSHT20 . . . 40

**5 LHC phenomenology at aN$^3$LO accuracy** **45**
5.1 Inclusive Drell-Yan production . . . 45
5.2 Inclusive Higgs production . . . 47

**6 Summary and outlook** **52**

**A Explicit expressions of anomalous dimensions** **55**


# 1 Introduction

Calculations of hard-scattering cross-sections at fourth perturbative order in the strong coupling, i.e. at next-to-next-to-next-to-leading order (N$^3$LO), have been available for a long time for massless deep-inelastic scattering (DIS) [1–4], and have more recently become available for a rapidly growing set of hadron collider processes. These include inclusive Higgs production in gluon-fusion [5, 6], bottom-fusion [7], in association with vector bosons [8], and in vector-boson-fusion [9], Higgs pair production [10], inclusive Drell-Yan production [11,12], differential Higgs production [13–17], and differential Drell-Yan distributions [18,19], see [20] for an overview.

In order to obtain predictions for hadronic observables with this accuracy, these partonic cross-sections must be combined with parton distribution functions (PDFs) determined at the same perturbative order. These, in turn, must be determined by comparing to experimental data theory predictions computed at the same accuracy. The main bottleneck in carrying out this programme is the lack of exact expressions for the N$^3$LO splitting functions that govern the scale dependence of the PDFs: for these only partial information is available [21–31]. This information includes a set of integer $N$-Mellin moments, terms proportional to $n_f^k$ with $k \geq 1$, and the large- and small-$x$ limits. By combining these partial results it is possible to attempt an approximate determination of the N$^3$LO splitting functions [30, 32], as was successfully done in the past at NNLO [33].

At present a global PDF determination at N$^3$LO must consequently be based on incomplete information: the approximate knowledge of splitting functions, and full knowledge of partonic cross-sections only for a subset of processes. A first attempt towards achieving this was recently made in Ref. [32], where the missing theoretical information on N$^3$LO calculations was parametrized in terms of a set of nuisance parameters, which were determined together with the PDFs from a fit to experimental data.

Here we adopt a somewhat different strategy. Namely, we use a theory covariance matrix formalism in order to account for the missing perturbative information. It was shown in Ref. [34] that nuclear uncertainties can be included through a theory covariance matrix, and it was further shown in Refs. [35, 36] how such a theory covariance matrix can be constructed to account for missing higher-order uncertainties (MHOUs), estimated through renormalization and factorization scale variation. Here we will use the same formalism in order to also construct a theory covariance matrix for incomplete higher-order uncertainties (IHOUs), namely, those related to incomplete knowledge of the N$^3$LO theory, specifically for the splitting functions and for the massive DIS coefficient functions. Equipped with such theory covariance matrices, we can perform a determination of PDFs at "approximate N$^3$LO" (hereafter denoted aN$^3$LO), in which the theory covariance matrix accounts both for incomplete knowledge of N$^3$LO splitting functions and massive coefficient functions (IHOUs), and for missing N$^3$LO corrections to the partonic cross-sections for hadronic processes (MHOUs).

We will thus present the aN$^3$LO NNPDF4.0 PDF determination, to be added to the existing LO, NLO and NNLO sets [37], as well as the more recent NNPDF4.0 MHOU PDFs [38] that also include MHOUs in the PDF uncertainty. Besides using a different methodology to the MSHT20 study [32], here we are also able to include more recent exact results [28–31] that stabilize the N$^3$LO splitting function parametrisation. Our construction is implemented in the open-source NNPDF framework [39]. Specifically, our aN$^3$LO evolution is implemented in EKO [40] and the N$^3$LO DIS coefficient functions, including the FONLL general-mass scheme, in YADISM [41]. With PDFs determined from the same global dataset and using the same methodology at four consecutive perturbative orders it is now possible to assess carefully perturbative stability and provide a reliable uncertainty estimation.

The outline of this paper is as follows. In Sect. 2 we construct an approximation to the N$^3$LO splitting functions based on all known exact results and limits. We compare it with the MSHT approximation [32] as well as with the more recent approximation of Refs. [28–30]. In Sect. 3 we discuss available and approximate N$^3$LO corrections to hard cross sections: specifically, DIS coefficient functions, including a generalization to this order of the FONLL [42–44] method for the inclusion of heavy quark mass effects, and the Drell-Yan cross-section. In Sect. 4 we present the main results of this work, namely the aN$^3$LO NNPDF4.0 PDF set, based on the results of Sects. 2 and 3. Perturbative convergence before and after the inclusion of MHOUs is discussed in detail, and results are compared to those of the MSHT group [32]. A first assessment of the impact of aN$^3$LO PDFs on Drell-Yan and Higgs production is presented in Sect. 5. Finally, a summary and outlook on future developments are presented in Sect. 6. Expressions for the anomalous dimensions parametrized in Sects. 2.3-2.4 are given in Appendix A.

## 2 Approximate N$^3$LO evolution

We proceed to the construction and implementation of aN$^3$LO evolution. We first describe our strategy to approximate the N$^3$LO evolution equations, the way this is used to construct aN$^3$LO anomalous dimensions and splitting functions, and to estimate the uncertainty in the approximation and its impact on theory predictions. We then use this strategy to construct an approximation in the nonsinglet sector, where accurate results have been available for a while [22], and benchmark it against these results. We then present our construction of aN$^3$LO singlet splitting functions, examine our results, their uncertainties and their perturbative behavior, and also how they relate to NLL small-$x$ resummation. We next describe our implementation of aN$^3$LO evolution and study the impact of aN$^3$LO on the perturbative evolution of PDFs. Finally, we compare our aN$^3$LO singlet splitting functions to those of the MSHT group and to the recent results of [28–30].

### 2.1 Construction of the approximation

We write the QCD evolution equations as

$$\mu^2 \frac{\partial f_i(x,\mu^2)}{\partial \mu^2} = \int_x^1 \frac{dz}{z} P_{ij}(x/z, a_s(\mu^2)) f_j(z,\mu^2)\,, \tag{2.1}$$

where $f_i(x,\mu^2)$ is a vector of PDFs and, with $n_f$ active quark flavors, the $(2n_f+1)\times(2n_f+1)$ splitting function matrix $P_{ij}(x,a_s(\mu^2))$ is expanded perturbatively as

$$P_{ij}(x,a_s(\mu^2)) = a_s P_{ij}^{(0)}(x) + a_s^2 P_{ij}^{(1)}(x) + a_s^3 P_{ij}^{(2)}(x) + a_s^4 P_{ij}^{(3)}(x) + \mathcal{O}\left(a_s^5\right)\,. \tag{2.2}$$

in powers of the strong coupling $a_s(\mu^2) = \alpha_s(\mu^2)/4\pi$.

Defining Mellin space PDFs $f_i(N, \mu^2)$ (denoted in a slight abuse of notation by the same symbol as the $x$-space PDFs), and anomalous dimensions $\gamma_{ij}(N, a_s(\mu^2))$ as minus the Mellin transforms of splitting functions,

$$f_i(N, \mu^2) = \mathcal{M}[f_i(x, \mu^2)](N) = \int_0^1 dx\, x^{N-1} f_i(x, \mu^2) \tag{2.3}$$

$$\gamma_{ij}(N, a_s(\mu^2)) = -\mathcal{M}[P_{ij}(x, a_s(\mu^2))](N) = -\int_0^1 dx\, x^{N-1} P_{ij}(x, a_s(\mu^2)) \tag{2.4}$$

the evolution equations become

$$\mu^2 \frac{\partial f_i(N, \mu^2)}{\partial \mu^2} = -\gamma_{ij}(N, a_s(\mu^2)) f_j(N, \mu^2)\,, \tag{2.5}$$

where the perturbative expansion of the anomalous dimensions is

$$\gamma_{ij}(N, a_s(\mu^2)) = a_s \gamma_{ij}^{(0)}(N) + a_s^2 \gamma_{ij}^{(1)}(N) + a_s^3 \gamma_{ij}^{(2)}(N) + a_s^4 \gamma_{ij}^{(3)}(N) + \mathcal{O}\left(a_s^5\right)\,. \tag{2.6}$$

The $(2n_f + 1) \times (2n_f + 1)$ matrix of anomalous dimensions has seven independent entries (see e.g. [45]), driving the evolution of various PDF combinations as follows:

- All nonsinglet combinations

$$q_{ij}^{\pm} = q_i^{\pm} - \bar{q}_j^{\pm}, \tag{2.7}$$
$$q_i^{\pm} = q_i \pm \bar{q}_i \tag{2.8}$$

  satisfy decoupled evolution equations with the same two anomalous dimension $\gamma_{\mathrm{ns},\pm}$; note that the plus and minus variants of $\gamma_{\mathrm{ns},\pm}$ start differing from each other already at NLO.

- The total valence combination

$$V = \sum_{i=1}^{n_f} q_i^- \tag{2.9}$$

  satisfies a decoupled evolution equation with an anomalous dimension

$$\gamma_{\mathrm{ns,v}} = \gamma_{\mathrm{ns,s}} + \gamma_{\mathrm{ns},-}\,; \tag{2.10}$$

  note that the flavor-independent "sea" contribution $\gamma_{\mathrm{ns,s}}$ starts being nonzero only at NNLO.

- The singlet combination

$$\Sigma = \sum_{i=1}^{n_f} q_i^+ \tag{2.11}$$

  mixes with the gluon

$$\mu^2 \frac{\partial}{\partial \mu^2} \begin{pmatrix} \Sigma(N, \mu^2) \\ g(N, \mu^2) \end{pmatrix} = -\begin{pmatrix} \gamma_{qq}(N, a_s(\mu^2)) & \gamma_{qg}(N, a_s(\mu^2)) \\ \gamma_{gq}(N, a_s(\mu^2)) & \gamma_{gg}(N, a_s(\mu^2)) \end{pmatrix} \begin{pmatrix} \Sigma(N, \mu^2) \\ g(N, \mu^2) \end{pmatrix}\,. \tag{2.12}$$

  The quark-quark entry of the anomalous dimension matrix can be further decomposed into nonsinglet and pure singlet contributions according to

$$\gamma_{qq} = \gamma_{\mathrm{ns},+} + \gamma_{qq,\mathrm{ps}}\,, \tag{2.13}$$

  where the pure singlet contribution $\gamma_{qq,\mathrm{ps}}$ starts at NLO.

There are thus seven independent contributions: three in the nonsinglet sector, $\gamma_{\rm ns,\pm}$ and $\gamma_{\rm ns,s}$, and four in the singlet sector, $\gamma_{qq,\rm ps}$, $\gamma_{qg}$, $\gamma_{gq}$, and $\gamma_{gg}$. In turn, each of these anomalous dimensions can be expanded according to Eq. (2.6). Our goal is to determine an approximate expression for the corresponding seven $\gamma_{ij}^{(3)}(N)$ N$^3$LO terms.

The information that can be exploited in order to achieve this goal comes from three different sources: (1) full analytic knowledge of contributions to the anomalous dimensions proportional to the highest powers of the number of flavors $n_f$; (2) large-$x$ and small-$x$ resummations provide all-order information on terms that are logarithmically enhanced by powers of $\ln(1-x)$ and $\ln x$ respectively; (3) analytic knowledge of a finite set of integer moments. We construct an approximation based on this information by first separating off the analytically known terms (1-2), then expanding the remainder on a set of basis functions and using the known moments to determine the expansion coefficients. Finally, we vary the set of basis functions in order to obtain an estimate of the uncertainties.

Schematically, we proceed as follows:

1. We include all terms in the expansion
$$\gamma_{ij}^{(3)}(N) = \gamma_{ij}^{(3,0)}(N) + n_f\gamma_{ij}^{(3,1)}(N) + n_f^2\gamma_{ij}^{(3,2)}(N) + n_f^3\gamma_{ij}^{(3,3)}(N)\,, \tag{2.14}$$
of the anomalous dimension in powers of $n_f$ that are fully or partially known analytically. We collectively denote such terms as $\gamma_{ij,n_f}^{(3)}(N)$.

2. We include all terms from large-$x$ and small-$x$ resummation, to the highest known logarithmic accuracy, including all known subleading power corrections in both limits. We denote these terms as $\gamma_{ij,N\to\infty}^{(3)}(N)$ and $\gamma_{ij,N\to 0}^{(3)}(N)$, $\gamma_{ij,N\to 1}^{(3)}(N)$ respectively. Possible double counting coming from the overlap of these terms with $\gamma_{ij,n_f}^{(3)}(N)$ is removed.

3. We write the approximate anomalous dimension matrix element $\gamma_{ij}^{(3)}(N)$ as the sum of the terms which are known exactly and a remainder $\widetilde{\gamma}_{ij}^{(3)}(N)$ according to
$$\gamma_{ij}^{(3)}(N) = \gamma_{ij,n_f}^{(3)}(N) + \gamma_{ij,N\to\infty}^{(3)}(N) + \gamma_{ij,N\to 0}^{(3)}(N) + \gamma_{ij,N\to 1}^{(3)}(N) + \widetilde{\gamma}_{ij}^{(3)}(N)\,. \tag{2.15}$$
We determine the remainder as a linear combination of a set of $n^{ij}$ interpolating functions $G_\ell^{ij}(N)$ (kept fixed) and $H_\ell^{ij}(N)$ (to be varied)
$$\widetilde{\gamma}_{ij}^{(3)}(N) = \sum_{\ell=1}^{n^{ij}-n_H} b_\ell^{ij} G_\ell^{ij}(N) + \sum_{\ell=1}^{n_H} b_{n^{ij}-2+\ell}^{ij} H_\ell^{ij}(N)\,, \tag{2.16}$$
with $n^{ij}$ equal to the number of known Mellin moments of $\gamma_{ij}^{(3)}(N)$. We determine the coefficients $b_\ell^{ij}$ by equating the evaluation of $\widetilde{\gamma}_{ij}^{(3)}(N)$ to the known moments of the splitting functions.

4. In the singlet sector, we take $n_H = 2$ and we make $\widetilde{N}_{ij}$ different choices for the two functions $H_\ell^{ij}(N)$, by selecting them out of a list of distinct basis functions (see Sect. 2.4 below). Thereby, we obtain $\widetilde{N}_{ij}$ expressions for the remainder $\widetilde{\gamma}_{ij}^{(3)}(N)$ and accordingly for the N$^3$LO anomalous dimension matrix element $\gamma_{ij}^{(3)}(N)$ through Eq. (2.15). These are used to construct the approximate anomalous dimension matrix and the uncertainty on it, in the way discussed in Sect. 2.2 below. In the nonsinglet sector instead, we take $n_H = 0$, i.e. we take a unique answer as our approximation, and we neglect the uncertainty on it, for reasons to be discussed in greater detail at the end of Sect. 2.3.

### 2.2 The approximate anomalous dimension matrix and its uncertainty

The procedure described in Sect. 2.1 provides us with an ensemble of $\widetilde{N}_{ij}$ different approximations to the N$^3$LO anomalous dimension, denoted $\gamma_{ij}^{(3),(k)}(N)$, $k = 1,\ldots \widetilde{N}_{ij}$ . Our best estimate for the approximate anomalous dimension is then their average
$$\gamma_{ij}^{(3)}(N) = \frac{1}{\widetilde{N}_{ij}}\sum_{k=1}^{\widetilde{N}_{ij}} \gamma_{ij}^{(3),(k)}(N). \tag{2.17}$$

We include the uncertainty on the approximation, and the ensuing uncertainty on N$^3$LO theory predictions, using the general formalism for the treatment of theory uncertainties developed in Refs. [34–36]. Namely, we consider the uncertainty on each anomalous dimension matrix element due to its incomplete knowledge as a source of uncertainty on theoretical predictions, uncorrelated from other sources of uncertainty, and neglecting possible correlations between our incomplete knowledge of each individual matrix element $\gamma_{ij}^{(3)}$. This uncertainty on the incomplete higher (N$^3$LO) order terms (incomplete higher order uncertainty, or IHOU) is then treated in the same way as the uncertainty due to missing higher order terms (missing higher order uncertainty, or MHOU).

Namely, we construct the shift of theory prediction for the $m$-th data point due to replacing the central anomalous dimension matrix element $\gamma_{ij}^{(3)}(N)$, Eq. (2.17), with each of the instances $\gamma_{ij}^{(3),(k)}(N)$, viewed as an independent nuisance parameter:

$$\Delta_m(ij,k) = T_m(ij,k) - \bar{T}_m, \tag{2.18}$$

where $\bar{T}_m$ is the prediction for the $m$-th datapoint obtained using the best estimate Eq. (2.17) for the full anomalous dimension matrix, while $T_m(ij,k)$ is the prediction obtained when the the $ij$ matrix element of our best estimate is replaced with the $k$-th instance $\gamma_{ij}^{(3),(k)}(N)$.

We then construct the covariance matrix over theory predictions for individual datapoints due to the IHOU on the $ij$ N$^3$LO matrix element as the covariance of the shifts $\Delta_m(ij,k)$ over all $\widetilde{N}_{ij}$ instances:

$$\text{cov}_{mn}^{(ij)} = \frac{1}{\widetilde{N}_{ij}-1}\sum_{k=1}^{\widetilde{N}_{ij}} \Delta_m(ij,k)\Delta_n(ij,k). \tag{2.19}$$

We recall that we do not associate an IHOU to the nonsinglet anomalous dimensions and we assume conservatively that there is no correlation between the different singlet anomalous dimension matrix elements. Thus we can write the total contribution to the theory covariance matrix due to IHOU as

$$\text{cov}_{mn}^{\text{IHOU}} = \text{cov}_{mn}^{(gg)} + \text{cov}_{mn}^{(gq)} + \text{cov}_{mn}^{(qg)} + \text{cov}_{mn}^{(qq)}. \tag{2.20}$$

The mean square uncertainty on the anomalous dimension matrix element itself is then determined, by viewing it as a pseudo-observable, as the variance

$$(\sigma_{ij}(N))^2 = \frac{1}{\widetilde{N}_{ij}-1}\sum_{k=1}^{\widetilde{N}_{ij}} \left(\gamma_{ij}^{(3),(k)}(N) - \gamma_{ij}^{(3)}(N)\right)^2. \tag{2.21}$$

### 2.3 aN$^3$LO anomalous dimensions: the nonsinglet sector

Information on the Mellin moments of nonsinglet anomalous dimensions is especially abundant, in that eight moments of $\gamma_{\text{ns},\pm}^{(3)}$ and nine moments of $\gamma_{\text{ns,s}}^{(3)}$ are known. An approximation based on this knowledge was given in Ref. [22]. More recently, further information on the small-$x$ behavior of $\gamma_{\text{ns},\pm}^{(3)}$ was derived in Ref. [23]. While for $\gamma_{\text{ns},s}^{(3)}$ we directly rely on the approximation of Ref. [22], which already includes all the available information, we construct an approximation to $\gamma_{\text{ns},\pm}^{(3)}$ based on the procedure described in Sect. 2.1, in order to include also this more recent information, and also as a warm-up for the construction of our approximation to the singlet sector anomalous dimension that we present in the next section.

Contributions to $\gamma_{\text{ns},\pm}^{(3)}$ proportional to $n_f^2$ and $n_f^3$ are known exactly [21] (in particular the $n_f^3$ contributions to $\gamma_{\text{ns},\pm}^{(3)}$ coincide), while $\mathcal{O}(n_f^0)$ and $\mathcal{O}(n_f)$ terms[1] are known in the large-$N_c$ limit [22] and we include these in $\gamma_{\text{ns},\pm,n_f}^{(3)}(N)$.

Small-$x$ contributions to $\gamma_{\text{ns},\pm}$ are double logarithmic, i.e. of the form $a_s^{n+1}\ln^{2n-k}(x)$, corresponding in Mellin space to poles of order $2n-k+1$ in $N=0$, i.e. $\frac{1}{N^{2n-k+1}}$, so at N$^3$LO we have $n=3$ and thus

$$P_{\text{ns},\pm}^{(3)}(x) = \sum_{k=1}^{6} c_{\text{ns},N\to 0}^{k}\ln^k(1/x) + O(x)\,. \tag{2.22}$$

[1]The $n_f C_F^3$ terms have also been published very recently [46], but we do not include them yet.

| | |
|---|---|
| $G_1^{\rm ns,\pm}(N)$ | $1$ |
| $G_2^{\rm ns,\pm}(N)$ | $\mathcal{M}[(1-x)\ln(1-x)](N)$ |
| $G_3^{\rm ns,\pm}(N)$ | $\mathcal{M}[(1-x)\ln^2(1-x)](N)$ |
| $G_4^{\rm ns,\pm}(N)$ | $\mathcal{M}[(1-x)\ln^3(1-x)](N)$ |
| $G_5^{\rm ns,\pm}(N)$ | $\frac{S_1(N)}{N^2}$ |
| $G_6^{\rm ns,\pm}(N)$ | $\frac{1}{(N+1)^2}$ |
| $G_7^{\rm ns,\pm}(N)$ | $\frac{1}{(N+1)^3}$ |
| $G_8^{\rm ns,+}(N)$, $G_8^{\rm ns,-}(N)$ | $\frac{1}{(N+2)}$, $\frac{1}{(N+3)}$ |

**Table 2.1.** The Mellin space interpolating functions $G_\ell^{\rm ns,\pm}(N)$ entering the parametrisation of the remainder term $\widetilde{\gamma}_{\rm ns\,\pm}^{(3)}(N)$ for the nonsinglet anomalous dimension expansion of Eq. (2.16).

The coefficients $c_{{\rm ns},\,N\to 0}^k$ are known [23] exactly up to NNLL accuracy ($k=4,\,5,\,6$), and approximately up to N$^6$LL ($k=1,\,2,\,3$). Hence, we let

$$\gamma_{{\rm ns},\pm,\,N\to 0}^{(3)}(N)=\sum_{k=1}^{6} c_{{\rm ns},\,N\to 0}^k(-1)^k\frac{k!}{N^{k+1}}. \tag{2.23}$$

Large-$x$ logarithmic contributions in the $\overline{\rm MS}$ scheme only appear in coefficient functions [47], and so the $x\to 1$ behaviour of splitting functions is provided by the cusp anomalous dimension $\sim\frac{1}{(1-x)_+}$, corresponding to a single $\ln(N)$ behavior in Mellin space as $N\to\infty$. This behavior is common to the pair of anomalous dimensions $\gamma_{{\rm ns},\pm}^{(3)}(N)$. Furthermore, several subleading power corrections as $N\to\infty$ can also be determined and we set

$$\gamma_{{\rm ns},\pm,\,N\to\infty}^{(3)}(N)=A_4^q S_1(N)+B_4^q+C_4^q\frac{S_1(N)}{N}+D_4^q\frac{1}{N}. \tag{2.24}$$

The coefficient of the $\ln(N)$ term $A_4^q$, is the quark cusp anomalous dimension [24]. The constant coefficient $B^q$ is determined by the integral of the nonsinglet splitting function, which was originally computed in [22] in the large-$N_c$ limit and recently updated to the full color expansion [25] as a result of computing different N$^3$LO cross sections in the soft limit. The coefficients of the terms suppressed by $1/N$ in the large-$N$ limit, $C^q$ and $D^q$, can be obtained directly from lower-order anomalous dimensions by exploiting large-$x$ resummation techniques [21]. The explicit expressions of $\gamma_{{\rm ns}\,\pm,\,N\to\infty}^{(3)}(N)$ and $\gamma_{{\rm ns}\,\pm,\,N\to 0}^{(3)}(N)$ are given in Appendix A.

The remainder terms, $\widetilde{\gamma}_{{\rm ns},\pm}^{(3)}(N)$, are expanded over the set of eight functions $G_\ell^{\rm ns,\pm}(N)$ listed in Table 2.1. The coefficients $b_\ell^{\rm ns,\pm}$ (defined in Eq.(2.16)) are determined by imposing that the values of the eight moments given in Ref. [22] be reproduced. The set of functions $G_\ell^{\rm ns,\pm}(N)$ is chosen to adjust the overall constant ($\ell=1$), model the large-$N$ behavior ($2\le\ell\le 5$) and model the small-$N$ behavior ($\ell=6,7$), consistent with the general analytic structure of fixed order anomalous dimensions [48]. Specifically, the large-$N$ functions are chosen as the logarithmically enhanced next-to-next-to-leading power terms ($\ln^k(N)/N^2$, $\ell=2,3,4,5$) and the small-$N$ functions are chosen as logarithmically enhanced subleading poles ($1/(N+1)^k$, $\ell=6,7$) and sub-subleading poles ($1/(N+2)$ or $1/(N+3)$, $\ell=8$). The last element, $\ell=8$, is chosen at a fixed distance from the lowest known moment, $N=2$ for $\gamma_{{\rm ns},+}^{(3)}(N)$ and $N=1$ for $\gamma_{{\rm ns},-}^{(3)}(N)$.

In Fig. 2.1 we plot the resulting splitting functions $P_{{\rm ns},\pm}^{(3)}(x)$, obtained by Mellin inversion of the anomalous dimension. We compare our approximation to the approximation of Ref. [22], for $\alpha_s=0.2$ and $n_f=4$, and also show the (exact) NNLO result for reference. Because the splitting function is a distribution at $x=1$ we plot $(1-x)P(x)$. The result of Ref. [22] also provides an estimate of the uncertainty related to the approximation, shown in the figure as a band, and we observe that this uncertainty is negligible except at very small $x$. As we include further constraints on the small-$x$ behavior, the uncertainty on the approximation becomes negligible, and consequently, as mentioned in Sect. 2.1 above, we take $n^{\rm ns,\pm}=k^{\rm ns,\pm}$ in Eq. (2.16).

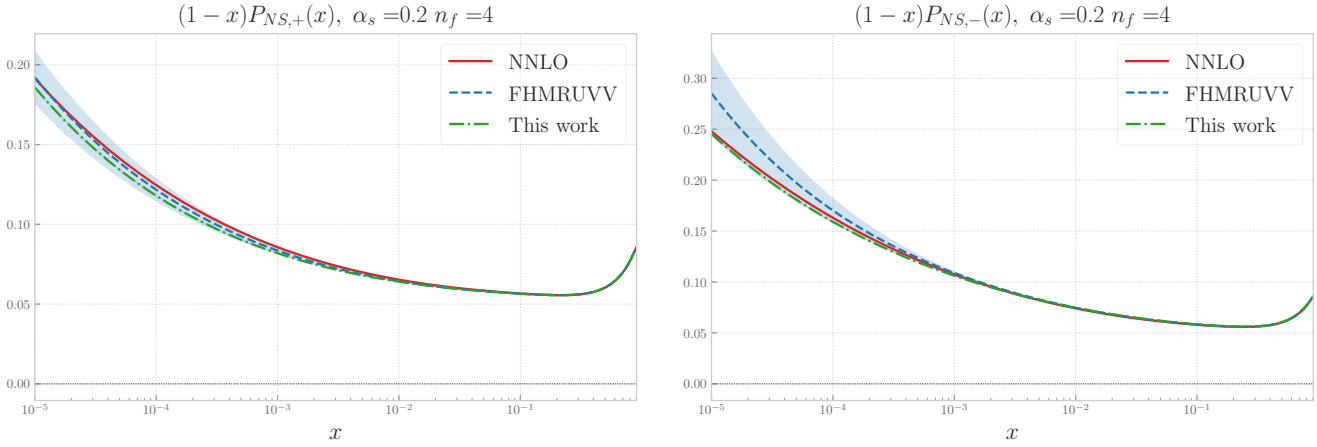



**Figure 2.1.** The aN$^3$LO nonsinglet splitting functions $(1-x)P_{\rm ns,+}(x,\alpha_s)$ and $(1-x)P_{\rm ns,-}(x,\alpha_s)$, evaluated as a function of $x$ for $n_f = 4$ and $\alpha_s = 0.2$ in our approximation compared to the previous approximation of Ref. [22] (denoted FHMRUVV), for which the approximation uncertainty, as estimated by its authors, is also displayed. For comparison, the (exact) NNLO result is also shown.

## 2.4 aN$^3$LO anomalous dimensions: singlet sector

In order to determine the singlet-sector anomalous dimension matrix entering Eq. (2.12), we must determine $\gamma_{qq,\rm ps}$ that, together with the previously determined nonsinglet anomalous dimension, contributes to the $qq$ entry, Eq. (2.13), and then also the three remaining matrix elements $\gamma_{qg}$, $\gamma_{gq}$, and $\gamma_{gg}$.

For all matrix elements, the leading large-$n_f$ $\mathcal{O}(n_f^3)$ contributions in Eq. (2.14) are known analytically [21], while for $\gamma_{qq,\rm ps}$ [49] and $\gamma_{gq}$ [31] the $\mathcal{O}(n_f^2)$ contributions are also known and we include all of them in $\gamma^{(3)}_{ij,n_f}(N)$.

Small-$x$ contributions in the singlet sector include, on top of the double-logarithmic contributions $a_s^{n+1}\ln^{2n-k}(x)$ that are present in the nonsinglet case, also single-logarithmic contributions $a_s^{n+1}\frac{1}{x}\ln^n(x)$. In Mellin space, this means that on top of order $2n-k+1$ subleading poles in $N=0$, there are also leading poles in $N=1$ of order $n-k+1$, i.e. $\frac{1}{(N-1)^{n-k+1}}$. The leading-power single logarithmic contributions can be extracted from the leading [50–54] and next-to-leading [55–59] high-energy resummation at LL$x$ [60] and NLL$x$ [61–63] accuracy. This allows for a determination of the coefficients of the leading $\frac{1}{(N-1)^4}$ and next-to-leading $\frac{1}{(N-1)^3}$ contributions to $\gamma^{(3)}_{gg}$ and of the next-to-leading $\frac{1}{(N-1)^3}$ contributions to $\gamma^{(3)}_{qg}$. The remaining entries can be obtained from these by using the color-charge (or Casimir scaling) relation $\gamma_{iq} = \frac{C_F}{C_A}\gamma_{ig}$ [63,64]. Hence, we set

$$\gamma^{(3)}_{gg,\,N\to1}(N) = c^4_{gg,\,N\to1}\frac{1}{(N-1)^4} + c^3_{gg,\,N\to1}\frac{1}{(N-1)^3}; \tag{2.25}$$

$$\gamma^{(3)}_{qg,\,N\to1}(N) = c^3_{qg,\,N\to1}\frac{1}{(N-1)^3}; \tag{2.26}$$

$$\gamma^{(3)}_{iq,\,N\to1}(N) = \frac{C_F}{C_A}\gamma^{(3)}_{ig,\,N\to1}(N), \quad i=q,\,g. \tag{2.27}$$

Although only the leading pole of $\gamma_{gq}$ satisfies Eq. (2.27) exactly, at NNLO this relation is only violated at the sub-percent level [65], so this is likely to be an adequate approximation also at this order: this approximation is also adopted in Ref. [30]. An important observation is that both NLO and NNLO coefficients of the leading poles, $\frac{1}{(N-1)^2}$ and $\frac{1}{(N-1)^3}$ respectively, vanish accidentally. Hence, at N$^3$LO the leading poles contribute for the first time beyond leading order. The subleading poles can be determined up to NNLL accuracy [23] and, thus, fix the coefficients of the $\frac{1}{N^7}$, $\frac{1}{N^6}$ and $\frac{1}{N^5}$ subleading poles for all entries of the singlet anomalous dimension matrix. All these contributions are included in $\gamma^{(3)}_{ij,N\to1}(N)$ and $\gamma^{(3)}_{ij,N\to0}(N)$.

In the singlet sector, large-$x$ contributions whose Mellin transform is not suppressed in the large-$N$ limit only appear in the diagonal $qq$ and $gg$ channels. In the quark channel these are already included, through Eq. (2.13) in $\gamma^{(3)}_{\rm ns,+,\,N\to\infty}(N)$, according to Eq. (2.24), while $\gamma^{(3)}_{qq,\rm ps}$ is suppressed in this limit. In the gluon-to-gluon channel they take the same form as in the nonsinglet and diagonal quark channel. Hence, we expand,

as in Eq. (2.24),

$$\gamma^{(3)}_{gg,\,N\to\infty}(N) = A^g_4 S_1(N) + B^g_4 + C^g_4 \frac{S_1(N)}{N} + D^g_4 \frac{1}{N}\,. \tag{2.28}$$

The coefficients $A^g_4$, $B^g_4$, $C^g_4$ and $D^g_4$ are the counterparts of those of Eq. (2.24): the gluon cusp anomalous dimension was determined in Ref. [24] and the constant in Ref. [25], while the $C^g_4$ and $D^g_4$ coefficients can be determined using results from Refs. [30, 66].

Off-diagonal $qg$ and $qg$ splitting functions have logarithmically enhanced next-to-leading power behavior at large-$x$:

$$P^{(3)}_{ij}(x) = \sum_{k=0}^{6}\sum_{l=0}^{\infty} c^{k,l}_{ij,\,N\to\infty}(1-x)^l \ln^k(1-x). \tag{2.29}$$

For $l=0$ the coefficients of the higher logs $k=4,5,6$ can be determined from N$^3$LO coefficient functions, based on a conjecture [27, 67] on the large-$x$ behavior of the physical evolution kernels that give the scale dependence of structure functions. The coefficient with the highest power $k=6$ cancels and thus we let

$$\gamma^{(3)}_{gq,\,N\to\infty}(N) = \sum_{k=4}^{5} c^{k,0}_{gq,\,N\to\infty} \mathcal{M}\left[\ln^k(1-x)\right](N), \tag{2.30}$$

$$\gamma^{(3)}_{qg,\,N\to\infty}(N) = \sum_{k=4}^{5} c^{k,0}_{qg,\,N\to\infty} \mathcal{M}\left[\ln^k(1-x)\right](N) + c^{k,1}_{qg,\,N\to\infty} \mathcal{M}\left[(1-x)\ln^k(1-x)\right](N), \tag{2.31}$$

$$\tag{2.32}$$

where in $\gamma^{(3)}_{qg,\,N\to\infty}$ we have retained also the $l=1$ terms [29].

Finally, the pure singlet quark-to-quark splitting function starts at next-to-next-to-leading power as $x\to 1$, i.e. it behaves as $(1-x)\ln^k(1-x)$, with $k\le 4$. The coefficients of the higher logs $k=3,4$ can be extracted by expanding the $x=1$ expressions from Refs. [27, 28]. Hence, we let

$$\gamma^{(3)}_{qq,\mathrm{ps},\,N\to\infty}(N) = \sum_{k=3}^{4}\left[c^{k,1}_{qq,\mathrm{ps},\,N\to\infty} \mathcal{M}\left[(1-x)\ln^k(1-x)\right](N) + c^{k,2}_{qq,\mathrm{ps},\,N\to\infty} \mathcal{M}\left[(1-x)^2\ln^k(1-x)\right](N)\right] \tag{2.33}$$

Note that for the $qq$ and $qg$ entries we also include the (known) next-to-leading power contributions, while we do not include them for $gq$ and $gg$ because for these anomalous dimension matrix elements a significantly larger number of higher Mellin moments is known, hence there is no risk that the inclusion of these contributions could contaminate the intermediate $x$ region where they are not necessarily dominant. The explicit expressions of $\gamma^{(3)}_{\mathrm{ns}\pm,\,N\to\infty}(N)$, $\gamma^{(3)}_{\mathrm{ns}\pm,\,N\to 0}(N)$ and $\gamma^{(3)}_{\mathrm{ns}\pm,\,N\to 1}(N)$ are all given in Appendix A.

As discussed in Sect. 2.1, the remainder contribution $\widetilde{\gamma}^{(3)}_{ij}(N)$, Eq. (2.16), is determined by expanding it over a set of basis functions and determining the expansion coefficients by demanding that the four components of the singlet anomalous dimension matrix reproduce the known moments. Specifically, these are the four moments computed in Ref. [26], the six additional moments for $\gamma_{qq,\mathrm{ps}}$ and $\gamma_{qg}$ computed in Ref. [28] and Ref. [29] respectively, and the additional moment $N=10$ for $\gamma_{gg}$ and $\gamma_{gq}$ evaluated in Ref. [30]. These constraints automatically implement momentum conservation:

$$\begin{aligned}\gamma_{qg}(N=2) + \gamma_{gg}(N=2) &= 0\,,\\ \gamma_{qq}(N=2) + \gamma_{gq}(N=2) &= 0\,.\end{aligned} \tag{2.34}$$

The set of basis functions is chosen based on the idea of constructing an approximation that reproduces the singularity structure of the Mellin transform of the anomalous dimension viewed as analytic functions in $N$ space [48], hence corresponding to the leading and subleading (i.e. rightmost) $N$-space poles with unknown coefficients as well as the leading unknown large-$N$ behavior. Specifically, the functions $G^{ij}$ are chosen as follows.

1. The function $G^{ij}_1(N)$ reproduces the leading unknown contribution in the large-$N$ limit, i.e. the unknown term in Eq. (2.29) with highest $k$ and lowest $l$.

2. The functions $G^{ij}_2(N)$ and $G^{ij}_3(N)$ reproduce the first two leading unknown contributions in the small-$N$ limit, i.e. the unknown $\frac{1}{(N-1)^k}$ leading poles with highest and next-to-highest values of $k$, i.e. $k=2$

| | | |
|---|---|---|
| $\gamma_{gg}^{(3)}(N)$ | $G_1^{gg}(N)$ | $\mathcal{M}[(1-x)\ln^3(1-x)](N)$ |
| | $G_2^{gg}(N)$ | $\frac{1}{(N-1)^2}$ |
| | $G_3^{gg}(N)$ | $\frac{1}{N-1}$ |
| | $\{H_1^{gg}(N),\ H_2^{gg}(N)\}$ | $\frac{1}{N^4},\ \frac{1}{N^3},\ \frac{1}{N^2},\ \frac{1}{N+1},\ \frac{1}{N+2},\ \mathcal{M}[(1-x)\ln^2(1-x)](N),\ \mathcal{M}[(1-x)\ln(1-x)](N)$ |
| $\gamma_{gq}^{(3)}(N)$ | $G_1^{gq}(N)$ | $\mathcal{M}[\ln^3(1-x)](N)$ |
| | $G_2^{gg}(N)$ | $\frac{1}{(N-1)^2}$ |
| | $G_3^{gq}(N)$ | $\frac{1}{N-1}$ |
| | $\{H_1^{gq}(N),\ H_2^{gq}(N)\}$ | $\frac{1}{N^4},\ \frac{1}{N^3},\ \frac{1}{N^2},\ \frac{1}{N+1},\ \frac{1}{N+2},\ \mathcal{M}[\ln^2(1-x)](N),\ \mathcal{M}[\ln(1-x)](N)$ |
| $\gamma_{qg}^{(3)}(N)$ | $G_1^{qg}(N)$ | $\mathcal{M}[\ln^3(1-x)](N)$ |
| | $G_2^{qg}(N)$ | $\frac{1}{(N-1)^2}$ |
| | $G_3^{qg}(N)$ | $\frac{1}{N-1}-\frac{1}{N}$ |
| | $G_{4,\ldots,8}^{qg}(N)$ | $\frac{1}{N^4},\ \frac{1}{N^3},\ \frac{1}{N^2},\ \frac{1}{N},\ \mathcal{M}[\ln^2(1-x)](N)$ |
| | $\{H_1^{qg}(N),\ H_2^{qg}(N)\}$ | $\mathcal{M}[\ln(x)\ln(1-x)](N),\ \mathcal{M}[\ln(1-x)](N),\ \mathcal{M}[(1-x)\ln^3(1-x)](N)$<br>$\mathcal{M}[(1-x)\ln^2(1-x)](N),\ \mathcal{M}[(1-x)\ln(1-x)](N),\ \frac{1}{1+N}$ |
| $\gamma_{qq,\rm ps}^{(3)}(N)$ | $G_1^{qq,\rm ps}(N)$ | $\mathcal{M}[(1-x)\ln^2(1-x)](N)$ |
| | $G_2^{qq,\rm ps}(N)$ | $-\frac{1}{(N-1)^2}+\frac{1}{N^2}$ |
| | $G_3^{qq,\rm ps}(N)$ | $-\frac{1}{(N-1)}+\frac{1}{N}$ |
| | $G_{4,\ldots,8}^{qq,\rm ps}(N)$ | $\frac{1}{N^4},\ \frac{1}{N^3},\ \mathcal{M}[(1-x)\ln(1-x)](N)$<br>$\mathcal{M}[(1-x)^2\ln(1-x)^2](N),\ \mathcal{M}[(1-x)\ln(x)](N)$ |
| | $\{H_1^{qq,\rm ps}(N),\ H_2^{qq,\rm ps}(N)\}$ | $\mathcal{M}[(1-x)(1+2x)](N),\ \mathcal{M}[(1-x)x^2](N),$<br>$\mathcal{M}[(1-x)x(1+x)](N),\ \mathcal{M}[(1-x)](N)$ |

**Table 2.2.** The set of basis functions $G_\ell^{ij}(N)$ and $H_\ell^{ij}(N)$ used to parametrize the singlet sector remainder anomalous dimensions matrix elements $\widetilde{\gamma}_{ij}^{(3)}(N)$ according to Eq. (2.16).

and $k=1$. For $\gamma_{qq,\rm ps}$ and $\gamma_{qg}$ a subleading small-$x$ pole with the same power and opposite sign is added to the leading pole with respectively $k=1$, 2 and $k=1$, so as to leave unaffected the respective large-$x$ leading power behavior Eqs. (2.31-2.33).

3. For $\gamma_{qq,\rm ps}$ and $\gamma_{qg}$, for which an additional five moments are known, the functions $G_{4,\ldots,8}^{qj}(N)$ reproduce subleading small- and large-$N$ terms.

The functions $H^{ij}$ are chosen as follows.

1. The functions $H_1^{gj}(N), H_2^{gj}(N)$ in the gluon sector, where only five moments are known exactly, are chosen to reproduce subleading small- and large-$N$ terms, i.e. similar to $G_{4,\ldots,8}^{qj}(N)$.

2. The functions $H_1^{qg}(N), H_2^{qg}(N)$ are chosen as subleading and next-to-leading power large-$x$ terms and the remaining unknown leading small-$N$ pole.

3. The functions $H_1^{qq,\rm ps}(N), H_2^{qq,\rm ps}(N)$ are chosen as low-order polynomials, i.e., sub-subleading small-$x$ poles.

The number of basis functions is greater for anomalous dimension matrix elements for which less exact information is available: 7 in the gluon sector (i.e. $gg$ and $gq$), 6 for the $qg$ entry and 4 for the pure singlet

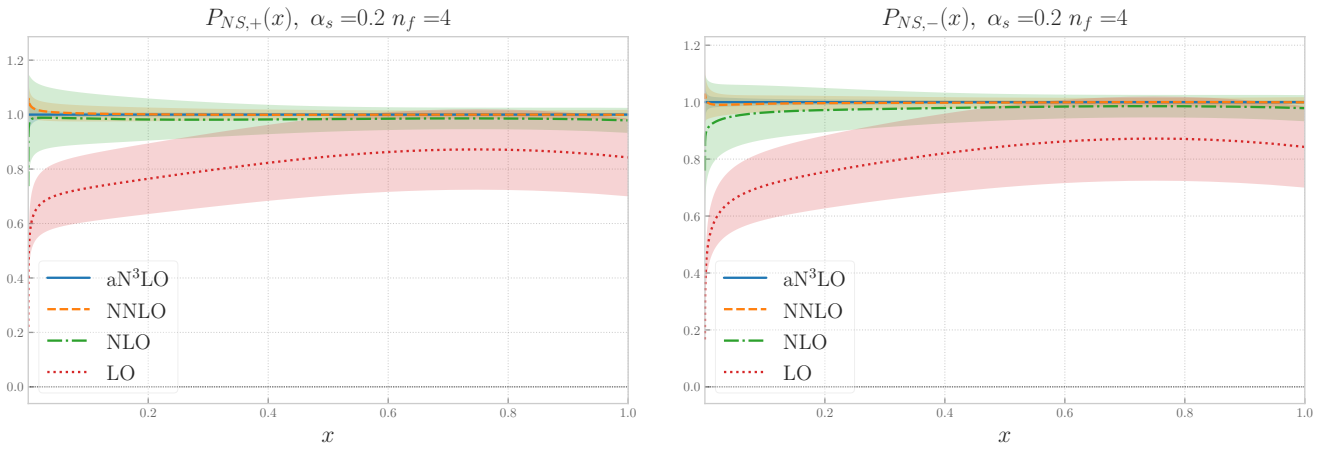



**Figure 2.2.** The nonsinglet splitting functions at LO, NLO, NNLO, and aN$^3$LO, normalized to the aN$^3$LO central value and with a linear scale on the $x$ axis. In each case we shown also the uncertainty due to missing higher orders (MHOU) estimated by scale variation according to Refs. [35, 36].

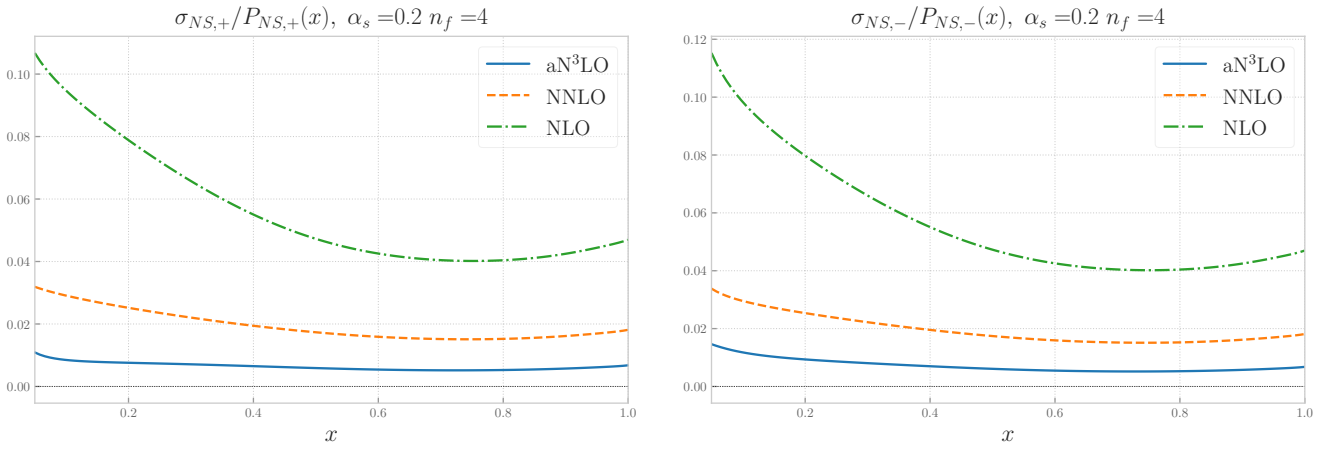



**Figure 2.3.** The relative size of the uncertainty due to missing higher orders (MHOU) on the splitting functions of Fig. 2.2.

entry. For the $gg$ entry two combinations are discarded as they lead to unstable (oscillating) results and we thus end up with $\widetilde{N}_{gg} = 19$, $\widetilde{N}_{gq} = 21$, $\widetilde{N}_{qg} = 15$, and $\widetilde{N}_{qq} = 6$ different parametrizations. The full set of basis functions $G^{ij}$ and $H^{ij}$ is listed in Table 2.2.

Upon combining the exactly known contributions with the $\widetilde{N}_{ij}$ remainder terms according to Eq. (2.15) we end up with an ensemble of $\widetilde{N}_{ij}$ instances of $\gamma_{ij}^{(3),(k)}(N)$ for each singlet anomalous dimension matrix element and the final matrix elements $\gamma_{ij}^{(3)}(N)$ and their uncertainties $\sigma_{ij}(N)$ are computed using Eqs. (2.17) and (2.21) respectively.

## 2.5 Results: aN$^3$LO splitting functions

We now present the aN$^3$LO splitting functions constructed following the procedure described in Sects. 2.1–2.4. The nonsinglet result, already compared in Fig. 2.1 to the previous approximation of Ref. [22], is shown in Fig. 2.2 at the first four perturbative orders as a ratio to the aN$^3$LO result. For each order we include the MHOU determined by scale variation according to Refs. [35, 36] and recall that there are no IHOU in the nonsinglet sector. As the nonsinglet splitting function are subdominant at small $x$ we only show the plot with a linear scale in $x$. The relative size of the MHOU is shown in Fig. 2.3.

Inspecting Figs. 2.2 and 2.3 reveals good perturbative convergence[2] for all values of $x$. Specifically, the differences between two subsequent perturbative orders are reduced as the accuracy of the calculation increases, and, correspondingly, the MHOUs associated to factorization scale variations decrease with the

[2]Here and henceforth by "convergence" we mean that the size of the missing N$^4$LO corrections is negligible compared to the target accuracy of theoretical predictions, i.e. at the sub-percent level.

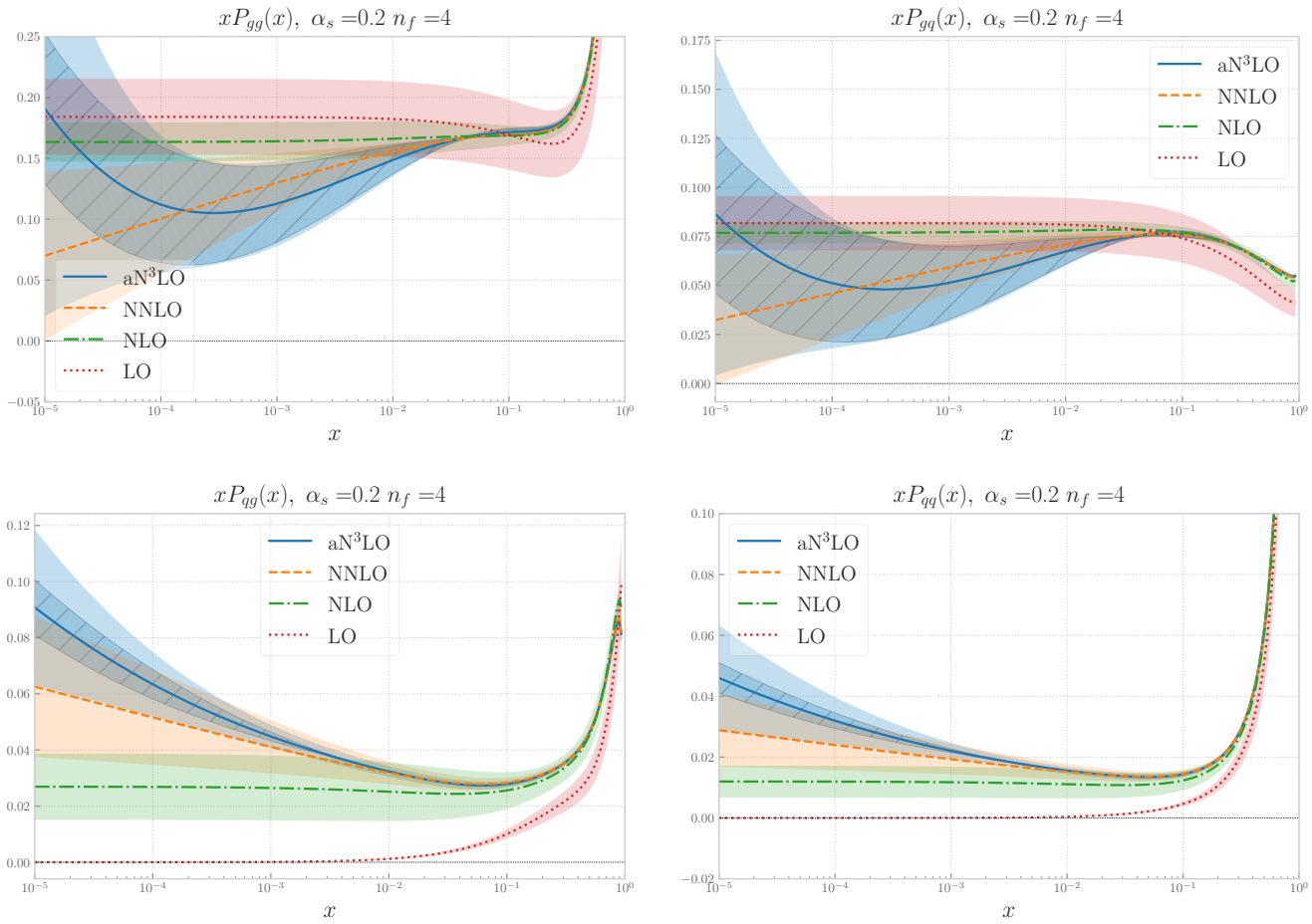



**Figure 2.4.** The singlet matrix of splitting functions $xP_{ij}$ at LO, NLO, NNLO and aN$^3$LO. From left to right and from top to bottom the $gg$, $gq$, $qg$ and $qq$ entries are shown. The MHOU estimated by scale variation is shown to all orders. At aN$^3$LO the dark blue band corresponds to IHOU only, while the light blue band is the sum in quadrature of IHOU and MHOU.

perturbative accuracy. Indeed, the MHOU appears to reproduce well the observed behaviour of the higher orders, with overlapping uncertainty bands between subsequent orders except at LO at the smallest $x$ values.

Based on these results it is clear that in the nonsinglet sector the N$^3$LO contribution to the splitting function is essentially negligible except at the smallest $x$ values, as shown in Fig. 2.1. Consequently, for all practical purposes we can consider the current approximation to the nonsinglet anomalous dimension to be essentially exact, and with negligible MHOU.

The situation in the singlet sector is more challenging. The singlet matrix of splitting functions is shown in Figs. 2.4 and 2.5, respectively with a logarithmic or linear scale on the $x$ axis. Because the diagonal splitting functions are distributions at $x = 1$ in the linear scale plots we display $x(1-x)P_{ii}$. The corresponding relative size of the MHOU is shown in Fig. 2.6 for the first four perturbative orders, along with the IHOU on the aN$^3$LO result, determined using Eq. (2.21).

A different behaviour is observed for the quark sector $P_{qi}$ and for the gluon sector $P_{gi}$. In the quark sector, the MHOU decreases with perturbative order for all $x$, but it remains sizable at aN$^3$LO for essentially all $x$, of order 5% for $10^{-2} \lesssim x \lesssim 10^{-1}$. In the gluon sector instead for $x \gtrsim 0.03$ the MHOU is negligible, but at smaller $x$ it grows rapidly, and in fact at very small $x$ it becomes larger than the NLO MHOU. This is due to the presence of leading small-$x$ logarithms, Eq. (2.25), which are absent at NLO. In fact the true gluon-sector MHOU at very small $x$ is likely to be underestimated by scale variation, because while it generates the fourth-order leading pole present in the N$^4$LO (the fifth-order pole vanishes), it fails to generate the sixth-order pole known to be present in the N$^5$LO splitting function.

We now turn to the IHOU and find again contrasting behaviour in the different sectors. In the quark sector, thanks to the large number of known Mellin moments and the copious information on the large-$x$ limit, the IHOU are significantly smaller than the MHOU, by about a factor three, and become negligible for $x \gtrsim 10^{-2}$. In the gluon sector instead the IHOU, while still essentially negligible for $x \gtrsim 0.1$, is larger

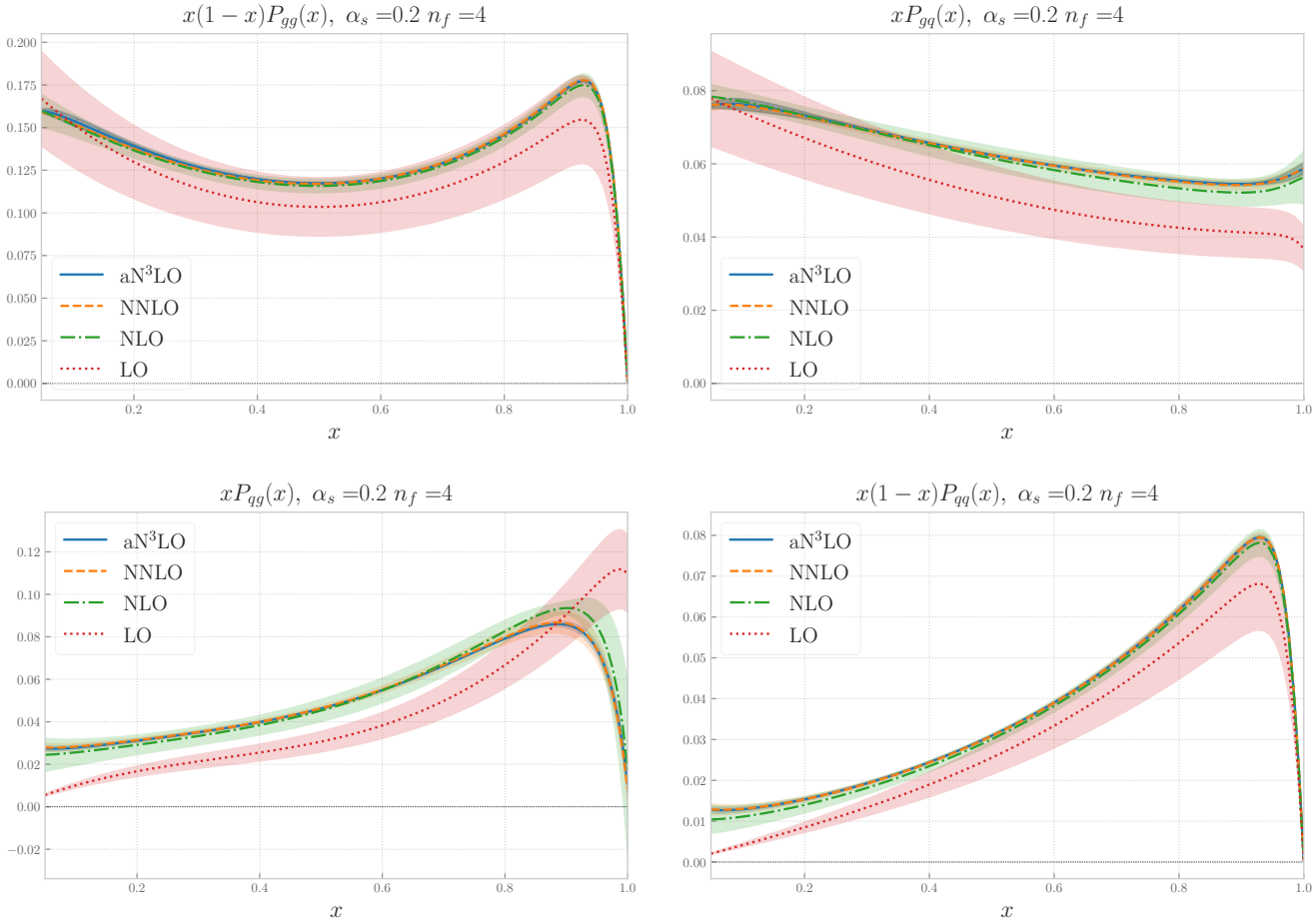



**Figure 2.5.** Same as Fig. 2.4 with a linear scale on the $x$ axis, and plotting $(1-x)xP_{ii}$ for diagonal entries.

than the MHOU except at very small $x \lesssim 10^{-4}$ where the MHOU dominates.

Consequently, for all matrix elements at large $x \gtrsim 0.1$ the behaviour of the singlet is similar to the behaviour of the nonsinglet: IHOU and MHOU are both negligible, meaning that aN$^3$LO results are essentially exact, and the perturbative expansion has essentially converged, see Fig. 2.5. At smaller $x$, while the aN$^3$LO and NNLO results agree within uncertainties, the uncertainties on the aN$^3$LO are sizable, dominated by MHOUs in the quark channel and by IHOUs in the gluon channel.

In the singlet sector the most dramatic impact of the aN$^3$LO correction is at small $x$. It is thus interesting to compare the aN$^3$LO singlet splitting functions with those obtained by the resummation of leading and next-to-leading order small-$x$ logarithms of Ref. [68], namely the two highest powers of $\ln x$ contained in the N$^3$LO result; this comparison is shown in Fig. 2.7. The agreement of all four entries $xP_{gg}$, $xP_{gq}$, $xP_{qg}$ and $xP_{qq}$ is remarkably good and well within the uncertainties in the two approaches. In particular the dip in $xP_{gg}$ at intermediate $x$ at aN$^3$LO (albeit with significant IHOU) is also a feature of the resummation. This is nontrivial, as the resummation includes only the asymptotic LL$x$ and NLL$x$ singularities at $N = 1$, but none of the subleading results incorporated at aN$^3$LO. Instead, it uses a symmetrization which resums collinear and anti-collinear logarithms in the small-$x$ expansion, and the effects of running coupling which change the nature of the small-$x$ singularity (from a fourth order pole at $N = 1$ in the fixed order N$^3$LO result to a simple pole a little further to the right on the real axis).

That both the resummed and fixed order approaches converge to very similar results, at least in the range of $x$ relevant for HERA and LHC, is very reassuring. It shows that in a global fit with current data, while NLL$x$ resummation significantly improves the quality of a fixed order NNLO fit [69], the same improvement should also be seen by adding aN$^3$LO corrections. Thus to find evidence for small-$x$ resummation at aN$^3$LO, it will probably be necessary to go to yet smaller values of $x$, e.g. below $10^{-5}$, where the fixed order and resummed results will eventually diverge again.

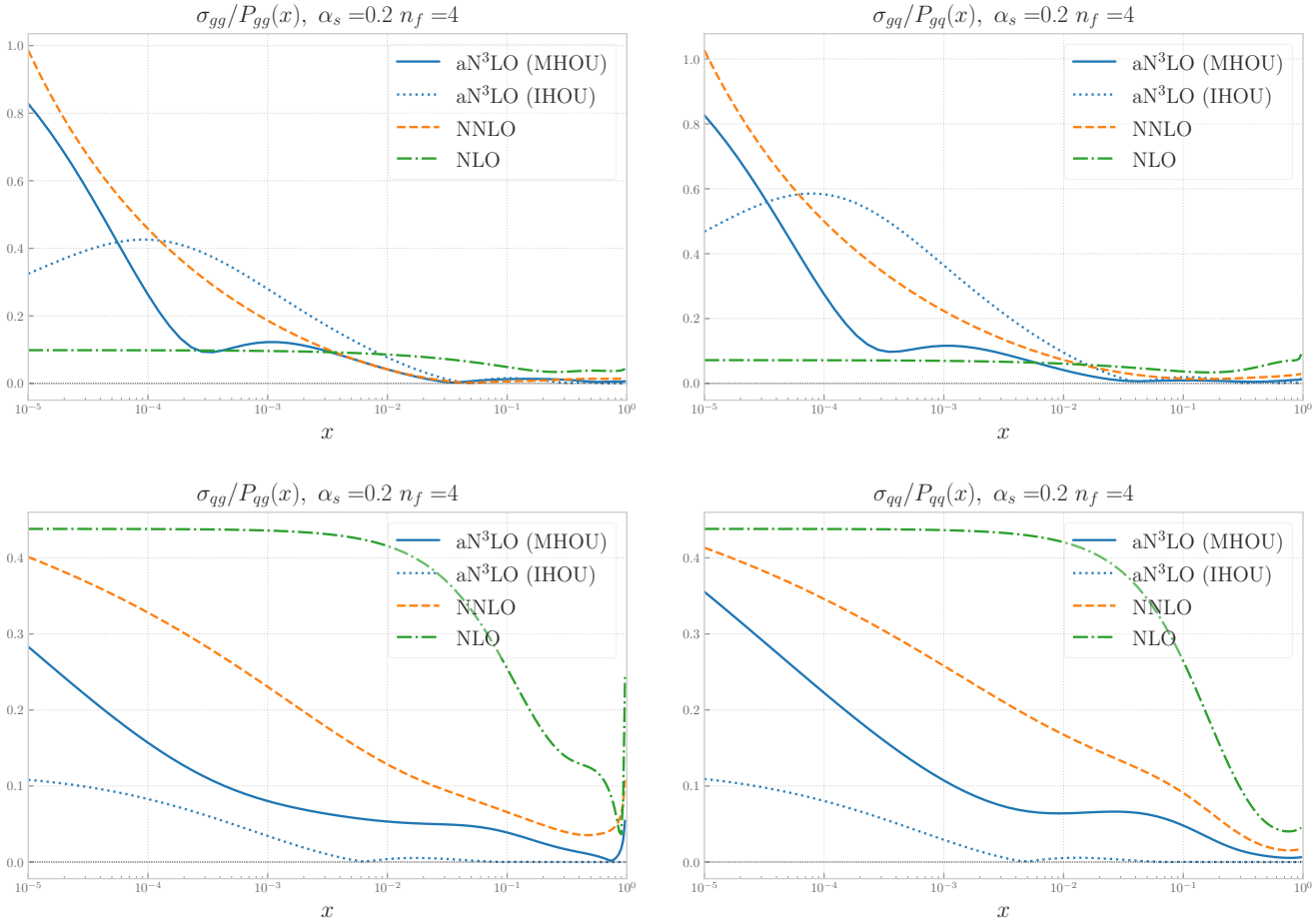



**Figure 2.6.** Same as Fig. 2.3 for the singlet splitting function matrix elements. At NLO and NNLO we show the MHOU, while at aN$^3$LO we also show the IHOU.

## 2.6 Results: aN$^3$LO evolution

The aN$^3$LO anomalous dimensions discussed in the previous sections have been implemented in the Mellin-space open-source evolution code EKO [40] which enters the new pipeline [70] adopted by NNPDF in order to produce theory predictions used for PDF determination. The parametrization is expressed in terms of a basis of Mellin space functions which are numerically efficient to evaluate. In order to achieve full aN$^3$LO accuracy, in addition to the anomalous dimensions, the four-loop running of the strong coupling constant $\alpha_s(Q)$ and the N$^3$LO matching conditions dictating the transitions between schemes with different numbers of active quark flavor have also been implemented.

The N$^3$LO matching conditions have been presented in Ref. [71] and subsequently computed analytically in Refs. [72–81]. The exception is the $a_{Hg}^{(3)}$ entry of the matching condition matrix, which is still unknown[3] and which instead is parametrized using the first 5 known moments [71] and the LL$x$ contribution as done in Ref. [83]. Also these matching conditions are implemented in EKO and thus it is possible to assess the impact of the inclusion of aN$^3$LO terms on perturbative evolution.

In Fig. 2.8 we compare the result of evolving a fixed set of PDFs from $Q_0 = 1.65$ GeV up to $Q = 100$ GeV at NLO, NNLO, and aN$^3$LO. We take as input the NNPDF4.0NNLO PDF set, and show results normalized to the aN$^3$LO evolution. Results are shown for all the combinations that evolve differently, as discussed in Sect. 2.1, namely the singlet, gluon, total valence and nonsinglet $\pm$ combinations, with a logarithmic scale on the $x$ axis for the singlet sector and a linear scale for the valence and nonsinglet combinations.

In all cases the perturbative expansion appears to have converged everywhere, with almost no difference between NNLO and aN$^3$LO except at small $x \lesssim 10^{-3}$, where singlet evolution is weaker at aN$^3$LO than at NNLO due to the characteristic dip seen in the gluon sector splitting functions of Fig. 2.4. Because the gluon-driven small-$x$ rise dominates small-$x$ evolution this is a generic feature of all quark and gluon PDFs in this small-$x$ region. It is interesting to observe that this is an all-order feature that persists upon small-$x$

[3]The terms recently computed in Ref. [82] are not yet included and left for future updates.

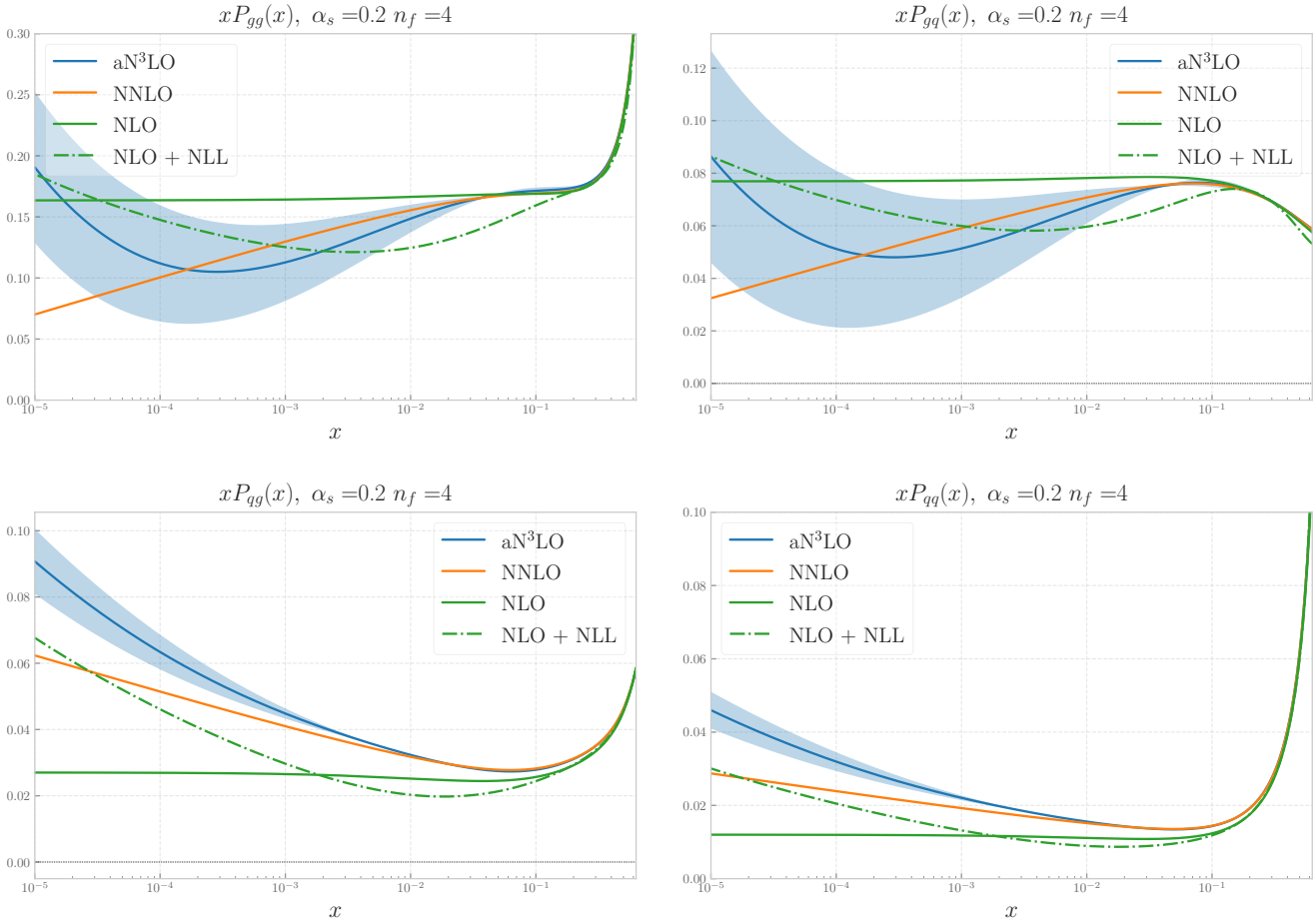



**Figure 2.7.** Comparison of the NLO, NNLO, and aN$^3$LO splitting functions (same as in Fig. 2.4) to the small-$x$ resummed NLO+NLL result of Ref. [68]. Only the IHOU on the aN$^3$LO result is shown.

resummation, as already discussed at the end of Sect. 2.5 and seen in Fig. 2.7. In fact, the total theory uncertainty at aN$^3$LO is at the sub-precent level for all $x \gtrsim 10^{-3}$. Hence, not only has the MHOU become negligible, but also the effect of IHOU on PDF evolution is only significant at small $x$.

## 2.7 Comparison to other groups

We finally compare our approximation of the N$^3$LO splitting functions to other recent results from Refs. [28–30, 32]. While the approach of Refs. [28–30] (FHMRUVV, henceforth) is very similar to our own, with differences only due to details of the choice of basis functions, a rather different approach is adopted in Ref. [32] (MSHT20, henceforth). There, the approximation is constructed from similar theoretical constraints (small-$x$, large-$x$ coefficients and Mellin moments), but supplementing the parametrization with additional nuisance parameters, which control the uncertainties arising from unknown N$^3$LO terms. However, these approximations are taken as a prior, and the nuisance parameters are fitted to the data along with the PDF parameters. The best-fit values of the parameters determine the posterior for the splitting function, and their uncertainties are interpreted as the final IHOU on it. A consequence of this procedure is that the posterior can reabsorb not only N$^3$LO corrections, but any other missing contribution, of theoretical or experimental origin.

The comparison is presented in Fig. 2.9, for all the four singlet splitting functions. For the MSHT20 results both prior and posterior are displayed. As expected, excellent agreement is found with the FHMRUVV result, for all splitting functions and for all $x$, especially for the $P_{qg}$ and $P_{qq}$ splitting functions, for which the highest number of Mellin moments is known. Good qualitative agreement is also found for $P_{gq}$ and $P_{gg}$, although at small $x$ IHOUs are larger and consequently central values differ somewhat more, though still in agreement within uncertainties. Coming now to MSHT20 results, good agreement is found with the prior, except for $P_{gq}$, for which MSHT20 shows a small-$x$ dip accompanied by a large-$x$ bump. The different small-$x$ behaviour is likely due to the fact that MSHT20 do not enforce the color-charge relation Eq. (2.27) at NLL$x$, with the large-$x$ bump then following from the constraints Eq. (2.34). Also, in the

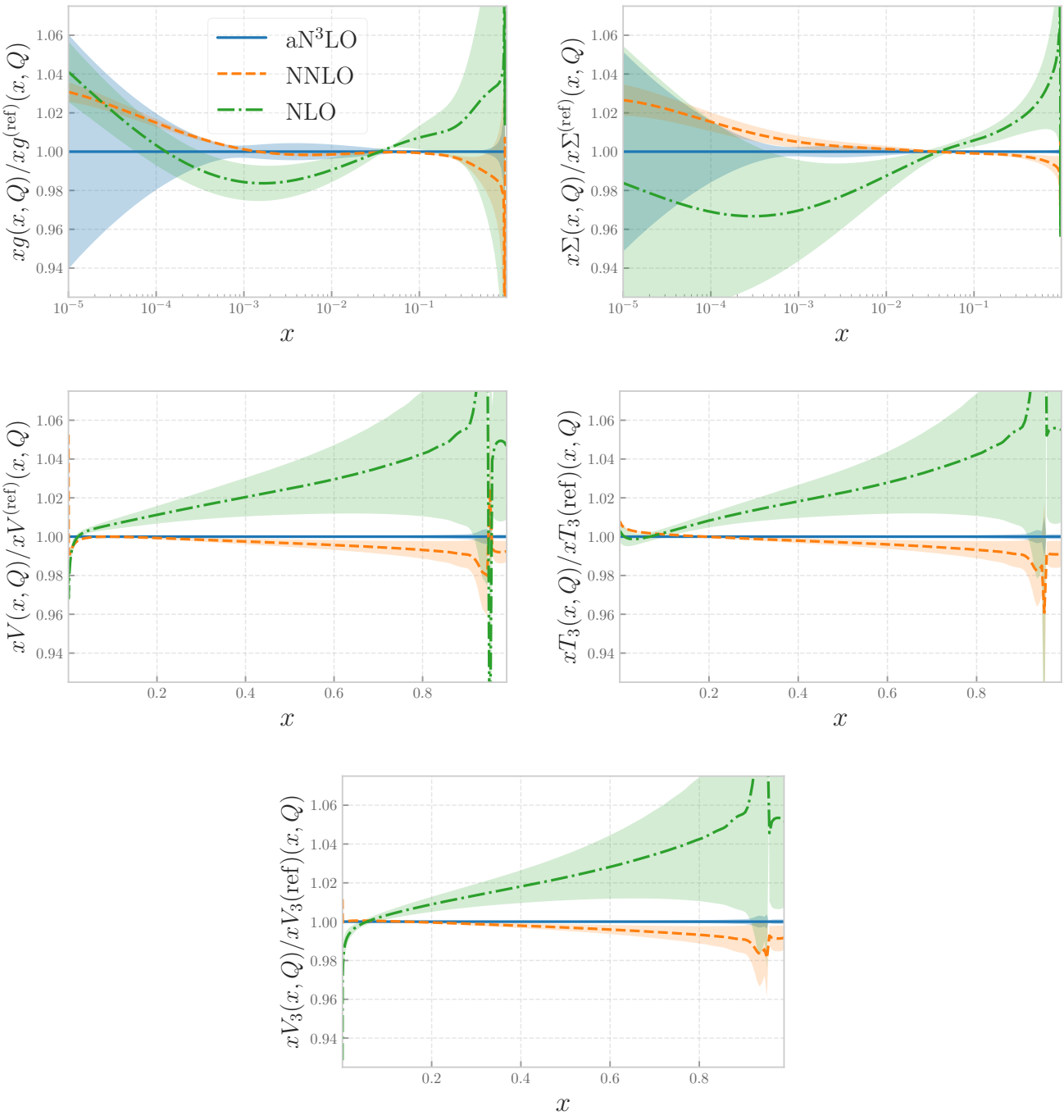



**Figure 2.8.** Comparison of the result obtained evolving from $Q_0 = 1.65$ GeV to $Q = 100$ GeV at NLO, NNLO, and aN$^3$LO using NNPDF4.0 NNLO as fixed starting PDF. Results are shown as ratio to the aN$^3$LO (from left to right and from top to bottom) for the gluon and singlet $\Sigma$, and for the $V$, $V_3$ and $T_3$ quark eigenstates of perturbative evolution (see Sect. 2.1). The total theory uncertainty is shown in all cases, i.e. the MHOU at NLO and NNLO, and the sum in quadrature of MHOU and IHOU at aN$^3$LO.

quark sector the MSHT20 prior has significantly larger IHOUs due to the fact that it does not include the more recent information on Mellin moments from Refs. [23, 28–31], which were not available at the time of the MSHT20 analysis [32]. At the level of posterior, however, significant differences appear also for $P_{gg}$, while persisting for $P_{gq}$. This means that the gluon evolution at aN$^3$LO is being significantly modified by the data entering the global fit, and it is not fully determined by the perturbative computation. Further benchmarks of aN$^3$LO splitting functions will be presented in Ref. [84].

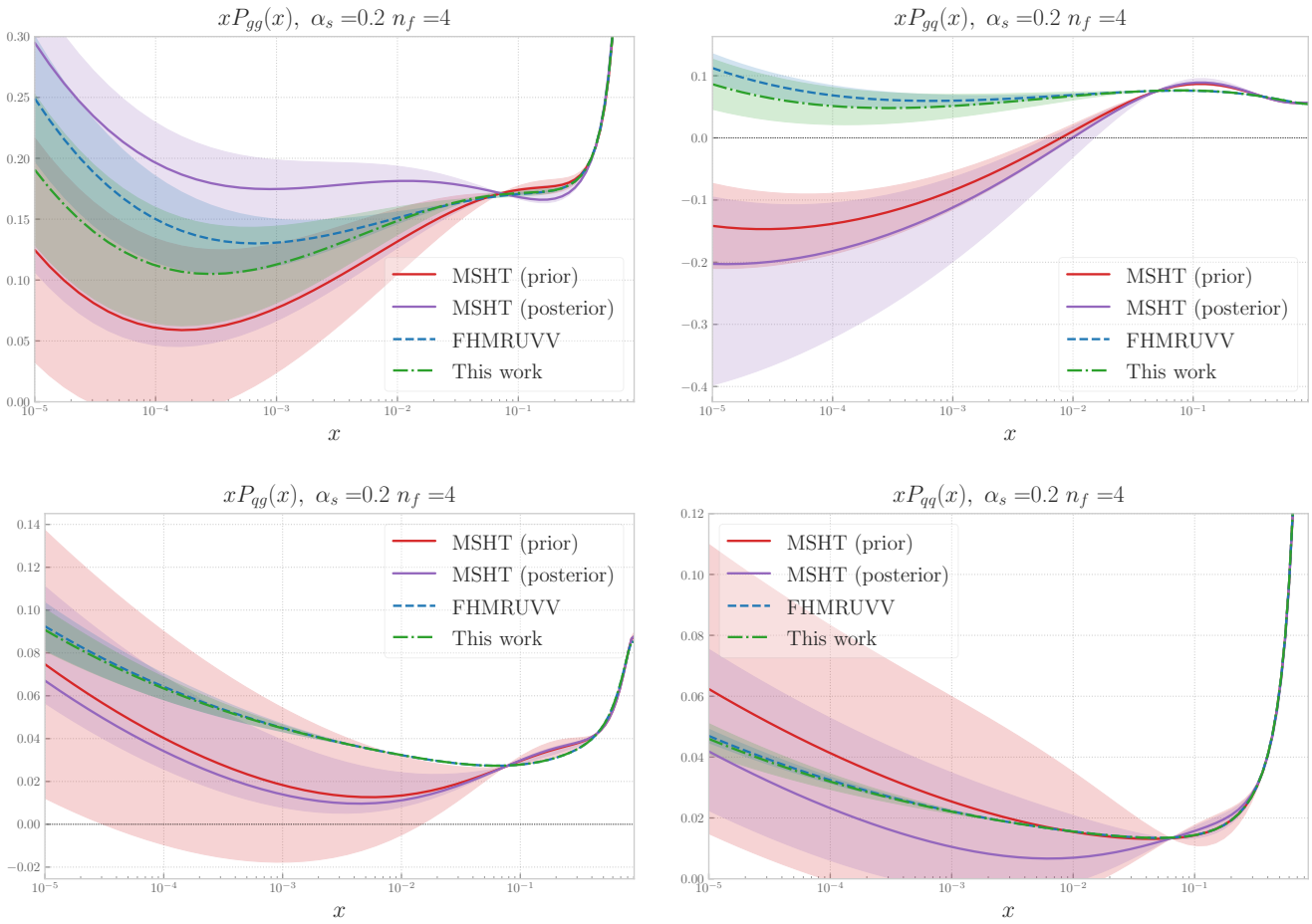



**Figure 2.9.** Same as Fig. 2.4, now comparing our aN$^3$LO result to those of Ref. [32] (MSHT20) and Refs. [28–30] (FHMRUVV). In all cases the uncertainty band correspond to the IHOU as estimated by the various groups. For the MSHT20 results, we display both the prior and the posterior parametrizations (see text).

# 3 N$^3$LO partonic cross-sections

A PDF determination at N$^3$LO requires, in addition to the splitting functions discussed in Sect. 2, hard cross sections at the same perturbative order. Exact N$^3$LO massless DIS coefficient functions have been known for several years [1–4,85,86], while massive coefficient functions are only available in various approximations [83, 87, 88]. For hadronic processes, N$^3$LO results are available for inclusive Drell-Yan production for the total cross-section [8,11,12] as well as for rapidity [18] and transverse momentum distributions [19], though neither of these is publicly available.

We now describe the implementation of these corrections. First, we review available results on DIS coefficient functions and summarize the main features of the approximation that we will use for massive coefficient functions [87,88]. Next we discuss how massless and massive DIS coefficient functions are combined to extend the FONLL general-mass variable-flavor number scheme to $\mathcal{O}\left(\alpha_s^3\right)$. Finally, we discuss N$^3$LO corrections for hadronic processes and different options for their inclusion in PDF determination.

## 3.1 N$^3$LO corrections to DIS structure functions

The DIS structure functions $F_i$ are evaluated from the convolution of PDFs and coefficient functions

$$\begin{aligned} F_i(x,Q^2) &= \sum_{k=1}^{n_f} C_{i,k}(x,\alpha_s) \otimes x q_k^+(x,Q^2) + C_{i,g}(x,\alpha_s) \otimes x g(x,Q^2)\,, \qquad i = \{2,L\}\,, \\ x F_3(x,Q^2) &= \sum_{k=1}^{n_f} C_{3,k}(x,\alpha_s) \otimes x q_k^-(x,Q^2)\,, \end{aligned} \tag{3.1}$$

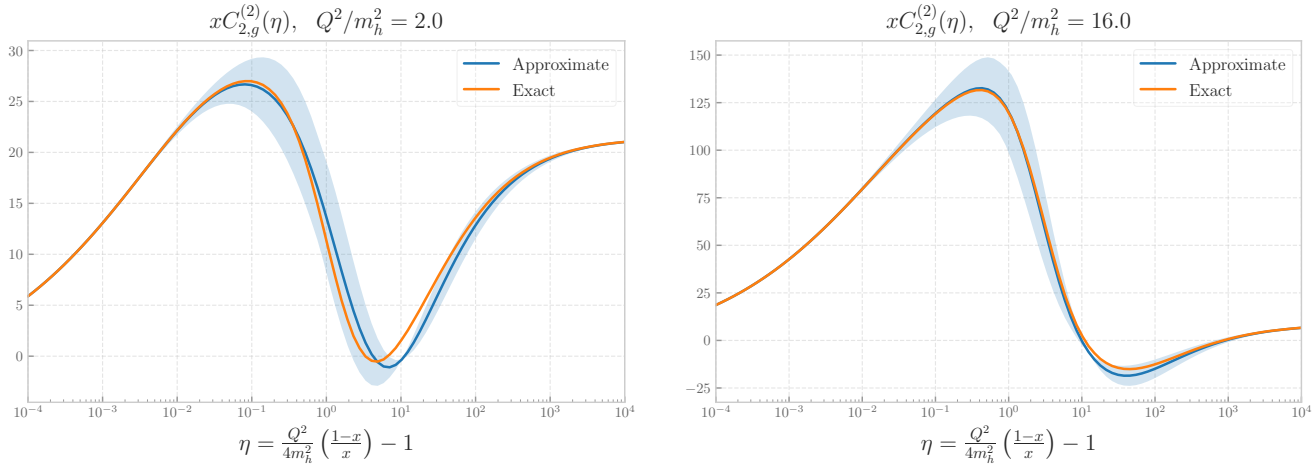



**Figure 3.1.** Comparison of the exact NNLO massive gluon-initiated coefficient function $xC^{(2)}_{2,g}(\eta)$ to the approximation Eq. (3.3) from Ref. [87], plotted as a function of $\eta$, Eq. (3.5), for fixed $Q^2$. Results are shown for two different values of $Q^2$, one close to threshold $Q^2 = 2m_h^2$ (left) and one at high scales $Q^2 = 16m_h^2$ (right). The uncertainty on the approximate result is obtained by varying the interpolating functions $f_1(x)$ and $f_2(x)$ in Eq. (3.3).

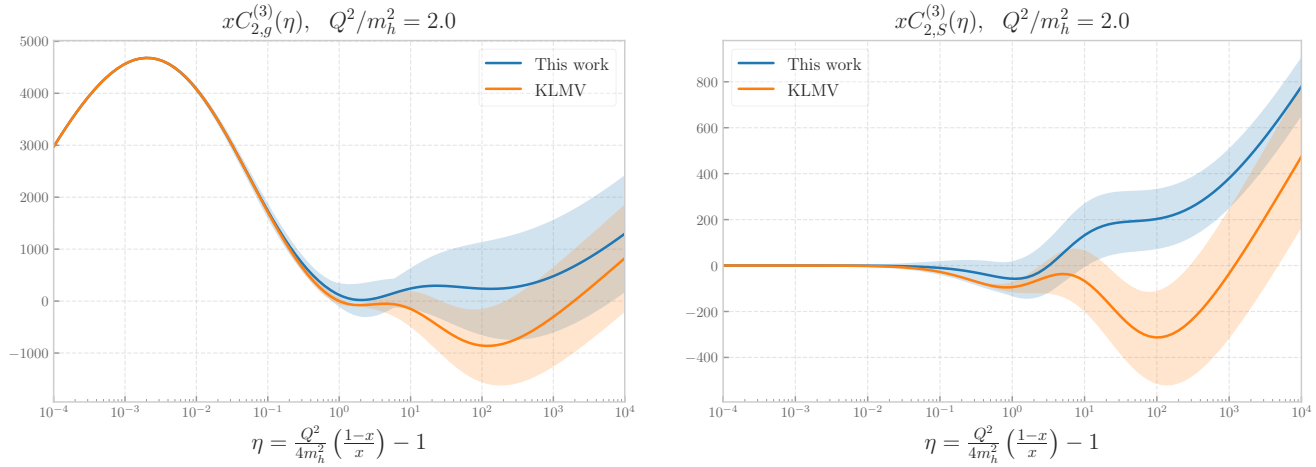



**Figure 3.2.** The approximate N$^3$LO massive gluon (left) and quark singlet (right) coefficient functions as a function of $\eta$ for fixed $Q^2 = 2m_h^2$. Our result based on the approximation of Ref. [87] is compared to the approximation of Ref. [83] (KLMV).

with the coefficient functions evaluated in a perturbative QCD expansion

$$C_{i,k}(x, \alpha_s(Q^2)) = \sum_{n=0} \alpha_s^n(Q^2) C^{(n)}_{i,k}(x)\,. \tag{3.2}$$

Coefficient functions with all quarks assumed to be massless were evaluated at N$^3$LO in [1, 2] for neutral-current charged-lepton scattering and recently independently benchmarked in [86]. The corresponding results for charged-current scattering were presented in [3, 4, 85].

For sufficiently low scale, some or all of the heavy quark masses cannot be neglected. Heavy quark contributions to structure functions may be treated in a decoupling scheme [89], in which heavy quarks do not contribute to the running of $\alpha_s$ and to PDF evolution, and coefficient functions acquire a dependence on the heavy quark mass $m_h$ [90]: $C_{i,k} = C_{i,k}(x, \alpha_s, m_h^2/Q^2)$ (massive coefficient functions, henceforth). The massive coefficient functions are known exactly up to NNLO for photon [91,92], Z [93,94] and W [95] exchange (for massless to massive transitions only) while at N$^3$LO only partial results are available [83, 88, 96, 97] or in the $Q^2 \gg m_h^2$ limit [71, 72, 74, 75, 77, 98].

We adopt an approximation for the N$^3$LO contribution $C^{(3)}_{i,k}(x, \alpha_s, m_h^2/Q^2)$ to massive coefficient functions for photon-induced DIS and neglect the axial-vector coupling of the Z boson, while we treat heavy quarks in the massless approximation for the W boson exchange. Such an approximation, based on known partial results, has been presented in Ref. [83], and recently revisited in Ref. [87]. The approaches of these

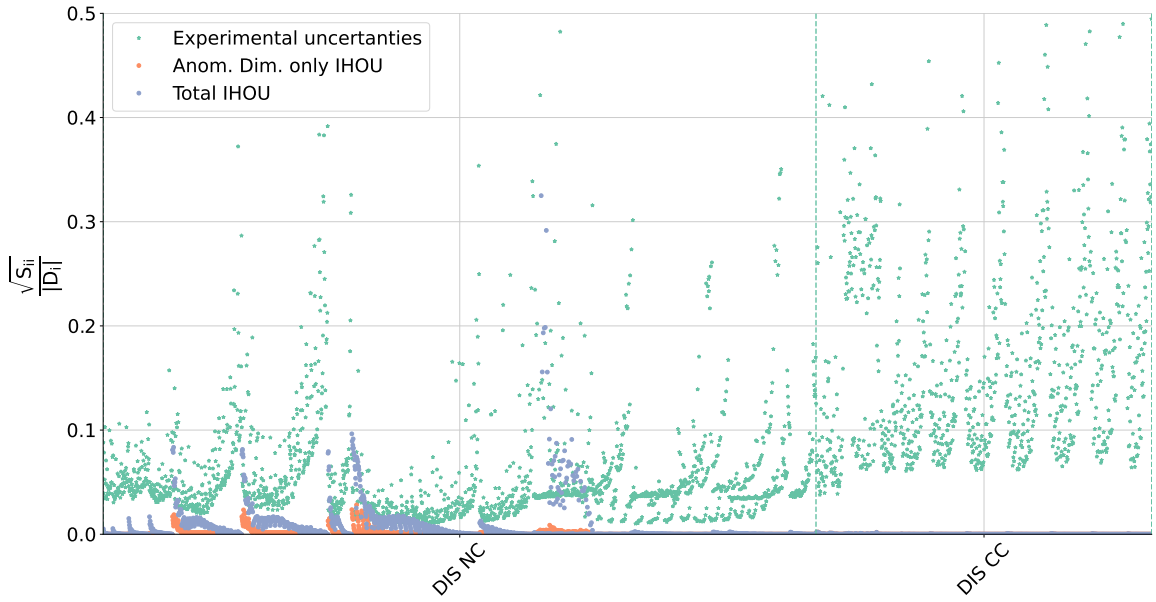



**Figure 3.3.** Square root of the diagonal entries of the IHOU covariance matrix for the DIS datasets normalized to the experimental central value $D_i$. We show the IHOU before and after adding to the covariance matrix Eq. (2.20) that accounts for uncertainty on anomalous dimensions the extra component Eq. (3.6) due to the massive coefficient function. The experimental uncertainty is also shown for comparison.

references rely on the same known exact results, and differ in the details of the way they are combined and interpolated. Here we will follow Ref. [87], see also Ref. [88], to which we refer for further details. Exact results come from threshold resummation and high-energy resummation, and are further combined with the asymptotic large-$Q^2$ limit, thereby ensuring that the approximate massive coefficient function reproduces the exact massless result in the $Q^2/m_h^2 \to \infty$ limit. In the approach of Refs. [87, 88] the massive coefficient functions are written as

$$C_{i,k}^{(3)}(x, m_h^2/Q^2) = C_{i,k}^{(3),\text{thr}}(x, m_h^2/Q^2) f_1(x) + C_{i,k}^{(3),\text{asy}}(x, m_h^2/Q^2) f_2(x)\,, \tag{3.3}$$

where $C_{i,k}^{(3),\text{thr}}$ and $C_{i,k}^{(3),\text{asy}}$ correspond to the contributions coming from differently resummations, and $f_1(x)$ and $f_2(x)$ are two suitable matching functions.

For massive quarks the threshold limit is $x \to x_{\max}$ with $x_{\max} = \frac{Q^2}{4m_h^2+Q^2}$ or $\beta \to 0$, with $\beta \equiv \sqrt{1 - \frac{4m_h^2}{s}}$ and $s = Q^2 \frac{1-x}{x}$ the centre-of-mass energy of the partonic cross-section. In this limit, the coefficient function contains logarithmically enhanced terms of the form $\alpha_s^n \ln^m \beta$ with $m \leq 2n$ due to soft gluon emission, which are predicted by threshold resummation [99]. Further contributions of the form $\alpha_s^n \beta^{-m} \ln^l \beta$, with $m \leq n$, arise from Coulomb exchange between the heavy quark and antiquark, and can also be resummed using non-relativistic QCD methods [100]. At N$^3$LO all these contributions are known and can be extracted from available resummed results [83]; they are included in $C_{i,k}^{(3),\text{thr}}$.

In the high-energy limit, the coefficient function contains logarithmically enhanced terms of the form $\alpha_s^n \ln^m x$ with $m \leq n-2$, which are determined at all orders through small-$x$ resummation at the LL level [97], from which the N$^3$LO expansion can be extracted [83]. This result can be further improved [87, 88] by including a particular class of NLL terms related to NLL perturbative evolution and the running of the coupling. In the approach of Refs. [87, 88] the high-energy contributions are combined into $C_{i,k}^{(3),\text{asy}}$ with the asymptotic $Q^2 \gg m_h^2$ limit of the coefficient function in the decoupling scheme [71, 72, 74, 75, 77, 98], while subtracting overlap terms. This ensures that in the $Q^2 \gg m_h^2$ limit, the structure function, computed from $C_{i,k}^{(3),\text{asy}}$ combined with decoupling-scheme PDFs, coincides with the structure function computed in the limit in which the heavy quark mass is neglected and the heavy quark is treated as a massless parton. However, the asymptotic limit can only be determined approximately since in particular some of the matching conditions are not fully known.

The interpolating functions, used to combine the two contributions in Eq. (3.3), are chosen to satisfy

the requirements

$$\begin{aligned} f_1(x) &\xrightarrow[x\to 0]{} 0, \quad f_1(x) \xrightarrow[x\to x_{\max}]{} 1\,, \\ f_2(x) &\xrightarrow[x\to 0]{} 1, \quad f_1(x) \xrightarrow[x\to x_{\max}]{} 0\,, \end{aligned} \tag{3.4}$$

which ensure that the threshold contribution vanishes in the small-$x$ limit and conversely. This guarantees that the approximation Eq. (3.3) is reliable in a broad kinematic range in the $(x, Q^2)$ plane: $C_{i,k}^{(3),\rm asy}$ reproduces the massless limit for large $Q^2$ values and for all values of $x$, including the small-$x$ limit, while $C_{i,k}^{(3),\rm thr}$ describes the threshold limit, with $x$ close to $x_{\max}$. An uncertainty on the approximate coefficient function can be constructed varying the functional form of the interpolating functions, as well as that of terms which are not fully known. This includes the NLL small-$x$ resummation, the matching functions that enter the asymptotic high $Q^2$ limit and constant (i.e. $\beta$-independent) terms in the threshold limit. This uncertainty vanishes in the $x \to x_{\max}$ limit, for which the exact known limit is reproduced, and becomes larger in the intermediate $\eta$ region. We refer to Ref. [87] for a detailed discussion of the construction of this uncertainty.

As a validation of this approximation, in Fig. 3.1 we compare a NNLO approximation constructed using the same methodology to the exact result for the massive gluon-initiated coefficient function $xC_{2,g}^{(2)}(\eta)$, expressed in terms of the variable

$$\eta = \frac{Q^2(1-x)}{4m_h^2 x} - 1\,. \tag{3.5}$$

Results are shown for two different values of the $Q^2/m_h^2$ ratio, close to threshold and at higher scales. Note that $\eta \to 0$ corresponds to $x \to x_{\max}$ (threshold limit), while $\eta \to \infty$ corresponds to either $Q^2/m_h^2 \to \infty$ for fixed $x$ (asymptotic limit), or $x \to 0$ for fixed $Q^2$ (high-energy limit). In this case the uncertainty band is obtained by varying the interpolating functions only.

The results found using the same procedure for the gluon and quark singlet coefficient functions at N$^3$LO are displayed in Fig. 3.2, compared to the approximation of Ref. [83], each shown with the respective uncertainty estimate. Good agreement between the different approximations is found, especially for the dominant gluon coefficient function. The approximations agree in the asymptotic $\eta \to 0$ and $\eta \to \infty$ limits, but differ somewhat at intermediate $\eta$.

The uncertainty involved in the approximation can be included as a further IHOU, alongside that discussed in Sect. 2.2, through an additional contribution to the theory covariance matrix. Namely, we define

$$\mathrm{cov}_{mn}^{C} = \frac{1}{2}\left(\Delta_m(+)\Delta_n(+) + \Delta_m(-)\Delta_n(-)\right). \tag{3.6}$$

Here $\Delta_m(\pm)$ is the shift in the prediction for the $m$-th DIS data point obtained by replacing the central approximation to the massive coefficient function with the upper or lower edge of the uncertainty range determined in Ref. [87] and shown as an uncertainty band in Fig. 3.2. Note that unlike in Eq. (2.20), we divide by the number of independent variations, without decreasing it by one, because the central value is not the average of the variations, and thus is independent. The contribution Eq. (3.6) is then added to the IHOU covariance matrix as a further term on the right-hand side of Eq. (2.20).

The impact of this contribution to the IHOU is assessed in Fig. 3.3, where the square root of the diagonal component of the covariance matrix is shown for all the DIS data points in our dataset, comparing the IHOU before and after adding to Eq. (2.20) the extra component Eq. (3.6) due to the IHOU on the massive coefficient function. It is clear that the impact of IHOUs due to perturbative evolution is generally negligible, in agreement with the results discussed in Sect. 2.6 and shown in Fig. 2.8: IHOUs on splitting functions are only significant at small $x$, but available small-$x$ data are at relatively low scale where the evolution length is small. The impact of IHOUs on massive coefficient functions is relevant for data on tagged bottom and charm structure functions, but otherwise moderate and only significant for structure function data close to the heavy quark production thresholds.

### 3.2 A general-mass variable flavor number scheme at N$^3$LO

The N$^3$LO DIS coefficients functions described in the previous section enable the extension to $\mathcal{O}\left(\alpha_s^3\right)$ of the FONLL [101] general-mass variable flavor number scheme for DIS [42–44]. The goal of the FONLL strategy is results that are accurate and reliable for all values of $Q^2$ from the production threshold $Q^2 \sim m_h^2$ to the asymptotic limit $Q^2 \gg m_h^2$.

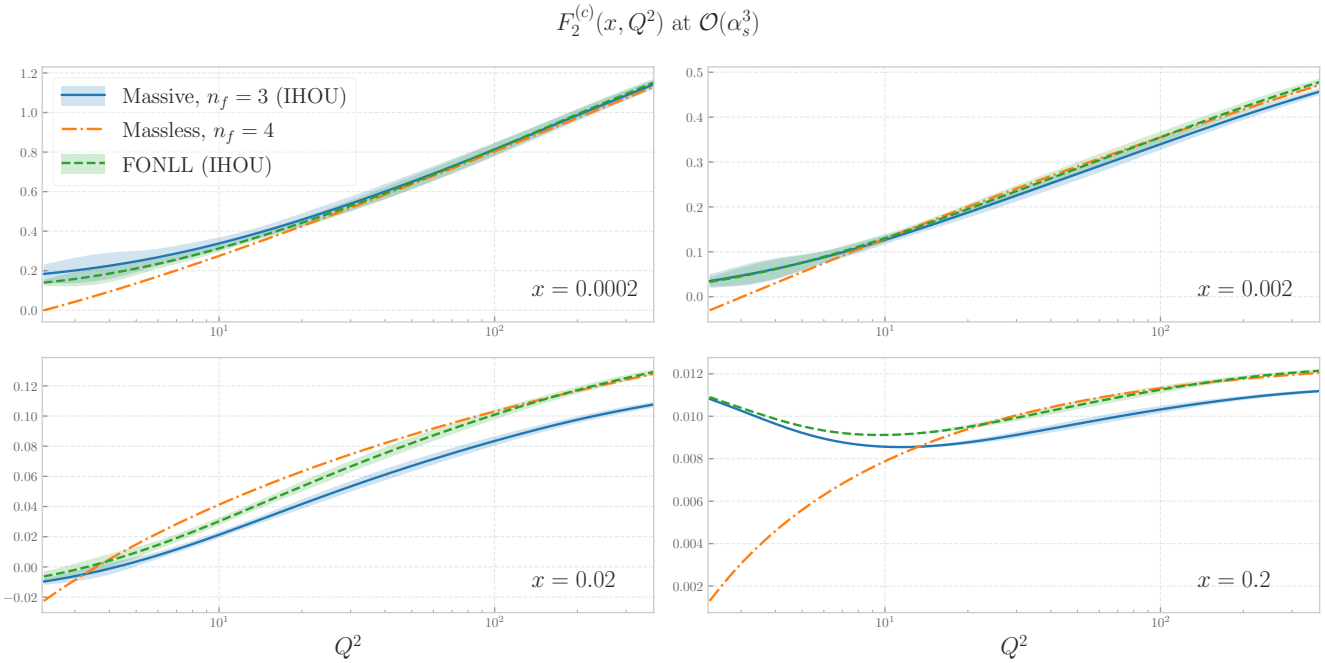



**Figure 3.4.** The charm structure function $F_2^{(c)}(x, Q^2, m_c^2)$ in the FONLL-E scheme, compared to the massive and massless scheme results (see text). Results are shown as a function of $Q^2$ for $x = 2 \times 10^{-4}$ (top left), $x = 2 \times 10^{-3}$ (top right), $x = 2 \times 10^{-2}$ (bottom left), and $x = 2 \times 10^{-1}$ (bottom right). The uncertainty shown on the FONLL and massive curves is the IHOU on the heavy quark coefficient functions Eq. (3.6).

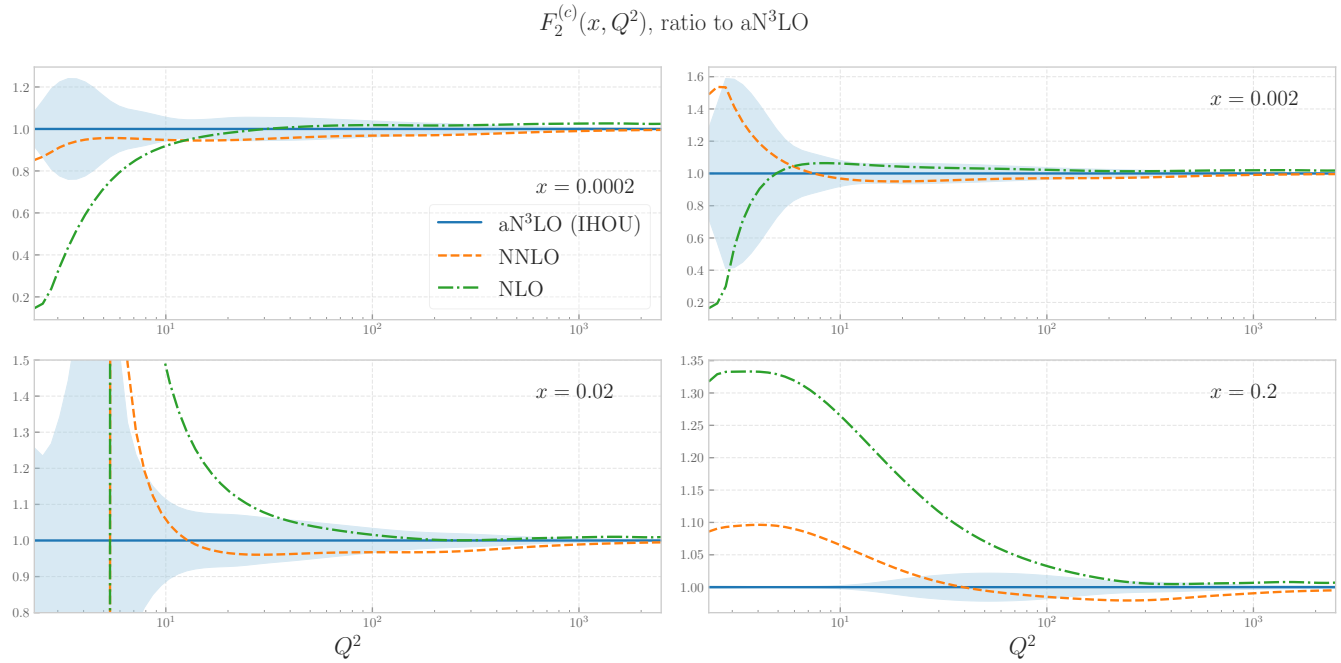



**Figure 3.5.** Same as Fig. 3.4, now comparing the FONLL-A (used at NLO $\mathcal{O}(\alpha_s)$), FONLL-C (used at NNLO $\mathcal{O}(\alpha_s^2)$), and FONLL-E (used at N$^3$LO $\mathcal{O}(\alpha_s^3)$), all shown as a ratio to FONLL-E. The FONLL-E result includes the IHOU on the heavy quark coefficient functions Eq. (3.6).

Assuming a single heavy quark, calculations performed in a decoupling scheme with $n_\ell$ light quarks retain the full dependence on the heavy quark mass and include the contribution of heavy quarks at a fixed perturbative order (massive scheme, henceforth). Calculations performed in a scheme in which the heavy quark is treated as massless (massless scheme, henceforth), and endowed with a PDF that satisfies perturbative matching conditions, resums logarithms of $Q^2/m_h^2$ to all orders through the running of the coupling and the evolution of PDFs, but does not include terms that are suppressed as powers of $\frac{m_h^2}{Q^2}$ and

thus become relevant when $Q^2 \gtrsim m_h^2$. The FONLL prescription matches the two calculations by defining

$$F_i^{\rm FONLL} = F_i^{(n)}(x,Q^2,m_h^2) + F_i^{(n+1)}(x,Q) - F_i^{(n,0)}(x,\ln(Q^2/m_h^2))\,, \qquad i = 2,L,3\,, \tag{3.7}$$

where $F_i^{(n)}$ denotes the massive computation in which the $(n+1)$-th (heavy) flavor decouples, $F_i^{(n+1)}$ the one in which it is treated as massless, and $F_i^{(n,0)}$ is the asymptotic large-$Q^2$ limit of the massive scheme calculation, which depends only logarithmically on the heavy quark mass. This construction reduces to the decoupling calculation for $Q^2 \approx m_h^2$ and to the massless one for $Q^2 \gg m_h^2$.

The FONLL prescription of Eq. (3.7) was implemented in Ref. [42–44] for DIS to NNLO, by expressing all terms on the right-hand side in terms of $\alpha_s$ and PDFs all defined in the massless scheme. This has the advantage of providing an expression that can used with externally provided PDFs, that are typically available only in a single factorization scheme for each value of the scale $Q$.

However, the recent EKO code [40] allows, at any given scale, the coexistence of PDFs defined in schemes with a different number of massless flavors. Furthermore, the recent YADISM program [41] implements DIS coefficient functions corresponding to all three contributions on the right-hand side of Eq. (3.7). It is then possible to implement the FONLL prescription Eq. (3.7) by simply combining expressions computed in different schemes [102]. This formalism is especially advantageous at higher perturbative orders, where the analytic expressions relating PDFs in different scheme grow in complexity, and also above bottom threshold, where the iteration of Eq. (3.7) on charm and bottom heavy quarks leads to coexisting $n_f = 3$, 4, 5 PDFs, while the method of Ref. [42–44] would require re-expressing the massive scheme PDFs into massless scheme PDFs twice.

In the FONLL method, Eq. (3.7), the first two terms on the right-hand side may be computed at different perturbative orders, provided one ensures that the third term correctly includes only their common contributions. In Ref. [42] some natural choices were discussed, based on the observation that in the massive scheme, the heavy quark contributes to the structure functions only at $\mathcal{O}\left(\alpha_s\right)$ and beyond, while in the massless scheme it already contributes at $\mathcal{O}\left(\alpha_s^0\right)$. Hence natural choices are to combine both the massive and massless contributions at $\mathcal{O}\left(\alpha_s\right)$ (FONLL-A), or else the massive contribution at $\mathcal{O}\left(\alpha_s^2\right)$ and the massless contribution at $\mathcal{O}\left(\alpha_s\right)$, i.e. both at second nontrivial order (FONLL-B). The corresponding two options at the next order are called FONLL-C and -D.

Here, we will consider FONLL-E, in which both the massless and massive contributions are determined at $\mathcal{O}\left(\alpha_s^3\right)$. The charm structure function $F_2^{(c)}(x,Q^2)$, computed in this scheme, is displayed in Fig. 3.4 as a function of $Q^2$ for four values of $x$ (with $m_c = 1.51$ GeV), and compared to the massive and massless scheme results, with the IHOU on the massive coefficient function shown for the first two cases. The structure functions are computed using the NNPDF4.0 aN$^3$LO PDF set (to be discussed in Sect. 4 below) which satisfies aN$^3$LO evolution equations, as is necessary for consistency with the massless scheme result at high scale. It is clear that the FONLL results interpolate between the massive and massless calculations as the scale grows. The $Q^2$ value at which either of the massive or massless results dominate depend strongly on $x$. Except for the lowest $Q^2$ values, the IHOUs associated with the calculation remain moderate.

The perturbative convergence of the charm structure function is assessed in Fig. 3.5, where we compare the FONLL-A, FONLL-C and FONLL-E results, all shown as a ratio to FONLL-E, the latter also including the IHOU as in Fig. 3.4. Clearly, convergence is faster at higher scales due to asymptotic freedom, and it appears that the perturbative expansion has essentially converged for $Q^2 \gtrsim 10$ GeV$^2$. On the other hand, the impact of aN$^3$LO at low scale is sizable, up to 50% for small $Q^2$ and $x = 2\times10^{-3}$. The IHOUs are correspondingly sizable at low scale, and in fact always larger than the difference between the NNLO and aN$^3$LO results except at the highest $x$ values and the lowest scales, implying that for the charm structure function aN$^3$LO may be more accurate, but possibly not more precise than NNLO.

An analogous study of perturbative convergence of the inclusive structure function is shown in Fig. 3.6 (note the different scale on the $y$ axis). Interestingly, the effect of the aN$^3$LO corrections changes sign when going from $x = 2\times10^{-4}$ to larger values of $x$. In general, N$^3$LO corrections are smaller at the inclusive level: specifically, aN$^3$LO corrections to the inclusive structure function are below 2% for $Q^2 \gtrsim 10$ GeV$^2$, and at most of order 10% around the charm mass scale. The impact of the IHOUs on the heavy coefficient is further reduced due to the fact that charm contributes at most one quarter of the total structure function. Consequently, the aN$^3$LO correction to the NNLO result is now larger than the IHOU in a significant kinematic region. This, together with the fact that aN$^3$LO corrections are comparable or larger than typical experimental uncertainties on structure function data, motivates their inclusion in a global PDF determination.

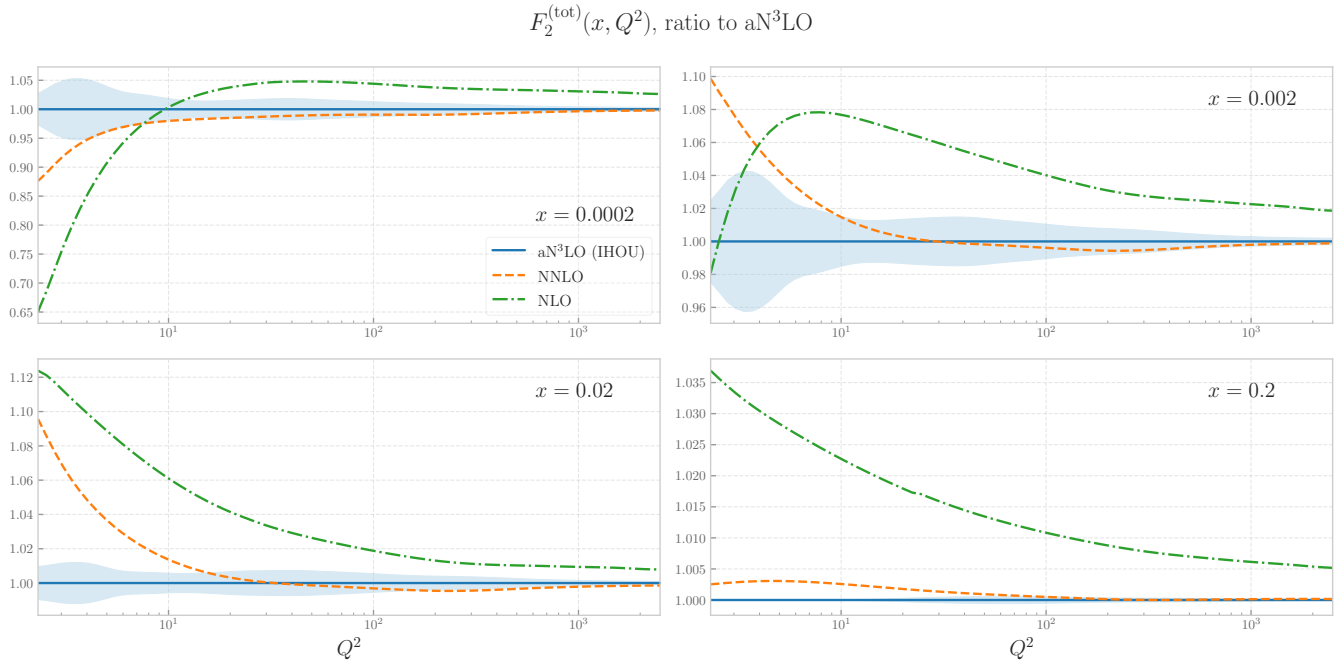



**Figure 3.6.** Same as Fig. 3.5 for the inclusive structure function $F_2^{(\mathrm{tot})}(x, Q^2)$. Note the different scale on the $y$ axis.

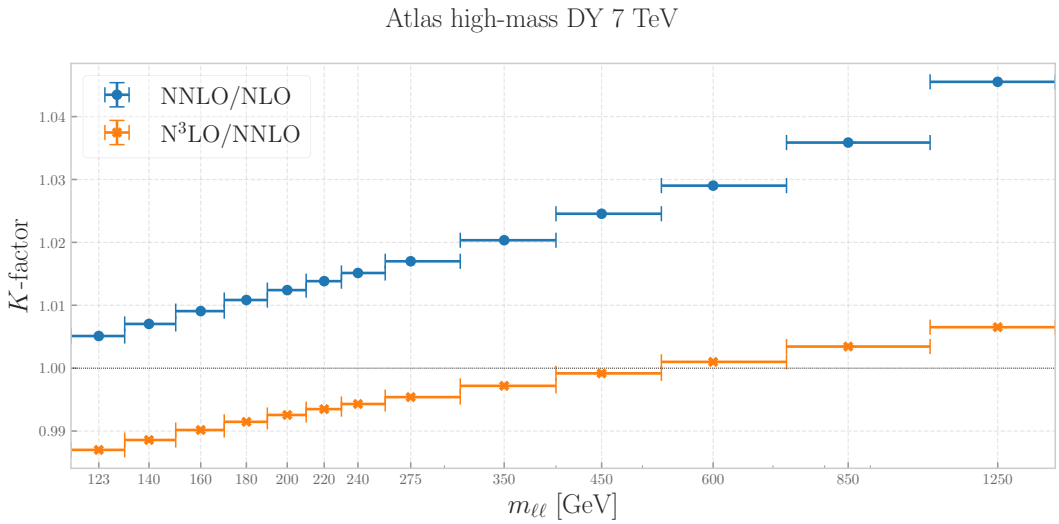



**Figure 3.7.** Ratio of the total NC Drell-Yan cross-section as a function of the NNLO/NLO and N$^3$LO/NNLO calculations for the inclusive NC Drell-Yan cross-section in bins of $m_{\ell\ell}$, the invariant mass of the final-state dilepton pair, using the N3LOXS code, integrated over all other kinematic variables. The $m_{\ell\ell}$ binning is chosen to be that of the ATLAS 7 TeV high-mass DY measurement [103]. The same PDF set is used in the numerator and denominator, namely NNPDF4.0 NNLO (for NNLO/NLO) and aN$^3$LO (for N$^3$LO/NNLO).

## 3.3 N$^3$LO corrections to hadronic processes

N$^3$LO corrections to the total cross-section for inclusive neutral- (NC) and charged-current (CC) Drell-Yan production [11, 12] are available through the N3LOXS public code [8], both for on-shell $W$ and $Z$ and as a function of the dilepton invariant mass $m_{\ell\ell}$. Differential distributions at the level of leptonic observables for the same processes have also been computed [18, 19], but are not publicly available. No N$^3$LO calculations are available for other processes included in the NNPDF4.0 dataset.

The ratio of the NC total cross-section evaluated at two subsequent perturbative orders with a fixed set of PDFs, chosen as NNPDF4.0NNLO is shown in Fig. 3.7. Results are shown in the high-mass region, as a function of $m_{\ell\ell}$, with the same binning as the ATLAS 7 TeV measurement [103]. Perturbative convergence is apparent, with the N$^3$LO/NNLO ratio closer to unity and smoother than its NNLO/NLO counterpart: while NNLO corrections range between +0.5% and +4%, at N$^3$LO they are reduced to −1.2% and +0.5%.

Total cross-section data are obtained by extrapolating measurements performed in a fiducial region.

| Dataset | Ref. | $n_{\rm dat}$ | Kin$_1$ | Kin$_2$ [GeV] |
|---|---|---|---|---|
| ATLAS high-mass DY 7 TeV | [103] | 13 | $\lvert\eta_\ell\rvert \leq 2.1$ | $116 \leq m_{\ell\ell} \leq 1500$ |
| ATLAS $Z$ 7 TeV ($\mathcal{L} = 35$ pb$^{-1}$) | [104] | 8 | $\lvert\eta_\ell, y_Z\rvert \leq 3.2$ | $Q = m_Z$ |
| ATLAS $Z$ 7 TeV ($\mathcal{L} = 4.6$ fb$^{-1}$) | [105] | 39 | $\lvert\eta_\ell, y_Z\rvert \leq 2.5, 3.6$ | $Q = m_Z$ |
| ATLAS $\sigma^{\rm tot}_{W,Z}$ 13 TeV | [106] | 3 | — | $Q = m_W, m_Z$ |

**Table 3.1.** The LHC NC DY production datasets in NNPDF4.0 for which an N$^3$LO $K$-factor has been included in a variant of the default aN$^3$LO PDF determination (see Sect. 4). For each dataset we indicate the published references, the number of datapoints and the kinematic variables.

Whereas for NC Drell-Yan production in the central rapidity region and for dilepton invariant masses around the $Z$-peak, the N$^3$LO/NNLO cross-section ratio depends only mildly on the dilepton rapidity $y_{\ell\ell}$ [18, 19], it is unclear whether this is the case also off-peak or at very large and very small rapidities. Hence, the inclusion of N$^3$LO corrections for hadronic processes is, at present, not fully reliable. We have consequently not included them in our default determination, but only in a dedicated variant, with the goal of assessing their impact.

The datasets for which N$^3$LO corrections have been included in this variant are listed in Table 3.1. We include the high-mass NC cross section, the $Z$ rapidity distribution in the central rapidity region for on-shell $Z$-production, and the total $W$ and $Z$ cross-sections. For all these processes the N$^3$LO cross-section is determined by multiplying the NNLO result by a $K$-factor determined using a fixed underlying PDF set, namely the aN$^3$LO NNPDF4.0 PDF set to be discussed in Sect. 4 below. Specifically, for the rapidity distribution we take the same fixed $K$-factor as for the total cross-section. We do not include off-shell or double-differential rapidity distributions (specifically from CMS), off-forward rapidity distributions (specifically from LHCb) and low-mass total cross-sections, for all of which the approximation of assuming the $K$-factor to be independent of rapidity and/or amenable to fiducial extrapolation is even less reliable. The datasets are labeled as in Table 2.4 of Ref. [37] [4].

Despite the fact that we are not yet able to determine reliably N$^3$LO corrections for currently available LHC measurements, we wish to include the full NNPDF4.0 dataset in our aN$^3$LO PDF determination. To this purpose, we endow all data for which N$^3$LO are not included with an extra uncertainty that accounts for the missing N$^3$LO terms. This is estimated using the methodology of Refs. [35, 36], recently used in Ref. [38] to produce a variant of the NNPDF4.0 PDF sets that includes MHOUs.

Thus, when not including N$^3$LO corrections to the hard cross-section, the theory prediction is evaluated by combining aN$^3$LO evolution with the NNLO cross sections. The prediction is then supplemented with a theory covariance matrix, computed varying the renormalization scale $\mu_R$ using a three-point prescription [35, 36]:

$$\mathrm{cov}^{\rm NNLO}_{mn} = \frac{1}{2}\left(\Delta_m(+)\Delta_n(+) + \Delta_m(-)\Delta_n(-)\right), \tag{3.8}$$

analogous to Eq. (3.6), but now with $\Delta_m(\pm)$ the shift in the prediction for the $m$-th data point obtained by replacing the coefficient functions with those obtained by performing upper or lower renormalization scale variation using the methodology of Ref. [36] (as implemented and discussed in Ref. [38], Eq. (2.9)). This MHOU covariance matrix is then added to the IHOU covariance matrix as a further term on the right-hand side of Eq. (2.20).

The impact of this uncertainty is shown in Fig. 3.8, where we show for all hadronic datasets the square root of the diagonal entries of the MHOU covariance matrix Eq. (3.8), compared to those of the IHOU covariance matrix Eq. (2.20), and to the experimental uncertainties, all normalized to the central theory prediction. The MHOU is generally larger than the IHOU, indicating that the missing N$^3$LO terms in the hard cross-sections are larger than the IHOU uncertainty in N$^3$LO perturbative evolution. The experimental uncertainties are generally larger still.

In addition to the NNPDF4.0 aN$^3$LO baseline PDF set obtained in this manner, we will also produce a NNPDF4.0 MHOU aN$^3$LO set, in analogy to the NLO and NNLO MHOU sets recently presented in

[4] The number of datapoints for the rapidity distributions differs from the numbers in this table because here we only include $Z$ distributions

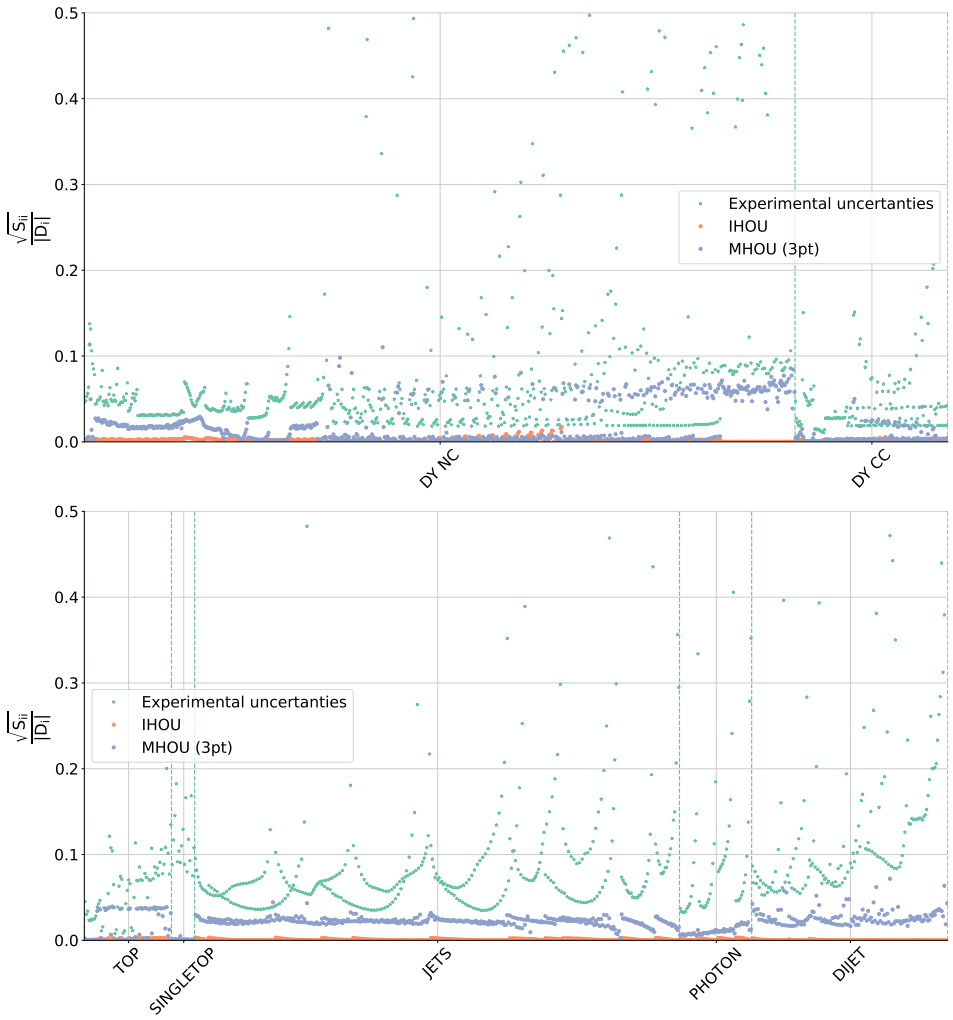



**Figure 3.8.** Same as Fig. 3.3 now comparing the IHOU from Eq. (2.20) with the MHOU from Eq. (3.8) due to the missing N$^3$LO correction to the matrix element. Results are shown for all hadronic data in the NNPDF4.0 dataset: specifically Drell-Yan (top) and top pair, single top, single-inclusive jet, prompt photon and dijet production (bottom).

Ref. [38]. For this set, MHOUs on both perturbative evolution and on the hard matrix elements are included using the methodology of Refs. [35, 36] with a theory covariance matrix determined performing combined correlated renormalization and factorization scale variations with a 7-point prescription, as discussed in detail in Ref. [38]. In this case, we simply perform scale variation on the expressions at the order at which they are being computed, namely aN$^3$LO for anomalous dimensions and DIS coefficient functions and NNLO for hadronic processes. The scale variation then is automatically larger and suitable deweights processes for which N$^3$LO corrections are not available. The possibility of simultaneously including in a PDF determination processes for which theory predictions are only available at different perturbative orders is an advantage of the inclusion of MHOUs in the PDF determination, as already pointed out in Refs. [107, 108].

# 4 NNPDF4.0 at aN$^3$LO

We now present the aN$^3$LO NNPDF4.0 PDF sets. They have been obtained by using the dataset and methodology discussed in [37] and used for the construction of the LO, NLO, and NNLO NNPDF4.0 presented there, now extended to aN$^3$LO. The aN$^3$LO results are obtained using the approximate N$^3$LO splitting function of Sect. 2, the exact massless and approximate massive N$^3$LO coefficient functions of Sect. 3.1, and NNLO hadronic cross-sections supplemented by an extra uncertainty as per Sect. 3.3.

Theoretical predictions are obtained using the new theory pipeline of Ref. [70], which relies on the EKO evolution code [40] and on the YADISM DIS module [41]. As discussed in Sect. 3.2, this pipeline in particular includes a new FONLL implementation, that differs from the previous one by subleading terms. A further small difference in comparison to Ref. [37] is the correction of a few minor bugs in the data implementation. The overall impact of all these changes was assessed in Appendix A of Ref. [109], and was found to be very

| Dataset | NLO | | | NNLO | | | aN$^3$LO | | |
|---|---|---|---|---|---|---|---|---|---|
| | $N_{\rm dat}$ | no MHOU | MHOU | $N_{\rm dat}$ | no MHOU | MHOU | $N_{\rm dat}$ | no MHOU | MHOU |
| DIS NC | 1980 | 1.30 | 1.22 | 2100 | 1.22 | 1.20 | 2100 | 1.22 | 1.20 |
| DIS CC | 988 | 0.92 | 0.87 | 989 | 0.90 | 0.90 | 989 | 0.91 | 0.92 |
| DY NC | 667 | 1.49 | 1.32 | 736 | 1.20 | 1.15 | 736 | 1.17 | 1.16 |
| DY CC | 193 | 1.31 | 1.27 | 157 | 1.45 | 1.37 | 157 | 1.37 | 1.36 |
| Top pairs | 64 | 1.90 | 1.24 | 64 | 1.27 | 1.43 | 64 | 1.23 | 1.41 |
| Single-inclusive jets | 356 | 0.86 | 0.82 | 356 | 0.94 | 0.81 | 356 | 0.84 | 0.83 |
| Dijets | 144 | 1.55 | 1.81 | 144 | 2.01 | 1.71 | 144 | 1.78 | 1.67 |
| Prompt photons | 53 | 0.58 | 0.47 | 53 | 0.76 | 0.67 | 53 | 0.72 | 0.68 |
| Single top | 17 | 0.35 | 0.34 | 17 | 0.36 | 0.38 | 17 | 0.35 | 0.36 |
| Total | 4462 | 1.24 | 1.16 | 4616 | 1.17 | 1.13 | 4616 | 1.15 | 1.14 |

**Table 4.1.** The number of data points and the $\chi^2$ per data point obtained in the NLO, NNLO, and aN$^3$LO NNPDF4.0 fits without and with MHOUs, see text for details. The datasets are grouped according to the same process categorisation as that used in Ref. [38].

limited, so that the new and old implementations can be considered equivalent, and the PDF sets presented here can be considered the extension to aN$^3$LO of the NNPDF4.0 PDF sets of Ref. [37].

In addition to the default NNPDF4.0 aN$^3$LO PDF determination, we also present an aN$^3$LO PDF determination that includes MHOUs on all the theory predictions used in the PDF determination. This is constructed using the same methodology recently used to produce the NNPDF4.0MHOU NNLO PDF set in Ref. [38]. In order to be able to discuss perturbative convergence and the impact of MHOUs we will also present a NNPDF4.0MHOU NLO PDF set constructed using the same methodology, and exactly the same dataset as the default NNPDF4.0 NLO PDF set (which differs from the NNPDF4.0 NNLO dataset).

We finally construct two variants of the aN$^3$LO PDF sets (both with and without MHOUs) with modified N$^3$LO theory. In a first variant, we replace our own approximation to the N$^3$LO anomalous dimensions, discussed in Sect. 2, with that of Refs. [28–30]. In the second variant, we will also include N$^3$LO corrections for the processes listed in Tab. 3.1, as discussed in Sect. 3.3.

We first assess the fit quality, then present the PDFs and their uncertainties, and study perturbative convergence and the effect on it of the inclusion of MHOUs. We then specifically study the impact of aN$^3$LO corrections on intrinsic charm. We then turn to the variants, and finally compare our results to the recent MSHT20 aN$^3$LO PDFs.

## 4.1 Fit quality

Tables 4.1-4.4 display the number of data points and the $\chi^2$ per data point obtained in the NLO, NNLO, and aN$^3$LO NNPDF4.0 fits with and without MHOUs. In Table 4.1 the datasets are grouped according to the process categorisation used in Ref. [38]. Results for individual datasets are displayed in Table 4.2, (NC and CC DIS), in Table 4.3 (NC and CC DY), and in Table 4.4 (top pairs, single-inclusive jets, dijets, isolated photons, and single top). The naming of the datasets follows Ref. [37]. The value of the total $\chi^2$ per data point is also shown as a function of the perturbative order in Fig. 4.1.

The NLO and NNLO results without MHOUs are obtained using the NLO and NNLO NNPDF4.0 PDF sets [37]. The NNLO result with MHOUs is obtained using the NNPDF4.0MHOU NNLO set from Ref. [38], while, as already mentioned, the NNPDF4.0MHOU NLO presented here for the first time uses an identical methodology to NNPDF4.0MHOU NNLO [38], but the same dataset as NNPDF4.0 NLO [37]. Hence, the datasets with and without MHOU are always the same, but the NLO and NNLO datasets are not the same but rather follow Ref. [37]. The N$^3$LO dataset is the same as NNLO. In all cases, the theoretical predictions entering the computation of the $\chi^2$ are obtained with the new theory pipeline. The covariance matrix, whenever needed, is computed as described in Sect. 4.1 of Ref. [38]. The N$^3$LO predictions are computed with the aforementioned aN$^3$LO PDF sets. These are based on the same datasets and kinematic cuts as the NNPDF4.0 NNLO PDF sets, use the theoretical predictions discussed in Sects. 2-3, and are supplemented with a IHOU covariance matrix as discussed in Sects. 2.2-3.1 and a MHOU for hadronic processes for which N$^3$LO hard cross sections are not available as discussed in Sect. 3.3.

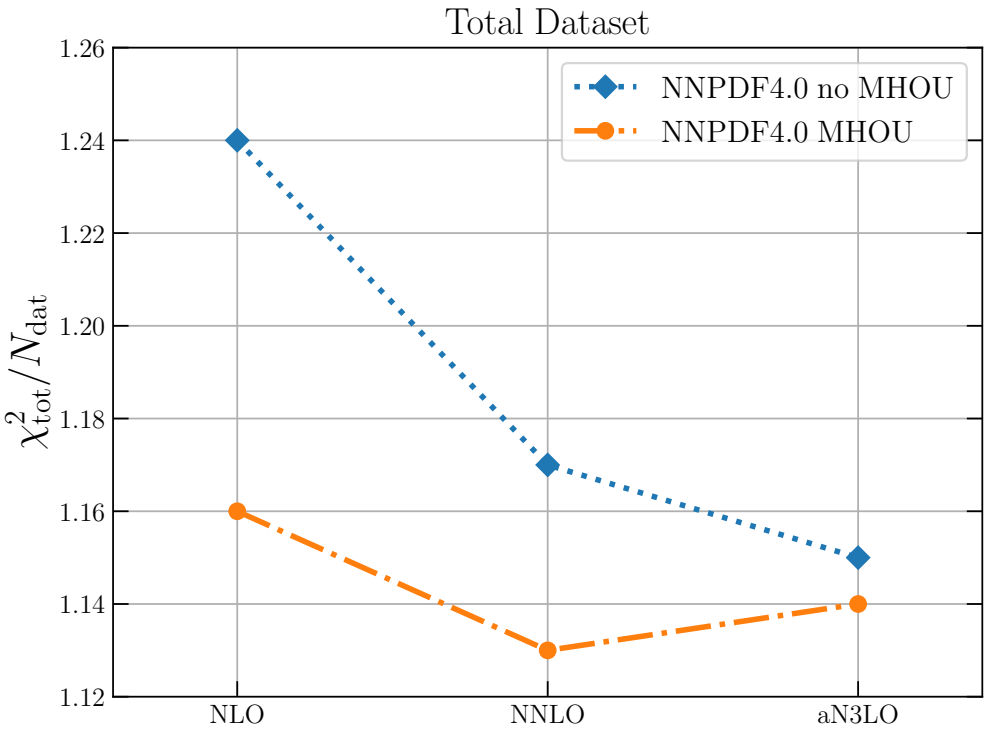



**Figure 4.1.** The values of the total $\chi^2$ per data point in the NNPDF4.0 NLO, NNLO, and aN$^3$LO fits without and with MHOUs.

| | | NLO | | | NNLO | | | aN$^3$LO | |
|---|---|---|---|---|---|---|---|---|---|
| Dataset | $N_{\rm dat}$ | no MHOU | MHOU | $N_{\rm dat}$ | no MHOU | MHOU | $N_{\rm dat}$ | no MHOU | MHOU |
| NMC $F_2^d/F_2^p$ | 121 | 0.87 | 0.86 | 121 | 0.87 | 0.88 | 121 | 0.88 | 0.88 |
| NMC $\sigma^{\rm NC,p}$ | 203 | 1.82 | 1.30 | 204 | 1.57 | 1.33 | 204 | 1.57 | 1.36 |
| SLAC $F_2^p$ | 33 | 1.64 | 0.74 | 33 | 0.91 | 0.68 | 33 | 0.93 | 0.72 |
| SLAC $F_2^d$ | 34 | 0.90 | 0.68 | 34 | 0.61 | 0.54 | 34 | 0.62 | 0.58 |
| BCDMS $F_2^p$ | 333 | 1.62 | 1.24 | 333 | 1.40 | 1.29 | 333 | 1.39 | 1.40 |
| BCDMS $F_2^d$ | 248 | 1.05 | 1.00 | 248 | 1.01 | 0.99 | 248 | 1.04 | 1.03 |
| HERA I+II $\sigma_{\rm NC}$ $e^-p$ | 159 | 1.44 | 1.40 | 159 | 1.40 | 1.39 | 159 | 1.45 | 1.40 |
| HERA I+II $\sigma_{\rm NC}$ $e^+p$ ($E_p = 460$ GeV) | 192 | 1.12 | 1.05 | 204 | 1.09 | 1.04 | 204 | 1.07 | 1.05 |
| HERA I+II $\sigma_{\rm NC}$ $e^+p$ ($E_p = 575$ GeV) | 236 | 0.85 | 0.84 | 254 | 0.93 | 0.88 | 254 | 0.87 | 0.88 |
| HERA I+II $\sigma_{\rm NC}$ $e^+p$ ($E_p = 820$ GeV) | 54 | 1.15 | 0.85 | 70 | 1.12 | 0.95 | 70 | 0.96 | 0.86 |
| HERA I+II $\sigma_{\rm NC}$ $e^+p$ ($E_p = 920$ GeV) | 317 | 1.30 | 1.21 | 377 | 1.31 | 1.25 | 377 | 1.27 | 1.24 |
| HERA I+II $\sigma^c_{\rm NC}$ | 24 | 2.18 | 1.40 | 37 | 1.96 | 1.75 | 37 | 1.86 | 1.57 |
| HERA I+II $\sigma^b_{\rm NC}$ | 26 | 1.42 | 1.05 | 26 | 1.44 | 1.11 | 26 | 1.26 | 1.07 |
| CHORUS $\sigma^{\nu}_{CC}$ | 416 | 0.96 | 0.95 | 416 | 0.96 | 0.97 | 416 | 0.97 | 0.98 |
| CHORUS $\sigma^{\bar{\nu}}_{CC}$ | 416 | 0.90 | 0.88 | 416 | 0.88 | 0.87 | 416 | 0.88 | 0.88 |
| NuTeV $\sigma^{\nu}_{CC}$ (dimuon) | 39 | 0.24 | 0.22 | 39 | 0.33 | 0.33 | 39 | 1.27 | 1.28 |
| NuTeV $\sigma^{\bar{\nu}}_{CC}$ (dimuon) | 36 | 0.43 | 0.35 | 37 | 0.56 | 0.64 | 37 | 0.63 | 0.59 |
| HERA I+II $\sigma_{\rm CC}$ $e^-p$ | 42 | 1.34 | 1.19 | 42 | 1.25 | 1.29 | 42 | 1.29 | 1.34 |
| HERA I+II $\sigma_{\rm CC}$ $e^+p$ | 39 | 1.26 | 1.22 | 39 | 1.23 | 1.25 | 39 | 1.27 | 1.28 |

**Table 4.2.** Same as Table 4.1 for DIS NC (top) and DIS CC (bottom) datasets.

Table 4.1 and Fig. 4.1 show that without MHOUs fit quality improves as the perturbative order increases. On the other hand, when MHOUs are included, fit quality becomes independent of perturbative order within uncertainties (note that, with $N_{\rm dat} = 4462$, $\sigma_{\chi^2} = 0.03$). This means that the MHOU covariance matrix estimated through scale variation is correctly reproducing the true MHOUs. Also, at aN$^3$LO the fit quality is the same within uncertainties irrespective of whether MHOUs are included or not. This strongly

| Dataset | NLO | | | NNLO | | | aN$^3$LO | | |
|---|---|---|---|---|---|---|---|---|---|
| | $N_{\rm dat}$ | no MHOU | MHOU | $N_{\rm dat}$ | no MHOU | MHOU | $N_{\rm dat}$ | no MHOU | MHOU |
| E866 $\sigma^d/2\sigma^p$ (NuSea) | 15 | 0.59 | 0.47 | 15 | 0.52 | 0.51 | 15 | 0.53 | 0.51 |
| E866 $\sigma^p$ (NuSea) | 89 | 1.33 | 0.86 | 29 | 1.59 | 1.00 | 29 | 1.08 | 1.03 |
| E605 $\sigma^d/2\sigma^p$ (NuSea) | 85 | 0.43 | 0.42 | 85 | 0.46 | 0.45 | 85 | 0.45 | 0.45 |
| E906 $\sigma^d/2\sigma^p$ (SeaQuest) | 6 | 1.47 | 0.74 | 6 | 0.97 | 0.90 | 6 | 0.82 | 0.88 |
| CDF $Z$ differential | 28 | 1.23 | 1.24 | 28 | 1.23 | 1.18 | 28 | 1.23 | 1.22 |
| D0 $Z$ differential | 28 | 0.69 | 0.71 | 28 | 0.64 | 0.64 | 28 | 0.64 | 0.63 |
| ATLAS low-mass DY 7 TeV | 4 | 0.69 | 0.66 | 6 | 0.88 | 0.78 | 6 | 0.78 | 0.76 |
| ATLAS high-mass DY 7 TeV | 5 | 1.74 | 1.66 | 5 | 1.60 | 1.67 | 5 | 1.64 | 1.68 |
| ATLAS $Z$ 7 TeV ($\mathcal{L} = 35$ pb$^{-1}$) | 8 | 0.67 | 0.44 | 8 | 0.58 | 0.57 | 8 | 0.56 | 0.61 |
| ATLAS $Z$ 7 TeV ($\mathcal{L} = 4.6$ fb$^{-1}$) CC | 16 | 3.82 | 2.99 | 24 | 1.80 | 1.68 | 24 | 1.66 | 1.69 |
| ATLAS $Z$ 7 TeV ($\mathcal{L} = 4.6$ fb$^{-1}$) CF | 15 | 1.77 | 1.22 | 15 | 1.07 | 1.02 | 15 | 1.02 | 0.99 |
| ATLAS low-mass DY 2D 8 TeV | 47 | 1.38 | 0.94 | 60 | 1.23 | 1.08 | 60 | 1.17 | 1.13 |
| ATLAS high-mass DY 2D 8 TeV | 48 | 1.52 | 1.38 | 48 | 1.11 | 1.08 | 48 | 1.09 | 1.09 |
| ATLAS $\sigma_Z^{\rm tot}$ 13 TeV | 1 | 0.12 | 0.41 | 1 | 0.24 | 0.60 | 1 | 0.24 | 0.66 |
| ATLAS $Z$ $p_T$ 8 TeV ($p_T, m_{\ell\ell}$) | 41 | 1.08 | 0.91 | 44 | 0.91 | 0.91 | 44 | 0.89 | 0.89 |
| ATLAS $Z$ $p_T$ 8 TeV ($p_T, y_Z$) | 28 | 0.87 | 0.52 | 48 | 0.90 | 0.70 | 48 | 0.77 | 0.68 |
| CMS DY 2D 7 TeV | 88 | 1.29 | 1.11 | 110 | 1.34 | 1.32 | 110 | 1.34 | 1.36 |
| CMS $Z$ $p_T$ 8 TeV | 28 | 1.66 | 1.47 | 28 | 1.40 | 1.41 | 28 | 1.35 | 1.39 |
| LHCb $Z \to ee$ 7 TeV | 9 | 1.47 | 1.18 | 9 | 1.65 | 1.53 | 9 | 1.48 | 1.46 |
| LHCb $Z \to \mu$ 7 TeV | 15 | 1.03 | 0.87 | 15 | 0.80 | 0.73 | 15 | 0.77 | 0.73 |
| LHCb $Z \to ee$ 8 TeV | 17 | 1.58 | 1.38 | 17 | 1.24 | 1.26 | 17 | 1.31 | 1.27 |
| LHCb $Z \to \mu$ 8 TeV | 15 | 1.25 | 1.06 | 15 | 1.44 | 1.59 | 15 | 1.60 | 1.60 |
| LHCb $Z \to ee$ 13 TeV | 15 | 1.68 | 1.60 | 15 | 1.72 | 1.80 | 15 | 1.78 | 1.76 |
| LHCb $Z \to \mu\mu$ 13 TeV | 16 | 1.10 | 1.11 | 16 | 0.94 | 0.99 | 16 | 0.99 | 0.94 |
| D0 $W$ muon asymmetry | 8 | 2.42 | 2.17 | 9 | 1.86 | 1.95 | 9 | 2.07 | 2.03 |
| ATLAS $W$ 7 TeV ($\mathcal{L} = 35$ pb$^{-1}$) | 22 | 1.20 | 1.13 | 22 | 1.11 | 1.12 | 22 | 1.09 | 1.12 |
| ATLAS $W$ 7 TeV ($\mathcal{L} = 4.6$ fb$^{-1}$) | 22 | 2.18 | 2.13 | 22 | 2.08 | 2.16 | 22 | 2.16 | 2.10 |
| ATLAS $\sigma_W^{\rm tot}$ 13 TeV | 2 | 0.16 | 0.54 | 2 | 1.21 | 1.60 | 2 | 1.38 | 1.67 |
| ATLAS $W^+$+jet 8 TeV | 15 | 0.26 | 0.28 | 15 | 0.79 | 0.79 | 15 | 0.73 | 0.73 |
| ATLAS $W^-$+jet 8 TeV | 15 | 0.98 | 1.27 | 15 | 1.49 | 1.45 | 15 | 1.41 | 1.41 |
| CMS $W$ electron asymmetry 7 TeV | 11 | 0.92 | 1.03 | 11 | 0.84 | 0.85 | 11 | 0.82 | 0.86 |
| CMS $W$ muon asymmetry 7 TeV | 11 | 2.03 | 1.77 | 11 | 1.71 | 1.73 | 11 | 1.70 | 1.71 |
| CMS $W$ rapidity 8 TeV | 22 | 0.93 | 0.74 | 22 | 1.33 | 1.03 | 22 | 1.11 | 1.08 |
| LHCb $W \to \mu$ 7 TeV | 14 | 1.63 | 1.26 | 14 | 2.78 | 1.99 | 14 | 2.12 | 2.03 |
| LHCb $W \to \mu$ 8 TeV | 14 | 0.60 | 0.44 | 14 | 0.97 | 0.92 | 14 | 0.80 | 0.84 |

**Table 4.3.** Same as Table 4.1 for DY NC (top) and DY CC (bottom) datasets.

suggests that within current experimental uncertainties, the current methodology and the current dataset the perturbative expansion has converged, meaning that inclusion of higher-order QCD correction beyond N$^3$LO would not lead to a further improvement in fit quality to current data.

## 4.2 Parton distributions

We now examine the NNPDF4.0 aN$^3$LO parton distributions. We compare the NLO, NNLO and aN$^3$LO NNPDF4.0 PDFs, obtained without and with inclusion of MHOUs, in Figs. 4.2-4.3 and in Figs. 4.4-4.5, respectively. Specifically, we show the up, antiup, down, antidown, strange, antistrange, charm and gluon PDFs at $Q = 100$ GeV, normalized to the aN$^3$LO result, as a function of $x$ in logarithmic and linear scale. Error bands correspond to one sigma PDF uncertainties, which do (MHOU sets) or do not (no MHOU sets) include MHOUs on all theory predictions used in the fit. The PDF sets, with and without MHOUs, are the same used to compute the values of the $\chi^2$ in Tables 4.1-4.4.

The excellent perturbative convergence seen in the fit quality is also manifest at the level of PDFs. In

| Dataset | NLO | | | NNLO | | | aN$^3$LO | | |
|---|---|---|---|---|---|---|---|---|---|
| | $N_{\rm dat}$ | no MHOU | MHOU | $N_{\rm dat}$ | no MHOU | MHOU | $N_{\rm dat}$ | no MHOU | MHOU |
| ATLAS $\sigma_{tt}^{\rm tot}$ 7 TeV | 1 | 10.4 | 0.96 | 1 | 4.50 | 2.40 | 1 | 2.78 | 2.05 |
| ATLAS $\sigma_{tt}^{\rm tot}$ 8 TeV | 1 | 1.74 | 0.59 | 1 | 0.02 | 0.03 | 1 | 0.04 | 0.08 |
| ATLAS $\sigma_{tt}^{\rm tot}$ 13 TeV ($\mathcal{L}$=139 fb$^{-1}$) | 1 | 3.82 | 0.96 | 1 | 0.49 | 0.41 | 1 | 0.51 | 0.44 |
| ATLAS $t\bar{t}$ $\ell$+jets 8 TeV ($1/\sigma d\sigma/dy_t$) | 4 | 4.16 | 1.79 | 4 | 3.13 | 3.70 | 4 | 2.98 | 3.64 |
| ATLAS $t\bar{t}$ $\ell$+jets 8 TeV ($1/\sigma d\sigma/dy_{t\bar{t}}$) | 4 | 8.93 | 3.73 | 4 | 4.50 | 5.80 | 4 | 4.26 | 4.92 |
| ATLAS $t\bar{t}$ 2$\ell$ 8 TeV ($1/\sigma d\sigma/dy_{t\bar{t}}$) | 4 | 1.94 | 1.76 | 4 | 1.60 | 1.86 | 4 | 1.66 | 1.80 |
| CMS $\sigma_{tt}^{\rm tot}$ 5 TeV | 1 | 0.61 | 0.73 | 1 | 0.02 | 0.01 | 1 | 0.03 | 0.02 |
| CMS $\sigma_{tt}^{\rm tot}$ 7 TeV | 1 | 5.27 | 1.30 | 1 | 1.01 | 0.50 | 1 | 0.60 | 0.34 |
| CMS $\sigma_{tt}^{\rm tot}$ 8 TeV | 1 | 3.50 | 0.85 | 1 | 0.26 | 0.17 | 1 | 0.21 | 0.10 |
| CMS $\sigma_{tt}^{\rm tot}$ 13 TeV | 1 | 0.75 | 0.26 | 1 | 0.06 | 0.01 | 1 | 0.04 | 0.05 |
| CMS $t\bar{t}$ $\ell$+jets 8 TeV ($1/\sigma d\sigma/dy_{t\bar{t}}$) | 9 | 1.87 | 1.59 | 9 | 1.21 | 1.59 | 9 | 1.31 | 1.52 |
| CMS $t\bar{t}$ 2D 2$\ell$ 8 TeV ($1/\sigma d\sigma/dy_t dm_{t\bar{t}}$) | 15 | 2.03 | 1.89 | 15 | 1.30 | 1.25 | 15 | 1.28 | 1.37 |
| CMS $t\bar{t}$ 2$\ell$ 13 TeV ($d\sigma/dy_t$) | 10 | 0.78 | 0.69 | 10 | 0.51 | 0.59 | 10 | 0.55 | 0.60 |
| CMS $t\bar{t}$ $\ell$+jet 13 TeV ($d\sigma/dy_t$) | 11 | 0.66 | 0.25 | 11 | 0.60 | 0.66 | 11 | 0.52 | 0.71 |
| ATLAS incl. jets 8 TeV, $R = 0.6$ | 171 | 0.67 | 0.74 | 171 | 0.68 | 0.64 | 171 | 0.68 | 0.64 |
| CMS incl. jets 8 TeV | 185 | 0.95 | 0.83 | 185 | 1.19 | 0.95 | 185 | 0.97 | 0.99 |
| ATLAS dijets 7 TeV, $R = 0.6$ | 90 | 1.47 | 1.72 | 90 | 2.14 | 1.69 | 90 | 1.76 | 1.63 |
| CMS dijets 7 TeV | 54 | 1.57 | 2.01 | 54 | 1.79 | 1.74 | 54 | 1.84 | 1.78 |
| ATLAS isolated $\gamma$ prod. 13 TeV | 53 | 0.57 | 0.47 | 53 | 0.76 | 0.67 | 53 | 0.72 | 0.68 |
| ATLAS single $t$ $R_t$ 7 TeV | 1 | 0.43 | 0.29 | 1 | 0.50 | 0.57 | 1 | 0.51 | 0.58 |
| ATLAS single $t$ $R_t$ 13 TeV | 1 | 0.04 | 0.03 | 1 | 0.06 | 0.07 | 1 | 0.06 | 0.07 |
| ATLAS single $t$ 7 TeV ($1/\sigma d\sigma/dy_t$) | 3 | 0.83 | 0.84 | 3 | 0.96 | 0.94 | 3 | 0.97 | 0.97 |
| ATLAS single $t$ 7 TeV ($1/\sigma d\sigma/dy_{\bar{t}}$) | 3 | 0.06 | 0.06 | 3 | 0.06 | 0.06 | 3 | 0.06 | 0.06 |
| ATLAS single $t$ 8 TeV ($1/\sigma d\sigma/dy_t$) | 3 | 0.38 | 0.31 | 3 | 0.25 | 0.26 | 3 | 0.22 | 0.24 |
| ATLAS single $t$ 8 TeV ($1/\sigma d\sigma/dy_{\bar{t}}$) | 3 | 0.19 | 0.21 | 3 | 0.19 | 0.19 | 3 | 0.20 | 0.20 |
| CMS single $t$ $\sigma_t + \sigma_{\bar{t}}$ 7 TeV | 1 | 0.89 | 0.88 | 1 | 0.74 | 0.84 | 1 | 0.39 | 0.43 |
| CMS single $t$ $R_t$ 8 TeV | 1 | 0.15 | 0.08 | 1 | 0.18 | 0.20 | 1 | 0.18 | 0.21 |
| CMS single $t$ $R_t$ 13 TeV | 1 | 0.33 | 0.27 | 1 | 0.36 | 0.38 | 1 | 0.36 | 0.38 |

**Table 4.4.** Same as Table 4.1 for (from top to bottom) top pair, single-inclusive jet, isolated photon and single top production datasets.

particular, the NNLO PDFs are either very close to or indistinguishable from their aN$^3$LO counterparts. Inclusion of MHOUs further improves the consistency between NNLO and aN$^3$LO PDFs, which lie almost on top of each other. This means that the NNLO PDFs are made more accurate by the inclusion of MHOUs, and that the aN$^3$LO PDFs have converged, in the sense discussed above. Exceptions to this stability are the charm and gluon PDFs, for which aN$^3$LO corrections have a sizeable impact. In the case of charm, they lead to an enhancement of the central value of about 4% for $x \sim 0.05$; in the case of gluon, to a suppression of about 2-3% for $x \sim 0.005$. In both cases, inclusion of MHOUs leads to an increase in PDF uncertainties by about 1-2%. This makes the NNLO and aN$^3$LO charm PDFs with MHOUs compatible within uncertainties, and the NNLO and aN$^3$LO gluon PDFs with MHOU almost compatible.

Figure 4.6 presents a comparison similar to that of Figs. 4.2-4.4 for the gluon-gluon, gluon-quark, quark-quark, and quark-antiquark parton luminosities. These are shown integrated in rapidity as a function of the invariant mass of the final state $m_X$ for a center-of-mass energy $\sqrt{s} = 14$ TeV. Their definition follows Eqs. (1)-(4) of Ref. [110].

As already observed for PDFs, perturbative convergence is excellent, and improves upon inclusion of MHOUs. The NNLO and aN$^3$LO results are compatible within uncertainties for the gluon-quark, quark-quark, and quark-antiquark luminosities. Some differences are seen for the gluon-gluon luminosity, consistent with the differences seen in the gluon PDF. Specifically, the aN$^3$LO corrections lead to a suppression of the gluon-gluon luminosity of 2-3% for $m_X \sim 100$ GeV. This effect is somewhat compensated by an increase in uncertainty of about 1% upon inclusion of MHOUs. Indeed, the NNLO and aN$^3$LO gluon-gluon luminosities

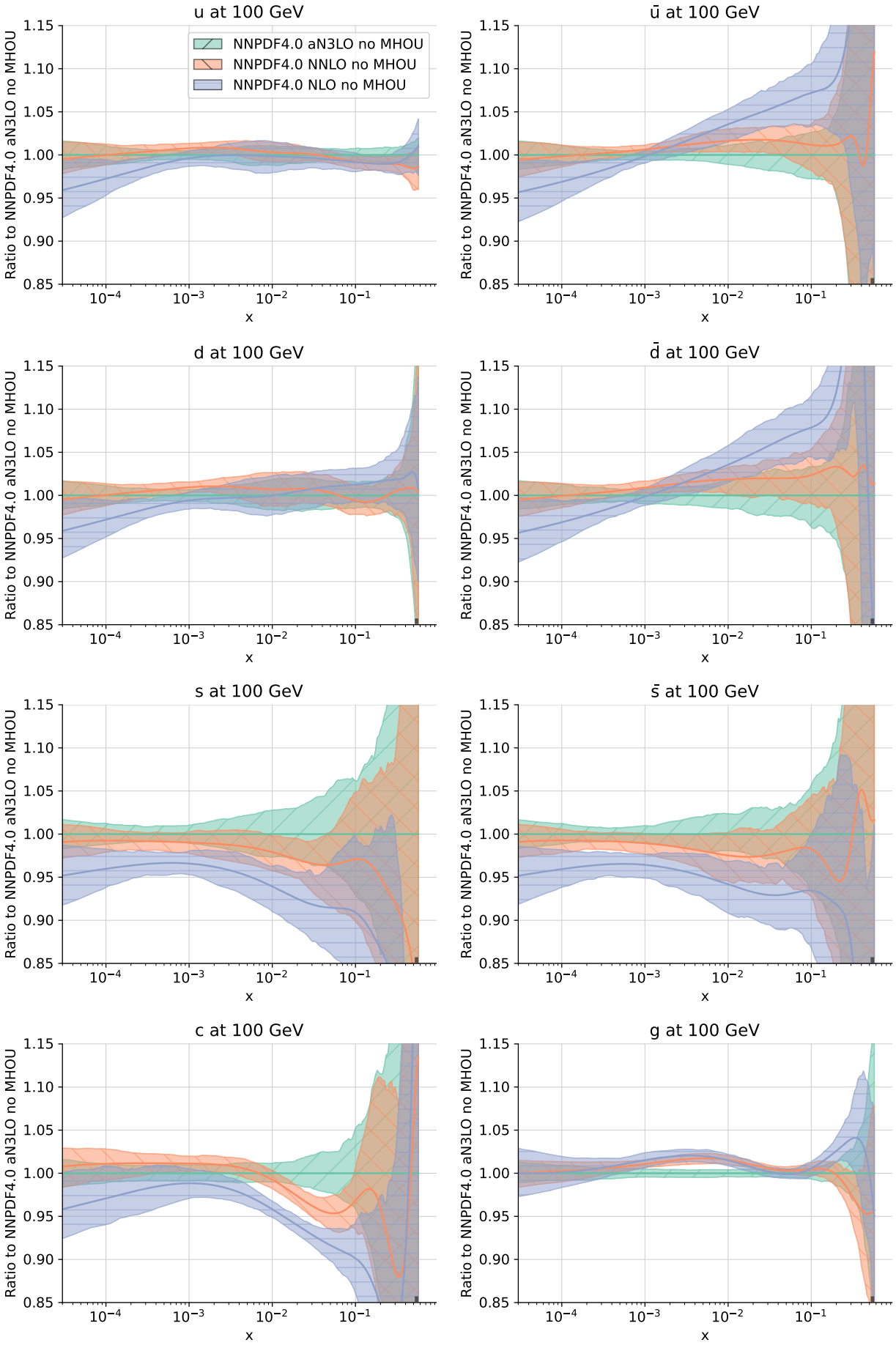



**Figure 4.2.** The NLO, NNLO and aN$^3$LO NNPDF4.0 PDFs at $Q = 100$ GeV. We display the up, antiup, down, antidown, strange, antistrange, charm and gluon PDFs normalized to the aN$^3$LO result. Error bands correspond to one sigma PDF uncertainties, not including MHOUs on the theory predictions used in the fit.

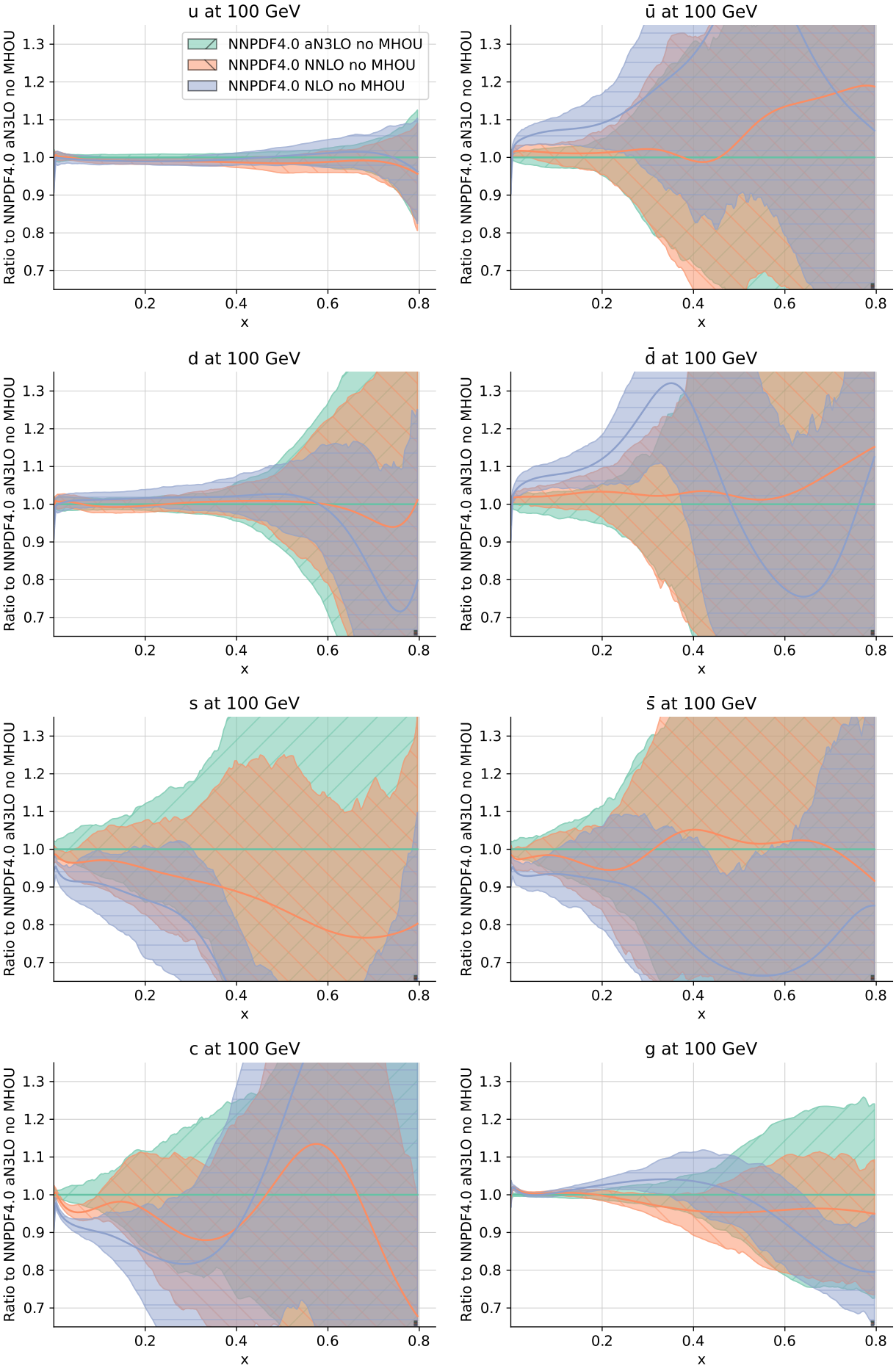



**Figure 4.3.** Same as Fig. 4.2 in linear scale.

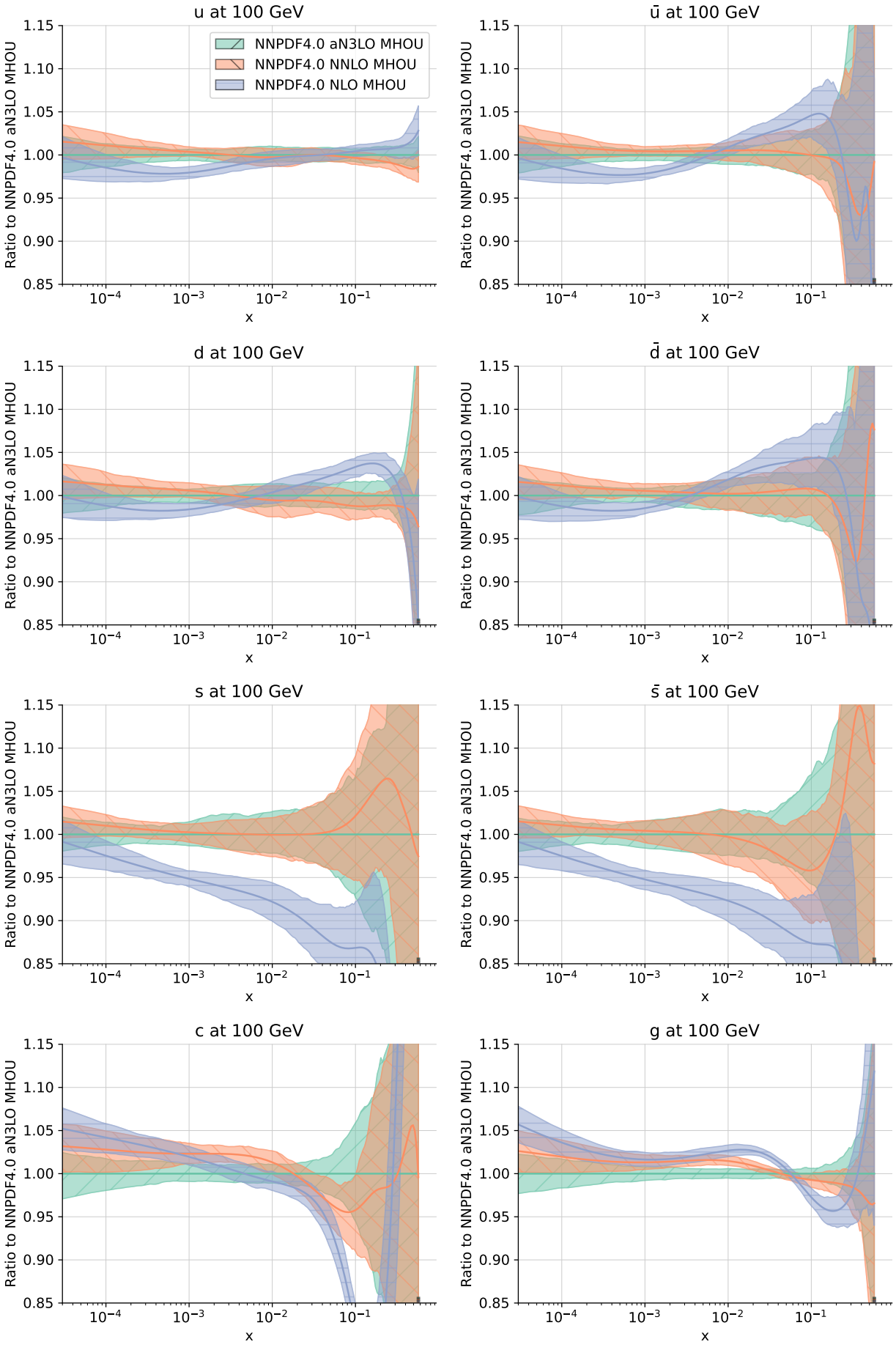



**Figure 4.4.** Same as Fig. 4.2 for NNPDF4.0MHOU PDF sets. Error bands correspond to one sigma PDF uncertainties also including MHOUs on the theory predictions used in the fit.

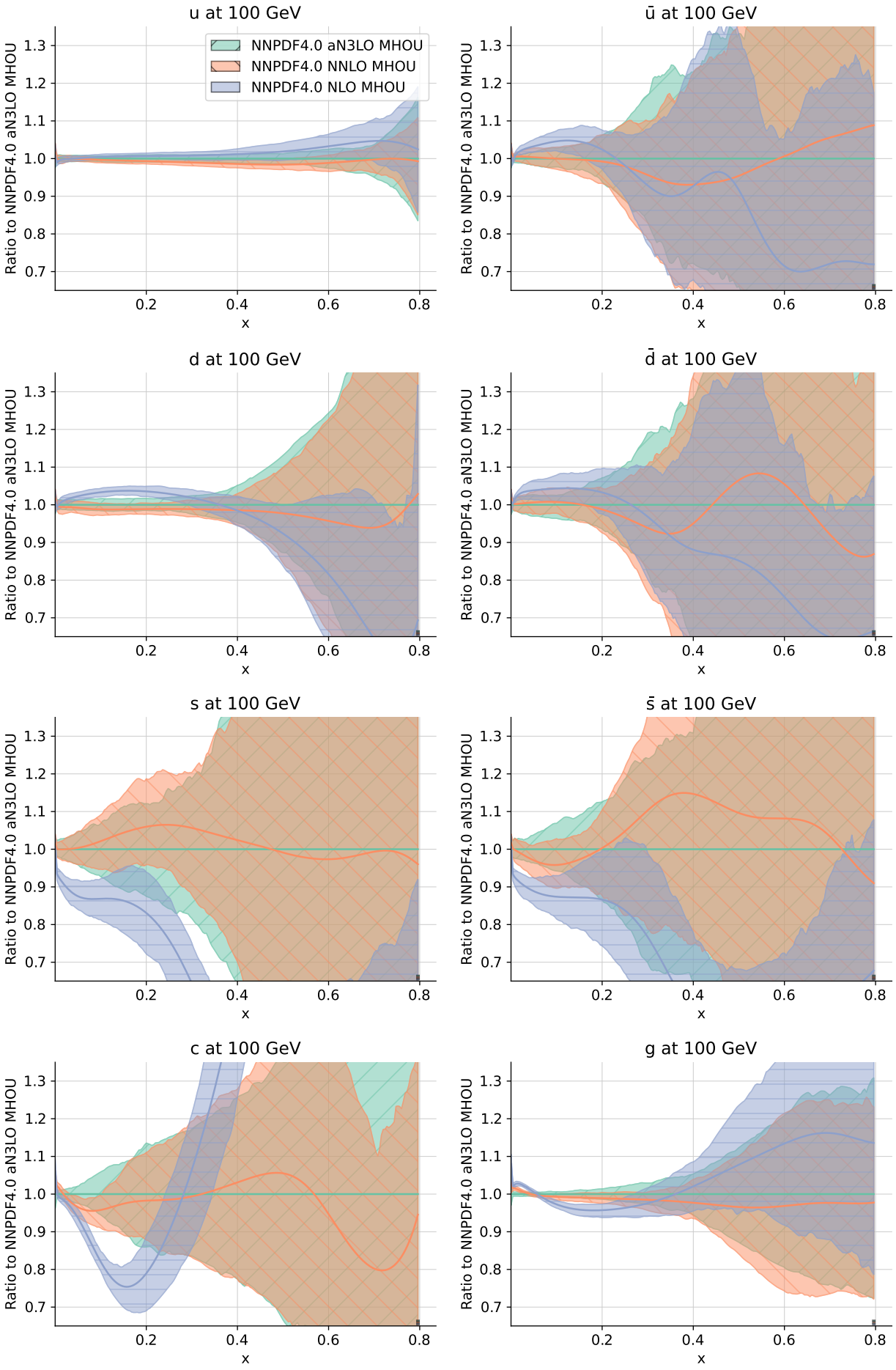



**Figure 4.5.** Same as Fig. 4.4 in linear scale.

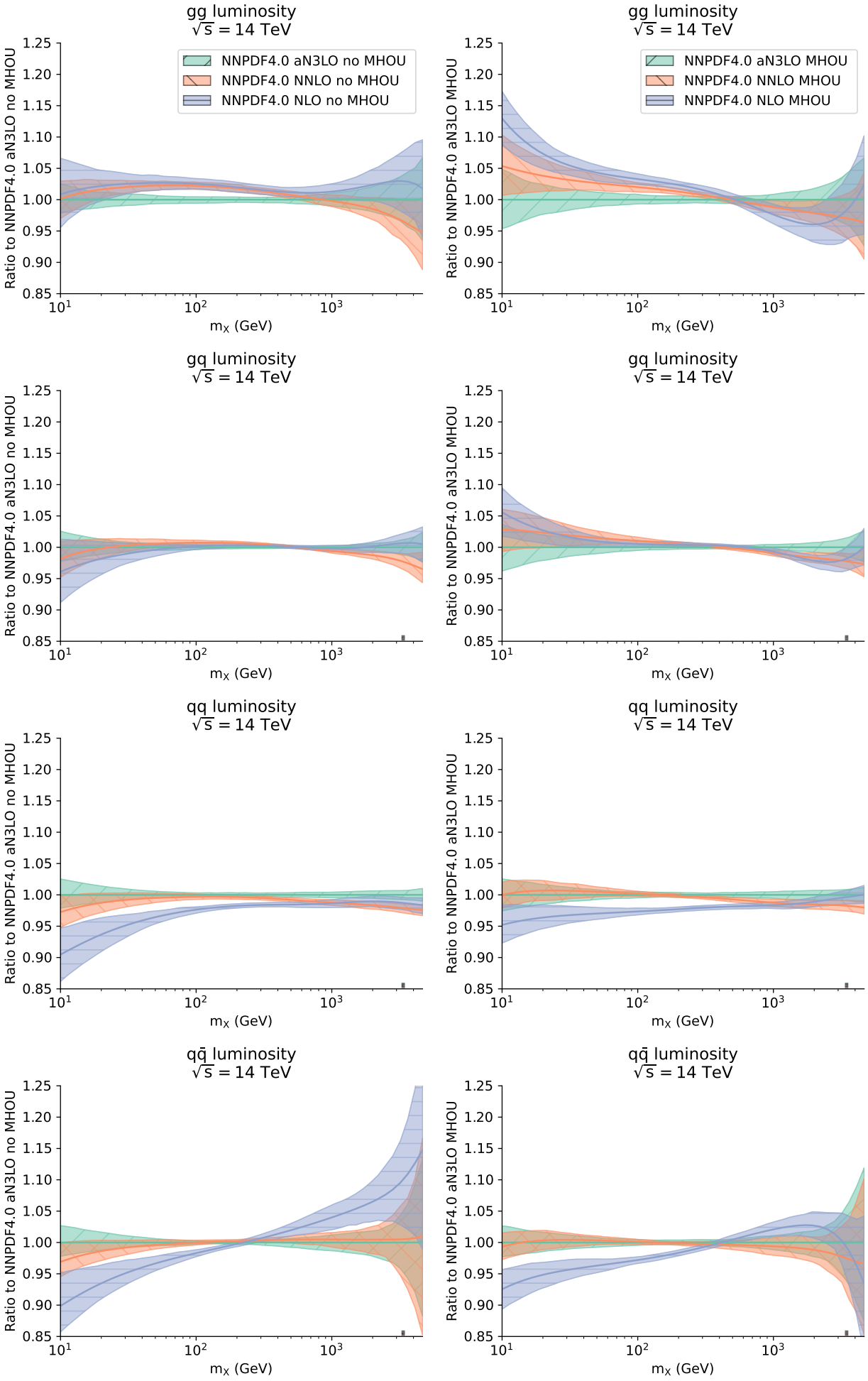



**Figure 4.6.** The gluon-gluon, gluon-quark, quark-quark, and quark-antiquark parton luminosities as a function of $m_X$ at $\sqrt{s} = 14$ TeV, computed with NLO, NNLO and aN$^3$LO NNPDF4.0 PDFs without MHOUs (left) and with MHOUs (right), all shown as a ratio to the respective aN$^3$LO results. Uncertainties are as in Figs. 4.2-4.4

| Dataset | NLO | | NNLO | | N$^3$LO | |
|---|---|---|---|---|---|---|
| | no MHOU | MHOU | no MHOU | MHOU | no MHOU | MHOU |
| DIS NC | 0.14 | 0.13 | 0.15 | 0.13 | 0.13 | 0.13 |
| DIS CC | 0.11 | 0.11 | 0.12 | 0.12 | 0.12 | 0.12 |
| DY NC | 0.19 | 0.17 | 0.18 | 0.17 | 0.17 | 0.18 |
| DY CC | 0.33 | 0.27 | 0.35 | 0.32 | 0.31 | 0.32 |
| Top pairs | 0.18 | 0.17 | 0.17 | 0.17 | 0.16 | 0.19 |
| Single-inclusive jets | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 | 0.13 |
| Dijets | 0.10 | 0.10 | 0.11 | 0.10 | 0.10 | 0.10 |
| Prompt photons | 0.06 | 0.07 | 0.06 | 0.06 | 0.05 | 0.05 |
| Single top | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| Total | 0.18 | 0.15 | 0.16 | 0.15 | 0.15 | 0.15 |

**Table 4.5.** The $\phi$ uncertainty estimator for NNPDF4.0 PDFs at NLO, NNLO and aN$^3$LO without and with MHOUs for the process categories as in Table 4.1.

for $m_X \sim 100$ GeV differ by about $2.5\sigma$ without MHOU, but become almost compatible within uncertainties when MHOUs are included.

All in all, these results show that aN$^3$LO corrections are generally small, except for the gluon PDF, and that at aN$^3$LO the perturbative expansion has all but converged, with NNLO and aN$^3$LO PDFs very close to each other, especially upon inclusion of MHOUs. They also show that MHOUs generally improve the accuracy of PDFs, though at aN$^3$LO they have a very small impact. The phenomenological consequences of this state of affairs will be further discussed in Sect. 5.

### 4.3 PDF uncertainties

We now take a closer look at PDF uncertainties. In Fig. 4.7 we display one sigma uncertainties for the NNPDF4.0 NLO, NNLO, and aN$^3$LO PDFs with and without MHOUs at $Q = 100$ GeV. All uncertainties are normalized to the central value of the NNPDF4.0 aN$^3$LO PDF set with MHOUs. The NLO uncertainty is generally the largest of all in the absence of MHOUs, and for quark distributions the smallest once MHOUs are included. All other uncertainties, at NNLO and aN$^3$LO, with and without MHOUs, are quite similar to each other, especially for quark PDFs. The fact that upon inclusion of an extra source of uncertainty, namely the MHOU, PDF uncertainties are reduced (at NLO) or unchanged (at NNLO and aN$^3$LO) may look counter-intuitive. However, as already pointed out in Refs. [34, 38, 111], this can be understood to be a consequence of the increased compatibility of the data due to inclusion of MHOUs and of higher-order perturbative corrections.

The impact of MHOUs on NLO and NNLO PDFs was extensively assessed in Ref. [38]. In a similar vein, here we focus on the impact of MHOUs on aN$^3$LO PDFs. To this purpose, in Fig. 4.8 we compare the NNPDF4.0 aN$^3$LO PDFs with and without MHOUs. The related comparison for parton luminosities is presented in Fig. 4.9. Again, aN$^3$LO PDFs and luminosities with and without MHOU are very compatible with each other. This evidence reinforces the expectation that perturbative corrections beyond N$^3$LO will not alter PDFs significantly, at least with current data and methodology.

In analogy with Ref. [38], we also compare the $\phi$ estimator introduced in Ref. [112] (see Eq. (4.6) there). The estimator gives the ratio of the average correlated PDF uncertainty to the data uncertainty. As such, it provides an estimate of the consistency of the data: consistent data are combined by the underlying theory and lead to an uncertainty in the prediction which is smaller than that of the original data. The value of $\phi$ obtained in the NLO, NNLO, and aN$^3$LO NNPDF4.0 fits with and without MHOUs (as in Table 4.1) is reported in Table 4.5. It is clear that $\phi$ converges to very similar values with the increase of the perturbative order and/or with inclusion of MHOUs for both the total dataset and for most of the data categories. This fact is further quantitative evidence of the excellent perturbative convergence of the PDF uncertainties.

### 4.4 Implications for intrinsic charm

The availability of the aN$^3$LO PDFs discussed in Sects. 4.2-4.3 allows us to revisit and consolidate our recent results on intrinsic charm. Specifically, based on the NNPDF4.0 NNLO PDF determination, we have found

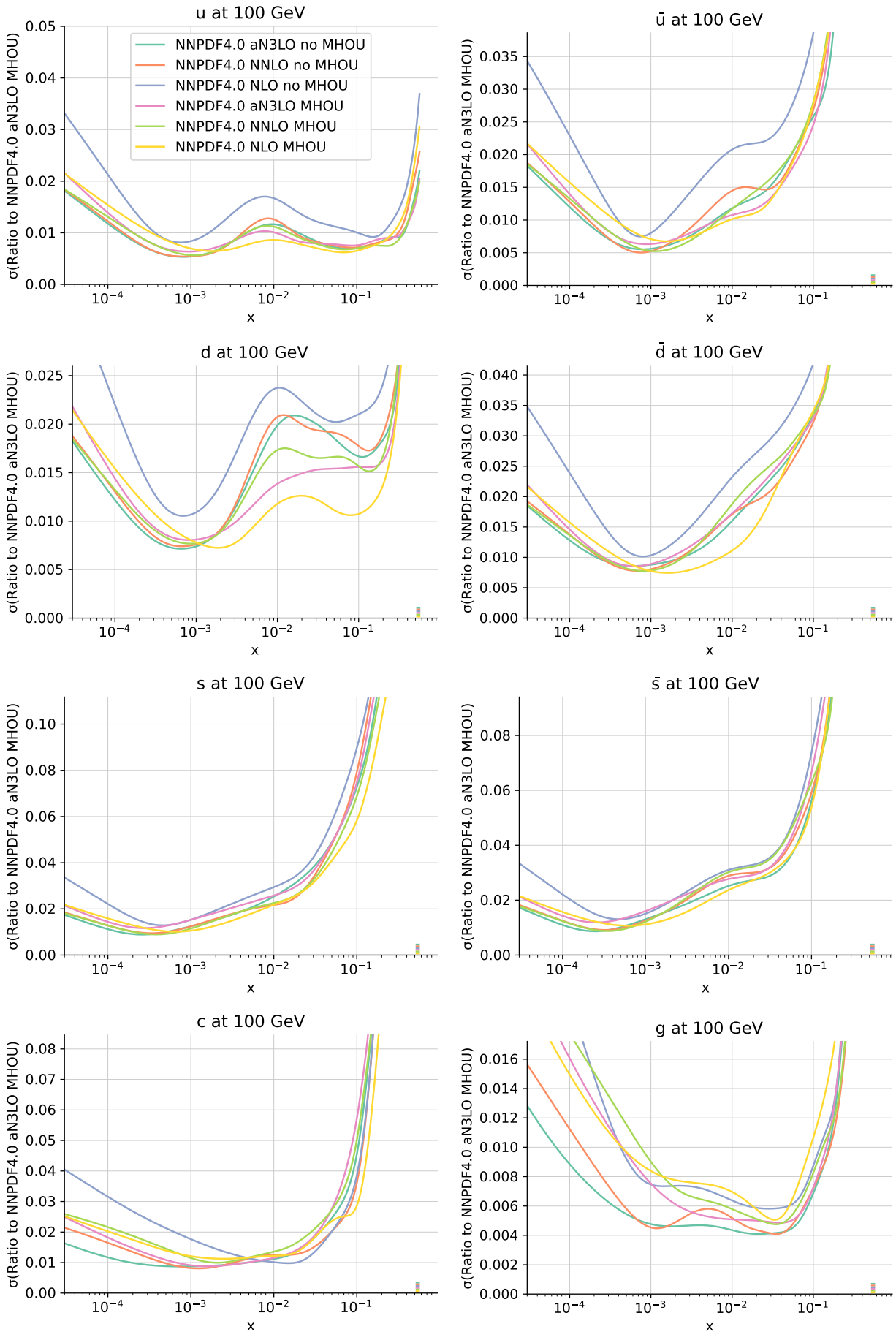



**Figure 4.7.** Relative one sigma uncertainties for the PDFs shown in Figs. 4.2-4.4. All uncertainties are normalized to the central value of the NNPDF4.0 aN$^3$LO set with MHOUs.

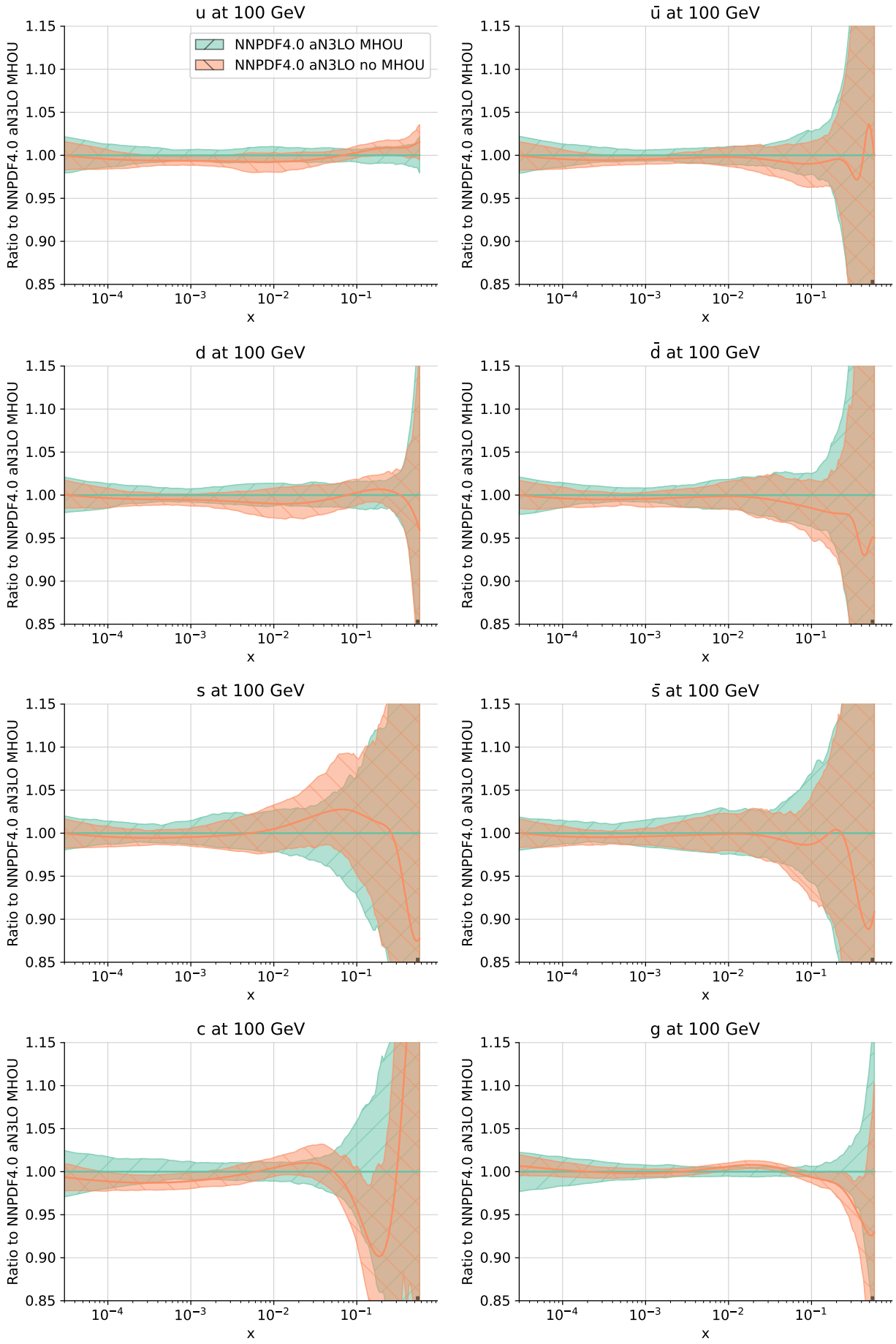



**Figure 4.8.** Same as Figs. 4.2-4.4, now comparing NNPDF4.0 aN$^3$LO PDFs without and with MHOUs.

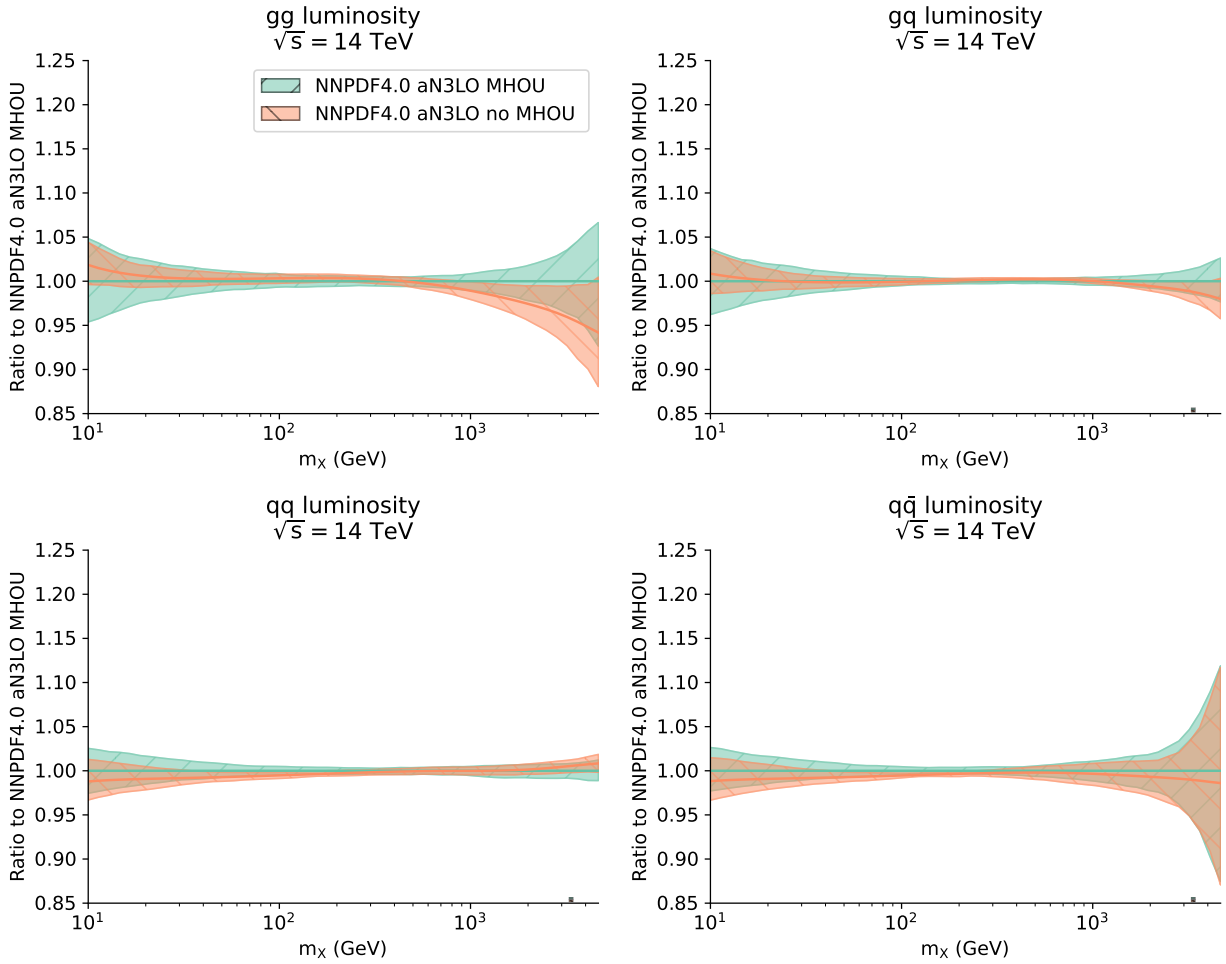



**Figure 4.9.** Same as Fig. 4.6, now comparing NNPDF4.0 aN$^3$LO PDFs without and with MHOUs.

evidence for intrinsic charm [113] and an indication for a non-vanishing valence charm component [114]. In these analyses, the dominant source of theory uncertainty was estimated to come from the matching conditions that are used in order to obtain PDFs in a three-flavor charm decoupling scheme from high-scale data, while MHOUs were assumed to be subdominant. The uncertainty in the matching conditions was in turn estimated by comparing results obtained using NNLO matching and the best available aN$^3$LO matching conditions, both applied to NNLO PDFs.

It is now possible to improve these results on three counts. First, we can now fully include MHOUs. Second, we can consistently combine aN$^3$LO matching conditions and aN$^3$LO PDFs, and perform a consistent comparison of NNLO and aN$^3$LO results. Finally, knowledge of aN$^3$LO matching conditions themselves is now improved thanks to recent results [81] that were not available at the time of the analysis of Ref. [113]. We will specifically discuss the determination of the total intrinsic charm component and we do not consider the valence component, because effects of MHOUs and of the flavor scheme transformation are already very small at NNLO [114].

To this purpose, in Fig. 4.10 we show the total charm PDF, $xc^+(x,Q^2)$, in the 4FNS at $Q = 1.65$ GeV and in the 3FNS, as obtained from using NNPDF4.0 NLO, NNLO and aN$^3$LO without and with MHOUs. Note that in the 3FNS the charm PDF does not depend on scale. Error bands correspond to one sigma PDF uncertainties. The 4FNS results share the general features discussed in Sect. 4.2: the perturbative expansion converges nicely, with the aN$^3$LO result very close to the NNLO. The convergence is further improved by the inclusion of MHOUs, which move the NNLO yet closer to the aN$^3$LO. The 3FNS result is especially remarkable: whereas the combination of aN$^3$LO matching with NNLO PDFs, used in Ref. [113] to conservatively estimate MHOUs, was somewhat unstable, now results display complete stability, and in particular the NNLO and aN$^3$LO results completely coincide.

In order to assess the impact of MHOUs more clearly, in Fig. 4.11 we compare the total charm PDF in the 3FNS with and without MHOUs, respectively at NNLO and aN$^3$LO. At NNLO MHOUs have a small but non-negligible impact on central values, with almost unchanged uncertainty, but at aN$^3$LO they have

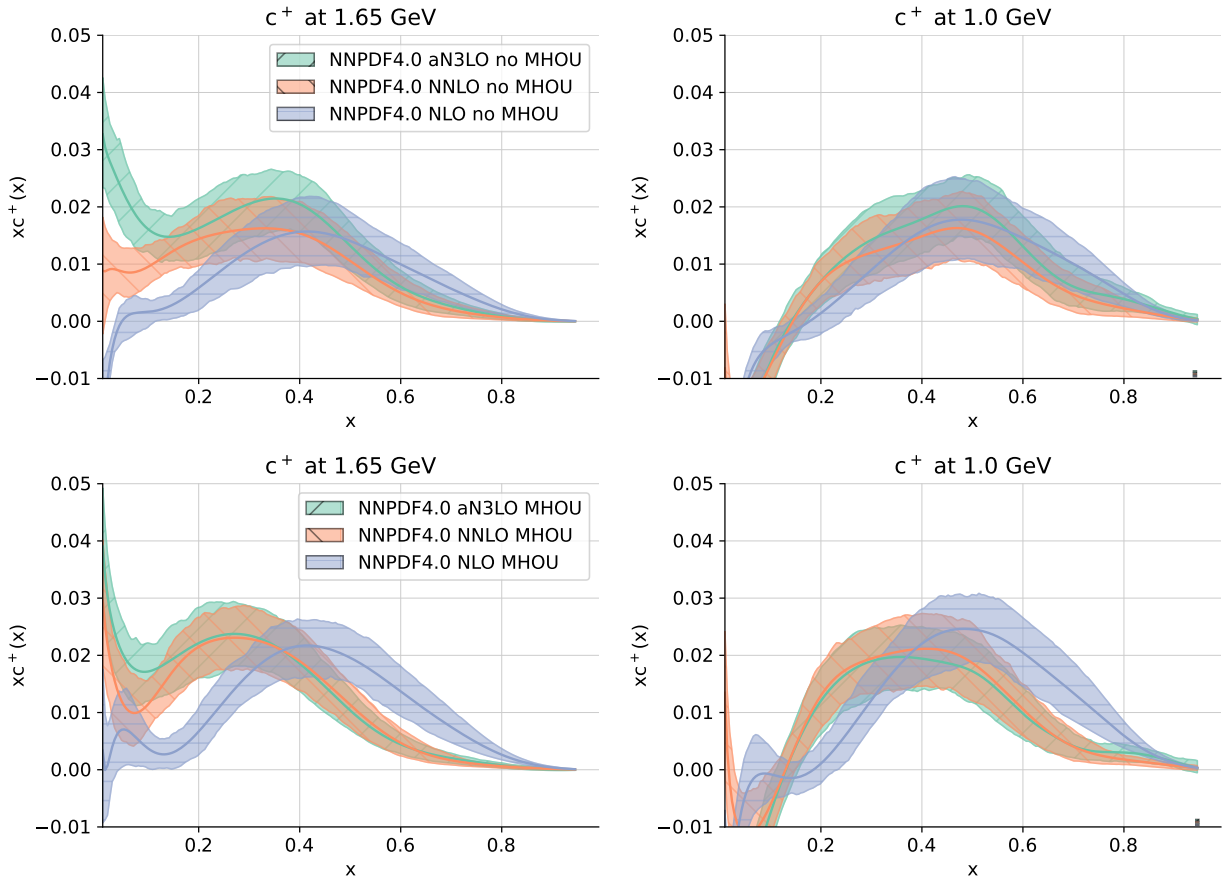



**Figure 4.10.** The total charm PDF, $xc^+(x, Q^2)$, in the 4FNS at $Q = 1.65$ GeV (left) and 3FNS (right), as obtained from the NNPDF4.0 NLO, NNLO, and aN$^3$LO fits without (top) and with (bottom) MHOUs. Error bands correspond to one sigma PDF uncertainties. Note that in the 3FNS the charm PDF does not depend on scale.

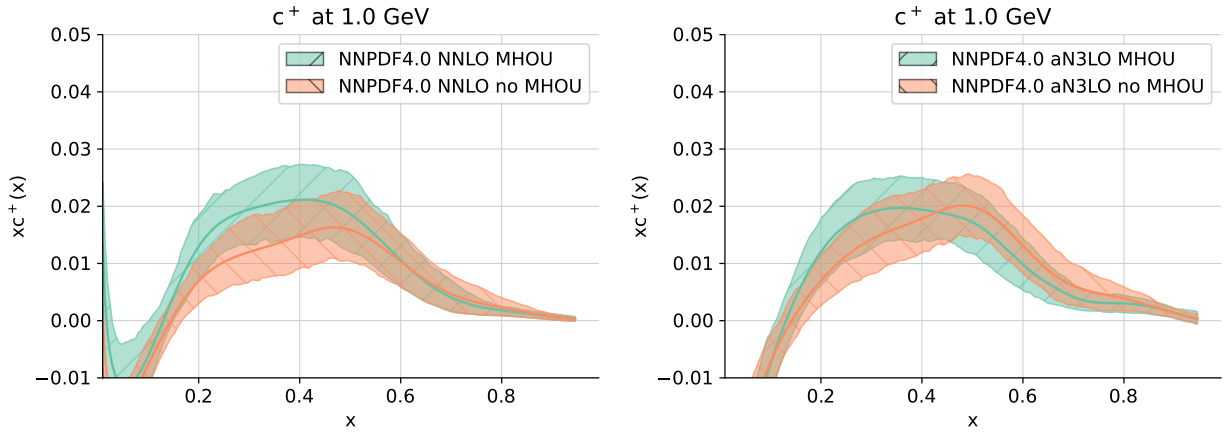



**Figure 4.11.** Same as Fig. 4.10, now comparing the total charm PDF in the 3FNS with and without MHOUs, respectively at NNLO (left) and aN$^3$LO (right).

essentially no impact, confirming the perturbative convergence of the result.

We thus proceed to a final re-assessment of the significance of intrinsic charm through the pull, defined as the central value divided by total uncertainty, using NNPDF4.0MHOU NNLO and aN$^3$LO PDFs. We estimate the total uncertainty by adding in quadrature to the PDF uncertainty (which already includes the MHOU from the theory predictions used in the fit) a further theory uncertainty, taken equal to the difference between the central value at given perturbative order, and that at the previous perturbative order (so at NNLO from the difference to NLO, and so on). This now also includes the MHOU due to change in the matching from 4FNS to 3FNS, but also the shift in the 4FNS result that is in principle already accounted for by the MHOU. Also, it conservatively assumes that the shift between the current order and the next is equal to that from the previous order, rather than smaller. Results obtained with this conservative error

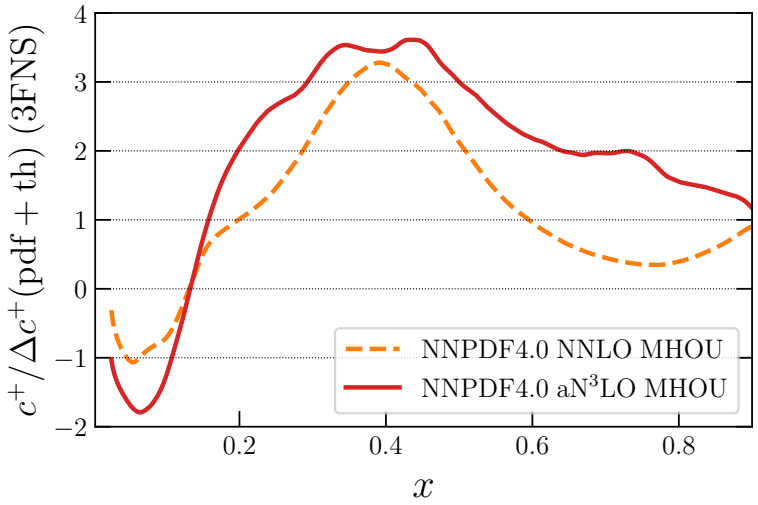



**Figure 4.12.** The pull (central value divided by total uncertainty) for the total charm PDF in the 3FNS obtained in the NNPDF4.0 NNLO and aN$^3$LO fits with MHOUs.

estimate are shown in Fig. 4.12. It is clear that the significance of intrinsic charm is increased somewhat when going from NNLO to aN$^3$LO. It is now also somewhat increased already at NNLO in comparison to the result of Ref. [113], despite the more conservative uncertainty estimate, thanks to the increased accuracy of MHOU PDFs and the consistent and improved treatment of matching aN$^3$LO conditions. Indeed, local significance at the peak is now more than three sigma for the default fit.

## 4.5 Dependence on the treatment of aN$^3$LO corrections.

We now discuss two variants in the treatment of aN$^3$LO corrections: a different approximation to perturbative evolution, and a different treatment of hadronic cross-sections.

The aN$^3$LO PDF sets presented in Sect. 4.2 are based on our approximation to the full set of N$^3$LO splitting functions presented in Sect. 2. This approximation was compared in Sect. 2.7 to that recently presented in Refs. [22, 28–30] (FHMRUVV). In order to fully assess the impact of this different approximation to N$^3$LO perturbative evolution we have repeated the aN$^3$LO PDF determination, without and with MHOUs, but now using the FHMRUVV approximation to N$^3$LO perturbative evolution, with everything else unchanged. PDFs obtained using the FHMRUVV approximation vs our own are compared in Fig. 4.13 without MHOUs and in Fig. 4.14 with MHOUs. PDFs are displayed at $Q = 100$ GeV, normalized to the central value of our default result. Differences turn out to be completely negligible, as might have been expected given the good agreement at the level of splitting functions seen in Fig. 2.9.

As discussed in Sect. 3.3 N$^3$LO corrections to hadronic hard cross-sections are not included in our default aN$^3$LO PDF determination. We can however assess the impact of the inclusion of all the publicly available corrections for all relevant data in the NNPDF4.0 dataset, listed in Tab. 3.1 and discussed in Sect. 3.3. To this purpose, we have repeated the aN$^3$LO PDF determination, with and without MHOUs, now with N$^3$LO for these datasets included. The ensuing PDFs are compared to the default NNPDF4.0 aN$^3$LO in Fig. 4.15 without MHOUs and in Fig. 4.16 with MHOUs. In both cases, PDFs are displayed at $Q = 100$ GeV and are normalized to the central value of the corresponding default NNPDF4.0 aN$^3$LO set. The impact of these N$^3$LO corrections is only (barely) visible in the PDFs without MHOUs, and even in this case it only significantly impacts the strange and antistrange PDFs, for which it leads to an enhancement of 4-5% for $x \sim 0.1$, though well within the PDF uncertainty. Even this small effect is absent in the PDFs with MHOUs. We conclude that at present available N$^3$LO corrections for hadronic processes have no effect on PDF determination.

## 4.6 Comparison with MSHT20

We compare the NNPDF4.0 aN$^3$LO PDF set to the only other existing aN$^3$LO PDF set, MSHT20 aN$^3$LO [32]. As already discussed in Sect. 2.7, MSHT20 aN$^3$LO PDFs are determined by fitting to the data the nuisance parameters that parametrize the IHOU uncertainty on a prior approximation to splitting functions. It follows that the ensuing central value is partly determined by the data, and the IHOU is entirely data-driven. When comparing NNPDF4.0 and MSHT20 aN$^3$LO PDF sets it should of course be borne in mind

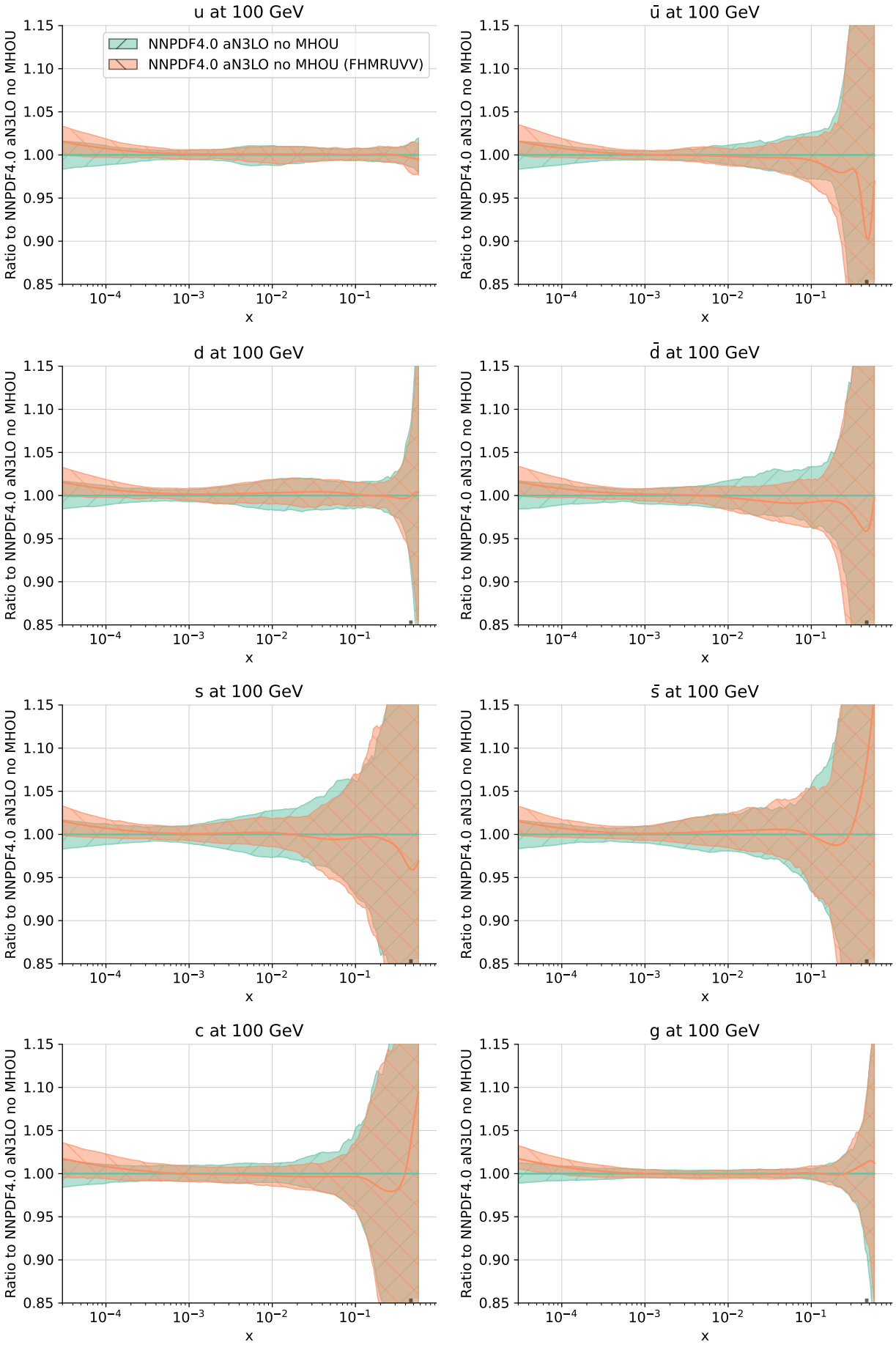



**Figure 4.13.** Same as Fig. 4.2, now comparing the NNPDF4.0 aN$^3$LO PDFs and PDFs obtained using the FHMRUVV approximation to N$^3$LO perturbative evolution.

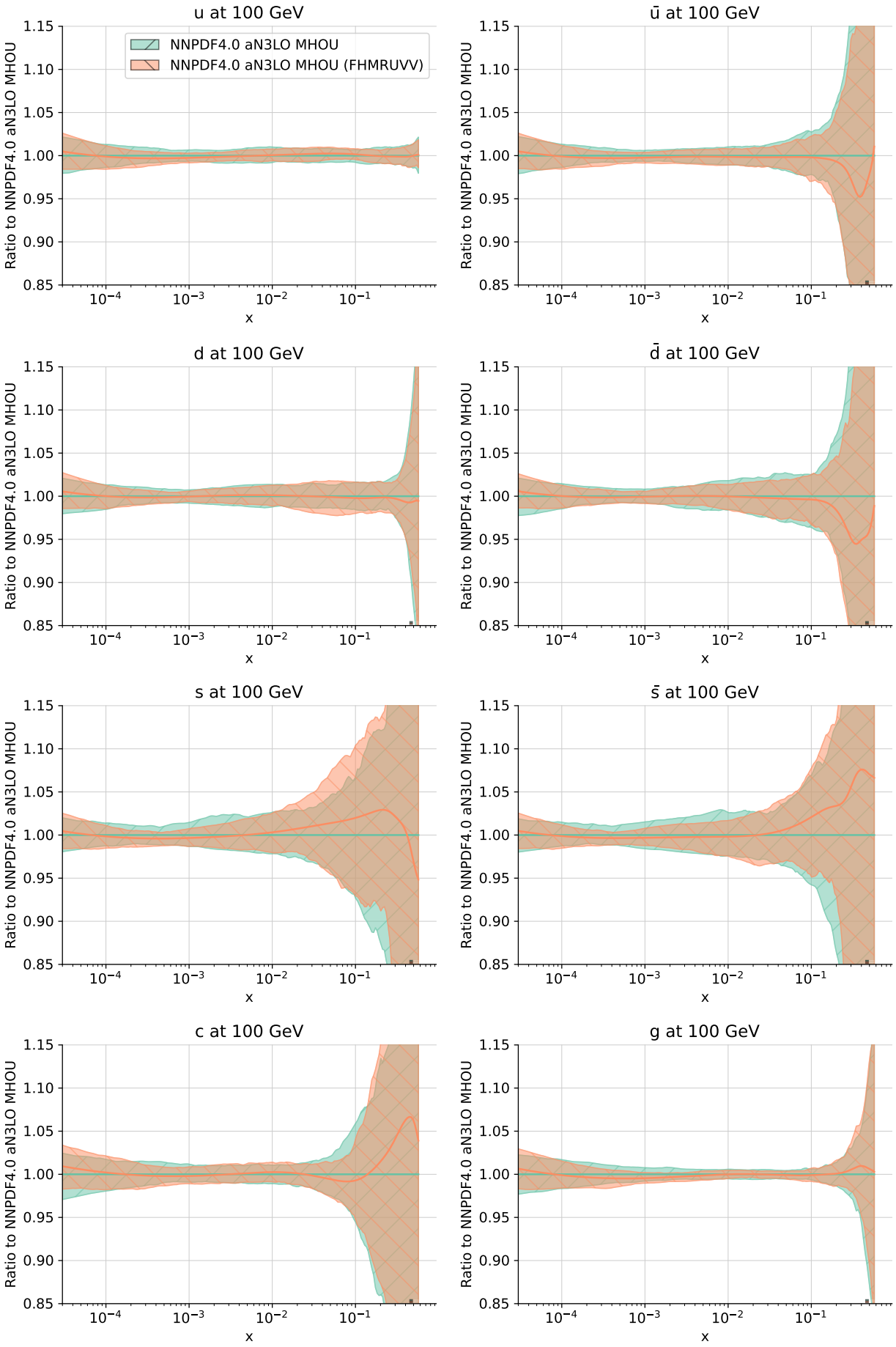



**Figure 4.14.** Same as Fig. 4.13 for aN$^3$LO PDF sets with MHOUs.

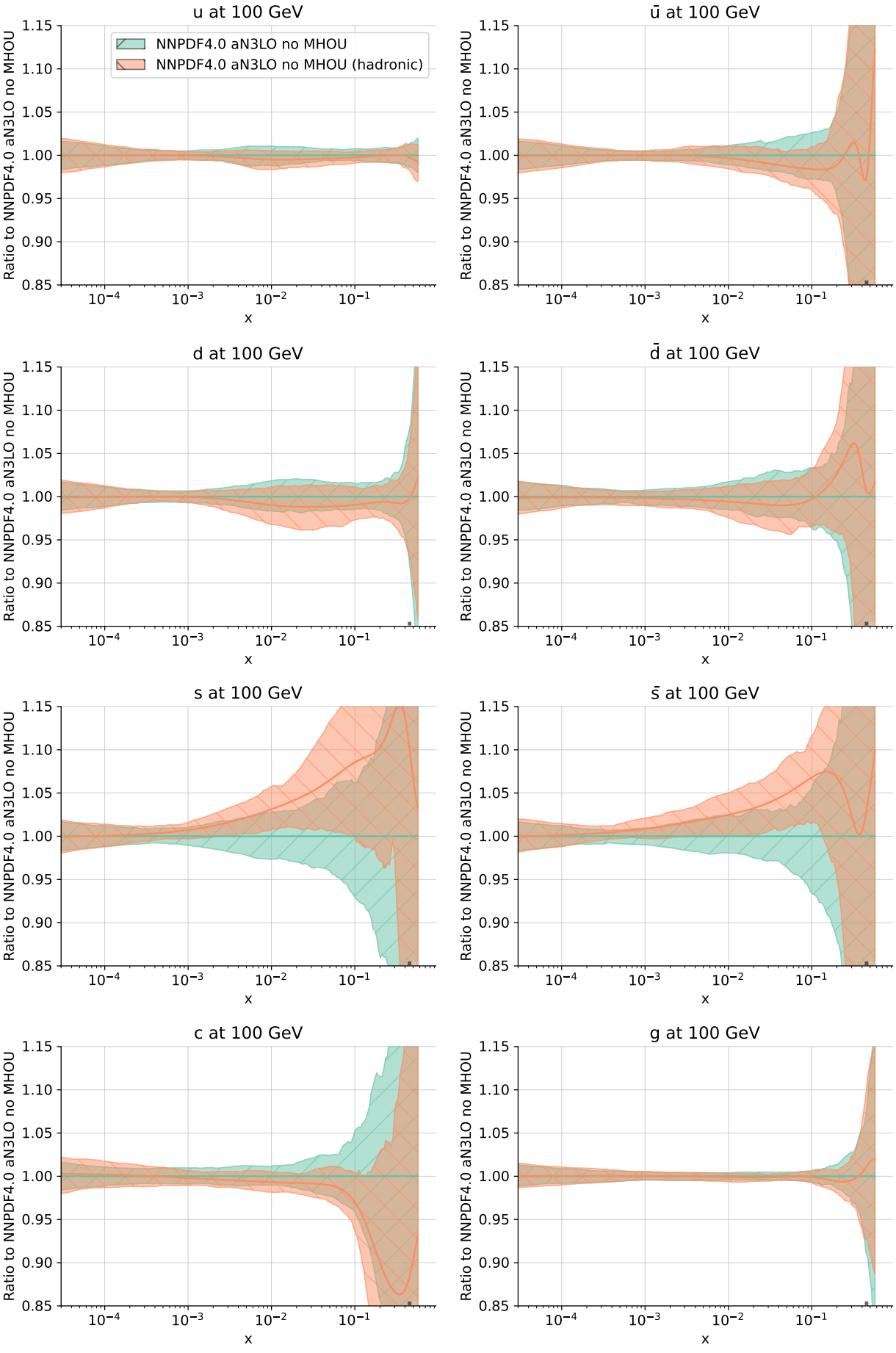



**Figure 4.15.** Same as Fig. 4.2, now comparing the NNPDF4.0 aN$^3$LO baseline PDF set without MHOUs to a variant obtained with inclusion of N$^3$LO corrections for the hadronic processes of Tab. 3.1.

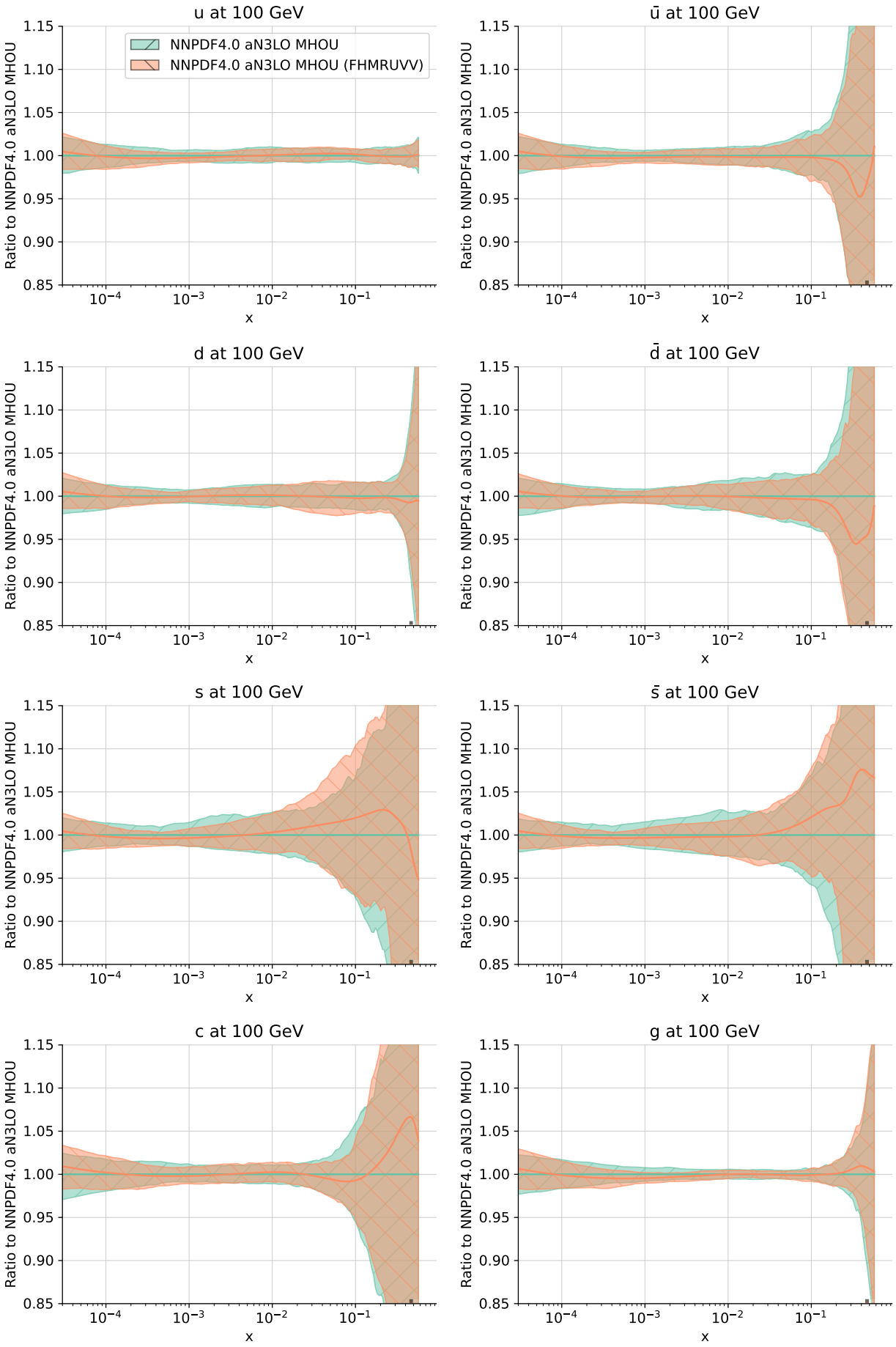



**Figure 4.16.** Same as Fig. 4.15 for aN$^3$LO PDF sets with MHOUs.

that the sets already differ at NNLO due to differences in dataset and methodology. The NNLO MSHT20 and NNPDF4.0 PDF sets were compared in Fig. 21 and the corresponding parton luminosity in Fig. 60 of Ref. [37], while a detailed benchmarking was presented in Ref. [115].

The comparison of the aN$^3$LO sets is presented in Fig. 4.17, where we show the NNPDF4.0 no MHOU set and the MSHT20 set recommended as baseline in Ref. [32] at $Q = 100$ GeV, normalized to the NNPDF4.0 central value. All error bands are one sigma uncertainties. The dominant differences between the PDF sets are the same as already observed at NNLO, with the largest difference observed for the charm PDF, which is independently parametrized in NNPDF4.0, but not in MSHT20, where it is determined by perturbative matching conditions. However, the differences, while remaining qualitatively similar, are slightly reduced (by 1-2%) when moving from NNLO to aN$^3$LO. Exceptions are the charm and especially the gluon PDF, which differ more at aN$^3$LO. Specifically, the gluon PDF, while reasonably compatible for $x \lesssim 0.07$ at NNLO, disagrees at aN$^3$LO, with the MSHT20 result suppressed by 3-4% in the region $10^{-3} \lesssim x \lesssim 10^{-1}$, with a PDF uncertainty of 1-2%. This suppression of the MSHT20 gluon can likely be traced to the behaviour of the $P_{gq}$ splitting function seen in Fig. 2.9.

Parton luminosities are compared in Fig. 4.18. Again the pattern is similar to that seen at NNLO, but now with a considerable suppression of the gluon-gluon and gluon-quark luminosities in the $M_X \sim 100$ GeV region that can be traced to the behaviour of the gluon PDF seen in Fig. 4.17. The quark-quark luminosity remains similar in MSHT20 and NNPDF4.0 both at NNLO and aN$^3$LO. The impact of these effects on the computation of precision LHC cross sections will be addressed in Sect. 5.

In order to understand better the comparative impact of aN$^3$LO corrections, we compare for each set the NNLO and aN$^3$LO luminosities. Results are shown in Fig. 4.19, normalized to the aN$^3$LO result. The qualitative impact of the aN$^3$LO corrections on either set is similar, but with a stronger aN$^3$LO suppression of gluon luminosities for MSHT20. In particular the gluon-gluon luminosity is suppressed for $10^2 \lesssim m_X \lesssim 10^3$ GeV by about 3% in NNPDF4.0 and 6% in MSHT20 and the gluon-quark luminosity is suppressed in the same region by about 1% in NNPDF4.0 and 3% in MSHT20. In the case of the gluon-gluon luminosity the differences between NNLO and aN$^3$LO are larger than the respective PDF uncertainties (that do not include MHOUs in either case). As already mentioned in Sect. 2.7, a dedicated benchmark of aN$^3$LO results is ongoing and will be presented in Ref. [84].

# 5 LHC phenomenology at aN$^3$LO accuracy

We present a first assessment of the implications of aN$^3$LO PDFs for LHC phenomenology, by looking at processes for which N$^3$LO results are publicly available, namely the Drell-Yan and Higgs total inclusive cross-sections. We present predictions at NLO, NNLO, and aN$^3$LO using both NNPDF4.0 and MSHT20 PDFs, consistently matching the perturbative order of the PDF and matrix element. At N$^3$LO we also show results obtained with the currently common approximation of using NNLO PDFs with aN$^3$LO matrix elements.

At each perturbative order, the uncertainty on the cross-section is determined by adding in quadrature the PDF uncertainty to the MHOU on the hard matrix element determined performing 7-point renormalization and factorization scale variation and taking the envelope of the results. This is the procedure that is most commonly used for the estimation of the total uncertainty on hadron collider processes, and we follow it here for ease of comparison with available results. In a more refined treatment, MHOUs on the hard cross section can be included through a theory covariance matrix for the hard cross section itself, like the MHOUs and IHOUs on the PDF. This would then make it possible to keep track of the correlation between these different sources of uncertainty [116–118]. Here, in order to disentangle the contribution of the MHOU in the processes used for PDF determination from that in the matrix element, we show results with the PDF uncertainty evaluated with both MHOU and no-MHOU PDF sets.

## 5.1 Inclusive Drell-Yan production

We show results for inclusive charged-current and neutral-current gauge boson production cross-sections followed by their decays into the dilepton final state. Cross-sections are evaluated using the N3LOXS code [8] for different ranges in the final-state dilepton invariant mass, $Q = m_{\ell\ell}$ for neutral-current and $Q = m_{\ell\nu}$ for charged-current scattering. Fig. 5.1 displays the inclusive neutral-current Drell-Yan cross-section $pp \to \gamma^*/Z \to \ell^+\ell^-$ and Figs. 5.2-5.3 the charged-current cross-sections $pp \to W^\pm \to \ell^\pm\nu_\ell$. We consider one low-mass bin (30 GeV $\leq Q \leq$ 60 GeV), the mass peak bin (60 GeV $\leq Q \leq$ 120 GeV), and two high-mass

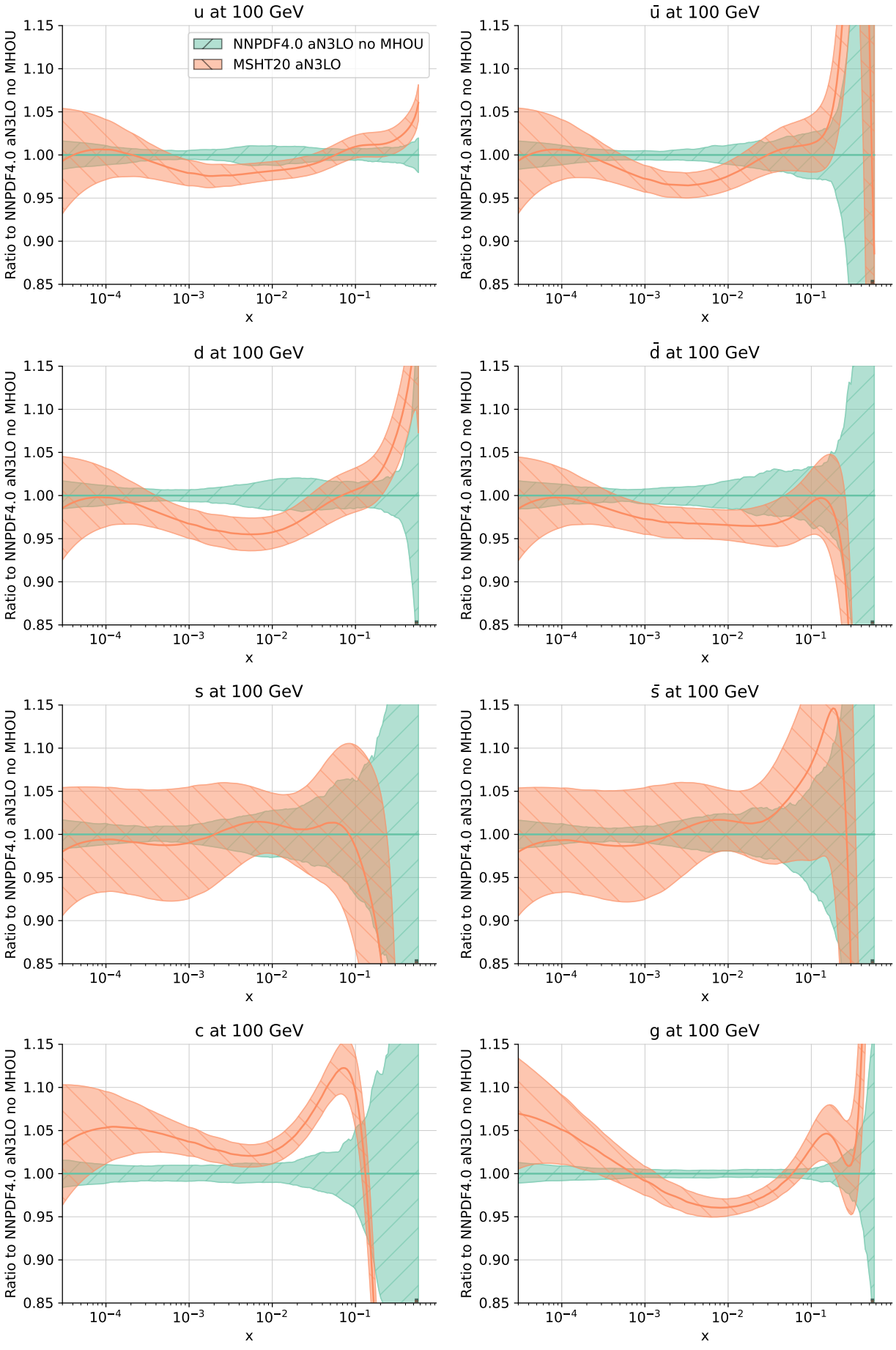



**Figure 4.17.** Same as Fig. 4.2, now comparing the NNPDF4.0 aN$^3$LO baseline PDF set without MHOUs to the MSHT20 set recommended as baseline in Ref. [32].

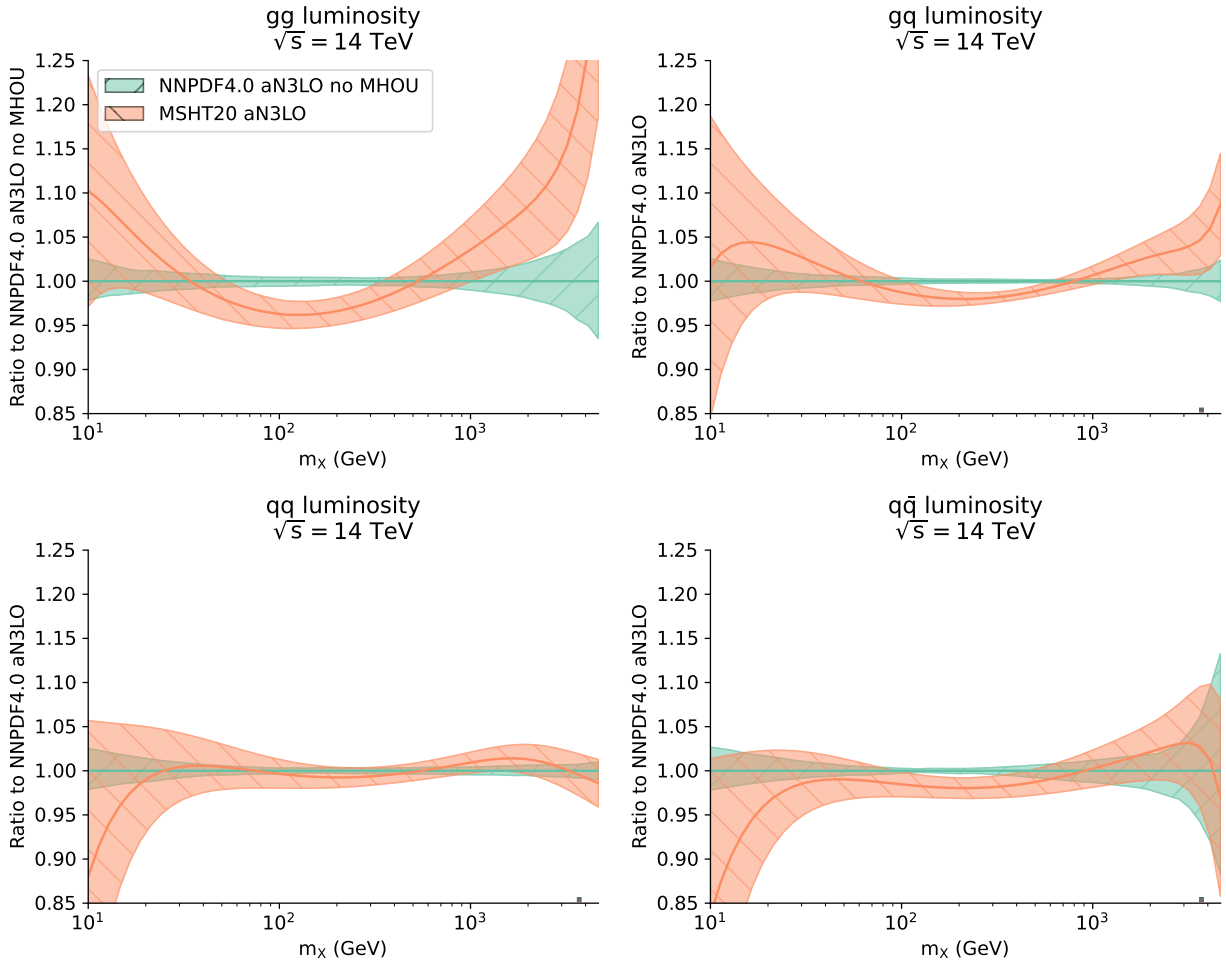



**Figure 4.18.** Same as Fig. 4.17 for parton luminosities as in Fig. 4.6.

bins (120 GeV $\leq Q \leq$ 300 GeV and 2 TeV $\leq Q \leq$ 3 TeV), relevant for high-mass new physics searches [119]. In all cases, we compare the NLO, NNLO, and aN$^3$LO predictions using NNPDF4.0 and MSHT20 PDFs determinations, with the same perturbative order in matrix element and PDFs, and also the aN$^3$LO result with NNLO PDFs, and then we compare the aN$^3$LO with NNPDF4.0 aN$^3$LO PDFs with and without MHOUs.

In general we observe a good perturbative convergence, with predictions at two subsequent orders in agreement within uncertainties, and generally improved convergence upon including MHOUs on the PDF. Predictions based on NNPDF4.0 and MSHT20 are always consistent with each other within uncertainties. From these comparisons we can draw two main conclusions. First, in many cases differences between the NNLO and N$^3$LO predictions tend to be reduced when using consistently the appropriate PDFs at each order, rather than NNLO PDFs with N$^3$LO matrix elements (though in some cases the results are unchanged). For instance, for the two lowest $m_{\ell\ell}$ bins for NC production aN$^3$LO PDFs drive upwards the N$^3$LO prediction, making it closer to the NNLO result. Second, the difference between PDFs with and without MHOUs, while moderate, remains non-negligible even at N$^3$LO, where it starts being comparable to the overall uncertainty, and thus it must be included in precision calculations.

## 5.2 Inclusive Higgs production

We now consider Higgs production in gluon fusion, in associated production with vector bosons, and via vector-boson fusion (VBF). Predictions are obtained using the ggHiggs code [120] for gluon fusion, n3loxs for associate production, and proVBFH code [121] for VBF. Results are shown in Fig. 5.4 for gluon-fusion and VBF, and Fig. 5.5 for associate production with $W^+$ and $Z$.

Here too we observe generally good perturbative convergence, even for gluon fusion, that notoriously has a very slowly converging expansion. Also in this case, there is generally better agreement between NNPDF4.0 and MSHT20 as the perturbative order increases, except for gluon fusion where the agreement is similar at all orders. Indeed, in all cases MSHT20 and NNPDF4.0 results agree within uncertainties at aN$^3$LO, while

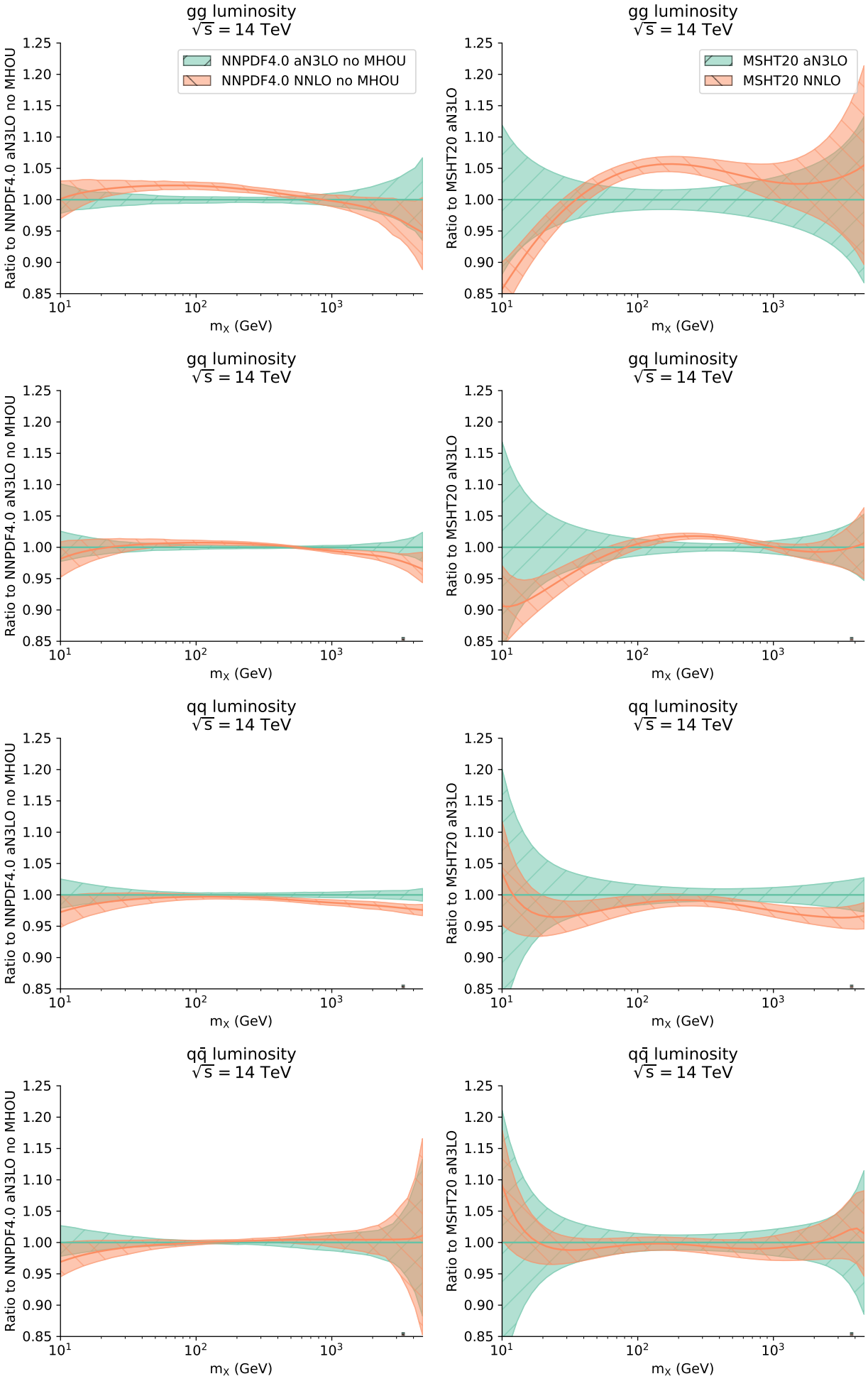



**Figure 4.19.** Same as Fig. 4.18, now comparing aN$^3$LO and NNLO parton luminosities, separately for the NNPDF4.0 (left) and MSHT20 (right) PDF sets, normalized to the aN$^3$LO result.

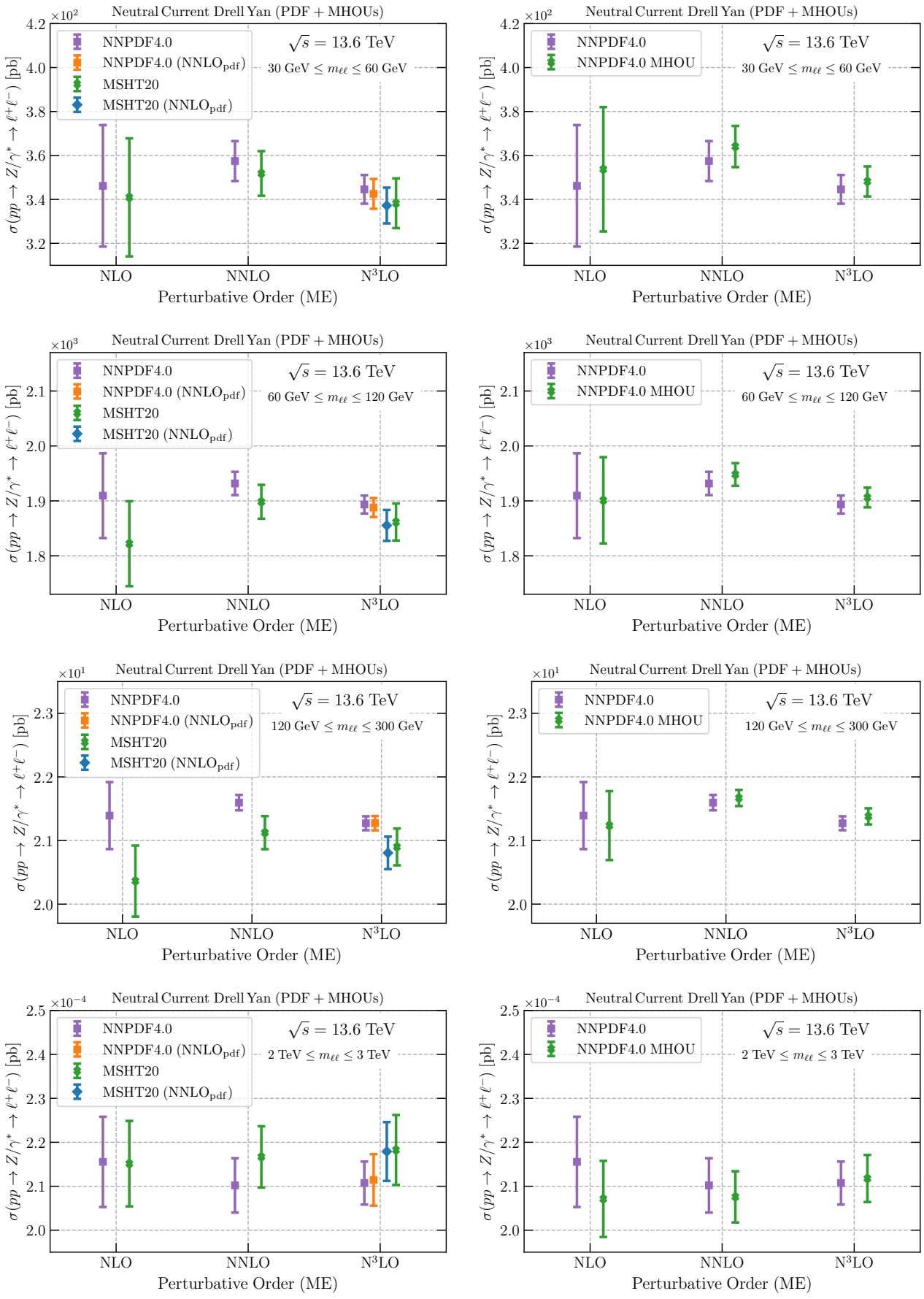



**Figure 5.1.** The inclusive neutral-current Drell-Yan production cross-section, $pp \to \gamma^*/Z \to \ell^+\ell^-$, for different ranges of the dilepton invariant mass $Q = m_{\ell\ell}$, from low to high invariant masses (top to bottom). Results are shown (left) comparing NLO, NNLO and aN$^3$LO with matched perturbative order in the matrix element and PDF, and also at aN$^3$LO with NNLO PDFs using NNPDF4.0 and MSHT20 PDFs and (right) at aN$^3$LO with PDFs without and with MHOUs.

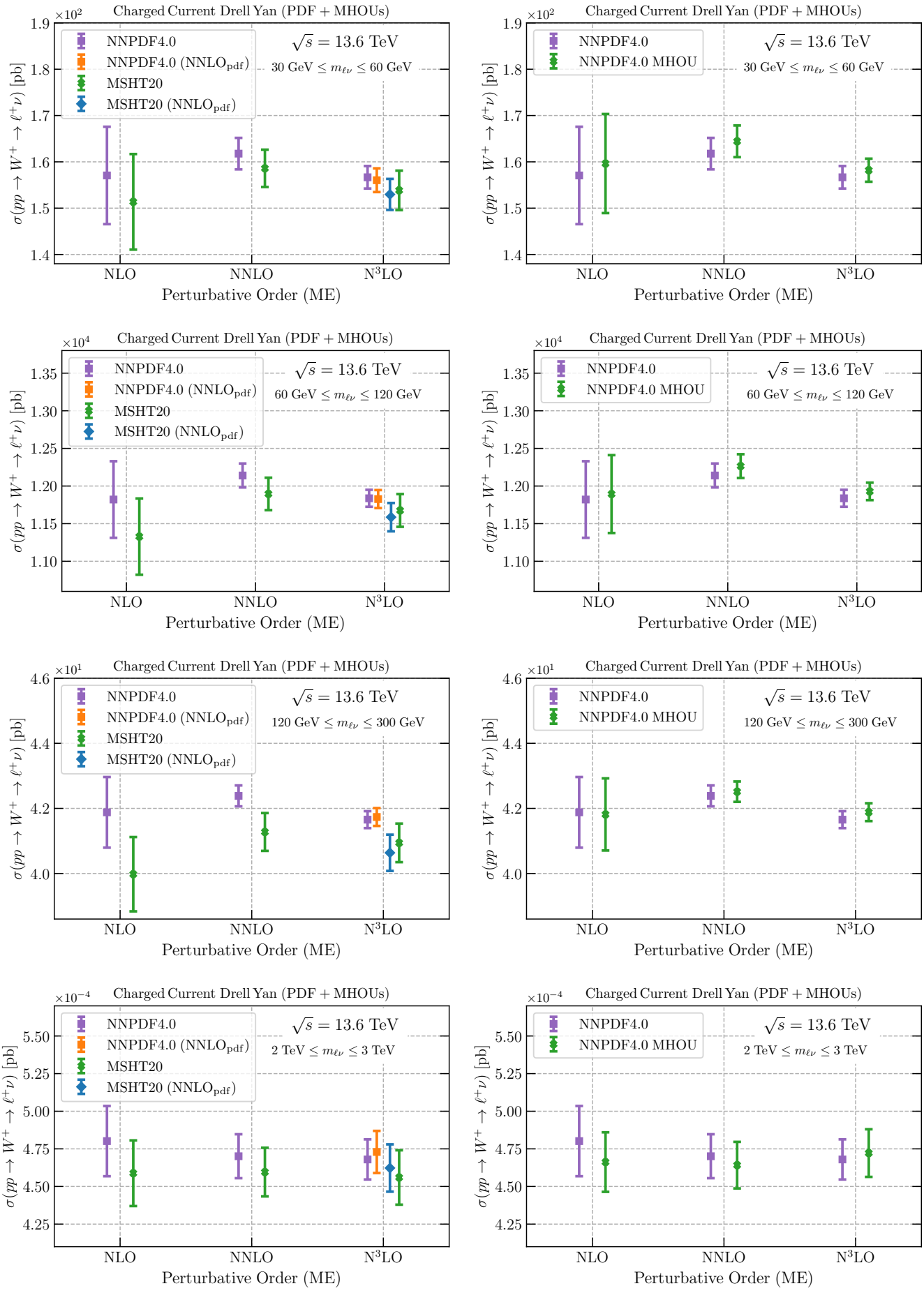



**Figure 5.2.** Same as Fig. 5.1 for the inclusive charged-current Drell-Yan production cross-section, $pp \to W^+ \to \ell^+\nu_\ell$.

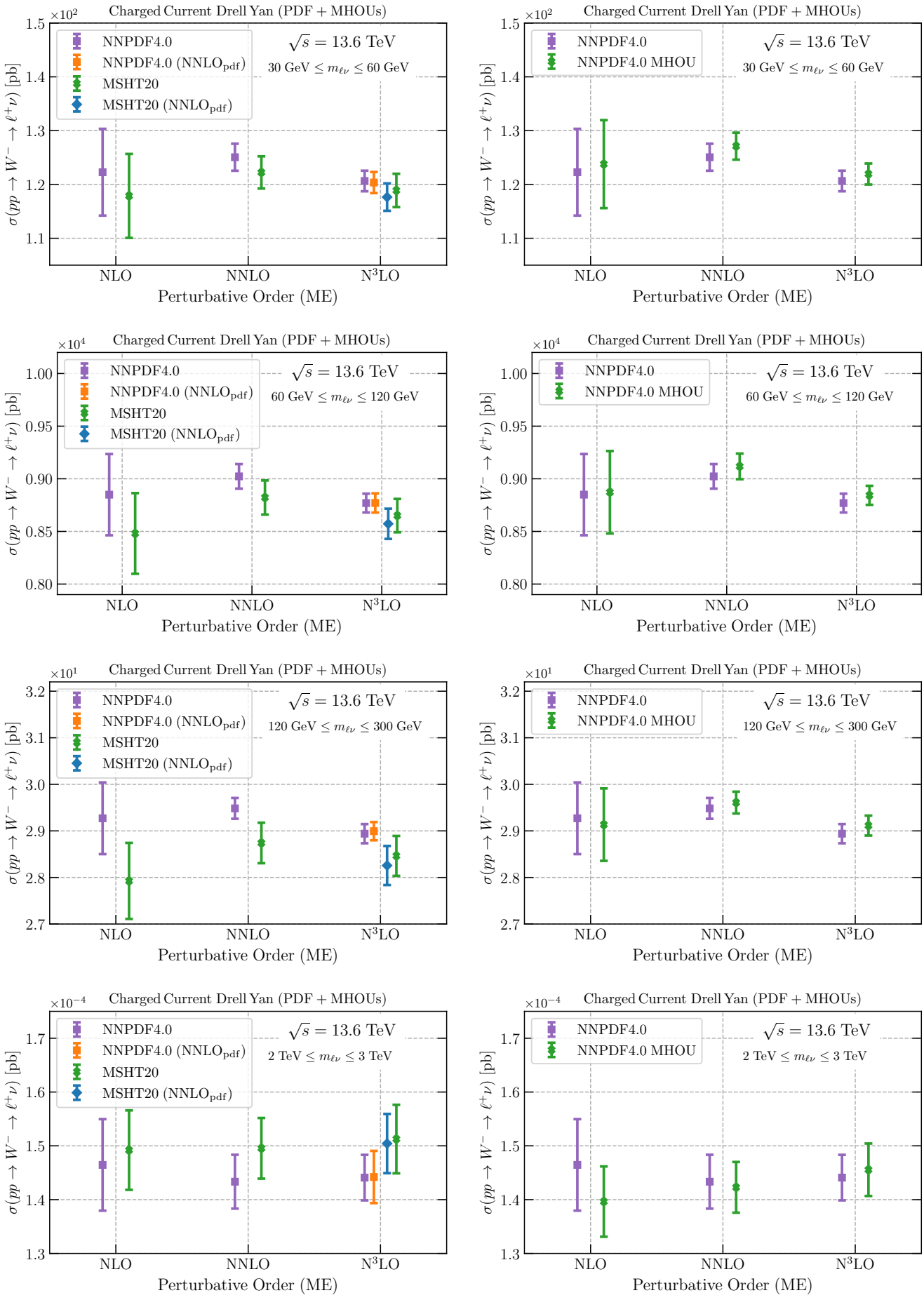



**Figure 5.3.** Same as Fig. 5.1 for the inclusive charged-current Drell-Yan production cross-section, $pp \to W^- \to \ell^-\bar{\nu}_\ell$.

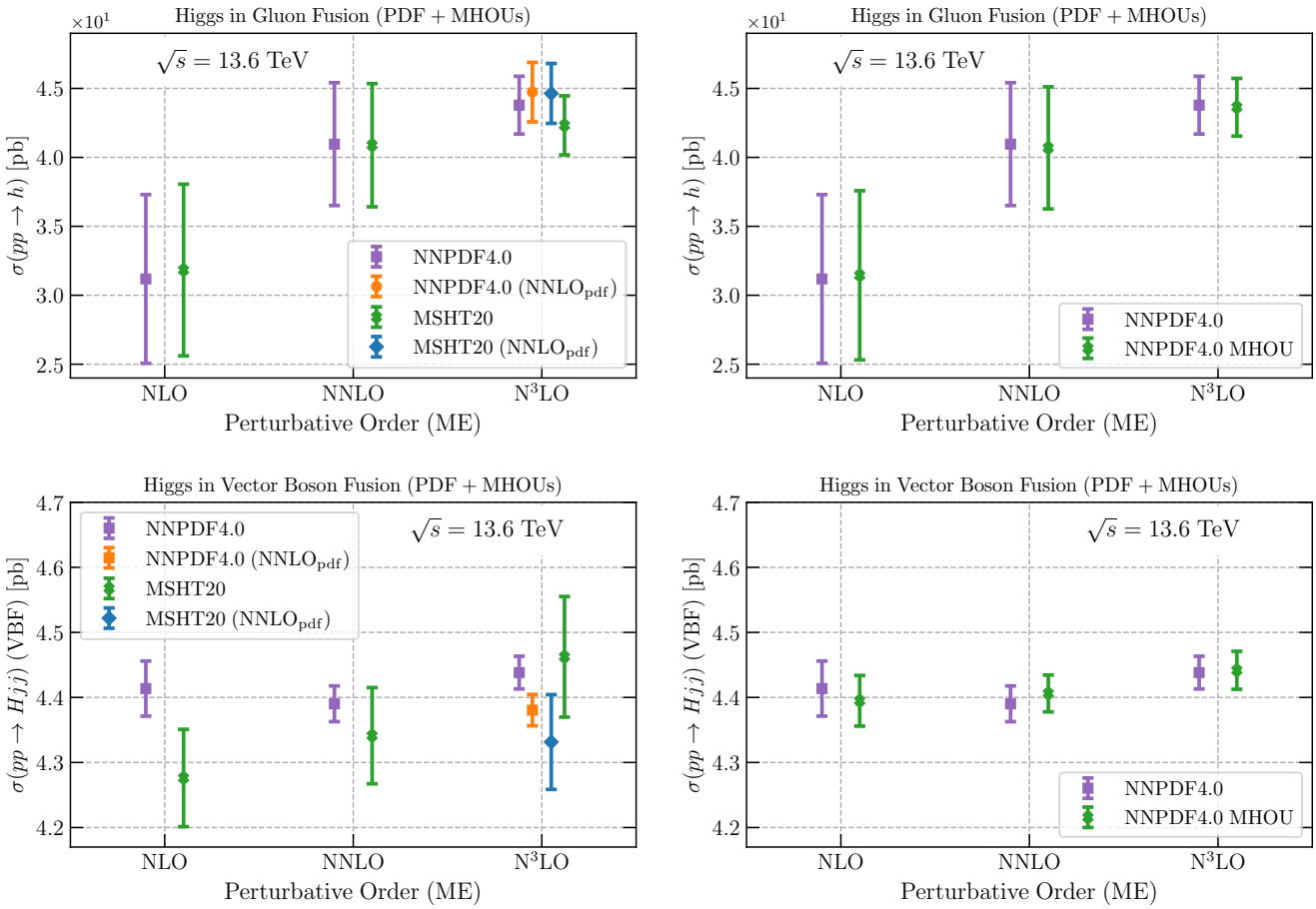



**Figure 5.4.** Same as Fig. 5.1 for Higgs production in gluon-fusion and via vector-boson fusion.

they do not at NLO for VBF, nor at NLO and NNLO for associated production. The impact of using aN$^3$LO PDFs instead of NNLO PDFs at N$^3$LO for NNPDF4.0 is very moderate for gluon fusion, somewhat more significant for associated production, and more significant for VBF, in which it is comparable to the PDF uncertainty. For MSHT20 instead a significant from using aN$^3$LO instead of NNLO PDFs is also observed for gluon fusion, where suppression of the cross-sections is seen when replacing NNLO with aN$^3$LO PDFs. This follows from the behaviour of the gluon luminosity seen in Fig. 4.19. The impact of MHOUs on the PDFs is generally quite small on the scale of the PDF uncertainty at all perturbative orders, and essentially absent for gluon fusion. For associated production it marginally improves perturbative convergence.

# 6 Summary and outlook

We have presented the first aN$^3$LO PDF sets within the NNPDF framework, by constructing a full set of approximate N$^3$LO splitting functions based on available partial results and known limits, approximate massive DIS coefficient functions, and extending to this order the FONLL general-mass scheme for DIS coefficient functions. We now summarize the new PDF sets that we are releasing, our main conclusions on their features, and our plans for future developments.

The NNPDF4.0 aN$^3$LO PDF sets are available via the LHAPDF6 interface,

http://lhapdf.hepforge.org/ .

Specifically, we provide an aN$^3$LO NNPDF4.0 set

`NNPDF40_an3lo_as_01180`

that supplements the LO, NLO and NNLO sets of Ref. [37].
We also provide NLO and aN$^3$LO NNPDF4.0MHOU sets

`NNPDF40_nlo_as_01180_mhou`
`NNPDF40_an3lo_as_01180_mhou`

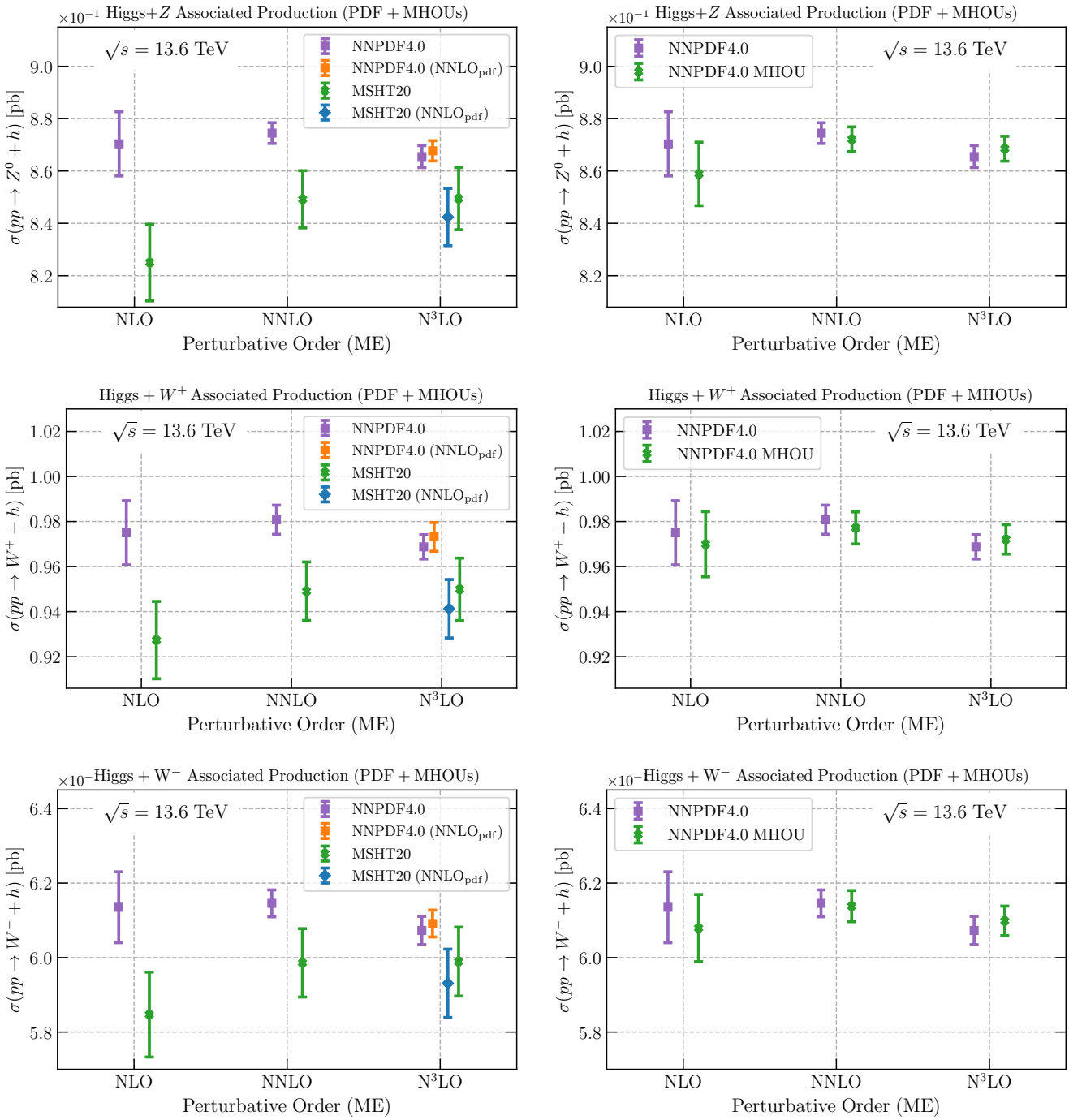



**Figure 5.5.** Same as Fig. 5.1 for Higgs production in association with $W^+$ and $Z$ gauge bosons: from top to bottom, $Zh$, $W^+h$, and $W^-h$.

that supplement the NNLO NNPDF4.0MHOU PDF set of Ref. [38]. These sets include in the PDF uncertainty the MHOU on the processes used for PDF determination, but in all other respects (including the dataset) follow the default sets to which they can be directly compared.

For both aN$^3$LO sets, we also release the corresponding sets including $\alpha_s$ variations,

```
NNPDF40_an3lo_as_01180_pdfas
NNPDF40_an3lo_as_01180_mhou_pdfas
```

in which replicas 101 and 102 correspond to fits with $\alpha_s(m_Z) = 0.117$ and 0.119 respectively, in order to evaluate the combined PDF+$\alpha_s$ uncertainties following the prescription of [115, 122, 123].

All these sets are delivered as ensembles of $N_{\rm rep} = 100$ Monte Carlo replicas. However, we also make available Hessian variants following [124, 125] and denoted

```
NNPDF40_an3lo_as_01180_hessian
```

`NNPDF40_an3lo_as_01180_mhou_hessian`
`NNPDF40_an3lo_as_01180_pdfas_hessian`
`NNPDF40_an3lo_as_01180_mhou_pdfas_hessian`

each set comprising $N_{\rm eig} = 100$ eigenvectors.

All these sets are also available on the NNPDF Collaboration website,

https://nnpdf.mi.infn.it/nnpdf4-0-n3lo/ .

where we also give the PDF sets discussed in Sect. 4.5 based on variant treatments of the aN$^3$LO corrections. In addition to the LHAPDF grids themselves, all the results obtained in this work are reproducible by means of the open-source NNPDF code [39] and the related suite of theory tools.

We have provided a first assessment of these PDF sets by comparing them to their NLO and NNLO counterparts with and without MHOUs. Our main conclusions are the following

- For all PDFs good perturbative convergence is observed, with differences decreasing as the perturbative order increases, and the aN$^3$LO result always compatible with the NNLO within uncertainties.
- For quark PDFs the difference between NNLO and aN$^3$LO results is extremely small, suggesting that with current data and methodology the effect of yet higher orders is negligible.
- For the gluon PDF a more significant shift is observed between NNLO and N$^3$LO, thus making the inclusion of N$^3$LO important for precision phenomenology.
- The inclusion of MHOUs improves perturbative convergence, mostly by shifting central values at each order towards the higher-order result, by an amount that decreases with increasing perturbative order.
- Upon inclusion of MHOUs the fit quality becomes all but independent of perturbative order, and PDF uncertainties generally decrease (or remain unchanged) due to the improved data compatibility.
- The effect of MHOUs at N$^3$LO is very small for quarks but not negligible for the gluon PDF.
- Evidence for intrinsic charm is somewhat increased already at NNLO by the inclusion of MHOUs, and somewhat increased again when going from NNLO to N$^3$LO.
- The impact of N$^3$LO corrections on the total cross-section for Higgs in gluon fusion is very small on the scale of the PDF uncertainty.

All in all, these results underline the importance of the inclusion of N$^3$LO corrections and MHOUs for precision phenomenology at sub-percent accuracy.

Future NNPDF releases will include by default MHOUs, will be at all orders up to aN$^3$LO, and will include a photon PDF. Specifically, we aim to extend to aN$^3$LO with MHOUs our recent construction of NNPDF4.0QED PDFs [109]. Indeed, aN$^3$LO PDFs including a photon PDF (such as those recently released by MSHT20 [126]) will be a necessary ingredient for theory predictions based on state-of-the art QCD and electroweak (EW) corrections. In fact, we are working towards the consistent inclusion of combined QCD$\times$EW corrections also in the theory predictions used for PDF determination.

Another important line of future development involves the all-order resummation of potentially large perturbative contributions in the large $x$ and small $x$ regions [69,127]. This will involve matching resummed and fixed-order cross-sections and (at small $x$) perturbative evolution in the new streamlined NNPDF theory pipeline. Such resummed PDFs will be instrumental for precision phenomenology: specifically at small $x$, forward neutrino production at the LHC and scattering processes for high-energy astroparticle physics, and at large $x$, searches for new physics in high-mass final states at the LHC and future hadron colliders.

**Acknowledgments**

We thank James McGowan, Thomas Cridge, Lucian Harland-Lang, and Robert Thorne for discussions on MSHT20 PDFs. We are grateful to Sven Moch and Joshua Davies for discussion and for communications concerning their N$^3$LO results..

R.D.B, L.D.D, and R.S. are supported by the U.K. Science and Technology Facility Council (STFC) consolidated grants ST/T000600/1 and ST/X000494/1. F.H. is supported by the Academy of Finland project 358090 and is funded as a part of the Center of Excellence in Quark Matter of the Academy of Finland, project 346326. E.R.N. is supported by the Italian Ministry of University and Research (MUR) through the "Rita Levi-Montalcini" Program. M.U. and Z.K. are supported by the European Research Council under the European Union's Horizon 2020 research and innovation Programme (grant agreement n.950246), and partially by the STFC consolidated grant ST/T000694/1 and ST/X000664/1. J.R. is partially supported by NWO, the Dutch Research Council. C.S. is supported by the German Research Foundation (DFG) under reference number DE 623/6-2.

# A Explicit expressions of anomalous dimensions

We provide here explicit expressions for the $\gamma^{(3)}_{\mathrm{ns}\pm,\,N\to\infty}(N)$ and $\gamma^{(3)}_{\mathrm{ns}\pm,\,N\to0}(N)$ anomalous dimension discussed in Sect. 2.3 and the $\gamma^{(3)}_{\mathrm{ns}\pm,\,N\to\infty}(N)$, $\gamma^{(3)}_{\mathrm{ns}\pm,\,N\to0}(N)$ and $\gamma^{(3)}_{\mathrm{ns}\pm,\,N\to1}(N)$ discussed in Sect. 2.4.

$\gamma^{(3)}_{\mathrm{ns},\pm}(N)$**:**

$$\begin{aligned}\gamma^{(3)}_{\mathrm{ns},\pm,N\to0}(N) = &-\frac{252.84}{N^7}+\frac{1580.25-126.42n_f}{N^6}\\ &+\frac{-5806.8+752.198n_f-18.963n_f^2}{N^5}\\ &+\frac{14899.9-2253.11n_f+99.1605n_f^2-0.790123n_f^3}{N^4}\\ &+\frac{-28546.4+5247.18n_f-226.441n_f^2+2.89712n_f^3}{N^3}\\ &+\frac{50759.7-8769.15n_f+395.605n_f^2-3.16049n_f^3}{N^2}\end{aligned} \tag{A.1}$$

$$\begin{aligned}\gamma^{(3)}_{\mathrm{ns},\pm N\to\infty}(N) = &\left(+20702.4-5171.92n_f+195.577n_f^2+3.27234n_f^3\right)S_1(N)\\ &-23393.8+5550.04n_f-193.855n_f^2-3.01498n_f^3\\ &+\left(16950.9-2741.83n_f+26.6886n_f^2\right)\frac{S_1(N)}{N}\\ &+\frac{+11126.6-3248.4n_f+180.432n_f^2+0.526749n_f^3}{N}\end{aligned} \tag{A.2}$$

$\gamma^{(3)}_{gg}(N)$**:**

$$\begin{aligned}\gamma^{(3)}_{gg,N\to0}(N) = &\frac{-103680+20005.9n_f-568.889n_f^2}{N^7}\\ &+\frac{-17280-19449.7n_f+1725.63n_f^2}{N^6}\\ &+\frac{-627979+80274.1n_f-2196.54n_f^2+4.74074n_f^3}{N^5}\end{aligned} \tag{A.3}$$

$$\gamma^{(3)}_{gg,N\to1}(N) = -\frac{49851.7}{(N-1)^4}+\frac{213824+1992.77n_f}{(N-1)^3} \tag{A.4}$$

$$\begin{aligned}\gamma^{(3)}_{gg,N\to\infty}(N) = &\left(+40880.3 - 11714.2 n_f + 440.049 n_f^2 + 7.36278 n_f^3\right) S_1(N)\\ &- 68587.9 + 18144 n_f - 423.811 n_f^2 - 0.906722 n_f^3\\ &- \left(-85814.1 + 13880.5 n_f - 135.111 n_f^2\right) \frac{S_1(N)}{N}\\ &+ \frac{-54482.8 + 4341.13 n_f + 21.3333 n_f^2}{N}\end{aligned} \tag{A.5}$$

$\gamma^{(3)}_{gq}(N)$**:**

$$\begin{aligned}\gamma^{(3)}_{gq,N\to 0}(N) = &\frac{-37609.9 + 5309.63 n_f}{N^7}\\ &+ \frac{-35065.7 + 221.235 n_f}{N^6}\\ &+ \frac{-175455 + 9092.91 n_f + 778.535 n_f^2}{N^5}\end{aligned} \tag{A.6}$$

$$\gamma^{(3)}_{gq,N\to 1}(N) = -\frac{22156.3}{(N-1)^4} + \frac{95032.9 + 885.674 n_f}{(N-1)^3} \tag{A.7}$$

$$\begin{aligned}\gamma^{(3)}_{gq,N\to\infty}(N) = &\left(-13.4431 + 0.548697 n_f\right) L_{5,0}(N)\\ &+ \left(-375.398 + 34.4947 n_f - 0.877915 n_f^2\right) L_{4,0}(N)\end{aligned} \tag{A.8}$$

$\gamma^{(3)}_{qg}(N)$**:**

$$\begin{aligned}\gamma^{(3)}_{qg,N\to 0}(N) = &\frac{14103.7 n_f - 1991.11 n_f^2}{N^7}\\ &+ \frac{2588.84 n_f + 2069.33 n_f^2}{N^6}\\ &+ \frac{68802.3 n_f - 7229.38 n_f^2 - 99.1605 n_f^3}{N^5}\end{aligned} \tag{A.9}$$

$$\gamma^{(3)}_{qg,N\to 1}(N) = -\frac{7871.52 n_f}{(N-1)^3} \tag{A.10}$$

$$\begin{aligned}\gamma^{(3)}_{qg,N\to\infty}(N) = &\left(-1.85185 n_f + 0.411523 n_f^2\right) L_{5,0}(N)\\ &+ \left(-35.6878 n_f + 3.51166 n_f^2 + 0.0823045 n_f^3\right) L_{4,0}(N)\\ &+ \left(-2.88066 n_f - 0.823045 n_f^2\right) L_{5,1}(N)\\ &+ \left(+40.5114 n_f - 5.54184 n_f^2 - 0.164609 n_f^3\right) L_{4,1}(N)\end{aligned} \tag{A.11}$$

$\gamma^{(3)}_{qq,\mathrm{ps}}(N)$**:**

$$\begin{aligned}\gamma^{(3)}_{qq,\mathrm{ps},N\to 0}(N) = &\frac{5404.44 n_f - 568.889 n_f^2}{N^7}\\ &+ \frac{3425.98 n_f + 455.111 n_f^2}{N^6}\\ &+ \frac{20515.2 n_f - 1856.79 n_f^2 + 4.74074 n_f^3}{N^5}\end{aligned} \tag{A.12}$$

$$\gamma^{(3)}_{qq,\mathrm{ps},N\to 1}(N) = -\frac{3498.45 n_f}{(N-1)^3}, \tag{A.13}$$

$$\begin{aligned}\gamma^{(3)}_{qq,\mathrm{ps},N\to\infty}(N) = &\left(+56.4609n_f - 3.6214n_f^2\right) L_{4,1}(N)\\ &+ \left(+247.551n_f - 40.5597n_f^2 + 1.58025n_f^3\right) L_{3,1}(N)\\ &+ 13.1687n_f L_{4,2}(N)\\ &+ \left(+199.111n_f - 13.6955n_f^2\right) L_{3,2}(N).\end{aligned} \tag{A.14}$$

The functions $L_{k,j}(N)$ are defined as the Mellin transform of $(1-x)^j \ln^k(1-x)$:

$$L_{k,0}(N) = \mathcal{M}\left[\ln^k(1-x)\right](N) = (-1)^k k! \frac{S_{1_k,\ldots,1_1}(N)}{N} \tag{A.15}$$

$$L_{k,1}(N) = \mathcal{M}\left[(1-x)\ln^k(1-x)\right] = L_{k,0}(N) - L_{k,0}(N+1) \tag{A.16}$$

$$L_{k,2}(N) = \mathcal{M}\left[(1-x)^2\ln^k(1-x)\right] = L_{k,0}(N) - 2L_{k,0}(N+1) + L_{k,0}(N+2) \tag{A.17}$$

with the multi-index harmonics of weight-$k$ defined recursively as:

$$S_{1_k,\ldots,1_1}(N) = \sum_{j=1}^{N} \frac{S_{1_{k-1},\ldots,1_1}(j)}{j}\ . \tag{A.18}$$

## References


[1] J. A. M. Vermaseren, A. Vogt, and S. Moch, *The third-order QCD corrections to deep-inelastic scattering by photon exchange*, *Nucl. Phys.* **B724** (2005) 3, [hep-ph/0504242].

[2] S. Moch, J. A. M. Vermaseren, and A. Vogt, *The longitudinal structure function at the third order*, *Phys. Lett.* **B606** (2005) 123, [hep-ph/0411112].

[3] S. Moch, M. Rogal, and A. Vogt, *Differences between charged-current coefficient functions*, *Nucl. Phys.* **B790** (2008) 317–335, [arXiv:0708.3731].

[4] S. Moch, J. A. M. Vermaseren, and A. Vogt, *Third-order QCD corrections to the charged-current structure function F(3)*, *Nucl. Phys. B* **813** (2009) 220–258, [arXiv:0812.4168].

[5] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, *Phys. Rev. Lett.* **114** (2015), no. 21 212001, [arXiv:1503.06056].

[6] B. Mistlberger, *Higgs boson production at hadron colliders at $N^3LO$ in QCD*, *JHEP* **05** (2018) 028, [arXiv:1802.00833].

[7] C. Duhr, F. Dulat, and B. Mistlberger, *Higgs Boson Production in Bottom-Quark Fusion to Third Order in the Strong Coupling*, *Phys. Rev. Lett.* **125** (2020), no. 5 051804, [arXiv:1904.09990].

[8] J. Baglio, C. Duhr, B. Mistlberger, and R. Szafron, *Inclusive production cross sections at $N^3LO$*, *JHEP* **12** (2022) 066, [arXiv:2209.06138].

[9] F. A. Dreyer and A. Karlberg, *Vector-Boson Fusion Higgs Production at Three Loops in QCD*, *Phys. Rev. Lett.* **117** (2016), no. 7 072001, [arXiv:1606.00840].

[10] L.-B. Chen, H. T. Li, H.-S. Shao, and J. Wang, *Higgs boson pair production via gluon fusion at $N^3LO$ in QCD*, *Phys. Lett. B* **803** (2020) 135292, [arXiv:1909.06808].

[11] C. Duhr, F. Dulat, and B. Mistlberger, *Charged current Drell-Yan production at $N^3LO$*, *JHEP* **11** (2020) 143, [arXiv:2007.13313].

[12] C. Duhr and B. Mistlberger, *Lepton-pair production at hadron colliders at $N^3LO$ in QCD*, *JHEP* **03** (2022) 116, [arXiv:2111.10379].

[13] F. Dulat, B. Mistlberger, and A. Pelloni, *Differential Higgs production at $N^3LO$ beyond threshold*, *JHEP* **01** (2018) 145, [arXiv:1710.03016].

[14] F. Dulat, B. Mistlberger, and A. Pelloni, *Precision predictions at $N^3LO$ for the Higgs boson rapidity distribution at the LHC*, *Phys. Rev. D* **99** (2019), no. 3 034004, [arXiv:1810.09462].

[15] X. Chen, T. Gehrmann, E. W. N. Glover, A. Huss, B. Mistlberger, and A. Pelloni, *Fully Differential Higgs Boson Production to Third Order in QCD*, *Phys. Rev. Lett.* **127** (2021), no. 7 072002, [arXiv:2102.07607].

[16] G. Billis, B. Dehnadi, M. A. Ebert, J. K. L. Michel, and F. J. Tackmann, *Higgs $pT$ Spectrum and Total Cross Section with Fiducial Cuts at Third Resummed and Fixed Order in QCD*, *Phys. Rev. Lett.* **127** (2021), no. 7 072001, [arXiv:2102.08039].

[17] S. Camarda, L. Cieri, and G. Ferrera, *Drell–Yan lepton-pair production: $qT$ resummation at N3LL accuracy and fiducial cross sections at N3LO*, *Phys. Rev. D* **104** (2021), no. 11 L111503, [arXiv:2103.04974].

[18] X. Chen, T. Gehrmann, N. Glover, A. Huss, T.-Z. Yang, and H. X. Zhu, *Dilepton Rapidity Distribution in Drell-Yan Production to Third Order in QCD*, *Phys. Rev. Lett.* **128** (2022), no. 5 052001, [arXiv:2107.09085].

[19] X. Chen, T. Gehrmann, N. Glover, A. Huss, T.-Z. Yang, and H. X. Zhu, *Transverse mass distribution and charge asymmetry in $W$ boson production to third order in QCD*, *Phys. Lett. B* **840** (2023) 137876, [arXiv:2205.11426].

[20] F. Caola, W. Chen, C. Duhr, X. Liu, B. Mistlberger, F. Petriello, G. Vita, and S. Weinzierl, *The Path forward to $N^3LO$*, in *Snowmass 2021*, 3, 2022. arXiv:2203.06730.

[21] J. Davies, A. Vogt, B. Ruijl, T. Ueda, and J. A. M. Vermaseren, *Large-nf contributions to the four-loop splitting functions in QCD*, *Nucl. Phys. B* **915** (2017) 335–362, [arXiv:1610.07477].

[22] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *Four-Loop Non-Singlet Splitting Functions in the Planar Limit and Beyond*, *JHEP* **10** (2017) 041, [arXiv:1707.08315].

[23] J. Davies, C. H. Kom, S. Moch, and A. Vogt, *Resummation of small-x double logarithms in QCD: inclusive deep-inelastic scattering*, *JHEP* **08** (2022) 135, [arXiv:2202.10362].

[24] J. M. Henn, G. P. Korchemsky, and B. Mistlberger, *The full four-loop cusp anomalous dimension in $\mathcal{N} = 4$ super Yang-Mills and QCD*, *JHEP* **04** (2020) 018, [arXiv:1911.10174].

[25] C. Duhr, B. Mistlberger, and G. Vita, *Soft integrals and soft anomalous dimensions at $N^3LO$ and beyond*, *JHEP* **09** (2022) 155, [arXiv:2205.04493].

[26] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *Low moments of the four-loop splitting functions in QCD*, *Phys. Lett. B* **825** (2022) 136853, [arXiv:2111.15561].

[27] G. Soar, S. Moch, J. A. M. Vermaseren, and A. Vogt, *On Higgs-exchange DIS, physical evolution kernels and fourth-order splitting functions at large x*, *Nucl. Phys. B* **832** (2010) 152–227, [arXiv:0912.0369].

[28] G. Falcioni, F. Herzog, S. Moch, and A. Vogt, *Four-loop splitting functions in QCD – The quark-quark case*, *Phys. Lett. B* **842** (2023) 137944, [arXiv:2302.07593].

[29] G. Falcioni, F. Herzog, S. Moch, and A. Vogt, *Four-loop splitting functions in QCD – The gluon-to-quark case*, *Phys. Lett. B* **846** (2023) 138215, [arXiv:2307.04158].

[30] S. Moch, B. Ruijl, T. Ueda, J. Vermaseren, and A. Vogt, *Additional moments and x-space approximations of four-loop splitting functions in QCD*, *Phys. Lett. B* **849** (2024) 138468, [arXiv:2310.05744].

[31] G. Falcioni, F. Herzog, S. Moch, J. Vermaseren, and A. Vogt, *The double fermionic contribution to the four-loop quark-to-gluon splitting function*, *Phys. Lett. B* **848** (2024) 138351, [arXiv:2310.01245].

[32] J. McGowan, T. Cridge, L. A. Harland-Lang, and R. S. Thorne, *Approximate $N^3LO$ parton distribution functions with theoretical uncertainties: MSHT20a$N^3$LO PDFs*, *Eur. Phys. J. C* **83** (2023), no. 3 185, [arXiv:2207.04739]. [Erratum: Eur.Phys.J.C 83, 302 (2023)].

[33] W. L. van Neerven and A. Vogt, *Improved approximations for the three loop splitting functions in QCD*, *Phys. Lett. B* **490** (2000) 111–118, [hep-ph/0007362].

[34] **NNPDF** Collaboration, R. D. Ball, E. R. Nocera, and R. L. Pearson, *Nuclear Uncertainties in the Determination of Proton PDFs*, *Eur. Phys. J.* **C79** (2019), no. 3 282, [arXiv:1812.09074].

[35] **NNPDF** Collaboration, R. Abdul Khalek et al., *A first determination of parton distributions with theoretical uncertainties*, *Eur. Phys. J.* **C** (2019) 79:838, [arXiv:1905.04311].

[36] **NNPDF** Collaboration, R. Abdul Khalek et al., *Parton Distributions with Theory Uncertainties: General Formalism and First Phenomenological Studies*, *Eur. Phys. J. C* **79** (2019), no. 11 931, [arXiv:1906.10698].

[37] **NNPDF** Collaboration, R. D. Ball et al., *The path to proton structure at 1% accuracy*, *Eur. Phys. J. C* **82** (2022), no. 5 428, [arXiv:2109.02653].

[38] **NNPDF** Collaboration, R. D. Ball et al., *Determination of the theory uncertainties from missing higher orders on NNLO parton distributions with percent accuracy*, arXiv:2401.10319.

[39] **NNPDF** Collaboration, R. D. Ball et al., *An open-source machine learning framework for global analyses of parton distributions*, *Eur. Phys. J. C* **81** (2021), no. 10 958, [arXiv:2109.02671].

[40] A. Candido, F. Hekhorn, and G. Magni, *EKO: evolution kernel operators*, *Eur. Phys. J. C* **82** (2022), no. 10 976, [arXiv:2202.02338].

[41] A. Candido, F. Hekhorn, G. Magni, T. R. Rabemananjara, and R. Stegeman, *Yadism: Yet Another Deep-Inelastic Scattering Module*, arXiv:2401.15187.

[42] S. Forte, E. Laenen, P. Nason, and J. Rojo, *Heavy quarks in deep-inelastic scattering*, *Nucl. Phys.* **B834** (2010) 116–162, [arXiv:1001.2312].

[43] R. D. Ball, M. Bonvini, and L. Rottoli, *Charm in Deep-Inelastic Scattering*, *JHEP* **11** (2015) 122, [arXiv:1510.02491].

[44] R. D. Ball, V. Bertone, M. Bonvini, S. Forte, P. Groth Merrild, J. Rojo, and L. Rottoli, *Intrinsic charm in a matched general-mass scheme*, *Phys. Lett.* **B754** (2016) 49–58, [arXiv:1510.00009].

[45] A. Vogt, *Efficient evolution of unpolarized and polarized parton distributions with QCD-PEGASUS*, *Comput. Phys. Commun.* **170** (2005) 65–92, [hep-ph/0408244].

[46] T. Gehrmann, A. von Manteuffel, V. Sotnikov, and T.-Z. Yang, *The $N_f C_F^3$ contribution to the non-singlet splitting function at four-loop order*, *Phys. Lett. B* **849** (2024) 138427, [arXiv:2310.12240].

[47] S. Albino and R. D. Ball, *Soft resummation of quark anomalous dimensions and coefficient functions in MS-bar factorization*, *Phys. Lett. B* **513** (2001) 93–102, [hep-ph/0011133].

[48] R. D. Ball, M. Bonvini, S. Forte, S. Marzani, and G. Ridolfi, *Higgs production in gluon fusion beyond NNLO*, *Nucl.Phys.* **B874** (2013) 746, [arXiv:1303.3590].

[49] T. Gehrmann, A. von Manteuffel, V. Sotnikov, and T.-Z. Yang, *Complete $N_f^2$ contributions to four-loop pure-singlet splitting functions*, *JHEP* **01** (2024) 029, [arXiv:2308.07958].

[50] V. S. Fadin, E. A. Kuraev, and L. N. Lipatov, *On the Pomeranchuk Singularity in Asymptotically Free Theories*, *Phys. Lett. B* **60** (1975) 50–52.

[51] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, *Multi - Reggeon Processes in the Yang-Mills Theory*, *Sov. Phys. JETP* **44** (1976) 443–450.

[52] L. N. Lipatov, *Reggeization of the Vector Meson and the Vacuum Singularity in Nonabelian Gauge Theories*, *Sov. J. Nucl. Phys.* **23** (1976) 338–345.

[53] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, *The Pomeranchuk Singularity in Nonabelian Gauge Theories*, *Sov. Phys. JETP* **45** (1977) 199–204.

[54] I. I. Balitsky and L. N. Lipatov, *The Pomeranchuk Singularity in Quantum Chromodynamics*, *Sov. J. Nucl. Phys.* **28** (1978) 822–829.

[55] V. S. Fadin and L. N. Lipatov, *Next-to-leading corrections to the BFKL equation from the gluon and quark production*, *Nucl. Phys. B* **477** (1996) 767–808, [hep-ph/9602287].

[56] V. S. Fadin, R. Fiore, A. Flachi, and M. I. Kotsky, *Quark - anti-quark contribution to the BFKL kernel*, *Phys. Lett. B* **422** (1998) 287–293, [hep-ph/9711427].

[57] V. S. Fadin, M. I. Kotsky, and L. N. Lipatov, *One-loop correction to the BFKL kernel from two gluon production*, *Phys. Lett. B* **415** (1997) 97–103.

[58] G. Camici and M. Ciafaloni, *Irreducible part of the next-to-leading BFKL kernel*, *Phys. Lett. B* **412** (1997) 396–406, [hep-ph/9707390]. [Erratum: Phys.Lett.B 417, 390 (1998)].

[59] V. S. Fadin and L. N. Lipatov, *BFKL pomeron in the next-to-leading approximation*, *Phys. Lett. B* **429** (1998) 127–134, [hep-ph/9802290].

[60] T. Jaroszewicz, *Gluonic Regge Singularities and Anomalous Dimensions in QCD*, *Phys. Lett. B* **116** (1982) 291–294.

[61] R. D. Ball and S. Forte, *Summation of leading logarithms at small x*, *Phys. Lett. B* **351** (1995) 313–324, [hep-ph/9501231].

[62] R. D. Ball and S. Forte, *The Small x behavior of Altarelli-Parisi splitting functions*, *Phys. Lett. B* **465** (1999) 271–281, [hep-ph/9906222].

[63] M. Bonvini and S. Marzani, *Four-loop splitting functions at small x*, *JHEP* **06** (2018) 145, [arXiv:1805.06460].

[64] S. Catani and F. Hautmann, *High-energy factorization and small x deep inelastic scattering beyond leading order*, *Nucl. Phys. B* **427** (1994) 475–524, [hep-ph/9405388].

[65] A. Vogt, S. Moch, and J. A. M. Vermaseren, *The Three-loop splitting functions in QCD: The Singlet case*, *Nucl. Phys.* **B691** (2004) 129–181, [hep-ph/0404111].

[66] Y. L. Dokshitzer, G. Marchesini, and G. P. Salam, *Revisiting parton evolution and the large-x limit*, *Phys. Lett. B* **634** (2006) 504–507, [hep-ph/0511302].

[67] A. A. Almasy, G. Soar, and A. Vogt, *Generalized double-logarithmic large-x resummation in inclusive deep-inelastic scattering*, *JHEP* **03** (2011) 030, [arXiv:1012.3352].

[68] G. Altarelli, R. D. Ball, and S. Forte, *Small x Resummation with Quarks: Deep-Inelastic Scattering*, *Nucl. Phys.* **B799** (2008) 199–240, [arXiv:0802.0032].

[69] R. D. Ball, V. Bertone, M. Bonvini, S. Marzani, J. Rojo, and L. Rottoli, *Parton distributions with small-x resummation: evidence for BFKL dynamics in HERA data*, *Eur. Phys. J.* **C78** (2018), no. 4 321, [arXiv:1710.05935].

[70] A. Barontini, A. Candido, J. M. Cruz-Martinez, F. Hekhorn, and C. Schwan, *Pineline: Industrialization of high-energy theory predictions*, *Comput. Phys. Commun.* **297** (2024) 109061, [arXiv:2302.12124].

[71] I. Bierenbaum, J. Blümlein, and S. Klein, *Mellin Moments of the $O(\alpha_s^3)$ Heavy Flavor Contributions to unpolarized Deep-Inelastic Scattering at $Q^2 \gg m^2$ and Anomalous Dimensions*, *Nucl. Phys. B* **820** (2009) 417–482, [arXiv:0904.3563].

[72] I. Bierenbaum, J. Blumlein, S. Klein, and C. Schneider, *Two-Loop Massive Operator Matrix Elements for Unpolarized Heavy Flavor Production to O(epsilon)*, *Nucl. Phys. B* **803** (2008) 1–41, [arXiv:0803.0273].

[73] I. Bierenbaum, J. Blümlein, and S. Klein, *The Gluonic Operator Matrix Elements at $O(\alpha_s^2)$ for DIS Heavy Flavor Production*, *Phys. Lett. B* **672** (2009) 401–406, [arXiv:0901.0669].

[74] J. Ablinger, J. Blümlein, S. Klein, C. Schneider, and F. Wissbrock, *The $O(\alpha_s^3)$ Massive Operator Matrix Elements of $O(n_f)$ for the Structure Function $F_2(x, Q^2)$ and Transversity*, *Nucl. Phys. B* **844** (2011) 26–54, [arXiv:1008.3347].

[75] J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. Hasselhuhn, A. von Manteuffel, M. Round, C. Schneider, and F. Wißbrock, *The 3-Loop Non-Singlet Heavy Flavor Contributions and Anomalous Dimensions for the Structure Function $F_2(x, Q^2)$ and Transversity*, *Nucl. Phys. B* **886** (2014) 733–823, [arXiv:1406.4654].

[76] J. Ablinger, J. Blümlein, A. De Freitas, A. Hasselhuhn, A. von Manteuffel, M. Round, and C. Schneider, *The $O(\alpha_s^3 T_F^2)$ Contributions to the Gluonic Operator Matrix Element*, *Nucl. Phys. B* **885** (2014) 280–317, [arXiv:1405.4259].

[77] A. Behring, I. Bierenbaum, J. Blümlein, A. De Freitas, S. Klein, and F. Wißbrock, *The logarithmic contributions to the $O(\alpha_s^3)$ asymptotic massive Wilson coefficients and operator matrix elements in deeply inelastic scattering*, *Eur. Phys. J. C* **74** (2014), no. 9 3033, [arXiv:1403.6356].

[78] J. Blümlein, J. Ablinger, A. Behring, A. De Freitas, A. von Manteuffel, C. Schneider, and C. Schneider, *Heavy Flavor Wilson Coefficients in Deep-Inelastic Scattering: Recent Results*, *PoS* **QCDEV2017** (2017) 031, [arXiv:1711.07957].

[79] J. Ablinger, J. Blümlein, A. De Freitas, A. Hasselhuhn, A. von Manteuffel, M. Round, C. Schneider, and F. Wißbrock, *The transition matrix element $a_{gq}(n)$ of the variable flavor number scheme at $o(\alpha_s^3)$*, *Nuclear Physics B* **882** (May, 2014) 263–288.

[80] J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. von Manteuffel, and C. Schneider, *The 3-loop pure singlet heavy flavor contributions to the structure function f2(x,q2) and the anomalous dimension*, *Nuclear Physics B* **890** (Jan, 2015) 48–151.

[81] J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. Goedicke, A. von Manteuffel, C. Schneider, and K. Schönwald, *The unpolarized and polarized single-mass three-loop heavy flavor operator matrix elements $A_{gg,Q}$ and $\Delta A_{gg,Q}$*, *JHEP* **12** (2022) 134, [arXiv:2211.05462].

[82] J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. von Manteuffel, C. Schneider, and K. Schönwald, *The first–order factorizable contributions to the three–loop massive operator matrix elements $AQg(3)$ and $\Delta AQg(3)$*, *Nucl. Phys. B* **999** (2024) 116427, [arXiv:2311.00644].

[83] H. Kawamura, N. A. Lo Presti, S. Moch, and A. Vogt, *On the next-to-next-to-leading order QCD corrections to heavy-quark production in deep-inelastic scattering*, *Nucl. Phys. B* **864** (2012) 399–468, [arXiv:1205.5727].

[84] T. Cridge et al., *Approximate $N^3LO$ splitting functions benchmarking*, *in preparation* (2024).

[85] J. Davies, A. Vogt, S. Moch, and J. A. M. Vermaseren, *Non-singlet coefficient functions for charged-current deep-inelastic scattering to the third order in QCD*, *PoS* **DIS2016** (2016) 059, [arXiv:1606.08907].

[86] J. Blümlein, P. Marquard, C. Schneider, and K. Schönwald, *The massless three-loop Wilson coefficients for the deep-inelastic structure functions $F_2$, $F_L$, $xF_3$ and $g_1$*, *JHEP* **11** (2022) 156, [arXiv:2208.14325].

[87] N. Laurenti, *Construction of a next-to-next-to-next-to-leading order approximation for heavy flavour production in deep inelastic scattering with quark masses*, Master's thesis, Rome Sapienza U., 2021. arXiv:2401.12139.

[88] A. Barontini, M. Bonvini, and N. Laurenti, *Implementation of DIS at N3LO for PDF determinations, in preparation* (2024).

[89] J. C. Collins, F. Wilczek, and A. Zee, *Low-Energy Manifestations of Heavy Particles: Application to the Neutral Current*, *Phys. Rev. D* **18** (1978) 242.

[90] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, *Charm electroproduction viewed in the variable flavor number scheme versus fixed order perturbation theory*, *Eur. Phys. J. C* **1** (1998) 301–320, [hep-ph/9612398].

[91] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, *Complete O (alpha-s) corrections to heavy flavor structure functions in electroproduction*, *Nucl. Phys. B* **392** (1993) 162–228.

[92] S. I. Alekhin and J. Blumlein, *Mellin representation for the heavy flavor contributions to deep inelastic structure functions*, *Phys. Lett. B* **594** (2004) 299–307, [hep-ph/0404034].

[93] F. Hekhorn and M. Stratmann, *Next-to-Leading Order QCD Corrections to Inclusive Heavy-Flavor Production in Polarized Deep-Inelastic Scattering*, *Phys. Rev. D* **98** (2018), no. 1 014018, [arXiv:1805.09026].

[94] F. Hekhorn, *Next-to-Leading Order QCD Corrections to Heavy-Flavour Production in Neutral Current DIS*. PhD thesis, Tubingen U., Math. Inst., 2019. arXiv:1910.01536.

[95] J. Gao, *Massive charged-current coefficient functions in deep-inelastic scattering at NNLO and impact on strange-quark distributions*, *JHEP* **02** (2018) 026, [arXiv:1710.04258].

[96] E. Laenen and S.-O. Moch, *Soft gluon resummation for heavy quark electroproduction*, *Phys. Rev. D* **59** (1999) 034027, [hep-ph/9809550].

[97] S. Catani, M. Ciafaloni, and F. Hautmann, *High-energy factorization and small x heavy flavor production*, *Nucl. Phys. B* **366** (1991) 135–188.

[98] J. Ablinger, A. Behring, J. Blümlein, A. De Freitas, A. von Manteuffel, and C. Schneider, *The 3-loop pure singlet heavy flavor contributions to the structure function $F_2(x, Q^2)$ and the anomalous dimension*, *Nucl. Phys. B* **890** (2014) 48–151, [arXiv:1409.1135].

[99] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, *NLL resummation of the heavy quark hadroproduction cross-section*, *Nucl. Phys. B* **529** (1998) 424–450, [hep-ph/9801375]. [Erratum: Nucl.Phys.B 803, 234 (2008)].

[100] A. Pineda and A. Signer, *Heavy Quark Pair Production near Threshold with Potential Non-Relativistic QCD*, *Nucl. Phys. B* **762** (2007) 67–94, [hep-ph/0607239].

[101] M. Cacciari, M. Greco, and P. Nason, *The p(T) spectrum in heavy-flavour hadroproduction*, *JHEP* **05** (1998) 007, [hep-ph/9803400].

[102] A. Barontini, A. Candido, F. Hekhorn, G. Magni, and R. Stegeman, *An FONLL prescription with coexisting flavor number PDFs, in preparation* (2024).

[103] **ATLAS** Collaboration, G. Aad et al., *Measurement of the high-mass Drell–Yan differential cross-section in pp collisions at $\sqrt{s}$=7 TeV with the ATLAS detector*, *Phys.Lett.* **B725** (2013) 223, [arXiv:1305.4192].

[104] **ATLAS** Collaboration, G. Aad et al., *Measurement of the inclusive $W^{\pm}$ and $Z/\gamma^*$ cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s}$= 7 TeV with the ATLAS detector*, *Phys.Rev.* **D85** (2012) 072004, [arXiv:1109.5141].

[105] **ATLAS** Collaboration, M. Aaboud et al., *Precision measurement and interpretation of inclusive $W^+$ , $W^-$ and $Z/\gamma^*$ production cross sections with the ATLAS detector*, *Eur. Phys. J.* **C77** (2017), no. 6 367, [arXiv:1612.03016].

[106] **ATLAS** Collaboration, G. Aad et al., *Measurement of $W^\pm$ and $Z$-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett.* **B759** (2016) 601–621, [arXiv:1603.09222].

[107] F. Faura, S. Iranipour, E. R. Nocera, J. Rojo, and M. Ubiali, *The Strangest Proton?*, *Eur. Phys. J. C* **80** (2020), no. 12 1168, [arXiv:2009.00014].

[108] S. Amoroso et al., *Snowmass 2021 Whitepaper: Proton Structure at the Precision Frontier*, *Acta Phys. Polon. B* **53** (2022), no. 12 12–A1.63–67, [arXiv:2203.13923].

[109] **NNPDF** Collaboration, R. D. Ball et al., *Photons in the proton: implications for the LHC*, arXiv:2401.08749.

[110] M. L. Mangano et al., *Physics at a 100 TeV pp collider: Standard Model processes*, arXiv:1607.01831.

[111] R. D. Ball, E. R. Nocera, and R. L. Pearson, *Deuteron Uncertainties in the Determination of Proton PDFs*, *Eur. Phys. J. C* **81** (2021), no. 1 37, [arXiv:2011.00009].

[112] **NNPDF** Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [arXiv:1410.8849].

[113] **NNPDF** Collaboration, R. D. Ball, A. Candido, J. Cruz-Martinez, S. Forte, T. Giani, F. Hekhorn, K. Kudashkin, G. Magni, and J. Rojo, *Evidence for intrinsic charm quarks in the proton*, *Nature* **608** (2022), no. 7923 483–487, [arXiv:2208.08372].

[114] R. D. Ball, A. Candido, J. Cruz-Martinez, S. Forte, T. Giani, F. Hekhorn, G. Magni, E. R. Nocera, J. Rojo, and R. Stegeman, *The intrinsic charm quark valence distribution of the proton*, arXiv:2311.00743.

[115] **PDF4LHC Working Group** Collaboration, R. D. Ball et al., *The PDF4LHC21 combination of global PDF fits for the LHC Run III*, *J. Phys. G* **49** (2022), no. 8 080501, [arXiv:2203.05506].

[116] L. A. Harland-Lang and R. S. Thorne, *On the Consistent Use of Scale Variations in PDF Fits and Predictions*, *Eur. Phys. J.* **C79** (2019), no. 3 225, [arXiv:1811.08434].

[117] R. D. Ball and R. L. Pearson, *Correlation of theoretical uncertainties in PDF fits and theoretical uncertainties in predictions*, *Eur. Phys. J. C* **81** (2021), no. 9 830, [arXiv:2105.05114].

[118] Z. Kassabov, M. Ubiali, and C. Voisey, *Parton distributions with scale uncertainties: a Monte Carlo sampling approach*, *JHEP* **03** (2023) 148, [arXiv:2207.07616].

[119] R. D. Ball, A. Candido, S. Forte, F. Hekhorn, E. R. Nocera, J. Rojo, and C. Schwan, *Parton distributions and new physics searches: the Drell–Yan forward–backward asymmetry as a case study*, *Eur. Phys. J. C* **82** (2022), no. 12 1160, [arXiv:2209.08115].

[120] M. Bonvini, R. D. Ball, S. Forte, S. Marzani, and G. Ridolfi, *Updated Higgs cross section at approximate $N^3LO$*, *J. Phys.* **G41** (2014) 095002, [arXiv:1404.3204].

[121] F. A. Dreyer and A. Karlberg, *Vector-Boson Fusion Higgs Pair Production at $N^3LO$*, *Phys. Rev. D* **98** (2018), no. 11 114016, [arXiv:1811.07906].

[122] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys.* **G43** (2016) 023001, [arXiv:1510.03865].

[123] F. Demartin, S. Forte, E. Mariani, J. Rojo, and A. Vicini, *The impact of PDF and $\alpha_s$ uncertainties on Higgs Production in gluon fusion at hadron colliders*, *Phys. Rev.* **D82** (2010) 014002, [arXiv:1004.0962].

[124] S. Carrazza, S. Forte, Z. Kassabov, J. I. Latorre, and J. Rojo, *An Unbiased Hessian Representation for Monte Carlo PDFs*, *Eur. Phys. J.* **C75** (2015), no. 8 369, [arXiv:1505.06736].

[125] S. Carrazza, S. Forte, Z. Kassabov, and J. Rojo, *Specialized minimal PDFs for optimized LHC calculations*, *Eur. Phys. J.* **C76** (2016), no. 4 205, [arXiv:1602.00005].

[126] T. Cridge, L. A. Harland-Lang, and R. S. Thorne, *Combining QED and Approximate $N^3LO$ QCD Corrections in a Global PDF Fit: MSHT20qed_an3lo PDFs*, arXiv:2312.07665.

[127] M. Bonvini, S. Marzani, J. Rojo, L. Rottoli, M. Ubiali, R. D. Ball, V. Bertone, S. Carrazza, and N. P. Hartland, *Parton distributions with threshold resummation*, *JHEP* **09** (2015) 191, [arXiv:1507.01006].

Prepared for submission to JHEP ALBERTA-THY-2-24

# Light quark mediated Higgs boson production in association with a jet at the next-to-next-leading order and beyond

**Tao Liu,**[a,b] **Alexander A. Penin,**[c] **Abdur Rehman**[c]

[a] *Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*
[b] *University of Chinese Academy of Sciences, Beijing 100049, China*
[c] *Department of Physics, University of Alberta, Edmonton AB T6G 2J1, Canada*

*E-mail:* liutao86@ihep.ac.cn, penin@ualberta.ca, rehman3@ualberta.ca

Abstract: We study the light quark effect on the Higgs boson production in association with a jet at the LHC in the intermediate transverse momentum region between the quark and the Higgs boson mass scales. Though the effect is suppressed by the small Yukawa coupling, it is enhanced by large logarithms of the quark mass ratio to the Higgs boson mass or transverse momentum. Following a remarkable success of the logarithmic expansion [39] for the prediction of the next-to-next-to-leading bottom quark contribution to the total cross section of the Higgs boson production we extend the analysis to its kinematical distributions. A new factorization formula is derived for the light quark mediated $gg \to Hg$ amplitudes and the differential cross section of the process is computed in the logarithmic approximation, which is used for an estimate of the bottom quark effect at the next-to-next-to-leading order.

arXiv:2402.18625v1 [hep-ph] 28 Feb 2024

## Contents


**1 Introduction** 1

**2 Setup and notations** 2

**3 Factorization and resummation of the leading logarithms** 4
- 3.1 Factorization of the one-loop amplitudes 4
- 3.2 All-order analysis 8

**4 Logarithmic corrections to the kinematical distributions** 11

**5 Conclusion** 14


## 1 Introduction

Determination of the Higgs boson properties with high precision is a major part of the Large Hadron Collider physics program [1, 2]. A significant deviation of the measured couplings from the theory predictions may indicate a signal of the physics beyond the standard model. The transverse momentum distribution of the Higgs boson produced in association with a jet is an important kinematical observable sensitive to its Yukawa couplings and can be used to test different extensions of the standard model in different regions of the transverse momentum [3–6]. The dominant contribution to the process cross section is given by the top quark loop mediated gluon fusion. It is known through the next-to-next-to-leading order (NNLO) in the strong coupling constant in the heavy top quark effective theory [7–11] and recently has been evaluated to next-to-leading order (NLO) with the full top quark mass dependence [12]. The contribution of the light quarks to the process is strongly suppressed by their masses. This suppression is partially compensated by the enhancement of the leading order amplitude by the second power of the large logarithm $\ln(m_H^2/m_q^2)$ of the Higgs boson to the light quark mass ratio [13], and the bottom quark effect can be as large as 5% [14]. Moreover, the contribution can be distinguished by its transverse momentum dependence for $p_\perp \ll m_H$, where the cross section is large. This makes the differential distribution sensitive to the light quark Yukawa couplings [15, 16]. Thus, it is important to get an accurate prediction for the differential cross section in perturbative QCD. The NLO result in the small quark mass limit has been derived in [17, 18]. The higher order corrections may also be important due to the presence of the double-logarithmic terms with the second power of the large logarithm per each power of the strong coupling constant. The resummation of the logarithmically enhanced terms poses a serious theoretical challenge for the effective field theory methods [19–22] due to their non-standard origin [23–25]. In contrast to the standard Sudakov case [26–31], the logarithmic correction to the mass-suppressed amplitudes are determined by the eikonal color charge nonconservation in the

process of the soft quark exchange [32]. The comprehensive analysis of this phenomena in the leading (double) logarithmic approximation has been performed for the three-point amplitudes in QED and QCD [33, 34]. For the four-point amplitudes the abelian part of the all-order result is available for the $gg \to Hg$ production [35] and only the two-loop result for Bhabha scattering [36, 37].[1] The resummation of the subleading logarithms has been also completed for the $gg \to H$ amplitudes [39–41]. The logarithmic expansion turned out to be remarkably successful for the estimate of the bottom quark contribution to the total cross section of the Higgs boson production in the NNLO. The next-to-leading-logarithmic (NLL) approximation for this quantity gives $2.18 \pm 0.20\ pb$ [39], while the recently computed full result reads $1.99(1)\,^{+0.30}_{-0.15}\ pb$ [42].[2] Hence, the generalization of the analysis to the case of the final state jet seems quite important both for general theory of renormalization group at subleading power and from purely phenomenological point of view.

In this paper we extend the abelian result [35] to the full QCD description of the light quark contribution to the differential cross section of the Higgs boson production in association with a jet for the intermediate values of transverse momentum. This requires a deeper understanding of the origin of the non-Sudakov logarithms and a systematic derivation of the factorization formula. The paper is organized as follows. In the next section we introduce our notations. In Sect. 3 we start with the factorization analysis of the double-logarithmic contribution to the leading order one-loop amplitude, then derive the general all-order resummation formula and compare it to the available two-loop expressions [45]. In Sect. 4 we apply this result to the analysis of the $pp \to Hj + X$ differential cross section, which determines the NNLO bottom quark contribution in terms of the known $K$-factors obtained in the heavy quark limit. Sect. 5 is our conclusion.

## 2 Setup and notations

We consider a light quark mediated Higgs boson production in the process

$$g(p_1) + g(p_2) \to g(p_3) + H(p_H) \tag{2.1}$$

for the intermediate values of the transverse momentum

$$m_q^2 \ll p_\perp^2 \ll s, m_H^2\,, \tag{2.2}$$

where $p_\perp^2 = ut/s$, $s = (p_1 + p_2)^2, t = (p_1 - p_3)^2, u = (p_2 - p_3)^2$. We focus on the near-threshold production for partonic center-of-mass energy $s \approx m_H^2$ so that the additional soft gluon emission energy is small compared to the leading jet. This, in particular, means that the transverse momenta of the Higgs boson and the jet are approximately equal in magnitude. The above kinematical region gives the dominant contribution to the hadronic cross section. For $p_\perp^2 \ll s$ the process is described by four helicity amplitudes $M_{\lambda_1\lambda_2\lambda_3}$

[1] The all-order leading logarithmic result is known for the subleading power contribution in the Regge limit [38].

[2] The discrepancy in the next-to-next-to-leading order terms quoted in [42] is mainly due to the different bottom quark mass renormalization scheme adopted in the paper.

with the same helicity of colliding gluons $\lambda_1 = \lambda_2$, which are pairwise connected by spatial inversion. We take $M_{++\pm}$ as the two independent amplitudes and only keep the leading terms which grow as $1/p_\perp$ with decreasing transverse momentum. In the spinor notations the corresponding expressions read

$$M_{+++} = -\sqrt{2} f^{a_1 a_2 a_3} \frac{g_s}{v} \frac{\alpha_s}{4\pi} \frac{\langle 12\rangle^2}{[12]\langle 23\rangle\langle 13\rangle} Z_{3g} \sum_q A^{(q)}_{+++} \,, \tag{2.3}$$

$$M_{++-} = -\sqrt{2} f^{a_1 a_2 a_3} \frac{g_s}{v} \frac{\alpha_s}{4\pi} \frac{\langle 12\rangle}{[23][13]} Z_{3g} \sum_q A^{(q)}_{++-} \,, \tag{2.4}$$

where $\alpha_s = g_s^2/(4\pi)$ is the strong coupling constant, $v$ is the Higgs field vacuum expectation value, $f^{a_1 a_2 a_3}$ is the $SU(3)$ structure constant, the sum goes over the quark flavors, $Z_{3g}$ is the Sudakov factor incorporating the soft and collinear divergences of the amplitudes, and $A^{(q)}_{++-}$ are the infrared finite form factors. Let us first discuss the leading contribution of the top quark which can be expanded in the inverse powers of $m_t$. The perturbative series for the form factors then take the following form

$$A^{(t)}_{++\pm} = C_t \sum_{n=0} \left(\frac{\alpha_s}{2\pi}\right)^n \tilde{A}^{(n)}_{++\pm} + \mathcal{O}(m_t^{-2}) \,, \tag{2.5}$$

where the Wilson coefficient $C_t = 1 + \frac{11}{4}\frac{\alpha_s}{\pi} + \ldots$ is known through $\alpha_s^4$ [43, 44], $\tilde{A}^{(0)}_{++\pm} = \pm 2/3$, and $\tilde{A}^{(n)}_{++\pm}$ stands for the $n$-loop correction computed in the heavy top effective theory, which is available up to $n = 2$ [46]. For light quarks the amplitudes can be expanded in powers of $m_q$ and the perturbative series for the form factors reads

$$A^{(q)}_{++\pm} = \frac{m_q^2}{s} \sum_{n=0} \left(\frac{\alpha_s}{2\pi}\right)^n A^{(n+1)}_{++\pm} + \mathcal{O}(m_q^4) \,, \tag{2.6}$$

where $A^{(n)}_{++\pm}$ stands for the $n$-loop QCD contribution also available up to $n = 2$ [45]. These coefficients include the terms enhanced by the large logarithms of the scale ratios $p_\perp^2/m_q^2$, $m_H^2/m_q^2$ or $t/u$, which becomes relevant at large rapidity. The goal of this work is to compute the coefficients $A^{(n)}_{++\pm}$ in the leading (double) logarithmic approximation, *i.e.* keeping the highest power of the large logarithms equal to $2n$, for all $n$.

In the leading logarithmic approximation the Sudakov factor can be written as follows $Z_{3g}^{LL} = e^{\frac{\alpha_s}{2\pi} I^{(1)}}$ where $I^{(1)}$ is an analog of the one-loop Catani's operator [47] with the trivial structure in the color space. It is a function of the Mandelstam variables and is scheme dependent. To make the factorization of the Sudakov and non-Sudakov logarithms explicit it is convenient to define this function in such a way that it incorporates the nonsingular double logarithmic dependence of the one-loop corrections to the effective theory amplitudes with the local $ggH$ interaction on the kinematical invariants. By using the result [48] it is straightforward to get the leading singularity subtraction operator in the "physical" scheme

$$I^{(1)}_{\rm ph} = -\frac{C_A}{2\epsilon^2} \left[ 2 \left(\frac{-s}{\mu^2}\right)^{-\epsilon} + \left(\frac{-tu}{s\mu^2}\right)^{-\epsilon} \right] , \tag{2.7}$$

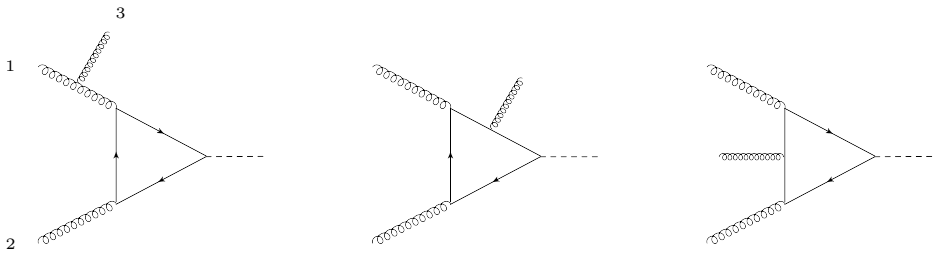



**Figure 1**. One-loop diagrams representing the leading order light quark contribution to $gg \to Hg$ process. Wide and narrow gluon lines correspond to the initial and final gluon states, respectively. Symmetric diagrams corresponding to the opposite direction of the quark flow and to the soft emissions off the opposite gluon/quark line are not shown. In the first diagram the gluons are explicitly enumerated according to their momentum index.

where $\epsilon = 2 - d/2$ is the parameter of dimensional regularization and $C_A = N_c$ for the $SU(N_c)$ color gauge group. Note that the canonical symmetric form of the subtraction operator has been used in [45], which differs from Eq. (2.7) by a non-singular double logarithmic terms, as discussed in Sect. 3.2.

## 3 Factorization and resummation of the leading logarithms

### 3.1 Factorization of the one-loop amplitudes

A detailed discussion of the evaluation of the leading order one-loop amplitude in the double-logarithmic approximation can be found in [35]. It, however, does not reveal a deeper factorization structure, which is crucial for the full QCD analysis and is discussed in this section. Following [35] we choose the gluon polarization vectors $\epsilon_i$ satisfying the gauge conditions

$$\epsilon_i p_i = 0 \quad (i = 1, 2, 3)\,, \qquad \epsilon_1 p_2 = 0\,, \qquad \epsilon_2 p_1 = 0\,, \qquad \epsilon_3 p_2 = 0\,, \tag{3.1}$$

so that the jet cannot be emitted off the gluon or quark line carrying the large light-cone momentum $p_2$. It is convenient to choose a frame where the light-cone components $p_{1\perp} = p_{2\perp} = p_1^+ = p_2^- = \epsilon_3^- = 0$. Then, up to the direction of the closed quark line the leading order process is described by the three diagrams in Fig. 1. The diagrams Fig. 1(a,b) have the standard structure of soft emission from a highly energetic eikonal line. In the diagram Fig. 1(c), however, the jet is emitted from the soft quark line which does not carry the large momenta $p_{1,2}$. We focus on the specific structure of the corresponding double logarithmic contribution. Due to the helicity conservation in the small mass limit and the helicity flipping scalar Yukawa interaction, the light quark contribution is mass suppressed. The mass-suppressed double-logarithmic corrections are known to be generated by the on-shell soft quark contribution. Thus the quark propagator in the expression for the Feynman

diagrams in Fig. 1 can be approximated as follows

$$S(l) \equiv \frac{1}{\not{l} - m_q + i0} \to -i m_q \pi \delta(l^2 - m_q^2)\,, \tag{3.2}$$

where $l$ is the soft loop momentum and the mass factor accounts for the required helicity flip on the virtual quark line.

Each of the two t-channel quark propagators in Fig. 1(c) can go on-shell which defines two soft virtual momentum regions. Let us consider the one with the upper on-shell propagator so that the jet momentum $p_3$ flows through the lower one. After omitting irrelevant terms, the off-shell quark propagators become

$$S(p_1 + l) \to \frac{\not{p}_1}{2p_1 l}\,, \tag{3.3}$$

$$S(p_3 + l) \to \frac{\not{p}_3 + \not{l}}{2p_3 l}\,, \tag{3.4}$$

$$S(p_3 - p_2 + l) \to \frac{\not{p}_2}{2p_2 p_3 + 2p_2 l}\,. \tag{3.5}$$

Then the integral develops the double logarithmic scaling either with the first term in the numerator of Eq. (3.4) for $|p_2 l| < |p_2 p_3|$ or with the second term for $|p_2 p_3| < |p_2 l|$. Thus we can write

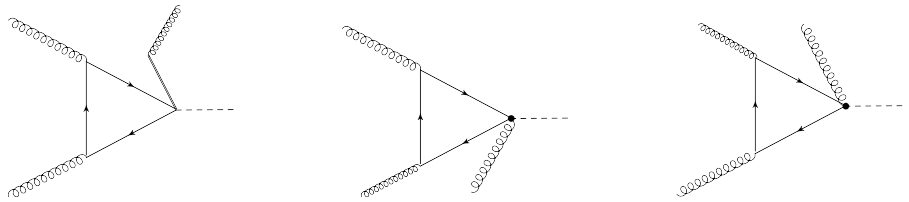
$$\begin{aligned} S(p_3 + l)\not{\epsilon}_2 S(p_3 - p_2 + l) \to\; & \theta\left(|p_2 p_3| - |p_2 l|\right) \frac{\not{p}_3}{2p_3 l}\not{\epsilon}_2 \frac{\not{p}_2}{2p_2 p_3} \\ & + \theta(|p_2 l| - |p_2 p_3|) \frac{\not{l}}{2p_3 l}\not{\epsilon}_2 \frac{\not{p}_2}{2p_2 l}\,. \end{aligned} \tag{3.6}$$

This decomposition, however, does not reflect the factorization structure in the effective theory terms and we rewrite it as follows

$$\begin{aligned} S(p_3 + l)\not{\epsilon}_2 S(p_3 - p_2 + l) \to\; & \theta\left(|p_2 p_3| - |p_2 l|\right) \left( \frac{\not{p}_3}{2p_3 l}\not{\epsilon}_2 \frac{\not{p}_2}{2p_2 p_3} - \frac{\not{l}}{2p_3 l}\not{\epsilon}_2 \frac{\not{p}_2}{2p_2 l} \right) \\ & + \frac{\not{l}}{2p_3 l}\not{\epsilon}_2 \frac{\not{p}_2}{2p_2 l}\,, \end{aligned} \tag{3.7}$$

where the integration region is extended in the second term and its variation is subtracted from the first term.

Let us first consider the contribution proportional to the theta-function in Eq. (3.7). It includes two distinct structures corresponding to two terms in the brackets. For the second term in the logarithmic integration region $\frac{\not{l}}{2p_3 l}$ can be replaced by $\frac{\not{p}_1}{2p_3 p_1}$, which results in the amplitude proportional to $\frac{\epsilon_3 p_1}{p_3 p_1}$ characteristic to the color-dipole emission from the eikonal line of the momentum $p_1$, and gives $A_{+++} = -A_{++-}$. Moreover, the only dependence on the momentum $p_3$ is through the argument of the theta-function and the integrand coincides with the one of the one-loop $ggH$ amplitude. Thus this *soft dipole* contribution reduces to the effective diagram Fig. 2(a) with the decoupled jet and a constraint $|p_2 l| < |p_2 p_3|$ on the loop momentum. The double line in this diagram stands for the eikonal (Wilson) line defined by the light-like momentum $p_1$, with the propagator $-p_1/(2p_3 p_1)$. The calculation of this



(a) (b) (c)

**Figure 2**. The effective representation of the sum of the one-loop diagrams in Fig. 1 after decomposition into the *dipole* and *symmetric* structures. The double line represents the eikonal line defined by the momentum $p_1$, as explained in the text. The black circle corresponds to the effective vertex.

contribution is straightforward by using the classical Sudakov technique which converts the integral over $l$ into the integral over the logarithmic Sudakov variables $\eta = \ln(|lp_3|/|p_1p_3|)/L$ and $\xi = \ln(|lp_1|/|p_2p_1|)/L$ (see [35] for the details)[3]

$$\left[A^{(1)}_{++\pm}\right]_{\text{s.d.}} = \mp L^2 \int_{\tau_t-\tau}^{\tau_t} \mathrm{d}\eta \int_{1-\tau_t}^{1-\eta} \mathrm{d}\xi = \mp L^2 \frac{\tau^2}{2}\,, \tag{3.8}$$

where $L = \ln(s/m_q^2)$, $\tau_t = \ln\left(|t|/m_q^2\right)/L$, and $\tau = \ln\left(p_\perp^2/m_q^2\right)/L$. Note that a dipole contribution from the virtual momentum configuration where the lower t-channel propagator goes on-shell and the jet momentum $p_3$ flows through the upper one is proportional to $\epsilon_3 p_2$ and vanishes for our gauge condition.

The first term in the brackets in Eq. (3.7) proportional to $p_3$ results in a new symmetric tensor structure described by a gauge invariant operator $G^a_{\mu\nu}G^b_{\nu\lambda}G^c_{\lambda\mu}f^{abc}$, which does not contribute to the all-plus helicity amplitude. It can be represented by the effective theory diagram Fig. 2(a) with the propagator $S(p_2 - p_3 + l)$ collapsed to an effective local vertex. Due to its symmetry and explicit gauge invariance this structure does get the equal contribution from the crossing-symmetric configuration with the opposite direction of the jet momentum flow represented by the effective diagram in Fig. 2(b). The sum gives the *symmetric* contribution

$$\left[A^{(1)}_{++-}\right]_{\text{sym.}} = -L^2\left(\int_{\tau_t-\tau}^{\tau_t} \mathrm{d}\eta \int_{1-\tau_t}^{1-\eta} \mathrm{d}\xi + (\tau_t \leftrightarrow \tau_u)\right) = -L^2\tau^2\,, \tag{3.9}$$

where $\tau_u = \ln\left(|u|/m_q^2\right)/L$. The total contribution of the theta-function term then reads

$$\left[A^{(1)}_{++\pm}\right]_{\text{s.d.}} + \left[A^{(1)}_{++\pm}\right]_{\text{sym.}} = -\frac{\tau^2}{2}L^2\,. \tag{3.10}$$

It is helicity independent and vanishes at the boundary of the transverse momentum region $p_\perp \to m_q$ where $\tau \to 0$.

To account for the unconstrained dipole contribution of the second line in Eq. (3.7) we use the method of [33] and move the jet vertex from the soft quark propagator to the

[3] In [35] the second logarithmic variable is defined as $\xi = \ln(|lp_1|/|p_3p_1|)/L$.

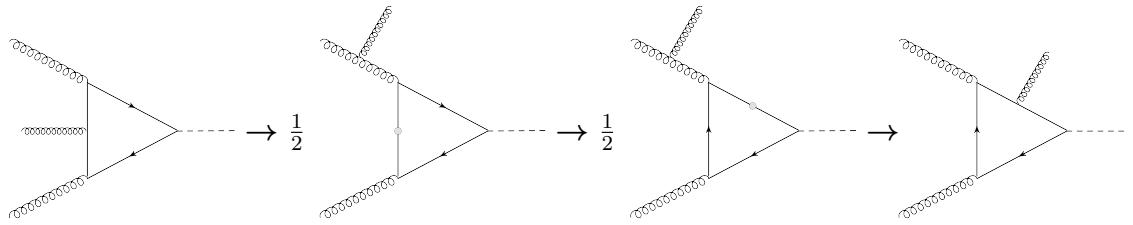



**Figure 3**. Diagrammatic representation of the sequence of the Ward identities and momentum shifts which relates the unconstrained dipole contribution of the diagram Fig. 1(c) to the diagram Fig. 1(b), as explained in the text. The empty circle marks the subtracted quark propagator.

eikonal line, Fig 3. This can be done in three steps. First we transform the soft line as follows

$$S(l)\not{\epsilon}_3 S(l+p_3) = S(l)\not{\epsilon}_3 S(l+p_3^+) + \ldots = \frac{\epsilon_3 p_1}{p_3 p_1}\left(S(l) - S(l+p_3^+)\right) + \ldots\,, \tag{3.11}$$

where we used the fact that $\epsilon_3^- = 0$, and the omitted terms do not contribute to the leading logarithms. This transformation, being a particular realization of the Slavnov-Ward identities, is graphically represented by the diagram Fig. 3(b), where the empty circle on the quark propagator represents the replacement $S(l) \to S(l) - S(l+p_3^+)$ and the factor $1/2$ comes from the color weight of the diagrams. By the momentum shift $l \to l - p_3^+$ in the second term of the above expression the crossed circle can be moved to the upper eikonal quark line which becomes $S(p_2+l) - S(p_2+l-p_3^+)$, Fig. 3(c). The opposite eikonal line is not sensitive to this shift since $p_2^- = 0$. On the final step we use the "inverted identity" on the upper eikonal quark line

$$\frac{\epsilon_3 p_1}{p_3 p_1}\left(S(p_2+l) - S(p_2+l-p_3^+)\right) = -S(p_2+l)\not{\epsilon}_3 S(p_2+l-p_3) + \ldots\,, \tag{3.12}$$

to transform the diagram Fig. 2(c) into Fig. 2(d) equal to Fig. 1(b). Now we can combine this dipole part of the diagram Fig. 1(c) with the diagrams Fig. 1(a,b) to get the standard eikonal factorization Fig. 1(a)+2 Fig. 1(b)= Fig. 2(a) where the factor 2 accounts for the difference in the color weights. In Fig. 2(a) the jet is emitted by the eikonal Wilson line of momentum $p_1$ and is completely decoupled from the quark loop without the constraint on the loop momentum present in the soft dipole contribution. Thus this *eikonal dipole* contribution reduces to the one-loop $g(p_1)g(p_2)H$ form factor

$$\left[A^{(1)}_{++\pm}\right]_{\text{e.d.}} = \pm 2L^2 \int_0^1 \mathrm{d}\eta \int_0^{1-\eta} \mathrm{d}\xi = \pm L^2\,, \tag{3.13}$$

and is independent of the jet kinematics. By adding up Eqs. (3.10) and (3.13) we get

$$A^{(1)}_{+++} = L^2\left(1 - \frac{\tau^2}{2}\right), \qquad A^{(1)}_{++-} = -L^2\left(1 + \frac{\tau^2}{2}\right), \tag{3.14}$$

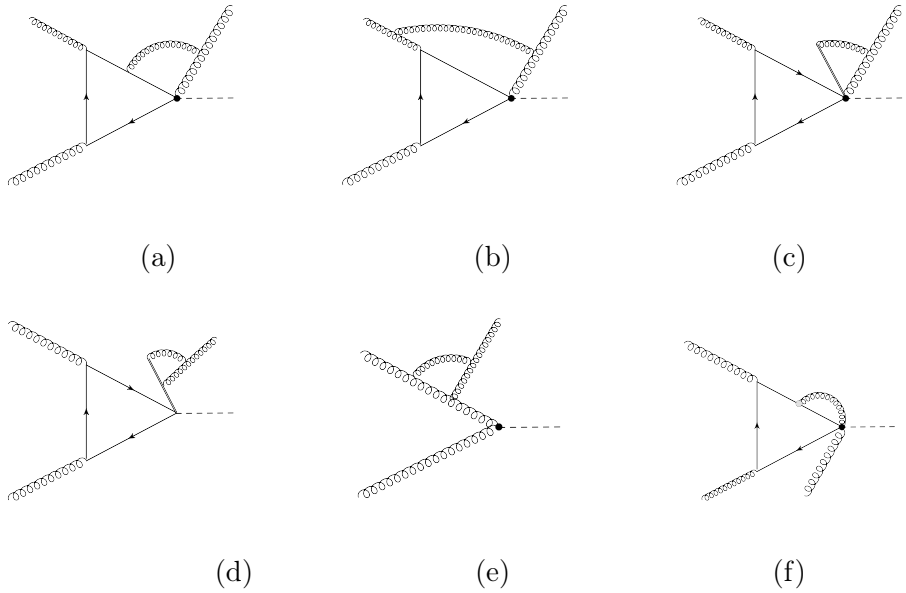



**Figure 4**. (a,b,c) Factorization of the soft interaction of the initial state gluon in the symmetric contribution. The eikonal double line carries the jet momentum. (d) The factorized soft interaction of the final state gluon in the dipole contribution. The eikonal double line carries the initial state momentum. (e) The heavy quark limit diagram with the same structure of the infrared Sudakov logarithms as the diagrams (c) and (d). (f) The Sudakov-Wilson contribution with the soft gluon momentum resolving the collapsed propagator. The empty circle represents the dipole interaction with the adjusted color weight.

in agreement with [13].

The main result of this section is the decomposition of the one-loop amplitudes into the eikonal dipole, soft dipole, and symmetric parts. In the first one the jet emission factors out from the quark loop, as it happens for $p_\perp \ll m_q$, *i.e.* in the heavy quark limit with the effective $g(p_1)g(p_2)H$ vertex. The last two describe the deviation from this factorization, accommodate all the dependence on the jet kinematical variables, and vanish when $p_\perp \to m_q$. The soft dipole contribution reduces to the quark-loop mediated $g(p_1)g(p_2)H$ form factor with a new *ultraviolet* constraint on the double logarithmic region of loop momentum depending on jet kinematics. The symmetric contribution is characterized by a new local effective vertex involving one of the initial state gluons and reduces to the quark-loop mediated $g(p_3)g(p_{1,2})H$ form factors with one of the gluons being the jet and with the same constraint on the double logarithmic region.

### 3.2 All-order analysis

The result of the previous section makes the all-order resummation of the double logarithmic radiative corrections rather straightforward. These corrections are generated by the soft virtual gluons with the well known factorization and exponentiation properties [26, 27] which should be adjusted to account for the non-Sudakov logarithms in the soft quark initiated processes [32, 33]. The latter are characteristic to the interaction of the eikonal lines

connected by the soft quark. Thus we do not expect the non-Sudakov contribution from the interaction of the decoupled gluons to the quark loop. An example of the eikonal factorization of this interaction for the symmetric contribution is shown in Fig. 4. If the soft gluon momentum does not resolve the local vertex we have Fig. 4(a)+Fig. 4(b)=Fig. 4(c). Note that the interaction of the virtual gluon with the soft quark has the vanishing color factor. For the dipole contribution similar factorization results in the diagram Fig. 4(d). Both diagrams Fig. 4(c,d) have the same structure of the Sudakov logarithms as the effective theory diagram Fig. 4(e) with the local $g(p_1)g(p_2)H$ vertex. Hence with our definition of the operator $I^{(1)}_{\rm ph}$ all the corresponding singular and finite double logarithmic terms are included in $Z_{3g}$ factor and we are left with the soft corrections to a quark loop mediated $ggH$ amplitude, where one of the gluon can be the jet. The structure of the double-logarithmic corrections to such an amplitude is well understood too. In general, after factoring out Sudakov corrections absorbed into $Z_{3g}$ we get the non-Sudakov double logarithmic contribution associated with $A_{++\pm}$. The structure of the factorization and the non-Sudakov logarithms, however, is different for the dipole and symmetric contributions. Let us first discuss the eikonal dipole part of the result. In this case the jet completely decouples from the quark loop and one can use the result [32, 33] for the $g(p_1)g(p_2)H$ amplitude, which gives

$$\left[A^{\rm LL}_{++\pm}\right]_{\rm e.d.} = \pm 2L^2 \int_0^1 {\rm d}\eta \int_0^{1-\eta} {\rm d}\xi\, e^{2z\eta\xi} = \pm L^2 \int_0^1 \frac{{\rm d}\eta}{z\eta}\left(e^{2z\eta(1-\eta)} - 1\right) = \pm L^2 g(z)\,, \tag{3.15}$$

where $g(z) = {}_2F_2\,(1,1;3/2,2;z/2)$ is the hypergeometric function, $z = (C_A - C_F)x$, $C_F = \frac{N_c^2-1}{2N_c}$, and $x = \frac{\alpha_s(\mu_s)}{4\pi}L^2$. The exponent $e^{2z\eta\xi}$ in Eq. (3.15) is the off-shell $\bar{q}qH$ Sudakov form factor with the color weight $C_F$ replaced by $C_F - C_A$ accounting for the variation of the color group representation for a particle moving along the eikonal lines from adjoint to fundamental in the process of the soft quark emission, which is the physical origin of the non-Sudakov logarithms.

The soft dipole contribution differs from the eikonal one by the cutoff of the double-logarithmic region of virtual momentum. This is an *ultraviolet* cutoff which does not change the infrared structure of the factorization and therefore can be imposed directly in Eq. (3.15). This gives

$$\left[A^{\rm LL}_{++\pm}\right]_{\rm s.d.} = \mp L^2 \int_{\tau_t-\tau}^{\tau_t} {\rm d}\eta \int_{1-\tau_t}^{1-\eta} {\rm d}\xi\, e^{2z\eta\xi} = \mp L^2 \int_{\tau_t-\tau}^{\tau_t} \frac{{\rm d}\eta}{2z\eta}\left(e^{2z\eta(1-\eta)} - e^{2z\eta(1-\tau_t)}\right)\,, \tag{3.16}$$

*cf.* Eqs. (3.8, 3.13).

For the symmetric contribution the eikonal lines carry the momenta $p_3$ and $p_1$ or $p_2$. Thus for *i.e.* $g(p_1)g(p_3)H$ amplitude the exponential factor of Eq. (3.15) should be changed to $e^{2z(\eta-\tau_t+\tau)(\xi-1+\tau_t)}$, while the integration limits are the same as in Eq. (3.16). However, if a light-cone component of the virtual soft gluon momentum exceeds the one of $p_3$, it may resolve the local effective vertex and such a correction is not taken into account by the above exponent. In this case the jet becomes soft with respect to the virtual gluon momentum and factors out of the corresponding loop integral. Thus, the factorization of

the Sudakov corrections is realized in the same way as for the $g(p_1)g(p_2)H$ amplitude. The resulting non-Sudakov contribution is accounted for by introducing a "Sudakov-Wilson" factor $e^{2z(\tau_t-\tau)\xi}$ local along one of the eikonal lines, Fig. 4(f), with the adjusted $C_F - C_A$ color weight. Compared to Eq. (3.15), in this expression the $\tau_t - \tau$ factor replaces the logarithmic variable $\eta$ as the infrared cutoff for the collapsed quark propagator in the effective vertex. Combining the corrections we get

$$\begin{aligned}\left[A^{\rm LL}_{++-}\right]_{\rm sym.} &= -L^2\left[\int_{\tau_t-\tau}^{\tau_t} d\eta \int_{1-\tau_t}^{1-\eta} d\xi\, e^{2z(\eta-\tau_t+\tau)(\xi-1+\tau_t)} e^{2z(\tau_t-\tau)\xi} + (\tau_t \to \tau_u)\right] \\ &= -L^2\left[\int_{\tau_t-\tau}^{\tau_t} \frac{d\eta}{2z\eta} e^{2z(\tau_t-\tau)(1-\tau_t)}\left(e^{2z\eta(\tau_t-\eta)}-1\right) + (\tau_t \to \tau_u)\right]. \end{aligned} \tag{3.17}$$

Thus the effective theory decomposition of the symmetric contribution has a more complex structure with the short-distance effects described by the Sudakov-Wilson factor.

We have verified the above factorization structure of the amplitudes by explicit identification and evaluation of the double-logarithmic contributions to the two-loop diagrams using the technique [33].

The total all-order leading logarithmic approximation for the helicity form factors can be written in a compact form

$$A^{LL}_{+++} = L^2\left[g(z) - \int_{(1-\tau+\zeta)/2}^{(1+\tau+\zeta)/2} \frac{d\eta}{2z\eta}\left(e^{2z\eta(1-\eta)} - e^{z\eta(1-\tau-\zeta))}\right)\right], \tag{3.18}$$

$$A^{LL}_{++-} = -A^{LL}_{+++} - L^2\left[\int_{(1-\tau+\zeta)/2}^{(1+\tau+\zeta)/2} \frac{d\eta}{2z\eta}\, e^{z((1-\tau)^2-\zeta^2)/2}\left(e^{z\eta(1+\tau+\zeta-2\eta)}-1\right) + (\zeta \to -\zeta)\right], \tag{3.19}$$

where $\zeta = \ln(t/u)/L$ is the rapidity variable. At the boundary values of the allowed transverse momentum interval Eqs. (3.18, 3.19) simplify and can be obtained in a closed analytic form. For $p_\perp \to m_b$ corresponding to $\tau \to 0$ the soft dipole and symmetric contributions vanish and we get

$$A^{LL}_{++\pm}|_{\tau=0} = \pm L^2 g(z)\,. \tag{3.20}$$

At the same time for central rapidity $\zeta = 0$ and $p_\perp \to m_H$ corresponding to $\tau \to 1$ we have $|s| \sim |t| \sim |u|$ and up to an overall factor the symmetric, soft and eikonal dipole contributions are determined by the same $ggH$ form factor. This gives

$$A^{LL}_{+++}|_{\tau=1,\,\zeta=0} = \frac{1}{2}L^2 g(z)\,, \qquad A^{LL}_{++-}|_{\tau=1,\,\zeta=0} = -\frac{3}{2}L^2 g(z)\,. \tag{3.21}$$

By expanding Eqs. (3.18, 3.19) in $z$ we get the two-loop amplitudes

$$A^{(2)}_{+++} = \frac{(C_A-C_F)L^4}{24}\left(2-3\tau^2+2\tau^3+3\tau^2\zeta^2\right), \tag{3.22}$$

$$A^{(2)}_{++-} = -\frac{(C_A-C_F)L^4}{24}\left(2+3\tau^2-6\tau^3+4\tau^4-3\tau^2\zeta^2\right), \tag{3.23}$$

which can be compared to the result of the explicit calculation [45]. The comparison, however, is not straightforward since in [45] the canonical symmetric form of the Catani's operator has been adopted

$$I^{(1)}_{\rm sym} = -\frac{C_A}{2\epsilon^2}\left[\left(\frac{-s}{\mu^2}\right)^{-\epsilon} + \left(\frac{-t}{\mu^2}\right)^{-\epsilon} + \left(\frac{-u}{\mu^2}\right)^{-\epsilon}\right], \tag{3.24}$$

which differs from Eq. (2.7) by

$$\delta I^{(1)} = -\frac{C_A}{8} L^2 \left[(1-\tau)^2 - \zeta^2\right]. \tag{3.25}$$

Thus for the helicity form factors defined in [45] we need to perform the infrared matching [49, 50]

$$\bar{\Omega}^{(2l),\rm fin}_{+++} = A^{(2)}_{+++} + \delta I^{(1)} A^{(1)}_{+++}, \qquad -\bar{\Omega}^{(2l),\rm fin}_{++-} = A^{(2)}_{++-} + \delta I^{(1)} A^{(1)}_{++-}, \tag{3.26}$$

which gives

$$\begin{aligned}\bar{\Omega}^{(2l),\rm fin}_{+++} &= -\frac{L^4}{24}\left[C_A + 2C_F - 6C_A\tau + \left(\frac{9}{2}C_A - 3C_F\right)\tau^2 + (C_A + 2C_F)\,\tau^3\right.\\ &\quad \left. -3C_A\zeta^2 + \left(-\frac{3}{2}C_A + 3C_F\right)\tau^2\zeta^2\right], \end{aligned}\tag{3.27}$$

$$\begin{aligned}\bar{\Omega}^{(2l),\rm fin}_{++-} &= -\frac{L^4}{24}\left[C_A + 2C_F - 6C_A\tau + \left(\frac{3}{2}C_A + 3C_F\right)\tau^2 + (3C_A + 6C_F)\,\tau^3\right.\\ &\quad \left. + \left(-\frac{5}{2}C_A + 4C_F\right)\tau^4 - 3C_A\zeta^2 + \left(\frac{3}{2}C_A - 3C_F\right)\tau^2\zeta^2\right]. \end{aligned}\tag{3.28}$$

Upon changing the notations $\zeta \to \xi$, $\tau \to -\tau$, for $C_A = 3$, $C_F = 4/3$, Eqs. (3.27, 3.28) agree with Eqs. (6.11, 6.12) of [45], providing a nontrivial check of our analysis. Note that for the symmetric choice of the infrared subtraction the physical $C_F - C_A$ color scaling is missing.

## 4 Logarithmic corrections to the kinematical distributions

We now can apply the result of the previous section to estimate the bottom quark effect to the differential cross section of the Higgs boson production in association with a jet. The dominant contribution is due to the interference of the top and bottom quark mediated amplitudes. We discard the partonic channel with the initial state quarks and consider the numerically dominant contribution of the gluon fusion. In the physical inclusive cross section the processes with additional real emission should be taken into account. If the real emission energy $E_{\rm real}$ exceeds $m_b$, it resolves the bottom quark loop and may generate a new type of the double logarithms. We, however, treat all the real emission as in the heavy quark limit with the local $ggH$ interaction. This approximation is discussed at the end of the section. Then both the virtual Sudakov and the soft real corrections are the same in the bottom and top quark mediated processes and the correction to the differential partonic cross section can be written as follows

$$d\sigma^{tb}_{gg\to Hg+X} = -\frac{3m_b^2}{m_H^2} L^2 C_t C_b(\tau,\zeta)\, d\tilde{\sigma}^{\rm eff}_{gg\to Hg+X}\,. \tag{4.1}$$

where $d\tilde{\sigma}^{\rm eff}_{gg\to Hg+X}$ is computed in the heavy top effective theory and

$$
\begin{aligned}
C_b(\tau,\zeta) = \frac{A_{+++}-A_{++-}}{2L^2} &= 1+\frac{z}{6}\left(1-\tau^3+\tau^4\right)\\
&+ z^2\left[\frac{1}{45}-\frac{\tau^3}{12}+\frac{\tau^4}{6}-\frac{7\tau^5}{60}+\frac{\tau^6}{30}+\frac{\zeta^2}{12}(\tau^3-\tau^4)\right]\\
&+ z^3\left[\frac{1}{420}-\frac{\tau^3}{48}+\frac{\tau^4}{16}-\frac{\tau^5}{12}+\frac{23\tau^6}{360}-\frac{143\tau^7}{5040}+\frac{31\tau^8}{5040}\right.\\
&\left.+\ \zeta^2\left(\frac{\tau^3}{24}-\frac{\tau^4}{12}+\frac{\tau^5}{20}-\frac{\tau^6}{180}\right)-\frac{\zeta^4}{48}(\tau^3-\tau^4)\right]+\dots
\end{aligned}
\tag{4.2}
$$

is a function of the transverse momentum and rapidity. Convolution of Eq. (4.1) with the parton distribution functions gives the correction to the kinematical distributions of the $pp\to Hj+X$ production in the threshold approximation.

However, within a well-motivated approximation the analysis of the hadronic cross section can be significantly simplified [35]. Indeed, the dependence of Eq. (4.2) on the jet rapidity is very weak. It starts with $\mathcal{O}(z^2)$ and the rapidity-dependent terms include at least the second power of the variable $\zeta$ and the third power of the variable $\tau$. If the soft jet is emitted at large rapidity, we have $|\zeta|\approx 1$ and $\tau\ll 1$. On the contrary, central emission with the large transverse momentum implies $|\zeta|\ll 1$ and $\tau\approx 1$. Therefore, the rapidity-dependent terms are small everywhere and can be neglected. Then $C_b(\tau,0)$ factors out from the parton distribution function integral. At the same time the transverse momentum dependence of the coefficients in Eq. (4.2) turned out to be very weak too, Fig. 5, and it can be approximated by $C_b=C_b(0,\zeta)=C_b(1,0)$ with a very high accuracy. This is quite a nontrivilal fact since the individual helicity form factors strongly depend on the transverse momentum, Eqs. (3.20, 3.21). The coefficient $C_b=1+\sum_{n=1}^{\infty}c_n$ is saturated by the eikonal dipole part of the factorization formula with the jet emission decoupled from the quark loop and describes the non-Sudakov logarithmic corrections to the total cross section of the Higgs boson production. It is known through the NLL approximation [39]. The first three terms of the series read

$$
\begin{aligned}
c_1 &= \frac{z}{6}+C_F\frac{\alpha_s L}{4\pi}\,,\\
c_2 &= \frac{z^2}{45}+\frac{x}{5}\frac{\alpha_s L}{4\pi}\left[\frac{3}{2}C_F-\beta_0\left(\frac{5}{6}\frac{L_\mu}{L}-\frac{1}{3}\right)\right],\\
c_3 &= \frac{z^3}{420}+\frac{x^2}{5}\frac{\alpha_s L}{4\pi}\left[\frac{5}{21}C_F-\beta_0\left(\frac{2}{9}\frac{L_\mu}{L}-\frac{2}{21}\right)\right],
\end{aligned}
\tag{4.3}
$$

where $\beta_0=\frac{11}{3}C_A-\frac{4}{3}T_F n_l$, $n_l=5$, $L_\mu=\ln(s/\mu^2)$ and $\mu$ is the renormalization scale of $\alpha_s$ in the double-logarithmic variable $z$. Thus neglecting the tiny soft dipole and symmetric contributions to the factorization formula we can extend the NLL analysis of the total threshold cross section to the process with the final jet. This gives

$$
d\sigma^{tb}_{pp\to Hj+X}=-\frac{3m_b^2}{m_H^2}\left(\frac{\alpha_s(m_H)}{\alpha_s(m_b)}\right)^{\gamma_m^{(1)}/\beta_0}L^2C_bC_t\,d\tilde{\sigma}^{\rm eff}_{pp\to Hj+X}\,,
\tag{4.4}
$$

where $\gamma_m^{(1)} = 3C_F$ is the one-loop mass anomalous dimension and the renormalization group factor sets the physical scale of the bottom quark Yukawa coupling to $m_H$. Clearly, in this approximation we miss the NLL terms in $C_b(\tau, \zeta)$ with the non-trivial dependence on the kinematical variables. However, the analysis of the total cross section [39] indicates that the main effect of the non-Sudakov NLL contribution is in fixing the physical renormalization scales of the strong and Yukawa couplings in the leading-logarithmic result, which is included into Eq. (4.4).

It is instructive to rewrite this formula in a way which relates the $K$-factors for the top-bottom interference contribution and the top quark mediated cross section in the heavy top limit

$$d\sigma^{tb}_{pp\to Hj+X} = \left[\frac{C_b}{C_t}\left(\frac{\alpha_s(m_H)}{\alpha_s(m_b)}\right)^{\gamma_m^{(1)}/\beta_0}\right]\left(\frac{d\sigma^{tb}_{pp\to Hj+X}}{d\sigma^{tt}_{pp\to Hj+X}}\right)^{\rm LO} d\sigma^{tt}_{pp\to Hj+X}\,. \qquad (4.5)$$

Here the first factor accounts for the difference between the QCD corrections to the top and bottom quark contributions. In the above approximation this difference does not depend on the kinematical variables and does not exceed twenty percent in the NLO, in agreement with the full calculation [17]. The NNLO term in the expansion of this factor amounts of a few percent so that the total NNLO correction to the bottom quark contribution is dominated by the K-factor of the top quark mediated cross section [7–11]. Note that in the NLL approximation there is no difference between the pole mass $m_b$ and the $\overline{\rm MS}$ mass $\bar{m}_b(\bar{m}_b)$ but the use of the latter as the bottom quark mass parameter in the available exact fixed order results [39] boosts the convergence of the perturbative expansion (*cf.* [42]).

An important property of our result is that at $p_\perp \to m_b$ the soft dipole and symmetric parts of the factorization formula vanish and in the remaining eikonal part the jet is completely decoupled from the quark loop. Thus, the transverse momentum evolution of the cross section matches the one for $p_\perp \ll m_b$ derived in the heavy bottom effective theory with $C_b$ playing the role of the Wilson coefficient of the local $ggH$ interaction. Moreover, up to a tiny corrections even for $m_q \ll p_\perp$ the jet emission factors out as in the large quark mass limit. We may assume a similar behavior of the unresolved real emission with the energy below the energy of the jet. Then our result, being formally valid for $E_{\rm real} \ll m_b$, can be extended to the entire energy region where the real emission is not kinematically suppressed. Such a behavior is supported by the observed structure of the NLO corrections [17] and explains the accuracy of the NNLO result for the total cross section [39], where the (soft) real emission was approximated by the heavy quark limit expression.

Thus our result can be used to complete the renormalization group analysis of the bottom quark contribution to the transverse momentum observables. The well-understood logarithms of the ratio $p_\perp/m_H$ in the effective theory cross section $d\tilde{\sigma}^{\rm eff}_{pp\to Hj+X}$ are known to very high orders of the logarithmic expansion [51, 52]. At the same time the non-Sudakov logarithms discussed above are missing in the existing analysis of the bottom quark effects [21, 22]. They can be fully taken into account in the leading-logarithmic approximation through Eq. (4.1). However, as it has been explained above, their effect can be very well approximated by Eq. (4.4) with the all-order NLL expression for $C_b$ available in [39]. Note

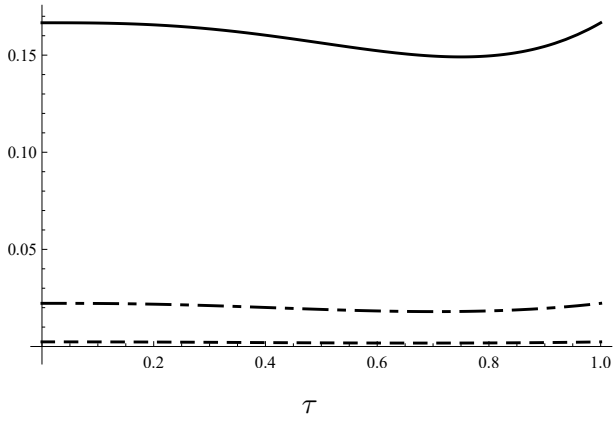



**Figure 5**. The transverse momentum dependence of $z^n$ coefficients in Eq. (4.2) at central rapidity $\zeta = 0$ for $n = 1$ (solid line), $n = 2$ (dot-dashed line) and $n = 3$ (dashed line). For each coefficient the boundary values at $\tau = 0,\ 1$ are equal.

that the multiple soft emission with the transverse momenta of order $p_\perp$ can in general resolve the bottom quark loop and the corresponding logarithmic corrections differ from the effective theory result. However, the arguments given above indicate that the difference is numerically small. Moreover, with the increasing number of soft real partons their energies in the kinematically unsuppressed region are reduced and the heavy bottom effective theory becomes more and more accurate.

## 5 Conclusion

In this paper, we have studied a light quark mediated production of the Higgs boson in association with a jet in gluon fusion in the double-logarithmic approximation for the intermediate values of transverse momentum $m_q \ll p_\perp \ll m_H$. The factorization formula has been obtained for the helicity amplitudes, which separates the universal infrared Sudakov renormalization of the external on-shell gluon states and the non-Sudakov logarithms characteristic to the power suppressed contributions. The physical origin of the latter is the color charge nonconservation due to exchange of a soft quark between each pair of the eikonal lines associated with the external on-shell gluons, as for the $gg \to H$ amplitude. Thus, despite significantly more complex geometry of the double-logarithmic region, with the physical definition of the infrared subtraction operator the full QCD result can be obtained from its abelian part by the color factor adjustment $C_F \to C_F - C_A$. The non-Sudakov double-logarithmic corrections to the $gg \to Hg$ partonic cross section show very weak dependence on the jet transverse momentum and rapidity in the whole allowed region of the kinematical variables and for the both boundary values $p_\perp \to m_q$, $m_H$ reduce precisely to the $gg \to H$ case.[4]

This indicates that at least in the leading logarithmic approximation up to a tiny corrections the jet emission factors out even for $m_q \ll p_\perp$ as in the large quark mass limit

[4]This is not valid for the individual helicity amplitudes at $p_\perp \to m_H$.

but with a different mass dependence of the Wilson coefficient. This, in particular, explains the observed structure of the NLO result [17] as well as the success of the NNLO analysis of the total cross section [39], with the real emission being treated in the same way as for the top quark mediated process.

Relying on this property we have derived the NNLO bottom quark contribution to the kinematical distributions of the $pp \to Hj + X$ production with the logarithmic accuracy in terms of the transverse momentum dependent $K$-factors previously computed in the heavy top quark limit [7–11]. In fact, the NNLO $K$-factors for the top-bottom interference and the top quark contribution to the kinematical distributions should agree within a few percent in the given region of transverse momentum.

Our all-order result for the amplitudes also provides the last missing ingredient of the renormalization group analysis of the bottom quark effect in the Higgs boson plus jet production at the LHC [21, 22]. To our knowledge this is the first double-logarithmic asymptotic result for a QCD amplitude which captures all-order dependence on two kinematical variables, *i.e.* simultaneously sums up three different types of the large logarithms.

## Acknowledgments

A.P is grateful to Babis Anastasiou and Kirill Melnikov for many helpful discussions and comments. A.R. would like to thank Lorenzo Tancredi for useful communications. The work of T.L. is supported in part by IHEP under Grants No. Y9515570U1, the National Natural Science Foundation of China (NNSFC) under grant No.12375082 and No.12135013. The work of A.P. is supported in part by NSERC and Perimeter Institute for Theoretical Physics. The work of A.R. is supported by NSERC.

## References

[1] G. Aad *et al.* [ATLAS and CMS], JHEP **08**, 045 (2016).

[2] M. Cepeda, S. Gori, P. Ilten, M. Kado, F. Riva, R. Abdul Khalek, A. Aboubrahim, J. Alimena, S. Alioli and A. Alves, *et al.* CERN Yellow Rep. Monogr. **7**, 221 (2019).

[3] C. Arnesen, I. Z. Rothstein and J. Zupan, Phys. Rev. Lett. **103**, 151801 (2009).

[4] E. Bagnaschi, G. Degrassi, P. Slavich and A. Vicini, JHEP **02**, 088 (2012).

[5] S. Dawson, I. M. Lewis and M. Zeng, Phys. Rev. D **90**, 093007 (2014).

[6] M. Grazzini, A. Ilnicka, M. Spira and M. Wiesemann, JHEP **03**, 115 (2017).

[7] X. Chen, T. Gehrmann, E. W. N. Glover and M. Jaquier, Phys. Lett. B **740**, 147 (2015).

[8] R. Boughezal, F. Caola, K. Melnikov, F. Petriello and M. Schulze, Phys. Rev. Lett. **115**, 082003 (2015).

[9] R. Boughezal, C. Focke, W. Giele, X. Liu and F. Petriello, Phys. Lett. B **748**, 5 (2015).

[10] F. Caola, K. Melnikov and M. Schulze, Phys. Rev. D **92**, 074032 (2015).

[11] X. Chen, J. Cruz-Martinez, T. Gehrmann, E. W. N. Glover and M. Jaquier, JHEP **10**, 066 (2016).

[12] S. P. Jones, M. Kerner and G. Luisoni, Phys. Rev. Lett. **120**, 162001 (2018).

[13] U. Baur and E. W. N. Glover, Nucl. Phys. B **339**, 38-66 (1990).

[14] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos and B. Mistlberger, JHEP **1605**, 058 (2016).

[15] F. Bishara, U. Haisch, P. F. Monni and E. Re, Phys. Rev. Lett. **118**, 121801 (2017).

[16] Y. Soreq, H. X. Zhu and J. Zupan, JHEP **12**, 045 (2016).

[17] J. M. Lindert, K. Melnikov, L. Tancredi and C. Wever, Phys. Rev. Lett. **118**, 252002 (2017).

[18] R. Bonciani, V. Del Duca, H. Frellesvig, M. Hidding, V. Hirschi, F. Moriello, G. Salvatori, G. Somogyi and F. Tramontano, Phys. Lett. B **843**, 137995 (2023).

[19] H. Mantler and M. Wiesemann, Eur. Phys. J. C **73**, 2467 (2013).

[20] M. Grazzini and H. Sargsyan, JHEP **09**, 129 (2013).

[21] A. Banfi, P. F. Monni and G. Zanderighi, JHEP **01**, 097 (2014).

[22] F. Caola, J. M. Lindert, K. Melnikov, P. F. Monni, L. Tancredi and C. Wever, JHEP **09**, 035 (2018).

[23] M. I. Kotsky and O. I. Yakovlev, Phys. Lett. B **418**, 335 (1998).

[24] A. A. Penin, Phys. Lett. B **745**, 69 (2015).

[25] T. Liu, A. A. Penin and N. Zerf, Phys. Lett. B **771**, 492 (2017).

[26] V. V. Sudakov, Sov. Phys. JETP **3**, 65 (1956) [Zh. Eksp. Teor. Fiz. **30**, 87 (1956)].

[27] J. Frenkel and J. C. Taylor, Nucl. Phys. B **116**, 185 (1976).

[28] A. H. Mueller, Phys. Rev. D **20**, 2037 (1979).

[29] J. C. Collins, Phys. Rev. D **22**, 1478 (1980).

[30] A. Sen, Phys. Rev. D **24**, 3281 (1981).

[31] G. F. Sterman, Nucl. Phys. B **281**, 310 (1987).

[32] T. Liu and A. A. Penin, Phys. Rev. Lett. **119**, 262001 (2017).

[33] T. Liu and A. Penin, JHEP **1811**, 158 (2018).

[34] T. Liu, S. Modi and A. A. Penin, JHEP **02**, 170 (2022).

[35] K. Melnikov and A. Penin, JHEP **1605**, 172 (2016).

[36] A. A. Penin and N. Zerf, Phys. Lett. B **760**, 816 (2016).

[37] M. Delto, C. Duhr, L. Tancredi and Y. J. Zhu, [arXiv:2311.06385 [hep-ph]].

[38] A. A. Penin, JHEP **2004**, 156 (2020).

[39] C. Anastasiou and A. Penin, JHEP **2007**, 195 (2020).

[40] Z. L. Liu, B. Mecaj, M. Neubert and X. Wang, JHEP **2101**, 077 (2021).

[41] Z. L. Liu, M. Neubert, M. Schnubel and X. Wang, JHEP **06**, 183 (2023).

[42] M. Czakon, F. Eschment, M. Niggetiedt, R. Poncelet and T. Schellenberger, [arXiv:2312.09896 [hep-ph]].

[43] Y. Schroder and M. Steinhauser, JHEP **0601**, 051 (2006).

[44] K. G. Chetyrkin, J. H. Kuhn and C. Sturm, Nucl. Phys. B **744**, 121 (2006).

[45] K. Melnikov, L. Tancredi and C. Wever, JHEP **1611**, 104 (2016).

[46] T. Gehrmann, M. Jaquier, E. W. N. Glover and A. Koukoutsakis, JHEP **02**, 056 (2012).

[47] S. Catani, Phys. Lett. B **427**, 161 (1998).

[48] C. R. Schmidt, Phys. Lett. B **413**, 391-395 (1997).

[49] A. A. Penin, Phys. Rev. Lett. **95**, 010408 (2005).

[50] A. A. Penin, Nucl. Phys. B **734**, 185 (2006).

[51] A. Banfi, F. Caola, F. A. Dreyer, P. F. Monni, G. P. Salam, G. Zanderighi and F. Dulat, JHEP **04**, 049 (2016).

[52] W. Bizon, P. F. Monni, E. Re, L. Rottoli and P. Torrielli, JHEP **02**, 108 (2018).