SPHERICAL COVARIANCE REPRESENTATIONS

SERGEY G. BOBKOV[1] AND DEVRAJ DUGGAL[2]

ABSTRACT. Covariance representations are developed for the uniform distributions on the Euclidean spheres in terms of spherical gradients and Hessians. They are applied to derive a number of Sobolev type inequalities and to recover and refine the concentration of measure phenomenon, including second order concentration inequalities. A detail account is also given in the case of the circle, with a short overview of Höffding's kernels and covariance identities in the class of periodic functions.

1. Introduction

Let γ_n denote the standard Gaussian measure on $\mathbb{R}^n$ with density

$$\varphi_n(x) = (2\pi)^{-n/2} e^{-|x|^2/2}, \quad x \in \mathbb{R}^n.$$

The space $\mathbb{R}^n$ is equipped with the Euclidean norm $|\cdot|$ and the inner product $\langle \cdot, \cdot \rangle$.

It is well known that the covariance functional with respect to this measure

$$\mathrm{cov}_{\gamma_n}(f, g) = \int_{\mathbb{R}^n} fg \, d\gamma_n - \int_{\mathbb{R}^n} f \, d\gamma_n \int_{\mathbb{R}^n} g \, d\gamma_n$$

admits the representation

$$\mathrm{cov}_{\gamma_n}(f, g) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \langle \nabla f(x), \nabla g(y) \rangle \, d\pi_n(x, y) \tag{1.1}$$

for a certain probability measure π_n on $\mathbb{R}^n \times \mathbb{R}^n$. Here f and g may be arbitrary smooth functions on $\mathbb{R}^n$ with γ_n-square integrable gradients ∇f and ∇g.

In an equivalent form, this identity was derived in the works by Houdré and Pérez-Abreu using an interpolation argument in an infinite-dimensional setting for functionals of the Wiener process and by Ledoux by means of the Ornstein-Uhlenbeck semi-group, cf. [9], [11]. The form (1.1) with an explicit description of the mixing measure π_n was later proposed in [2]. It was shown that (1.1) can serve as a convenient tool to recover a number of Sobolev-type inequalities in Gauss space such as the Poincaré-type inequality and exponential bounds, as well as to refine the classical dimension-free concentration in the form of a deviation inequality.

The measure π_n in (1.1) is unique. Indeed, extending this relation by linearity to complex-valued functions and applying it to $f(x) = e^{i\langle t,x \rangle}$, $g(x) = e^{i\langle s,x \rangle}$ with parameters $t, s \in \mathbb{R}^n$,

we get an explicit expression for the Fourier-Stieltjes transform $\hat{\pi}_n$ of π_n, namely

$$\hat{\pi}_n(t,s) = e^{-\frac{1}{2}(|t|^2 + |s|^2)} \frac{1 - e^{-\langle t, s \rangle}}{\langle t, s \rangle}.$$

It determines π_n in a unique way. Although this measure is not Gaussian, it has γ_n as marginals, which is a crucial property for many applications.

What is also interesting, only Gaussian measures may satisfy a representation such as (1.1) in dimension $n \geq 2$: If for a probability measure μ on $\mathbb{R}^n$ there exists an identity

$$\text{cov}_\mu(f,g) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \langle \nabla f(x), \nabla g(y) \rangle \, d\pi(x,y) \tag{1.2}$$

with some (signed) measure π on $\mathbb{R}^n \times \mathbb{R}^n$, then necessarily μ must be a Gaussian measure with covariance matrix proportional to the identity matrix ([2], Theorem 2.4). Thus, (1.2) characterizes the class of Gaussian measures.

Nevertheless, it is natural to wonder whether or not one can obtain (1.2) with a proper modification of the Euclidean gradient. This turns out to be the case where μ is a uniform distribution on the hypercube $M = \{-1, 1\}^n$, that is, the product n-th power of the symmetric Bernoulli measure. Then we have (1.2) for the discrete gradient of functions on M ([2]).

In this paper, we consider a similar question for the uniform distribution $\mu = \sigma_{n-1}$ on the unit sphere

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}, \quad n \geq 2.$$

Any smooth function f on $\mathbb{S}^{n-1}$ has a (continuous) spherical gradient $\nabla_{\mathbb{S}} f(x)$ which may be defined for every point x on the sphere as the shortest vector $w \in \mathbb{R}^n$ satisfying the Taylor expansion up to the linear term

$$f(x') = f(x) + \langle w, x' - x \rangle + o(|x' - x|), \quad x' \to x \ (x, x' \in \mathbb{S}^{n-1}). \tag{1.3}$$

As we will see, (1.2) does exist for the spherical gradient.

Theorem 1.1. *On $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ there exists a positive measure ν_n such that, for all smooth functions f, g on $\mathbb{S}^{n-1}$,*

$$\text{cov}_{\sigma_{n-1}}(f,g) = \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle \, d\nu_n(x,y). \tag{1.4}$$

Moreover, $\nu_n = c_n \mu_n$ for a probability measure μ_n on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ with marginals σ_{n-1} and with a constant satisfying

$$\frac{1}{n-1} < c_n < \frac{1}{n-2} \quad (n \geq 3).$$

Here ν_n has density with respect to the product measure $\sigma_{n-1} \otimes \sigma_{n-1}$ of the form $\psi_n(x,y) = \psi_n(\langle x, y \rangle)$ for a certain positive function ψ_n on $[-1, 1]$. This ensures that the marginals of ν_n represent (equal) multiples of σ_{n-1}.

One should however emphasize that ν_n in (1.4) is not unique, in contrast with the Gaussian case. Let us explain this in the case of the circle. Let κ be any signed measure on the torus $\mathbb{S}^1 \times \mathbb{S}^1$ supported on

$$\Delta = \{(x,y) \in \mathbb{S}^1 \times \mathbb{S}^1 : \langle x, y \rangle = 0\}.$$

On this set necessarily $\langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle = 0$, since the vector $\nabla_{\mathbb{S}} f(x)$ is orthogonal to x, while $\nabla_{\mathbb{S}} g(y)$ is orthogonal to y. Hence, if (1.4) holds true for ν, this relation continues to

hold for the measure $\nu + \kappa$. In addition, κ may have σ_1 as marginals, for example, when κ is the image of σ_1 under the mapping $x \to (x, e^i x)$ from $\mathbb{S}^1$ to $\mathbb{S}^1 \times \mathbb{S}^1$.

Despite the non-uniqueness issue, the measure ν_n in Theorem 1.1 can be constructed in rather natural ways leading to the same result. The first approach is based on the Gaussian covariance identity (1.1) which we apply to 0-homogeneous functions f and g. For this aim, we give an integrable description of the density of π_n (Section 2) and then integrate in polar coordinates in order to reduce the right-hand side of (1.1) to the form (1.4) and thus to obtain an explicit integrable description for the density of ν_n (Sections 3-4). This will also allow us to analyze an asymptotic behavior of the function $\psi_n(x, y)$ when the points x, y are close to each other and to show that

$$\psi_n(x, y) \sim \frac{1}{|x - y|^{n-3}} \quad (n \geq 4), \qquad \psi_3(x, y) \sim \log \frac{1}{|x - y|}, \tag{1.5}$$

where the equivalence is understood as two sided bounds with n-dependent factors (Section 5). This relation shows that for the growing dimension n the mixing measure π_n is almost concentrated on the diagonal $x = y$.

In the case of the circle, which is treated separately in Section 6, the density ψ_2 is bounded. We will return to this case in the end of the paper in order to clarify the meaning of (1.1), connect it with the class of periodic covariance identities on the line, and eventually evaluate the density on the torus (addressing the uniqueness problem as well).

In Sections 7-8 we collect basic definitions and identities related to the differentiation on the sphere, the spherical Hessian and Laplacian $\Delta_{\mathbb{S}}$. Then we discuss a heat semi-group approach to the covariance functional with respect to σ_{n-1}, by involving the semi-group $P_t = e^{t\Delta_{\mathbb{S}}}$ (Section 9). This may be done in analogy with M. Ledoux' approach to the Gaussian covariance representation, which leads to the following counterpart of Theorem 1.1.

Theorem 1.2. *For all smooth functions f, g on $\mathbb{S}^{n-1}$ $(n \geq 3)$,*

$$\mathrm{cov}_{\sigma_{n-1}}(f, g) = \int_0^\infty \int_{\mathbb{S}^{n-1}} \langle \nabla_{\mathbb{S}} f, \nabla_{\mathbb{S}} P_t g \rangle \, d\sigma_{n-1} \, dt. \tag{1.6}$$

The relationship between the two representations (1.4) and (1.6) are briefly discussed in Section 10. Let us note here that (1.6) is no longer true on the circle.

Then we turn to applications of (1.4) including the spherical concentration. Let us recall that, for any function f on $\mathbb{S}^{n-1}$ with Lipschitz semi-norm $\|f\|_{\mathrm{Lip}} \leq 1$ and with σ_{n-1}-mean m, there is a deviation inequality

$$\sigma_{n-1}\{|f - m| \geq r\} \leq 2e^{-(n-1)\,r^2/2}, \quad r \geq 0. \tag{1.7}$$

This classical result can be proved by applying the spherical isoperimetric inequality (a theorem due to P. Lévy and Schmidt), or using the logarithmic Sobolev inequality on the sphere (due to Mueller and Weissler, cf. [13], [12]). In Section 11 the following assertion is proved.

Theorem 1.3. *For any function f on $\mathbb{S}^{n-1}$ with $\|f\|_{\mathrm{Lip}} \leq 1$ and σ_{n-1}-mean m,*

$$\sigma_{n-1}\{|f - m| \geq r\} \leq \frac{1}{r}\, e^{-(n-2)\,r^2/2}\, \mathbb{E}_{\sigma_{n-1}} |f - m|, \quad r \geq 0. \tag{1.8}$$

By the (spherical) Poincaré inequality, the expectation $\mathbb{E}_{\sigma_{n-1}} |f - m| = \int_{\mathbb{S}^{n-1}} |f - m| \, d\sigma_{n-1}$ is bounded by $\frac{1}{\sqrt{n-1}}$. Hence (1.8) yields

$$\sigma_{n-1}\{|f - m| \geq r\} \leq \frac{1}{r\sqrt{n-1}} \, e^{-(n-2)\, r^2/2},$$

which may be viewed as a sharpening of (1.7), at least for the values in the range $r \geq 1/\sqrt{n}$. Another point of sharpening is that the factor $\mathbb{E}_{\sigma_{n-1}} |f - m|$ may be much smaller than $1/\sqrt{n}$.

Covariance representations in Theorems 1.1-1.2 can be further developed in terms of derivatives of higher orders such as the spherical Hessian $f''_{\mathbb{S}}$, which is well motivated from the point of view of higher order concentration phenomena. In particular, it will be shown that on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ there exists a probability measure μ_n with marginals σ_{n-1} such that

$$\mathrm{cov}_{\sigma_{n-1}}(f, g) = c_n \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \langle Df(x), Dg(y) \rangle \, d\mu_n(x, y) \qquad (1.9)$$

for all C^2-smooth f, g that are orthogonal to all linear functions in $L^2(\sigma_{n-1})$, and with constants behaving like $c_n \sim 1/n^2$ for the growing dimension n. In this representation, we use the linear differential operator

$$Df(x) = f''_{\mathbb{S}}(x) - 2\,\nabla_{\mathbb{S}} f(x) \otimes x$$

and involve the canonical inner product for symmetric $n \times n$ matrices in the integrand in (1.9). Such results are discussed in Sections 12-13. First applications to bounding the covariance by quantities of order $1/n^2$ are given in Section 14, and then a spherical second order deviation inequality is derived in Section 15.

In the end, we give a short overview of one-dimensional covariance representations and then focus on the case of the circle in Theorem 1.1. The whole plan is as follows.

2. Gaussian covariance representation

The mixing measure in the covariance representation (1.1) may be constructed as the mixture

$$\pi_n = \int_0^1 \mathsf{L}\big(X, tX + \sqrt{1-t^2}\, Z\big)\, dt, \tag{2.1}$$

where X and Z are independent standard normal random vectors in $\mathbb{R}^n$ with density φ_n, and where L refers to the joint distribution. Since $tX + \sqrt{1-t^2}\, Z$ and X are distributed according to γ_n, it follows from this description that π_n has marginals γ_n.

Let us state the identity (1.1) once more in a way which will be needed later on.

Theorem 2.1. *Let u and v be C^1-smooth complex-valued functions on an open set G in $\mathbb{R}^n$ of γ_n-measure 1 such that*

$$\mathbb{E}\,|\nabla u(X)|^2 < \infty, \quad \mathbb{E}\,|\nabla v(X)|^2 < \infty. \tag{2.2}$$

Then the random variables $u(X)$ and $v(X)$ have finite second moments, and

$$\mathrm{cov}(u(X), v(X)) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \langle \nabla u(x), \nabla v(y) \rangle\, d\pi_n(x,y). \tag{2.3}$$

The double integral in (2.3) is bounded in absolute value by

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |\langle \nabla u(x), \nabla v(y) \rangle|\, d\pi_n(x,y) \le \frac{1}{2}\,\mathbb{E}\,|\nabla u(X)|^2 + \frac{1}{2}\,\mathbb{E}\,|\nabla v(X)|^2,$$

so that it is finite. Also, the Poincaré-type inequality

$$\mathrm{Var}(u(X)) \le \mathbb{E}\,|\nabla u(X)|^2 \tag{2.4}$$

(which readily follows from (1.1)) and the assumption (2.2) ensures that $u(X)$ and $v(X)$ have finite second moments, so the covariance is well-defined and finite as well.

The equality (2.3) is easily verified to be true for all exponential functions $u(x) = e^{i\langle t,x\rangle}$, $v(x) = e^{i\langle s,x\rangle}$. Therefore, it holds true for finite linear combinations of such functions and actually for Fourier-Stieltjes transforms of signed measures on $\mathbb{R}^n$ with finite absolute moments. This class is sufficiently rich to properly approximate all smooth functions as in Theorem 2.1 and establish (2.3) in full generality by an approximation argument.

As we have mentioned before, one may also state the identity (2.3) by involving the Ornstein-Uhlenbeck operators

$$P_t v(x) = \int_{\mathbb{R}^n} v\Big(e^{-t}x + \sqrt{1 - e^{-2t}}\, y\Big)\, d\gamma_n(y), \quad x \in \mathbb{R}^n,\ t \ge 0. \tag{2.5}$$

A similar definition is applied to vector-valued functions. These operators act on Lebesgue spaces $L^p(\gamma_n)$ as contractions and form a semi-group, that is, $P_t(P_s v) = P_{t+s} v$, $t, s \ge 0$.

If v is integrable and has an integrable gradient ∇v with respect to γ_n, one may differentiate under the integral sign in (2.5), which leads to

$$\nabla P_t v(x) = e^{-t} \int_{\mathbb{R}^n} \nabla v\Big(e^{-t}x + \sqrt{1 - e^{-2t}}\, y\Big)\, d\gamma_n(y) = e^{-t} P_t \nabla v(x).$$

An equivalent formulation of (2.3) is the following:

Theorem 2.2. *Under the assumptions of Theorem 2.1,*

$$\mathrm{cov}(u(X), v(X)) = \int_0^\infty \mathbb{E}\,\langle \nabla u(X), \nabla P_t v(X)\rangle\, dt, \tag{2.6}$$

where the integral is absolutely convergent.

From the definition (2.1) it follows that the probability measure π_n is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^n \times \mathbb{R}^n$ and has some density

$$p_n(x, y) = \frac{d\pi_n(x, y)}{dx\, dy}.$$

To write it down explicitly, let us write $s = \sqrt{1 - t^2}$. For any bounded Borel measurable function $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, changing the variable $z = \frac{y - tx}{s}$, we have

$$
\begin{aligned}
\mathbb{E}\, h(X, tX + sZ) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} h(x, tx + sz)\, \varphi_n(x)\varphi_n(z)\, dx\, dz \\
&= \frac{1}{s^n} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} h(x, y)\, \varphi_n(x)\varphi_n\Big(\frac{y - tx}{s}\Big)\, dx\, dy,
\end{aligned}
$$

which means that

$$p_n(x, y) = \int_0^1 s^{-n}\varphi_n(x)\varphi_n\Big(\frac{y - tx}{s}\Big)\, dt.$$

But

$$\varphi_n(x)\varphi_n\Big(\frac{y - tx}{s}\Big) = \frac{1}{(2\pi)^n}\,\exp\Big[-\frac{|x|^2 + |y|^2 - 2t\,\langle x, y\rangle}{2s^2}\Big].$$

Thus, we arrive at:

Proposition 2.3. *The mixing probability measure π_n in the Gaussian covariance representation (2.3) has density*

$$p_n(x, y) = \frac{1}{(2\pi)^n}\int_0^1 s^{-n}\,\exp\Big[-\frac{|x|^2 + |y|^2 - 2t\,\langle x, y\rangle}{2s^2}\Big]\, dt, \quad x, y \in \mathbb{R}^n, \tag{2.7}$$

where we write $s = \sqrt{1 - t^2}$.

If $n \geq 2$, the density p_n is unbounded, since $p_n(0, 0) = (2\pi)^{-n}\int_0^1 s^{-n}\, dt = \infty$. On the other hand, in the case $n = 1$, we have $p_1(x, y) \leq \frac{1}{4}$.

3. From the Gauss space to the sphere

Starting from (2.3), one can develop a spherical variant of the Gaussian covariance identity

$$\mathrm{cov}_{\gamma_n}(u, v) = \int_G\int_G \langle \nabla u(x), \nabla v(y)\rangle\, p_n(x, y)\, dx\, dy. \tag{3.1}$$

Here G is an open subset of $\mathbb{R}^n$ of a full γ_n-measure, and p_n is the density of the probability measure π_n described in Proposition 2.3. This section is devoted to the first step in this reduction. Putting

$$x = r\theta,\ y = r'\theta',\ r, r' > 0,\ \theta, \theta' \in \mathbb{S}^{n-1},$$

the formula (2.7) takes the form in polar coordinates

$$p_n(r\theta, r'\theta') = \frac{1}{(2\pi)^n} \int_0^1 s^{-n} \exp\left[-\frac{r^2 + r'^2 - 2rr't\langle\theta,\theta'\rangle}{2s^2}\right] dt, \tag{3.2}$$

where we use the notation $s = \sqrt{1-t^2}$ in the integrand, as before.

With two smooth functions $f, g : \mathbb{S}^{n-1} \to \mathbb{R}$, we associate the functions $u(x) = f(\theta)$, $v(y) = g(\theta')$, where

$$\theta = \frac{x}{r}, \ \theta' = \frac{y}{r'}, \qquad r = |x|, \ r' = |y|.$$

They are defined and C^1-smooth on $G = \mathbb{R}^n \setminus \{0\}$ and have (Euclidean) gradients

$$\nabla u(x) = \frac{1}{r} \nabla_{\mathbb{S}} f(\theta), \quad \nabla v(y) = \frac{1}{r'} \nabla_{\mathbb{S}} g(\theta'). \tag{3.3}$$

As random variables on the sphere, f and g have the same distributions under the normalized Lebesgue measure σ_{n-1} on $\mathbb{S}^{n-1}$ as u and v have under γ_n. In particular,

$$\int_{\mathbb{R}^n} u \, d\gamma_n = \int_{\mathbb{S}^{n-1}} f \, d\sigma_{n-1}, \quad \int_{\mathbb{R}^n} v \, d\gamma_n = \int_{\mathbb{S}^{n-1}} g \, d\sigma_{n-1}.$$

A similar identity is also true for the products uv ang fg. Hence from (3.1) we obtain that

$$\mathrm{cov}_{\sigma_{n-1}}(f, g) = \mathrm{cov}_{\gamma_n}(u, v) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle\nabla u(x), \nabla v(y)\rangle \, p_n(x, y) \, dx \, dy. \tag{3.4}$$

In addition, since r and θ are independent under γ_n, and r has the same distribution as $|Z|$ (where Z is a standard normal random vector in $\mathbb{R}^n$), we have

$$\int |\nabla u|^2 \, d\gamma_n = \mathbb{E}\, \frac{1}{|Z|^2} \int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} f|^2 \, d\sigma_{n-1}$$

and

$$\int |\nabla v|^2 \, d\gamma_n = \mathbb{E}\, \frac{1}{|Z|^2} \int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} g|^2 \, d\sigma_{n-1}.$$

Here, while the spherical integrals are finite by the smoothness of f and g, the expectation on the right-hand side is finite for $n \geq 3$, only. So, this has to be assumed. For further developments, we will need a recursive formula on similar expectations.

Lemma 3.1. *For any real number $m > -(n-2)$,*

$$\mathbb{E}\, |Z|^m = (m + n - 2)\, \mathbb{E}\, |Z|^{m-2},$$

where Z is a standard normal random vector in $\mathbb{R}^n$. In particular, $\mathbb{E}\, |Z|^{-2} = \frac{1}{n-2}$ for $n \geq 3$.

Proof. One may use a general formula

$$\mathbb{E}\, w(|Z|) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} w(|x|)\, e^{-|x|^2/2} \, dx = \frac{\omega_{n-1}}{(2\pi)^{n/2}} \int_0^\infty w(r) r^{n-1}\, e^{-r^2/2} \, dr,$$

where we integrated in polar coordinates in the last step. Here and in the sequel, we denote by ω_{n-1} the surface area of the sphere $\mathbb{S}^{n-1}$, that is,

$$\omega_{n-1} = \frac{(2\pi)^{\frac{n}{2}}}{2^{\frac{n}{2}-1}\, \Gamma(\frac{n}{2})}. \tag{3.5}$$

Integrating by parts for the polynomial function $w(r) = r^m$ with $m > -(n-2)$ (this condition is posed for the integrability reason), we get

$$
\begin{aligned}
\mathbb{E}\,|Z|^m &= \frac{\omega_{n-1}}{(2\pi)^{n/2}} \int_0^\infty r^{m+n-1}\, e^{-r^2/2}\, dr \\
&= (m+n-2)\, \frac{\omega_{n-1}}{(2\pi)^{n/2}} \int_0^\infty r^{m+n-3}\, e^{-r^2/2}\, dr = (m+n-2)\,\mathbb{E}\,|Z|^{m-2}.
\end{aligned}
$$

$\square$

4. Proof of Theorem 1.1 for $n \geq 3$

Keeping the same notations as before, let us return to (3.4). To continue, one may integrate on the right-hand of this representation in polar coordinates using the general formula

$$
\begin{aligned}
\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} h(x,y)\, dxdy &= \int_0^\infty\int_0^\infty \int_{\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}} h(r\theta, r'\theta')\,(rr')^{n-1}\, dr\, dr'\, d\theta\, d\theta' \\
&= \omega_{n-1}^2 \int_0^\infty\int_0^\infty \int_{\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}} h(r\theta, r'\theta')\,(rr')^{n-1}\, dr\, dr'\, d\sigma_{n-1}(\theta)\, d\sigma_{n-1}(\theta'),
\end{aligned}
$$

where ω_{n-1} stands for the area size of the sphere $\mathbb{S}^{n-1}$ described in (3.5). Hence, applying (3.2)-(3.3), the double integral in (3.4) may be rewritten, and we obtain that

$$
\mathrm{cov}_{\sigma_{n-1}}(f,g) = \int_{\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}} \big\langle \nabla_{\mathbb{S}} f(\theta), \nabla_{\mathbb{S}} g(\theta') \big\rangle\, \psi_n(\theta, \theta')\, d\sigma_{n-1}(\theta)\, d\sigma_{n-1}(\theta') \tag{4.1}
$$

with

$$
\begin{aligned}
\psi_n(\theta,\theta') &= \frac{1}{2^{n-2}\,\Gamma(\frac{n}{2})^2} \int_0^1 s^{-n}\left[\int_0^\infty\int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\,\langle\theta,\theta'\rangle}{2s^2}\right](rr')^{n-2}\, dr\, dr'\right] dt \\
&= \frac{1}{2^{n-2}\,\Gamma(\frac{n}{2})^2} \int_0^1 s^{n-2}\left[\int_0^\infty\int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\,\langle\theta,\theta'\rangle}{2}\right](rr')^{n-2}\, dr\, dr'\right] dt,
\end{aligned}
$$

where changed the variables in the last step by replacing r with sr and r' with sr'.

Let us see that ψ_n is integrable over the product measure $\sigma_{n-1} \otimes \sigma_{n-1}$, so that it serves as density of some finite Borel positive measure ν_n on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ with total mass

$$
c_n = \nu_n(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}) = \int_{\mathbb{S}^{n-1}}\int_{\mathbb{S}^{n-1}} \psi_n(\theta, \theta')\, d\sigma_{n-1}(\theta)\, d\sigma_{n-1}(\theta').
$$

Repeating integration in polar coordinates as before, one may notice that

$$
\mathbb{E}\int_0^1 \frac{1}{|X|\,|tX + sZ|}\, dt = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{1}{|x|\,|y|}\, p_n(x,y)\, dx\, dy = c_n.
$$

But, for any fixed $0 < t < 1$, by Cauchy's inequality and applying Lemma 3.1,

$$
\begin{aligned}
\mathbb{E}\,\frac{1}{|X|\,|tX + sZ|} &< \left(\mathbb{E}\,\frac{1}{|X|^2}\right)^{1/2}\left(\mathbb{E}\,\frac{1}{|tX + sZ|^2}\right)^{1/2} \\
&= \mathbb{E}\,\frac{1}{|X|^2} = \frac{1}{n-2} < \infty.
\end{aligned}
$$

Thus, $\nu_n = c_n\mu_n$ for some Borel probability measure μ_n on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ with $c_n < \frac{1}{n-2}$.

Note also that $\psi_n(\theta, \theta') = \psi_n(\langle \theta, \theta' \rangle)$ depends on θ and θ' via the inner product only. This implies that the integral

$$\int_{\mathbb{S}^{n-1}} \psi_n(\theta, \theta') \, d\sigma_{n-1}(\theta') = \int_{\mathbb{S}^{n-1}} \psi_n(\langle \theta, \theta' \rangle) \, d\sigma_{n-1}(\theta')$$

does not depend on θ and is therefore equal to c_n. As a consequence, we conclude that ν_n is a positive, finite, symmetric measure whose marginal distributions are spherically invariant measures. Hence they coincide with a multiple of σ_{n-1}. Returning to (4.1), one can summarize in the following refinement of Theorem 1.1 for the case $n \geq 3$.

Theorem 4.1. *For all smooth functions f, g on $\mathbb{S}^{n-1}$ ($n \geq 3$),*

$$\mathrm{cov}_{\sigma_{n-1}}(f, g) = \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle \, d\nu_n(x, y), \qquad (4.2)$$

where ν_n is a positive measure on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ with density $\psi_n(\langle x, y \rangle)$ with respect to the product measure $\sigma_{n-1} \otimes \sigma_{n-1}$, where

$$\psi_n(\alpha) = \frac{1}{2^{n-2} \, \Gamma(\frac{n}{2})^2} \int_0^1 s^{n-2} \left[\int_0^\infty \int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\alpha}{2} \right] (rr')^{n-2} \, dr \, dr' \right] dt \quad (4.3)$$

defined for $|\alpha| \leq 1$. Moreover, $\nu_n = c_n \mu_n$ for a probability measure μ_n on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ with marginals σ_{n-1} and with total mass satisfying

$$\frac{1}{n-1} < c_n < \frac{1}{n-2}. \qquad (4.4)$$

Proof. It remains to explain the lower bound in (4.3). From (4.2) in the case of equal functions $f = g$ we get

$$\mathrm{Var}(f(\xi)) = c_n \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \langle \nabla_{\mathbb{S}} f(\theta), \nabla_{\mathbb{S}} f(\theta') \rangle \, d\mu_n(\theta, \theta'), \qquad (4.5)$$

where ξ is a random vector with distribution σ_{n-1}. By Cauchy's inequality,

$$\begin{aligned}
\mathrm{Var}(f(\xi)) & \leq c_n \left(\int\!\!\int |\nabla_{\mathbb{S}} f(x)|^2 \, d\mu_n(x, y) \right)^{1/2} \left(\int\!\!\int |\nabla_{\mathbb{S}} f(y)|^2 \, d\mu_n(x, y) \right)^{1/2} \quad (4.6) \\
& = c_n \int |\nabla_{\mathbb{S}} f|^2 \, d\sigma_{n-1}.
\end{aligned}$$

Thus, we arrive at the Poincaré-type inequality $\mathrm{Var}(f(\xi)) \leq c_n \, \mathbb{E} \, |\nabla_{\mathbb{S}} f(\xi)|^2$ on the unit sphere. Here, the constant may not be better than the optimal constant $\frac{1}{n-1}$.

Moreover, the equality in (4.6) is only possible if and only if

$$\langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} f(y) \rangle = |\nabla_{\mathbb{S}} f(x)| \, |\nabla_{\mathbb{S}} f(y)|$$

for μ_n-almost all $(x, y) \in \mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$, hence for all $x, y \in \mathbb{S}^{n-1}$ (by the continuity of the gradient). Equivalently, one should have $\nabla_{\mathbb{S}} f(x) = \nabla_{\mathbb{S}} f(y)$ for all $x, y \in \mathbb{S}^{n-1}$, which means that f must be a constant. This implies that $c_n > \frac{1}{n-1}$. $\qquad \square$

Remark 4.2. In order to estimate c_n and prove (4.4), it is also sufficient to apply the identity (4.5) to linear functions $f(x) = \langle x, v \rangle$, $v \in \mathbb{S}^{n-1}$. In this case $\mathrm{Var}(f(\xi)) = \frac{1}{n}$ and

$\nabla_{\mathbb{S}} f(\theta) = v - \langle v, \theta \rangle \theta$, so that

$$\langle \nabla_{\mathbb{S}} f(\theta), \nabla_{\mathbb{S}} f(\theta') \rangle = 1 - \langle v, \theta \rangle^2 - \langle v, \theta' \rangle^2 + \langle v, \theta \rangle \langle v, \theta' \rangle \langle \theta, \theta' \rangle.$$

Hence, integrating this expression over μ_n and using the property that μ_n has marginals σ_{n-1}, (4.5) yields

$$\frac{1}{nc_n} = 1 - \frac{2}{n} + \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \langle v, \theta \rangle \langle v, \theta' \rangle \langle \theta, \theta' \rangle \, d\mu_n(\theta, \theta').$$

Here, the left-hand side does not depend on v, while formally the right-hand side does. Integrating this equality over $d\sigma_{n-1}(v)$, we then get

$$\frac{1}{nc_n} = 1 - \frac{2}{n} + \frac{1}{n} \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \langle \theta, \theta' \rangle^2 \, d\mu_n(\theta, \theta').$$

The latter implies $\frac{1}{nc_n} > 1 - \frac{2}{n}$, that is, $c_n < \frac{1}{n-2}$. On the other hand, since $\langle \theta, \theta' \rangle^2 < 1$ for μ_n-almost all couples (θ, θ'), the double integral is smaller than 1. Hence, $\frac{1}{nc_n} < 1 - \frac{1}{n}$, that is, we also have a lower bound $c_n > \frac{1}{n-1}$.

5. Asymptotic behaviour of mixing densities

Since $\langle \theta, \theta' \rangle = 1 - |\theta - \theta'|^2$, the property that the density function $\psi_n(\theta, \theta')$ on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ depends on the inner product is equivalent to the assertion that it depends on the distance function $|\theta - \theta'|$. We will now examine an asymptotic behaviour of this function when this distance is close to zero.

The previous results are applicable in dimension $n = 2$ as well (with a slightly modified argument), that is, for the circle $\mathbb{S}^1$ which we equip with the uniform probability measure σ_1. In this case, according to (4.3), the measure ν_2 has density

$$\psi_2(\theta, \theta') = \psi_2(\langle \theta, \theta' \rangle) = \int_0^1 \left[\int_0^\infty \int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't \langle \theta, \theta' \rangle}{2} \right] dr dr' \right] dt, \qquad (5.1)$$

where we use the notation $s = \sqrt{1 - t^2}$. In the previous section, it was however not checked, whether ν_2 is finite or not. In fact, although the above triple integral is apparently not easy to evaluate explicitly, one can show that the density ψ_2 is bounded.

Since $\sigma_1 \otimes \sigma_1$ is invariant under transformations $(\theta, \theta') \to (\pm\theta, \pm\theta')$, we have

$$\nu_2(\mathbb{S}^1 \times \mathbb{S}^1) = \frac{1}{4} \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \Psi(\alpha) \, d\sigma_1(\theta) \, d\sigma_1(\theta'), \quad \alpha = \langle \theta, \theta' \rangle,$$

where

$$\Psi(\alpha) = \int_0^1 \left[\int_{-\infty}^\infty \int_{-\infty}^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\alpha}{2} \right] dr dr' \right] dt.$$

In addition,

$$\psi_2(\alpha) \le \Psi(\alpha). \qquad (5.2)$$

The inner integral with respect to r in the definition of $\Psi(\alpha)$ is equal to

$$\int_{-\infty}^\infty \exp\left[-\frac{(r - r't\alpha)^2 - (r't\alpha)^2 + r'^2}{2} \right] dr = \sqrt{2\pi} \exp\left[-\frac{1 - (t\alpha)^2}{2} r'^2 \right].$$

The next integration with respect to r' yields the value $\frac{2\pi}{\sqrt{1-(t\alpha)^2}}$. Integrating over $0 < t < 1$, we then get

$$\Psi(\alpha) = 2\pi \int_0^1 \frac{dt}{\sqrt{1-(t\alpha)^2}} = 2\pi \frac{\arcsin(\alpha)}{\alpha}.$$

In view of the concavity of the sine function on $[0, \frac{\pi}{2}]$, we have $\frac{\pi}{2} \sin(t) \geq t$ on this interval, or equivalently $\arcsin(\alpha) \leq \frac{\pi}{2} \alpha$ for all $0 \leq \alpha \leq 1$. Thus, $\Psi(\alpha) \leq \pi^2$ implying the following:

Lemma 5.1. *We have $\psi_2(\theta, \theta') \leq \pi^2$ for all $\theta, \theta' \in \mathbb{S}^1$ and*

$$c_2 = \nu_2(\mathbb{S}^1 \times \mathbb{S}^1) \leq \frac{\pi^2}{4}.$$

However, the density of the mixing measure ν_n is not bounded for $n \geq 3$. Let us give natural lower and upper bounds for $\psi_n(\alpha)$ in terms of α and n as α approaches 1. To this aim, consider the functions

$$R_m(\alpha) = \int_0^1 s^m \left[\int_0^\infty \int_0^\infty \exp \left[-\frac{r^2 + r'^2 - 2rr't\alpha}{2} \right] (rr')^m \, dr \, dr' \right] dt, \tag{5.3}$$

for $-1 \leq \alpha < 1$ with a real parameter $m \geq 0$ which may be used with $m = n - 2$ for the sphere $\mathbb{S}^{n-1}$, $n \geq 3$, and with $0 < m < 1$ for $n = 2$.

Put $\beta = 1 - \alpha t$ and first consider the inner double integral

$$\begin{aligned}
I_m(\beta) &= \int_0^\infty \int_0^\infty \exp \left[-\frac{r^2 + r'^2 - 2(1-\beta)rr'}{2} \right] (rr')^m \, dr \, dr' \\
&= \int_0^\infty \int_0^\infty \exp \left[-\frac{1}{2}(r-r')^2 - \beta rr' \right] (rr')^m \, dr \, dr'.
\end{aligned}$$

Changing the variables $x = r - r'$, $y = rr'$, the last integral takes the form

$$\int_0^\infty \int_0^\infty \exp \left[-\frac{1}{2} x^2 - \beta y \right] y^m \, dr \, dr'.$$

The next change $r' = y/x$, $x = r - y/x$ with arbitrary $x \in \mathbb{R}$ and $y > 0$ leads to

$$I_m(\beta) = \int_{-\infty}^\infty \int_0^\infty \exp \left[-\frac{1}{2} x^2 - \beta y \right] \frac{y^m}{\sqrt{x^2 + 4y}} \, dx \, dy, \tag{5.4}$$

which is bounded from above by

$$\frac{1}{2} \int_{-\infty}^\infty \int_0^\infty \exp \left[-\frac{1}{2} x^2 - \beta y \right] y^{m-1/2} \, dx \, dy = \frac{\sqrt{2\pi}}{2} \beta^{-m-1/2} \, \Gamma(m + 1/2).$$

For a similar lower bound, one may restrict integration in (5.4) to the regions $|x| < 1$ and $y > 1$, in which case $x^2 + 4y \leq 5y$. Since also $0 < \beta < 1$, this leads to

$$\begin{aligned}
I_m(\beta) &\geq \frac{2}{\sqrt{5e}} \int_1^\infty e^{-\beta y} y^{m-1/2} \, dy = \frac{2}{\sqrt{5e}} \beta^{-m-1/2} \int_\beta^\infty e^{-z} z^{m-1/2} \, dz \\
&\geq b_0 \, \beta^{-m-1/2} \, \Gamma(m + 1/2)
\end{aligned}$$

for some absolute constant $b_0 > 0$. Thus,

$$b_0 \, \beta^{-m-1/2} \, \Gamma(m + 1/2) \leq I_m(\beta) \leq b_1 \, \beta^{-m-1/2} \, \Gamma(m + 1/2), \tag{5.5}$$

where one may take $b_1 = \frac{\sqrt{2\pi}}{2}$ and where we recall that $\beta = 1 - \alpha t$.

Let us turn to (5.3) and apply the upper bound in (5.5) which gives

$$
\begin{aligned}
R_m(\alpha) &= \int_0^1 s^m \, I_m(1 - \alpha t) \, dt \\
&\leq \frac{\sqrt{2\pi}}{2} \, \Gamma(m + 1/2) \int_0^1 s^m \, (1 - \alpha t)^{-m-1/2} \, dt.
\end{aligned}
\tag{5.6}
$$

Note that $R_m(\alpha)$ is increasing in $-1 \leq \alpha \leq 1$. Using $s \leq \sqrt{2} \, (1 - t)^{1/2}$ and assuming $0 \leq m < 1$, this gives

$$
R_m(\alpha) \leq R_m(1) \leq \sqrt{\pi} \, \Gamma(m + 1/2) \int_0^1 (1 - t)^{-\frac{m+1}{2}} \, dt = \sqrt{\pi} \, \Gamma(m + 1/2) \, \frac{2}{1 - m}.
$$

That is, we proved:

Lemma 5.2. *For $0 \leq m < 1$ and $|\alpha| \leq 1$,*

$$
R_m(\alpha) \leq 2\sqrt{\pi} \, \frac{\Gamma(m + 1/2)}{1 - m}.
\tag{5.7}
$$

In particular, choosing $m = 0$, we have $\psi_2(\theta, \theta') = R_0(\alpha) \leq 2\pi$ for all $\theta, \theta' \in \mathbb{S}^1$.

The last bound is slightly better then the uniform bound of Lemma 5.1.

However, Lemma 5.2 is not applicable for $m \geq 1$, and we have to return to the bound (5.6). We need to estimate the integral

$$
J_m(\alpha) = \int_0^1 s^m \, (1 - \alpha t)^{-m-1/2} \, dt
$$

as α approaching 1. Let us write $Q_1 \sim Q_2$ for two positive quantities, if the ratio Q_1/Q_2 is bounded away from zero and infinity by some m-dependent constants. Note that $J_m(\alpha) \leq J_m(1/2) \leq 2^{m+1/2}$ for $\alpha \leq 1/2$, so, we may assume that $\frac{1}{2} \leq \alpha < 1$. Since the above integral remains bounded when integrating over $0 < t < 1/2$, and $s \sim (1 - t)^{1/2}$, we have, changing the variable and putting $\varepsilon = 1 - \alpha$,

$$
\begin{aligned}
J_m(\alpha) &\sim \int_0^{1/2} \frac{t^{m/2}}{(1 - \alpha(1 - t))^{m+1/2}} \, dt = \int_0^{1/2} \frac{t^{m/2}}{(\varepsilon + (1 - \varepsilon) \, t)^{m+1/2}} \, dt \\
&= \varepsilon^{-\frac{m-1}{2}} \int_0^{1/2\varepsilon} \frac{x^{m/2}}{(1 + (1 - \varepsilon) \, x)^{m+1/2}} \, dt \sim \varepsilon^{-\frac{m-1}{2}} \int_0^{1/2\varepsilon} \frac{x^{m/2}}{(1 + x)^{m+1/2}} \, dx \\
&\sim \varepsilon^{-\frac{m-1}{2}} \int_{1/2}^{1/2\varepsilon} \frac{1}{(1 + x)^{\frac{m+1}{2}}} \, dx.
\end{aligned}
$$

The last integral is bounded for $m > 1$ and is equivalent to $\log(1/\varepsilon)$ for $m = 1$. Therefore,

$$
R_m(\alpha) \sim (1 - \alpha)^{-\frac{m-1}{2}} \quad (m > 1), \qquad R_1(\alpha) \sim \log \frac{1}{1 - \alpha}.
$$

Using this with $m = n - 2$ according to the formula (4.3), we obtain a similar conclusion about the densities $\psi_n(\alpha) = \psi_n(\langle \theta, \theta' \rangle)$ of ν_n and thus prove the asymptotic relations (1.5):

Proposition 5.3. *For* $-1 \le \alpha < 1$,

$$\psi_n(\alpha) \sim (1-\alpha)^{-\frac{n-3}{2}} \quad (n \ge 4), \qquad \psi_3(\alpha) \sim \log \frac{1}{1-\alpha}.$$

6. Proof of Theorem 1.1 for the circle

Here we extend the assertion of Theorem 4.1 to the case of the circle, using the same formula (4.3) for the density ψ_2 for the mixing measure ν_2 which is also described in (5.1).

For this aim, the Gaussian covariance identity may still be used, however, with slightly different functions u and v. Namely, with a given smooth function $f : \mathbb{S}^{n-1} \to \mathbb{R}$, we associate the homogeneous functions of order $\varepsilon > 0$

$$u_\varepsilon(x) = r^\varepsilon f(r^{-1}x) = r^\varepsilon f(\theta), \quad \text{where} \quad r = |x| = \sqrt{x_1^2 + \cdots + x_n^2}, \ \theta = r^{-1}x.$$

They are defined, C^1-smooth in the whole space except for the origin, and have gradients

$$\nabla u_\varepsilon(x) = r^{\varepsilon-1} \left(\varepsilon f(\theta)\theta + \nabla_\mathbb{S} f(\theta) \right).$$

Define similarly the homogeneous functions v_ε for a smooth function $g : \mathbb{S}^{n-1} \to \mathbb{R}$, so that

$$\nabla v_\varepsilon(x) = r^{\varepsilon-1} \left(\varepsilon g(\theta)\theta + \nabla_\mathbb{S} g(\theta) \right).$$

Since r and θ are independent under γ_n, while r has the same distribution as $|Z|$, we have

$$\mathrm{cov}_{\gamma_n}(u_\varepsilon, v_\varepsilon) = \mathbb{E} |Z|^{2\varepsilon} \, \mathrm{cov}_{\sigma_{n-1}}(f, g) \ \to \ \mathrm{cov}_{\sigma_{n-1}}(f, g) \tag{6.1}$$

as $\varepsilon \to 0$. In addition,

$$\int |\nabla u_\varepsilon|^2 \, d\gamma_n = \mathbb{E} |Z|^{2\varepsilon-2} \int_{\mathbb{S}^{n-1}} (\varepsilon^2 f^2 + |\nabla_\mathbb{S} f|^2) \, d\sigma_{n-1},$$

which is finite, and similarly for v_ε. Hence, we may apply Theorem 2.1 to the couple $(u_\varepsilon, v_\varepsilon)$, which gives

$$\mathrm{cov}_{\gamma_n}(u_\varepsilon, v_\varepsilon) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \langle \nabla u_\varepsilon(x), \nabla v_\varepsilon(y) \rangle \, p_n(x, y) \, dx \, dy. \tag{6.2}$$

Putting

$$\begin{aligned} w_\varepsilon(\theta, \theta') \ = \ & \varepsilon^2 f(\theta) g(\theta') \langle \theta, \theta' \rangle + \varepsilon f(\theta) \langle \theta, \nabla_\mathbb{S} g(\theta') \rangle \\ & + \varepsilon g(\theta') \langle \theta', \nabla_\mathbb{S} f(\theta) \rangle + \langle \nabla_\mathbb{S} f(\theta), \nabla_\mathbb{S} g(\theta') \rangle, \end{aligned}$$

one may integrate in polar coordinates separately with respect to x and y in (6.2) and represent the double integral similarly as in the proof of Theorem 4.1 as

$$\int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} w_\varepsilon(\theta, \theta') \, \psi_{n,\varepsilon}(\theta, \theta') \, d\sigma_{n-1}(\theta) d\sigma_{n-1}(\theta') \tag{6.3}$$

with $\psi_{n,\varepsilon}(\theta, \theta')$ given by

$$\frac{1}{2^{n-2}\,\Gamma(\frac{n}{2})^2} \int_0^1 s^{-n} \left[\int_0^\infty \int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\,\langle\theta, \theta'\rangle}{2s^2} \right] (rr')^{n+\varepsilon-2} \, dr dr' \right] dt,$$

where we use the notation $s = \sqrt{1-t^2}$. Changing the variable and using $\alpha = \langle \theta, \theta' \rangle$, this expression is simplified to

$$\frac{1}{2^{n-2}\,\Gamma(\frac{n}{2})^2} \int_0^1 s^{n+2\varepsilon-2} \left[\int_0^\infty \int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\alpha}{2} \right] (rr')^{n+\varepsilon-2}\, dr dr' \right] dt.$$

As we know from Section 4, one may continue with $\varepsilon = 0$ in the case $n \geq 3$. If $n = 2$, then

$$\psi_{2,\varepsilon}(\alpha) = \int_0^1 s^{2\varepsilon} \left[\int_0^\infty \int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\alpha}{2} \right] (rr')^\varepsilon\, dr dr' \right] dt,$$

and by (6.2),

$$\mathrm{cov}_{\gamma_2}(u_\varepsilon, v_\varepsilon) = \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} w_\varepsilon(\theta, \theta')\, \psi_{2,\varepsilon}(\theta, \theta')\, d\sigma_1(\theta) d\sigma_1(\theta'). \tag{6.4}$$

Here, the covariance part is convergent as $\varepsilon \to 0$ to $\mathrm{cov}_{\sigma_1}(f, g)$, according to (6.1).

In order to turn to the limit on the right-hand side, note that the functions $w_\varepsilon(\theta, \theta')$ remain bounded uniformly over all $\theta, \theta' \in \mathbb{S}^1$, $0 < \varepsilon \leq 1/4$, and are convergent to $\langle \nabla_{\mathbb{S}} f(\theta), \nabla_{\mathbb{S}} g(\theta') \rangle$ as $\varepsilon \to 0$. By Lemma 6.2 with $m = \varepsilon$, the quantities $\psi_{2,\varepsilon}(\alpha) = R_{2\varepsilon}(\alpha)$ are also bounded by a constant. Hence, if we show that $\psi_{2,\varepsilon}(\alpha) \to \psi_2(\alpha)$ for all $\alpha \in [-1, 1)$, one may apply the Lebesgue dominated convergence theorem, so that to derive from (6.4) the limit case

$$\mathrm{cov}_{\sigma_1}(f, g) = \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \langle \nabla_{\mathbb{S}} f(\theta), \nabla_{\mathbb{S}} g(\theta') \rangle\, \psi_2(\theta, \theta')\, d\sigma_1(\theta) d\sigma_1(\theta').$$

As we did in the previous section, cf. (5.3)-(5.4) with $m = 2\varepsilon$, we have

$$\psi_{2,\varepsilon}(\alpha) = R_{2\varepsilon}(\alpha) = \int_0^1 \int_{-\infty}^\infty \int_0^\infty \exp\left[-\frac{1}{2} x^2 - (1-t\alpha)y \right] \frac{(sy)^{2\varepsilon}}{\sqrt{x^2+4y}}\, dx\, dy\, dt,$$

of which the limit case is given by

$$\psi_2(\alpha) = R_0(\alpha) = \int_0^1 \int_{-\infty}^\infty \int_0^\infty \exp\left[-\frac{1}{2} x^2 - (1-t\alpha)y \right] \frac{1}{\sqrt{x^2+4y}}\, dx\, dy\, dt. \tag{6.5}$$

In order to apply the Lebesgue dominated convergence theorem and show that $\psi_{2,\varepsilon}(\alpha) \to \psi_2(\alpha)$, it is sufficient to see that the integrands

$$K_{\varepsilon,\alpha}(x, y, t) = \exp\left[-\frac{1}{2} x^2 - (1-t\alpha)y \right] \frac{(sy)^{2\varepsilon}}{\sqrt{x^2+4y}}$$

are bounded by an integrable function $K_\alpha(x, y, t)$ on $\mathbb{R} \times (0, \infty) \times (0, 1)$ with respect to the Lebesgue measure on this product space uniformly over all $0 < \varepsilon \leq \frac{1}{4}$. Clearly, for a fixed value of $\alpha \in [-1, 1)$, one should examine the majorant

$$K_\alpha(x, y, t) = \sup_{0 < \varepsilon \leq \frac{1}{4}} K_{\varepsilon,\alpha}(x, y, t) = \exp\left[-\frac{1}{2} x^2 - (1-t\alpha)y \right] \frac{\max(1, \sqrt{sy})}{\sqrt{x^2+4y}}.$$

Clearly, the last expression is maximized for $\alpha = 1$. Since also $s \leq 2(1-t)$, we have

$$K_\alpha(x, y, t) \leq K_1(x, y, t) \leq \exp\left[-\frac{1}{2} x^2 - (1-t)y \right] \frac{\max(1, \sqrt{(1-t)y})}{\sqrt{2y}},$$

so,

$$\int_0^1 \int_{-\infty}^{\infty} \int_0^{\infty} K_\alpha(x, y, t)\, dx\, dy\, dt \;\leq\; \sqrt{2\pi} \int_0^1 \int_0^{\infty} e^{-(1-t)y}\, \frac{\max(1, \sqrt{(1-t)y})}{\sqrt{2y}}\; dy\, dt$$

$$= \;\sqrt{2\pi} \int_0^1 \int_0^{\infty} \frac{1}{\sqrt{1-t}}\, e^{-z}\, \frac{\max(1, \sqrt{z})}{\sqrt{2z}}\; dz\, dt \;<\; \infty.$$

One may summarize.

Proposition 6.1. *On the torus $\mathbb{S}^1 \times \mathbb{S}^1$ there exists a positive measure ν_2 with marginals $c\sigma_1$ for some constant $c > 0$ such that, for all smooth functions f, g on $\mathbb{S}^1$, we have the covariance representation*

$$\mathrm{cov}_{\sigma_1}(f, g) = \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \big\langle \nabla_{\mathbb{S}} f(\theta), \nabla_{\mathbb{S}} g(\theta') \big\rangle\, d\nu_2(\theta, \theta'). \tag{6.6}$$

Moreover, the mixing measure ν_2 has density $\psi_2(\langle \theta, \theta' \rangle)$ for the function ψ_2 with respect to the product measure $\sigma_1 \otimes \sigma_1$ given in (6.5). We also have $\psi_2 \leq 2\pi$ and $c = \nu_2(\mathbb{S}^1 \times \mathbb{S}^1) < \pi$.

The triple integral in (6.5) may be simplified by changing the variable $y = \frac{1}{4} x^2 z$ and then integrating over x and t. Then it becomes the one-dimensional integral

$$\psi_2(\alpha) = \frac{2}{\alpha} \int_0^{\infty} \Big(\log(1 + z/2) - \log(1 + (1-\alpha)z/2) \Big) \frac{dz}{z\sqrt{1+z}}.$$

7. Differentiation on the sphere

Before turning to Theorem 1.2 and applications of the spherical covariance representation, let us recall a few basic formulas related to differentiation on the sphere.

The spherical gradient $w = \nabla_{\mathbb{S}} f(\theta)$ of a smooth function f on $\mathbb{S}^{n-1}$ may be defined as the shortest vector in $\mathbb{R}^n$ satisfying the Taylor formula (1.3). If f is defined and smooth on the whole space $\mathbb{R}^n$ or a smaller open subset G containing the unit sphere, its spherical gradient at a point $\theta \in \mathbb{S}^{n-1}$ is related to the usual gradient $\nabla f(\theta)$ by

$$\nabla_{\mathbb{S}} f(\theta) = \nabla f(\theta) - \langle \nabla f(\theta), \theta \rangle\, \theta = P_{\theta^\perp} \nabla f(\theta), \tag{7.1}$$

where $P_{\theta^\perp}$ denotes the projection operator in $\mathbb{R}^n$ onto the orthogonal complement of the line containing θ (the shifted tangent space). Alternatively, one may start with a smooth function $f : \mathbb{S}^{n-1} \to \mathbb{R}^n$ and extend it to $G = \mathbb{R}^n \setminus \{0\}$, for example, by the formula

$$u(x) = f(r^{-1}x), \quad r = |x|. \tag{7.2}$$

Then we obtain a smooth function on G with

$$\nabla u(x) = r^{-1}\, \nabla_{\mathbb{S}} f(r^{-1}x), \tag{7.3}$$

which coincides with the spherical gradient when it is restricted to the sphere.

From (7.1) it follows that the spherical derivative $\nabla_{\mathbb{S}}$ is a linear operator satisfying the usual rule of differentiation of products

$$\nabla_{\mathbb{S}}(fg) = f\nabla_{\mathbb{S}}g + g\nabla_{\mathbb{S}}f. \tag{7.4}$$

Here and in the sequel, we endow the real linear space of $n \times n$ matrices with the standard inner product

$$\langle A, B \rangle = \sum_{i,j=1}^{n} a_{ij} b_{ij} = \mathrm{Tr}(AB), \quad A = (a_{ij})_{i,j=1}^{n}, \ B = (a_{ij})_{i,j=1}^{n},$$

(where Tr means the trace). The associated Hilbert-Schmidt norm is

$$\|A\|_{\mathrm{HS}} = \sqrt{\langle A, A \rangle} = \Big(\sum_{i,j=1}^{n} a_{ij}^2 \Big)^{1/2}.$$

Given a C^2-smooth function f on $\mathbb{S}^{n-1}$, its second derivative or the Hessian $B = f_S''(\theta)$ at the point $\theta \in \mathbb{S}^{n-1}$ may be defined intrinsically using the Taylor formula

$$
\begin{aligned}
f(\theta') - f(\theta) &= \langle \nabla_{\mathbb{S}} f(\theta), \theta' - \theta \rangle \\
&\quad + \frac{1}{2} \langle B(\theta' - \theta), \theta' - \theta \rangle + o(|\theta' - \theta|^2), \quad \theta' \to \theta \ (\theta' \in \mathbb{S}^{n-1}),
\end{aligned}
\tag{7.5}
$$

where B is required to be a symmetric $n \times n$ matrix with the smallest Hilbert-Schmidt norm. One can show that, for all $v \in \mathbb{R}^n$ and $\theta \in \mathbb{S}^{n-1}$,

$$f_{\mathbb{S}}''(\theta) v = \nabla_{\mathbb{S}} \langle \nabla_{\mathbb{S}} f(\theta), v \rangle + \langle v, \theta \rangle \nabla_{\mathbb{S}} f(\theta), \tag{7.6}$$

which means that the second order derivative represents the double application of the spherical differentiation, however, for v in the tangent space, only.

For a matrix description in terms of (any) smooth extension u of f from the sphere to its neighborhood, we have

$$f_{\mathbb{S}}''(\theta) = P_{\theta^{\perp}} A P_{\theta^{\perp}}, \quad A = u''(\theta) - \langle \nabla u(\theta), \theta \rangle I_n, \tag{7.7}$$

where I_n denotes the identity $n \times n$ matrix, and where u'' is the usual (Euclidean) Hessian of u, that is, the matrix of second order partial derivatives of u (cf. e.g. [4]).

In view of the orthogonality of $\nabla_{\mathbb{S}} f(\theta)$ and θ, we always have $f_{\mathbb{S}}''(\theta) \theta = 0$ (since $P_{\theta^{\perp}} \theta = 0$). Hence, by the symmetry of the second spherical derivative,

$$\langle f_{\mathbb{S}}''(\theta) h, \theta \rangle = \langle h, f_{\mathbb{S}}''(\theta) \theta \rangle = 0$$

for any $h \in \mathbb{R}^n$. Hence the image $f_{\mathbb{S}}''(\theta) h$ lies in $\theta^{\perp}$.

Let us relate the usual second derivative of the classical extension u as in (7.2) to the spherical derivative $f_{\mathbb{S}}''$. By (7.3), $\langle \nabla u(\theta), \theta \rangle = 0$, so that $A = u''(\theta)$ in (7.7). Hence, we get:

Proposition 7.1. *If u is the 0-homogeneous extension of a C^2-smooth function f on $\mathbb{S}^{n-1}$ defined in (7.2), then*

$$f_{\mathbb{S}}''(\theta) = P_{\theta^{\perp}} u''(\theta) P_{\theta^{\perp}}, \quad \theta \in \mathbb{S}^{n-1}.$$

To argue in the opposite direction from f to u, fix a point $x \in \mathbb{R}^n$, $x \neq 0$, and put

$$\theta' = \frac{x + h}{|x + h|}, \quad \theta = \frac{x}{|x|} = r^{-1} x, \quad r = |x|,$$

with $h \in \mathbb{R}^n$ small enough. Using Taylor's formula, it is easy to check that

$$
\begin{aligned}
\theta' - \theta \;=\; & r^{-1}h - r^{-3} \langle x, h \rangle \, x - r^{-3} \langle x, h \rangle \, h \\
& - \frac{1}{2}\, r^{-3} |h|^2 \, x + \frac{3}{2}\, r^{-5} \langle x, h \rangle^2 \, x + O(|h|^3).
\end{aligned}
$$

In particular, $|\theta' - \theta| = O(|h|)$. Since $\langle \nabla_{\mathbb{S}} f(\theta), x \rangle = r \langle \nabla_{\mathbb{S}} f(\theta), \theta \rangle = 0$, we get

$$
\langle \nabla_{\mathbb{S}} f(\theta), \theta' - \theta \rangle = r^{-1} \langle \nabla_{\mathbb{S}} f(\theta), h \rangle - r^{-2} \langle \nabla_{\mathbb{S}} f(\theta), h \rangle \langle \theta, h \rangle + O(|h|^3).
$$

Recall that the matrix $B = f_{\mathbb{S}}''(\theta)$ satisfies $Bx = rB\theta = 0$. Using a shorter expansion

$$
\theta' - \theta = r^{-1}h - r^{-3} \langle x, h \rangle \, x + O(|h|^2),
$$

we also have $B(\theta' - \theta) = r^{-1}Bh + O(|h|^2)$ and therefore

$$
\langle B(\theta' - \theta), \theta' - \theta \rangle = r^{-2} \langle Bh, h \rangle + O(|h|^3).
$$

Hence, by Taylor's expansion (7.5),

$$
\begin{aligned}
u(x + h) - u(x) \;=\; & f(\theta') - f(\theta) \\
=\; & r^{-1} \langle \nabla_{\mathbb{S}} f(\theta), h \rangle - r^{-2} \langle \nabla_{\mathbb{S}} f(\theta), h \rangle \langle \theta, h \rangle + \frac{1}{2}\, r^{-2} \langle Bh, h \rangle + o(|h|^2).
\end{aligned}
$$

Thus, according to the usual Taylor expansion for the function $u(x)$ up to the quadratic term, necessarily $\nabla u(x) = r^{-1} \nabla_{\mathbb{S}} f(\theta)$ and, for all $h \in \mathbb{R}^n$,

$$
\langle u''(x)h, h \rangle = -2r^{-2} \langle \nabla_{\mathbb{S}} f(\theta), h \rangle \langle \theta, h \rangle + r^{-2} \langle Bh, h \rangle. \tag{7.8}
$$

To give an equivalent matrix description, one may use the symmetrized tensor product.

Definition 7.2. Given vectors $v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n)$ in $\mathbb{R}^n$, the symmetrized tensor product $v \otimes w$ is an $n \times n$ symmetric matrix with entries

$$
(v \otimes w)_{ij} = \frac{1}{2}\, (v_i w_j + w_i v_j), \quad 1 \le i, j \le n.
$$

Note that

$$
\langle v, h \rangle \langle w, h \rangle = \sum_{i,j=1}^{n} v_i \, w_j h_i h_j = \langle (v \otimes w)h, h \rangle, \quad h = (h_1, \ldots, h_n) \in \mathbb{R}^n.
$$

Using this equality, from (7.8) we obtain:

Proposition 7.3. *The extension $u(x)$ in (7.2) of a C^2-smooth function $f(\theta)$ on $\mathbb{S}^{n-1}$ has the matrix of second derivatives*

$$
u''(x) = r^{-2} \left[f_{\mathbb{S}}''(\theta) - 2 \nabla_{\mathbb{S}} f(\theta) \otimes \theta \right], \quad \theta = r^{-1}x, \ x \ne 0.
$$

In particular, on the unit sphere

$$
f_{\mathbb{S}}''(\theta) = u''(\theta) + 2 \nabla_{\mathbb{S}} f(\theta) \otimes \theta.
$$

Example. For the linear function $f(\theta) = \langle v, \theta \rangle$ with a fixed $v \in \mathbb{R}^n$, we have, by (7.1),

$$
\nabla_{\mathbb{S}} f(\theta) = P_{\theta^\perp} \nabla f(\theta) = v - \langle v, \theta \rangle \, \theta.
$$

Using the extension (7.2) and applying (7.6) or (7.7) together with Proposition 7.3, one can show that

$$f''_{\mathbb{S}}(\theta) = -\langle v, \theta \rangle P_{\theta^\perp} = \langle v, \theta \rangle [\theta \otimes \theta - I_n].$$

8. Spherical Laplacian

Here we collect basic formulas related to the spherical Laplacian with reference to [1, 3, 4] for more details and proofs. This operator is defined on the class of all C^2-smooth functions f on $\mathbb{S}^{n-1}$ by the equality

$$\Delta_{\mathbb{S}} f = \operatorname{Tr} f''_{\mathbb{S}}.$$

Introduce the "spherical partial derivatives" $D_i f(\theta) = \langle \nabla_{\mathbb{S}} f(\theta), e_i \rangle$, where $e_1, \dots, e_n$ is the canonical basis in $\mathbb{R}^n$, so that

$$\nabla_{\mathbb{S}} f(\theta) = \sum_{i=1}^n D_i f(\theta)\, e_i.$$

As further partial derivatives, one may define the "second order" differential operators

$$D_{ij} f = D_i(D_j f) = \langle \nabla_{\mathbb{S}} \langle \nabla_{\mathbb{S}} f, e_j \rangle, e_i \rangle, \quad i, j = 1, \dots, n$$

(note that the identity $D_{ij} f = D_{ji} f$ is no longer true in the entire class C^2). Then

$$\Delta_{\mathbb{S}} = \sum_{i=1}^n D_{ii}.$$

In fact, any orthonormal basis in $\mathbb{R}^n$ could be used in place of the e_i's in the definition of D_{ii}, and the above statement will continue to hold.

The next statement indicates how one may evaluate the Laplacian of a given function in terms of usual (Euclidean) derivatives.

Proposition 8.1. *If f is C^2-smooth in an open region G containing the unit sphere, then for any $\theta \in \mathbb{S}^{n-1}$,*

$$\Delta_{\mathbb{S}} f(\theta) = \Delta f(\theta) - (n-1) \langle \nabla f(\theta), \theta \rangle - \langle f''(\theta)\theta, \theta \rangle. \tag{8.1}$$

The Laplacian appears naturally in the classical formula for spherical integration by parts,

$$\int \langle \nabla_{\mathbb{S}} f, \nabla_{\mathbb{S}} g \rangle\, d\sigma_{n-1} = - \int f \Delta_{\mathbb{S}} g\, d\sigma_{n-1}, \tag{8.2}$$

which may be taken as an equivalent definition of $\Delta_{\mathbb{S}}$. It yields the following characterization of the orthogonality to linear functions in terms of the spherical gradient.

Proposition 8.2. *For any C^1-smooth function f on $\mathbb{S}^{n-1}$,*

$$\int f(\theta)\theta\, d\sigma_{n-1}(\theta) = \frac{1}{n-1} \int \nabla_{\mathbb{S}} f(\theta)\, d\sigma_{n-1}(\theta).$$

In particular, f is orthogonal to all linear functions in $L^2(\mathbb{S}^{n-1})$ if and only if the linear forms $\langle \nabla_{\mathbb{S}} f(\theta), v \rangle$ have σ_{n-1}-mean zero for any $v \in \mathbb{R}^n$.

Applying the rule (7.4), we have another identity

$$\frac{1}{2}\left[\Delta_{\mathbb{S}}(fg) - f\Delta_{\mathbb{S}}g - g\Delta_{\mathbb{S}}f\right] = \langle\nabla_{\mathbb{S}}f, \nabla_{\mathbb{S}}g\rangle \tag{8.3}$$

in the class of all C^2-smooth functions f and g on the sphere. For the proof, it is sufficient to multiply this equality by an arbitrary smooth function h and integrate according to (8.2).

It is well-known (cf. e.g. [16]) that the Hilbert space $L^2(\mathbb{S}^{n-1}, \sigma_{n-1})$ can be decomposed into the sum of orthogonal linear subspaces H_d, $d = 0, 1, 2, \ldots$, consisting of all d-homogeneous harmonic polynomials (more precisely – restrictions of such polynomials to the sphere). In particular, H_0 is the space of constant functions, H_1 is the space of linear functions, H_2 consists of quadratic harmonics, and so on. Any element f_d of H_d represents an eigenfunction of the Laplacian with the eigenvalue $-d(n+d-2)$. That is, any function f in $L^2(\mathbb{S}^{n-1}, \sigma_{n-1})$ admits a unique orthogonal (Fourier) expansion in spherical harmonics,

$$f = \sum_{d=0}^{\infty} f_d \quad (f_d \in H_d), \tag{8.4}$$

and $\Delta_{\mathbb{S}}f_d = -d(n + d - 2)\,f_d$. Hence, if f is C^2-smooth, we get another representation for the Laplacian,

$$\Delta_{\mathbb{S}}f = -\sum_{d=1}^{\infty} d(n + d - 2)f_d. \tag{8.5}$$

9. Semi-group approach to the covariance

The heat semi-group $P_t = e^{t\Delta_{\mathbb{S}}}$, $t \geq 0$, on the sphere $\mathbb{S}^{n-1}$ has the generator $\Delta_{\mathbb{S}}$, so that

$$\frac{d}{dt}\,P_t f = \Delta_{\mathbb{S}}P_t f, \quad t > 0. \tag{9.1}$$

For every $t > 0$ and $f \in L^2(\sigma_{n-1})$, the operator P_t assigns a C^∞-smooth function $P_t f$, which can be defined using the orthogonal decomposition (8.4) into spherical harmonics by

$$P_t f = \sum_{d=0}^{\infty} e^{-d(n+d-2)t}\,f_d. \tag{9.2}$$

Let us list a few properties of this semi-group. To simplify the notations, in this section the gradients $\nabla = \nabla_{\mathbb{S}}$ and the Laplacian $\Delta = \Delta_{\mathbb{S}}$ are understood in the spherical sense.

1) (commutativity) $\Delta P_t f = P_t \Delta f$ for any C^2-smooth f on $\mathbb{S}^{n-1}$.

2) $\int (P_t f)\, g \, d\sigma_{n-1} = \int f\, P_t g \, d\sigma_{n-1}$.

3) $P_t f \to \int f \, d\sigma_{n-1}$ as $t \to \infty$.

4) (Jensen's inequality) $\Psi(P_t f) \leq P_t \Psi(f)$ for any convex function Ψ.

There is another important property related to the geometry of the sphere known as the curvature-dimension condition (for details we refer an interested reader to [1]). According to (8.3), for the generator $L = \Delta$, the carré du champ operator is given by

$$\Gamma(f, g) = \frac{1}{2}\left[L(fg) - fLg - gLf\right] = \langle\nabla f, \nabla g\rangle.$$

Analogously, one may speak of an iteration of Γ

$$\Gamma_2(f, g) = \frac{1}{2} \left[L\Gamma(f, g) - \Gamma(f, Lg) - \Gamma(g, Lf) \right].$$

In general, a semi-group is said to satisfy the curvature dimension condition $CD(\rho, n)$ with constant ρ if

$$\Gamma_2(f) \geq \rho \, \Gamma(f) + \frac{1}{n} \, (Lf)^2,$$

where $\Gamma(f) = \Gamma(f, f)$. It is a classic result that the heat semi-group on $\mathbb{S}^{n-1}$ satisfies the curvature condition $CD(n - 2, n - 1)$, so that

$$\Gamma_2(f) \geq (n - 2) \, \Gamma(f) + \frac{1}{n - 1} \, (Lf)^2.$$

A general characterization commonly known as the Bakry-Emery condition indicates that

$$CD(\rho, \infty) \iff \sqrt{\Gamma(P_t f)} \leq e^{-\rho t} P_t \sqrt{\Gamma(f)} \quad \text{for all } t \geq 0.$$

Hence, the heat semi-group on the sphere satisfies

$$\sqrt{\Gamma(P_t f)} \leq e^{-(n-2)t} P_t \sqrt{\Gamma(f)},$$

that is, we have a pointwise bound

$$|\nabla P_t f| \leq e^{-(n-2)t} P_t |\nabla f|. \tag{9.3}$$

These properties are sufficient to derive a covariance representation in terms of the semi-group operators. We start with a variance representation.

Lemma 9.1. *For any smooth function f on $\mathbb{S}^{n-1}$ $(n \geq 3)$,*

$$\mathrm{Var}_{\sigma_{n-1}}(f) = \int_0^\infty \int_{\mathbb{S}^{n-1}} \langle \nabla P_t f, \nabla f \rangle \, d\sigma_{n-1} \, dt. \tag{9.4}$$

Proof. We may assume that f is C^2-smooth and has mean $P_\infty f = \int_{\mathbb{S}^{n-1}} f \, d\sigma_{n-1} = 0$. Note that the double integral in (9.4) is absolutely convergent due to the exponential decay in (9.3) with respect to t. Using $P_0 f = f$ and applying (9.1) together with properties 3) and 1), we can write

$$f^2 = -\int_0^\infty \frac{d}{dt} \, (P_t f)^2 \, dt = -2 \int_0^\infty P_t f \, P_t \Delta f \, dt.$$

Integrating this equality over σ_{n-1} and applying 2) together with Fubini's theorem, we get

$$\begin{aligned}
\mathrm{Var}_{\sigma_{n-1}}(f) &= -2 \int_0^\infty \left[\int_{\mathbb{S}^{n-1}} P_{2t} f \, \Delta f \, d\sigma_{n-1} \right] dt \\
&= -\int_0^\infty \left[\int_{\mathbb{S}^{n-1}} P_t f \, \Delta f \, d\sigma_{n-1} \right] dt \\
&= \int_0^\infty \left[\int_{\mathbb{S}^{n-1}} \langle \nabla P_t f, \nabla f \rangle \, d\sigma_{n-1} \right] dt,
\end{aligned}$$

where in the last step we employed the formula (8.2) for spherical integration by parts. $\square$

Proof of Theorem 1.2. In view of the identity

$$2\operatorname{cov}_{\sigma_{n-1}}(f, g) = \operatorname{Var}_{\sigma_{n-1}}(f + g) - \operatorname{Var}_{\sigma_{n-1}}(f) - \operatorname{Var}_{\sigma_{n-1}}(g), \tag{9.5}$$

it makes sense to apply Lemma 8.1 to f, g and $f + g$. By linearity of the semi-group and spherical gradient,

$$\begin{aligned}
\langle \nabla P_t(f + g), \nabla(f + g) \rangle &= \langle \nabla P_t f, \nabla f \rangle + \langle \nabla P_t g, \nabla g \rangle \\
&\quad + \langle \nabla P_t f, \nabla g \rangle + \langle \nabla P_t g, \nabla f \rangle.
\end{aligned}$$

Hence, after integration over $d\sigma_{n-1}$ and dt, (8.4) yields

$$\begin{aligned}
\operatorname{Var}_{\sigma_{n-1}}(f + g) &= \operatorname{Var}_{\sigma_{n-1}}(f) + \operatorname{Var}_{\sigma_{n-1}}(g) + \int_0^\infty \left[\int_{\mathbb{S}^{n-1}} \langle \nabla P_t f, \nabla g \rangle \, d\sigma_{n-1} \right] dt \\
&\quad + \int_0^\infty \left[\int_{\mathbb{S}^{n-1}} \langle \nabla P_t g, \nabla f \rangle \, d\sigma_{n-1} \right] dt.
\end{aligned}$$

Here the last inner integral is equal to

$$\begin{aligned}
- \int_{\mathbb{S}^{n-1}} f \, \Delta P_t g \, d\sigma_{n-1} &= - \int_{\mathbb{S}^{n-1}} f \, P_t \Delta g \, d\sigma_{n-1} = - \int_{\mathbb{S}^{n-1}} P_t f \, \Delta g \, d\sigma_{n-1} \\
&= \int_{\mathbb{S}^{n-1}} \langle \nabla P_t f, \nabla g \rangle \, d\sigma_{n-1}.
\end{aligned}$$

Thus

$$\operatorname{Var}_{\sigma_{n-1}}(f + g) = \operatorname{Var}_{\sigma_{n-1}}(f) + \operatorname{Var}_{\sigma_{n-1}}(g) + 2 \int_0^\infty \left[\int_{\mathbb{S}^{n-1}} \langle \nabla P_t f, \nabla g \rangle \, d\sigma_{n-1} \right] dt.$$

It remains to apply (8.5) which leads to the desired representation

$$\operatorname{cov}_{\sigma_{n-1}}(f, g) = \int_0^\infty \left[\int_{\mathbb{S}^{n-1}} \langle \nabla P_t f, \nabla g \rangle \, d\sigma_{n-1} \right] dt. \tag{9.6}$$

$\square$

10. Comparison of the two representations

In the Gauss space there is a similar heat semi-group description of the covariance functional

$$\operatorname{cov}_{\gamma_n}(u, v) = \int_0^\infty \left[\int_{\mathbb{R}^n} \langle \nabla T_t u, \nabla v \rangle \, d\gamma_n \right] dt, \tag{10.1}$$

where

$$T_t u(x) = \int_{\mathbb{R}^n} u\big(e^{-t} x + \sqrt{1 - e^{-2t}}\, y\big) \, d\gamma_n(y), \quad x \in \mathbb{R}^n, \ t \geq 0, \tag{10.2}$$

denote the Ornstein-Uhlenbeck operators. The identity (10.1) can be obtained with similar arguments used in the proof of Theorem 1.2. This certainly begs the following question: May one start from the Gaussian covariance representation (10.1) and obtain the semi-group representation (9.6) of the covariance on the sphere? This can be indeed done by applying (10.1) to functions of the form $u(x) = f(\frac{x}{|x|}) = f(\theta)$, $v(x) = g(\frac{x}{|x|}) = g(\theta)$ by means of orthogonal polynomials. All of the following on Chebyshev-Hermite polynomials can be found in the detailed survey by Bogachev on the Ornstein-Uhlenbeck operators [6].

In dimension one, the normalized Chebyshev-Hermite polynomial of degree $m \geq 0$ is defined as

$$H_m(x) = \frac{(-1)^m}{\sqrt{m!}} e^{x^2/2} \frac{d^m}{dx^m} e^{-x^2/2}.$$

Similarly, given a multi-index $m = (k_1, \ldots, k_n)$ with integer components $k_i \geq 0$, the corresponding Chebyshev-Hermite polynomial of degree $k = |m| = k_1 + \cdots + k_n$ is defined by

$$H_m(x) = \prod_{i=1}^{n} H_{k_i}(x_i), \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

These polynomials constitute an orthonormal basis of $L^2(\gamma_n)$.

The Ornstein-Uhlenbeck operator (10.2) has the following representation

$$T_t u = \sum_{k=0}^{\infty} e^{-kt} I_k(u), \tag{10.3}$$

where I_k denotes the projection on the space spanned by the $\begin{pmatrix} k+n-1 \\ k \end{pmatrix}$ multi-index Chebyshev-Hermite polynomials H_m of degree k.

They also serve as eigenfunctions for the generator $L = \Delta - x \cdot \nabla$ of the semi-group $(T_t)_{t \geq 0}$, and more precisely,

$$-L H_m(x) = k H_m(x). \tag{10.4}$$

Let us mention the advantage presented by an examination of the Ornstein-Uhlenbeck operators from this angle. The domain of the generator is given as follows

$$\mathcal{D}(L) = \left\{ u \in L^2(\gamma_n) : \sum_{k=0}^{\infty} k^2 \, \|I_k(u)\|_{L^2(\gamma_n)} < \infty \right\}.$$

Turning to the heat semi-group on $\mathbb{S}^{n-1}$, recall that f in $L^2(\sigma_{n-1})$ admits a unique orthogonal decomposition $f = \sum_{d=0}^{\infty} f_d$ over spherical harmonics of degree d. We will need the following technical result whose proof we leave for an interested reader as an exercise.

Lemma 10.1. *Let f and g be in $L^2(\sigma_{n-1})$. For the spherical harmonics f_d, define their extensions $u_d(x) = f_d(\frac{x}{|x|}) = f_d(\theta)$ and let $v(x) = g(\frac{x}{|x|}) = g(\theta)$ $(x \in \mathbb{R}^n, x \neq 0)$. Then*

$$\int_0^{\infty} \left[\int_{\mathbb{R}^n} v \sum_{k=0}^{\infty} k e^{-kt} \sum_{d=0}^{\infty} I_k(u_d) \, d\gamma_n \right] dt$$

$$= \sum_{d=0}^{\infty} \left[d(d+n-2) \, e^{-d(d+n-2)t} \int_0^{\infty} \int_{\mathbb{S}^{n-1}} g f_d \, d\sigma_{n-1} \right] dt.$$

Applying this result, let us show that

$$\mathrm{cov}_{\sigma_{n-1}}(f, g) = \int_0^{\infty} \int_{\mathbb{S}^{n-1}} \langle \nabla_{\mathbb{S}} P_t f, \nabla_{\mathbb{S}} g \rangle \, d\sigma_{n-1} \, dt.$$

Integrating by parts with respect to the Gaussian measure, we may rewrite (10.1) as

$$\mathrm{cov}_{\gamma_n}(u, v) = -\int_0^{\infty} \int_{\mathbb{R}^n} v L(T_t u) \, d\gamma_n \, dt.$$

Using a decomposition of u into Chebyshev-Hermite polynomials and applying (8.3)-(8.4), we see that the last double integral is equal to

$$-\int_0^\infty \int_{\mathbb{R}^n} v L\Big(\sum_{k=0}^\infty e^{-kt} I_k(u)\Big)\, d\gamma_n\, dt \;=\; -\int_0^\infty \int_{\mathbb{R}^n} v \sum_{k=0}^\infty e^{-kt} L(I_k(u))\, d\gamma_n\, dt$$

$$=\; \int_0^\infty \int_{\mathbb{R}^n} v \sum_{k=0}^\infty k e^{-kt}\, I_k(u)\, d\gamma_n\, dt.$$

Next, using the decomposition of u into spherical harmonics and applying Lemma 10.1 together with the orthogonal decomposition (8.5) for the Laplacian, the latter expression is

$$\int_0^\infty \int_{\mathbb{R}^n} v \sum_{k=0}^\infty k e^{-kt} I_k\Big(\sum_{d=0}^\infty u_d\Big)\, d\gamma_n\, dt \;=\; \int_0^\infty \int_{\mathbb{R}^n} v \sum_{k=0}^\infty k e^{-kt} \sum_{d=0}^\infty I_k(u_d)\, d\gamma_n\, dt$$

$$=\; \sum_{d=0}^\infty \Big[d(d+n-2)\, e^{-d(d+n-2)t} \int_0^\infty \int_{\mathbb{S}^{n-1}} g f_d\, d\sigma_{n-1} \Big] dt$$

$$=\; -\int_0^\infty \int_{\mathbb{S}^{n-1}} g \sum_{d=0}^\infty e^{-d(d+n-2)t} \Delta_{\mathbb{S}} f_d\, d\sigma_{n-1}\, dt$$

$$=\; -\int_0^\infty \int_{\mathbb{S}^{n-1}} g \,\Delta_{\mathbb{S}} \sum_{d=0}^\infty e^{-d(d+n-2)t} f_d\, d\sigma_{n-1}\, dt$$

$$=\; -\int_0^\infty \int_{\mathbb{S}^{n-1}} g \,\Delta_{\mathbb{S}} P_t f\, d\sigma_{n-1}\, dt$$

$$=\; \int_0^\infty \int_{\mathbb{S}^{n-1}} \langle \nabla_{\mathbb{S}} P_t f, \nabla_{\mathbb{S}} g\rangle\, d\sigma_{n-1}\, dt.$$

11. Applications to spherical concentration

We now return to Theorem 1.1 and develop several applications of the covariance identity

$$\mathrm{cov}_{\sigma_{n-1}}(f,g) = c_n \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y)\rangle\, d\mu_n(x,y). \tag{11.1}$$

All of them are entirely based on the property that the probability measure μ_n on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ (explicitly described in Theorem 4.1) has marginals σ_{n-1}, while the constants satisfy

$$\frac{1}{n-1} < c_n < \frac{1}{n-2} \quad (n \geq 3).$$

As a particular case, (11.1) implies the Poincaré-type inequality

$$\mathrm{Var}_{\sigma_{n-1}}(f) \leq c_n \int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} f|^2\, d\sigma_{n-1}$$

with an asymptotically correct constant (the optimal one is $\frac{1}{n-1}$). More generally, a similar bound can be given in terms of L^p-norms of gradients

$$\|\nabla_{\mathbb{S}} f\|_p = \Big(\int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} f|^p\, d\sigma_{n-1} \Big)^{1/p}, \quad 1 \leq p \leq \infty,$$

where the case $p = \infty$ corresponds to the maximum (for smooth f).

Corollary 11.1. *Let $p, q \geq 1$, $\frac{1}{p} + \frac{1}{q} = 1$, For all smooth functions f and g on $\mathbb{S}^{n-1}$,*

$$|\mathrm{cov}_{\sigma_{n-1}}(f, g)| \leq c_n \, \|\nabla_{\mathbb{S}} f\|_p \, \|\nabla_{\mathbb{S}} g\|_q. \tag{11.2}$$

In particular, if f has a Lipschitz semi-norm $\|f\|_{\mathrm{Lip}} \leq 1$, then

$$|\mathrm{cov}_{\sigma_{n-1}}(f, g)| \leq c_n \int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} g| \, d\sigma_{n-1}. \tag{11.3}$$

Proof. By Hölder's inequality, the absolute value of the double integral in (10.1) does not exceed

$$\int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} f(x)| \, |\nabla_{\mathbb{S}} g(y)| \, d\mu_n(x, y) \quad \leq$$

$$\left(\int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} f(x)|^p \, d\mu_n(x, y) \right)^{1/p} \left(\int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} g(y)|^q \, d\mu_n(x, y) \right)^{1/q} =$$

$$\left(\int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} f(x)|^p \, d\sigma_{n-1}(x) \right)^{1/p} \left(\int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} g(y)|^q \, d\sigma_{n-1}(y) \right)^{1/q} = \|\nabla_{\mathbb{S}} f\|_p \, \|\nabla_{\mathbb{S}} g\|_q.$$

The inequality (11.3) corresponds to (11.2) with $p = \infty$, $q = 1$. $\qquad\square$

The representation (11.1) may also be used to recover the spherical concentration.

Corollary 11.2. *For any smooth function f on $\mathbb{S}^{n-1}$ with σ_{n-1}-mean zero,*

$$\int_{\mathbb{S}^{n-1}} e^f \, d\sigma_{n-1} \leq \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \exp\left\{ c_n \left\langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \right\rangle \right\} d\mu_n(x, y). \tag{11.4}$$

In particular,

$$\int_{\mathbb{S}^{n-1}} e^f \, d\sigma_{n-1} \leq \int_{\mathbb{S}^{n-1}} e^{c_n |\nabla_{\mathbb{S}} f|^2} \, d\sigma_{n-1}. \tag{11.5}$$

Proof. The argument involves the entropy functional

$$\mathrm{Ent}(\xi) = \mathbb{E}\,\xi \log \xi - \mathbb{E}\,\xi \log \mathbb{E}\,\xi.$$

It is well-defined and non-negative for every random variable $\xi \geq 0$, and is known to admit the variational description

$$\mathrm{Ent}(\xi) = \sup \mathbb{E}\,\xi \eta,$$

where the supremum is taken over all random variables η such that $\mathbb{E}\,e^\eta \leq 1$. In particular,

$$\mathbb{E}\,e^\eta = 1 \implies \mathbb{E}\,\xi\eta \leq \mathrm{Ent}(\xi). \tag{11.6}$$

We consider the expectations (integrals) and entropy on the probability spaces $(\mathbb{S}^{n-1}, \sigma_{n-1})$ and $(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}, \mu_n)$ and write correspondingly $\mathbb{E}_{\sigma_{n-1}}\xi$, $\mathrm{Ent}_{\sigma_{n-1}}(\xi)$, and $\mathbb{E}_{\mu_n}\xi$, $\mathrm{Ent}_{\mu_n}(\xi)$.

Define the constant β as the logarithm of the right-hand side in (11.4), so that, by (11.6),

$$\mathbb{E}_{\mu_n}\left[\left(c_n \left\langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \right\rangle - \beta \right) \xi \right] \leq \mathrm{Ent}_{\mu_n}(\xi) \tag{11.7}$$

for any bounded measurable function $\xi(x, y)$ on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$. Choosing $\xi(x, y) = g(y) = e^{f(y)}$, we have $\mathrm{Ent}_{\mu_n}(\xi) = \mathrm{Ent}_{\sigma_{n-1}}(g)$, and (11.1) together with (11.7) give

$$
\begin{aligned}
\mathbb{E}_{\sigma_{n-1}} f e^f - \beta \, \mathbb{E}_{\sigma_{n-1}} e^f &= \mathrm{cov}_{\sigma_{n-1}}(f, e^f) - \beta \, \mathbb{E}_{\sigma_{n-1}} e^f \\
&= c_n \, \mathbb{E}_{\mu_n} \langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle \, e^{f(y)} - \beta \, \mathbb{E}_{\mu_n} e^f \\
&\leq \mathrm{Ent}_{\mu_n}(e^f) = \mathrm{Ent}_{\sigma_{n-1}}(e^f) \\
&= \mathbb{E}_{\sigma_{n-1}}(f e^f) - \mathbb{E}_{\sigma_{n-1}} e^f \log \mathbb{E}_{\sigma_{n-1}} e^f.
\end{aligned}
$$

Hence $\beta \geq \log \mathbb{E}_{\sigma_{n-1}} e^f$, which is the relation (11.4). As for (11.5), it follows from (11.4) using

$$
\langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle \leq \frac{1}{2} |\nabla_{\mathbb{S}} f(x)|^2 + \frac{1}{2} |\nabla_{\mathbb{S}} f(y)|^2
$$

and applying the Cauchy inequality. $\qquad\square$

Proof of Theorem 1.3. The argument is based on the relation (11.2), where the smoothness condition on f may be removed, just keeping the assumption $\|f\|_{\mathrm{Lip}} \leq 1$. We may further assume that f has σ_{n-1}-mean zero. Consider $g = T(f)$, where T is a non-decreasing (piecewise) differentiable function. Then from (11.2) we get

$$
\mathbb{E}_{\sigma_{n-1}} f T(f) \leq c_n \, \mathbb{E}_{\sigma_{n-1}} T'(f).
$$

Without loss of generality, assume that f has a continuous density p under the measure σ_{n-1}. Let $F(x) = \sigma_{n-1}\{f \leq x\}$ denote the distribution function of f. Choosing $T(x) = \min((x - r)^+, \varepsilon)$ with parameters $r > 0$ and $\varepsilon > 0$, the above inequality becomes

$$
\int_r^{r+\varepsilon} x(x - r) \, dF(x) + \varepsilon \int_{r+\varepsilon}^{\infty} x \, dF(x) \leq c_n \, (F(r + \varepsilon) - F(r)).
$$

Dividing by ε and letting ε tend to zero, we obtain that for all $r > 0$,

$$
V(r) \equiv \int_r^{\infty} x p(x) \, dx = \int_r^{\infty} x \, dF(x) \leq c_n p(r).
$$

Hence, the function V satisfies the differential inequality $V(r) \leq -c_n V'(r)/r$, that is,

$$
(\log V(r))' \leq (-r^2/2c_n)'.
$$

This is the same as saying that $\log V(r) + r^2/2c_n$ is non-increasing, and so is the function $V(r) \exp(r^2/2c_n)$. Since $V(0) = \mathbb{E} f^+$, we get

$$
\mathbb{E}_{\sigma_{n-1}} f^+ \geq \exp(c_n r^2/2) \int_r^{\infty} x p(x) \, dx \geq \exp(r^2/2c_n) \, r(1 - F(r)).
$$

Thus,

$$
\sigma_{n-1}\{f \geq r\} \leq \frac{1}{r} e^{-r^2/2c_n} \, \mathbb{E}_{\sigma_{n-1}} f^+.
$$

A similar inequality holds when replacing f with $-f$, that is,

$$
\sigma_{n-1}\{f \leq -r\} \leq \frac{1}{r} e^{-r^2/2c_n} \, \mathbb{E}_{\sigma_{n-1}} f^-.
$$

Summing the two bounds, we arrive at (1.8). $\qquad\square$

12. **Second order covariance identity in Gauss space**

For short, we write $\mathbb{E}_{\gamma_n}$ to denote the expectation, that is, the integral with respect to the Gaussian measure γ_n on $\mathbb{R}^n$.

Applying the Gaussian covariance identity (1.1) to partial derivatives of u and v of the first order, we are led to the second covariance identity

$$\mathrm{cov}_{\gamma_n}(u,v) = \langle \mathbb{E}_{\gamma_n} \nabla u, \mathbb{E}_{\gamma_n} \nabla v \rangle + \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \langle u''(x), v''(y) \rangle \, d\kappa_n(x,y), \qquad (12.1)$$

where the integrand represents the inner product of square symmetric matrices, which is the trace $\mathrm{Tr}(u''(x)\,v''(y))$ of the product of the two Hessians. Here, $2\kappa_n$ is a probability measure on $\mathbb{R}^n \times \mathbb{R}^n$ which may be defined by means of the mixture

$$\kappa_n = \int_0^1 (1-t)\, \mathsf{L}\big(X, tX + sZ\big)\, dt \qquad (12.2)$$

with convention that $s = \sqrt{1-t^2}$. This shows that $2\kappa_n$ has γ_n as marginals, similarly to the mixing measure π_n in the first order covariance representation for the Gaussian measure γ_n.

In an equivalent form, the identity (12.1) was emphasized in [10]. It may be stated in the class of all C^2-smooth functions u and v on $\mathbb{R}^n \setminus \{0\}$ such that

$$\mathbb{E}_{\gamma_n} \|u''\|_{\mathrm{HS}}^2 < \infty, \quad \mathbb{E}_{\gamma_n} \|v''\|_{\mathrm{HS}}^2 < \infty.$$

Then the expectations

$$\mathbb{E}_{\gamma_n}(u^2 + v^2) \quad \text{and} \quad \mathbb{E}_{\gamma_n}(|\nabla u|^2 + |\nabla v|^2)$$

are also finite, by the Poincaré-type inequality (2.3) for γ_n, so that (12.1) makes sense.

In addition, (12.2) implies that κ_n is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^n \times \mathbb{R}^n$ and has some density

$$q_n(x,y) = \frac{d\kappa_n(x,y)}{dx\,dy}.$$

To write it explicitly, recall that for any bounded Borel measurable function $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, changing the variable $tx + sz = y$, or $z = \frac{y-tx}{s}$, we have

$$\begin{aligned}
\mathbb{E}\, h(X, tX + sZ) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} h(x, tx + sz)\, \varphi_n(x)\varphi_n(z)\, dx\, dz \\
&= s^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} h(x,y)\, \varphi_n(x)\varphi_n\Big(\frac{y-tx}{s}\Big)\, dx\, dy,
\end{aligned}$$

and this means that

$$q_n(x,y) = \int_0^1 (1-t)s^{-n}\, \varphi_n(x)\varphi_n\Big(\frac{y-tx}{s}\Big)\, dt.$$

But

$$\varphi_n(x)\varphi_n\Big(\frac{y-tx}{s}\Big) = \frac{1}{(2\pi)^n}\, \exp\left[-\frac{|x|^2 + |y|^2 - 2t\,\langle x,y \rangle}{2s^2} \right].$$

Hence, one may complement (12.1) with the following assertion, where we write $s = \sqrt{1-t^2}$.

Lemma 12.1. *The mixing measure κ_n in the Gaussian covariance representation (12.1) has density*

$$q_n(x,y) = \frac{1}{(2\pi)^n} \int_0^1 (1-t)s^{-n} \exp\left[-\frac{|x|^2 + |y|^2 - 2t\,\langle x,y\rangle}{2s^2} \right] dt, \quad x,y \in \mathbb{R}^n.$$

Putting $x = r\theta$, $y = r'\theta'$ ($r,r' > 0$, $\theta,\theta' \in \mathbb{S}^{n-1}$), this density may be written in polar coordinates as

$$q_n(r\theta, r'\theta') = \frac{1}{(2\pi)^n} \int_0^1 (1-t)s^{-n} \exp\left[-\frac{r^2 + r'^2 - 2rr't\,\langle \theta,\theta'\rangle}{2s^2} \right] dt. \qquad (12.3)$$

13. Second order covariance identities on the sphere

We now develop a spherical variant of the second order covariance identity (12.1). With any C^2-smooth f on $\mathbb{S}^{n-1}$ we associate the function

$$u(x) = f(r^{-1}x) = f(\theta), \quad \text{where } r = |x|,\ \theta = r^{-1}x,$$

which is defined and C^2-smooth in $\mathbb{R}^n \setminus \{0\}$. By (7.3) and Proposition 7.3, it has derivatives

$$\begin{aligned}
\nabla u(x) &= r^{-1}\nabla_\mathbb{S} f(\theta), \\
u''(x) &= r^{-2} Df(\theta) = r^{-2}\left(f''_\mathbb{S}(\theta) - 2\nabla_\mathbb{S} f(\theta) \otimes \theta \right)
\end{aligned}$$

(cf. Definition 7.2 for the symmetrized tensor product). Hence, as in Section 4,

$$\int \nabla u\, d\gamma_n = \mathbb{E}\, \frac{1}{|Z|} \int_{\mathbb{S}^{n-1}} \nabla_\mathbb{S} f\, d\sigma_{n-1}$$

and

$$\int \|u''\|^2_{\mathrm{HS}}\, d\gamma_n = \mathbb{E}\, \frac{1}{|Z|^4} \int_{\mathbb{S}^{n-1}} \|f''_\mathbb{S}(\theta) - 2\nabla_\mathbb{S} f(\theta) \otimes \theta\|^2_{\mathrm{HS}}\, d\sigma_{n-1},$$

where Z is a standard normal random vector in $\mathbb{R}^n$. By the smoothness assumption, the last integrand is bounded (so, the last integral is finite), while the expectation is finite for $n \geq 5$, only, in which case, by Lemma 3.1,

$$\mathbb{E}\, \frac{1}{|Z|^4} = \frac{1}{n-4}\, \mathbb{E}\, \frac{1}{|Z|^2} = \frac{1}{(n-2)(n-4)}.$$

Now, let g be another C^2-smooth function on $\mathbb{S}^{n-1}$. Define $v(y) = g(\theta')$ for $y \in \mathbb{R}^n \setminus \{0\}$, $\theta' = \frac{y}{r'}$, $r' = |y|$, so that

$$\nabla v(y) = \frac{1}{r'}\nabla_\mathbb{S} g(\theta') \quad \text{and} \quad v''(y) = \frac{1}{r'^2} Dg(\theta').$$

To simplify the resulting formulas, let us assume that f and g are orthogonal to linear functions in $L^2(\sigma_{n-1})$. This is equivalent to the similar property of u and v in $L^2(\gamma_n)$, which is the same as saying that ∇u and ∇v have γ_n-mean zero (compare with Proposition 8.2). We may now apply (12.1) together with Lemma 12.1, which give

$$\mathrm{cov}_{\sigma_{n-1}}(f,g) = \mathrm{cov}_{\gamma_n}(u,v) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle u''(x), v''(y)\rangle\, q_n(x,y)\, dx\, dy. \qquad (13.1)$$

To continue, let us integrate in (13.1) in polar coordinates separately along the x and y variables in the same manner as we did in Section 4. Then, according to (12.3), we obtain

$$\mathrm{cov}_{\sigma_{n-1}}(f,g) = \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \left\langle f_S''(\theta), g_S''(\theta') \right\rangle \psi_n(\theta, \theta') \, d\sigma_{n-1}(\theta) \, d\sigma_{n-1}(\theta') \qquad (13.2)$$

with $\psi_n(\theta, \theta')$ representable as

$$\frac{1}{2^{n-2}\,\Gamma(\frac{n}{2})^2} \int_0^1 (1-t)s^{-n} \left[\int_0^\infty \int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\,\langle \theta, \theta' \rangle}{2s^2} \right] (rr')^{n-3} \, dr\, dr' \right] dt.$$

Changing the variable, this expression is simplified to

$$\frac{1}{2^{n-2}\,\Gamma(\frac{n}{2})^2} \int_0^1 (1-t)s^{n-4} \left[\int_0^\infty \int_0^\infty \exp\left[-\frac{r^2 + r'^2 - 2rr't\,\langle \theta, \theta' \rangle}{2} \right] (rr')^{n-3} \, dr\, dr' \right] dt. \quad (13.3)$$

Let us see that ψ_n is integrable over the product measure $\sigma_{n-1} \otimes \sigma_{n-1}$ and serves as density of some finite positive measure on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ with total mass

$$c_n = \kappa_n(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}) = \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \psi_n(\theta, \theta') \, d\sigma_{n-1}(\theta) \, d\sigma_{n-1}(\theta').$$

Note that, for the finiteness of (13.3), it is necessary that the integral $\int_0^1 (1-t)s^{n-4}\,dt$ be convergent, that is, $n \geq 3$. Repeating integration in polar coordinates as before, we have

$$\mathbb{E} \int_0^1 \frac{1}{|X|^2\,|tX + sZ|^2} \, dt = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{1}{|x|^2\,|y|^2} \, q_n(x, y) \, dx\, dy = c_n,$$

where X and Z are independent standard normal random vectors in $\mathbb{R}^n$. But, for any fixed $0 < t < 1$, by Cauchy's inequality,

$$\mathbb{E} \frac{1}{|X|^2\,|tX + sZ|^2} \quad \leq \quad \left(\mathbb{E}\,|X|^{-4} \right)^{1/2} \left(\mathbb{E}\,|tX + sZ|^{-4} \right)^{1/2}$$

$$= \quad \mathbb{E}\,|X|^{-4} \quad = \quad \frac{1}{(n-2)(n-4)}.$$

Thus, $c_n < \frac{1}{(n-2)(n-4)}$ and $\kappa_n = c_n \mu_n$ for some probability measure μ_n on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$.

We have also a natural lower bound on this constant. Indeed, for a fixed value of X,

$$\mathbb{E}\,|tX + sZ|^2 = \mathbb{E}\,(t^2|X|^2 + s^2|Z|^2 + 2ts\,\langle X, Z \rangle) = t^2|X|^2 + s^2 n.$$

Hence, using $\mathbb{E}\,|X|^4 = n^2 + 2n$ and $t^2 + s^2 = 1$, we get

$$\mathbb{E}\,|X|^2\,|tX + sZ|^2 \quad = \quad \mathbb{E}_X\,|X|^2 \left(\mathbb{E}_Z\,|tX + sZ|^2 \right)$$

$$= \quad \mathbb{E}_X\,|X|^2\,(t^2|X|^2 + s^2 n) \quad = \quad n^2 + 2t^2 n \quad \leq \quad n^2 + 2n.$$

Hence, by Jensen's inequality, using the convexity of the function $1/r$ in $r > 0$,

$$c_n = \int_0^1 \mathbb{E} \frac{1}{|X|^2\,|tX + sZ|^2} \, dt \geq \int_0^1 \frac{1}{\mathbb{E}\,|X|^2|tX + sZ|^2} \, dt \geq \frac{1}{n(n+2)}.$$

Note also that $\psi_n(\theta, \theta') = \psi_n(\langle \theta, \theta' \rangle)$ depends on θ and θ' via the inner product only. This implies that the integral

$$\int_{S^{n-1}} \psi_n(\theta, \theta') \, d\sigma_{n-1}(\theta') = \int_{S^{n-1}} \psi_n(\langle \theta, \theta' \rangle) \, d\sigma_{n-1}(\theta')$$

does not depend on θ and is therefore equal to c_n. In other words, the marginal distributions of μ_n represent spherically invariant measures, hence coincide with σ_{n-1}.

As a result, we obtain the following spherical analog of (12.1), using the second order linear matrix-valued differential operator

$$Df(x) = f''_{\mathbb{S}}(x) - 2\,\nabla_{\mathbb{S}}f(x) \otimes x, \quad x \in \mathbb{S}^{n-1}. \tag{13.4}$$

Theorem 13.1. *On $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ $(n \geq 5)$ there exist a probability measure μ_n with marginals σ_{n-1} and a constant $c_n > 0$ such that*

$$\mathrm{cov}_{\sigma_{n-1}}(f,g) = c_n \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \langle Df(x), Dg(y) \rangle \, d\mu_n(x,y) \tag{13.5}$$

for all C^2-smooth f, g on $\mathbb{S}^{n-1}$ orthogonal to all linear functions in $L^2(\sigma_{n-1})$. Moreover, μ_n has density with respect to the product measure $\sigma_{n-1} \otimes \sigma_{n-1}$ of the form $\psi_n(\langle x, y \rangle)$ for the positive function ψ_n given in (13.3), and

$$\frac{1}{n(n+2)} < c_n < \frac{1}{(n-2)(n-4)}. \tag{13.6}$$

One may also develop a heat semi-group variant of the identity (13.5) in a full analogy between the covariance representations (1.4) and (1.7). As a consequence of (1.7) we have the following identity where the orthogonality hypothesis is not needed.

Theorem 13.2. *For all C^2-smooth f, g on $\mathbb{S}^{n-1}$ $(n \geq 3)$,*

$$\mathrm{cov}_{\sigma_{n-1}}(f,g) = \int_0^\infty \left[\int_{\mathbb{S}^{n-1}} t\,\Delta_{\mathbb{S}} P_t f\,\Delta_{\mathbb{S}} g\, d\sigma_{n-1} \right] dt. \tag{13.7}$$

Proof. Applying the pointwise bound (9.3) together with Cauchy's inequality, we have that, for all $t \geq 0$,

$$\begin{aligned}
t\,|\langle \nabla_{\mathbb{S}} P_t f, \nabla_{\mathbb{S}} g \rangle| &\leq t\,|\nabla P_t f|\,|\nabla_{\mathbb{S}} g| \\
&\leq t e^{-(n-2)t} P_t |\nabla f|\,|\nabla_{\mathbb{S}} g|.
\end{aligned}$$

Here, the latter expression is vanishing at zero, and the same is true at $t = \infty$, since $P_t |\nabla f| \to \int_{\mathbb{S}^{n-1}} |\nabla f|\, d\sigma_{n-1}$ as $t \to \infty$. This justifies integration by parts in the representation (1.4):

$$\begin{aligned}
\mathrm{cov}_{\sigma_{n-1}}(f,g) &= \int_{\mathbb{S}^{n-1}} \left[\int_0^\infty \langle \nabla_{\mathbb{S}} P_t f, \nabla_{\mathbb{S}} g \rangle \, dt \right] d\sigma_{n-1} \\
&= -\int_{\mathbb{S}^{n-1}} \left[\int_0^\infty t\, \frac{d}{dt} \langle \nabla_{\mathbb{S}} P_t f, \nabla_{\mathbb{S}} g \rangle \, dt \right] d\sigma_{n-1} \\
&= -\int_0^\infty t \left[\int_{\mathbb{S}^{n-1}} \frac{d}{dt} \langle \nabla_{\mathbb{S}} P_t f, \nabla_{\mathbb{S}} g \rangle \, d\sigma_{n-1} \right] dt.
\end{aligned}$$

Moreover, differentiating under the scalar product sign according to the semi-group identity (9.1) and applying the formula (8.2) on the spherical integration by parts, we see that the

last inner integral is equal to

$$\int_{\mathbb{S}^{n-1}} \langle \nabla_{\mathbb{S}} \Delta_{\mathbb{S}} P_t f, \nabla_{\mathbb{S}} g \rangle \, d\sigma_{n-1} = -\int_{\mathbb{S}^{n-1}} \Delta_{\mathbb{S}} P_t f \, \Delta_{\mathbb{S}} g \, d\sigma_{n-1}.$$

$\square$

Remark 13.3. The last step in this derivation can be changed to yield an equivalent representation in place of (12.7), namely

$$\mathrm{cov}_{\sigma_{n-1}}(f, g) = \int_0^\infty \left[\int_{\mathbb{S}^{n-1}} t P_t f \, \Delta_{\mathbb{S}}^2 g \, d\sigma_{n-1} \right] dt, \tag{13.8}$$

where $\Delta_{\mathbb{S}}^2 f = \Delta_{\mathbb{S}} \Delta_{\mathbb{S}} f$ is the square of the spherical Laplacian.

14. Upper bounds on covariance of order $1/n^2$

Note that the constants (13.6) are of order $1/n^2$ in a big contrast with constants in the covariance representation of Theorem 1.1 which have the rate $1/n$ (for growing n). It is therefore reasonable to derive from (13.5) a corresponding analog of Corollary 11.1 in terms of L^p-norms of spherical derivatives and Hessians

$$\|\nabla_{\mathbb{S}} f\|_p = \left(\int_{\mathbb{S}^{n-1}} |\nabla_{\mathbb{S}} f|^p \, d\sigma_{n-1} \right)^{1/p}, \quad \|f_{\mathbb{S}}''\|_p = \left(\int_{\mathbb{S}^{n-1}} \|f_{\mathbb{S}}''\|_{\mathrm{HS}}^p \, d\sigma_{n-1} \right)^{1/p} \tag{14.1}$$

in the class of functions as in Theorem 12.1.

Since the inner product of $n \times n$ matrices satisfies $|\langle A, B \rangle| \leq \|A\|_{\mathrm{HS}} \|B\|_{\mathrm{HS}}$ (Cauchy's inequality), we have, for all $x, y \in \mathbb{S}^{n-1}$,

$$|\langle Df(x), Dg(y) \rangle| \leq \|Df(x)\|_{\mathrm{HS}} \|Df(y)\|_{\mathrm{HS}}. \tag{14.2}$$

Now, by the triangle inequality for the Hilbert-Schmidt norm, from (13.4) it follows that

$$\|Df(x)\|_{\mathrm{HS}} \leq \|f_S''(x)\|_{\mathrm{HS}} + 2 \|\nabla_{\mathbb{S}} f(x) \otimes x\|_{\mathrm{HS}}.$$

In turn, according to Definition 7.2, the symmetrized tensor product $\nabla_{\mathbb{S}} f(x) \otimes x$ has entries

$$(\nabla_{\mathbb{S}} f(x) \otimes x)_{ij} = \frac{1}{2} \left(x_i D_j f(x) + x_j D_i f(x) \right), \quad 1 \leq i, j \leq n,$$

so that

$$(\nabla_{\mathbb{S}} f(x) \otimes x)_{ij}^2 \leq \frac{1}{2} \left(x_i^2 (D_j f(x))^2 + x_j^2 (D_i f(x))^2 \right)$$

and

$$\|\nabla_{\mathbb{S}} f(x) \otimes x)\|_{\mathrm{HS}}^2 \leq \sum_{i=1}^n (D_i f(x))^2 = |\nabla_{\mathbb{S}} f(x)|^2$$

for all $x \in \mathbb{S}^{n-1}$. Thus,

$$\|Df(x)\|_{\mathrm{HS}} \leq \|f_S''(x)\|_{\mathrm{HS}} + 2 |\nabla_{\mathbb{S}} f(x)|, \tag{14.3}$$

and (14.2) gives

$$|\langle Df(x), Dg(y) \rangle| \leq \left(\|f_S''(x)\|_{\mathrm{HS}} + 2 |\nabla_{\mathbb{S}} f(x)| \right) \left(\|g_S''(y)\|_{\mathrm{HS}} + 2 |\nabla_{\mathbb{S}} g(y)| \right).$$

As a consequence from (13.5), $\mathrm{cov}_{\sigma_{n-1}}(f, g)$ is bounded in absolute value by

$$c_n \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \left(\|f_S''(x)\|_{\mathrm{HS}} + 2 \left| \nabla_{\mathbb{S}} f(x) \right| \right) \left(\|g_S''(y)\|_{\mathrm{HS}} + 2 \left| \nabla_{\mathbb{S}} g(y) \right| \right) d\mu_n(x, y). \qquad (14.4)$$

Moreover, since the marginals of the probability measure μ_n in Theorem 13.1 coincides with σ_{n-1}, the L^p-norm of the expression in the first bracket in (14.4) with respect to μ_n and σ_{n-1} are equal to each other and does not exceed $\|f_{\mathbb{S}}''\|_p + 2 \|\nabla_{\mathbb{S}} f\|_p$. A similar conclusion is true about the L^q-norm of the expression in the second bracket, which is needed when q is the conjugate power. Applying Hölder's inequality, we may conclude.

Corollary 14.1. *Let* $p, q \geq 1$, $\frac{1}{p} + \frac{1}{q} = 1$. *Given* C^2*-smooth functions* f, g *on* $\mathbb{S}^{n-1}$ $(n \geq 5)$ *which are orthogonal to all linear functions in* $L^2(\sigma_{n-1})$, *we have*

$$|\mathrm{cov}_{\sigma_{n-1}}(f, g)| \leq \frac{1}{(n-2)(n-4)} \left(\|f_{\mathbb{S}}''\|_p + 2 \|\nabla_{\mathbb{S}} f\|_p \right) \left(\|g_{\mathbb{S}}''\|_q + 2 \|\nabla_{\mathbb{S}} g\|_q \right), \qquad (14.5)$$

where the L^p *and* L^q *norms are defined in* (14.1).

If these norms are of order at most 1, the covariance f and g will be therefore of order at most $1/n^2$ for the growing dimension n. As easy to see, the inequality such as (14.5) remains to hold for $2 \leq n \leq 4$ with an absolute constant in front of the norms.

In the case $p = q = 2$, (14.5) can be simplified by removing the norms for the spherical gradients. Indeed, as was shown in [3], $\|\nabla_{\mathbb{S}} f\|_2 \leq \|f_{\mathbb{S}}''\|_2$ for any C^2-smooth function f on $\mathbb{S}^{n-1}$ (cf. also [4], Proposition 10.1.2). In fact, if f is orthogonal to all linear functions in $L^2(\sigma_{n-1})$, a much stronger inequality holds true, namely

$$\|\nabla_{\mathbb{S}} f\|_2^2 \leq \frac{1}{n+2} \|f_{\mathbb{S}}''\|_2^2.$$

Applying this in (14.5) with $f = g$, we obtain a second order Poincaré-type inequality

$$\mathrm{Var}_{\sigma_{n-1}}(f) \leq \frac{c}{n^2} \|f_{\mathbb{S}}''\|_2^2$$

up to an absolute constant c.

In order to get a more precise relation in this particular case, one may actually employ the following identity for the L^2-norm of the Hessian: Given a C^4-smooth function f on $\mathbb{S}^{n-1}$,

$$\|f_{\mathbb{S}}''\|_2^2 = \int_{\mathbb{S}^{n-1}} f \left(\Delta_{\mathbb{S}}^2 f + (n-2) \Delta_{\mathbb{S}} f \right) d\sigma_{n-1},$$

where we recall that $\Delta_{\mathbb{S}}^2 f = \Delta_{\mathbb{S}} \Delta_{\mathbb{S}} f$. Using the representation (8.5) of this operator in terms of spherical harmonics, it follows from this identity that

$$\mathrm{Var}_{\sigma_{n-1}}(f) \leq \frac{1}{2n(n+2)} \|f_{\mathbb{S}}''\|_2^2 \qquad (14.6)$$

for any C^2-smooth function f on $\mathbb{S}^{n-1}$ which is orthogonal to all linear functions in $L^2(\sigma_{n-1})$. Here, an equality is attained for quadratic harmonics, cf. [3].

Another approach to (14.6) can be based on the semi-group covariance representation (13.8), cf. Remark 13.3.

15. Second order concentration on the sphere

Here we employ Theorem 13.1 to strengthen Theorem 1.3 with respect to the growing dimension n for a certain class of functions. The next assertion is analogous to Corollary 10.2; however, we will use the constants c_n from (13.6) which are of the order $1/n^2$.

Corollary 15.1. *Given a C^2-smooth function f on the sphere $\mathbb{S}^{n-1}$ ($n \geq 5$) which is orthogonal to all affine functions in $L^2(\sigma_{n-1})$, we have*

$$\int_{\mathbb{S}^{n-1}} e^f \, d\sigma_{n-1} \leq \int_{\mathbb{S}^{n-1}} \exp\left\{ \frac{1}{(n-2)(n-4)} \left(2 \, \|f_{\mathbb{S}}''\|_{\mathrm{HS}}^2 + 8 \, |\nabla_{\mathbb{S}} f|^2 \right) \right\} d\sigma_{n-1}. \tag{15.1}$$

A similar property was proved in [3] under the following assumptions:

a) f is orthogonal to all affine functions in $L^2(\sigma_{n-1})$;

b) $\|f_{\mathbb{S}}''\| \leq 1$ pointwise on the sphere where $\|f_{\mathbb{S}}''\|$ denotes the operator norm;

c) $\int \|f_{\mathbb{S}}''\|_{\mathrm{HS}}^2 \, d\sigma_{n-1} \leq b$.

Then it was shown that

$$\int_{\mathbb{S}^{n-1}} \exp\left\{ \frac{n-1}{2(1+b)} \, |f| \right\} d\sigma_{n-1} \leq 2. \tag{15.2}$$

This implies that f is of the order at most $1/n$, and we have a deviation inequality

$$\sigma_{n-1}\big\{ (n-1)|f| \geq r \big\} \leq 2e^{-r/2(1+b)}, \quad r \geq 2.$$

A similar conclusion can also be made on the basis of (15.1) if the conditions $b) - c)$ are replaced with, for example, $\|f_{\mathbb{S}}''\|_{\mathrm{HS}} + |\nabla_{\mathbb{S}} f| \leq 1$ on the sphere.

Proof of Corollary 15.1. It is rather similar to the proof of Corollary 11.2. Replacing the spherical gradient with the operator D defined in (13.4) and repeating the same arguments on the basis of the covariance identity (13.5), we obtain the analog of (11.4), namely

$$\int_{\mathbb{S}^{n-1}} e^f \, d\sigma_{n-1} \leq \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \exp\{ c_n \, \langle Df(x), Df(y) \rangle \} \, d\mu_n(x,y) \tag{15.3}$$

with constants c_n satisfying (13.6). Moreover, using

$$\langle Df(x), Df(y) \rangle \;\leq\; \|Df(x)\|_{\mathrm{HS}} \, \|Df(y)\|_{\mathrm{HS}}$$
$$\leq\; \frac{1}{2} \, \|Df(x)\|_{\mathrm{HS}}^2 + \frac{1}{2} \, \|Df(y)\|_{\mathrm{HS}}^2$$

and the property that the measure μ_n has σ_{n-1} as marginals, the double integral in (15.3) can be bounded by

$$\int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} \exp\left\{ \frac{c_n}{2} \, \|Df(x)\|_{\mathrm{HS}}^2 + \frac{c_n}{2} \, \|Df(y)\|_{\mathrm{HS}}^2 \right\} d\mu_n(x,y)$$
$$\leq \left(\int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} e^{c_n \, \|Df(x)\|_{\mathrm{HS}}^2} \, d\mu_n(x,y) \right)^{1/2} \left(\int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} e^{c_n \, \|Df(y)\|_{\mathrm{HS}}^2} \, d\mu_n(x,y) \right)^{1/2}$$
$$= \int_{\mathbb{S}^{n-1}} e^{c_n \, \|Df\|_{\mathrm{HS}}^2} \, d\sigma_{n-1},$$

where we also applied Cauchy's inequality. It remains to note that, by (14.3),

$$\|Df(x)\|_{\mathrm{HS}}^2 \ \leq\ \left(\|f_{\mathbb{S}}''(x)\|_{\mathrm{HS}} + 2\,|\nabla_{\mathbb{S}}f(x)|\right)^2 \ \leq\ 2\,\|f_{\mathbb{S}}''(x)\|_{\mathrm{HS}} + 8\,|\nabla_{\mathbb{S}}f(x)|^2.$$

$\square$

16. Covariance representations on the line

As we have already emphasized, covariance identities on the Euclidean space $\mathbb{R}^n$ of dimension $n \geq 2$ with the usual gradient exist for Gaussian measures, only. However, in dimension $n = 1$ the situation is completely different. In fact, for any probability measure μ on the real line a covariance identity such as

$$\mathrm{cov}_\mu(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u'(x)v'(y)\,d\lambda(x,y) \tag{16.1}$$

exists for a suitable measure λ on the plane $\mathbb{R}\times\mathbb{R}$. In this and next sections, we collect several results in this direction and refer an interested reader to [5] for more details and historical references. Denote by C_b^∞ the class of all functions $u : \mathbb{R} \to \mathbb{R}$ having C^∞-smooth, compactly supported derivatives (in which case u are bounded).

Proposition 16.1. *Given a probability measure μ on the real line, (16.1) holds true for all $u,v \in C_b^\infty$ with a unique locally finite measure $\lambda = \lambda_\mu$. This measure is non-negative and absolutely continuous with respect to the Lebesgue measure on the plane.*

Here "locally finite" means that λ is finite on compact sets in the plane. In this case, both sides of (16.1) are well-defined and finite for all functions in C_b^∞. Moreover, (16.1) is extended to all locally absolutely continuous, complex-valued functions u,v such that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u'(x)|\,|u'(y)|\,d\lambda(x,y) < \infty, \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |v'(x)|\,|v'(y)|\,d\lambda(x,y) < \infty,$$

where the derivatives u' and v' are understood in the Radon-Nikodym sense. This condition guarantees that the double integral in (16.1) is finite and that u and v belong to $L^2(\mu)$.

The equality (16.1) can be further generalized as a covariance identity

$$\mathrm{cov}(u(X),v(Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u'(x)v'(y)\,H(x,y)\,dx\,dy \tag{16.2}$$

for arbitrary random variables X and Y, where

$$H(x,y) = \mathbb{P}\{X \leq x, Y \leq y\} - \mathbb{P}\{X \leq x\}\,\mathbb{P}\{Y \leq y\}, \quad x,y \in \mathbb{R}.$$

Here, the particular case of the identical functions $u(x) = x$ and $v(y) = y$ corresponds to the observation by Höffding [8] (provided that X and Y have finite second moments).

Moreover, when $X = Y$, and the random variable X is distributed according to μ, (16.2) is reduced to (16.1). One may therefore conclude that the mixing measure in (16.1) has density

$$H_\mu(x,y) = \frac{d\lambda(x,y)}{dx\,dy} = F(x \wedge y)\,(1 - F(x \vee y)), \quad x,y \in \mathbb{R}, \tag{16.3}$$

where

$$F(x) = \mathbb{P}\{X \leq x\} = \mu((-\infty,x]), \quad x \in \mathbb{R},$$

is the distribution function associated to μ, with notations $x \wedge y = \min(x, y)$, $x \vee y = \max(x, y)$.

Definition 16.2. Following [5], we call $\lambda = \lambda_\mu$ the Höffding measure and its density $H = H_\mu$ the Höffding kernel associated to μ.

For example, if μ is the Bernoulli measure assigning the weights $p \in (0, 1)$ and $q = 1 - p$ to the points $a < b$, then $\lambda = pqU$ where U is the uniform distribution on the square $(a, b) \times (a, b)$.

Being applied with $u(x) = v(x) = x$, the identity (16.1) shows that the total mass of the Höffding measure is the variance

$$\lambda(\mathbb{R} \times \mathbb{R}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y)\, dx\, dy = \mathrm{Var}(X).$$

Thus, λ is finite, if and only if μ has finite second moment.

Once the measure λ is finite, it may also be described via its Fourier-Stieltjes transform in terms of the characteristic function of the random variable X,

$$f(t) = \mathbb{E}\, e^{itX} = \int_{-\infty}^{\infty} e^{itx}\, d\mu(x), \quad t \in \mathbb{R}.$$

Namely, applying (16.1) to the exponential functions $u(x) = e^{itx}$ and $v(y) = e^{isy}$ $(t, s \in \mathbb{R})$, we obtain an explicit formula

$$\hat{\lambda}(t, s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{itx+isy}\, d\lambda(x, y) = \frac{f(t)f(s) - f(t+s)}{ts} \quad (t, s \neq 0). \tag{16.4}$$

In particular, this provides the uniqueness part in Theorem 1.1.

Thus, the expression on the above right-hand side represents a positive definite function in two real variables, as long as the characteristic function f is twice differentiable.

Moreover, there is a similar property of Höffding's kernels: every such function $H = H_\mu$ defined in (16.3) is positive definite on the plane, that is, $\sum_{i,j=1}^{n} a_i a_j H(x_i, x_j) \geq 0$ for any collection $a_i, x_i \in \mathbb{R}$. More generally,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)f(y)H(x, y)\, dx\, dy \geq 0,$$

for any (bounded) measurable function f on the real line.

Being positive definite, every Höffding kernel satisfies $H(x, y)^2 \leq H(x, x)H(y, y)$, which may be used to construct a pseudometric

$$d(x, y) = \big(H(x, x) - 2H(x, y) + H(y, y)\big)^{1/2}.$$

This property can be strengthened in terms of the Höffding measure. In particular,

$$\lambda(A \times B)^2 \leq \lambda(A \times A)\, \lambda(B \times B)$$

for all Borel sets $A, B \subset \mathbb{R}$.

Since the kernel $H = H_\mu$ is symmetric about the diagonal $x = y$, the Höffding measure $\lambda = \lambda_\mu$ has equal marginals $\Lambda = \Lambda_\mu$ defined by

$$\Lambda(A) = \lambda(A \times \mathbb{R}) = \int_{A} \int_{-\infty}^{\infty} H(x, y)\, dx\, dy, \quad A \subset \mathbb{R} \text{ (Borel)}. \tag{16.5}$$

Obviously, it is absolutely continuous with respect to the Lebesgue measure. More precisely, we have the following density description. As before, assume that the random variable X is distributed according to μ.

Proposition 16.3. *If X has finite first absolute moment, then the marginal Λ is finite and has density*

$$h(x) = \frac{d\Lambda(x)}{dx} = \int_x^\infty (y - a)\, dF(y), \quad a = \mathbb{E}X. \tag{16.6}$$

In particular, it is unimodal with mode at the point a, that is, $h(x)$ is non-decreasing on the half-axis $x < a$ and is non-increasing for $x > a$. Moreover, it is continuous at $x = a$ with

$$h(a-) = h(a+) = \frac{1}{2}\,\mathbb{E}\,|X - a|. \tag{16.7}$$

In the case $\mathbb{E}\,|X| = \infty$, the density of Λ is a.e. infinite on the support interval of μ.

Proposition 16.4. *Assuming that X has finite first absolute moment, the marginal Λ is a multiple of μ, if and only if μ is Gaussian.*

Once $\Lambda = \sigma^2 \mu$ with some constant σ^2, the measures Λ and λ are necessarily finite, so that μ must have a finite second moment. If $\mathbb{E}X = 0$ (without loss of generality), it follows from (16.4) that the Fourier-Stieltjes transform of Λ is given by

$$\hat{\Lambda}(t) = \hat{\lambda}(t,0) = \int_{-\infty}^\infty e^{itx}\, h(x)\, dx = -\frac{f'(t)}{t}, \quad t \in \mathbb{R},\ t \neq 0,$$

where f is the characteristic function of X. Hence, $\Lambda = \sigma^2 \mu$ if and only if $f'(t) = -\sigma^2 t f(t)$, $t \in \mathbb{R}$. But this is only possible when μ is Gaussian with mean zero and variance σ^2.

Often, the marginals of Höffding's measures appear in the particular case of the covariance representation (16.1) with the function $v(x) = x$. Then we arrive at

$$\mathrm{cov}(X, u(X)) = \int_{-\infty}^\infty u'(x)\, h(x)\, dx,$$

holding true as long as the integral is convergent. If μ is supported on an interval Δ and has there an a.e. positive density p, this formula may be rewritten as

$$\mathrm{cov}(X, u(X)) = \mathbb{E}\,\tau(X) u'(X). \tag{16.8}$$

Here the function

$$\tau(x) = \frac{h(x)}{p(x)} = \frac{1}{p(x)} \int_x^\infty (y - a)\, p(y)\, dy, \quad x \in \Delta,$$

is called the Stein kernel. We have $\tau(x) = 1$ (a.e. on Δ), if and only if μ is the standard Gaussian measure, in which case (16.8) becomes Stein's equation

$$\mathbb{E}\,X u(X) = \mathbb{E}\,u'(X).$$

After the pioneering work [15], the identity (16.8) served as a starting point for the extensive development of Stein's method as an approach to various forms of the central limit

theorem and estimating the distances to the normal law. For a comprehensive exposition of this theory, we refer an interested reader to the book [7] and survey [14].

17. Periodic covariance representations

The spherical covariance representation (1.4) of Theorem 1.3 in dimension $n = 2$, that is, on the circle $\mathbb{S}^1$, can be reduced to the covariance representation

$$\mathrm{cov}_\mu(u, v) = \int_0^1 \int_0^1 u'(x) v'(y) \, d\lambda(x, y) \tag{17.1}$$

for the uniform distribution $\mu = m$ on $(0, 1)$ in the class of all 1-periodic smooth functions.

Definition 17.1. We call a signed symmetric measure λ on $[0, 1) \times [0, 1)$ a mixing measure for a given probability measure μ on $[0, 1)$, if (17.1) holds true for all 1-periodic smooth functions u and v on the real line.

Following [5], let us now describe several results about this type of covariance representations. As we discussed before, the identity (17.1) always holds with the Höffding measure $\lambda = \lambda_\mu$. But, its marginals may be a multiple of μ in the Gaussian case only. This motivates the following question. Given μ, is it possible to choose a mixing measure λ whose marginals are multiples of μ? If so, how to describe all of them and choose a best one (in some sense)?

Proposition 17.2. *Let μ be a probability measure on $[0, 1)$ with Höffding's measure λ_μ. Subject to the constraint that the marginal distribution of λ in (17.1) is equal to $c\mu$ for a prescribed value $c \in \mathbb{R}$, the mixing measure λ exists, is unique, and is given by*

$$\begin{aligned} \lambda \;=\;& \lambda_\mu + (\sigma^2 - c)\, m \otimes m \\ & + c\,(\mu \otimes m + m \otimes \mu) - (\Lambda_\mu \otimes m + m \otimes \Lambda_\mu), \end{aligned}$$

where Λ_μ is the marginal of λ_μ and σ^2 is the variance of μ.

Here and in the sequel, m denotes the uniform distribution on $(0, 1)$.
One can specialize this characterization to the measue m and consider identities

$$\mathrm{cov}_m(u, v) = \int_0^1 \int_0^1 u'(x) v'(y) \, d\lambda(x, y). \tag{17.2}$$

Using Proposition 16.1, Proposition 17.2 leads to the following assertion which will be needed in the study of covariance identities on the circle.

Corollary 17.3. *Subject to the constraint that the marginal distribution of a mixing measure λ in (17.2) is equal to cm, $c \in \mathbb{R}$, the measure λ is unique and has density*

$$\frac{\lambda(x, y)}{dx \, dy} = Q(|x - y|) + \left(c - \frac{1}{24}\right), \quad x, y \in (0, 1),$$

where

$$Q(h) = \frac{1}{8} \left[\, 1 - 4h(1 - h)\,\right], \quad 0 \le h \le 1. \tag{17.3}$$

Note that $Q(h) \geq 0$ for all $h \in [0,1]$ and the inequality becomes an equality for $h = \frac{1}{2}$.

Moreover, the mixing measure λ is non-negative, if and only if $c \geq \frac{1}{24}$. Hence, the smallest positive measure (for the usual comparison) corresponds to the parameter $c = \frac{1}{24}$. In this sense, the optimal variant of (18.2) is given by the covariance representation

$$\mathrm{cov}_m(u,v) = \frac{1}{24} \int_0^1 \int_0^1 u'(x)v'(y) \, d\nu(x,y)$$

with a probability measure $d\nu(x,y) = 24\,Q(|x-y|)\,dx\,dy$ on $(0,1) \times (0,1)$. It has the uniform distribution m on $(0,1)$ as a marginal one.

If we want to write down a similar representation on the interval $(0,T)$, $T > 0$, one may use a linear transform. Let m_T denote the uniform distribution on $(0,T)$. Then we get that, for all smooth T-periodic functions u and v, and for all $c \geq 1/24$,

$$\mathrm{cov}_{m_T}(u,v) = \int_0^T \int_0^T u'(x)v'(y) \, d\lambda_T(x,y) \tag{17.4}$$

with a positive measure having the density

$$\frac{d\lambda_T(x,y)}{dx\,dy} = Q\left(\left|\frac{x}{T} - \frac{y}{T}\right|\right) + \left(c - \frac{1}{24}\right) \tag{17.5}$$

on $(0,T) \times (0,T)$. It has the marginal $cT m_T(dx) = c\,dx$ on $(0,T)$.

18. From the circle to the interval

Let us now return to the covariance representations on the circle

$$\mathrm{cov}_{\sigma_1}(f,g) = \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle \, d\nu(x,y), \tag{18.1}$$

where we admit that ν may be a signed measure on the torus $\mathbb{S}^1 \times \mathbb{S}^1$. The question of whether such a measure ν exists is settled in Proposition 6.6 with $\nu = \nu_2$. Here we relate (18.1) to the similar covariance identity

$$\mathrm{cov}_{m_{2\pi}}(u,v) = \int_0^{2\pi} \int_0^{2\pi} u'(t)v'(s) \, d\lambda(t,s) \tag{18.2}$$

on the semi-open interval $[0, 2\pi)$ with respect to the uniform measure $m_{2\pi}$ in the class of all 2π-periodic functions u and v on the real line.

Identifying $\mathbb{R}^2$ with the complex plane $\mathbb{C}$, to every smooth function f on $\mathbb{S}^1$ one associates

$$u(t) = f(e^{it}) = f(\cos t, \sin t), \quad t \in \mathbb{R}, \tag{18.3}$$

which is a smooth, 2π-periodic function on the real line. Conversely, starting with such $u(t)$, (18.3) defines a smooth function $f(x)$ on the circle in a unique way. Given another smooth function g on $\mathbb{S}^1$, define similarly $v(t) = g(e^{it}) = g(\cos t, \sin t)$, $t \in \mathbb{R}$, so that

$$\mathrm{cov}_{\sigma_1}(f,g) = \mathrm{cov}_{m_{2\pi}}(u,v). \tag{18.4}$$

One can also rewrite the double integral in (18.1) explicitly in terms of u and v. Recall that the spherical gradient $w = \nabla_{\mathbb{S}} f(x)$ at the point $x \in \mathbb{S}^{n-1}$ for a smooth function f

represents the unique vector in $\mathbb{R}^n$ with the smallest Euclidean norm such that

$$f(x') - f(x) = \langle w, x' - x \rangle + o(|x' - x|), \quad x' \to x \ (x' \in \mathbb{S}^{n-1}).$$

In the case of the circle, one can put $x = e^{it}$, $x' = e^{i(t+\varepsilon)}$ and rewrite the above as

$$u(t + \varepsilon) - u(t) = \left\langle w, e^{i(t+\varepsilon)} - e^{it} \right\rangle + o(\varepsilon) = \left\langle w, ie^{it} \right\rangle \varepsilon + o(\varepsilon), \quad \varepsilon \to 0. \tag{18.5}$$

Since $ie^{it} = (-\sin t, \cos t)$, we have

$$\left\langle w, ie^{it} \right\rangle = -w_1 \sin t + w_2 \cos t, \quad w = (w_1, w_2).$$

Thus, by (18.5),

$$u'(t) = -w_1 \sin t + w_2 \cos t, \tag{18.6}$$

implying that $|u'(t)| \leq |w|$. Moreover, the equality here is attained for $w_1 = -u'(t) \sin t$ and $w_2 = u'(t) \cos t$. These numbers are therefore the coordinates of the shortest vector w satisfying (18.5). One may conclude that

$$\nabla_{\mathbb{S}} f(x) = u'(t) \, (-\sin t, \cos t), \quad x = (\cos t, \sin t),$$

and similarly

$$\nabla_{\mathbb{S}} g(y) = v'(s) \, (-\sin s, \cos s), \quad y = (\cos s, \sin s),$$

which gives

$$\langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle = u'(t) v'(s) \, \cos(t - s).$$

Denote by $\tilde{\nu}$ the measure on $[0, 2\pi) \times [0, 2\pi)$ such that ν is the image of $\tilde{\nu}$ under the map $T(t, s) = (e^{it}, e^{is})$. In view of (18.4), one may rewrite (18.1) as a covariance identity

$$\mathrm{cov}_{m_{2\pi}}(u, v) = \int_0^{2\pi} \int_0^{2\pi} u'(t) v'(s) \, \cos(t - s) \, d\tilde{\nu}(t, s),$$

Thus, we obtain:

Lemma 18.1. *The representation (18.1) on the circle with a signed measure ν on $\mathbb{S}^1 \times \mathbb{S}^1$ is equivalent to the covariance representation (18.2) in the class of all 2π-periodic smooth functions u and v with a signed measure λ on $[0, 2\pi) \times [0, 2\pi)$ defined by*

$$d\lambda(t, s) = \cos(t - s) \, d\tilde{\nu}(t, s), \quad t, s \in [0, 2\pi). \tag{18.7}$$

Note that $\cos(t - s) = 0$ if and only if $\langle x, y \rangle = 0$ for $x = e^{it}$, $y = e^{is}$. Hence, the map $\nu \sim \tilde{\nu} \to \lambda$ in (18.7) transfers the part of the measure ν supported on the set

$$\Delta = \left\{ (x, y) \in \mathbb{S}^1 \times \mathbb{S}^1 : \langle x, y \rangle = 0 \right\} \tag{18.8}$$

to zero. With this in mind, one may always assume that ν is supported outside Δ.

To proceed, we will need to clarify the correspondence $\nu \sim \tilde{\nu}$ for an important class of measures on the torus.

Lemma 18.2. *A finite measure ν on the torus $\mathbb{S}^1 \times \mathbb{S}^1$ has density $\psi(x, y) = \psi(\langle x, y \rangle)$ with respect to $\sigma_1 \otimes \sigma_1$ depending on the inner product $\langle x, y \rangle$ if and only if $\tilde{\nu}$ has density*

$$\frac{d\tilde{\nu}(t, s)}{dt \, ds} = \frac{1}{(2\pi)^2} \, \psi(\cos(t - s)), \quad 0 < t, s < 2\pi, \tag{18.9}$$

with respect to the Lebesgue measure on the square $[0, 2\pi) \times [0, 2\pi)$.

Proof. The assumption that ν appears as the image of $\tilde{\nu}$ under the map $(t, s) \rightarrow (e^{it}, e^{is})$ may be equivalently stated as an integral identity

$$\int_{\mathbb{S}^1} \int_{\mathbb{S}^1} w(x, y) \, d\nu(x, y) = \int_0^{2\pi} \int_0^{2\pi} w(e^{it}, e^{is}) \, d\tilde{\nu}(t, s), \qquad (18.10)$$

holding for all Borel measurable functions $w : \mathbb{S}^1 \times \mathbb{S}^1 \rightarrow \mathbb{R}$ (assuming that the integrals exist in the Lebesgue sense). Since this map transfers the uniform distribution on $[0, 2\pi) \times [0, 2\pi)$ to the uniform distribution on the torus, we also have a similar general identity

$$\int_{\mathbb{S}^1} \int_{\mathbb{S}^1} w(x, y) \, d\sigma_1(x) \, d\sigma_1(y) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} w(e^{it}, e^{is}) \, dt \, ds. \qquad (18.11)$$

If ν has density $\psi(x, y) = \psi(\langle x, y \rangle)$ with respect to $\sigma_1 \otimes \sigma_1$, then (18.10) becomes

$$\int_{\mathbb{S}^1} \int_{\mathbb{S}^1} w(x, y) \, \psi(\langle x, y \rangle) \, d\sigma_1(x) \, d\sigma_1(y) = \int_0^{2\pi} \int_0^{2\pi} w(e^{it}, e^{is}) \, d\tilde{\nu}(t, s). \qquad (18.12)$$

On the other hand, applying (18.11) to functions of the form $w(x, y) \, \psi(\langle x, y \rangle)$ and using $\langle x, y \rangle = \cos(t - s)$ for $x = e^{it}$, $y = e^{is}$, we then obtain that

$$\int_{\mathbb{S}^1} \int_{\mathbb{S}^1} w(x, y) \, \psi(\langle x, y \rangle) \, d\sigma_1(x) \, d\sigma_1(y) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} w(e^{it}, e^{is}) \, \psi(\cos(t - s)) \, dt \, ds.$$

The right-hand side in this equality should be equalized with the right-hand side in (18.12). Since in these equalities w may be an arbitrary bounded, Borel measurable function on the torus, the desired relation (18.9) readily follows.

The argument in the opposite direction is similar. $\qquad \square$

19. Mixing measures on the circle and the interval

How can one describe all measures ν in (18.1) under natural constraints? As we have already explained, ν is not unique: If a signed measure κ is supported on the set Δ defined in (18.8), and (18.1) holds true with ν, this relation will continue to hold for the measure $\nu + \kappa$. In addition, κ may have σ_1 as marginals. However, being supported on the set of measure zero, any such measure κ may not be absolutely continuous with respect to $\sigma_1 \otimes \sigma_1$.

Since (18.1) has been related to (18.2), the question of possible measures ν may be reformulated in terms of the latter covariance representation in terms of λ. If we require that ν should have marginals being multiples of σ_1, one may then apply Corollary 17.3.

Moreover, let us require that the measure ν in (18.1) has density $\psi(x, y) = \psi(\langle x, y \rangle)$ with respect to $\sigma_1 \otimes \sigma_1$ depending on the inner product $\langle x, y \rangle$ on the torus $\mathbb{S}^1 \times \mathbb{S}^1$. Recall that $\tilde{\nu}$ was defined as a signed measure on $[0, 2\pi) \times [0, 2\pi)$ such that ν appears as the distribution of $\tilde{\nu}$ under the map $(t, s) \rightarrow (e^{it}, e^{is})$. Then, by Lemma 18.2, $\tilde{\nu}$ has density $\frac{1}{(2\pi)^2} \psi(\cos(t - s))$ with respect to the Lebesgue measure on the square $[0, 2\pi) \times [0, 2\pi)$. Hence, by Lemma 18.1, the measure λ has density

$$\frac{d\lambda(t, s)}{dt \, ds} = \frac{1}{(2\pi)^2} \cos(t - s) \, \psi(\cos(t - s)), \quad t, s \in (0, 2\pi).$$

In particular, this measure is symmetric and has a marginal distribution proportional to $m_{2\pi}$.

On the other hand, we are in position to apply Corollary 17.3 in the form (17.5) with $T = 2\pi$ and conclude that, for some constant $c \in \mathbb{R}$,

$$\frac{d\lambda(t, s)}{dt\, ds} = Q\Big(\frac{|t - s|}{2\pi}\Big) + c',$$

where

$$Q(h) = \frac{1}{8}\Big[1 - 4h(1 - h)\Big], \quad c' = c - \frac{1}{24}.$$

In this case, λ has as a marginal c times the Lebesgue measure on $[0, 2\pi)$. Equalizing the two formulas, we arrive at

$$\frac{1}{(2\pi)^2}\,(\cos h)\,\psi(\cos h) = Q\Big(\frac{h}{2\pi}\Big) + c', \tag{19.1}$$

which holds for almost all $h \in (0, 2\pi)$, or equivalently a.e.

$$\psi(\sin h) = \frac{(2\pi)^2}{\sin h}\Big(Q\Big(\frac{h + \frac{\pi}{2}}{2\pi}\Big) + c'\Big), \quad -\frac{\pi}{2} < h < \frac{3\pi}{2}.$$

Here, $\sin h \sim h$ in the neighborhood of zero, over which the left-hand side is integrable. However, as $h \to 0$,

$$Q\Big(\frac{h + \frac{\pi}{2}}{2\pi}\Big) + c' \to Q(1/4) + c' = \frac{1}{32} + c',$$

which shows that, for the integrability of the right-hand side, it is necessary that $c' = -\frac{1}{32}$. Similarly, as $h \to \pi$,

$$Q\Big(\frac{h + \frac{\pi}{2}}{2\pi}\Big) + c' \to Q(3/4) + c' = \frac{1}{32} + c',$$

so that we do not obtain a new restriction. Note that

$$\begin{aligned}
K(h) \equiv Q(h) - \frac{1}{32} &= \frac{1}{8}\Big[1 - 4h(1 - h)\Big] - \frac{1}{32} \\
&= \frac{3}{32} - \frac{1}{2}h(1 - h) = \frac{1}{2}\Big(h - \frac{1}{4}\Big)\Big(h - \frac{3}{4}\Big).
\end{aligned}$$

This new kernel is symmetric about the point $1/2$, with

$$K(0) = K(1) = \frac{3}{32}, \quad K(1/2) = -\frac{1}{32}, \quad K(1/4) = K(3/4) = 0,$$

so that it is positive in $0 < h < 1/4$ and $3/4 < h < 1$, and is negative in the interval $1/4 < h < 3/4$. This behavior is rather similar to the one of the function $\cos(\frac{h}{2\pi})$.

Thus, with the necessary value $c' = -\frac{1}{32}$ from (19.1) we obtain that

$$\psi(\cos h) = (2\pi)^2\,\frac{K\big(\frac{h}{2\pi}\big)}{\cos h}, \quad 0 \le h \le 2\pi, \tag{19.2}$$

In particular, applying this equality with $h = \pi$, $h = \frac{\pi}{2}$, and $h = 0$, we have

$$\psi(-1) = \frac{\pi^2}{8}, \quad \psi(0) = \frac{\pi}{2}, \quad \psi(1) = \frac{3\pi^2}{8}.$$

In this case, $c = c' + \frac{1}{24} = -\frac{1}{32} + \frac{1}{24} = \frac{1}{96}$ which is the total mass of the measure λ.

Let us now describe the total mass of the measure ν, that is, $\tilde{\nu}$. Applying (18.12) with $w = 1$, we have

$$
\begin{aligned}
\nu(\mathbb{S}^1 \times \mathbb{S}^1) &= \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \psi(\langle x, y \rangle) \, d\sigma_1(x) \, d\sigma_1(y) \\
&= \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \psi(\cos(t - s)) \, dt \, ds = \int_0^1 \int_0^1 \psi(\cos(2\pi(t - s))) \, dt \, ds \\
&= \int_{-1}^1 \psi(\cos(2\pi h)) \, (1 - |h|) \, dh = 2 \int_0^1 \psi(\cos(2\pi h)) \, (1 - h) \, dh \\
&= 2 \, (2\pi)^2 \int_0^1 \frac{K(h)}{\cos(2\pi h)} \, (1 - h) \, dh.
\end{aligned}
\tag{19.3}
$$

One can now refine Proposition 6.1, in which necessarily $\psi_2 = \psi$.

Proposition 19.1. *On the torus $\mathbb{S}^1 \times \mathbb{S}^1$ there exists a unique measure ν with density of the form $\psi(\langle x, y \rangle)$ with respect to $\sigma_1 \otimes \sigma_1$ such that, for all smooth functions f, g on $\mathbb{S}^1$,*

$$
\mathrm{cov}_{\sigma_1}(f, g) = \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle \, d\nu(x, y).
\tag{19.4}
$$

The function $\psi(\alpha)$ is positive, increasing, continuous for $|\alpha| \leq 1$, and is given by (19.2) with

$$
K(h) = \frac{1}{32} (4h - 1)(4h - 3), \quad 0 \leq h \leq 1.
$$

In addition $\frac{\pi^2}{8} \leq \psi(\alpha) \leq \frac{3\pi^2}{8}$, where equalities are attained for $\alpha = -1$ and $\alpha = 1$.

One may rewrite (19.4) equivalently as

$$
\mathrm{cov}_{\sigma_1}(f, g) = c \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \langle \nabla_S f(x), \nabla_S g(y) \rangle \, d\mu(x, y)
$$

with the assumption that μ is a probability measure on $\mathbb{S}^1 \times \mathbb{S}^1$ with marginals equal to σ_1. Here the constant c is described in (19.3).

As an alternative variant for (19.5), we have the following modified representation.

Proposition 19.2. *On the torus $\mathbb{S}^1 \times \mathbb{S}^1$ there exists a finite positive measure ν such that, for all smooth functions f, g on S^1,*

$$
\mathrm{cov}_{\sigma_1}(f, g) = \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \frac{\langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle}{\langle x, y \rangle} \, d\nu(x, y).
\tag{19.5}
$$

Proof. Recall that in the case $\langle x, y \rangle = 0$, necessarily $\langle \nabla_{\mathbb{S}} f(x), \nabla_{\mathbb{S}} g(y) \rangle = 0$. Hence, there is no uncertainty in the integrand in (19.5). Moreover, for $x = e^{it}$, $y = e^{is}$, we have $\langle x, y \rangle = \cos(t - s)$. Transferring the circle to $[0, 2\pi)$ and the torus to the square $[0, 2\pi) \times [0, 2\pi)$ via the inverse of the map T, the representation (19.5) will take the form

$$
\mathrm{cov}_{m_{2\pi}}(u, v) = \int_0^{2\pi} \int_0^{2\pi} u'(t) v'(s) \, d\tilde{\nu}(t, s)
$$

for the functions $u(t) = f(e^{it})$ and $v(s) = g(e^{is})$.

Here for $\tilde{\nu}$ one may take the Höffding measure on the square $(0, 2\pi) \times (0, 2\pi)$ in the covariance representation for the uniform measure $m_{2\pi}$, that is, with positive density

$$\frac{d\tilde{\nu}(t, s)}{dt \, ds} = \psi(t, s) = F(t \wedge s)\,(1 - F(t \vee s)), \quad t, s \in (0, 2\pi),$$

where $F(t) = t/(2\pi)$ is the distribution function for $m_{2\pi}$. The measure ν in (19.5) will then appear as the image of $\tilde{\nu}$ under the map T.

Remark 19.3. As was mentioned in section 16, the measure ν has a total mass

$$\nu(\mathbb{S}^1 \times \mathbb{S}^1) = \tilde{\nu}\big((0, 2\pi) \times (0, 2\pi)\big) = \int_0^{2\pi} \int_0^{2\pi} \psi(t, s)\, dt \, ds = \mathrm{Var}(U) = \frac{\pi^2}{3},$$

where U is a random variable distributed according to $m_{2\pi}$. The marginal distributions of $\tilde{\nu}$ coincide and have a positive density on $(0, 2\pi)$ given by

$$\int_0^{2\pi} \psi(t, s)\, ds = \frac{1}{4\pi}\, t(2\pi - t), \quad 0 < t < 2\pi.$$

After transferring this marginal distribution to the circle, we will obtain the marginal distribution of ν which is however not the uniform measure σ_1.

Note also that ψ is vanishing on the boundary of the square $[0, 2\pi] \times [0, 2\pi]$. Hence ν is absolutely continuous and has a continuous density on the torus with respect to the product measure $\sigma_1 \otimes \sigma_1$. However, in contrast with higher dimensions, this density does not represent a function of the inner product $\langle x, y \rangle$ (since $\psi(t, s)$ is not a function of $\cos(t - s)$).

REFERENCES

[1] Bakry, D.; Gentil, I.; Ledoux, M. Analysis and geometry of Markov diffusion operators. Fundamental Principles of Mathematical Sciences, vol. 348, Springer, Cham, 2014, xx+552 pp.

[2] Bobkov, S. G.; Götze, F.; Houdré, C. On Gaussian and Bernoulli covariance representations. Bernoulli 7 (2001), no. 3, 439–451.

[3] Bobkov, S. G.; Chistyakov, G. P.; Götze, F. Second-order concentration on the sphere. Commun. Contemp. Math. 19 (2017), no. 5, 1650058, 20 pp.

[4] Bobkov, S. G.; Chistyakov, G. P.; Götze, F. Concentration and Gaussian approximation for randomized sums. Probability Theory and Stochastic Modelling, pp. 1–426, Vol. 104. Springer Nature.

[5] Bobkov, S. G.; Duggal, D. Höffding's kernels and periodic covariance representations. Preprint (2023).

[6] Bogachev, V. I. Ornstein–Uhlenbeck operators and semigroups. Uspekhi Mat. Nauk 73:2(440) (2018), 3–74; Russian Math. Surveys 73:2 (2018), 191–260.

[7] Chen, L. H. Y.; Goldstein, L; Shao, Q.-M. Normal approximation by Stein's method. Probability and its Applications (New York). Springer, Heidelberg, 2011. xii+405 pp.

[8] Höffding, K. Maszstabinvariante Korrelationstheorie. Schr. Math. Inst. U. Inst. Angew. Math. Univ. Berlin 5 (1940), 181–233.

[9] Houdré, C.; Pérez-Abreu, V. Covariance identities and inequalities for functionals on Wiener and Poisson spaces. Ann. Probab. 23 (1995), 400–419.

[10] Houdré, C.; Pérez-Abreu, V.; Surgailis, D. Interpolation, correlation identities, and inequalities for infinitely divisible variables. J. Fourier Anal. Appl. 4(1998), no.6, 651–668.

[11] Ledoux, M. L'algébre de Lie des gradients itérés d'un générateur Markovien – développements de moyennes et entropies. Ann. Sci. École. Norm. Sup. 28 (1995), 435–460.

[12] Ledoux, M. The concentration of measure phenomenon. Math. Surveys Monogr., 89 American Mathematical Society, Providence, RI, 2001, x+181 pp.

[13] Mueller, C. E.; Weissler, F. B. Hypercontractivity for the heat semigroup for ultraspherical polynomials and on the n-sphere J. Funct. Anal. 48 (1982), no. 2, 252–283.

[14] Shao, Q.-M. Stein's method, self-normalized limit theory and applications. Proceedings of the International Congress of Mathematicians. Vol. IV, 2325—2350, Hindustan Book Agency, New Delhi, 2010.

[15] Stein, C. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In Proc. of the Sixth Berkeley Symposium on Math. Stat. and Probab. (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory 583–602, 1972, Berkeley, CA.

[16] Stein, E. M.; Weiss, G. Introduction to Fourier Analysis on Euclidean spaces. Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J. (1971). x+297 pp.