SIAM J. NUMER. ANAL.
Vol. 56, No. 6, pp. 3407–3429

© 2018 Society for Industrial and Applied Mathematics

DISCRETE H^1-INEQUALITIES FOR SPACES ADMITTING M-DECOMPOSITIONS*

BERNARDO COCKBURN[†], GUOSHENG FU[‡], AND WEIFENG QIU[§]

Abstract. We find new discrete H^1- and Poincaré–Friedrichs inequalities by studying the invertibility of the discontinuous Galkerkin (DG) approximation of the flux for local spaces admitting M-decompositions. We then show how to use these inequalities to define and analyze new, superconvergent hybridizable DG (HDG) and mixed methods for which the stabilization function is defined in such a way that the approximations satisfy new H^1-stability results with which their error analysis is greatly simplified. We apply this approach to define a wide class of energy-bounded, superconvergent HDG and mixed methods for the incompressible Navier–Stokes equations defined on unstructured meshes using, in two dimensions, general polygonal elements and, in three dimensions, general, flat-faced tetrahedral, prismatic, pyramidal, and hexahedral elements.

Key words. discontinuous Galerkin, hybridization, stability, superconvergence, Navier–Stokes

AMS subject classifications. 65N30, 65M60, 35L65

DOI. 10.1137/17M1144830

1. Introduction.

In this paper, we obtain new discrete stability inequalities with which we carry out the first a priori error analysis of a wide class of hybridizable discontinuous Galerkin (HDG) and mixed methods for the Navier–Stokes equations. The methods are defined on unstructured meshes using, in two dimensions (2D), general polygonal elements and, in 3D, general, flat-faced tetrahedral, prismatic, pyramidal, and hexahedral elements. They are a direct extension of the corresponding methods introduced for the Stokes flow in [13]. We prove optimal error estimates in all the unknowns as well as superconvergence results for the approximate velocity. By this, we mean that a new approximation for the velocity can be obtained in an elementwise manner which converges faster than the original velocity approximation.

The unifying feature of the above-mentioned class of methods is that they are defined by using the theory of M-decompositions. Using this theory, superconvergent HDG and mixed methods have been devised for diffusion [14, 10, 11], for linear incompressible flow [13], and for linear elasticity [9]. The theory of M-decompostions has also been used to obtain commuting de Rham sequences [12]. Here, we use it to obtain the above-mentioned new discrete inequalities.

To better explain our results, we introduce the HDG and mixed methods for steady-state diffusion

$$\mathrm{c}\boldsymbol{q} + \nabla u = 0, \quad \nabla \cdot \boldsymbol{q} = f \text{ in } \Omega, \quad \text{ and } \quad u = g \text{ on } \partial\Omega,$$

---

*Received by the editors August 25, 2017; accepted for publication (in revised form) October 18, 2018; published electronically December 4, 2018.
http://www.siam.org/journals/sinum/56-6/M114483.html

Funding: The work of the first author was supported in part by the National Science Foundation (grant DMS-1522657) and by the University of Minnesota Supercomputing Institute. The work of the third author was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project CityU 11302014).

[†] School of Mathematics, University of Minnesota, Vincent Hall, Minneapolis, MN 55455 (cockburn@math.umn.edu).

[‡] Division of Applied Mathematics, Brown University, Providence, RI 02912 (guosheng_fu@brown.edu).

[§] Corresponding author. Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong, China (weifeqiu@cityu.edu.hk).

and introduce the concept of an M-decomposition. We then describe the inequalities we want to obtain and, finally, describe how we are going to apply them to the analysis of HDG and mixed methods for the Navier–Stokes equations.

HDG methods and M-decompositions. To define the HDG methods, we follow [15]. Thus, we take the domain $\Omega \subset \mathbb{R}^d$ to be a polygon if $d = 2$ and a polyhedron if $d = 3$. We triangulate it with a conforming mesh $\mathcal{T}_h := \{K\}$ made of shape-regular polygonal/polyhedral elements K. We set $\partial\mathcal{T}_h := \{\partial K : K \in \mathcal{T}_h\}$, and denote by $\mathcal{F}_h$ the set of faces F of the elements $K \in \mathcal{T}_h$. We also denote by $\mathcal{F}(K)$ the set of faces F of the element K.

The HDG method seeks an approximation to $(u, \boldsymbol{q}, u|_{\mathcal{F}_h})$, $(u_h, \boldsymbol{q}_h, \widehat{u}_h)$, in the finite dimensional space $W_h \times \boldsymbol{V}_h \times M_h$, where

$$\begin{aligned}\boldsymbol{V}_h &:= \{\boldsymbol{v} \in \boldsymbol{L}^2(\mathcal{T}_h) : \boldsymbol{v}|_K \in \boldsymbol{V}(K),\ K \in \mathcal{T}_h\},\\ W_h &:= \{w \in L^2(\mathcal{T}_h) : w|_K \in W(K), K \in \mathcal{T}_h\},\\ M_h &:= \{\widehat{w} \in L^2(\mathcal{F}_h) : \widehat{w}|_F \in M(F),\ F \in \mathcal{F}_h\},\end{aligned}$$

and determines it as the only solution of the following weak formulation:

$$\begin{aligned}
&\text{(1.1a)} && (\mathrm{c}\,\boldsymbol{q}_h\,,\,\boldsymbol{v})_{\mathcal{T}_h} - (u_h\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle \widehat{u}_h\,,\,\boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} &&= 0,\\
&\text{(1.1b)} && -(\boldsymbol{q}_h\,,\,\nabla w)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}\,,\,w\rangle_{\partial\mathcal{T}_h} &&= (f\,,\,w)_{\mathcal{T}_h},\\
&\text{(1.1c)} && \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n} = \boldsymbol{q}_h\cdot\boldsymbol{n} + \alpha(u_h - \widehat{u}_h) && \text{on} \quad \partial\mathcal{T}_h,\\
&\text{(1.1d)} && \langle \widehat{\boldsymbol{q}}_h\cdot\boldsymbol{n}, \widehat{w}\rangle_{\partial\mathcal{T}_h\setminus\partial\Omega} &&= 0,\\
&\text{(1.1e)} && \langle \widehat{u}_h, \widehat{w}\rangle_{\partial\Omega} &&= \langle u_D \widehat{w}\rangle_{\partial\Omega}
\end{aligned}$$

for all $(w, \boldsymbol{v}, \widehat{w}) \in W_h \times \boldsymbol{V}_h \times M_h$. Here we write $(\eta\,,\,\zeta)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}(\eta,\zeta)_K$, where $(\eta,\zeta)_D$ denotes the integral of $\eta\zeta$ over the domain $D \subset \mathbb{R}^n$. We also write $\langle\eta\,,\,\zeta\rangle_{\partial\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h}\langle\eta\,,\,\zeta\rangle_{\partial K}$, where $\langle\eta\,,\,\zeta\rangle_D$ denotes the integral of $\eta\zeta$ over the domain $D \subset \mathbb{R}^{n-1}$. When vector-valued functions are involved, we use a similar notation.

The different HDG methods are obtained by choosing the local spaces $\boldsymbol{V}(K)$, $W(K)$, and

$$M(\partial K) := \{\widehat{w} \in L^2(\partial K) :\ \widehat{w}|_F \in M(F)\ \forall F \in \mathcal{F}(K)\},$$

and the *linear local stabilization* function α. It turns out [14] that if we can decompose $\boldsymbol{V}(K) \times W(K)$ in such a way that

$$\begin{aligned}\boldsymbol{V}(K) &= \widetilde{\boldsymbol{V}}(K) \oplus \widetilde{\boldsymbol{V}}^{\perp}(K),\\ W(K) &= \widetilde{W}(K) \oplus \widetilde{W}^{\perp}(K),\\ M(\partial K) &= \widetilde{\boldsymbol{V}}^{\perp}(K)\cdot\boldsymbol{n}|_{\partial K} \oplus \widetilde{W}^{\perp}(K)|_{\partial K},\end{aligned}$$

and use a couple of simple inclusion properties, that it is possible to find a stabilization function α such that the resulting HDG ($\alpha \neq 0$) or mixed method ($\alpha = 0$) is superconvergent. Since this decomposition is essentially induced by the space $M(\partial K)$, it is called an $M(\partial K)$-decomposition of the space $\boldsymbol{V}(K) \times W(K)$. The explicit construction of those spaces for general polygonal elements was carried in [10] (see the main examples in Table 2.1) and for flat-faced general pyramids, prisms, and hexahedral elements in [11].

Invertibility of the discrete gradient operator. In this paper, we study the invertibility properties of the mapping

$$W(K)\times M(\partial K) \longrightarrow \boldsymbol{V}(K), \tag{1.2a}$$

$$(u_h, \widehat{u}_h) \qquad \longmapsto \boldsymbol{q}_h, \tag{1.2b}$$

where

$$(\mathrm{c}\,\boldsymbol{q}_h, \boldsymbol{v})_K = (u_h, \nabla\cdot\boldsymbol{v})_K - \langle \widehat{u}_h, \boldsymbol{n}\cdot\boldsymbol{v}\rangle_{\partial K} \quad \forall\, \boldsymbol{v}\in \boldsymbol{V}(K) \tag{1.2c}$$

for spaces $\boldsymbol{V}(K)\times W(K)$ admitting an $M(\partial K)$-decomposition [14]. This mapping is a discrete version of the *constitutive* equation relating a vector-valued function $\boldsymbol{q}$ and a scalar-valued function u:

$$\mathrm{c}\,\boldsymbol{q} = -\nabla u,$$

where c and c^{-1} are bounded, symmetric, and uniformly positive definite matrix-valued functions, and have been used in, arguably, *all* DG and hybridized versions of mixed methods. In particular, it captures the first equation defining the HDG method for steady-state diffusion. We present new discrete versions of the estimates

$$\begin{aligned} \|\nabla u\|_K^2 &= \|\mathrm{c}\,\boldsymbol{q}\|_K^2 && \text{(trivial)},\\ h_K^{-2}\|u-\overline{u}^K\|_K^2 &\le C\,\|\mathrm{c}\,\boldsymbol{q}\|_K^2 && \text{(Poincaré–Friedrichs)}, \end{aligned}$$

where $\overline{\zeta}^D$ denotes the average of ζ on D and $\|\cdot\|_D$ is the $L^2(D)$-norm. They are expressed in terms of the (equivalent) seminorms

$$|(u_h,\widehat{u}_h)|_{1,K}^2 := \|\nabla u_h\|_K^2 + h_K^{-1}\,\|u_h-\widehat{u}_h\|_{\partial K}^2, \tag{1.3a}$$

$$|(u_h,\widehat{u}_h)|_{\mathrm{PF},K}^2 := \|u_h-\overline{\widehat{u}_h}^{\partial K}\|_K^2 + h_K\,\|\widehat{u}_h-\overline{\widehat{u}_h}^{\partial K}\|_{\partial K}^2 \tag{1.3b}$$

and are, essentially, of the form

$$\begin{aligned} |\,(u_h,\widehat{u}_h)\,|_{1,K}^2 &\le C\,\left(\|\mathrm{c}\,\boldsymbol{q}_h\|_K^2 + h_K^{-1}\|P_{M_S}(u_h-\widehat{u}_h)\|_{\partial K}^2\right) && (H^1),\\ h_K^{-2}\,|\,(u_h,\widehat{u}_h)\,|_{\mathrm{PF},K}^2 &\le C\,\left(\|\mathrm{c}\,\boldsymbol{q}_h\|_K^2 + h_K^{-1}\|P_{M_S}(u_h-\widehat{u}_h)\|_{\partial K}^2\right) && \text{(Poincaré–Friedrichs)}, \end{aligned}$$

where $M_S = M_S(\partial K)$, referred to as the *stabilization* space, is an easy-to-compute subspace of the space $M(\partial K)$ whose dimension is chosen to be *minimal*, and P_{M_S} is its corresponding L^2-projection. These inequalities, which are nothing but *stabilized* versions of inf-sup conditions [1], are the key ingredients for our analysis of HDG and mixed methods. They generalize, to *all* spaces admitting M-decompositions, the H^1-inequality obtained with $M_S=\emptyset$ in [19, Proposition 3.2], for the well known Raviart–Thomas spaces for simplexes, and, for smaller spaces, in [7, Theorem 3.2] with M_S equal to the restriction of $M(\partial K)$ onto an arbitrary face F_K on which $\widehat{u}_h$ was set to coincide with u_h.

Application to the Navier–Stokes equations. We show how to do that, not in the relatively simple case of convection-diffusion equations, but in the more difficult case of the velocity gradient-velocity-pressure formulation of the steady-state incompressible Navier–Stokes equations in two- and three-space dimensions:

$$\mathrm{L} = \nabla \boldsymbol{u} \quad \text{in} \quad \Omega, \tag{1.4a}$$

$$-\nu \boldsymbol{\nabla}\cdot \mathrm{L} + \nabla\cdot(\boldsymbol{u}\otimes\boldsymbol{u}) + \nabla p = \boldsymbol{f} \quad \text{in} \quad \Omega, \tag{1.4b}$$

$$\nabla\cdot\boldsymbol{u} = 0 \quad \text{in} \quad \Omega, \tag{1.4c}$$

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on} \quad \partial\Omega, \tag{1.4d}$$

$$\int_\Omega p = 0\ , \tag{1.4e}$$

where L is the velocity gradient, $\boldsymbol{u}$ is the velocity, p is the pressure, ν is the kinematic viscosity, and $\boldsymbol{f} \in L^2(\Omega)^d$ is the external body force.

Let us compare our results with those in [4] where the only error analysis for HDG methods for the Navier–Stokes equations has been recently carried out. Let $(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)$ be an approximation of the velocity $(\boldsymbol{u}|_\Omega, \boldsymbol{u}|_{\mathcal{F}_h})$, where $\mathcal{F}_h$ denotes the set of faces of the mesh $\mathcal{T}_h$ of the domain Ω, and let L_h be an approximation of the velocity gradient $\mathrm{L}|_\Omega$. In [4], the authors considered unstructured meshes made of simplexes, spaces of polynomials of degree k, and a stabilization function α such that

$$\langle \alpha(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h), \boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h\rangle_{\partial K} = h_K^{-1}\|(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n}\|_{\partial K}^2.$$

For this HDG method, optimal convergence order for all unknowns as well as the superconvergence of the velocity was obtained by using the novel upper bound

$$|\!|\!|(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)|\!|\!|_{1,\mathcal{T}_h}^2 \le C \sum_{K\in\mathcal{T}_h} (\|\mathrm{L}_h\|_K^2 + h_K^{-1}\|(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\cdot\boldsymbol{n}\|_{\partial K}^2),$$

where the discrete H^1-norm $|\!|\!|\cdot|\!|\!|_{\mathcal{T}_h}$ is given by

$$|\!|\!|(u_h, \widehat{u}_h)|\!|\!|_{1,\mathcal{T}_h} := \left(\sum_{K\in\mathcal{T}_h} |(u_h, \widehat{u}_h)|_{1,K}^2\right)^{1/2}.$$

In contrast, in this paper, stronger results are obtained for a wide class of HDG and mixed methods defined on a variety of element shapes: general polgonal elements in 2D, and tetrahedral, pyramidal, prismatic, and hexahedral elements in 3D. The local spaces defining these methods are those used for the corresponding methods for the Stokes equations of incompressible flow proposed in [13]; the stabilization function is not the same though. The spaces are constructed by using, as building blocks, the local spaces $\boldsymbol{V}(K)\times W(K)$ admitting an $M(\partial K)$-decomposition introduced in [14] for steady-state diffusion.

To obtain the new discrete inequalities, we proceed in two steps. First, we show that for all these methods, we have the discrete $\boldsymbol{H}^1$-inequality

$$|\!|\!|(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)|\!|\!|_{1,\mathcal{T}_h}^2 \le C \sum_{K\in\mathcal{T}_h} (\|\mathrm{L}_h\|_K^2 + h_K^{-1}\|P_{M_S}(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\|_{\partial K}^2).$$

We then show that if we *define* a stabilization function α such that

$$\langle \alpha(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h), \boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h\rangle_{\partial K} = h_K^{-1}\|P_{M_S}(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\|_{\partial K}^2,$$

we obtain new $\boldsymbol{H}^1$-boundedness results for the approximation, and new $\boldsymbol{H}^1$-stability inequalities, with which can easily obtain the above-mentioned convergence properties.

Organization of the paper. The rest of the paper is organized as follows. In section 2, we present the general properties of the local spaces admitting

M-decompositions and those of the the stabilization subspaces M_S; specific choices of M_S for the main spaces admitting M-decompositions are also provided. We then present and discuss our main result, namely, the new discrete inequalities of Theorem 2.3 which we prove in section 3. In section 4, we define our HDG and mixed methods for the incompressible Navier–Stokes equations and present their energy-boundedness and superconvergence properties; their proofs are provided in section 5. We end with some concluding remarks in section 6.

2. The main result.

In this section, we present and discuss our main result, namely, the discrete $\boldsymbol{H}^1$- and Poincaré–Friedrichs inequalities of Theorem 2.3; their proof is postponed to section 3. We first present the two ingredients needed to obtain these inequalities, namely, the spaces admitting M-decompositions and a stabilization subspace of the trace space $M(\partial K)$.

2.1. Notation.

Given a domain $D \subset \mathbb{R}^n$, we denote by $\mathcal{P}_k(D)$ and $\widetilde{\mathcal{P}}_k(D)$ the space of polynomials of degree no greater than k, and the space of homogeneous polynomials of degree k, respectively, defined on the domain D. When D is a unit square with coordinates (x, y), we denote by $\mathcal{Q}_k(D) := \mathcal{P}_k(x) \otimes \mathcal{P}_k(y)$ and $\widetilde{\mathcal{Q}}_k(D) := \widetilde{\mathcal{P}}_k(x) \otimes \widetilde{\mathcal{P}}_k(y)$ the space of tensor-product polynomials of degree no greater than k, and the space of homogeneous tensor-product polynomials of degree k, respectively. We use a similar notation on tensor-product polynomial spaces on the unit cube. When $D := B \otimes I$ is a unit prism having a triangular base B with coordinates (x, y) and a z-directional edge I, we denote by $\mathcal{P}_{k|k}(D) := \mathcal{P}_k(x, y) \otimes \mathcal{P}_k(z)$ and $\widetilde{\mathcal{P}}_{k|k}(D) := \widetilde{\mathcal{P}}_k(x, y) \otimes \widetilde{\mathcal{P}}_k(z)$ the space of tensor-product polynomials of degree no greater than k, and the space of homogeneous tensor-product polynomials of degree k, respectively. Vector-valued spaces are denoted with a superscript d (the space dimension); for example, $\mathcal{P}_k(K)^d$ is the space of vectors whose entries lie in $\mathcal{P}_k(K)$.

We denote by $\|\cdot\|_{W^{m,p}(D)}$ the standard $W^{m,p}$-Sobolev norm on the domain $D \subset \mathbb{R}^d$. For the Hilbert space $H^m(D) := W^{m,2}(D)$, we simply write $\|\cdot\|_{m,D}$ instead of $\|\cdot\|_{H^m(D)}$, and $\|\cdot\|_D$ instead of $\|\cdot\|_{0,D}$. Similarly, when $p = \infty$, we write $\|\cdot\|_{m,\infty,D}$ instead of $\|\cdot\|_{W^{m,\infty}(D)}$, and $\|\cdot\|_{\infty,D}$ instead of $\|\cdot\|_{0,\infty,D}$. For a given a second-order tensor c, we denote by $\|\cdot\|_{\mathrm{c},D}$ the c-weighted L^2-norm on the domain D.

Finally, we denote by $\lambda_{\mathrm{c}}^{\max}(K)$ the $L^\infty(K)$-norm of the maximum eigenvalue of the tensor c.

2.2. Examples of spaces $\boldsymbol{V}(K) \times W(K)$ admitting $M(\partial K)$-decompositions.

An M-decomposition relates the trace of the normal component of the space of approximate fluxes $\boldsymbol{V}(K)$ and the trace of the space of approximate scalars $W(K)$ with the space of approximate traces $M(\partial K)$. To define it, we need to consider the combined trace operator

$$\begin{aligned} \mathrm{tr} : \boldsymbol{V}(K) \times W(K) &\longrightarrow L^2(\partial K), \\ (\boldsymbol{v}, w) &\longmapsto (\boldsymbol{v} \cdot \boldsymbol{n} + w)|_{\partial K}. \end{aligned}$$

Definition 2.1 (The M-decomposition). *We say that $\boldsymbol{V}(K) \times W(K)$ admits an M-decomposition when*

(a) $\mathrm{tr}(\boldsymbol{V}(K) \times W(K)) \subset M(\partial K)$,

and there exists a subspace $\widetilde{\boldsymbol{V}}(K) \times \widetilde{W}(K)$ of $\boldsymbol{V}(K) \times W(K)$ satisfying

(b) $\nabla W(K) \times \nabla \cdot \boldsymbol{V}(K) \subset \widetilde{\boldsymbol{V}}(K) \times \widetilde{W}(K)$,

(c) $\mathrm{tr} : \widetilde{\boldsymbol{V}}^\perp(K) \times \widetilde{W}^\perp(K) \to M(\partial K)$ *is an isomorphism.*

TABLE 2.1
Spaces $\boldsymbol{V}(K) \times W(K)$ admitting an $M(\partial K)$-decomposition [14].

$\boldsymbol{V}(K)$	$W(K)$	Method
$M(\partial K) = \mathcal{P}_k(\partial K)$, K is a square.		
$\boldsymbol{\mathcal{Q}}_k \oplus \mathbf{curl}\,\mathrm{span}\{x^{k+1}y, x\,y^{k+1}\} \oplus \mathrm{span}\{\boldsymbol{x}\,x^k y^k\}$	$\mathcal{Q}_k$	$\mathbf{TNT}_{[k]}$ [16]
$\boldsymbol{\mathcal{Q}}_k \oplus \mathbf{curl}\,\mathrm{span}\{x^{k+1}y, x\,y^{k+1}\}$	$\mathcal{Q}_k$	$\mathbf{HDG}^{Q}_{[k]}$[16]
$\boldsymbol{\mathcal{Q}}_k \oplus \mathbf{curl}\,\mathrm{span}\{x^{k+1}y, x\,y^{k+1}\}$	$\mathcal{Q}_k \setminus \{x^k\,y^k\}$	$\mathbf{BDM}_{[k]}$
$M = \mathcal{P}_k(\partial K)$, K is a triangle.		
$\boldsymbol{\mathcal{P}}_k \oplus \boldsymbol{x}\,\widetilde{\mathcal{P}}_k$	$\mathcal{P}_k$	$\mathbf{RT}_k$ [20]
$\boldsymbol{\mathcal{P}}_k$	$\mathcal{P}_k$	$\mathbf{HDG}_k$[16]
$\boldsymbol{\mathcal{P}}_k$	$\mathcal{P}_{k-1}$	$\mathbf{BDM}_k$ [2]
$M = \mathcal{P}_k(\partial K)$, K is a square.		
$\boldsymbol{\mathcal{P}}_k \oplus \mathbf{curl}\,\mathrm{span}\{x^{k+1}y, x\,y^{k+1}\} \oplus \boldsymbol{x}\,\widetilde{\mathcal{P}}_k$	$\mathcal{P}_k$	[10]
$\boldsymbol{\mathcal{P}}_k \oplus \mathbf{curl}\,\mathrm{span}\{x^{k+1}y, x\,y^{k+1}\}$	$\mathcal{P}_k$	[10]
$\boldsymbol{\mathcal{P}}_k \oplus \mathbf{curl}\,\mathrm{span}\{x^{k+1}y, x\,y^{k+1}\}$	$\mathcal{P}_{k-1}$	$\mathbf{BDM}_{[k]}$ [2]
$M = \mathcal{P}_k(\partial K)$, K is a quadrilateral.		
$\boldsymbol{\mathcal{P}}_k \oplus_{i=1}^{ne} \mathbf{curl}\,\mathrm{span}\{\xi_4\,\lambda_3^k, \xi_4\,\lambda_4^k\} \oplus \boldsymbol{x}\,\widetilde{\mathcal{P}}_k$	$\mathcal{P}_k$	[10]
$\boldsymbol{\mathcal{P}}_k \oplus_{i=1}^{ne} \mathbf{curl}\,\mathrm{span}\{\xi_4\,\lambda_3^k, \xi_4\,\lambda_4^k\}$	$\mathcal{P}_k$	[10]
$\boldsymbol{\mathcal{P}}_k \oplus_{i=1}^{ne} \mathbf{curl}\,\mathrm{span}\{\xi_4\,\lambda_3^k, \xi_4\,\lambda_4^k\}$	$\mathcal{P}_{k-1}$	[10]

Here $\widetilde{\boldsymbol{V}}^{\perp}(K)$ and $\widetilde{W}^{\perp}(K)$ are the $L^2(K)$-orthogonal complements of $\widetilde{\boldsymbol{V}}(K)$ in $\boldsymbol{V}(K)$, and of $\widetilde{W}(K)$ in $W(K)$, respectively.

Local spaces $\boldsymbol{V}(K) \times W(K)$ admitting $M(\partial K)$-decompositions have been explicitly constructed in 2D for general polygonal elements K (see some examples in Table 2.1) in [10] and in 3D for four types of polyhedral elements K, namely, tetrahedra, pyramids, prisms, and hexahedra in [11]. As pointed out in the introduction, the main interest of these spaces is that they generate superconvergent HDG and mixed methods; see [14].

Let us explain the notation used in Table 2.1. By $\mathbf{curl}\,p$ we mean the vector $(-p_y, p_x)$. By $\{\mathbf{v}_i\}_{i=1}^4$ (and $\mathbf{v}_5 := \mathbf{v}_1$), we mean the four vertices of a quadrilateral; the vertices are ordered in a counterclockwise manner. We denote by $\mathbf{e}_i$ the edge connecting the vertices $\mathbf{v}_i$ and $\mathbf{v}_{i+1}$. Then, we set

$$\xi_i := \eta_{i-1}\frac{\lambda_{i-2}}{\lambda_{i-2}(\mathbf{v}_i)} + \eta_i\frac{\lambda_{i+1}}{\lambda_{i+1}(\mathbf{v}_i)} \qquad \text{and} \qquad \eta_i := \Pi_{\substack{j=1\\ j\neq i}}^4 \frac{\lambda_j}{\lambda_j + \lambda_i},$$

where λ_i is the linear function that vanishes on the edge $\mathbf{e}_i$ and reaches the maximum value 1 in the closure of K. For details, see [14, 10].

2.3. The stabilization subspace $\boldsymbol{M_S(\partial K)}$. We also need to introduce the stabilization space $M_S(\partial K)$. This is a subspace of $M(\partial K)$ satisfying the following two conditions inspired from [14, Proposition 3.2]:

$$\dim M_S(\partial K) = \dim \widetilde{W}^{\perp}(K) = \dim W(K) - \dim \nabla\cdot\boldsymbol{V}(K), \tag{2.1a}$$

$$\|P_{M_S}(\cdot)\|_{\partial K} \text{ is a norm on the space } \widetilde{W}^{\perp}(K). \tag{2.1b}$$

Here, P_{M_S} denotes the $L^2(\partial K)$-projection into the space $M_S(\partial K)$. Examples of $M_S(\partial K)$ for various element shapes are collected in the following proposition, whose proof is given in section 3.

PROPOSITION 2.2. *Let the space* $\boldsymbol{V}(K)\times W(K)$ *admit an* $M(\partial K)$*-decomposition. Then, conditions* (2.1) *are satisfied*

(1) *if* $\nabla\cdot\boldsymbol{V}(K)=W(K)$ *and* $M_S(\partial K)=\emptyset$*;*

(2) *if* $\nabla\cdot\boldsymbol{V}(K)=\mathcal{P}_{k-1}(K), W(K)=\mathcal{P}_k(K)$ *and*

$$M_S(\partial K):=\{\widehat{w}\in L^2(\partial K):\ \ \widehat{w}|_{F^*}\in\mathcal{P}_k(F^*),\ \ \widehat{w}|_{\partial K\setminus F^*}=0\};$$

here F^* *is a fixed face of the element* K *such that* K *lies in one side of the hyperplane containing* F^**;*

(3) *if* K *is a square or cube,* $\nabla\cdot\boldsymbol{V}(K)=\nabla\cdot\mathcal{Q}_k(K)^d, W(K)=\mathcal{Q}_k(K)$ *and*

$$M_S(\partial K):=\{\widehat{w}\in L^2(\partial K):\ \ \widehat{w}|_{F^*}\in\widetilde{\mathcal{Q}}_k(F^*),\ \ \widehat{w}|_{\partial K\setminus F^*}=0\};$$

here F^* *is any fixed face of the square or cubic element* K*;*

(4) *if* K *is a prism with tensor product structure,* $\nabla\cdot\boldsymbol{V}(K)=\nabla\cdot\mathcal{P}_{k|k}(K)^d$*,* $W(K)=\mathcal{P}_{k|k}(K)$*, and*

$$M_S(\partial K):=\{\widehat{w}\in M:\ \ \widehat{w}|_{F^*}\in\widetilde{\mathcal{P}}_k(F^*);\ \ \widehat{w}|_{\partial K\setminus F^*}=0\};$$

here F^* *is a triangular base of the prism* K*.*

2.4. Discrete H^1- and Poincaré–Friedrichs inequalities.

Our main result is the following.

THEOREM 2.3 (local, discrete H^1- and Poincaré–Friedrichs inequalities). - - - - *Let* K *be any element of the mesh* $\mathcal{T}_h$*. Consider the mapping* $(u_h,\widehat{u}_h)\in W(K)\times M(\partial K)\longmapsto\ \boldsymbol{q}_h\in\boldsymbol{V}(K)$ *given by* (1.2)*. Then, if* $\boldsymbol{V}(K)\times W(K)$ *admits an* $M(\partial K)$*-decomposition, and*

$$\Theta_K:=\left(\lambda_{\rm c}^{\max}(K)\,\|\boldsymbol{q}_h\|_{{\rm c},K}^2+h_K^{-1}\|P_{M_S}(u_h-\widehat{u}_h)\|_{\partial K}^2\right),$$

where $M_S(\partial K)$ *is any subspace of* $M(\partial K)$ *satisfying conditions* (2.1)*, we have the inequalities*

$$\begin{aligned} |\,(u_h,\widehat{u}_h)\,|_{1,K}^2&\le C\,\Theta_K\quad(H^1),\\ h_K^{-2}\,|\,(u_h,\widehat{u}_h)\,|_{{\rm PF},K}^2&\le C\,\Theta_K\quad(\text{Poincaré–Friedrichs}),\end{aligned}$$

where the constant C *only depends on the finite element spaces* $\boldsymbol{V}(K)$*,* $W(K)$*, and* $M_S(\partial K)$*, and on the shape-regularity properties of the element* K*.*

A detailed proof of this result is given in the next section. Here, let us briefly discuss it:

(1) First, note that it is not very difficult to obtain these inequalities if the projection operator P_{M_S} is replaced by the identity. Indeed, if we *only* assume that $\nabla W(K)\subset\boldsymbol{V}(K)$, we can take $\boldsymbol{v}:=\nabla u_h$ in the equation defining $\boldsymbol{q}_h$, (1.2c), to immediately obtain

$$\|\nabla u_h\|_K^2\le\|{\rm c}\,\boldsymbol{q}_h\|_K^2+C\,h_K^{-1}\,\|u_h-\widehat{u}_h\|_{\partial K}^2.$$

The wanted inequality now easily follows. However, such a choice might degrade the accuracy of the HDG method, as is typical of DG methods; see, for example, [3]. To avoid this, we must chose a *minimal* space M_S such that the inequalities in Theorem 2.3 still hold.

(2) The inequalities of the above result are nothing but *stabilized* versions of inf-sup conditions for the bilinear form defining $\boldsymbol{q}_h$ (see (1.2)), since

$$\|\mathrm{c}\,\boldsymbol{q}_h\|_K \geq \sup_{\boldsymbol{v}\in\boldsymbol{V}(K)\setminus\{\boldsymbol{0}\}} \frac{(u_h, \nabla\cdot\boldsymbol{v})_K - \langle \widehat{u}_h, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial K}}{\|\boldsymbol{v}\|_K};$$

see [1, section 6.3]. For this reason, the subspace $M_S(\partial K)$ is called a *stabilization* subspace.

(3) Let us argue that the dimension of the stabilization space $M_S(\partial K)$ is actually minimal. It is obvious that the influence of u_h on $\boldsymbol{q}_h$ is only through its L^2-projection into $\nabla\cdot\boldsymbol{V}(K)$. As a consequence, the part of u_h lying on the $L^2(K)$-orthogonal complement of $\nabla\cdot\boldsymbol{V}(K)$ in $W(K)$ *cannot* be controled by the size of $\boldsymbol{q}_h$. Since the dimension of such space is $\dim W(K) - \dim \nabla\cdot\boldsymbol{V}(K)$ and this number, by the first of conditions (2.1), is equal to $\dim M_S(K)$, we see that the dimension of $M_S(\partial K)$ cannot be smaller for the inequalities under consideration to hold.

(4) The above H^1-inequality has been explicitly obtained in the literature for two cases [19, 7]. The first [19] is the case of the Raviart–Thomas elements on a simplex in which the spaces, using our notation,

$$\begin{gathered}\boldsymbol{V}(K) = \mathcal{P}_k(K)^d + \boldsymbol{x}\,\mathcal{P}_k(K), \quad W(K) := \mathcal{P}_k(K),\\ M(\partial K) := \{\mu\in L^2(\partial K):\ \mu|_F \in \mathcal{P}_k(F)\ \forall\, F\in\mathcal{F}(K)\}, M_S(\partial K) = \emptyset,\end{gathered}$$

see [19, Proposition 3.2]; the second [7] is the case for the staggered DG method in which the spaces (defined on a simplex) are given as follows:

$$\begin{aligned}\boldsymbol{V}(K) = \mathcal{P}_k(K)^d, &\quad W(K) := \mathcal{P}_k(K),\\ M(\partial K) &:= \{\mu\in L^2(\partial K):\ \mu|_F \in \mathcal{P}_k(F)\ \forall\, F\in\mathcal{F}(K)\},\\ M_S(\partial K) &:= \{\mu\in M(\partial K) : \mu = 0 \text{ on } \partial K\setminus F_K\},\end{aligned}$$

where F_K is a single face of the simplex K; see [7, Theorem 3.2].

(5) Given data $\widehat{u}_h$ and f, let $(\boldsymbol{q}_h, u_h)\in\boldsymbol{V}(K)\times W(K)$ be the solution to the local problem (1.1a)–(1.1b) with the space $\boldsymbol{V}(K)\times W(K)$ admitting an $M(\partial K)$-decomposition. The following inequalities were obtained in [14, Theorem 4.3]:

$$\begin{aligned}\|\nabla u_h\|_K^2 &\leq C\,\left(\lambda_{\mathrm{c}}^{\max}(K)\,\|\boldsymbol{q}_h\|_{\mathrm{c},K}^2 + \|P_{\widetilde{W}^\perp} f\|_K^2\right),\\ h_K^{-1}\|u_h - \widehat{u}_h\|_{\partial K}^2 &\leq C\,\left(\lambda_{\mathrm{c}}^{\max}(K)\,\|\boldsymbol{q}_h\|_{\mathrm{c},K}^2 + \|P_{\widetilde{W}^\perp} f\|_K^2\right).\end{aligned}$$

Our result replaces the quantity $\|P_{\widetilde{W}^\perp} f\|_K^2$ on the above right-hand side with

$$h_K^{-1}\|P_{M_S}(u_h - \widehat{u}_h)\|_{\partial K}^2.$$

It is this small change that significantly facilitates the analysis of HDG schemes for the incompressible Navier–Stokes equation considered in this paper.

(6) The dependence of the constant C in the estimates on the local spaces $\boldsymbol{V}(K), W(K)$, and $M_S(\partial K)$, and on the shape regularity of the element K, remains to be studied. It is reasonable to believe that C can be uniformly bounded by a function of the maximum degree of the polynomial functions belonging to the local spaces and by a suitable measure of the element shape-regularity.

2.5. Choosing the stabilization function $\boldsymbol{\alpha}$ to get $\boldsymbol{H^1}$-stability. We end this section by illustrating the fact that the stabilization subspace $M_S(\partial K)$ can be actually used, when defining HDG methods, to obtain what we could call the *minimal* stabilization function α needed to achieve a new H^1-stability result. Let us do that in the framework of HDG approximations for steady-state diffusion problems.

So, if $(\boldsymbol{q}_h, u_h) \in \boldsymbol{V}(K) \times W(K)$ is the solution of the local problem (1.1a)–(1.1b), we have the discrete energy identity

$$\mathsf{E}_K(\boldsymbol{q}_h; u_h, \widehat{u}_h) = (f, u_h)_K - \langle \widehat{\boldsymbol{q}}_h \cdot \boldsymbol{n}, \widehat{u}_h \rangle_{\partial K},$$

where

$$\mathsf{E}_K(\boldsymbol{q}_h; u_h, \widehat{u}_h) := (\mathrm{c}\, \boldsymbol{q}_h, \boldsymbol{q}_h)_K + \langle \alpha(u_h - \widehat{u}_h), u_h - \widehat{u}_h \rangle_{\partial K},$$

is the *energy* associated to the element K. We immediately see that

$$\|\mathrm{c}\, \boldsymbol{q}_h\|_K^2 + h_K^{-1} \|P_{M_S}(u_h - \widehat{u}_h)\|_{\partial K}^2 \le C\, \mathsf{E}_K(\boldsymbol{q}_h; u_h, \widehat{u}_h)$$

if we pick the stabilization function α as

$$\alpha(\widehat{\omega}) := h_K^{-1} P_{M_S}(\widehat{\omega}) \ \ \forall\, \widehat{\omega} \in L^2(\partial K), \tag{2.2}$$

the case in which we say that this stabilization function α is *minimal*. Thus, by establishing this link between the HDG stabilization function α and the stabilization subspace $M_S(\partial K)$, an estimate of the energy immediately implies an estimate on the discrete seminorms under consideration, that is,

$$\max\{h_K^{-2} |(u_h, \widehat{u}_h)|_{\mathrm{PF},K}^2, |(u_h, \widehat{u}_h)|_{1,K}^2\} \le C\, \mathsf{E}_K(\boldsymbol{q}_h; u_h, \widehat{u}_h).$$

Now considering the full HDG scheme (1.1) for diffusion, we easily obtain a discrete H^1-stability result of the approximation with respect to the data f by summing the above inequality over all elements:

$$|\!|\!|(u_h, \widehat{u}_h)|\!|\!|_{1,\mathcal{T}_h}^2 = \sum_{K \in \mathcal{T}_h} |(u_h, \widehat{u}_h)|_{1,K}^2 \le C \sum_{K \in \mathcal{T}_h} \mathsf{E}_K(\boldsymbol{q}_h; u_h, \widehat{u}_h) = C\,(f, u_h)_{\mathcal{T}_h}.$$

This stability result can be similarly obtained for the HDG method for the convection-diffusion equation in which convection is treated with the standard *upwinding* technique. We use this approach in section 4 to deal with the HDG and mixed methods for the Navier–Stokes equations.

3. Proofs of the results of section 2.

In this section, we give a proof of the properties of the stabilization spaces $M_S(\partial K)$, and then a proof of the discrete H^1- and the discrete Poincaré–Friedrichs inequalities.

3.1. Proof of Proposition 2.2. Let us first prove Proposition 2.2 on the properties of the stabilization spaces $M_S(\partial K)$. We just prove the second case since the proofs for the other three are similar and simpler.

For this case, we have $\nabla\cdot \boldsymbol{V} = \mathcal{P}_{k-1}(K)$, $W = \mathcal{P}_k(K)$, and

$$M_S = \{\widehat{w} \in L^2(\partial K) : \ \widehat{w}|_{F^*} \in \mathcal{P}_k(F^*), \widehat{w}|_{\partial K \setminus F^*} = 0\},$$

where F^* is a face of the element K such that K lies on one side of the hyperplane containing F^*. Hence, we have

$$\begin{aligned} \dim M_S = \dim \mathcal{P}_k(F^*) &= \dim \mathcal{P}_k(K) - \dim \mathcal{P}_{k-1}(K) \\ &= \dim W - \dim \nabla\cdot \boldsymbol{V} = \dim W - \dim \widetilde{W} = \dim \gamma(\widetilde{W}^{\perp}). \end{aligned}$$

This proves the first condition for M_S.

To prove the second condition, we only need to show that for any function $\widehat{w} \in \gamma(\widetilde{W}^\perp)$, $P_{M_S}(\widehat{w}) = 0$ implies $\widehat{w} = 0$. Now, let $\widehat{w}$ be a function in $\gamma(\widetilde{W}^\perp)$ such that $P_{M_S}(\widehat{w}) = 0$. By the definition of $\gamma(\widetilde{W}^\perp)$, there exists a function $w \in \widetilde{W}^\perp$ such that $\gamma(w) = \widehat{w}$. Hence, $P_{M_S}(\gamma(w)) = 0$. By the definition of M_S and W, we have $w = \lambda\widetilde{w}$, where $\lambda \in \mathcal{P}_1(K)$ is the linear function vanishing on F^* and $\widetilde{w} \in \mathcal{P}_{k-1}(K) = \widetilde{W}$. By L^2-orthogonality of the spaces $\widetilde{W}$ and $\widetilde{W}^\perp$, we have

$$(w, \widetilde{w})_K = (\lambda\widetilde{w}, \widetilde{w})_K = 0,$$

which immediately implies $w = 0$ by the assumption on the face F^*. This completes the proof of Proposition 2.2.

3.2. Proof of Theorem 2.3.

Here, we prove the inequalities of Theorem 2.3. Although it is enough to prove only one since the seminorms $|(\cdot,\cdot)|_{1,K}$ and $|(\cdot,\cdot)|_{\mathrm{PF},K}$ are equivalent, we provide a different proof for each of them, as they put in evidence different properties of the M-decompositions.

3.2.1. Proof of the first inequality.

To prove the first inequality, it is convenient to first carry out a simple integration-by-parts in the equation defining $\boldsymbol{q}_h$, (1.2c):

$$(\mathrm{c}\,\boldsymbol{q}_h, \boldsymbol{v})_K = -(\nabla u_h, \boldsymbol{v})_K + \langle u_h - \widehat{u}_h\,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial K} \quad \forall\, \boldsymbol{v} \in \boldsymbol{V}(K).$$

By property (b) of an M-decomposition, we can now set $\boldsymbol{v} := \nabla u_h$ to get

$$\|\nabla u_h\|_K^2 = -\,(\mathrm{c}\,\boldsymbol{q}_h, \boldsymbol{v})_K + \langle u_h - \widehat{u}_h\,,\, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial K},$$

and conclude that

$$\|\nabla u_h\|_K \le (\lambda_{\mathrm{c}}^{\max})^{1/2}\|\boldsymbol{q}_h\|_{\mathrm{c},K} + C_{\nabla W}\, h_K^{-1/2}\,\|u_h - \widehat{u}_h\|_{\partial K},$$

$$C_{\nabla W} := \sup_{\boldsymbol{v}\in\nabla W(K)\setminus\{0\}} \frac{h_K^{1/2}\|\boldsymbol{v}\cdot\boldsymbol{n}\|_{\partial K}}{\|\boldsymbol{v}\|_K}.$$

Let us now estimate the jump $u_h - \widehat{u}_h \in M(\partial K)$. By property (c) of an M-decomposition, we can write that $u_h - \widehat{u}_h = P_{\gamma\widetilde{W}^\perp}(u_h - \widehat{u}_h) + P_{\gamma\widetilde{\boldsymbol{V}}^\perp}(u_h - \widehat{u}_h)$. Now, by the second of conditions (2.1), there is a constant C_{M_S} such that

$$\begin{aligned}\|P_{\gamma\widetilde{W}^\perp}(u_h - \widehat{u}_h)\|_{\partial K} &\le C_{M_S}\|P_{M_S}\left(P_{\gamma\widetilde{W}^\perp}(u_h - \widehat{u}_h)\right)\|_{\partial K}\\ &\le C_{M_S}\big(\|P_{M_S}\,(u_h - \widehat{u}_h)\|_{\partial K} + \|P_{M_S}\left(P_{\gamma\widetilde{\boldsymbol{V}}^\perp}(u_h - \widehat{u}_h)\right)\|_{\partial K}\big)\\ &\le C_{M_S}\big(\|P_{M_S}\,(u_h - \widehat{u}_h)\|_{\partial K} + \|P_{\gamma\widetilde{\boldsymbol{V}}^\perp}(u_h - \widehat{u}_h)\|_{\partial K}\big).\end{aligned}$$

It remains to estimate $\|P_{\gamma\widetilde{\boldsymbol{V}}^\perp}(u_h - \widehat{u}_h)\|_{\partial K}$. Taking $\boldsymbol{v} \in \widetilde{\boldsymbol{V}}^\perp(K)$ such that $\boldsymbol{v}\cdot\boldsymbol{n}|_{\partial K} = P_{\gamma\widetilde{\boldsymbol{V}}^\perp}(u_h - \widehat{u}_h)$ in the definition of $\boldsymbol{q}_h$, and using the fact that $\nabla u_h \in \widetilde{\boldsymbol{V}}(K)$ is L^2-orthogonal to $\boldsymbol{v} \in \widetilde{\boldsymbol{V}}^\perp(K)$, we get

$$\|P_{\gamma\widetilde{\boldsymbol{V}}^\perp}(u_h - \widehat{u}_h)\|_{\partial K}^2 = (\mathrm{c}\,\boldsymbol{q}_h, \boldsymbol{v})_K,$$

and conclude that

$$\|P_{\gamma\widetilde{\boldsymbol{V}}^\perp}(u_h - \widehat{u}_h)\|_{\partial K} \le C_{\widetilde{\boldsymbol{V}}^\perp}(\lambda_{\mathrm{c}}^{\max})^{1/2}\, h_K^{1/2}\,\|\boldsymbol{q}_h\|_{\mathrm{c},K},$$

$$C_{\widetilde{\boldsymbol{V}}^\perp} := \sup_{\boldsymbol{v}\in\widetilde{\boldsymbol{V}}^\perp(K)\setminus\{0\}} \frac{\|\boldsymbol{v}\|_K}{h_K^{1/2}\|\boldsymbol{v}\cdot\boldsymbol{n}\|_{\partial K}}.$$

The first inequality now easily follows.

3.2.2. Proof of the second inequality. To prove this inequality, it is convenient to rewrite the equation defining $\boldsymbol{q}_h$, (1.2c), as follows:

$$(\mathrm{c}\,\boldsymbol{q}_h, \boldsymbol{v})_K - (u_h - \overline{\widehat{u}_h}^{\partial K}, \nabla\cdot\boldsymbol{v})_K + \langle \widehat{u}_h - \overline{\widehat{u}_h}^{\partial K}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial K} = 0 \quad \forall\, \boldsymbol{v}\in\boldsymbol{V}(K).$$

By [14, Theorem 2.4], since $\boldsymbol{V}(K)\times W(K)$ admits an $M(\partial K)$ decomposition, we have the identity

$$\{\mu\in M(\partial K):\ \langle\mu,1\rangle_{\partial K}=0\} = \{\boldsymbol{v}\cdot\boldsymbol{n}|_{\partial K}:\ \boldsymbol{v}\in\boldsymbol{V}(K),\ \nabla\cdot\boldsymbol{v}=0\}. \tag{3.1}$$

This means that there is a function $\boldsymbol{v}\in\boldsymbol{V}(K)$ such that $\boldsymbol{v}\cdot\boldsymbol{n}|_{\partial K} = \widehat{u}_h - \overline{\widehat{u}_h}^{\partial K}$ and $\nabla\cdot\boldsymbol{v}=0$. Using this function as test function, we get

$$\|\widehat{u}_h - \overline{\widehat{u}_h}^{\partial K}\|^2_{\partial K} = -(\mathrm{c}\,\boldsymbol{q}_h, \boldsymbol{v})_K,$$

and so

$$\begin{aligned}
&\|\widehat{u}_h - \overline{\widehat{u}_h}^{\partial K}\|_{\partial K} \le (\lambda_{\mathrm{c}}^{\max}(K))^{1/2}\|\boldsymbol{q}_h\|_{c,K} C_{\boldsymbol{V}\cdot\boldsymbol{n}}\, h_K^{1/2},\\
&C_{\boldsymbol{V}\cdot\boldsymbol{n}} := \sup_{\substack{\mu\,\in\, M(\partial K)\\ \langle\mu,1\rangle_{\partial K}\,=\,0}} \ \inf_{\substack{\boldsymbol{v}\,\in\,\boldsymbol{V}(K)\setminus\{0\}\\ \nabla\cdot\boldsymbol{v}\,=\,0\\ \boldsymbol{v}\cdot\boldsymbol{n}\,=\,\mu}} \frac{\|\boldsymbol{v}\|_K}{h_K^{1/2}\|\boldsymbol{v}\cdot\boldsymbol{n}\|_{\partial K}}.
\end{aligned}$$

It remains to estimate $\|u_h - \overline{\widehat{u}_h}^{\partial K}\|_K$. We define a test function $\boldsymbol{v}\in\boldsymbol{V}(K)$ such that $\nabla\cdot\boldsymbol{v} = P_{\nabla\cdot\boldsymbol{V}}(u_h - \overline{\widehat{u}_h}^{\partial K})$, which we can assume to be different from zero. Obviously, we get

$$\|P_{\nabla\cdot\boldsymbol{V}}(u_h - \overline{\widehat{u}_h}^{\partial K})\|^2_K = (\mathrm{c}\,\boldsymbol{q}_h, \boldsymbol{v})_K + \langle \widehat{u}_h - \overline{\widehat{u}_h}^{\partial K}, \boldsymbol{v}\cdot\boldsymbol{n}\rangle_{\partial K},$$

and so

$$\begin{aligned}
&\|P_{\nabla\cdot\boldsymbol{V}}(u_h - \overline{\widehat{u}_h}^{\partial K})\|_K \le \big(\|\mathrm{c}\boldsymbol{q}_h\|_K + h_K^{-1/2}\,\|\widehat{u}_h - \overline{\widehat{u}_h}^{\partial K}\|_{\partial K}\big)\, C_{\nabla\cdot\boldsymbol{V}}\, h_K,\\
&C_{\nabla\cdot\boldsymbol{V}} := \sup_{g\in\nabla\cdot\boldsymbol{V}(K)\setminus\{0\}} \ \inf_{\substack{\boldsymbol{v}\,\in\,\boldsymbol{V}(K)\\ \nabla\cdot\boldsymbol{v}\,=\,g}} \frac{(\|\boldsymbol{v}\|_K + h_K^{1/2}\|\boldsymbol{v}\cdot\boldsymbol{n}\|_{\partial K})}{h_K\|\nabla\cdot\boldsymbol{v}\|_K}.
\end{aligned}$$

Finally, let us estimate $(\mathrm{Id} - P_{\nabla\cdot\boldsymbol{V}})(u_h - \overline{\widehat{u}_h}^{\partial K})$. Since this function coincides with $P_{\widetilde{W}^\perp}(u_h - \overline{\widehat{u}_h}^{\partial K})$ because $\widetilde{W}(K) = \nabla\cdot\boldsymbol{V}(K)$, we get

$$\begin{aligned}
\|P_{\widetilde{W}^\perp}(u_h - \overline{\widehat{u}_h}^{\partial K})\|_K &\le C_K\, h_K^{1/2}\,\|P_{\gamma\widetilde{W}^\perp}(u_h - \overline{\widehat{u}_h}^{\partial K})\|_{\partial K}\\
&\le C_M\, C_K\, h_K^{1/2}\,\|P_{M_S}(u_h - \overline{\widehat{u}_h}^{\partial K})\|_{\partial K}\\
&\le C_M\, C_K\, h_K^{1/2}\,(\|P_{M_S}(u_h - \widehat{u}_h)\|_{\partial K} + \|P_{M_S}(\widehat{u}_h - \overline{\widehat{u}_h}^{\partial K})\|_{\partial K}),\\
&\le C_M\, C_K\, h_K^{1/2}\,(\|P_{M_S}(u_h - \widehat{u}_h)\|_{\partial K} + \|\widehat{u}_h - \overline{\widehat{u}_h}^{\partial K}\|_{\partial K}),
\end{aligned}$$

and the estimate follows. This completes the proof of Theorem 2.3.

4. Application: HDG methods for the Navier–Stokes equations.

In this section, we introduce and analyze new HDG and mixed methods for the steady-state incompressible Navier–Stokes equation with velocity gradient-velocity-pressure formulation described by (1.4).

We proceed as follows. After defining the methods, we show that their approximate solution exists, is unique, and satisfies an energy-boundedness property under a *smallness* assumption on the data. We then provide results on the convergence properties.

Some of the errors involving the velocities are measured in the norms and seminorms defined as follows. For any $(\boldsymbol{v}, \widehat{\boldsymbol{v}}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h$, we set

$$|\!|\!|(\boldsymbol{v}, \widehat{\boldsymbol{v}})|\!|\!|^2_{\ell,\mathcal{T}_h} := \sum_{i=1}^{d} \sum_{K\in\mathcal{T}_h} |(\boldsymbol{v}_i, \widehat{\boldsymbol{v}}_i)|^2_{\ell,K} \qquad \text{for } \ell = 0, 1, \text{PF},$$

where $|(\cdot,\cdot)|_{1,K}$ and $|(\cdot,\cdot)|_{\text{PF},K}$ are defined by (1.3), and

$$|(\boldsymbol{v}_i, \widehat{\boldsymbol{v}}_i)|^2_{0,K} := \|\boldsymbol{v}_i\|^2_K + h_K(\|\widehat{\boldsymbol{v}}_i\|^2_{\partial K} + \|\boldsymbol{v}_i - \widehat{\boldsymbol{v}}_i\|^2_{\partial K}).$$

4.1. Definition of the methods.

4.1.1. The general form of the methods.

The HDG and mixed methods for (1.4) seek an approximation to $(\mathrm{L}, \boldsymbol{u}, p, \boldsymbol{u}|_{\mathcal{F}_h})$, $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h)$, in the space $\mathcal{G}_h \times \boldsymbol{V}_h \times \mathring{Q}_h \times \boldsymbol{M}_h(0)$ given by

$$\begin{aligned}
&(4.1\text{a}) && \mathcal{G}_h := \{\mathrm{G} \in L^2(\mathcal{T}_h)^{d\times d} : \mathrm{G}|_K \in \mathcal{G}(K),\ K \in \mathcal{T}_h\},\\
&(4.1\text{b}) && \boldsymbol{V}_h := \{\boldsymbol{v} \in L^2(\mathcal{T}_h)^d : \quad \boldsymbol{v}|_K \in \boldsymbol{V}(K),\ K \in \mathcal{T}_h\},\\
&(4.1\text{c}) && \mathring{Q}_h := \{q \in L^2(\mathcal{T}_h) : \quad q|_K \in Q(K),\ K \in \mathcal{T}_h, (q,1)_\Omega = 0\},\\
&(4.1\text{d}) && \boldsymbol{M}_h := \{\widehat{\boldsymbol{v}} \in L^2(\mathcal{F}_h)^d : \quad \widehat{\boldsymbol{v}}|_F \in \boldsymbol{M}(F), F \in \mathcal{F}_h\},\\
&(4.1\text{e}) && \boldsymbol{M}_h(0) := \{\widehat{\boldsymbol{v}} \in \boldsymbol{M}_h : \quad \widehat{\boldsymbol{v}}|_{\partial\Omega} = 0\},
\end{aligned}$$

where the local spaces $\mathcal{G}(K), \boldsymbol{V}(K), Q(K)$, and $\boldsymbol{M}(F)$ are suitably defined finite dimensional spaces, and determine it as the only solution of the following weak formulation:

$$\begin{aligned}
&(4.2\text{a}) && (\nu\,\mathrm{L}_h\,,\,\mathrm{G})_{\mathcal{T}_h} + (\boldsymbol{u}_h\,,\,\nu\,\boldsymbol{\nabla\cdot}\mathrm{G})_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}_h\,,\,\nu\,\mathrm{G}\,\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0,\\
& && (\nu\,\mathrm{L}_h\,,\,\boldsymbol{\nabla v})_{\mathcal{T}_h} + \langle -\nu\,\mathrm{L}_h\,\boldsymbol{n} + \alpha_v(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\,,\,\boldsymbol{v} - \widehat{\boldsymbol{v}}\rangle_{\partial\mathcal{T}_h}\\
& && - (p_h\,,\,\nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle p_h\,\boldsymbol{n}\,,\,\boldsymbol{v} - \widehat{\boldsymbol{v}}\rangle_{\partial\mathcal{T}_h}\\
&(4.2\text{b}) && - (\boldsymbol{u}_h \otimes \boldsymbol{\beta}\,,\,\boldsymbol{\nabla v})_{\mathcal{T}_h} + \langle(\boldsymbol{\beta}\cdot\boldsymbol{n})\,\widehat{\boldsymbol{u}}_h + \alpha_c(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\,,\,\boldsymbol{v} - \widehat{\boldsymbol{v}}\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{f}\,,\,\boldsymbol{v})_{\mathcal{T}_h},\\
&(4.2\text{c}) && - (\boldsymbol{u}_h\,,\,\nabla q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}_h\cdot\boldsymbol{n}\,,\,q\rangle_{\partial\mathcal{T}_h} = 0
\end{aligned}$$

for all $(\mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}}) \in \mathcal{G}_h \times \boldsymbol{V}_h \times \mathring{Q}_h \times \boldsymbol{M}_h(0)$, where

$$\alpha_v : L^2(\partial K)^d \longrightarrow L^2(\partial K)^d \quad \text{and} \quad \alpha_c : L^2(\partial K)^d \longrightarrow L^2(\partial K)^d$$

are the *local stabilization operators* related to the viscous and convective parts, respectively. To complete the definition of the method, we have to define the local spaces, the divergence-free post-processed velocity $\boldsymbol{\beta}$, and the stabilization operators. We do this next.

4.1.2. The local spaces. The finite element spaces are the ones used in [13] for Stokes flow. Let the space $\boldsymbol{V}^{\mathrm{D}}(K)\times W^{\mathrm{D}}(K)\times M^{\mathrm{D}}(\partial K)$ be such that $\boldsymbol{V}^{\mathrm{D}}(K)\times W^{\mathrm{D}}(K)$ admits an $M^{\mathrm{D}}(\partial K)$-decomposition; see Definition 2.1. Moreover, we assume that

$$W^{\mathrm{D}}(K) \text{ is a polynomial space such that } \sum_{i=1}^{d} \partial_i W^{\mathrm{D}}(K) \subset W^{\mathrm{D}}(K). \tag{4.3}$$

Then, the local spaces $\mathcal{G}(K)$, $\boldsymbol{V}(K)$, and $Q(K)$ and the local trace space $\boldsymbol{M}(\partial K)$ are defined as follows:

$$\mathcal{G}_i(K)\times \boldsymbol{V}_i(K)\times \boldsymbol{M}_i(K) := \boldsymbol{V}^{\mathrm{D}}(K)\times W^{\mathrm{D}}(K)\times M^{\mathrm{D}}(\partial K) \quad i=1,\dots,d, \tag{4.4a}$$

$$Q(K) := W^{\mathrm{D}}(K). \tag{4.4b}$$

4.1.3. The postprocessed velocity $\boldsymbol{\beta}$. On the element K, the postprocessed velocity $\boldsymbol{\beta}$ is taken in a finite dimensional space $\boldsymbol{V}^*(K)$ satisfying the conditions

$$\boldsymbol{V}^{\mathrm{D}}(K) \subset \boldsymbol{V}^*(K), \nabla\cdot \boldsymbol{V}^*(K) = W^{\mathrm{D}}(K), \tag{4.5a}$$

$$\boldsymbol{V}^*(K)\times W^{\mathrm{D}}(K) \text{ admits an } M^{\mathrm{D}}(\partial K)\text{-decomposition.} \tag{4.5b}$$

This vector-valued space can be easily constructed from $\boldsymbol{V}^{\mathrm{D}}(K)$, as shown in [14, Proposition 5.3].

On the element K, the postprocessed velocity $\boldsymbol{\beta} := \boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) \in \boldsymbol{V}_h^*$ is defined as the function in $\boldsymbol{V}^*(K)$ such that

$$(\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), \boldsymbol{v})_K = (\boldsymbol{u}_h, \boldsymbol{v})_K \qquad \forall\, \boldsymbol{v} \in \widetilde{\boldsymbol{V}^*}(K), \tag{4.6a}$$

$$\langle \boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)\cdot \boldsymbol{n}\,,\, \widehat{v}\rangle_{\partial K} = \langle \widehat{\boldsymbol{u}}_h\cdot \boldsymbol{n}\,,\, \widehat{v}\rangle_{\partial K} \quad \forall\, \widehat{v} \in M^{\mathrm{D}}(\partial K). \tag{4.6b}$$

Here $\widetilde{\boldsymbol{V}^*}(K) := \nabla W^{\mathrm{D}}(K) \oplus \{\boldsymbol{v}\in \boldsymbol{V}^*(K) :\ \nabla\cdot\boldsymbol{v} = 0,\ \boldsymbol{v}\cdot\boldsymbol{n}|_{\partial K} = 0\}$.

We gather the main properties of this mapping in the next result, which we prove in Appendix A.

PROPOSITION 4.1. *Let $(\boldsymbol{v}, \widehat{\boldsymbol{v}}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h$. Then, for any element $K\in \mathcal{T}_h$, we have*

$$\begin{aligned} \lVert\!| (\boldsymbol{P}_h(\boldsymbol{v},\widehat{\boldsymbol{v}}),\, \{\boldsymbol{P}_h(\boldsymbol{v},\widehat{\boldsymbol{v}})\}) |\!\rVert_{\ell,K} &\le C\, \lVert\!| (\boldsymbol{v},\widehat{\boldsymbol{v}}) |\!\rVert_{\ell,K} \qquad \text{for } \ell = 0,1,\\ \lVert \boldsymbol{P}_h(\boldsymbol{v},\widehat{\boldsymbol{v}}) \rVert_{\infty,K} &\le C\, \lVert\!| (\boldsymbol{v},\widehat{\boldsymbol{v}}) |\!\rVert_{\infty,K} \end{aligned}$$

with a constant C depending only on the space $\boldsymbol{V}(K)\times \boldsymbol{M}(\partial K)$ and the shape regularity of the element K. Moreover, if $(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) \in \boldsymbol{V}_h \times \boldsymbol{M}_h(0)$ satisfies the weak incompressibility condition given by (4.2c)*, then*

$$\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) \in H(\mathrm{div},\Omega) \text{ and } \nabla\cdot \boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) = 0.$$

4.1.4. The stabilization operators. For the convective stabilization operator, we take the choice leading to the classic upwinding:

$$\alpha_c(\widehat{\boldsymbol{v}}) := \max\{\boldsymbol{\beta}\cdot\boldsymbol{n}, 0\}\,\widehat{\boldsymbol{v}} \quad \forall\, \widehat{\boldsymbol{v}} \in L^2(\partial K)^d, \tag{4.7a}$$

where $\boldsymbol{\beta} = \boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)$ is given in (4.6). For the viscous stabilization operator, we take

$$\alpha_v(\widehat{\boldsymbol{v}}) := \frac{\nu}{h_K} P_{\boldsymbol{M_S}}(\widehat{\boldsymbol{v}}) \quad \forall\, \widehat{\boldsymbol{v}} \in L^2(\partial K)^d, \tag{4.7b}$$

where $P_{\boldsymbol{M_S}}$ is the projection onto the space $\boldsymbol{M_S}(\partial K)$, whose ith component is taken to be $M_S^{\mathrm{D}}(\partial K)$.

4.2. Existence, uniqueness, and boundedness. Now that we have completed the definition of the methods, we must ask ourselves if the approximate solutions actually exist and are unique. The next result shows that this is the case under a standard *smallness* condition on the data.

THEOREM 4.2 (existence, uniqueness, and boundedness). *If $\nu^{-2}\|\boldsymbol{f}\|_\Omega$ is small enough, then the HDG method (4.2) has a unique solution. Furthermore, for the component $(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) \in \boldsymbol{V}_h \times \boldsymbol{M}_h(0)$ of the approximate solution the following stability bound is satisfied:*

$$|\!|\!|(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)|\!|\!|_{1,\mathcal{T}_h} \le C\nu^{-1}\,\|\boldsymbol{f}\|_\Omega$$

for a constant C that depends only on the finite element spaces, the shape-regularity of the mesh, and the domain.

4.3. Convergence properties. Having shown that the approximate solutions are well defined, we next measure how well they approximate the exact solution by comparing them with suitably chosen projections of the exact solution.

4.3.1. Projections of the errors. Let us define the projections we are going to use in our a priori error analysis. We denote $P_{\mathcal{G}}$, $P_{\boldsymbol{V}}$, P_Q, $P_{\boldsymbol{M}}$ to be the L^2-projections onto $\mathcal{G}_h$, $\boldsymbol{V}_h$, $\mathring{Q}_h$, and $\boldsymbol{M}_h$. We also define the projection $\Pi_{\boldsymbol{V}}$ into the space $\boldsymbol{V}_h$ as follows. On the element K, $\Pi_{\boldsymbol{V}}\boldsymbol{u} \in \boldsymbol{V}(K)$ is defined as follows:

$$(\Pi_{\boldsymbol{V}}\boldsymbol{u}, \boldsymbol{v})_K = (\boldsymbol{u}, \boldsymbol{v})_K \quad \forall\, w \in \boldsymbol{\nabla}\cdot\mathcal{G}, \tag{4.8a}$$

$$\langle \Pi_{\boldsymbol{V}}\boldsymbol{u}\,,\, \widehat{\boldsymbol{v}}\rangle_{\partial K} = \langle \boldsymbol{u}\,,\, \widehat{\boldsymbol{v}}\rangle_{\partial K} \quad \forall\, \widehat{\boldsymbol{v}} \in \boldsymbol{M}_{\boldsymbol{S}}. \tag{4.8b}$$

Our strategy is to first estimate the size of the projection of the errors

$$\mathrm{e}_L = P_{\mathcal{G}}\mathrm{L} - \mathrm{L}_h,\ \boldsymbol{e}_u = \Pi_{\boldsymbol{V}}\boldsymbol{u} - \boldsymbol{u}_h,\ e_p = P_Q p - p_h,\ \boldsymbol{e}_{\widehat{u}} = P_{\boldsymbol{M}}\boldsymbol{u} - \widehat{\boldsymbol{u}}_h,$$

and then use the triangle inequality to estimate the size of the actual errors. To do that, we need to use the well-known approximation properties of the various L^2-projections. We also need the approximation properties of the projection $\Pi_{\boldsymbol{V}}$, which we show depend on the L^2-projection $P_{\boldsymbol{V}}$. The following result, proven in Appendix B, is a direct consequence of the assumption on the stabilization space $\boldsymbol{M}_{\boldsymbol{S}}$.

PROPOSITION 4.3. *For the projection $\Pi_{\boldsymbol{V}}\boldsymbol{u} \in \boldsymbol{V}(K)$ defined above, we have*

$$\begin{aligned}\|\Pi_{\boldsymbol{V}}\boldsymbol{u} - \boldsymbol{u}\|_K &\le C\,\left(\|P_{\boldsymbol{V}}\boldsymbol{u} - \boldsymbol{u}\|_K + h_K^{1/2}\|P_{\boldsymbol{V}}\boldsymbol{u} - \boldsymbol{u}\|_{\partial K}\right)\\ \|\Pi_{\boldsymbol{V}}\boldsymbol{u}\|_{\infty,K} &\le C\,\|\boldsymbol{u}\|_{\infty,K},\end{aligned}$$

where the constant C only depends on the spaces $\boldsymbol{V}(K)$ and $\boldsymbol{M}_{\boldsymbol{S}}(K)$.

4.3.2. A priori error estimates. Next, we state our main convergence result.

THEOREM 4.4. *Let $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h) \in \mathcal{G}_h \times \boldsymbol{V}_h \times \mathring{Q}_h \times \boldsymbol{M}_h(0)$ be the numerical solution of (4.2). Assume that*

$$\begin{aligned}\mathcal{P}_k(K)^{d\times d} \times \mathcal{P}_k(K)^d \times \mathcal{P}_k(K) &\subset \mathcal{G}(K) \times \boldsymbol{V}(K) \times Q(K) &&\forall\, K \in \mathcal{T}_h,\\ \mathcal{P}_k(F)^d &\subset \boldsymbol{M}(F) &&\forall\, F \in \mathcal{F}_h.\end{aligned}$$

Then, for $\nu^{-2}\|\boldsymbol{f}\|_\Omega$ and $\nu^{-1}\|\boldsymbol{u}\|_{\infty,\Omega}$ sufficiently small, we have

$$\|\mathrm{e}_L\|_{\mathcal{T}_h} + \|e^p\|_{\mathcal{T}_h} + |\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{1,\mathcal{T}_h} + h^{-1}\,|\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{\mathrm{PF},\mathcal{T}_h} + \|e_u\|_{\mathcal{T}_h} \le C\,h^{k+1}\,\Xi, \tag{4.9}$$

where $\Xi := \|\mathrm{L}\|_{k+1} + \nu^{-1}\|\boldsymbol{\beta}\|_{\infty,\Omega}\|\boldsymbol{u}\|_{k+1} + \nu^{-1}\|p\|_{k+1}$ *and the constant* C *only depends on the finite element spaces, the shape-regularity of the mesh, and the domain* Ω*.*

Moreover, if $\nu^{-1}\|\nabla \boldsymbol{u}\|_\Omega$ *is small enough,* $\boldsymbol{u} \in \boldsymbol{W}^{1,\infty}(\Omega)$*, and the regularity estimate in* [4, equation (2.3)] *holds, then*

$$\|e_u\|_\Omega \leq C\, h^{k+2} \quad \forall k \geq 1. \tag{4.10}$$

Finally, if $\boldsymbol{u}_h^* \in H(\mathrm{div}, \Omega)$ *is the postprocessed approximate velocity introduced in* [13, equation (2.9)]*, then we have* $\nabla \cdot \boldsymbol{u}_h^* = 0$ *in* Ω*, and*

$$\|\boldsymbol{u}_h^* - \boldsymbol{u}\|_\Omega \leq C\, h^{k+2} \quad \forall k \geq 1. \tag{4.11}$$

Note that this result gives optimal convergence of the velocity gradient L_h, the velocity $\boldsymbol{u}_h$, and the pressure p_h approximations. It also gives two superconvergence results. The first is the one of the projections of the error in the velocity, which are of order $k+1$ for $|\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{1,\mathcal{T}_h}$ and of order $k+2$ for $|\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{\mathrm{PF},\mathcal{T}_h}$. The second is also for the projection of the error in the velocity. The only difference is that the first superconvergence estimate does not say anything about the convergence properties of the local averages, whereas the second does. Moreover, the second superconvergence result allows the local postprocessing of the velocity $\boldsymbol{u}_h^*$ to be an $\boldsymbol{H}(\mathrm{div})$, globally divergence-free approximation to the velocity converging faster than the original approximation $\boldsymbol{u}_h$.

5. Proofs of the results of section 4. In this section, we prove Theorem 4.2 on the existence, uniqueness, and boundedness of the approximate solution, and the convergence properties of Theorem 4.4.

We would like to emphasize that, due to the existence of the discrete-H^1 stability results in Theorem 2.3, the proofs in this section can be considered as a word-by-word "translation" of the corresponding proofs in [4], where, for the first time, a superconvergent HDG method was analyzed for the incompressible Navier–Stokes equations

To simplify the notation, we write $A \lesssim B$ to indicate that $A \leq C\,B$ with a constant C that only depends on the finite element spaces, the shape-regularity of the mesh, and the domain.

5.1. Preliminaries.

Rewriting the method in a compact form. To facilitate the analysis, we rewrite the formulation of the methods under consideration by using the bilinear form associated to the Stokes system,

$$\begin{aligned} B_h(\mathrm{L}, \boldsymbol{u}, p, \widehat{\boldsymbol{u}};\ \mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}}) := {} & \nu(\mathrm{L}, \mathrm{G})_{\mathcal{T}_h} + \nu(\boldsymbol{u}, \boldsymbol{\nabla}\cdot\mathrm{G})_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{u}}\,,\, \nu\,\mathrm{G}\,\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} \\ & + (\nu\,\mathrm{L}, \boldsymbol{\nabla}\boldsymbol{v})_{\mathcal{T}_h} + \langle -\nu\,\mathrm{L}\,\boldsymbol{n} + \alpha_v(\boldsymbol{u} - \widehat{\boldsymbol{u}})\,,\, \boldsymbol{v} - \widehat{\boldsymbol{v}}\rangle_{\partial\mathcal{T}_h} \\ & - (p, \nabla\cdot\boldsymbol{v})_{\mathcal{T}_h} + \langle p\,\boldsymbol{n}\,,\, \boldsymbol{v} - \widehat{\boldsymbol{v}}\rangle_{\partial\mathcal{T}_h} \\ & - (\boldsymbol{u}, \nabla q)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{u}}\cdot\boldsymbol{n}\,,\, q\rangle_{\partial\mathcal{T}_h}, \end{aligned} \tag{5.1a}$$

and the bilinear form associated to the convection,

$$\begin{aligned} \mathcal{O}_h(\boldsymbol{\beta}; (u, \widehat{u}), (w, \widehat{w})) := {} & -(\boldsymbol{u}\otimes\boldsymbol{\beta}, \boldsymbol{\nabla}\boldsymbol{v})_{\mathcal{T}_h} \\ & + \langle (\boldsymbol{\beta}\cdot\boldsymbol{n})\,\widehat{\boldsymbol{u}} + \alpha_c(\boldsymbol{u} - \widehat{\boldsymbol{u}})\,,\, \boldsymbol{v} - \widehat{\boldsymbol{v}}\rangle_{\partial\mathcal{T}_h}, \end{aligned} \tag{5.1b}$$

where $(\mathrm{L}, \boldsymbol{u}, p, \widehat{\boldsymbol{u}})$ and $(\mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}})$ lie in the space $\big(H^1(\mathcal{T}_h)^{d\times d} + \mathcal{G}_h\big) \times H^1(\mathcal{T}_h)^d \times H^1(\mathcal{T}_h) \times L^2(\mathcal{F}_h; 0)^d$, and $\boldsymbol{\beta} \in \boldsymbol{V}_\beta \cap \boldsymbol{V}_h^*$, where

$$\begin{aligned}
\boldsymbol{V}_\beta &:= \{\boldsymbol{v} \in H(\mathrm{div}, \Omega) : \nabla\cdot\boldsymbol{v} = 0, \boldsymbol{v}\cdot\boldsymbol{n}|_{\partial K} \in L^2(\partial K),\ K \in \mathcal{T}_h\},\\
\boldsymbol{V}_h^* &:= \{\boldsymbol{v} \in L^2(\mathcal{T}_h)^d : \quad \boldsymbol{v}|_K \in \boldsymbol{V}^*(K),\ K \in \mathcal{T}_h\}.
\end{aligned}$$

Now, the equations defining the HDG method (4.2) can be recast as

$$B_h(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h;\ \mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}}) + \mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (\boldsymbol{v}, \widehat{\boldsymbol{v}})) = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h} \tag{5.2}$$

with $\boldsymbol{\beta} = \boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)$ defined in (4.6). Consistency of the HDG method (4.2) implies that, for the exact solution $(\mathrm{L}, \boldsymbol{u}, p) \in H^1(\Omega)^{d\times d} \times H^2(\Omega)^d \times H^1(\Omega)$ of (1.4) (assuming H^2-regularity),

$$B_h(\mathrm{L}, \boldsymbol{u}, p, \boldsymbol{u};\ \mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}}) + \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}), (\boldsymbol{v}, \widehat{\boldsymbol{v}})) = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h} \tag{5.3}$$

for all $(\mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}}) \in \mathcal{G}_h \times \boldsymbol{V}_h \times \mathring{Q}_h \times \boldsymbol{M}_h(0)$.

An inequality for the viscous energy. Next, we obtain a key inequality for the viscous energy associated the discrete Stokes operator associated with the HDG method (4.2), namely,

$$\begin{aligned}
\mathsf{E}(\mathrm{L}, \boldsymbol{u}, \widehat{\boldsymbol{u}}) &:= B_h(\mathrm{L}, \boldsymbol{u}, p, \widehat{\boldsymbol{u}};\ \mathrm{L}, \boldsymbol{u}, p, \widehat{\boldsymbol{u}})\\
&= \nu(\mathrm{L}, \mathrm{L})_{\mathcal{T}_h} + \Big\langle \frac{\nu}{h_K} P_{\boldsymbol{M_S}}(\boldsymbol{u} - \widehat{\boldsymbol{u}})\,,\, P_{\boldsymbol{M_S}}(\boldsymbol{u} - \widehat{\boldsymbol{u}}) \Big\rangle_{\partial\mathcal{T}_h}.
\end{aligned} \tag{5.4}$$

Lemma 5.1. *Let* $(\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h) \in \mathcal{G}_h \times \boldsymbol{V}_h \times \mathring{Q}_h \times \boldsymbol{M}_h(0)$ *be the numerical solution of the linear system* (4.2) *with a prescribed velocity* $\boldsymbol{\beta} \in \boldsymbol{V}_\beta$*, then we have*

$$\mathsf{E}(\mathrm{L}_h, \boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) \le (f, u_h)_{\mathcal{T}_h}.$$

Proof. By (5.2) with $(\mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}}) := (\mathrm{L}_h, \boldsymbol{u}_h, p_h, \widehat{\boldsymbol{u}}_h)$, we get the energy identity

$$\mathsf{E}(\mathrm{L}_h, \boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) + \mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)) = (\boldsymbol{f}, \boldsymbol{v})_{\mathcal{T}_h},$$

and since $\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h), (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)) = \frac{1}{2}\langle|\boldsymbol{\beta}\cdot\boldsymbol{n}|(\boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h)\,,\, \boldsymbol{u}_h - \widehat{\boldsymbol{u}}_h\rangle_{\partial\mathcal{T}_h} \ge 0$, the inequality follows. This completes the proof. ☐

The new discrete inequalities. Next, we relate the viscous energy of the discrete Stokes operator with our new discrete inequalitites of Theorem 2.3.

Theorem 5.2 (global, discrete $\boldsymbol{H}^1$- and Poincaré–Friedrichs inequalities). *Let* $(\mathrm{r}_h, \boldsymbol{z}_h, \widehat{\boldsymbol{z}}_h) \in \mathcal{G}_h \times W_h \times M_h$ *satisfy*

$$(\mathrm{r}_h, \mathrm{G})_{\mathcal{T}_h} - (\boldsymbol{z}_h, \boldsymbol{\nabla}\cdot\mathrm{G})_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{z}}_h\,,\, \mathrm{G}\,\boldsymbol{n}\rangle_{\partial\mathcal{T}_h} = 0 \qquad \forall\, \mathrm{G} \in \mathcal{G}_h.$$

Then,

$$\begin{aligned}
\|\!|(\boldsymbol{z}_h, \widehat{\boldsymbol{z}}_h)|\!\|_{1,\mathcal{T}_h}^2 &\le C\,\Theta_h \quad (\boldsymbol{H}^1),\\
h^{-2}\,\|\!|(\boldsymbol{z}_h, \widehat{\boldsymbol{z}}_h)|\!\|_{\mathrm{PF},\mathcal{T}_h}^2 &\le C\,\Theta_h \quad (\text{Poincaré–Friedrichs}),
\end{aligned}$$

where

$$\Theta_h := \sum_{K\in\mathcal{T}_h} (\|\mathrm{r}_h\|_K^2 + h_K^{-1}\|P_{\boldsymbol{M}_S}(\boldsymbol{z}_h - \widehat{\boldsymbol{z}}_h)\|_{\partial K}^2) = \nu^{-1}\,\mathsf{E}(\mathrm{r}_h, \boldsymbol{z}_h, \widehat{\boldsymbol{z}}_h).$$

Here, the constant C *only depends on the finite element spaces* $\boldsymbol{V}(K)$*,* $W(K)$*, and* $M_S(\partial K)$*, and on the shape-regularity properties of the elements* $K \in \mathcal{T}_h$*.*

Proof. This result follows from the local discrete inequalities of Theorem 2.3. For $i = 1, \dots, d$, let r_i denote the ith row of the matrix r, and let $\boldsymbol{v}_i$ denote the ith component of the vector $\boldsymbol{v}$. Then, by the choice of the local spaces (4.4a), we have that, on the element K,

$$((\mathrm{r}_h)_i, (\boldsymbol{z}_h)_i, (\widehat{\boldsymbol{z}}_h)_i) \in \boldsymbol{V}^{\mathrm{D}}(K) \times W^{\mathrm{D}}(K) \times M^{\mathrm{D}}(\partial K),$$

and since $\boldsymbol{V}^{\mathrm{D}}(K) \times W^{\mathrm{D}}(K)$ admits an $M^{\mathrm{D}}(\partial K)$-decomposition, we can apply Theorem 2.3 with $\mathrm{c} = \mathrm{Id}$ and $(\boldsymbol{q}_h, u_h, \widehat{u}_h) := ((\mathrm{r}_h)_i, (\boldsymbol{z}_h)_i, (\widehat{\boldsymbol{z}}_h)_i)$. The inequalities now follow by adding over all element $K \in \mathcal{T}_h$ and then over the components $i = 1, \dots, d$. This completes the proof. □

Properties of the convective form $\mathcal{O}_h$. In the next result, we gather some properties of the convective form $\mathcal{O}_h$.

LEMMA 5.3 (properties of the nonlinear term $\mathcal{O}_h$ [4, Propositions 3.4 and 3.5]). *For any $(\boldsymbol{v}_h, \widehat{\boldsymbol{v}}_h) \in \boldsymbol{V}_h \times \boldsymbol{M}_h(0)$, we have*

$$|\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}_h, \widehat{\boldsymbol{v}}))| \lesssim |\!|\!|(\boldsymbol{\beta}, \{\boldsymbol{\beta}\})|\!|\!|_{1,\mathcal{T}_h} |\!|\!|(\boldsymbol{u}, \widehat{\boldsymbol{u}})|\!|\!|_{1,\mathcal{T}_h} |\!|\!|(\boldsymbol{v}_h, \widehat{\boldsymbol{v}}_h)|\!|\!|_{1,\mathcal{T}_h} \tag{5.5a}$$

for all $\boldsymbol{\beta} \in \boldsymbol{V}_h^*$ and $(\boldsymbol{u}, \widehat{\boldsymbol{u}}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h(0)$,

$$|\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}_h, \widehat{\boldsymbol{v}}))| \lesssim \|\boldsymbol{\beta}\|_{\infty,\Omega} |\!|\!|(\boldsymbol{u}, \widehat{\boldsymbol{u}})|\!|\!|_{0,\mathcal{T}_h} |\!|\!|(\boldsymbol{v}_h, \widehat{\boldsymbol{v}}_h)|\!|\!|_{1,\mathcal{T}_h} \tag{5.5b}$$

for all $\boldsymbol{\beta} \in L^\infty(\Omega)^d \cap \boldsymbol{V}_h^*$, and $(\boldsymbol{u}, \widehat{\boldsymbol{u}}) \in H^1(\mathcal{T}_h)^d \times L^2(\mathcal{F}_h, 0)^d$, and

$$\begin{aligned} &|\mathcal{O}_h(\boldsymbol{\beta}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}_h, \widehat{\boldsymbol{v}})) - \mathcal{O}_h(\boldsymbol{\gamma}; (\boldsymbol{u}, \widehat{\boldsymbol{u}}), (\boldsymbol{v}_h, \widehat{\boldsymbol{v}}))| \\ &\qquad \lesssim |\!|\!|(\boldsymbol{\beta} - \boldsymbol{\gamma}, 0)|\!|\!|_{0,\mathcal{T}_h} |\!|\!|(\boldsymbol{u}, \widehat{\boldsymbol{u}})|\!|\!|_{\infty,\mathcal{T}_h} |\!|\!|(\boldsymbol{v}_h, \widehat{\boldsymbol{v}}_h)|\!|\!|_{1,\mathcal{T}_h} \end{aligned} \tag{5.5c}$$

for all $\boldsymbol{\beta} \in H^1(\mathcal{T}_h)^d + \boldsymbol{V}_h^*$ and $(\boldsymbol{u}, \widehat{\boldsymbol{u}}) \in L^\infty(\mathcal{T}_h)^d \times L^\infty(\mathcal{F}_h)^d$.

5.2. Proof of Theorem 4.2.

Now we are ready to prove the existence and uniqueness of the approximation in Theorem 4.2. The proof is almost identical to that in [4, section 5].

We use a Banach fixed-point theorem by constructing a contraction mapping $\mathcal{F} : Z_h \to Z_h$, where

$$Z_h := \{(\boldsymbol{v}, \widehat{\boldsymbol{v}}) \in \boldsymbol{V}_h \times \boldsymbol{M}_h(0) : (\boldsymbol{v}_h, \nabla q)_{\mathcal{T}_h} - \langle \widehat{\boldsymbol{v}} \cdot \boldsymbol{n} \,,\, q \rangle_{\partial \mathcal{T}_h} = 0 \ \ \forall\, q \in \mathring{Q}_h\}.$$

Let us show that there is a ball K_h inside Z_h such that $\mathcal{F}$ maps K_h into K_h. For a pair $(\boldsymbol{w}_h, \widehat{\boldsymbol{w}}_h) \in Z_h$, the mapping is defined by $\mathcal{F}(\boldsymbol{w}_h, \widehat{\boldsymbol{w}}_h) := (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)$ with $(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)$ being part of the numerical solution to the linear system (4.2) with $\boldsymbol{\beta} = \boldsymbol{P}_h(\boldsymbol{w}_h, \widehat{\boldsymbol{w}}_h)$. By Lemma 4.1, we have that $\boldsymbol{\beta} \in \boldsymbol{V}_\beta$. Then,

$$\begin{aligned} |\!|\!|(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)|\!|\!|_{1,\mathcal{T}_h}^2 &\lesssim \nu^{-1} \mathsf{E}(\mathrm{L}_h, \boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h) && \text{by Theorem 5.2} \\ &\lesssim \nu^{-1} (\boldsymbol{f}, \boldsymbol{u}_h)_{\mathcal{T}_h} && \text{by Lemma 5.1} \\ &\lesssim \nu^{-1} \|\boldsymbol{f}\|_{\mathcal{T}_h} \|\boldsymbol{u}_h\|_{\mathcal{T}_h} \\ &\lesssim \nu^{-1} \|\boldsymbol{f}\|_{\mathcal{T}_h} |\!|\!|(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)|\!|\!|_{1,\mathcal{T}_h}, \end{aligned}$$

and we get

$$|\!|\!|(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h)|\!|\!|_{1,\mathcal{T}_h} \lesssim \nu^{-1} \|\boldsymbol{f}\|_{\mathcal{T}_h}.$$

Then, defining

$$K_h := \{(\boldsymbol{v}, \widehat{\boldsymbol{v}}) \in Z_h : \ \ \| (\boldsymbol{v}, \widehat{\boldsymbol{v}}) \|_{1,\mathcal{T}_h} \leq C_{\rm sm}\, \nu^{-1} \| \boldsymbol{f} \|_{\mathcal{T}_h} \}$$

with a positive constant $C_{\rm sm}$ big enough, we conclude that $\mathcal{F}$ maps K_h into itself.

Now we only have to show that $\mathcal{F}$ is a contraction in K_h. Set $(\boldsymbol{u}_h^1, \widehat{\boldsymbol{u}}_h^1) := \mathcal{F}(\boldsymbol{w}_h^1, \widehat{\boldsymbol{w}}_h^1)$ and $(\boldsymbol{u}_h^2, \widehat{\boldsymbol{u}}_h^2) := \mathcal{F}(\boldsymbol{w}_h^2, \widehat{\boldsymbol{w}}_h^2)$ with $(\boldsymbol{w}_h^i, \widehat{\boldsymbol{w}}_h^i) \in K_h$ for $i = 1, 2$. Now, let $(\mathrm{L}_h^i, \boldsymbol{u}_h^i, p_h^i, \widehat{\boldsymbol{u}}_h^i)$ be the solution to (4.2) with $\boldsymbol{\beta}^i := \boldsymbol{P}_h(\boldsymbol{w}_h, \widehat{\boldsymbol{w}}_h)$. Using $\mathrm{d}_L := \mathrm{L}_h^1 - \mathrm{L}_h^2$ and similar definitions for $\mathrm{d}_u, \mathrm{d}_p, \mathrm{d}_{\widehat{u}}, \mathrm{d}_\beta, \mathrm{d}_w$, and $\mathrm{d}_{\widehat{w}}$ and the fact that (5.2) is satisfied for $i = 1, 2$, we conclude that

$$\begin{aligned}
\mathsf{E}(\mathrm{d}_L, \mathrm{d}_u, \mathrm{d}_{\widehat{u}}) =& -\mathcal{O}_h(\boldsymbol{\beta}^1; (\boldsymbol{u}_h^1, \widehat{\boldsymbol{u}}_h^1), (\mathrm{d}_u, \mathrm{d}_{\widehat{u}})) + \mathcal{O}_h(\boldsymbol{\beta}^2; (\boldsymbol{u}_h^2, \widehat{\boldsymbol{u}}_h^2), (\mathrm{d}_u, \mathrm{d}_{\widehat{u}}))\\
=& -\mathcal{O}_h(\mathrm{d}_\beta, ; (\boldsymbol{u}_h^1, \widehat{\boldsymbol{u}}_h^1), (\mathrm{d}_u, \mathrm{d}_{\widehat{u}})) - \mathcal{O}_h(\boldsymbol{\beta}^2; (\mathrm{d}_u, \mathrm{d}_{\widehat{u}}), (\mathrm{d}_u, \mathrm{d}_{\widehat{u}}))\\
\leq& -\mathcal{O}_h(\mathrm{d}_\beta; (\boldsymbol{u}_h^1, \widehat{\boldsymbol{u}}_h^1), (\mathrm{d}_u, \mathrm{d}_{\widehat{u}})).
\end{aligned}$$

By Lemma 5.3, we easily get that

$$\begin{aligned}
\mathsf{E}(\mathrm{d}_L, \mathrm{d}_u, \mathrm{d}_{\widehat{u}}) &\lesssim \| (\mathrm{d}_\beta, \{\mathrm{d}_\beta\}) \|_{1,\mathcal{T}_h} \left\| (\boldsymbol{u}_h^1, \widehat{\boldsymbol{u}}_h^1) \right\|_{1,\mathcal{T}_h} \| (\mathrm{d}_u, \mathrm{d}_{\widehat{u}}) \|_{1,\mathcal{T}_h}\\
&\lesssim \| (\mathrm{d}_w, \mathrm{d}_{\widehat{w}}) \|_{1,\mathcal{T}_h} \left\| (\boldsymbol{u}_h^1, \widehat{\boldsymbol{u}}_h^1) \right\|_{1,\mathcal{T}_h} \| (\mathrm{d}_u, \mathrm{d}_{\widehat{u}}) \|_{1,\mathcal{T}_h} \quad \text{by Proposition 4.1,}\\
&\lesssim \nu^{-1} \| \boldsymbol{f} \|_{\mathcal{T}_h} \| (\mathrm{d}_w, \mathrm{d}_{\widehat{w}}) \|_{1,\mathcal{T}_h} \| (\mathrm{d}_u, \mathrm{d}_{\widehat{u}}) \|_{1,\mathcal{T}_h},
\end{aligned}$$

by Theorem 4.2. Combining this result with Theorem 5.2, we immediately get

$$\| (\mathrm{d}_u, \mathrm{d}_{\widehat{u}}) \|_{1,\mathcal{T}_h} \lesssim \ \nu^{-2} \| \boldsymbol{f} \|_{\mathcal{T}_h} \| (\mathrm{d}_w, \mathrm{d}_{\widehat{w}}) \|_{1,\mathcal{T}_h}.$$

Hence, for $\nu^{-2} \| \boldsymbol{f} \|_{\mathcal{T}_h}$ sufficiently small, the mapping $\mathcal{F}$ is a contraction in K_h. This completes the proof of Theorem 4.2.

5.3. Proof of estimate (4.9) in Theorem 4.4. The energy estimate (4.9) in Theorem 4.4 directly follows from Proposition 4.3, the approximation properties of the finite element spaces, and Theorem 5.4 below. To simplify the notation, we introduce the following approximation errors:

$$\delta_L := \mathrm{L} - P_{\mathcal{G}} \mathrm{L}, \ \ \boldsymbol{\delta}_u := \boldsymbol{u} - \Pi_{\boldsymbol{V}} \boldsymbol{u}, \ \ \delta_p := p - P_Q p, \ \ \boldsymbol{\delta}_{\widehat{u}} := \boldsymbol{u} - P_{\boldsymbol{M}} \boldsymbol{u}.$$

THEOREM 5.4. *Under the assumptions of Theorem* 4.4, *we have*

$$\| \mathrm{e}_L \|_{\mathcal{T}_h} + \| (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}) \|_{1,\mathcal{T}_h} + h^{-1} \| (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}) \|_{0,\mathcal{T}_h} + \| e_u \|_{\mathcal{T}_h} \leq C\, \nu^{-1}\, \Theta_{ns}^{1/2},$$

where

$$\Theta_{ns} := \sum_{K \in \mathcal{T}_h} h_K \left(\| \nu\, \delta_L\, \boldsymbol{n} \|_{\partial K}^2 + \| \delta_p \|_{\partial K}^2 \right) + \| \boldsymbol{u} \|_{\infty,\Omega}^2 \, \| (\boldsymbol{\delta}_u, \boldsymbol{\delta}_{\widehat{u}}) \|_{0,\mathcal{T}_h}^2.$$

Here, the constant C depends only on the finite element spaces, the shape-regularity of the mesh $\mathcal{T}_h$, and the domain Ω.

The rest of this subsection is devoted to the proof of Theorem 5.4. We need the following two auxiliary results.

Lemma 5.5. *We have*

$$\begin{aligned} B_h(\mathrm{e}_L, \boldsymbol{e}_u, e_p, \boldsymbol{e}_{\widehat{u}};\, \mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}}) = {}& \langle \nu\, \delta_L\, \boldsymbol{n} - \delta_p\, \boldsymbol{n}\,,\, \boldsymbol{v} - \widehat{\boldsymbol{v}} \rangle_{\partial \mathcal{T}_h} \\ &+ \mathcal{O}_h(\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h),\, (\boldsymbol{v}, \widehat{\boldsymbol{v}})) \\ &- \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}),\, (\boldsymbol{v}, \widehat{\boldsymbol{v}})) \end{aligned}$$

for all $(\boldsymbol{v}, w_h, \widehat{w}_h) \in \boldsymbol{V}_h \times W_h \times M_h(0)$.

Proof. It is a direct consequence of the definition of the numerical method (5.2), the consistency of the method (5.3), and the definition of the projections in subsection 4.3.1. In particular, note that, by the definition of $\Pi_{\boldsymbol{V}}$, there holds

$$\left\langle \frac{\nu}{h_K} P_{\boldsymbol{M_S}}(\boldsymbol{\delta}_u - \boldsymbol{\delta}_{\widehat{u}})\,,\, \boldsymbol{v} - \widehat{\boldsymbol{v}} \right\rangle_{\partial K} = \frac{\nu}{h_K} \langle \boldsymbol{\delta}_u - \boldsymbol{\delta}_{\widehat{u}}\,,\, P_{\boldsymbol{M_S}}(\boldsymbol{v} - \widehat{\boldsymbol{v}}) \rangle_{\partial K} = 0. \qquad \square$$

Lemma 5.6. *We have*

$$\begin{aligned} &\mathcal{O}_h(\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})) \\ &\qquad - \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})) \lesssim \|\boldsymbol{u}\|_{\infty,\Omega}\, \Phi\, \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h}, \end{aligned}$$

where

$$\Phi := \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{0,\mathcal{T}_h} + \|\!|(\boldsymbol{\delta}_u, \boldsymbol{\delta}_{\widehat{u}})|\!\|_{0,\mathcal{T}_h} + \|\!|(\boldsymbol{P}_h(\Pi_{\boldsymbol{V}} \boldsymbol{u}, P_{\boldsymbol{M}} \boldsymbol{u}) - \boldsymbol{u}, 0)|\!\|_{0,\mathcal{T}_h}.$$

For a proof, see Appendix C; see also [4, section 6].

We are now ready to prove Theorem 5.4. Since the following estimates hold,

$$\begin{aligned} \|e_u\|_{\mathcal{T}_h} &\le \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h}, && \text{by [18, Theorem 2.1]}, \\ \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h} &\le C\, \nu^{-1}\, \mathsf{E}(\mathrm{e}_L, \boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}), && \text{by Theorem 5.2}, \\ h^{-1}\, \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{\mathrm{PF},\mathcal{T}_h} &\le \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h}, \end{aligned}$$

the left-hand side of the inequality in Theorem 5.4 is smaller than

$$C \nu^{-1} \mathsf{E}(\mathrm{e}_L, \boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}).$$

We turn to estimate the above term next using a standard energy argument.

To do that, we take

$$(\mathrm{G}, \boldsymbol{v}, q, \widehat{\boldsymbol{v}}) := (\mathrm{e}_L, \boldsymbol{e}_u, e_p, \boldsymbol{e}_{\widehat{u}})$$

in Lemma 5.5, to get

$$\begin{aligned} \mathsf{E}(\mathrm{e}_L, \boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}) = {}& \langle \nu\, \delta_L\, \boldsymbol{n} - \delta_p\, \boldsymbol{n}\,,\, \boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}} \rangle_{\partial \mathcal{T}_h} \\ &+ \mathcal{O}_h(\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h),\, (\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}})) \\ &- \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}),\, (\boldsymbol{e}_u - \boldsymbol{e}_{\widehat{u}})). \end{aligned}$$

Then, applying the Cauchy–Schwarz inequality and using Lemma 5.3, we obtain

$$\begin{aligned} \mathsf{E}(\mathrm{e}_L, \boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}) \lesssim {}& \left(\sum_{K \in \mathcal{T}_h} h_K\, \|\nu\, \delta_L\, \boldsymbol{n} - \delta_p\, \boldsymbol{n}\|^2_{\partial K} \right)^{1/2} \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h} \\ &+ \|\boldsymbol{u}\|_{\infty,\Omega} \|\!|(\boldsymbol{\delta}_u, \boldsymbol{\delta}_{\widehat{u}})|\!\|_{0,\mathcal{T}_h}\, \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h} \\ &+ \|\boldsymbol{u}\|_{\infty,\Omega}\, \|\!|(\boldsymbol{P}_h(\Pi_{\boldsymbol{V}} \boldsymbol{u}, P_{\boldsymbol{M}} \boldsymbol{u})) - \boldsymbol{u}, 0)|\!\|_{0,\mathcal{T}_h}\, \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h} \\ &+ \|\boldsymbol{u}\|_{\infty,\Omega}\, \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{0,\mathcal{T}_h}\, \|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h}. \end{aligned}$$

Now, assuming $\nu^{-1}\|\boldsymbol{u}\|_{\infty,\Omega}$ sufficiently small such that

$$\|\boldsymbol{u}\|_{\infty,\Omega}\,\|\!|(\boldsymbol{e}_u,\boldsymbol{e}_{\widehat{u}})|\!\|_{0,\mathcal{T}_h}\,\|\!|(\boldsymbol{e}_u,\boldsymbol{e}_{\widehat{u}})|\!\|_{1,\mathcal{T}_h}\le\frac{1}{2}\mathsf{E}(\mathrm{e}_L,\boldsymbol{e}_u,\boldsymbol{e}_{\widehat{u}}),$$

we get

$$\mathsf{E}(\mathrm{e}_L,\boldsymbol{e}_u,\boldsymbol{e}_{\widehat{u}})\lesssim\sum_{K\in\mathcal{T}_h}h_K\,\|\nu\,\delta_L\,\boldsymbol{n}-\delta_p\,\boldsymbol{n}\|^2_{\partial K}+\|\boldsymbol{u}\|^2_{\infty,\Omega}\,\|\!|(\boldsymbol{\delta}_u,\boldsymbol{\delta}_{\widehat{u}})|\!\|^2_{0,\mathcal{T}_h}$$

since we have that

$$\|\!|(\boldsymbol{P}_h(\Pi_{\boldsymbol{V}}\boldsymbol{u},P_{\boldsymbol{M}}\boldsymbol{u}))-\boldsymbol{u},0)|\!\|_{0,K}\lesssim\|\!|(\boldsymbol{\delta}_u,\boldsymbol{\delta}_{\widehat{u}})|\!\|_{0,K},$$

by the approximation properties of $\boldsymbol{P}_h$; see Proposition 4.1. This completes the proof of Theorem 5.4.

5.4. Proofs of estimates (4.10) and (4.11) in Theorem 4.4.

The superconvergent velocity estimates in L^2-norm in (4.10) and (4.11) follow from a standard duality argument. For a detailed proof, we refer interested reader to [4].

6. Concluding remarks.

As we pointed out in section 2.5, the application of our approach to the steady-state diffusion problem gives rise to the first superconvergent HDG method, namely, the so-called SFH method proposed in [8] when its nonzero stabilization is taken to be of order $1/h$. As shown in [8], the convergence properties of the SFH method remain unchanged when the stabilization function increases. A similar phenomenon takes place for all the methods considered here.

The extension of the techniques developed in this paper to other nonlinear partial differential equations constitutes the subject of ongoing work.

Appendix A. Proof of Proposition 4.1.

Here we give a proof of Proposition 4.1 on the properties of the convective velocity $\boldsymbol{P}_h(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$.

The well-posedness of the projection $\boldsymbol{P}_h(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)\in\boldsymbol{V}^*(K)$ is due to properties (4.5) on the space $\boldsymbol{V}^*(K)$ since we have $\gamma((\widetilde{\boldsymbol{V}^*}(K))^\perp)=M^{\mathrm{D}}(\partial K)$; see [14, Proposition 6.4]. Then, the first two estimates directly follows from scaling and norm-equivalence on finite dimensional spaces.

Now, assume $(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ satisfies (4.2c) for all $q\in\mathring{Q}_h$. By (4.6b) and property (4.5b) on the space $\boldsymbol{V}^*(K)$, we immediately have $\boldsymbol{P}_h(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)\in H(\mathrm{div};\Omega)$.

Let us now prove that it is divergence-free. Obviously, (4.2c) is satisfied for the constant test function $q=1$. Hence, we have, on each element K,

$$-(\boldsymbol{u}_h,\nabla q)_K+\langle\widehat{\boldsymbol{u}}_h\cdot\boldsymbol{n}\,,\,q\rangle_{\partial K}=0\quad\forall\,q\in Q(K)=W^{\mathrm{D}}(K).$$

Next, by the definition of $\widetilde{\boldsymbol{V}^*}(K)$, we have $\nabla Q(K)=\nabla W^{\mathrm{D}}(K)\subset\widetilde{\boldsymbol{V}^*}(K)$. Hence, using the definition of $\boldsymbol{P}_h(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)$ in the above equation, and integrating by parts, we get

$$(\nabla\cdot\boldsymbol{P}_h(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h),q)_K=0\quad\forall\,q\in W^{\mathrm{D}}(K).$$

This implies $\nabla\cdot\boldsymbol{P}_h(\boldsymbol{u}_h,\widehat{\boldsymbol{u}}_h)=0$ by (4.5a). This concludes the proof of Proposition 4.1.

Appendix B. Proof of Proposition 4.3.

Here, we give a proof of Proposition 4.3 on the approximation properties of the projection $\Pi_{\boldsymbol{V}}$. By defintion of $\Pi_{\boldsymbol{V}}\boldsymbol{u}$, (4.8), we have that, on the element K, its ith component, $(\Pi_{\boldsymbol{V}}\boldsymbol{u})_i$, is defined as the element of $W^{\text{D}}(K)$ such that

$$((\Pi_{\boldsymbol{V}}\boldsymbol{u})_i, w)_K = (\boldsymbol{u}_i, w)_K \quad \forall\, w \in W^{\text{D}}(K), \tag{B.1a}$$

$$\langle(\Pi_{\boldsymbol{V}}\boldsymbol{u})_i\,,\, \widehat{w}\rangle_{\partial K} = \langle \boldsymbol{u}_i\,,\, \widehat{w}\rangle_{\partial K} \quad \forall\, \widehat{w} \in M_S(\partial K). \tag{B.1b}$$

Thus, if we set $\Pi_W \boldsymbol{u}_i := (\Pi_{\boldsymbol{V}}\boldsymbol{u})_i$, to prove our result, we only have to prove a similar result for the projection Π_W.

By (B.1a), we have $\Pi_W u - P_W u \in \widetilde{W}^{\perp}(K)$. By (B.1b), we have

$$\langle \Pi_W u - P_W u\,,\, \widehat{w}\rangle_{\partial K} = \langle u - P_W u\,,\, \widehat{w}\rangle_{\partial K} \quad \forall\, \widehat{w} \in M_S.$$

Hence,

$$\|\Pi_W u - P_W u\|_{\partial K} \le\, C_{M_S}\|P_{M_S}(\Pi_W u - P_W u)\|_{\partial K} \le C_{M_S}\,\|u - P_W u\|_{\partial K},$$

since the constant C_{M_S} exists by condition (2.1). Then, the first estimate follows directly by scaling and norm-equivalence of $\|w\|_K$ and $\|w\|_{\partial K}$ for functions $w \in \widetilde{W}^{\perp}(K)$.

Moreover, we have

$$\begin{aligned}
\|P_W u - u\|_K + h_K^{1/2}\|P_W u - u\|_{\partial K} &\le \|P_W u\|_K + \|u\|_K + h_K^{1/2}\|P_W u\|_{\partial K} + h_K^{1/2}\|u\|_{\partial K}\\
&\le C\,\|P_W u\|_K + \|u\|_K + h_K^{1/2}\|u\|_{\partial K}\\
&\le C\,\|u\|_K + h_K^{1/2}\|u\|_{\partial K}\\
&\le C\,h_K^{d/2}\|u\|_{\infty,K}.
\end{aligned}$$

The second estimate is obtained by scaling, norm-equivalence of $h_K^{d/2}\|w\|_{\infty,K}$ and $\|w\|_K$ for the finite dimensional space $W(K)$, the above estimate, and the first estimate of Proposition 4.3. This completes the proof of Proposition 4.3.

Appendix C. Proof of Lemma 5.6.

Here, we prove Lemma 5.6 on the properties of the convective term $\mathcal{O}_h$. The main idea is to first split the terms on the left-hand side of the estimate in Lemma 5.6 into the sum of the following four terms,

$$\begin{aligned}
T_1 :=\, &\mathcal{O}_h(\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}))\\
&- \mathcal{O}_h(\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\Pi_{\boldsymbol{V}}, P_{\boldsymbol{M}}\boldsymbol{u}),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})),\\
T_2 :=\, &\mathcal{O}_h(\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}))\\
&- \mathcal{O}_h(\boldsymbol{P}_h(\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}); (\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})),\\
T_3 :=\, &\mathcal{O}_h(\boldsymbol{P}_h(\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}); (\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}))\\
&- \mathcal{O}_h(\boldsymbol{P}_h(\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}); (\boldsymbol{u}, \boldsymbol{u}),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})),\\
T_4 :=\, &\mathcal{O}_h(\boldsymbol{P}_h(\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}); (\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}))\\
&- \mathcal{O}_h(\boldsymbol{u}; (\boldsymbol{u}, \boldsymbol{u}),\, (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})),
\end{aligned}$$

and then estimate each of them.

So, by (5.5c), we have that $T_1 = -\mathcal{O}_h(\boldsymbol{P}_h(\boldsymbol{u}_h, \widehat{\boldsymbol{u}}_h); (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}), (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})) \le 0$.

For the second term, we have

$$\begin{aligned} T_2 = &-\mathcal{O}_h(\boldsymbol{P}_h(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}); (\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}), (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})) \\ &\lesssim |\!|\!|(\boldsymbol{P}_h(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}}), \{\boldsymbol{P}_h(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})\})|\!|\!|_{0,\mathcal{T}_h} |\!|\!|(\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u})|\!|\!|_{\infty,\mathcal{T}_h} |\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{1,\mathcal{T}_h} \\ &\lesssim |\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{0,\mathcal{T}_h} \|\boldsymbol{u}\|_{\infty,\Omega} |\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{1,\mathcal{T}_h}, \end{aligned}$$

by Propositions 4.1 and 4.3. For the third term, we have, by (5.5b),

$$\begin{aligned} T_3 &= \mathcal{O}_h(\boldsymbol{P}_h(\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}); (\boldsymbol{\delta}_u, \boldsymbol{\delta}_{\widehat{u}}), (\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})) \\ &\lesssim \|\boldsymbol{P}_h(\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u})\|_{\infty,\mathcal{T}_h} |\!|\!|(\boldsymbol{\delta}_u, \boldsymbol{\delta}_{\widehat{u}})|\!|\!|_{0,\mathcal{T}_h} |\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{1,\mathcal{T}_h} \\ &\lesssim \|\boldsymbol{u}\|_{\infty,\Omega} |\!|\!|(\boldsymbol{\delta}_u, \boldsymbol{\delta}_{\widehat{u}})|\!|\!|_{0,\mathcal{T}_h} |\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{1,\mathcal{T}_h}, \end{aligned}$$

by Proposition 4.1. For the last term, we have

$$T_4 \lesssim \| (\boldsymbol{P}_h(\Pi_{\boldsymbol{V}}\boldsymbol{u}, P_{\boldsymbol{M}}\boldsymbol{u}) - \boldsymbol{u}, 0) \|_{0,\mathcal{T}_h} \|\boldsymbol{u}\|_{\infty,\Omega} |\!|\!|(\boldsymbol{e}_u, \boldsymbol{e}_{\widehat{u}})|\!|\!|_{1,\mathcal{T}_h},$$

by (5.5c). This concludes the proof of Lemma 5.6.

Acknowledgment. The authors would like to thank the reviewers for their constructive comments leading to a better presentation of the material of this paper.

REFERENCES

[1] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer Ser. Comput. Math. 44, Springer, New York, 2013.

[2] F. Brezzi, J. Douglas, Jr., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47 (1985), pp. 217–235.

[3] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau, *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems*, SIAM J. Numer. Anal., 38 (2000), pp. 1676–1706.

[4] A. Cesmelioglu, B. Cockburn, and W. Qiu, *Analysis of a hybridizable discontinuous Galerkin method for the steady-state incompressible Navier–Stokes equations*, Math. Comp., 86 (2017), pp. 1643–1670.

[5] Y. Chen and B. Cockburn, *Analysis of variable-degree HDG methods for convection-diffusion equations. Part* I*: General nonconforming meshes*, IMA J. Numer. Anal., 32 (2012), pp. 1267–1293.

[6] E. Chung, B. Cockburn, and G. Fu, *The staggered DG method is the limit of a hybridizable DG method*, SIAM J. Numer. Anal., 52 (2014), pp. 915–932.

[7] E. T. Chung and B. Engquist, *Optimal discontinuous Galerkin methods for the acoustic wave equation in higher dimensions*, SIAM J. Numer. Anal., 47 (2009), pp. 3820–3848.

[8] B. Cockburn, B. Dong, and J. Guzmán, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp., 77 (2008), pp. 1887–1916.

[9] B. Cockburn and G. Fu, *Devising superconvergent HDG methods with symmetric approximate stresses for linear elasticity*, IMA J. Numer. Anal., 38 (2018), pp. 566–604.

[10] B. Cockburn and G. Fu, *Superconvergence by M-decompositions. Part* II*: Construction of two-dimensional finite elements*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 165–186.

[11] B. Cockburn and G. Fu, *Superconvergence by M-decompositions. Part* III*: Construction of three-dimensional finite elements*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 365–398.

[12] B. Cockburn and G. Fu, *A systematic construction of finite element commuting exact sequences*, SIAM J. Numer. Anal., 55 (2017), pp. 1650–1688.

[13] B. Cockburn, G. Fu, and W. Qiu, *A note on the devising of superconvergent HDG methods for Stokes flow by M-decompositions*, IMA J. Numer. Anal., 37 (2017), pp. 730–749.

[14] B. Cockburn, G. Fu, and F. J. Sayas, *Superconvergence by M-decompositions. Part* I*: General theory for HDG methods for diffusion*, Math. Comp., 86 (2017), pp. 1609–1641.

[15] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1319–1365.

[16] B. Cockburn, W. Qiu, and K. Shi, *Conditions for superconvergence of HDG methods for second-order eliptic problems*, Math. Comp., 81 (2012), pp. 1327–1353.

[17] B. Cockburn and K. Shi, *Conditions for superconvergence of HDG methods for Stokes flow*, Math. Comp., 82 (2013), pp. 651–671.

[18] D. A. Di Pietro and A. Ern, *Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier-Stokes equations*, Math. Comp., 79 (2010), pp. 1303–1330.

[19] H. Egger and J. Schöberl, *A hybrid mixed discontinuous Galerkin finite-element method for convection-diffusion problems*, IMA J. Numer. Anal., 30 (2010), pp. 1206–1234.

[20] P. A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems*, in Mathematical Aspects of the Finite Element Method, I. Galligani and E. Magenes, eds., Lecture Notes in Math. 606, Springer, New York, 1977, pp. 292–315.