



UNIVERSITÀ DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI INGEGNERIA
QUADERNI DEL DIPARTIMENTO

Department of Engineering

Working Paper

Series “*Mathematics and Statistics*”

n. 08/MS – 2013

**QUADRATURE RULES AND DISTRIBUTION OF POINTS ON
MANIFOLDS**

by

L. Brandolini, C. Choirat, L. Colzani, G. Gigante, R. Seri, G. Travaglini

COMITATO DI REDAZIONE[§]

Series Mathematics and Statistics (MS): Luca Brandolini, Alessandro Fassò, Christian Vergara

[§] L'accesso alle *Series* è approvato dal Comitato di Redazione. I *Working Papers* della Collana dei Quaderni del Dipartimento di Ingegneria dell'Informazione e Metodi Matematici costituiscono un servizio atto a fornire la tempestiva divulgazione dei risultati dell'attività di ricerca, siano essi in forma provvisoria o definitiva.

Quadrature rules and distribution of points on manifolds

Luca Brandolini
Giacomo Gigante

Christine Choirat
Raffaello Seri

Leonardo Colzani
Giancarlo Travaglini

Abstract

We study the error in quadrature rules on a compact manifold. Our estimates are in the same spirit of the Koksma Hlawka inequality and they depend on a sort of discrepancy of the sampling points and a generalized variation of the function. In particular, we give sharp quantitative estimates for quadrature rules of functions in Sobolev classes.

Keywords. Quadrature, discrepancy, harmonic analysis

1 Introduction

In what follows, \mathcal{M} is a smooth compact d dimensional Riemannian manifold without boundary, with Riemannian measure dx , normalized so that the total volume of the manifold is 1, and Δ is the Laplace Beltrami operator. This operator is self-adjoint in $L^2(\mathcal{M})$, it has a sequence of eigenvalues $\{\lambda^2\}$ and an orthonormal complete system of eigenfunctions $\{\varphi_\lambda(x)\}$, $\Delta\varphi_\lambda(x) = \lambda^2\varphi_\lambda(x)$. The eigenvalues, possibly repeated, are ordered with increasing modulus. In particular, the first eigenvalue is 0 and the associated eigenfunction is 1. An example is the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ with the Laplace operator $-\sum \partial^2/\partial x_j^2$, eigenvalues $\{4\pi^2 |k|^2\}_{k \in \mathbb{Z}^d}$ and eigenfunctions $\{\exp(2\pi i kx)\}_{k \in \mathbb{Z}^d}$. Another example is the sphere $\mathbb{S}^d = \{x \in \mathbb{R}^{d+1}, |x| = 1\}$ with dx the normalized surface measure and with Δ the angular component of the Laplacian in the space \mathbb{R}^{d+1} , eigenvalues $\{n(n+d-1)\}_{n=0}^{+\infty}$ and eigenfunctions the restriction to the sphere of homogeneous harmonic polynomials in space. With a small abuse of notation and in analogy with the Euclidean space, the Riemannian distance between x and y will be denoted $|x - y|$.

A classical problem is to approximate an integral $\int_{\mathcal{M}} f(x)dx$ with Riemann sums $N^{-1} \sum_{j=1}^N f(z_j)$, or weighted analogues $\sum_{j=1}^N \omega_j f(z_j)$, and what follows will be concerned with the discrepancy between integrals and sums for functions in Sobolev classes

Mathematics Subject Classification (2010): Primary 41A55; Secondary 11K38, 42C15

$W^{\alpha,p}(\mathcal{M})$ with $1 \leq p \leq +\infty$ and $\alpha > d/p$. The assumption $\alpha > d/p$ guarantees the boundedness and continuity of the function $f(x)$, otherwise the point evaluations $f(z_j)$ may be not defined. As a motivation, assume there exists a decomposition of \mathcal{M} into N disjoint pieces $\mathcal{M} = U_1 \cup U_2 \cup \dots \cup U_N$ and these pieces have measures N^{-1} and diameters at most $cN^{-1/d}$. In what follows, as usual, the constants a, b, c, \dots may change from step to step. Choosing a point z_j in each U_j , one obtains the estimate

$$\begin{aligned} & \left| N^{-1} \sum_{j=1}^N f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \\ & \leq \sum_{j=1}^N \int_{U_j} |f(z_j) - f(x)| dx \leq \sup_{|y-x| \leq cN^{-1/d}} |f(y) - f(x)|. \end{aligned}$$

In particular, since functions in $W^{\alpha,p}(\mathcal{M})$ with $\alpha > d/p$ are Hölder continuous of degree $\alpha - d/p$, one obtains

$$\left| N^{-1} \sum_{j=1}^N f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq cN^{-(\alpha-d/p)/d} \|f\|_{W^{\alpha,p}(\mathcal{M})}.$$

On the other hand, it will be shown that suitable choices of the sampling points $\{z_j\}_{j=1}^N$ improve the exponent $1/p - \alpha/d$ to $-\alpha/d$ and this is best possible. More precisely, the main results of this paper are the following:

(A) For every $d/2 < \alpha < d/2 + 1$ there exists $c > 0$ such that if $\mathcal{M} = U_1 \cup U_2 \cup \dots \cup U_N$ is a decomposition of the manifold in disjoint pieces with measure $|U_j| = \omega_j$, then there exists a distribution of points $\{z_j\}_{j=1}^N$ with $z_j \in U_j$ such that for every function $f(x)$ in the Sobolev space $W^{\alpha,2}(\mathcal{M})$,

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq c \max_{1 \leq j \leq N} \text{diameter}(U_j)^\alpha \|f\|_{W^{\alpha,2}}.$$

(B) Assume that the points $\{z_j\}_{j=1}^N$ and the positive weights $\{\omega_j\}_{j=1}^N$ give an exact quadrature for all eigenfunctions with eigenvalues $\lambda^2 < r^2$, that is

$$\sum_{j=1}^N \omega_j \varphi_\lambda(z_j) = \int_{\mathcal{M}} \varphi_\lambda(x) dx = \begin{cases} 1 & \text{if } \lambda = 0, \\ 0 & \text{if } 0 < \lambda < r. \end{cases}$$

Then for every $1 \leq p \leq +\infty$ and $\alpha > d/p$ there exist $c > 0$, which may depend on \mathcal{M}, p, α , but is independent of $r, \{z_j\}_{j=1}^N$ and $\{\omega_j\}_{j=1}^N$, such that

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq cr^{-\alpha} \|f\|_{W^{\alpha,p}}.$$

(C) If $1 \leq p \leq +\infty$, if $\alpha > d/p$ and if $\kappa \geq 1/2$, then there exists $c > 0$ with the following property: let $\{z_j\}_{j=1}^N$ be a distribution of points on \mathcal{M} with

$$\frac{\sup_{x \in \mathcal{M}} \min_j |x - z_j|}{\min_{i \neq j} |z_i - z_j|} \leq \kappa.$$

Then there exist positive weights $\{\omega_j\}_{j=1}^N$ such that

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq c N^{-\alpha/d} \|f\|_{W^{\alpha,p}}.$$

(D) For every $1 \leq p \leq +\infty$ and $\alpha > d/p$ there exists $c > 0$ such that for every distribution of points $\{z_j\}_{j=1}^N$ and numbers $\{\omega_j\}_{j=1}^N$ there exists a function $f(x)$ in $W^{\alpha,p}(\mathcal{M})$ with

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \geq c N^{-\alpha/d} \|f\|_{W^{\alpha,p}}.$$

In (C) the quantity $\sup_{x \in \mathcal{M}} \min_j |x - z_j|$ is the mesh norm, $\min_{i \neq j} |z_i - z_j|$ is the separation distance, and their ratio is the mesh-separation ratio of the distribution of points $\{z_j\}_{j=1}^N$. See [16]. An explicit example is the following. The torus \mathbb{T}^d can be partitioned into $N = n^d$ congruent cubes with sides $1/n$ and this partition generates the mesh of points $(n^{-1}\mathbb{Z}^d) \cap \mathbb{T}^d$, which gives an exact quadrature at least for all exponentials $\exp(2\pi i k x)$ with k in the hypercube $\{\max_{j=1,\dots,d} |k_j| < n\}$. In this case, (A) and (B) give an upper bound for the error in numerical integration of the order of $N^{-\alpha/d}$. More generally, if a manifold is decomposed into N disjoint pieces $\mathcal{M} = U_1 \cup U_2 \cup \dots \cup U_N$ with diameters $\leq cN^{-1/d}$, then (A) gives the upper bound $N^{-\alpha/d}$. Moreover, by Weyl's estimates on the spectrum of an elliptic operator, for every $r > 1$ there are approximately cr^d eigenfunctions with eigenvalues $\lambda^2 < r^2$ and there exist $N \leq cr^d$ nodes $\{z_j\}_{j=1}^N$ and positive weights $\{\omega_j\}_{j=1}^N$ which give an exact quadrature for these eigenfunctions. Then in this case (B) gives the above upper bound $N^{-\alpha/d}$. Hence both (A) and (B) give the bound $N^{-\alpha/d}$, and by (D) this latter is optimal. Similarly, observe that if $r > 0$ and if $\{|x - z_j| < r\}_{j=1}^N$ is a maximal set of pairwise disjoint spheres in \mathcal{M} , then the centers $\{z_j\}_{j=1}^N$ satisfy the assumption of (C) with $\kappa = 1$ and $N \approx r^{-d}$. Hence, by (C) and (D), these nodes give an optimal cubature rule. When the manifold is a torus or a sphere these results are essentially known, and indeed there is a huge literature on this subject. See [29] for deterministic and stochastic error bounds in numerical analysis. In particular, (B) and (D) for $p = 2$ and for spheres are contained in [7], [17], [18], [19] and [20]. For Besov spaces on spheres some results slightly more precise than (B) and (D) are in [21], while a result slightly weaker than (D) for compact two point homogeneous spaces is in [25]. See also [10] and, for a survey on related results, [15] and [30]. Beside the proofs of (A), (B), (C), (D), which are contained in the following section, the paper contains also a final

section with a number of further results and remarks. Among them it is proved that if a quadrature rule achieves optimal error bounds in the Sobolev space $W^{\alpha,2}(\mathcal{M})$, then this quadrature rule is optimal also in all spaces $W^{\beta,2}(\mathcal{M})$ with $d/2 < \beta < \alpha$. Moreover, it is proved that there is a relation between quadrature rules and geometric discrepancy:

(E) If $d\nu(x)$ is a probability measure on \mathcal{M} , then the norm of the measure $d\nu(x) - dx$ as a linear functional on $W^{\alpha,2}(\mathcal{M})$ decreases as α increases. Moreover, if the norm of $d\nu(x) - dx$ on $W^{\alpha,2}(\mathcal{M})$ is $r^{-\alpha}$ for some $r > 1$,

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq r^{-\alpha} \|f\|_{W^{\alpha,2}},$$

then for every $d/2 < \beta < \alpha$ there exists a constant c which may depend on $\alpha, \beta, \mathcal{M}$, but is independent of r and $d\nu(x)$, such that

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq cr^{-\beta} \|f\|_{W^{\beta,2}}.$$

(F) Assume that for some $r \geq 1$ the discrepancy of the probability measure $d\nu(x)$ with respect to the balls $\mathcal{B}(y, \delta)$ with center y and radius δ satisfies the estimates

$$\left| \int_{\mathcal{B}(y,\delta)} d\nu(x) - \int_{\mathcal{B}(y,\delta)} dx \right| \leq \begin{cases} r^{-d} & \text{if } \delta \leq 1/r, \\ r^{-1}\delta^{d-1} & \text{if } \delta \geq 1/r. \end{cases}$$

Then for every $1 \leq p \leq +\infty$ and $\alpha > d/p$, there exists a constant c , which may depend on α and p , but is independent of $d\nu(x)$ and r , such that

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq \begin{cases} cr^{-\alpha} \|f\|_{W^{\alpha,p}} & \text{if } 0 < \alpha < 1, \\ cr^{-1} \log(1+r) \|f\|_{W^{\alpha,p}} & \text{if } \alpha = 1, \\ cr^{-1} \|f\|_{W^{\alpha,p}} & \text{if } \alpha > 1. \end{cases}$$

Observe that while (A) and (B) hold for specific quadrature rules, (E) is a result for arbitrary quadratures. Actually, (E) is only one way, from α to $\beta < \alpha$. The estimate $r^{-\alpha}$ for an α does not necessarily imply the estimate $cr^{-\beta}$ for $\beta > \alpha$. Moreover, the sets $\mathcal{B}(y, \delta)$ in (F) are not precisely geodesic balls, but level sets of suitable kernels on the manifold. However, for spheres or compact rank one symmetric spaces these sets are geodesic balls, and the discrepancy of the measure is the spherical cap discrepancy. See [4] or [28], and for other relations between quadrature and discrepancy on spheres see also [2]. Finally, we would like to point out that our paper is (almost) self-contained, it does not rely on explicit properties of manifolds or special functions, and it may provide a unified perspective and simple alternative proofs of some known results.

We would like to thank the referee for some useful suggestions and especially for bringing to our attention the papers [14] and [27], which have led us to improve the original draft, in particular Corollary 2.15.

2 Main results

The eigenfunction expansions of functions and operators are a basic tool in what follows. The system of eigenfunctions $\{\varphi_\lambda(x)\}$ is orthonormal complete in $L^2(\mathcal{M})$ and every square integrable function has a Fourier transform and a Fourier expansion,

$$\mathcal{F}f(\lambda) = \int_{\mathcal{M}} f(y)\overline{\varphi_\lambda(y)}dy, \quad f(x) = \sum_{\lambda} \mathcal{F}f(\lambda)\varphi_\lambda(x).$$

Since the Laplace operator is elliptic, the eigenfunctions are smooth and it is possible to extend the definition of Fourier transforms and series to distributions. In particular, these Fourier expansions are always convergent, at least in the topology of distributions. One can write the discrepancy between integral and Riemann sum as a single integral with respect to a signed measure $d\mu(x) = \sum_{j=1}^N \omega_j d\delta_{z_j}(x) - dx$, with $d\delta_y(x)$ the Dirac measure concentrated at the point y and dx the Riemannian measure,

$$\sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx = \int_{\mathcal{M}} f(x)d\mu(x).$$

Then the estimate of the numerical integration error reduces to the estimate of the norm of a linear functional $d\mu(x)$ on a space of test functions $f(x)$. Some of the results which follow will be stated for generic finite complex valued measures $d\mu(x)$, for signed measures of the form $d\mu(x) = d\nu(x) - dx$ with $d\nu(x)$ a probability measure, and also for atomic probability measures $d\nu(x) = \sum_{j=1}^N \omega_j d\delta_{z_j}(x)$. The following is an easy and straightforward extension to compact manifolds and p norms of some abstract results for reproducing kernel Hilbert spaces. See, e.g., [1], [6], [12], [13].

Theorem 2.1. *Let $\{\psi(\lambda)\}$ be a complex sequence indexed by the eigenvalues $\{\lambda^2\}$, with $\{\psi(\lambda)\}$ and $\{\psi(\lambda)^{-1}\}$ slowly increasing, that is $|\psi(\lambda)| \leq a(1 + \lambda^2)^{\alpha/2}$ and $|\psi(\lambda)^{-1}| \leq b(1 + \lambda^2)^{\beta/2}$. Let the operators A and B and the associated adjoints A^* and B^* be defined by*

$$\begin{aligned} Af(x) &= \sum_{\lambda} \psi(\lambda)\mathcal{F}f(\lambda)\varphi_\lambda(x), & A^*g(x) &= \sum_{\lambda} \psi(\lambda)\overline{\mathcal{F}(\overline{g})(\lambda)\varphi_\lambda(x)}, \\ Bf(x) &= \sum_{\lambda} \psi(\lambda)^{-1}\mathcal{F}f(\lambda)\varphi_\lambda(x), & B^*g(x) &= \sum_{\lambda} \psi(\lambda)^{-1}\overline{\mathcal{F}(\overline{g})(\lambda)\varphi_\lambda(x)}. \end{aligned}$$

All these operators are well defined and continuous on test functions, and they can be extended by duality to tempered distributions. Finally, let $f(x)$ be a continuous function and let $d\mu(x)$ be a finite complex measure on \mathcal{M} . If $1 \leq p, q \leq +\infty$ and $1/p + 1/q = 1$, then

$$\left| \int_{\mathcal{M}} f(x)d\mu(x) \right| \leq \left\{ \int_{\mathcal{M}} |Af(x)|^p dx \right\}^{1/p} \left\{ \int_{\mathcal{M}} |B^*\mu(x)|^q dx \right\}^{1/q}.$$

In particular, when $p = q = 2$, if $B^* \mu(x)$ is square integrable and if $f(x) = B(\overline{B^* \mu})(x)$ is continuous, then the above inequality reduces to an equality.

Proof. Integration by parts shows that $\lambda^{2n} \mathcal{F}f(\lambda) = \mathcal{F}(\Delta^n f)(\lambda)$. It follows that the space of test functions is characterized by the rapid decay of the Fourier transform. In particular, if $\{\psi(\lambda)\}$ is slowly increasing and $\{\mathcal{F}f(\lambda)\}$ is rapidly decreasing, then also $\{\psi(\lambda)\mathcal{F}f(\lambda)\}$ is rapidly decreasing, hence it is the Fourier transform of a test function. This implies that the operator A is well defined on test functions, and the same for A^* , B , B^* . In what follows the pairing between a test function and a distribution is denoted with an integral, even when the distribution is not a function and the integral is divergent. In particular, if $f(x)$ is a test function, by Hölder inequality with $1/p + 1/q = 1$, since the operators A and B are inverses of each other,

$$\begin{aligned} & \left| \int_{\mathcal{M}} f(x) d\mu(x) \right| = \left| \int_{\mathcal{M}} B A f(x) d\mu(x) \right| \\ &= \left| \int_{\mathcal{M}} A f(x) B^* \mu(x) dx \right| \leq \left\{ \int_{\mathcal{M}} |A f(x)|^p dx \right\}^{1/p} \left\{ \int_{\mathcal{M}} |B^* \mu(x)|^q dx \right\}^{1/q}. \end{aligned}$$

The general case of $f(x)$ continuous follows by approximation with test functions. Finally, when $p = q = 2$ the Cauchy inequality reduces to an equality if the functions $A f(x)$ and $\overline{B^* \mu(x)}$ are square integrable and proportional. \square

In what follows the operators A and B will be fractional powers of the Laplace Beltrami operator: $(I + \Delta)^{\pm\alpha/2}$.

Definition 2.2. The Sobolev space $W^{\alpha,p}(\mathcal{M})$, $-\infty < \alpha < +\infty$ and $1 \leq p \leq +\infty$, is the set of all distributions on \mathcal{M} with $(I + \Delta)^{\alpha/2} f(x)$ in $L^p(\mathcal{M})$, that is with

$$\begin{aligned} \|f\|_{W^{\alpha,p}} &= \left\{ \int_{\mathcal{M}} \left| \sum_{\lambda} (1 + \lambda^2)^{\alpha/2} \mathcal{F}f(\lambda) \varphi_{\lambda}(x) \right|^p dx \right\}^{1/p} < +\infty, \quad 1 \leq p < +\infty, \\ \|f\|_{W^{\alpha,\infty}} &= \sup_{x \in \mathcal{M}} \text{ess} \left| \sum_{\lambda} (1 + \lambda^2)^{\alpha/2} \mathcal{F}f(\lambda) \varphi_{\lambda}(x) \right| < +\infty. \end{aligned}$$

An equivalent definition is the following.

Definition 2.3. Let $B^{\alpha}(x, y)$, $-\infty < \alpha < +\infty$, be the Bessel kernel

$$B^{\alpha}(x, y) = \sum_{\lambda} (1 + \lambda^2)^{-\alpha/2} \varphi_{\lambda}(x) \overline{\varphi_{\lambda}(y)}.$$

A distribution $f(x)$ is in the Sobolev space $W^{\alpha,p}(\mathcal{M})$ if and only if it is a Bessel potential of a function $g(x)$ in $L^p(\mathcal{M})$,

$$f(x) = \int_{\mathcal{M}} B^{\alpha}(x, y) g(y) dy.$$

Moreover, $\|f\|_{W^{\alpha,p}} = \|g\|_{L^p}$.

In particular, when $p = 2$,

$$\|f\|_{W^{\alpha,2}} = \left\{ \sum_{\lambda} (1 + \lambda^2)^{\alpha} |\mathcal{F}f(\lambda)|^2 \right\}^{1/2}.$$

Another equivalent definition is a localization result: A distribution $f(x)$ is in $W^{\alpha,p}(\mathcal{M})$ if and only if for every smooth function $g(x)$ with support in a local card $x = \psi(y) : \mathbb{R}^d \rightsquigarrow \mathcal{M}$, the distribution $g(\psi(y))f(\psi(y))$ is in $W^{\alpha,p}(\mathbb{R}^d)$. In particular, if α is a positive even integer, then $f(x)$ is in $W^{\alpha,p}(\mathcal{M})$ if and only if the p -th power of $f(x)$ and of $\Delta^{\alpha/2}f(x)$ are integrable. Moreover, distributions in $W^{\alpha,p}(\mathcal{M})$ with $\alpha > d/p$ are Hölder continuous of degree $\alpha - d/p$. When applied to functions in Sobolev classes, Theorem 2.1 gives the following corollary.

Corollary 2.4. (1) If $B^{\alpha}(x, y)$ is the Bessel kernel, if $d\mu(x)$ is a finite complex measure on \mathcal{M} , and if $f(x)$ is a continuous function in $W^{\alpha,p}(\mathcal{M})$, with $1 \leq p, q \leq +\infty$ and $1/p + 1/q = 1$, then

$$\left| \int_{\mathcal{M}} f(x) d\mu(x) \right| \leq \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^{\alpha}(x, y) d\mu(x) \right|^q dy \right\}^{1/q} \|f\|_{W^{\alpha,p}}.$$

In particular, if $\alpha > d/p$ then every element in $W^{\alpha,p}(\mathcal{M})$ has a continuous representative and the above integrals are well defined and finite. On the contrary, the spaces $W^{\alpha,p}(\mathcal{M})$ with $\alpha \leq d/p$ contain unbounded functions and, if the measure $d\mu(x)$ does not vanish on the set where $f(x) = \infty$, then $\int_{\mathcal{M}} f(x) d\mu(x)$ may diverge.

(2) When $p = q = 2$ and $\alpha > d/2$, then the above inequality simplifies to

$$\left| \int_{\mathcal{M}} f(x) d\mu(x) \right| \leq \left\{ \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\mu(x) \overline{d\mu(y)} \right\}^{1/2} \|f\|_{W^{\alpha,2}}.$$

Equivalently, by the Fourier expansion of the Bessel kernel,

$$\left| \int_{\mathcal{M}} f(x) d\mu(x) \right| \leq \left\{ \sum_{\lambda} (1 + \lambda^2)^{-\alpha} |\mathcal{F}\mu(\lambda)|^2 \right\}^{1/2} \|f\|_{W^{\alpha,2}}.$$

Moreover, with $f(x) = \int_{\mathcal{M}} B^{2\alpha}(x, y) \overline{d\mu(y)}$ the above inequalities reduce to equalities.

(3) If $d\mu(x) = d\nu(x) - dx$ is the difference between a probability measure $d\nu(x)$ and the Riemannian measure dx , then

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq \left\{ \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\nu(x) d\nu(y) - 1 \right\}^{1/2} \|f\|_{W^{\alpha,2}}.$$

Equivalently,

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq \left\{ \sum_{\lambda > 0} (1 + \lambda^2)^{-\alpha} |\mathcal{F}\nu(\lambda)|^2 \right\}^{1/2} \|f\|_{W^{\alpha,2}}.$$

Proof. (1) is an immediate corollary of Theorem 2.1. In order to prove (2), observe that

$$\int_{\mathcal{M}} B^\alpha(x, y) B^\beta(y, z) dy = B^{\alpha+\beta}(x, z).$$

Moreover, this Bessel kernel is real and symmetric, see Lemma 2.6 below. Hence,

$$\begin{aligned} & \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^\alpha(x, y) d\mu(x) \right|^2 dy \\ &= \int_{\mathcal{M}} \int_{\mathcal{M}} \int_{\mathcal{M}} B^\alpha(x, y) B^\alpha(z, y) dy d\mu(x) \overline{d\mu(z)} \\ &= \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, z) d\mu(x) \overline{d\mu(z)}. \end{aligned}$$

(3) is a corollary of (1) and (2). Indeed, since $B^{2\alpha}(x, y) = B^{2\alpha}(y, x)$ and $\int_{\mathcal{M}} B^{2\alpha}(x, y) dy = 1$, it follows that

$$\begin{aligned} & \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) (d\nu(x) - dx) (d\nu(y) - dy) \\ &= \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\nu(x) d\nu(y) - \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\nu(x) dy \\ & \quad - \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) dx d\nu(y) + \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) dx dy \\ &= \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\nu(x) d\nu(y) - 1. \end{aligned}$$

Finally, by Sobolev imbedding theorem, functions in $W^{\alpha,p}(\mathcal{M})$ with $\alpha > d/p$ are continuous and all the above integrals are well defined and finite. The Sobolev imbedding also follows from the estimates on the Bessel kernel provided in Lemma 2.6, as explained in Remark 3.3 below. \square

The above corollary leads to estimate the energy integrals

$$\left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^\alpha(x, y) d\mu(x) \right|^q dy \right\}^{1/q},$$

which for $q = 2$ and $d\mu(x) = d\nu(x) - dx$ simplifies to

$$\left\{ \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\nu(x) d\nu(y) - 1 \right\}^{1/2} = \left\{ \sum_{\lambda > 0} (1 + \lambda^2)^{-\alpha} |\mathcal{F}\nu(\lambda)|^2 \right\}^{1/2}.$$

By the last formula, the energy attains a minimum if and only if $\mathcal{F}\nu(\lambda) = 0$ for all $\lambda > 0$. Hence $d\nu(x)$ has the eigenfunction expansion $\mathcal{F}\nu(0) \varphi_0(x)$, and since $\varphi_0(x) = 1$ this gives the Riemannian measure dx . The meaning of the corollary is that measures with low energy are close to the Riemannian measure and they give good quadrature rules. In

order to provide quantitative estimates for the above integrals, one has to collect some properties of the Bessel kernels. The norm of the function $y \rightsquigarrow B^\alpha(x, y)$ in $W^{\gamma,2}(\mathcal{M})$ is

$$\|B^\alpha(x, \cdot)\|_{W^{\gamma,2}} = \left\{ \sum_{\lambda} (1 + \lambda^2)^{\gamma-\alpha} |\varphi_\lambda(x)|^2 \right\}^{1/2}.$$

By Weyl's estimates on the spectrum of an elliptic operator, see [9, Chapter 6.4] and [22, Theorem 17.5.3 and Corollary 17.5.8], for every $r > 1$ there are approximately cr^d eigenfunctions $\varphi_\lambda(x)$ with eigenvalues $\lambda^2 < r^2$ and $\sum_{\lambda \leq r} |\varphi_\lambda(x)|^2 \leq cr^d$. It then follows that the norm in $W^{\gamma,2}(\mathcal{M})$ of $B^\alpha(x, y)$ is finite provided that $\gamma < \alpha - d/2$ and, by Sobolev imbedding theorem, it also follows that $B^\alpha(x, y)$ is Hölder continuous of degree $\delta < \alpha - d$. Indeed, we shall see that a bit more is true: $B^\alpha(x, y)$ is Hölder continuous of degree $\alpha - d$.

Lemma 2.5. *The heat kernel*

$$W(t, x, y) = \sum_{\lambda} \exp(-\lambda^2 t) \varphi_\lambda(x) \overline{\varphi_\lambda(y)},$$

which is the fundamental solution to the heat equation $\partial/\partial t = -\Delta$ on $\mathbb{R}_+ \times \mathcal{M}$, is symmetric real and positive: $W(t, x, y) = W(t, y, x) > 0$ for every $x, y \in \mathcal{M}$ and $t > 0$. Moreover, for all non negative integers m and n there exists c such that, if $|x - y|$ denotes the Riemannian distance between x and y , and ∇ the gradient,

$$\begin{cases} |\nabla^m W(t, x, y)| \leq ct^{-(d+m)/2} (1 + |x - y|/\sqrt{t})^{-n} & \text{if } 0 < t \leq 1, \\ |\nabla^m W(t, x, y)| \leq c & \text{if } 1 \leq t < +\infty. \end{cases}$$

Proof. All of this is well known, see, e.g., [9], [23], [33] and [35]. Anyhow we want to provide a proof for the torus and a hint for the general case. The idea is that heat has essentially a finite speed of propagation and diffusion on manifolds is comparable to diffusion on Euclidean spaces, at least for small times. The heat kernel in the Euclidean space \mathbb{R}^d is a Gaussian,

$$\begin{aligned} W(t, x, y) &= \int_{\mathbb{R}^d} \exp(-4\pi^2 t |\xi|^2) \exp(2\pi i (x - y) \xi) d\xi \\ &= (4\pi t)^{-d/2} \exp(-|x - y|^2 / 4t). \end{aligned}$$

By the Poisson summation formula, the heat kernel on the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ is the periodization of the kernel in the space,

$$\begin{aligned} &\sum_{k \in \mathbb{Z}^d} \exp(-4\pi^2 |k|^2 t) \exp(2\pi i k (x - y)) \\ &= \sum_{k \in \mathbb{Z}^d} (4\pi t)^{-d/2} \exp(-|x - y - k|^2 / 4t). \end{aligned}$$

Periodicity allows to assume that $x - y \in [-1/2, 1/2)^d$, and in this case the Riemannian distance between x and y coincides with the Euclidean distance. An explicit computation shows that the term $k = 0$ in the above series satisfies the required estimate,

$$\begin{aligned} & (4\pi t)^{-d/2} \exp(-|x - y|^2/4t) \\ & \leq \begin{cases} ct^{-d/2} (1 + |x - y|/\sqrt{t})^{-n} & \text{if } 0 < t \leq 1, \\ c & \text{if } 1 \leq t < +\infty. \end{cases} \end{aligned}$$

The sum for $k \neq 0$ is negligible. Indeed, since $\exp(-z) \leq cz^{-N}$ for $z > 0$,

$$\begin{aligned} & \sum_{k \neq 0} (4\pi t)^{-d/2} \exp(-|x - y - k|^2/4t) \\ & \leq ct^{N-d/2} \sum_{k \neq 0} |x - y - k|^{-2N} \leq ct^{N-d/2}, \end{aligned}$$

and this satisfies the required estimate when $0 < t \leq 1$. When $t > 1$ it suffices to observe that

$$\begin{aligned} & \sum_{k \neq 0} (4\pi t)^{-d/2} \exp(-|x - y - k|^2/4t) \\ & \leq c \int_{\mathbb{R}^d} (4\pi t)^{-d/2} \exp(-|z|^2/4t) dz \leq c. \end{aligned}$$

The estimates for the derivatives are analogous. This proves the lemma for the torus. The heat kernel on a compact manifold is similar, in particular it has an asymptotic expansion with Euclidean main term. See, e.g., [9, Chapter VI]. More precisely, by the Minakshisundaram Pleijel recursion formulas, there exist smooth functions $u_k(x, y)$ such that, if t is small and $|x - y|$ denotes the distance between x and y ,

$$W(t, x, y) \approx (4\pi t)^{-d/2} \exp(-|x - y|^2/4t) \left(\sum_{k=0}^n t^k u_k(x, y) + O(t^{n+1}) \right).$$

On the other hand, $W(t, x, y) \rightarrow 1$ as $t \rightarrow +\infty$. The estimates on the size of this kernel and its derivatives follow from this asymptotic expansion. The positivity $W(t, x, y) > 0$ is a consequence of the maximum principle for the heat equation and the symmetry $W(t, x, y) = W(t, y, x)$ follows from this positivity and the eigenfunction expansion. \square

Lemma 2.6. (1) *The Bessel kernel $B^\alpha(x, y)$ with $\alpha > 0$ is a superposition of heat kernels $W(t, x, y)$:*

$$B^\alpha(x, y) = \Gamma(\alpha/2)^{-1} \int_0^{+\infty} t^{\alpha/2-1} \exp(-t) W(t, x, y) dt.$$

(2) The Bessel kernel $B^\alpha(x, y)$ with $\alpha > 0$ is real and positive for every $x, y \in \mathcal{M}$, and it is smooth in $\{x \neq y\}$. Moreover, for suitable constants $0 < a < b$,

$$\begin{aligned} a|x-y|^{\alpha-d} &\leq B^\alpha(x, y) \leq b|x-y|^{\alpha-d} \quad \text{if } 0 < \alpha < d, \\ a \log(1+|x-y|^{-1}) &\leq B^\alpha(x, y) \leq b \log(1+|x-y|^{-1}) \quad \text{if } \alpha = d, \\ a &\leq B^\alpha(x, y) \leq b \quad \text{if } \alpha > d. \end{aligned}$$

(3) If $d < \alpha < d+1$, then $B^\alpha(x, y)$ is Hölder continuous of degree $\alpha - d$, that is there exists c such that for every $x, y, z \in \mathcal{M}$,

$$|B^\alpha(x, y) - B^\alpha(x, z)| \leq c|y - z|^{\alpha-d}.$$

(4) If $d < \alpha < d+2$, then there exists c such that for every $x, y \in \mathcal{M}$,

$$|B^\alpha(x, x) - B^\alpha(x, y)| \leq c|x - y|^{\alpha-d}.$$

Proof. When the manifold is a torus and the eigenfunctions are exponentials the proof is elementary. The Bessel kernel on the torus \mathbb{T}^d is an even function, and thus a sum of cosines,

$$\begin{aligned} B^\alpha(x, y) &= \sum_{k \in \mathbb{Z}^d} (1 + 4\pi^2 |k|^2)^{-\alpha/2} \exp(2\pi i k x) \exp(-2\pi i k y) \\ &= \sum_{k \in \mathbb{Z}^d} (1 + 4\pi^2 |k|^2)^{-\alpha/2} \cos(2\pi k(x - y)). \end{aligned}$$

Hence,

$$\begin{aligned} B^\alpha(x, x) - B^\alpha(x, y) &= 2 \sum_{k \in \mathbb{Z}^d} (1 + 4\pi^2 |k|^2)^{-\alpha/2} \sin^2(\pi k(x - y)) \\ &\leq 2\pi^2 |x - y|^2 \sum_{|k| \leq |x-y|^{-1}} |k|^2 (1 + 4\pi^2 |k|^2)^{-\alpha/2} + 2 \sum_{|k| > |x-y|^{-1}} (1 + 4\pi^2 |k|^2)^{-\alpha/2} \\ &\leq \begin{cases} c|x-y|^{\alpha-d} & \text{if } d < \alpha < d+2, \\ c|x-y|^2 \log(1+|x-y|^{-1}) & \text{if } \alpha = d+2, \\ c|x-y|^2 & \text{if } \alpha > d+2. \end{cases} \end{aligned}$$

Also observe that the series which defines $B^\alpha(x, x) - B^\alpha(x, y)$ has positive terms and the above inequalities can be reversed. This proves (4) for a torus, and the proof of (3) and (2) is similar. A proof for a generic manifold follows from the representation of Bessel kernels as superposition of heat kernels and the estimates in the previous lemma. In particular, (1) follows from the identity for the Gamma function

$$(1 + \lambda^2)^{-\alpha/2} = \Gamma(\alpha/2)^{-1} \int_0^{+\infty} t^{\alpha/2-1} \exp(-t(1 + \lambda^2)) dt.$$

Since the manifold is compact its diameter is bounded. For ease of notation, in what follows we shall assume that $|x - y| \leq 1$. By Lemma 2.5, for every n ,

$$0 < W(t, x, y) \leq \begin{cases} ct^{(n-d)/2} |x - y|^{-n} & \text{if } 0 < t \leq |x - y|^2, \\ ct^{-d/2} & \text{if } |x - y|^2 \leq t \leq 1, \\ c & \text{if } t \geq 1. \end{cases}$$

Hence, if $0 < \alpha < d$ and $n > d - \alpha$,

$$\begin{aligned} B^\alpha(x, y) &= \Gamma(\alpha/2)^{-1} \int_0^{+\infty} t^{\alpha/2-1} \exp(-t) W(t, x, y) dt \\ &\leq c |x - y|^{-n} \int_0^{|x-y|^2} t^{(\alpha+n-d)/2-1} dt + c \int_{|x-y|^2}^1 t^{(\alpha-d)/2-1} dt + \int_1^{+\infty} t^{\alpha/2-1} \exp(-t) dt \\ &\leq c |x - y|^{\alpha-d}. \end{aligned}$$

Indeed it follows from the estimates of the heat kernel from below (see [9] and [35]) that these inequalities can be reversed. Hence $B^\alpha(x, y) \approx c |x - y|^{\alpha-d}$. This proves (2) when $0 < \alpha < d$, and the proofs of the cases $\alpha = d$ and $\alpha > d$ are similar. Also the proof of (3) is similar. Write

$$\begin{aligned} &B^\alpha(x, y) - B^\alpha(x, z) \\ &= \Gamma(\alpha/2)^{-1} \int_0^{+\infty} t^{\alpha/2-1} \exp(-t) (W(t, x, y) - W(t, x, z)) dt. \end{aligned}$$

Then recall that, by Lemma 2.5,

$$|W(t, x, y) - W(t, x, z)| \leq \begin{cases} ct^{-d/2} & \text{if } 0 < t \leq |y - z|^2, \\ ct^{-(d+1)/2} |y - z| & \text{if } |y - z|^2 \leq t \leq 1, \\ c |y - z| & \text{if } t \geq 1. \end{cases}$$

Hence,

$$\begin{aligned} |B^\alpha(x, y) - B^\alpha(x, z)| &\leq c \int_0^{|y-z|^2} t^{(\alpha-d)/2-1} \exp(-t) dt \\ &+ c |y - z| \int_{|y-z|^2}^1 t^{(\alpha-d-1)/2-1} \exp(-t) dt + c |y - z| \int_1^{+\infty} t^{\alpha/2-1} \exp(-t) dt \\ &\leq c |y - z|^{\alpha-d}. \end{aligned}$$

Finally, the estimate for $|B^\alpha(x, x) - B^\alpha(x, y)|$ in (4) is analogous to the previous one, but it holds in a larger range of α . It suffices to observe that $W(t, x, y)$ is stationary at $x = y$ and it satisfies the estimates

$$|W(t, x, x) - W(t, x, y)| \leq \begin{cases} ct^{-d/2} & \text{if } 0 < t \leq |x - y|^2, \\ ct^{-d/2-1} |x - y|^2 & \text{if } |x - y|^2 \leq t \leq 1, \\ c |x - y|^2 & \text{if } t \geq 1. \end{cases}$$

□

The following is Result (A) in the Introduction.

Theorem 2.7. *For every $d/2 < \alpha < d/2 + 1$ there exists $c > 0$ with the following property: If $\mathcal{M} = U_1 \cup U_2 \cup \dots \cup U_N$ is a decomposition of \mathcal{M} in disjoint pieces with measures $|U_j| = \omega_j$, then there exists a distribution of points $\{z_j\}_{j=1}^N$ with $z_j \in U_j$ such that*

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq c \max_{1 \leq j \leq N} \text{diameter}(U_j)^\alpha \|f\|_{W^{\alpha,2}(\mathcal{M})}.$$

Proof. By Corollary 2.4 (3), with $d\nu(x) = \sum_{j=1}^N \omega_j d\delta_{z_j}(x)$,

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq \left\{ \sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j B^{2\alpha}(z_i, z_j) - 1 \right\}^{1/2} \|f\|_{W^{\alpha,2}}.$$

It suffices to compute the average value of $\sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j B^{2\alpha}(z_i, z_j) - 1$ on $U_1 \times U_2 \times \dots \times U_N$ with respect to the probability measures $\omega_j^{-1} dz_j$ uniformly distributed on U_j . First observe that

$$\begin{aligned} & \left(\prod_{k=1}^N \omega_k^{-1} \right) \int_{U_1} \dots \int_{U_N} dz_1 \dots dz_N = 1, \\ 1 &= \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) dx dy = \sum_{i=1}^N \sum_{j=1}^N \int_{U_i} \int_{U_j} B^{2\alpha}(x, y) dx dy. \end{aligned}$$

Then,

$$\begin{aligned} & \left(\prod_{k=1}^N \omega_k^{-1} \right) \int_{U_1} \dots \int_{U_N} \left(\sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j B^{2\alpha}(z_i, z_j) - 1 \right) dz_1 \dots dz_N \\ &= \sum_j \omega_j \int_{U_j} B^{2\alpha}(z_j, z_j) dz_j + \sum_{i \neq j} \int_{U_i} \int_{U_j} B^{2\alpha}(z_i, z_j) dz_i dz_j \\ & \quad - \sum_j \int_{U_j} \int_{U_j} B^{2\alpha}(x, y) dx dy - \sum_{i \neq j} \int_{U_i} \int_{U_j} B^{2\alpha}(x, y) dx dy \\ &= \sum_{j=1}^N \int_{U_j} \int_{U_j} (B^{2\alpha}(x, x) - B^{2\alpha}(x, y)) dx dy. \end{aligned}$$

Since, by Lemma 2.6 (4), $|B^{2\alpha}(x, x) - B^{2\alpha}(x, y)| \leq c|x - y|^{2\alpha-d}$ when $d < 2\alpha < d+2$,

and since $|U_j| \leq c \text{diameter}(U_j)^d$,

$$\begin{aligned} & \sum_{j=1}^N \int_{U_j} \int_{U_j} |B^{2\alpha}(x, x) - B^{2\alpha}(x, y)| \, dx dy \\ & \leq \sum_{j=1}^N |U_j|^2 \sup_{x, y \in U_j} |B^{2\alpha}(x, x) - B^{2\alpha}(x, y)| \\ & \leq c \sum_{j=1}^N |U_j|^2 \text{diameter}(U_j)^{2\alpha-d} \leq c \sum_{j=1}^N |U_j| \text{diameter}(U_j)^{2\alpha}. \end{aligned}$$

□

For the next result we shall need estimates for partial sums of Fourier expansions of the Bessel kernels.

Lemma 2.8. *Let $-\infty < \alpha < +\infty$, let $\chi(\lambda)$ be an even smooth function on $-\infty < \lambda < +\infty$ with support in $1/2 \leq |\lambda| \leq 2$, and let*

$$P^\alpha(r, x, y) = \sum_{\lambda} \chi(\lambda/r) (1 + \lambda^2)^{-\alpha/2} \varphi_\lambda(x) \overline{\varphi_\lambda(y)}.$$

Then for every positive integer n there exists c such that for every $r > 1$ and $x, y \in \mathcal{M}$,

$$|P^\alpha(r, x, y)| \leq cr^{d-\alpha} (1 + r|x - y|)^{-n}.$$

Proof. The numerology behind this estimate is quite simple. The approximation of the Bessel kernel $B^\alpha(x, y)$ by linear combinations of eigenfunctions with eigenvalues $\lambda^2 < r^2$ is localized and only points x and y with $|x - y| \leq 1/r$ really matter. In particular, since $B^\alpha(x, y)$ is smooth away from the diagonal, at distance $|x - y| \leq 1/r$ the approximation is rough, but at distance $|x - y| \geq 1/r$ it is quite good. The analogue of $P^\alpha(r, x, y)$ in the Euclidean setting is the kernel

$$\begin{aligned} Q(r, x - y) &= \int_{\mathbb{R}^d} \chi(2\pi|\xi|/r) (1 + 4\pi^2|\xi|^2)^{-\alpha/2} \exp(2\pi i(x - y)\xi) \, d\xi \\ &= r^d \int_{\mathbb{R}^d} \chi(2\pi|\xi|) (1 + 4\pi^2r^2|\xi|^2)^{-\alpha/2} \exp(2\pi ir(x - y)\xi) \, d\xi. \end{aligned}$$

Since $\chi(2\pi|\xi|)$ has support in $1/2 \leq 2\pi|\xi| \leq 2$, for every $r > 1$ and $x, y \in \mathbb{R}^d$ one has

$$\begin{aligned} & \left| r^d \int_{\mathbb{R}^d} \chi(2\pi|\xi|) (1 + 4\pi^2r^2|\xi|^2)^{-\alpha/2} \exp(2\pi ir(x - y)\xi) \, d\xi \right| \\ & \leq r^{d-\alpha} \int_{\mathbb{R}^d} (2\pi|\xi|)^{-\alpha} |\chi(2\pi|\xi|)| \, d\xi \leq cr^{d-\alpha}. \end{aligned}$$

This estimate can be improved in the range $|x - y| \geq 1/r$. Indeed, integration by parts gives

$$\begin{aligned} & r^d \int_{\mathbb{R}^d} \chi(2\pi|\xi|) (1 + 4\pi^2 r^2 |\xi|^2)^{-\alpha/2} \exp(2\pi i r (x - y) \xi) d\xi \\ &= r^d \int_{\mathbb{R}^d} \chi(2\pi|\xi|) (1 + 4\pi^2 r^2 |\xi|^2)^{-\alpha/2} \Delta_\xi^n \left((4\pi^2 r^2 |x - y|^2)^{-n} \exp(2\pi i r (x - y) \xi) \right) d\xi \\ &= r^d (4\pi^2 r^2 |x - y|^2)^{-n} \int_{\mathbb{R}^d} \exp(2\pi i r (x - y) \xi) \Delta_\xi^n \left(\chi(2\pi|\xi|) (1 + 4\pi^2 r^2 |\xi|^2)^{-\alpha/2} \right) d\xi. \end{aligned}$$

Hence,

$$\begin{aligned} & \left| r^d \int_{\mathbb{R}^d} \chi(2\pi|\xi|) (1 + 4\pi^2 r^2 |\xi|^2)^{-\alpha/2} \exp(2\pi i r (x - y) \xi) d\xi \right| \\ & \leq r^d (4\pi^2 r^2 |x - y|^2)^{-n} \int_{\mathbb{R}^d} \left| \Delta_\xi^n \left(\chi(2\pi|\xi|) (1 + 4\pi^2 r^2 |\xi|^2)^{-\alpha/2} \right) \right| d\xi \\ & \leq c r^{d-\alpha-2n} |x - y|^{-2n}. \end{aligned}$$

Now it suffices to transfer these estimates from the Euclidean space to the manifold. For the torus, this can be done via the Poisson summation formula. If $Q(r, x - y)$ is the truncated Bessel kernel in \mathbb{R}^d defined above, then the truncated Bessel kernel in \mathbb{T}^d is

$$\sum_{k \in \mathbb{Z}^d} \chi(2\pi|k|/r) (1 + 4\pi^2 |k|^2)^{-\alpha/2} \exp(2\pi i k (x - y)) = \sum_{k \in \mathbb{Z}^d} Q(r, x - y + k).$$

When $|x_j - y_j| \leq 1/2$, the main term in the last sum is the one with $k = 0$, while the contribution of terms with $k \neq 0$ is negligible,

$$\begin{aligned} |Q(r, x - y)| & \leq c r^{d-\alpha} (1 + r |x - y|)^{-n}, \\ \sum_{k \in \mathbb{Z}^d - \{0\}} |Q(r, x - y - k)| & \leq c r^{d-\alpha-n}. \end{aligned}$$

Finally, the estimate for the truncated Bessel kernel on a generic manifold can be obtained by transference from \mathbb{R}^d via pseudodifferential techniques. For more details, see, e.g., [34, Chapter XII], or [5]. For a more general approach on metric measure spaces see [14] and [27]. \square

The following is a result on the homogeneity of measures which appear in quadrature rules and it gives sharp estimates of the discrepancy of such measures. Similar estimates on spheres are in [2].

Lemma 2.9. *Assume that $d\nu(x)$ is a probability measure on \mathcal{M} with the property that for every eigenfunction $\varphi_\lambda(x)$ with eigenvalues $\lambda^2 < r^2$,*

$$\int_{\mathcal{M}} \varphi_\lambda(x) d\nu(x) = \int_{\mathcal{M}} \varphi_\lambda(x) dx.$$

Then for every positive integer n there exists c , which may depend on n and \mathcal{M} , but is independent of r and $d\nu(x)$, such that for every measurable set Ω in \mathcal{M} ,

$$\left| \int_{\Omega} d\nu(x) - \int_{\Omega} dx \right| \leq c \int_{\mathcal{M}} (1 + r \text{ distance } \{x, \partial\Omega\})^{-n} dx.$$

In particular, the discrepancy between the measures $d\nu(x)$ and dx with respect to balls $\{x : |x - y| \leq s\}$ is dominated by

$$\left| \int_{\{|x-y|\leq s\}} d\nu(x) - \int_{\{|x-y|\leq s\}} dx \right| \leq \begin{cases} cr^{-d} & \text{if } s \leq 1/r, \\ cr^{-1}s^{d-1} & \text{if } s \geq 1/r. \end{cases}$$

Proof. It is proved in [11] that given n , there exists c such that for every measurable set Ω in \mathcal{M} and every $r > 0$ there exist two linear combinations of eigenfunctions $A(x) = \sum_{\lambda < r} a(\lambda) \varphi_{\lambda}(x)$ and $B(x) = \sum_{\lambda < r} b(\lambda) \varphi_{\lambda}(x)$ which approximate the characteristic function $\chi_{\Omega}(x)$ from above and below,

$$A(x) \leq \chi_{\Omega}(x) \leq B(x), \quad B(x) - A(x) \leq c(1 + r \text{ distance } \{x, \partial\Omega\})^{-n}.$$

In particular, the properties of the function $A(x)$ and of the measure $d\nu(x)$ give

$$\begin{aligned} \int_{\Omega} d\nu(x) &\geq \int_{\mathcal{M}} A(x) d\nu(x) = \int_{\mathcal{M}} A(x) dx \\ &\geq \int_{\mathcal{M}} \chi_{\Omega}(x) dx - c \int_{\mathcal{M}} (1 + r \text{ distance } \{x, \partial\Omega\})^{-n} dx. \end{aligned}$$

Similarly, by the properties of $B(x)$ and $d\nu(x)$,

$$\begin{aligned} \int_{\Omega} d\nu(x) &\leq \int_{\mathcal{M}} B(x) d\nu(x) = \int_{\mathcal{M}} B(x) dx \\ &\leq \int_{\mathcal{M}} \chi_{\Omega}(x) dx + c \int_{\mathcal{M}} (1 + r \text{ distance } \{x, \partial\Omega\})^{-n} dx. \end{aligned}$$

In particular the choice of $\Omega = \{x : |x - y| \leq s\}$ gives the estimate for the discrepancy of balls. We omit the details. \square

Lemma 2.10. Assume that $d\nu(x)$ is a probability measure on \mathcal{M} which gives an exact quadrature for all eigenfunctions $\varphi_{\lambda}(x)$ with eigenvalues $\lambda^2 < r^2$,

$$\int_{\mathcal{M}} \varphi_{\lambda}(x) d\nu(x) = \int_{\mathcal{M}} \varphi_{\lambda}(x) dx.$$

If $1 \leq q \leq +\infty$ and $\alpha > d(1 - 1/q)$, then there exists c , which may depend on q , α , \mathcal{M} , but is independent of r and $d\nu(x)$, such that

$$\left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^{\alpha}(x, y) d\nu(x) - 1 \right|^q dy \right\}^{1/q} \leq cr^{-\alpha}.$$

Proof. Let $\chi(\lambda)$ be an even smooth function on $-\infty < \lambda < +\infty$, with support in $1/2 \leq |\lambda| \leq 2$ and with the property that $\sum_{j=-\infty}^{+\infty} \chi(2^{-j}\lambda) = 1$ for every $\lambda \neq 0$. Also, let

$$P^\alpha(s, x, y) = \sum_{\lambda} \chi(\lambda/s) (1 + \lambda^2)^{-\alpha/2} \varphi_\lambda(x) \overline{\varphi_\lambda(y)}.$$

Hence, $B^\alpha(x, y) = 1 + \sum_{j=-\infty}^{+\infty} P^\alpha(2^j, x, y)$. Since $d\nu(x)$ annihilates all eigenfunctions with $0 < \lambda < r$, it also annihilates all $P^\alpha(2^j, x, y)$ with $2^j \leq r/2$ and this gives

$$\int_{\mathcal{M}} B^\alpha(x, y) d\nu(x) - 1 = \sum_{2^j > r/2} \int_{\mathcal{M}} P^\alpha(2^j, x, y) d\nu(x).$$

When $q = 1$, by Lemma 2.8 with $n > d$,

$$\begin{aligned} & \int_{\mathcal{M}} \left| \int_{\mathcal{M}} P^\alpha(s, x, y) d\nu(x) \right| dy \\ & \leq cs^{d-\alpha} \int_{\mathcal{M}} \int_{\mathcal{M}} (1 + s|x-y|)^{-n} d\nu(x) dy \\ & \leq cs^{-\alpha} \sup_{x \in \mathcal{M}} \left\{ \int_{\mathcal{M}} s^d (1 + s|x-y|)^{-n} dy \right\} \leq cs^{-\alpha}. \end{aligned}$$

When $q = +\infty$ and $s \geq r$ and $n > d$, by Lemma 2.8 and Lemma 2.9,

$$\begin{aligned} & \sup_{y \in \mathcal{M}} \left\{ \left| \int_{\mathcal{M}} P^\alpha(s, x, y) d\nu(x) \right| \right\} \\ & \leq cs^{d-\alpha} \sup_{y \in \mathcal{M}} \left\{ \int_{\mathcal{M}} (1 + s|x-y|)^{-n} d\nu(x) \right\} \\ & \leq cs^{d-\alpha} \sup_{y \in \mathcal{M}} \left\{ \int_{\{|x-y| \leq 1/r\}} d\nu(x) \right\} + cs^{d-\alpha} \sup_{y \in \mathcal{M}} \left\{ \sum_{j=0}^{+\infty} (2^j s/r)^{-n} \int_{\{2^j/r < |x-y| \leq 2^{j+1}/r\}} d\nu(x) \right\} \\ & \leq cs^{d-\alpha} r^{-d} + cs^{d-\alpha-n} r^{n-d} \leq cs^{d-\alpha} r^{-d}. \end{aligned}$$

Hence, when $s \geq r$ and $1 < q < +\infty$, by interpolation between 1 and $+\infty$,

$$\begin{aligned} & \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} P^\alpha(s, x, y) d\nu(x) \right|^q dy \right\}^{1/q} \\ & \leq \sup_{y \in \mathcal{M}} \left\{ \left| \int_{\mathcal{M}} P^\alpha(s, x, y) d\nu(x) \right| \right\}^{(q-1)/q} \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} P^\alpha(s, x, y) d\nu(x) \right| dy \right\}^{1/q} \\ & \leq cs^{d(1-1/q)-\alpha} r^{-d(1-1/q)}. \end{aligned}$$

When $\alpha > d(1 - 1/q)$ these estimates sum to

$$\begin{aligned} & \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^\alpha(x, y) d\nu(x) - 1 \right|^q dy \right\}^{1/q} \\ & \leq \sum_{2^j > r/2} \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} P^\alpha(2^j, x, y) d\nu(x) \right|^q dy \right\}^{1/q} \\ & \leq cr^{-d(1-1/q)} \sum_{2^j > r/2} 2^{j(d(1-1/q)-\alpha)} \leq cr^{-\alpha}. \end{aligned}$$

□

Finally, the existence of exact quadrature rules associated to any system of continuous functions is a simple result in functional analysis, or in convex geometry.

Lemma 2.11. *Given any collection $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ of real continuous functions on \mathcal{M} , there exist an integer $N \leq n + 1$, points $\{z_j\}_{j=1}^N$ in \mathcal{M} and positive weights $\{\omega_j\}_{j=1}^N$ with $\sum_{j=1}^N \omega_j = 1$, such that for every such $\varphi_i(x)$,*

$$\int_{\mathcal{M}} \varphi_i(x) dx = \sum_{j=1}^N \omega_j \varphi_i(z_j).$$

If the functions $\varphi_i(x)$ are complex valued then the same holds with $N \leq 2n + 1$.

Proof. Define

$$\begin{aligned} \Phi(x) &= (\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)), \\ E &= \int_{\mathcal{M}} \Phi(x) dx = \left(\int_{\mathcal{M}} \varphi_1(x) dx, \int_{\mathcal{M}} \varphi_2(x) dx, \dots, \int_{\mathcal{M}} \varphi_n(x) dx \right). \end{aligned}$$

If all functions $\varphi_i(x)$ are real valued, then $\Phi(x)$ and E are vectors in \mathbb{R}^n . If the $\varphi_i(x)$ are complex, then $\Phi(x)$ and E can be seen as vectors in \mathbb{R}^{2n} . Moreover, E is in the convex hull of the vectors $\Phi(x)$ with $x \in \mathcal{M}$. It then follows from Carathéodory's theorem that E is also a convex combination of at most $n + 1$ vectors $\Phi(x)$ in the real case, or $2n + 1$ in the complex case, $E = \sum_{j=1}^N \omega_j \Phi(z_j)$ with $\omega_j > 0$ and $\sum_{j=1}^N \omega_j = 1$. □

The above result is simple but non constructive. See [32, Theorem 3.1.1], or [31], or [8] for explicit constructions on spheres. The following is Result (B) in the Introduction. Note that in the case of the sphere it is contained in [21].

Theorem 2.12. *Assume that the probability measure $d\nu(x)$ on \mathcal{M} gives an exact quadrature for all eigenfunctions $\varphi_\lambda(x)$ with eigenvalues $\lambda^2 < r^2$,*

$$\int_{\mathcal{M}} \varphi_\lambda(x) d\nu(x) = \int_{\mathcal{M}} \varphi_\lambda(x) dx.$$

Then, for every $1 \leq p \leq +\infty$ and $\alpha > d/p$ there exists a constant $c > 0$ independent of $d\nu(x)$ and r , such that for every function $f(x)$ in $W^{\alpha,p}(\mathcal{M})$,

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq cr^{-\alpha} \|f\|_{W^{\alpha,p}}.$$

Proof. By Corollary 2.4 (1) with $d\mu(x) = d\nu(x) - dx$ and $1/p + 1/q = 1$,

$$\left| \int_{\mathcal{M}} f(x) d\mu(x) \right| \leq \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^\alpha(x, y) d\nu(x) - 1 \right|^q dy \right\}^{1/q} \|f\|_{W^{\alpha,p}}.$$

By the assumption $\int_{\mathcal{M}} \varphi_\lambda(x) d\mu(x) = 0$ for every $\lambda < r$, and Lemma 2.10,

$$\left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^\alpha(x, y) d\nu(x) - 1 \right|^q dy \right\}^{1/q} \leq cr^{-\alpha}.$$

□

Corollary 2.13. *If $1 \leq p \leq +\infty$ and $\alpha > d/p$, then there exists $c > 0$ with the property that for every N there exist a point distribution $\{z_j\}_{j=1}^N$ and associated positive weights $\{\omega_j\}_{j=1}^N$, such that for every function $f(x)$ in $W^{\alpha,p}(\mathcal{M})$,*

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq cN^{-\alpha/d} \|f\|_{W^{\alpha,p}}.$$

Proof. It suffices to show that the above bound holds for infinitely many integers N , say $N_1 < N_2 < N_3 < \dots$ satisfying $N_{r+1} \leq cN_r$. Indeed, introducing multiple nodes and distributing the associated weights among them, gives the result for every positive integer N . Let $r = 1, 2, 3, \dots$ and let n_r be the number of eigenfunctions $\varphi_\lambda(x)$ with $\lambda^2 < r^2$. By Weyl's estimates on the spectrum of an elliptic operator, see [9, Chapter 6.4] or [22, Corollary 17.5.8], we have $c_1 r^d \leq n_r \leq c_2 r^d$. By Lemma 2.11, possibly introducing multiple nodes, there are $N_r = n_r + 1$ nodes $\{z_j\}_{j=1}^{N_r}$ and positive weights $\{\omega_j\}_{j=1}^{N_r}$ such that for all $\lambda^2 < r^2$,

$$\sum_{j=1}^{N_r} \omega_j \varphi_\lambda(z_j) = \int_{\mathcal{M}} \varphi_\lambda(x) dx.$$

Finally, by Theorem 2.12

$$\left| \sum_{j=1}^{N_r} \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq cr^{-\alpha} \|f\|_{W^{\alpha,p}} \leq cn_r^{-\alpha/d} \|f\|_{W^{\alpha,p}} \leq cN_r^{-\alpha/d} \|f\|_{W^{\alpha,p}}.$$

□

The above corollary relies on Lemma 2.11 and guarantees the existence of nodes and weights that give good bounds for the corresponding quadrature rule, but it gives no information on how to find these nodes and weights. In [14] there is a less elementary but somehow stronger result than Lemma 2.11, that essentially says that any choice of well distributed nodes works. For our purposes this result can be restated as follows.

Lemma 2.14. *Let $\{z_j\}_{j=1}^N$ be a distribution of points. Define the mesh norm δ and the minimal separation q of this collection by*

$$\delta = \sup_{x \in \mathcal{M}} \min_j |x - z_j|, \quad q = \min_{i \neq j} |z_i - z_j|.$$

Then there exist positive constants a and b depending only on \mathcal{M} and on $\kappa = \delta/q$, and positive weights $\{w_j\}_{j=1}^N$ with $w_j \geq a\delta^d$, such that for all eigenfunctions $\varphi_\lambda(x)$ with eigenvalues $\lambda^2 < b\delta^{-2}$,

$$\int_{\mathcal{M}} \varphi_\lambda(x) dx = \sum_{j=1}^N \omega_j \varphi_\lambda(z_j).$$

By applying Theorem 2.12 to a point distribution as in the above lemma, one obtains the following corollary, which is result (C) in the Introduction.

Corollary 2.15. *Let $1 \leq p \leq +\infty$ and $\alpha > d/p$. Let $\{z_j\}_{j=1}^N$ be a distribution of points with mesh norm δ and minimal separation distance q . Then there exist a positive constant c depending only on $\kappa = \delta/q$ and on \mathcal{M} , and positive weights $\{\omega_j\}_{j=1}^N$ such that*

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq c N^{-\alpha/d} \|f\|_{W^{\alpha,p}}.$$

Proof. By the above lemma and Theorem 2.12 with $r^2 = b\delta^{-2}$, there exists c_1 such that

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq c_1 \delta^\alpha \|f\|_{W^{\alpha,p}}.$$

By the definition of mesh norm δ and minimal separation distance q , the balls $\{|x - z_j| < \delta\}_{j=1}^N$ cover \mathcal{M} with finite overlapping, that is for some constant c_2 depending only on $\kappa = \delta/q$,

$$\sum_{j=1}^N \chi_{\{|x - z_j| < \delta\}}(x) \leq c_2.$$

See Lemma 4.4 in [14] for the details. It follows that

$$c_3 N \delta^d \leq \int_{\mathcal{M}} \sum_{j=1}^N \chi_{\{|x - z_j| < \delta\}}(x) dx \leq c_2.$$

Therefore $\delta \leq c_4 N^{-1/d}$. □

Finally, easy examples show that the above estimates for the error in approximate quadrature are, up to constants, best possible. Again, see [21] for the case of the sphere. In particular, the following is Result (D) in the Introduction.

Theorem 2.16. *For every $1 \leq p \leq +\infty$ and $\alpha > 0$ there exists $c > 0$ with the following property: For every distribution of points $\{z_j\}_{j=1}^N$ there exists a function $f(x)$ in $W^{\alpha,p}(\mathcal{M})$ which vanishes in a neighborhood of these points and satisfies*

$$\|f\|_{W^{\alpha,p}} \leq cN^{\alpha/d}, \quad \int_{\mathcal{M}} f(x)dx = 1.$$

As a consequence, for every distribution of points $\{z_j\}_{j=1}^N$ and complex weights $\{\omega_j\}_{j=1}^N$, there exists a function $f(x)$ with

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x)dx \right| \geq cN^{-\alpha/d} \|f\|_{W^{\alpha,p}}.$$

Proof. If ε is small, then one can choose $2N$ disjoint balls in \mathcal{M} with diameters $\varepsilon N^{-1/d}$ and, given N points $\{z_j\}_{j=1}^N$, at least N balls have no points inside. On each empty ball construct a bump function $\psi_j(x)$ supported on it with

$$\sup_x |\Delta^k \psi_j(x)| \leq cN^{2k/d}, \quad \int_{\mathcal{M}} \psi_j(x)dx = N^{-1}.$$

The construction of such functions in \mathbb{R}^d can be done by translating and dilating any fixed smooth function with compact support and integral 1. Then one can transport and normalize these functions to the local charts of the manifold and, by compactness, the constant c can be chosen independent of j and N . Define $f(x) = \sum_{j=1}^N \psi_j(x)$. Then,

$$\|f\|_{W^{\alpha,p}} \leq cN^{\alpha/d}, \quad \int_{\mathcal{M}} f(x)dx = 1.$$

The estimate of the $L^p(\mathcal{M})$ norms of $(I + \Delta)^{\alpha/2} f(x)$ when $\alpha/2$ is an integer follows from the fact that $(I + \Delta)^{\alpha/2}$ is a differential operator and the terms $(I + \Delta)^{\alpha/2} \psi_j(x)$ have disjoint supports. The case of $\alpha/2$ not an integer follows from the integer case. The proof when $p = 2$ is elementary. If n is an integer greater than $\alpha/2$, then by Hölder inequality,

$$\begin{aligned} \|f\|_{W^{\alpha,2}} &= \left\{ \sum_{\lambda} (1 + \lambda^2)^{\alpha} |\mathcal{F}f(\lambda)|^2 \right\}^{1/2} \\ &\leq \left\{ \sum_{\lambda} |\mathcal{F}f(\lambda)|^2 \right\}^{(2n-\alpha)/4n} \left\{ \sum_{\lambda} (1 + \lambda^2)^{2n} |\mathcal{F}f(\lambda)|^2 \right\}^{\alpha/4n} \leq cN^{\alpha/d}. \end{aligned}$$

In the general case, let $r > 0$ be a parameter to be fixed later, let n be an integer greater than $\alpha/2$ and, with the same notation as in Lemma 2.8 and Lemma 2.10, decompose $(I + \Delta)^{\alpha/2} f(x)$ into

$$\begin{aligned}
& (I + \Delta)^{\alpha/2} f(x) \\
&= \mathcal{F}f(0) + \sum_{2^j \leq r} \sum_{\lambda} \chi(2^{-j}\lambda) (1 + \lambda^2)^{\alpha/2} \mathcal{F}f(\lambda) \varphi_{\lambda}(x) \\
&+ \sum_{2^j > r} \sum_{\lambda} \chi(2^{-j}\lambda) (1 + \lambda^2)^{(\alpha-2n)/2} (1 + \lambda^2)^n \mathcal{F}f(\lambda) \varphi_{\lambda}(x) \\
&= \mathcal{F}f(0) + \sum_{2^j \leq r} \int_{\mathcal{M}} P^{-\alpha}(2^j, x, y) f(y) dy \\
&+ \sum_{2^j > r} \int_{\mathcal{M}} P^{2n-\alpha}(2^j, x, y) (I + \Delta)^n f(y) dy.
\end{aligned}$$

By Hölder inequality,

$$|\mathcal{F}f(0)| \leq \int_{\mathcal{M}} |f(x)| dx \leq \left\{ \int_{\mathcal{M}} |f(x)|^p dx \right\}^{1/p} \leq c.$$

By Lemma 2.8,

$$\begin{aligned}
& \left\{ \int_{\mathcal{M}} \left| \sum_{2^j \leq r} \int_{\mathcal{M}} P^{-\alpha}(2^j, x, y) f(y) dy \right|^p dx \right\}^{1/p} \\
&\leq \sum_{2^j \leq r} \left\{ \sup_y \int_{\mathcal{M}} |P^{-\alpha}(2^j, x, y)| dx \right\} \left\{ \int_{\mathcal{M}} |f(x)|^p dx \right\}^{1/p} \\
&\leq c \sum_{2^j \leq r} 2^{\alpha j} \left\{ \int_{\mathcal{M}} |f(x)|^p dx \right\}^{1/p} \leq cr^{\alpha}.
\end{aligned}$$

Again, by Lemma 2.8,

$$\begin{aligned}
& \left\{ \int_{\mathcal{M}} \left| \sum_{2^j > r} \int_{\mathcal{M}} P^{2n-\alpha}(2^j, x, y) (I + \Delta)^n f(y) dy \right|^p dx \right\}^{1/p} \\
&\leq \sum_{2^j > r} \left\{ \sup_y \int_{\mathcal{M}} |P^{2n-\alpha}(2^j, x, y)| dx \right\} \left\{ \int_{\mathcal{M}} |(I + \Delta)^n f(x)|^p dx \right\}^{1/p} \\
&\leq c \sum_{2^j > r} 2^{-(2n-\alpha)j} \left\{ \int_{\mathcal{M}} |(I + \Delta)^n f(x)|^p dx \right\}^{1/p} \leq cr^{\alpha-2n} N^{2n/d}.
\end{aligned}$$

Choosing $r = N^{1/d}$, so that $r^{\alpha} = r^{\alpha-2n} N^{2n/d} = N^{\alpha/d}$, one obtains the desired result. \square

3 Further results

The following is Result (E) in the Introduction and it states that a quadrature rule which gives an optimal error in the Sobolev space $W^{\alpha,2}(\mathcal{M})$ is also optimal in all spaces $W^{\beta,2}(\mathcal{M})$ with $d/2 < \beta < \alpha$.

Theorem 3.1. *If $d\nu(x)$ is a probability measure on \mathcal{M} , then the norm of the measure $d\nu(x) - dx$ as a linear functional on $W^{\alpha,2}(\mathcal{M})$ decreases as α increases. Moreover, if the norm of $d\nu(x) - dx$ on $W^{\alpha,2}(\mathcal{M})$ is $r^{-\alpha}$ for some $r > 1$,*

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq r^{-\alpha} \|f\|_{W^{\alpha,2}},$$

then for every $d/2 < \beta < \alpha$ there exists a constant c which may depend on $\alpha, \beta, \mathcal{M}$, but is independent of r and $d\nu(x)$, such that

$$\left| \int_{\mathcal{M}} f(x) d\nu(x) - \int_{\mathcal{M}} f(x) dx \right| \leq cr^{-\beta} \|f\|_{W^{\beta,2}}.$$

Proof. By Corollary 2.4 (2) and (3), the norm of the measure $d\nu(x) - dx$ as a linear functional on $W^{\alpha,2}(\mathcal{M})$ is

$$\left\{ \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\nu(x) d\nu(y) - 1 \right\}^{1/2} = \left\{ \sum_{\lambda > 0} (1 + \lambda^2)^{-\alpha} |\mathcal{F}\nu(\lambda)|^2 \right\}^{1/2}.$$

Since $(1 + \lambda^2)^{-\alpha} \leq (1 + \lambda^2)^{-\beta}$ when $\beta < \alpha$, it follows that this norm is a decreasing function of α . Write $d\nu(x) - dx = d\mu(x)$. By Lemma 2.6 (1), the norm of the functional $\int_{\mathcal{M}} f(x) d\mu(x)$ on $W^{\alpha,2}(\mathcal{M})$ can be written as

$$\begin{aligned} & \left\{ \int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\mu(x) \overline{d\mu(y)} \right\}^{1/2} \\ &= \left\{ \Gamma(\alpha)^{-1} \int_0^{+\infty} t^{\alpha-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t, x, y) d\mu(x) \overline{d\mu(y)} \right) dt \right\}^{1/2}. \end{aligned}$$

Note that

$$\int_{\mathcal{M}} \int_{\mathcal{M}} W(t, x, y) d\mu(x) \overline{d\mu(y)} = \sum_{\lambda} \exp(-\lambda^2 t) |\mathcal{F}\mu(\lambda)|^2 \geq 0.$$

Assuming that this norm is $r^{-\alpha}$, one has to show that the corresponding expression with β instead of α is at most $cr^{-\beta}$. Since $\beta < \alpha$, the integral over $1 \leq t < +\infty$ satisfies the estimate

$$\begin{aligned} & \int_1^{+\infty} t^{\beta-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t, x, y) d\mu(x) \overline{d\mu(y)} \right) dt \\ & \leq \int_1^{+\infty} t^{\alpha-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t, x, y) d\mu(x) \overline{d\mu(y)} \right) dt \\ & \leq \Gamma(\alpha) r^{-2\alpha}. \end{aligned}$$

Similarly, if $\beta < \alpha$ the integral over $r^{-2}/2 \leq t \leq 1$ satisfies the estimate

$$\begin{aligned} & \int_{r^{-2}/2}^1 t^{\beta-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t, x, y) d\mu(x) \overline{d\mu(y)} \right) dt \\ & \leq (r^{-2}/2)^{\beta-\alpha} \int_{r^{-2}/2}^1 t^{\alpha-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t, x, y) d\mu(x) \overline{d\mu(y)} \right) dt \\ & \leq 2^{\alpha-\beta} \Gamma(\alpha) r^{-2\beta}. \end{aligned}$$

Finally, by the small time Gaussian estimate on the heat diffusion kernel in [35], if $0 < t < r^{-2}/2$ then

$$t^{d/2} W(t, x, y) \leq cr^{-d} W(r^{-2}, x, y).$$

It then follows that if $\beta > d/2$ the integral over $0 \leq t \leq r^{-2}/2$ satisfies the estimate

$$\begin{aligned} & \int_0^{r^{-2}/2} t^{\beta-1} \exp(-t) \left(\int_{\mathcal{M}} \int_{\mathcal{M}} W(t, x, y) d\mu(x) \overline{d\mu(y)} \right) dt \\ & \leq cr^{-2\beta} \int_{\mathcal{M}} \int_{\mathcal{M}} W(r^{-2}, x, y) d|\mu|(x) d|\mu|(y). \end{aligned}$$

It remains to show that the last double integral is uniformly bounded in r . Since $d|\mu|(x) \leq d\nu(x) + dx$ and since $\int_{\mathcal{M}} W(r^{-2}, x, y) dx = 1$, replacing $d|\mu|(x)$ with $d\mu(x)$ it suffices to show that

$$\int_{\mathcal{M}} \int_{\mathcal{M}} W(r^{-2}, x, y) d\mu(x) d\mu(y) \leq c.$$

By the assumption on $d\mu(x)$ and the eigenfunction expansion of $W(r^{-2}, x, y)$,

$$\begin{aligned} & \int_{\mathcal{M}} \int_{\mathcal{M}} W(r^{-2}, x, y) d\mu(x) d\mu(y) \\ & \leq r^{-\alpha} \left\| \int_{\mathcal{M}} W(r^{-2}, \cdot, y) d\mu(y) \right\|_{W^{\alpha,2}} \\ & = r^{-\alpha} \left\{ \sum_{\lambda} (1 + \lambda^2)^{\alpha} \exp(-2(\lambda/r)^2) |\mathcal{F}\mu(\lambda)|^2 \right\}^{1/2} \\ & \leq r^{-\alpha} \left\{ \sum_{\lambda} (1 + \lambda^2)^{-\alpha} |\mathcal{F}\mu(\lambda)|^2 \right\}^{1/2} \sup_{\lambda} \left\{ (1 + \lambda^2)^{\alpha} \exp(-(\lambda/r)^2) \right\}. \end{aligned}$$

Finally, the last sum with $\{\mathcal{F}\mu(\lambda)\}$ is the norm of the measure $d\mu(x)$ as a functional on $W^{\alpha,2}(\mathcal{M})$, hence by assumption it equals $r^{-\alpha}$, and the last supremum is dominated by $cr^{2\alpha}$. \square

As we said, the above result is only one way, from α to $\beta < \alpha$. If the norm of $d\nu(x) - dx$ on $W^{\alpha,p}(\mathcal{M})$ is $r^{-\alpha}$ and if $\beta > \alpha$, then one cannot conclude that the norm of $d\nu(x) - dx$ on $W^{\beta,p}(\mathcal{M})$ is at most $cr^{-\beta}$. As a counterexample, it suffices to perturb a good

quadrature rule with nodes $\{z_j\}_{j=1}^N$ and weights $\{\omega_j\}_{j=1}^N$ by moving the last point z_N into a new point t_N , so that the new quadrature differs from the old one by the quantity $\omega_N |f(z_N) - f(t_N)|$. If $\alpha > d/p + 1$ then the function f is differentiable and, by the mean value theorem, $\omega_N |f(z_N) - f(t_N)| \approx \omega_N |z_N - t_N|$. Then, by choosing $|z_N - t_N| = r^{-\alpha}/\omega_N$ one obtains a quadrature rule which gives an error $\approx r^{-\alpha}$ in all spaces $W^{\beta,p}(\mathcal{M})$ with $\beta > \alpha$. The counterexample when $d/p < \alpha \leq d/p + 1$ is slightly more complicated but similar.

In all the above results, the accuracy in a quadrature rule has been estimated in terms of the energy of a measure. It is also possible to estimate this accuracy in terms of a geometric discrepancy. Let $\mathcal{B}(y, t)$ be the level sets of the Bessel kernel,

$$\mathcal{B}(y, t) = \{x \in \mathcal{M} : B^\alpha(x, y) > t\}.$$

Then the Bessel kernel can be decomposed as superposition of the characteristic functions of these level sets,

$$B^\alpha(x, y) = \int_0^{+\infty} \chi_{\mathcal{B}(y,t)}(x) dt.$$

If $1 \leq p, q \leq +\infty$ and $1/p + 1/q = 1$, by Corollary 2.4 (1) and Minkowski inequality, the following Koksma Hlawka type inequality holds:

$$\begin{aligned} \left| \int_{\mathcal{M}} f(x) d\mu(x) \right| &\leq \|f\|_{W^{\alpha,p}} \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} B^\alpha(x, y) d\mu(x) \right|^q dy \right\}^{1/q} \\ &\leq \|f\|_{W^{\alpha,p}} \int_0^{+\infty} \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right|^q dy \right\}^{1/q} dt. \end{aligned}$$

The quantity $\left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right|$ is the discrepancy of the measure $d\mu(x)$ with respect to the level set $\mathcal{B}(y, t)$. It can be proved that, for specific measures and at least in the range $0 < \alpha < 1$, the above estimates are sharp and they can lead to optimal quadrature rules. In particular, the following is Result (F) in the Introduction.

Theorem 3.2. *Denote by $\delta(t)$ the supremum with respect to y of the diameters of the level sets $\{\mathcal{B}(y, t)\}$ and assume that there exists $r \geq 1$ such that the discrepancy of the measure $d\mu(x)$ with respect to $\{\mathcal{B}(y, t)\}$ satisfies the estimates*

$$\left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right| \leq \begin{cases} r^{-d} & \text{if } \delta(t) \leq 1/r, \\ r^{-1} \delta(t)^{d-1} & \text{if } \delta(t) \geq 1/r. \end{cases}$$

Also assume that $1 \leq p \leq +\infty$ and $\alpha > d/p$. Then there exists a constant c , which may depend on α and p and on the total variation of the measure $|\mu|(\mathcal{M})$, but is independent of r , such that

$$\left| \int_{\mathcal{M}} f(x) d\mu(x) \right| \leq \begin{cases} cr^{-\alpha} \|f\|_{W^{\alpha,p}} & \text{if } 0 < \alpha < 1, \\ cr^{-1} \log(1+r) \|f\|_{W^{\alpha,p}} & \text{if } \alpha = 1, \\ cr^{-1} \|f\|_{W^{\alpha,p}} & \text{if } \alpha > 1. \end{cases}$$

Proof. If $1 \leq p, q \leq +\infty$ and $1/p + 1/q = 1$, by Corollary 2.4 (1) and Minkowski inequality,

$$\left| \int_{\mathcal{M}} f(x) d\mu(x) \right| \leq \|f\|_{W^{\alpha,p}} \int_0^{+\infty} \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right|^q dy \right\}^{1/q} dt.$$

By Lemma 2.6 (2), when $0 < \alpha < d$ then $B^\alpha(x, y) \approx |x - y|^{\alpha-d}$, and the level sets $\mathcal{B}(y, t)$ have diameters of order $\min\{1, t^{1/(\alpha-d)}\}$. Hence, writing $q = (q-1) + 1$, the estimate of the discrepancy of small level sets with $t \geq r^{d-\alpha}$ gives

$$\begin{aligned} & \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right|^q dy \right\}^{1/q} \\ & \leq \left\{ \sup_{y \in \mathcal{M}} \left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right| \right\}^{(q-1)/q} \left\{ \int_{\mathcal{M}} \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d|\mu|(x) dy \right\}^{1/q} \\ & \leq \left\{ \sup_{y \in \mathcal{M}} \left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right| \right\}^{(q-1)/q} \{c|\mu|(\mathcal{M})t^{d/(\alpha-d)}\}^{1/q} \\ & \leq cr^{-d(q-1)/q} t^{d/q(\alpha-d)}. \end{aligned}$$

Hence, if $\alpha > d/p$ the integral over $r^{d-\alpha} \leq t < +\infty$ satisfies the inequality

$$\begin{aligned} & \int_{r^{d-\alpha}}^{+\infty} \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right|^q dy \right\}^{1/q} dt \\ & \leq cr^{-d(q-1)/q} \int_{r^{d-\alpha}}^{+\infty} t^{d/q(\alpha-d)} dt \leq cr^{-\alpha}. \end{aligned}$$

Similarly, the integral over $0 \leq t \leq r^{d-\alpha}$, that is the discrepancy of large level sets, satisfies the inequality

$$\begin{aligned} & \int_0^{r^{d-\alpha}} \left\{ \int_{\mathcal{M}} \left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right|^q dy \right\}^{1/q} dt \\ & \leq r^{-1} \int_0^{r^{d-\alpha}} \min\{1, t^{(d-1)/(\alpha-d)}\} dt \leq \begin{cases} cr^{-\alpha} & \text{if } 0 < \alpha < 1, \\ cr^{-1} \log(1+r) & \text{if } \alpha = 1, \\ cr^{-1} & \text{if } \alpha > 1. \end{cases} \end{aligned}$$

The proof in the case $\alpha = d$ is similar and it follows from the estimate $B^\alpha(x, y) \approx -\log(|x - y|)$. The proof in the case $\alpha > d$ is even simpler, since in this case $B^\alpha(x, y)$ is bounded and it suffices to integrate on $0 \leq t \leq \sup_{x,y \in \mathcal{M}} B^\alpha(x, y)$ the inequality $\left| \int_{\mathcal{M}} \chi_{\mathcal{B}(y,t)}(x) d\mu(x) \right| \leq cr^{-1}$. \square

Observe that the hypotheses on the discrepancy in the above theorem match the estimates in Lemma 2.9. Indeed, by this lemma, the measures $d\nu(x)$ which give exact quadrature for eigenfunctions with eigenvalues $\lambda^2 < r^2$ have discrepancy

$$\left| \int_{\{|x-y| \leq s\}} d\nu(x) - \int_{\{|x-y| \leq s\}} dx \right| \leq \begin{cases} cr^{-d} & \text{if } s \leq 1/r, \\ cr^{-1}s^{d-1} & \text{if } s \geq 1/r. \end{cases}$$

Actually, these estimates hold not only for balls $\{|x - y| \leq s\}$, but also for sets with boundaries with finite $d - 1$ dimensional Minkowski measure, such as the level sets $\{\mathcal{B}(y, t)\}$. Also observe that these estimates are natural, since the discrepancy of large sets is qualitatively different from the one of small sets. In particular, it follows from Lemma 2.9, Theorem 2.12, Theorem 2.16, that, at least in the range $0 < \alpha < 1$, Theorem 3.2 gives an optimal quadrature. We conclude with a series of remarks.

Remark 3.3. As we said, the assumption $\alpha > d/2$ with $p = 2$ in Theorem 2.7, or $\alpha > d/p$ with $1 \leq p \leq +\infty$ in Theorem 2.12, guarantees the boundedness and continuity of $f(x)$, otherwise the point evaluation $f(z_j)$ may be not defined. This follows from the Sobolev imbedding theorem. Indeed, the imbedding is an easy corollary of Lemma 2.6. A function $f(x)$ is in the Sobolev space $W^{\alpha,p}(\mathcal{M})$ if and only if there exists a function $g(x)$ in $L^p(\mathcal{M})$ with

$$f(x) = \int_{\mathcal{M}} B^\alpha(x, y)g(y)dy.$$

When $1 \leq p, q \leq +\infty$, $1/p + 1/q = 1$, $d/p < \alpha < d$, then $B^\alpha(x, y) \leq c|x - y|^{\alpha-d}$ is in $L^q(\mathcal{M})$ and this implies that distributions in the Sobolev space $W^{\alpha,p}(\mathcal{M})$ with $\alpha > d/p$ are continuous functions. Indeed they are also Hölder continuous of order $\alpha - d/p$.

Remark 3.4. When the manifold is a Lie group or a homogeneous space, one can restate Theorem 2.1 in terms of convolutions. In the particular case of the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, let

$$A(x) = \sum_{k \in \mathbb{Z}^d} \psi(k) \exp(2\pi i k x), \quad B(x) = \sum_{k \in \mathbb{Z}^d} \psi(k)^{-1} \exp(2\pi i k x).$$

Then, if $1 \leq p, q, r \leq +\infty$ with $1/p + 1/q = 1/r + 1$,

$$\begin{aligned} \left\{ \int_{\mathbb{T}^d} \left| \int_{\mathbb{T}^d} f(x-y) d\mu(y) \right|^r dx \right\}^{1/r} &= \left\{ \int_{\mathbb{T}^d} |B * A * f * \mu(x)|^r dx \right\}^{1/r} \\ &\leq \left\{ \int_{\mathbb{T}^d} |A * f(x)|^p dx \right\}^{1/p} \left\{ \int_{\mathbb{T}^d} |B * \mu(x)|^q dx \right\}^{1/q}. \end{aligned}$$

In the case of the sphere $\mathbb{S}^d = \{x \in \mathbb{R}^{d+1}, |x| = 1\}$, let $\{Z_n(xy)\}$ be the system of zonal spherical harmonics polynomials and let

$$A(xy) = \sum_{n=0}^{+\infty} \psi(n) Z_n(xy), \quad B(xy) = \sum_{n=0}^{+\infty} \psi(n)^{-1} Z_n(xy).$$

Then, if $1 \leq p, q \leq +\infty$ with $1/p + 1/q = 1$,

$$\begin{aligned} &\left| \int_{\mathbb{S}^d} f(x) d\mu(x) \right| \\ &\leq \left\{ \int_{\mathbb{S}^d} \left| \int_{\mathbb{S}^d} A(xy) f(y) dy \right|^p dx \right\}^{1/p} \left\{ \int_{\mathbb{S}^d} \left| \int_{\mathbb{S}^d} B(xy) d\mu(y) \right|^q dx \right\}^{1/q}. \end{aligned}$$

Both results on the torus and the sphere follow from Young inequality for convolutions.

Remark 3.5. A result related to Theorem 2.1 and to the previous remark is the following. Identify \mathbb{T}^d with the unit cube $\{(x_1, \dots, x_d) : 0 \leq x_j < 1\}$ and denote by $\chi_{P(y)}(x)$ the characteristic function of the parallelepiped $P(y) = \{(x_1, \dots, x_d) : 0 \leq x_j < y_j\}$. Then define

$$\begin{aligned} B(x) &= \int_{\mathbb{T}^d} \chi_{P(y)}(x) dy - 2^{-d} = \prod_{j=1}^d (1 - x_j) - 2^{-d} \\ &= \sum_{k \in \mathbb{Z}^d - \{0\}} \left(\left(\prod_{k_j=0} 2 \right) \left(\prod_{k_j \neq 0} 2\pi i k_j \right) \right)^{-1} \exp(2\pi i k x). \end{aligned}$$

Also, define the differential integral operator

$$\begin{aligned} A * f(x) &= \sum_{k \neq 0} \left(\prod_{k_j=0} 2 \right) \left(\prod_{k_j \neq 0} 2\pi i k_j \right) \widehat{f}(k) \exp(2\pi i k x) \\ &= 2^{d-1} \sum_{1 \leq j \leq d} \int_{\mathbb{T}^{d-1}} \frac{\partial}{\partial x_j} f(x) \prod_{i \neq j} dx_i + 2^{d-2} \sum_{1 \leq i \neq j \leq d} \int_{\mathbb{T}^{d-2}} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \prod_{h \neq i, j} dx_h \\ &\quad + \dots + \frac{\partial^d}{\partial x_1 \dots \partial x_d} f(x). \end{aligned}$$

Observe that, as in Theorem 2.1, the Fourier coefficients of the distribution $A(x)$ and of the function $B(x)$ are one inverse to the other, however here the Fourier coefficients are indexed by the lattice points $2\pi i k$, and not by the eigenvalues $4\pi^2 |k|^2$. If $d\nu(x) = N^{-1} \sum_{j=1}^N d\delta_{z_j}(x)$, and if $1 \leq p, q, r \leq +\infty$ with $1/p + 1/q = 1/r + 1$, then

$$\begin{aligned} &\left\{ \int_{\mathbb{T}^d} \left| N^{-1} \sum_{j=1}^N f(x - z_j) - \int_{\mathbb{T}^d} f(y) dy \right|^r dx \right\}^{1/r} \\ &\leq \left\{ \int_{\mathbb{T}^d} |A * f(x)|^p dx \right\}^{1/p} \left\{ \int_{\mathbb{T}^d} |B * \nu(x)|^q dx \right\}^{1/q}. \end{aligned}$$

The norm of $A * f(x)$ is dominated by an analogue of the Hardy Krause variation,

$$\begin{aligned} &\left\{ \int_{\mathbb{T}^d} |A * f(x)|^p dx \right\}^{1/p} \\ &\leq 2^{d-1} \sum_{1 \leq j \leq d} \left\{ \int_{\mathbb{T}} \left| \int_{\mathbb{T}^{d-1}} \frac{\partial}{\partial x_j} f(x) \prod_{i \neq j} dx_i \right|^p dx_j \right\}^{1/p} \\ &+ 2^{d-2} \sum_{1 \leq i \neq j \leq d} \left\{ \int_{\mathbb{T}^2} \left| \int_{\mathbb{T}^{d-2}} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \prod_{h \neq i, j} dx_h \right|^p dx_i dx_j \right\}^{1/p} \\ &\quad + \dots + \left\{ \int_{\mathbb{T}^d} \left| \frac{\partial^d}{\partial x_1 \dots \partial x_d} f(x) \right|^p dx \right\}^{1/p}. \end{aligned}$$

The norm of $B * \nu(x)$ is dominated by the discrepancy of the points $\{z_j\}_{j=1}^N$ with respect to the family of boxes $P(y)$,

$$\begin{aligned} & \left\{ \int_{\mathbb{T}^d} |B * \nu(x)|^q dx \right\}^{1/q} \\ & \leq \int_{\mathbb{T}^d} \left\{ \int_{\mathbb{T}^d} \left| N^{-1} \sum_{j=1}^N \chi_{P(y)}(z_j + x) - \prod_{j=1}^d y_j \right|^q dx \right\}^{1/q} dy. \end{aligned}$$

In particular, the case $p = 1$ and $q = +\infty$ is an analogue of the Koksma Hlawka inequality. See [24]. A generalization of this classical inequality is contained in [6].

Remark 3.6. By Lemma 2.6 (1), the Bessel kernel $B^\alpha(x, y)$ with $\alpha > 0$ is a superposition of heat kernels $W(t, x, y)$. Indeed, it is possible to state an analogue of Corollary 2.4 in terms of the heat kernel, without explicit mention of Bessel potentials: If $\{z_j\}_{j=1}^N$ is a sequence of points in \mathcal{M} , if $\{\omega_j\}_{j=1}^N$ are positive weights with $\sum_j \omega_j = 1$, and if $f(x)$ is a function in $W^{\alpha,p}(\mathcal{M})$ with $\alpha > d/2$, then

$$\begin{aligned} & \left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \\ & \leq \left\{ \Gamma(\alpha)^{-1} \int_0^{+\infty} \left| \sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j W(t, z_i, z_j) - 1 \right| t^{\alpha-1} \exp(-t) dt \right\}^{1/2} \|f\|_{W^{\alpha,2}}. \end{aligned}$$

This suggests the following heuristic interpretation: Mathematically, a set of points on a manifold is well distributed if the associated Riemann sums are close to the integrals. Physically, a set of points is well distributed if the heat, initially concentrated on them, in a short time diffuses uniformly across the manifold.

Remark 3.7. In order to minimize the errors in the numerical integration in Corollary 2.4 (3), one has to minimize the energies

$$\int_{\mathcal{M}} \int_{\mathcal{M}} B^{2\alpha}(x, y) d\nu(x) d\nu(y), \quad \sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j B^{2\alpha}(z_i, z_j).$$

These are analogous to the energy integrals in potential theory

$$\int_{\mathcal{M}} \int_{\mathcal{M}} |x - y|^{-\varepsilon} d\nu(x) d\nu(y).$$

See [15]. When $d < \alpha < d + 1$ the kernel $B^{2\alpha}(x, y)$ is positive and bounded, with a maximum at $x = y$ and a spike $A - B|x - y|^{2\alpha-d}$ when $x \rightarrow y$. In particular, the gradient at $x = y$ is infinite. This implies that in order to minimize the discrete energy $\sum_{i,j} \omega_i \omega_j B^{2\alpha}(z_i, z_j)$ the points $\{z_j\}$ have to be well separated. This suggests the

following heuristic interpretation: Mathematically, a set of points on a manifold is well distributed if the energy is minimal. Physically, a set of points, free to move and repelling each other according to some law, is well distributed when they reach an equilibrium.

Remark 3.8. It can be proved that if $2\alpha > d + 2$ then

$$|B^{2\alpha}(x, x) - B^{2\alpha}(x, y)| \leq c|x - y|^2.$$

This estimate in the proof of Theorem 2.7 yields that for most choices of sampling points $z_j \in U_j$,

$$\left| \sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx \right| \leq c \max_{1 \leq j \leq N} \left\{ \text{diameter}(U_j)^{d/2+1} \right\} \|f\|_{W^{\alpha, 2}(\mathcal{M})}.$$

The same result holds if $2\alpha = d + 2$, with a logarithmic transgression. Observe that these estimates hold for most choices of sampling points, but not for all choices. Indeed, if the manifold \mathcal{M} is decomposed in disjoint pieces $\mathcal{M} = U_1 \cup U_2 \cup \dots \cup U_N$ with measure $a_1 N^{-1} \leq |U_j| = \omega_j \leq a_2 N^{-1}$ and $b_1 N^{-1/d} \leq \text{diameter}(U_j) \leq b_2 N^{-1/d}$, if $f(x)$ is a smooth non constant function and if the points $z_j \in U_j$ are the maxima of $f(x)$ in U_j , then $\sum_{j=1}^N \omega_j f(z_j)$ is an upper sum of the integral $\int_{\mathcal{M}} f(x) dx$ and

$$\sum_{j=1}^N \omega_j f(z_j) - \int_{\mathcal{M}} f(x) dx = \sum_{j=1}^N \int_{U_j} (f(z_j) - f(x)) dx \geq cN^{-1/d}.$$

Remark 3.9. Theorem 3.2 gives an estimate of the accuracy in a quadrature rule in terms of the discrepancy of a measure with respect to level sets of the Bessel kernel. The following argument shows that when the manifold is a sphere, or a rank-one compact symmetric space, then the level sets of the heat kernel $\{W(t, x, y) > s\}$, and hence of the Bessel kernels $\{B^\alpha(x, y) \leq t\}$, are geodesic balls $\{|x - y| \leq r\}$. The Laplace operator on the sphere \mathbb{S}^d with respect to a system of polar coordinates $x = (\vartheta, \sigma)$, with $0 \leq \vartheta \leq \pi$ the colatitude with respect to a given pole and $\sigma \in \mathbb{S}^{d-1}$ the longitude, is

$$\Delta_x = \Delta_{(\vartheta, \sigma)} = -\sin^{1-d}(\vartheta) \frac{\partial}{\partial \vartheta} \left(\sin^{d-1}(\vartheta) \frac{\partial}{\partial \vartheta} \right) + \sin^{-2}(\vartheta) \Delta_\sigma.$$

Let $u(t, x)$ be the solution of the Cauchy problem for the heat equation

$$\begin{cases} \frac{\partial}{\partial t} u(t, x) = -\Delta_x u(t, x), \\ u(0, x) = f(x). \end{cases}$$

If $f(x)$ depends only on the colatitude ϑ , and if it is even and decreasing in $0 < \vartheta < \pi$, then also $u(t, x)$ depends only on the colatitude and it is even and decreasing in $0 < \vartheta < \pi$.

π . In order to prove this, set $u(t, x) = U(t, \vartheta)$, $f(x) = F(\vartheta)$, and $\sin^{d-1}(\vartheta) \partial U(t, \vartheta) / \partial \vartheta = V(t, \vartheta)$. Then

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \vartheta} \frac{\partial}{\partial t} U(t, \vartheta) = \frac{\partial}{\partial \vartheta} \left\{ \sin^{1-d}(\vartheta) \frac{\partial}{\partial \vartheta} \left(\sin^{d-1}(\vartheta) \frac{\partial}{\partial \vartheta} U(t, \vartheta) \right) \right\}, \\ \frac{\partial}{\partial \vartheta} U(0, \vartheta) = \frac{\partial}{\partial \vartheta} F(\vartheta), \\ \left\{ \begin{array}{l} \frac{\partial}{\partial t} V(t, \vartheta) = \frac{\partial^2}{\partial \vartheta^2} V(t, \vartheta) + (1-d) \frac{\cos(\vartheta)}{\sin(\vartheta)} \frac{\partial}{\partial \vartheta} V(t, \vartheta), \\ V(0, \vartheta) = \sin^{d-1}(\vartheta) \frac{\partial}{\partial \vartheta} F(\vartheta), \\ V(t, 0) = V(t, \pi) = 0. \end{array} \right. \end{array} \right.$$

If $F(\vartheta)$ is decreasing in $0 < \vartheta < \pi$, then $V(0, \vartheta) \leq 0$ and, by the maximum principle, $V(t, \vartheta) \leq 0$, hence $U(t, \vartheta)$ is decreasing in $0 < \vartheta < \pi$. In particular, by considering a sequence of initial data $\{f_n(x)\}$ which depend only on the colatitude ϑ , even and decreasing in $0 < \vartheta < \pi$, and which converge to the Dirac $\delta(x)$, one proves that the heat kernel $W(t, \cos(\vartheta))$ is decreasing in $0 < \vartheta < \pi$. Since Bessel kernels are superpositions of heat kernels, they are also superpositions of spherical caps.

Remark 3.10. In [3] and [26] the discrepancy of orbits of discrete subgroups of rotations of a sphere are studied. Let \mathcal{G} be a compact Lie group, \mathcal{K} a closed subgroup, $\mathcal{M} = \mathcal{G}/\mathcal{K}$ a homogeneous space of dimension d . Also, let \mathcal{H} be a finitely generated free subgroup in \mathcal{G} and assume that the action of \mathcal{H} on \mathcal{M} is free. Given a positive integer n , let $\{\sigma_j\}_{j=1}^N$ be an ordering of the elements in \mathcal{H} with length at most n and for every function $f(x)$ on \mathcal{M} , define

$$Tf(x) = N^{-1} \sum_{j=1}^N f(\sigma_j x).$$

This operator is self-adjoint and it has eigenvalues and eigenfunctions in $L^2(\mathcal{M})$. Moreover, since the operators T and Δ commute, they have a common orthonormal system of eigenfunctions, $\Delta \varphi_\lambda(x) = \lambda^2 \varphi_\lambda(x)$ and $T \varphi_\lambda(x) = T(\lambda) \varphi_\lambda(x)$. All eigenvalues of T have modulus at most 1 and indeed 1 is an eigenvalue and the constants are eigenfunctions. Assume that all non constant eigenfunctions have eigenvalues much smaller than 1. Then, if $\alpha > d/2$,

$$\begin{aligned} & \left| N^{-1} \sum_{j=1}^N f(\sigma_j x) - \int_{\mathcal{M}} f(x) dx \right| = \left| \sum_{\lambda \neq 0} T(\lambda) \mathcal{F}f(\lambda) \varphi_\lambda(x) \right| \\ & \leq \left\{ \sup_{\lambda \neq 0} |T(\lambda)| \right\} \left\{ \sum_{\lambda} (1 + \lambda^2)^\alpha |\mathcal{F}f(\lambda)|^2 \right\}^{1/2} \left\{ \sum_{\lambda} (1 + \lambda^2)^{-\alpha} |\varphi_\lambda(x)|^2 \right\}^{1/2} \\ & \leq c \left\{ \sup_{\lambda \neq 0} |T(\lambda)| \right\} \left\{ \int_{\mathcal{M}} |(I + \Delta)^{\alpha/2} f(x)|^2 dx \right\}^{1/2}. \end{aligned}$$

The absolute convergence of the above series is consequence of the Sobolev's imbeddings, or the Weyl's estimates for eigenfunctions. In particular, when $\mathcal{M} = SO(3)/SO(2)$ is the two dimensional sphere and \mathcal{H} is the free group generated by rotations of angles $\arccos(-3/5)$ around orthogonal axes, it has been proved in [26] that the eigenvalues of the operator T satisfy the Ramanujan bounds

$$\sup_{\lambda \neq 0} \{|T(\lambda)|\} \leq cN^{-1/2} \log(N).$$

Hence, for the sphere,

$$\begin{aligned} & \left| N^{-1} \sum_{j=1}^N f(\sigma_j x) - \int_{\mathcal{M}} f(x) dx \right| \\ & \leq cN^{-1/2} \log(N) \left\{ \int_{\mathcal{M}} \left| (I + \Delta)^{\alpha/2} f(x) \right|^2 dx \right\}^{1/2}. \end{aligned}$$

All of this is essentially contained in [26]. Although this bound $N^{-1/2} \log(N)$ is worse than the bound $N^{-\alpha/2}$ in Corollary 2.13, the matrices $\{\sigma_j\}$ have rational entries and the sampling points $\{\sigma_j x\}$ are completely explicit.

References

- [1] C. Amstler and P. Zinterhof, *Uniform distribution, discrepancy, and reproducing kernel Hilbert spaces*, J. Complexity **17** (2001), 497-515.
- [2] V. V. Andrievskii, H.-P. Blatt and M. Götz, *Discrepancy estimates on the sphere*, Monatsh. Math. **128** (1999), 179-188.
- [3] V. I. Arnold and A. L. Krylov, *Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain*, Dokl. Akad. Nauk. SSSR **148** (1963), 9-12.
- [4] J. Beck and W. W. L. Chen, "Irregularities of distribution", Cambridge University Press, 1987.
- [5] L. Brandolini and L. Colzani, *Decay of Fourier transforms and summability of eigenfunction expansions*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **29** (2000), 611-638.
- [6] L. Brandolini, L. Colzani, G. Gigante and G. Travaglini, *On the Koksma-Hlawka inequality*, submitted to J. Complexity.
- [7] J. S. Brauchart and K. Hesse, *Numerical integration over spheres of arbitrary dimension*, Constr. Approx. **25** (2007), 41-71.

- [8] G. Brown, D. Feng and S. Y. Sheng, *Kolmogorov width of classes of smooth functions on the sphere S^{d-1}* , J. Complexity **18** (2002), 1001-1023.
- [9] I. Chavel, "Eigenvalues in Riemannian geometry", Academic Press (1984).
- [10] C. Choirat and R. Seri, *Numerical properties of generalized discrepancies on spheres of arbitrary dimension*, submitted to J. Complexity.
- [11] L. Colzani, G. Gigante and G. Travaglini, *Trigonometric approximation and a general form of the Erdős Turán inequality*, Trans. Amer. Math. Soc. **363** (2011), 1101-1123.
- [12] J. Cui and W. Freeden, *Equidistribution on the sphere*, SIAM J. Sci. Comput. **18** (1997), 595-609.
- [13] S. B. Damelin, J. Levesley, D. L. Ragozin and X. Sun, *Energies, group-invariant kernels and numerical integration on compact manifolds*, J. Complexity **25** (2009), 152-162.
- [14] F. Filbir and H. N. Mhaskar, *A quadrature formula for diffusion polynomials corresponding to a generalized heat kernel*, J. Fourier Anal. Appl. **16** (2010), 629-657
- [15] D. P. Hardin and E. B. Saff, *Discretizing manifolds via minimum energy points*, Notices Amer. Math. Soc. **51** (2004), 1186-1194.
- [16] D. P. Hardin, E. B. Saff and J. T. Whitehouse, *Quasi-uniformity of minimal weighted energy points on compact metric spaces*, J. Complexity **28** (2012), 177-191.
- [17] K. Hesse, *A lower bound for the worst-case cubature error on spheres of arbitrary dimensions*, Numer. Math. **103** (2006), 413-433.
- [18] K. Hesse and I. H. Sloan, *Optimal lower bounds for cubature error on the sphere S^2* , J. Complexity **21** (2005), 790-803.
- [19] K. Hesse and I. H. Sloan, *Cubature over the sphere S^2 in Sobolev spaces of arbitrary order*, J. Approx. Theory **141** (2006), 118-133.
- [20] K. Hesse and I. H. Sloan, *Worst-case errors in a Sobolev space setting for cubature over the sphere S^2* , Bull. Austral. Math. Soc. **71** (2005), 81-105.
- [21] K. Hesse, H. N. Mhaskar and I. H. Sloan, *Quadrature in Besov spaces on the Euclidean sphere*, J. Complexity **23** (2007), 528-552.
- [22] L. Hörmander, "The analysis of linear partial differential operators, I, II, III, IV", Springer Verlag (1983-1985).

- [23] E. P. Hsu, *Estimates of the derivatives of the heat kernel on a compact manifold*, Proc. Amer. Math. Soc. **127** (1999), 3739-3744
- [24] L. Kuipers and H. Niederreiter, “Uniform distribution of sequences”, Pure and Applied Mathematics, Wiley-Interscience (1974).
- [25] A. Kushpel, *Optimal cubature formulas on compact homogeneous manifolds*, J. Funct. Anal. **257** (2009), 1621-1629.
- [26] A. Lubotzky, R. Phillips and P. Sarnak, *Hecke operators and distributing points on the sphere I, II*, Comm. Pure Appl. Math. **39** (1986), 149-186, **40** (1987), 401-420.
- [27] M. Maggioni and H. N. Mhaskar, *Diffusion polynomial frames on metric measure spaces*, Appl. Comput. Harmon. Anal. **24** (2008), 329-353.
- [28] J. Matoušek, “Geometric discrepancy. An illustrated guide”, Springer Verlag (1999).
- [29] E. Novak, “Deterministic and stochastic error bounds in numerical analysis”, Lecture Notes in Mathematics 1349, Springer Verlag (1988).
- [30] E. B. Saff and A. B. J. Kuijlaars, *Distributing many points on a sphere*, Math. Intelligencer **19** (1997), 5-11.
- [31] P. D. Seymour and T. Zaslavsky, *Averaging sets: a generalization of mean values and spherical designs*, Adv. in Math. **52** (1984), 213-240.
- [32] H. S. Shapiro, “Topics in approximation theory”, Lecture Notes in Mathematics 187, Springer Verlag (1971).
- [33] D. W. Stroock, J. Turetsky, *Upper bounds on the derivatives of the logarithm of the heat kernel*, Comm. Anal. Geom. **6** (1998), 669-685
- [34] M. E. Taylor, “Pseudodifferential operators”, Princeton University Press (1981).
- [35] N. T. Varopoulos, *Small time Gaussian estimates of heat diffusion kernels. II. The theory of large deviations*, J. Funct. Anal. **93** (1990), 1-33.

L. BRANDOLINI: Dipartimento di Ingegneria, Università di Bergamo, Viale Marconi 5, 24044 Dalmine, Bergamo, Italia;

e-mail: luca.brandolini@unibg.it

C. CHOIRAT: Department of Economics, School of Economics and Business Management, Universidad de Navarra, Edificio de Bibliotecas (Entrada Este), 31080 Pamplona, Spain;

e-mail: cchoirat@unav.es

L. COLZANI: Dipartimento di Matematica e Applicazioni, Edificio U5, Università di Milano Bicocca, Via R. Cozzi 53, 20125 Milano, Italia;

e-mail: leonardo.colzani@unimib.it

G. GIGANTE: Dipartimento di Ingegneria, Università di Bergamo, Viale Marconi 5, 24044 Dalmine, Bergamo, Italia;

e-mail: giacomo.gigante@unibg.it

R. SERI: Università degli Studi dell'Insubria, Dipartimento di Economia, Via Monte Generoso 71, 21100 Varese, Italia;

e-mail: raffaello.seri@uninsubria.it

G. TRAVAGLINI (corresponding author): Dipartimento di Statistica, Edificio U7, Università di Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italia;

e-mail: giancarlo.travaglini@unimib.it