QMC DESIGNS: OPTIMAL ORDER QUASI MONTE CARLO INTEGRATION SCHEMES ON THE SPHERE

J. S. BRAUCHART, E. B. SAFF, I. H. SLOAN, AND R. S. WOMERSLEY

ABSTRACT. We study equal weight numerical integration, or Quasi Monte Carlo (QMC) rules, for functions in a Sobolev space $\mathbb{H}^s(\mathbb{S}^d)$ with smoothness parameter $s > d/2$ defined over the unit sphere $\mathbb{S}^d$ in $\mathbb{R}^{d+1}$. Focusing on N-point configurations that achieve optimal order QMC error bounds (as is the case for efficient spherical designs), we are led to introduce the concept of QMC designs: these are sequences of N-point configurations X_N on $\mathbb{S}^d$ such that the worst-case error satisfies

$$\sup_{\substack{f \in \mathbb{H}^s(\mathbb{S}^d), \\ \|f\|_{\mathbb{H}^s} \leq 1}} \left| \frac{1}{N} \sum_{\mathbf{x} \in X_N} f(\mathbf{x}) - \int_{\mathbb{S}^d} f(\mathbf{x}) \, d\sigma_d(\mathbf{x}) \right| = \mathcal{O}(N^{-s/d}), \qquad N \to \infty,$$

with an implied constant that depends on the $\mathbb{H}^s(\mathbb{S}^d)$-norm, but is independent of N. Here σ_d is the normalized surface measure on $\mathbb{S}^d$.

We provide methods for generation and numerical testing of QMC designs. An essential tool is an expression for the worst-case error in terms of a reproducing kernel for the space $\mathbb{H}^s(\mathbb{S}^d)$ with $s > d/2$. As a consequence of this and a recent result of Bondarenko et al. on the existence of spherical designs with appropriate number of points, we show that minimizers of the N-point energy for this kernel form a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$. Furthermore, without appealing to the Bondarenko et al. result, we prove that point sets that maximize the sum of suitable powers of the Euclidean distance between pairs of points form a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ with s in the interval $(d/2, d/2+1)$. For such spaces there exist reproducing kernels with simple closed forms that are useful for numerical testing of optimal order Quasi Monte Carlo integration.

Numerical experiments suggest that many familiar sequences of point sets on the sphere (equal area points, spiral points, minimal [Coulomb or logarithmic] energy points, and Fekete points) are QMC designs for appropriate values of s. For comparison purposes we show that configurations of random points that are independently and uniformly distributed on the sphere do not constitute QMC designs for any $s > d/2$.

If (X_N) is a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$, we prove that it is also a sequence of QMC designs for $\mathbb{H}^{s'}(\mathbb{S}^d)$ for all $s' \in (d/2, s)$. This leads to the question of determining the supremum of such s (here called the QMC strength of the sequence), for which we provide estimates based on computations for the aforementioned sequences.

1. Introduction

In this paper we introduce a new notion for sequences of finite point sets on the unit sphere $\mathbb{S}^d$ in the Euclidean space $\mathbb{R}^{d+1}$, $d \geq 2$, namely that of *sequences of QMC designs*. These are sequences that emulate spherical designs in that they provide optimal order equal weight numerical integration (or Quasi Monte Carlo) rules for certain Sobolev spaces of functions over the unit sphere $\mathbb{S}^d$.

A spherical t-design, a concept introduced in the groundbreaking paper [20] by Delsarte, Goethals and Seidel, is a finite subset $X_N \subset \mathbb{S}^d$ with the characterizing property that an equal weight integration rule with node set X_N integrates exactly all polynomials P with degree $\leq t$; that is,

$$(1) \qquad \frac{1}{N} \sum_{\mathbf{x} \in X_N} P(\mathbf{x}) = \int_{\mathbb{S}^d} P(\mathbf{x}) \, \mathrm{d}\,\sigma_d(\mathbf{x}), \qquad \deg P \leq t.$$

Here $N = |X_N|$ is the cardinality of X_N, or the number of points of the spherical design, while the integral is with respect to the normalized surface measure σ_d on $\mathbb{S}^d$, and the polynomials of degree $\leq t$ are the restrictions to $\mathbb{S}^d$ of the polynomials of degree $\leq t$ on $\mathbb{R}^{d+1}$.

Sequences of spherical designs have a known fast-convergence property in Sobolev spaces. (See Section 2.3 below for the definition of the Sobolev space $\mathbb{H}^s(\mathbb{S}^d)$.) This property, stated in the following theorem, was first proved for the particular case $s = 3/2$ and $d = 2$ in [25], then extended to all $s > 1$ for $d = 2$ in [26], and finally extended to all $s > d/2$ and all $d \geq 2$ in [16]. For the extension to Besov spaces, see [23]. (The results in those papers were proved for all positive-weight integration rules with an appropriate degree of polynomial accuracy in relation to the number of points, but here we restrict our attention to equal weight rules.)

Theorem 1. *Given $s > d/2$, there exists $C(s,d) > 0$ depending on the $\mathbb{H}^s(\mathbb{S}^d)$-norm such that for every N-point spherical t-design X_N on $\mathbb{S}^d$ it holds that*

$$(2) \qquad \sup_{\substack{f \in \mathbb{H}^s(\mathbb{S}^d), \\ \|f\|_{\mathbb{H}^s} \leq 1}} \left| \frac{1}{N} \sum_{\mathbf{x} \in X_N} f(\mathbf{x}) - \int_{\mathbb{S}^d} f(\mathbf{x}) \, \mathrm{d}\,\sigma_d(\mathbf{x}) \right| \leq \frac{C(s,d)}{t^s}.$$

Note that the constant $C(s,d)$ in this theorem does not depend on t or on N, or on the particular spherical design X_N. Note too that the condition $s > d/2$ is a natural one, since by the Sobolev embedding theorem this is the condition needed for $\mathbb{H}^s(\mathbb{S}^d)$ to be continuously embedded in the space of continuous functions $C(\mathbb{S}^d)$.

The relation between N and t in a spherical design is not fixed, but there are known lower bounds on N (see (7) below) that tell us that N is at least of order t^d, and a recent result [11] of Bondarenko et al. (see Theorem 8 below) asserts that given t there always exists a spherical design with $N \asymp t^d$. Here we write $a_n \asymp b_n$ to mean that there exist positive constants c_1 and c_2 independent of n such that $c_1 a_n \leq b_n \leq c_2 a_n$ for all n.

Motivated by these facts and the belief that the only interesting sequences of spherical designs are those with $N \asymp t^d$, we now define the notion of a sequence of QMC designs.

Definition 2. *Given $s > d/2$, a sequence (X_N) of N-point configurations on $\mathbb{S}^d$ with $N \to \infty$ is said to be a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ if there exists*

$c(s,d) > 0$, independent of N, such that

$$(3) \qquad \sup_{\substack{f \in \mathbb{H}^s(\mathbb{S}^d), \\ \|f\|_{\mathbb{H}^s} \leq 1}} \left| \frac{1}{N} \sum_{\mathbf{x} \in X_N} f(\mathbf{x}) - \int_{\mathbb{S}^d} f(\mathbf{x}) \, d\sigma_d(\mathbf{x}) \right| \leq \frac{c(s,d)}{N^{s/d}}.$$

In this definition X_N need not be defined for all natural numbers N: it is sufficient that X_N exists for an infinite subset of the natural numbers. By a special case of theorems in [24] and [22], the exponent of N in (3) cannot be larger than s/d:

Theorem 3. *Given $s > d/2$, there exists $c'(s,d) > 0$ depending on the $\mathbb{H}^s(\mathbb{S}^d)$-norm such that for any N-point configuration on $\mathbb{S}^d$, we have*

$$(4) \qquad \frac{c'(s,d)}{N^{s/d}} \leq \sup_{\substack{f \in \mathbb{H}^s(\mathbb{S}^d), \\ \|f\|_{\mathbb{H}^s} \leq 1}} \left| \frac{1}{N} \sum_{\mathbf{x} \in X_N} f(\mathbf{x}) - \int_{\mathbb{S}^d} f(\mathbf{x}) \, d\sigma_d(\mathbf{x}) \right|.$$

The following theorem, established in Section 9 by appealing to results of Brandolini et al. [13], asserts that if (X_N) is a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$, then it is also true for all coarser Sobolev spaces $\mathbb{H}^{s'}(\mathbb{S}^d)$ with $d/2 < s' < s$.

Theorem 4. *Given $s > d/2$, let (X_N) be a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$. Then (X_N) is a sequence of QMC designs for $\mathbb{H}^{s'}(\mathbb{S}^d)$, for all s' with $d/2 < s' \leq s$.*

It follows from this theorem that for every sequence of QMC designs (X_N) there is some number s^* such that (X_N) is a sequence of QMC designs for all s satisfying $d/2 < s < s^*$, and is not a QMC design for $s > s^*$; that is,

$$(5) \qquad s^* := s^*[(X_N)] := \sup \left\{ s : (X_N) \text{ is a sequence of QMC designs for } \mathbb{H}^s(\mathbb{S}^d) \right\}.$$

We call s^* the *QMC strength* of (X_N). If $s^* = +\infty$, we say the sequence of QMC designs is "generic".

Definition 5. A sequence of N-point configurations (X_N) on $\mathbb{S}^d$ is said to be a *sequence of generic QMC designs* if (3) holds for all $s > d/2$.

As in Definition 2, X_N need not be defined for all natural numbers N. Obviously, every sequence of spherical t-designs with $N \asymp t^d$ as $t \to \infty$ is a sequence of generic QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$, for **all** $s > d/2$. We noted already the existence of a sequence of spherical t-designs with $N \asymp t^d$. A simple application of that result yields the following.

Theorem 6. *For $N = 1, 2, 3, \ldots$ there exist N-point spherical t-designs Y_N on $\mathbb{S}^d$ with $t \asymp N^{1/d}$. These form a sequence of generic QMC designs.*

For **fixed** $s > d/2$, there exist many sequences of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ that are not composed of spherical designs. Indeed, if K is a reproducing kernel for the Sobolev space $\mathbb{H}^s(\mathbb{S}^d)$, $s > d/2$, we prove in Section 3 (by appealing to Theorem 6) that N-point configurations ($N \geq 2$) minimizing the K-energy functional

$$\sum_{j=1}^{N} \sum_{i=1}^{N} K(\mathbf{x}_j, \mathbf{x}_i)$$

form a sequence of QMC designs for this $\mathbb{H}^s(\mathbb{S}^d)$; cf. Theorem 12. For s in the interval $(d/2, d/2 + 1)$ we show in Section 5 that, for C a suitably large constant,

$C - |\mathbf{x} - \mathbf{y}|^{2s-d}$ is a reproducing kernel for $\mathbb{H}^s(\mathbb{S}^d)$, and therefore the maximizers of the generalized sum of distances

$$(6) \qquad \sum_{j=1}^{N} \sum_{i=1}^{N} |\mathbf{x}_j - \mathbf{x}_i|^{2s-d},$$

$N = 2, 3, 4, \ldots$, form a sequence of QMC designs for this $\mathbb{H}^s(\mathbb{S}^d)$.

Numerical evidence presented later in this paper suggests that many familiar sequences of point sets on $\mathbb{S}^2$ (such as minimal [Coulomb or logarithmic] energy points, generalized spiral points, equal area points, and Fekete points) form sequences of QMC designs for $\mathbb{H}^s(\mathbb{S}^2)$ for a range of values s that depends on the particular sequence. Some estimates of the QMC design strength s^* of these sequences are given in Section 8.

The fact that the QMC design property (3) is not satisfied by all sequences of point sets follows from a probabilistic argument.

Theorem 7. *Given $s > d/2$, the expected value of the squared worst-case error satisfies*

$$\sqrt{\mathbb{E}\left[\sup_{\substack{f \in \mathbb{H}^s(\mathbb{S}^d), \\ \|f\|_{\mathbb{H}^s} \leq 1}} \left| \frac{1}{N} \sum_{j=1}^{N} f(\mathbf{x}_j) - \int_{\mathbb{S}^d} f(\mathbf{x}) \, \mathrm{d}\,\sigma_d(\mathbf{x}) \right|^2 \right]} = \frac{b(s,d)}{N^{1/2}}$$

for some explicit constant $b(s,d) > 0$, where the points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are independently and uniformly distributed on $\mathbb{S}^d$.

Theorem 7 tells us that randomly chosen point sets give a slower rate of convergence than $N^{-s/d}$ for all $s > d/2$, and hence do not form QMC designs. (See Section 7 for a more complete discussion.) However, if we compartmentalize the random point selection process with respect to a partition of the sphere into N equal area regions with small diameter, then we do get an average worst-case error rate appropriate to QMC designs for $s \in (d/2, d/2 + 1)$; see Theorem 21. On the other hand, such randomized equal area point configurations will, on average, not form a sequence of QMC designs for $s > d/2 + 1$; see Theorem 22.

We shall also discuss "low-discrepancy sequences" on the sphere and estimates for their worst-case error when used for QMC rules. It turns out that the point sets of such a sequence almost satisfy the QMC design property for $s \in (d/2, (d+1)/2)$, except for a power of $\log N$.

To summarize, the key features and results in this paper are the following: we introduce (in Definition 2) the concept of *QMC design sequences for* $\mathbb{H}^s(\mathbb{S}^d)$; we show (in Theorem 4) that the QMC design property automatically holds for all smaller $s' \in (d/2, s)$ (that is, for coarser Sobolev spaces), leading to the definition of the *QMC strength s^** of a sequence of point sets (X_N); we characterize (in Theorem 12) QMC designs as minimizers of a quadratic functional of a reproducing kernel, which turns into a maximization problem for the sum of all pairwise distances if the reproducing kernel is the generalized distance kernel (Theorem 14); observe (in Section 6) that low-discrepancy sequences on the sphere almost (up to a power of $\log N$) satisfy the QMC design property for $s \in (d/2, d/2 + 1)$; and on the other hand, we show (in Theorem 7) that randomly selected point sets in general do not yield QMC design sequences for any $s > d/2$; however, randomized equal-area points do so on average for $s \in (d/2, d/2 + 1)$ (Theorem 21).

The structure of the paper is as follows: The next section provides background for spherical designs and for Sobolev spaces $\mathbb{H}^s(\mathbb{S}^d)$ and their associated reproducing kernels. Section 3 summarizes the reproducing kernel Hilbert space approach to worst-case error, and gives the application to QMC design sequences. Examples of Sobolev spaces and associated kernels are given in Sections 4 and 5, with particular emphasis on configurations maximizing sums of generalized distances. Section 6 concerns low-discrepancy sequences on the sphere and their quadrature properties. In Section 7 we analyze the quadrature error for randomly chosen points on the sphere and, in Section 8, we provide numerical results for worst-case errors and quadrature errors for certain familiar sequences of configurations. Most of the formal proofs are given in Section 9.

2. Background

2.1. **Spherical designs.** In the literature on spherical designs, and again in this paper, the relation between N and t in (1) plays an important role. It is known (Seymour and Zaslavsky [35]) that a spherical t-design always exists if N is sufficiently large, but that result says nothing about the size of N. In the important paper [20], lower bounds of exact order t^d were established; namely,

$$
(7) \qquad N \geq \begin{cases} \dbinom{d+t/2}{d} + \dbinom{d+t/2-1}{d} & \text{for } t \text{ even,} \\ 2\dbinom{d+\lfloor t/2 \rfloor}{d} & \text{for } t \text{ odd,} \end{cases}
$$

but it is known (see Bannai and Damerell [5, 6]) that these lower bounds can be achieved only for a few small values of t. Korevaar and Meyers [28] conjectured that there always exist spherical t-designs with $N \asymp t^d$ points. Bondarenko, Radchenko and Viazovska [11] have resolved this long-standing open problem, by establishing the following result.

Theorem 8. *For $d \geq 2$, there exists a constant c_d depending only on d such that for every $N \geq c_d\, t^d$ there exists a spherical t-design on $\mathbb{S}^d$ with N points.*

In fact, a more precise result seems possible, since in [18] interval analysis is used to establish rigorously the existence of spherical designs on $\mathbb{S}^2$ with $N = (t+1)^2$ for all values of $t \leq 100$.

2.2. **Spherical harmonics.** Recall that σ_d denotes the normalized (Hausdorff) surface measure on the unit sphere $\mathbb{S}^d$ in $\mathbb{R}^{d+1}$. The [non-normalized] surface area of $\mathbb{S}^d$ is denoted by ω_d. For future reference, we record the following facts:

$$
(8) \qquad \gamma_d := \frac{1}{d}\frac{\omega_{d-1}}{\omega_d}, \qquad \frac{\omega_{d-1}}{\omega_d} = \frac{\Gamma((d+1)/2)}{\sqrt{\pi}\,\Gamma(d/2)} \sim \frac{d^{1/2}}{\sqrt{2\pi}} \quad \text{as } d \to \infty.
$$

Here, $\Gamma(z)$ is the gamma function and $f(x) \sim g(x)$ as $x \to c$ means $f(x)/g(x) \to 1$ as $x \to c$. The asymptotic relation in (8) follows from (see [1, Eq. 5.11.12])

$$
\frac{\Gamma(z+a)}{\Gamma(z+b)} \sim z^{a-b} \qquad \text{as } z \to \infty \text{ in the sector } |\arg z| \leq \pi - \delta < \pi.
$$

We make use of the rising factorial, that is, the *Pochhammer symbol* defined by

$$
(9) \qquad (a)_0 = 1, \qquad (a)_{n+1} = (a)_n(n+a), \quad n = 0, 1, \ldots,
$$

which can be written in terms of the gamma function as $(a)_n = \Gamma(n+a)/\Gamma(a)$.

We denote, as usual, by $\{Y_{\ell,k} : k = 1, \ldots, Z(d, \ell)\}$ a collection of $\mathbb{L}_2$-orthonormal real spherical harmonics (homogeneous harmonic polynomials in $d + 1$ variables restricted to $\mathbb{S}^d$) of exact degree ℓ, where

$$(10) \qquad Z(d, \ell) = (2\ell + d - 1) \frac{\Gamma(\ell + d - 1)}{\Gamma(d) \, \Gamma(\ell + 1)} \sim \frac{2}{\Gamma(d)} \, \ell^{d-1} \quad \text{as } \ell \to \infty.$$

It is well known that the $Y_{\ell,k}$ satisfy the following identity known as the addition theorem:

$$(11) \qquad \sum_{k=1}^{Z(d,\ell)} Y_{\ell,k}(\mathbf{x}) Y_{\ell,k}(\mathbf{y}) = Z(d, \ell) \, P_\ell^{(d)}(\mathbf{x} \cdot \mathbf{y}), \qquad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d,$$

where $P_\ell^{(d)}$ is the normalized Gegenbauer (or Legendre) polynomial, orthogonal on the interval $[-1, 1]$ with respect to the weight function $(1-t^2)^{d/2-1}$, and normalized by $P_\ell^{(d)}(1) = 1$. Each spherical harmonic $Y_{\ell,k}$ of exact degree ℓ is an eigenfunction of the negative Laplace-Beltrami operator $-\Delta_d^*$ for $\mathbb{S}^d$ with eigenvalue

$$(12) \qquad\qquad\qquad\qquad \lambda_\ell := \ell \, (\ell + d - 1) \,.$$

(For further details see, e.g., [31].)

The family $\{Y_{\ell,k} : k = 1, \ldots, Z(d, \ell); \ell = 0, 1, \ldots\}$ forms a complete orthonormal (with respect to σ_d) system for the Hilbert space $\mathbb{L}_2(\mathbb{S}^d)$ of square-integrable functions on $\mathbb{S}^d$ endowed with the usual inner product and induced norm

$$(f, g)_{\mathbb{L}_2(\mathbb{S}^d)} := \int_{\mathbb{S}^d} f(\mathbf{x}) g(\mathbf{x}) \, \mathrm{d}\, \sigma_d(\mathbf{x}), \qquad \|f\|_{\mathbb{L}_2(\mathbb{S}^d)} := \sqrt{(f, f)_{\mathbb{L}_2(\mathbb{S}^d)}}.$$

We shall denote by $\mathbb{P}_t(\mathbb{S}^d)$ the space of all spherical polynomials of degree $\leq t$ (i.e., the restriction to $\mathbb{S}^d$ of all polynomials in $d + 1$ real variables of degree $\leq t$). The space $\mathbb{P}_t(\mathbb{S}^d)$ coincides with the span of all spherical harmonics up to (and including) degree t, and its dimension is $Z(d + 1, t)$.

We make frequent use of the following simple application of the addition theorem.

Lemma 9. *Let $d \geq 2$. For all integers $\ell \geq 0$ and all choices of points $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbb{S}^d$ it holds that*

$$\Phi_\ell(\mathbf{x}_1, \ldots, \mathbf{x}_N) := \frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} Z(d, \ell) \, P_\ell^{(d)}(\mathbf{x}_j \cdot \mathbf{x}_i) = \sum_{k=1}^{Z(d,\ell)} \left| \frac{1}{N} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) \right|^2 \geq 0.$$

2.3. **Sobolev spaces.** The Sobolev space $\mathbb{H}^s(\mathbb{S}^d)$ may be defined for $s \geq 0$ as the set of all functions $f \in \mathbb{L}_2(\mathbb{S}^d)$ whose Laplace-Fourier coefficients

$$\widehat{f}_{\ell,k} := (f, Y_{\ell,k})_{\mathbb{L}_2(\mathbb{S}^d)} = \int_{\mathbb{S}^d} f(\mathbf{x}) Y_{\ell,k}(\mathbf{x}) \, \mathrm{d}\, \sigma_d(\mathbf{x})$$

satisfy

$$(13) \qquad\qquad\qquad \sum_{\ell=0}^{\infty} \sum_{k=1}^{Z(d,\ell)} (1 + \lambda_\ell)^s \left| \widehat{f}_{\ell,k} \right|^2 < \infty,$$

where the λ_ℓ's are given in (12). On setting $s = 0$, we recover $\mathbb{H}^0(\mathbb{S}^d) = \mathbb{L}_2(\mathbb{S}^d)$.

The norm in $\mathbb{H}^s(\mathbb{S}^d)$ may of course be defined as the square root of the expression on the left-hand side of the last inequality. In this paper we shall, however, take advantage of the freedom to define equivalent Sobolev space norms. Let $s > d/2$

be fixed and suppose we are given a sequence of positive real numbers $(a_\ell^{(s)})_{\ell \geq 0}$ satisfying

$$(14) \qquad a_\ell^{(s)} \asymp (1 + \lambda_\ell)^{-s} \asymp (1 + \ell)^{-2s}.$$

Then we can define a norm in $\mathbb{H}^s(\mathbb{S}^d)$ by

$$(15) \qquad \|f\|_{\mathbb{H}^s} := \left(\sum_{\ell=0}^{\infty} \sum_{k=1}^{Z(d,\ell)} \frac{1}{a_\ell^{(s)}} \left| \widehat{f}_{\ell,k} \right|^2 \right)^{1/2}.$$

The norm therefore depends on the particular choice of the sequence $(a_\ell^{(s)})_{\ell \geq 0}$, but for notational simplicity we shall generally not show this dependence explicitly. Clearly, Definitions 2 and 5 are not tied to a particular Sobolev norm $\| \cdot \|_{\mathbb{H}^s}$, since a change to an *equivalent* Sobolev norm merely leads to a change of the constant $c(s, d)$. The corresponding inner product in the Sobolev space is

$$(16) \qquad (f, g)_{\mathbb{H}^s} := \sum_{\ell=0}^{\infty} \sum_{k=1}^{Z(d,\ell)} \frac{1}{a_\ell^{(s)}} \widehat{f}_{\ell,k} \, \widehat{g}_{\ell,k}.$$

It is well known that $\mathbb{H}^s(\mathbb{S}^d) \subset \mathbb{H}^{s'}(\mathbb{S}^d)$ whenever $s > s'$, and that $\mathbb{H}^s(\mathbb{S}^d)$ is continuously embedded in the space of k-times continuously differentiable functions $C^k(\mathbb{S}^d)$ if $s > k + d/2$ (e.g. [22]).

2.4. **Sobolev spaces as reproducing kernel Hilbert spaces.** Since the point-evaluation functional is bounded in $\mathbb{H}^s(\mathbb{S}^d)$ whenever $s > d/2$, the Riesz representation theorem assures the existence of a reproducing kernel $K^{(s)}(\mathbf{x}, \mathbf{y})$, which can be written as

$$(17) \qquad K^{(s)}(\mathbf{x}, \mathbf{y}) = \sum_{\ell=0}^{\infty} a_\ell^{(s)} Z(d,\ell) P_\ell^{(d)}(\mathbf{x} \cdot \mathbf{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{Z(d,\ell)} a_\ell^{(s)} Y_{\ell,k}(\mathbf{x}) Y_{\ell,k}(\mathbf{y}),$$

where the positive coefficients $a_\ell^{(s)}$ satisfy (14). It is easily verified that the above expression has the reproducing kernel properties:

$$K^{(s)}(\cdot, \mathbf{x}) \in \mathbb{H}^s(\mathbb{S}^d), \;\; \mathbf{x} \in \mathbb{S}^d, \quad (f, K^{(s)}(\cdot, \mathbf{x}))_{\mathbb{H}^s} = f(\mathbf{x}), \;\; \mathbf{x} \in \mathbb{S}^d, f \in \mathbb{H}^s(\mathbb{S}^d).$$

The kernel is a *zonal* function; that is, $K^{(s)}(\mathbf{x}, \mathbf{y})$ depends only on the inner product $\mathbf{x} \cdot \mathbf{y}$. We write for simplicity $K^{(s)}(\mathbf{x} \cdot \mathbf{y}) := K^{(s)}(\mathbf{x}, \mathbf{y})$. For the particular choice $a_\ell^{(s)} = (1 + \lambda_\ell)^{-s}$, we use the notation

$$(18) \qquad K_{\mathrm{can}}^{(s)}(\mathbf{x}, \mathbf{y}) := \sum_{\ell=0}^{\infty} (1 + \lambda_\ell)^{-s} Z(d,\ell) P_\ell^{(d)}(\mathbf{x} \cdot \mathbf{y}),$$

which we call the *canonical kernel* for $\mathbb{H}^s(\mathbb{S}^d)$.

Sections 4 and 5 contain explicit examples of Sobolev spaces and associated kernels.

3. NUMERICAL INTEGRATION AND WORST-CASE ERROR

In this section we first summarize some well-known essential facts pertaining to the reproducing kernel Hilbert space approach to the worst-case error analysis of numerical integration by QMC rules on the sphere. Theorems 12 and 13 then give the consequences for QMC design sequences.

3.1. **Worst-case error.** Our results are based on an explicit expression for the "worst-case error" that occurs on the left-hand side of (3):

Definition 10. For a Banach space B of continuous functions on $\mathbb{S}^d$ with norm $\|\cdot\|_B$, the *worst-case error* for the integration rule $Q[X_N]$ with node set $X_N = \{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$ approximating the integral $I(f)$, with $Q[X_N](f)$ and $I(f)$ defined by

$$Q[X_N](f):=\frac{1}{N}\sum_{j=1}^{N}f(\mathbf{x}_j), \qquad I(f):=\int_{\mathbb{S}^d}f(\mathbf{x})\,\mathrm{d}\,\sigma_d(\mathbf{x}),$$

is given by

$$\mathrm{wce}(Q[X_N];B):=\sup\left\{\left|Q[X_N](f)-I(f)\right|:f\in B,\|f\|_B\leq 1\right\}.$$

As a trivial consequence of the definition we have the following error bound for an arbitrary function $f \in B$:

$$(19) \qquad\qquad |Q[X_N](f)-I(f)| \leq \mathrm{wce}(Q[X_N];B)\,\|f\|_B.$$

Because of the similarity of (19) to the celebrated Koksma-Hlawka inequality, which involves the "star-discrepancy" of the node set, the worst-case error is sometimes referred to as a *generalized discrepancy*; see for example [19]. In this paper, however, we shall generally reserve the word "discrepancy" for quantities that have a geometric interpretation.

3.2. **Worst-case error in a reproducing kernel Hilbert space.** For most reproducing kernel Hilbert spaces there is a computable expression for the worst-case error, as shown by the following standard argument. With K the kernel of a reproducing kernel Hilbert space H with inner product $(\cdot,\cdot)_H$, the reproducing kernel property $f(\mathbf{x}) = (f, K(\cdot,\mathbf{x}))_H$ allows us to write

$$Q[X_N](f)-I(f) = (f, \mathcal{R}[X_N])_H, \qquad f\in H,$$

where $\mathcal{R}[X_N] \in H$ is the "representer" of the error, given by

$$\mathcal{R}[X_N](\mathbf{x}):=\frac{1}{N}\sum_{j=1}^{N}K(\mathbf{x},\mathbf{x}_j)-I_{\mathbf{y}}K(\mathbf{x},\cdot),$$

assuming that the integration functional $f \mapsto I(f)$ is bounded on H. Here $I_{\mathbf{y}}K$ means the integral functional I applied to the second variable in K (and later $I_{\mathbf{x}}K$ will mean the integral functional applied to the first variable).[*]

---

[*]The norm $\|\mathcal{R}[X_N]\|_H$ is also known as the *g-diaphony of X_N* with $g = I_{\mathbf{y}}K(\mathbf{x},\cdot)$; see [3].

It follows that

$$[\mathrm{wce}(Q[X_N]; H)]^2 := \sup\left\{\left|Q[X_N](f) - I(f)\right|^2 : f \in H, \|f\|_H \le 1\right\}$$

$$= (\mathcal{R}[X_N], \mathcal{R}[X_N])_H = \|\mathcal{R}[X_N]\|_H^2$$

$$= \frac{1}{N^2}\sum_{j=1}^N\sum_{i=1}^N K(\mathbf{x}_j, \mathbf{x}_i) - \frac{2}{N}\sum_{j=1}^N I_{\mathbf{y}}K(\mathbf{x}_j, \cdot) + I_{\mathbf{x}}I_{\mathbf{y}}K(\cdot, \cdot).$$

3.3. **Worst-case error in** $\mathbb{H}^s(\mathbb{S}^d)$. Now consider the special case of the reproducing kernel Hilbert space $\mathbb{H}^s(\mathbb{S}^d)$ with $s > d/2$, and with reproducing kernel given by (17). For this case it is easily seen that

$$I_{\mathbf{y}}K^{(s)}(\mathbf{x}, \cdot) = a_0^{(s)},$$

from which it follows that

$$\left[\mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\right]^2 = \left[\frac{1}{N^2}\sum_{j=1}^N\sum_{i=1}^N K^{(s)}(\mathbf{x}_j, \mathbf{x}_i)\right] - a_0^{(s)}$$

$$= \frac{1}{N^2}\sum_{j=1}^N\sum_{i=1}^N \mathcal{K}^{(s)}(\mathbf{x}_j \cdot \mathbf{x}_i),$$

where $\mathcal{K}^{(s)} : [-1, 1] \to \mathbb{R}$ is defined by

(20) $$\mathcal{K}^{(s)}(\mathbf{x} \cdot \mathbf{y}) := \sum_{\ell=1}^\infty a_\ell^{(s)} Z(d, \ell) P_\ell^{(d)}(\mathbf{x} \cdot \mathbf{y}).$$

The use of the calligraphic symbol here and for subsequent kernels indicates that the sum runs from $\ell = 1$ rather than $\ell = 0$. Note that the kernels depend on s through the sequence $(a_\ell^{(s)})_{\ell \ge 0}$.

We summarize these observations in the following proposition.

Proposition 11. *For $s > d/2$, let $\mathbb{H}^s(\mathbb{S}^d)$ be the Hilbert space with norm (15), where the sequence $(a_\ell^{(s)})_{\ell \ge 0}$ satisfies (14), and let $\mathcal{K}^{(s)}$ be given by (20). Then, for a rule $Q[X_N]$ with node set $X_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^d$,*

(21)
$$\mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) = \left(\frac{1}{N^2}\sum_{j=1}^N\sum_{i=1}^N \mathcal{K}^{(s)}(\mathbf{x}_j \cdot \mathbf{x}_i)\right)^{1/2}$$

$$= \left(\sum_{\ell=1}^\infty\sum_{k=1}^{Z(d,\ell)} a_\ell^{(s)}\left|\frac{1}{N}\sum_{j=1}^N Y_{\ell,k}(\mathbf{x}_j)\right|^2\right)^{1/2}.$$

From the first expression in (21), the squared worst-case error for the rule $Q[X_N]$ is the normalized $\mathcal{K}^{(s)}$-energy functional evaluated at the node set X_N. This expression can be computationally useful when the kernel $\mathcal{K}^{(s)}$ is available in closed form.

By comparison with a sequence satisfying Theorem 6, we deduce that the minimizers of (21) yield a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$.

Theorem 12. *Under the assumptions of Proposition* 11, *if* X_N^*, $N = 2, 3, 4, \dots$, *minimizes the energy functional*

$$\sum_{j=1}^{N} \sum_{i=1}^{N} \mathcal{K}^{(s)}(\mathbf{x}_j \cdot \mathbf{x}_i),$$

then there exists $c(s, d) > 0$ *depending on the* $\mathbb{H}^s(\mathbb{S}^d)$-*norm such that for all* $N \geq 2$,

$$\mathrm{wce}(Q[X_N^*]; \mathbb{H}^s(\mathbb{S}^d)) \leq \frac{c(s, d)}{N^{s/d}}.$$

Consequently, (X_N^*) *is a sequence of QMC designs for* $\mathbb{H}^s(\mathbb{S}^d)$.

The next result says, in essence, that the computed worst-case errors in $\mathbb{H}^s(\mathbb{S}^d)$ of a given sequence of QMC designs for $\mathbb{H}^{s'}(\mathbb{S}^d)$, where $d/2 < s' \leq s$, will always show a rate of decay at least $O(N^{-s'/d})$ and at most $O(N^{-s/d})$.

Theorem 13. *Under the assumptions of Proposition* 11 *for* $\mathbb{H}^s(\mathbb{S}^d)$ *and* $\mathbb{H}^{s'}(\mathbb{S}^d)$, *if* (X_N) *is a sequence of QMC designs for* $\mathbb{H}^{s'}(\mathbb{S}^d)$, *then*

$$(22) \qquad \mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) \leq \frac{c(s, s', d)}{N^{s'/d}}, \qquad d/2 < s' < s,$$

where $c(s, s', d) > 0$ *depends on the norms for* $\mathbb{H}^s(\mathbb{S}^d)$ *and* $\mathbb{H}^{s'}(\mathbb{S}^d)$, *but is independent of* N.

This result follows from the last expression in (21), since it implies that, with respect to the canonical kernels (18) for $\mathbb{H}^s(\mathbb{S}^d)$ and $\mathbb{H}^{s'}(\mathbb{S}^d)$, there holds for any N-point configuration X_N,

$$(23) \qquad \mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) < \mathrm{wce}(Q[X_N]; \mathbb{H}^{s'}(\mathbb{S}^d)), \qquad d/2 < s' < s.$$

We shall exploit Theorem 13 in Section 8 to determine empirical values of the supremum s^* in (5) for a number of putative sequences of QMC designs.

4. CUI AND FREEDEN KERNEL

For $\mathbb{S}^2$, Cui and Freeden [19] studied the kernel

$$(24) \quad K_{\mathrm{CF}}(\mathbf{x}, \mathbf{y}) := 1 + \sum_{\ell=1}^{\infty} \frac{1}{\ell(\ell + 1)} P_\ell(\mathbf{x} \cdot \mathbf{y}) = 2 - 2 \log\left(1 + \sqrt{\frac{1 - \mathbf{x} \cdot \mathbf{y}}{2}}\right), \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^2.$$

It was observed in [36] that this is a reproducing kernel for $\mathbb{H}^{3/2}(\mathbb{S}^2)$ as can be seen from the above Laplace-Fourier series expansion in terms of Legendre polynomials; cf. Section 2. Since the constant term in the series expansion (24) is 1, Proposition 11 asserts that the corresponding worst-case error for $\mathbb{H}^{3/2}(\mathbb{S}^2)$ equipped with this kernel is

$$\mathrm{wce}(Q[X_N]; \mathbb{H}^{3/2}(\mathbb{S}^2)) = \left(1 - \frac{2}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} \log\left(1 + \sqrt{\frac{1 - \mathbf{x}_j \cdot \mathbf{x}_i}{2}}\right)\right)^{1/2}.$$

The right-hand side, up to a constant factor, is known as the *Cui and Freeden (CF) discrepancy* of X_N. Note that by Theorem 12, *sequences of N-point configurations that minimize the CF discrepancy are sequences of QMC designs for* $\mathbb{H}^{3/2}(\mathbb{S}^2)$. This fact can also be seen without appealing to results for spherical designs by applying the independently derived Theorem 14 in the next section and by

using the equivalence of norms.[†] In [19] this discrepancy has been used to test for uniformity of a sequence of N-point configurations. Furthermore, the CF discrepancy was used in [36] for analyzing quadrature properties of so-called extremal or Fekete points (these are sets of $(t+1)^2$ points on $\mathbb{S}^2$ that maximize the determinant of the interpolation matrix for polynomials of degree t). Numerical data in [36] suggested that for $\mathbb{H}^{3/2}(\mathbb{S}^2)$, the CF discrepancy for spherical t-designs obtained using Fekete points as starting points, decays like $\mathcal{O}(t^{-3/2})$; this in turn led to the discovery of Theorem 1.

5. Generalized distance kernel

In the following we make use of the identity

$$|\mathbf{x} - \mathbf{y}|^2 = 2 - 2\mathbf{x} \cdot \mathbf{y}, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d.$$

Reproducing kernels for $\mathbb{H}^s(\mathbb{S}^d)$ for $s > d/2$ can be constructed utilizing powers of distances, provided the power $2s - d$ is not an even integer. Indeed, it is known (cf., e.g., [27]) that the signed power of the distance, with sign $(-1)^{L+1}$ with $L := L(s) := \lfloor s - d/2 \rfloor$, has the following Laplace-Fourier expansion: for $\mathbf{x}, \mathbf{y} \in \mathbb{S}^d$,

$$(25) \qquad (-1)^{L+1} |\mathbf{x} - \mathbf{y}|^{2s-d} = (-1)^{L+1} V_{d-2s}(\mathbb{S}^d) + \sum_{\ell=1}^{\infty} \alpha_\ell^{(s)} \, Z(d, \ell) \, P_\ell^{(d)}(\mathbf{x} \cdot \mathbf{y}),$$

where

$$(26) \qquad \begin{aligned} V_{d-2s}(\mathbb{S}^d) &:= \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} |\mathbf{x} - \mathbf{y}|^{2s-d} \, \mathrm{d}\,\sigma_d(\mathbf{x}) \, \mathrm{d}\,\sigma_d(\mathbf{y}) = 2^{2s-1} \frac{\Gamma((d+1)/2)\,\Gamma(s)}{\sqrt{\pi}\,\Gamma(d/2+s)}, \\ \alpha_\ell^{(s)} &:= V_{d-2s}(\mathbb{S}^d) \frac{(-1)^{L+1}(d/2-s)_\ell}{(d/2+s)_\ell}, \quad \ell \geq 1. \end{aligned}$$

From these formulas one can verify that all the coefficients $\alpha_\ell^{(s)}$ are positive for $\ell \geq L + 1$ and alternate in sign for $\ell \leq L$. Furthermore, the $\alpha_\ell^{(s)}$'s decay with the rate required for coefficients in the Laplace-Fourier expansion of a reproducing kernel for $\mathbb{H}^s(\mathbb{S}^d)$ as can be seen from the asymptotic expansion

$$\alpha_\ell^{(s)} \sim 2^{2s-1} \frac{\Gamma((d+1)/2)\,\Gamma(s)}{\sqrt{\pi}[(-1)^{L+1}\,\Gamma(d/2-s)]} \, \ell^{-2s} \qquad \text{as } \ell \to \infty.$$

Thus, by modifying if necessary some of the early coefficients, one can derive a reproducing kernel for $\mathbb{H}^s(\mathbb{S}^d)$ for $s > d/2$ and $2s - d$ not an even integer[‡] (cf. Section 2.4).

Case I. For $d/2 < s < d/2 + 1$ (in which case, $L(s) = 0$), only the constant term in (25) is negative and thus by adding any constant larger than $V_{d-2s}(\mathbb{S}^d)$, say $2V_{d-2s}(\mathbb{S}^d)$, we obtain the following reproducing kernel for $\mathbb{H}^s(\mathbb{S}^d)$ which we call the "generalized distance" kernel:

$$(27) \qquad K_{\mathrm{gd}}^{(s)}(\mathbf{x}, \mathbf{y}) := 2V_{d-2s}(\mathbb{S}^d) - |\mathbf{x} - \mathbf{y}|^{2s-d}, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d.$$

In particular, for $s = (d+1)/2$ we get the "distance kernel" for $\mathbb{H}^{(d+1)/2}(\mathbb{S}^d)$:

$$(28) \qquad K_{\mathrm{dist}}(\mathbf{x}, \mathbf{y}) := K_{\mathrm{gd}}^{((d+1)/2)}(\mathbf{x}, \mathbf{y}) = 2V_{-1}(\mathbb{S}^d) - |\mathbf{x} - \mathbf{y}|, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d,$$

[†]An even more direct proof can be given by applying [32, Theorem 2.2].

[‡]In the case of $2s - d$ is an even integer, the expansion (25) terminates after finitely many terms and so the $\alpha_\ell^{(s)}$'s do not satisfy the appropriate asymptotic behavior (14) for $\mathbb{H}^s(\mathbb{S}^d)$.

which for $d = 2$ is equivalent to the Cui and Freeden kernel in the sense that there exist positive constants c and C such that $c\,K_{\mathrm{dist}}(\mathbf{x}, \mathbf{y}) \leq K_{\mathrm{CF}}(\mathbf{x}, \mathbf{y}) \leq C\,K_{\mathrm{dist}}(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$.

With respect to the $K_{\mathrm{gd}}^{(s)}$ kernel, the worst-case error for $d/2 < s < d/2 + 1$ is, from Proposition 11, given by

$$(29) \qquad \mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) = \left(V_{d-2s}(\mathbb{S}^d) - \frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} |\mathbf{x}_j - \mathbf{x}_i|^{2s-d} \right)^{1/2}.$$

According to Theorem 12, for $d/2 < s < d/2 + 1$, minimizing the right-hand side above, or equivalently, maximizing the sum of generalized distances, yields QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$. This fact can also be established without appealing to any properties of spherical designs (and hence is independent of Theorem 8). Indeed, Wagner [40], extending a result of Stolarsky [38], showed that for $d/2 < s < d/2 + 1$ there exists a sequence of N-point configurations $\{\mathbf{x}_{1,N}^*, \dots, \mathbf{x}_{N,N}^*\}$ and a positive constant $\eta_{s,d}$ such that

$$(30) \qquad V_{d-2s}(\mathbb{S}^d) - \frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} |\mathbf{x}_{j,N}^* - \mathbf{x}_{i,N}^*|^{2s-d} \leq \eta_{s,d}\, N^{-2s/d}, \qquad N \geq 2.$$

(This fact also follows immediately from Theorem 21 below dealing with randomized equal area points on $\mathbb{S}^d$.) Consequently, we have provided an independent proof of the following result.

Theorem 14. *Given $s \in (d/2, d/2 + 1)$, a sequence of N-point sets X_N^* that maximize the generalized sum of Euclidean distances (6) is a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$.*

We remark that for $s > d/2 + 1$, N-point configurations on $\mathbb{S}^d$ with maximum generalized sum of distances (without further restrictions) will have a limit distribution that is concentrated in two opposite points on $\mathbb{S}^d$ (Björck [10, Remark 1 following Theorem 7]) and, clearly, do not lead to QMC designs.

Case II. For $s > d/2 + 1$ and L as defined as above (so that $d/2 + L < s < d/2 + L + 1$), the representation (25) gives rise to a reproducing kernel for $\mathbb{H}^s(\mathbb{S}^d)$ of the form

$$(31) \quad K_{\mathrm{gd}}^{(s)}(\mathbf{x}, \mathbf{y}) := \left(1 - (-1)^{L+1} \right) V_{d-2s}(\mathbb{S}^d) + \mathcal{Q}_L(\mathbf{x} \cdot \mathbf{y}) + (-1)^{L+1} |\mathbf{x} - \mathbf{y}|^{2s-d},$$

$\mathbf{x}, \mathbf{y} \in \mathbb{S}^d$, where $\mathcal{Q}_L$ is a polynomial of degree $L \geq 1$,

$$(32) \qquad \mathcal{Q}_L(\mathbf{x} \cdot \mathbf{y}) := \sum_{\ell=1}^{L} \left((-1)^{L+1-\ell} - 1 \right) \alpha_\ell^{(s)}\, Z(d, \ell)\, P_\ell^{(d)}(\mathbf{x} \cdot \mathbf{y}), \qquad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d,$$

that simply changes the signs of the negative coefficients $\alpha_\ell^{(s)}$, $\ell \geq 1$, in (25). As a consequence of Theorem 12, we obtain the following.

Theorem 15. *Given $s \in (d/2 + L, d/2 + L + 1)$, where L is a positive integer, the sequences of N-point sets X_N^* that minimize the worst-case error*

$$(33) \quad \mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) = \left(\frac{1}{N^2} \sum_{j=1}^N \sum_{i=1}^N \left[\mathcal{Q}_L(\mathbf{x}_j \cdot \mathbf{x}_i) + (-1)^{L+1} |\mathbf{x}_j - \mathbf{x}_i|^{2s-d} \right] \right.$$

$$\left. - (-1)^{L+1} V_{d-2s}(\mathbb{S}^d) \right)^{1/2},$$

form a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$.

For fixed $L \geq 1$, one can avoid the introduction of the $\mathcal{Q}_L$-energy term in the worst-case error formula above by restricting attention to node sets $X_{N,L} = \{\mathbf{x}_{1,L}, \ldots, \mathbf{x}_{N,L}\}$ that are spherical L-designs; i.e., satisfy

$$(34) \qquad \sum_{j=1}^N Y_{\ell,k}(\mathbf{x}_{j,L}) = 0, \qquad 1 \leq \ell \leq L, \, 1 \leq k \leq Z(d, \ell).$$

For such sequences the worst-case error formula reduces to

$$\left[\mathrm{wce}(Q[X_{N,L}]; \mathbb{H}^s(\mathbb{S}^d)) \right]^2$$

$$= \frac{1}{N^2} \sum_{j=1}^N \sum_{i=1}^N (-1)^{L+1} |\mathbf{x}_{j,L} - \mathbf{x}_{i,L}|^{2s-d} - (-1)^{L+1} V_{d-2s}(\mathbb{S}^d).$$

Thus, spherical L-design configurations that minimize

$$\sum_{j=1}^N \sum_{i=1}^N (-1)^{L+1} |\mathbf{x}_{j,L} - \mathbf{x}_{i,L}|^{2s-d}, \qquad N \geq c_d L^d,$$

yield sequences of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ whenever $2s - d$ is not an even integer.

Note that for even $2s - d = 2L = 2, 4, \ldots$, the expansion (25) terminates and is a polynomial of degree L in $\mathbf{x} \cdot \mathbf{y}$. In such a case the radial (signed) generalized distance $(-1)^{L+1} |\mathbf{x} - \mathbf{y}|^{2L} \log |\mathbf{x} - \mathbf{y}|$ can be used to define a reproducing kernel for $\mathbb{H}^{2L}(\mathbb{S}^d)$. This approach is explored in [17].

6. Uniform distribution and low-discrepancy sequences on the sphere

Uniform distribution. An infinite sequence (X_N) of N-point configurations on $\mathbb{S}^d$ is *asymptotically uniformly distributed on $\mathbb{S}^d$* if for every $f \in C(\mathbb{S}^d)$ it holds that

$$(35) \qquad Q[X_N](f) \to \int_{\mathbb{S}^d} f(\mathbf{x}) \, d\sigma_d(\mathbf{x}) \quad \text{as } N \to \infty.$$

For a sequence (X_N) of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$, $s > d/2$, it follows from (19) and (14) that (35) is satisfied for any polynomial on $\mathbb{S}^d$ and hence for all $f \in C(\mathbb{S}^d)$, since the polynomials are dense in $C(\mathbb{S}^d)$.

Proposition 16. *Given $s > d/2$, a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ is asymptotically uniformly distributed on $\mathbb{S}^d$.*

Asymptotic uniformity of point set sequences is a relatively weak property: Even if we restrict our attention to $f \in \mathbb{H}^s(\mathbb{S}^d)$ for all s, f non-constant, it is possible to construct asymptotically uniformly distributed sequences of node sets so that the convergence of the quadrature error is as slow as one likes. (Without loss of

generality, we can restrict attention to smooth functions f whose integral over $\mathbb{S}^d$ is zero. One can then assign too many points to regions where f is positive and too few where it is negative, correcting the imbalance as $N \to \infty$ as slowly as one wishes.)

Low-discrepancy sequences. Unlike the situation for the unit cube, on the sphere $\mathbb{S}^d$ there is no single Koksma-Hlawka inequality (cf., e.g., [21]), as evidenced by the many competing Koksma-Hlawka-like inequalities proposed in the literature (e.g. [19], [14]). The notion of worst-case error (Definition 10) provides a way to bound the error of numerical integration for sufficiently smooth functions f; see (19).

The Sobolev space $\mathbb{H}^{(d+1)/2}(\mathbb{S}^d)$ with (28) as reproducing kernel is special in the sense that the worst-case error $\mathrm{wce}(Q[X_N]; \mathbb{H}^{(d+1)/2}(\mathbb{S}^d))$ of a QMC rule with node set $X_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ has an interpretation as the *spherical cap $\mathbb{L}_2$-discrepancy* of X_N, defined by

$$(36) \quad D_{\mathbb{L}_2}^C(X_N) := \left\{ \int_0^\pi \int_{\mathbb{S}^d} \left| \frac{|X_N \cap \mathcal{C}(\mathbf{x}; \theta)|}{N} - \sigma_d(\mathcal{C}(\mathbf{x}; \theta)) \right|^2 \mathrm{d}\,\sigma_d(\mathbf{x}) \sin\theta \, \mathrm{d}\theta \right\}^{1/2},$$

where $\mathcal{C}(\mathbf{z}; \theta) := \{\mathbf{y} \in \mathbb{S}^d : \mathbf{y} \cdot \mathbf{z} \geq \cos\theta\}$ denotes a spherical cap. Indeed, *Stolarsky's invariance principle* [38] asserts that

$$(37) \qquad \left(V_{-1}(\mathbb{S}^d) - \frac{1}{N^2} \sum_{j=1}^N \sum_{i=1}^N |\mathbf{x}_j - \mathbf{x}_i| \right)^{1/2} = \frac{1}{\sqrt{\gamma_d}} D_{\mathbb{L}_2}^C(X_N),$$

where $V_{-1}(\mathbb{S}^d)$ is given in (26) and γ_d is given in (8). Recalling equation (29), we recognize that the left-hand side above is the worst-case error for $\mathbb{H}^{(d+1)/2}(\mathbb{S}^d)$ with kernel $K_{\mathrm{dist}}(\mathbf{x} \cdot \mathbf{y})$ (cf. (28)).[§] As a consequence of (37) and (29) for $s = (d+1)/2$, we obtain the following corollary to Theorem 14.

Corollary 17. *Minimizers of the spherical cap $\mathbb{L}_2$-discrepancy form a sequence of QMC designs for $\mathbb{H}^{(d+1)/2}(\mathbb{S}^d)$.*

A related concept is that of the *spherical cap $\mathbb{L}_\infty$-discrepancy* of an N-point set X_N on $\mathbb{S}^d$, defined by

$$(38) \qquad D_{\mathbb{L}_\infty}^C(X_N) := \sup \left\{ \left| \frac{|X_N \cap \mathcal{C}|}{N} - \sigma_d(\mathcal{C}) \right| : \mathcal{C} \text{ spherical cap in } \mathbb{S}^d \right\}.$$

Note that $D_{\mathbb{L}_2}^C(X_N) \leq \sqrt{2}\, D_{\mathbb{L}_\infty}^C(X_N)$. We shall establish in Section 9 the following.

Proposition 18. *Under the assumptions of Proposition 11, given $s \geq (d+1)/2$, every N-point configuration on $\mathbb{S}^d$ satisfies*

$$(39) \qquad\qquad \mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) \leq c_{s,d}\, D_{\mathbb{L}_\infty}^C(X_N),$$

where $c_{s,d} > 0$ depends on the $\mathbb{H}^s(\mathbb{S}^d)$-norm, but is independent of N.

Thus, node sets with small spherical cap discrepancy are of some interest with regard to numerical integration. We remark that a sequence (X_N) of N-point configurations on $\mathbb{S}^d$ is asymptotically uniformly distributed if and only if $D_{\mathbb{L}_\infty}^C(X_N) \to 0$ as $N \to \infty$ (see, e.g., [21]).

[§]This connection was proved in [15].

Beck [8] proved that there is a positive number c_1 such that for **any** N-point set Z_N on $\mathbb{S}^d$ there exists a spherical cap $\mathcal{C}_N \subset \mathbb{S}^d$ such that

$$\frac{c_1}{N^{[(d+1)/2]/d}} < \left| \frac{|Z_N \cap \mathcal{C}_N|}{N} - \sigma_d(\mathcal{C}_N) \right|$$

and, by employing a probabilistic argument, that there exist $c_2 > 0$ and N-point sets Z_N^* on $\mathbb{S}^d$ such that

$$D_{\mathbb{L}_\infty}^C(Z_N^*) < c_2 \frac{\sqrt{\log N}}{N^{[(d+1)/2]/d}}.$$

This motivates the following definition.

Definition 19. A sequence (Z_N) of N-point configurations on $\mathbb{S}^d$ is said to be a *low-discrepancy sequence*[¶] if there exists a positive number β_d, independent of N, such that for all Z_N,

$$D_{\mathbb{L}_\infty}^C(Z_N) \leq \beta_d \frac{\sqrt{\log N}}{N^{[(d+1)/2]/d}}.$$

Let (Z_N) be a low-discrepancy sequence on $\mathbb{S}^d$. Using (39) and Theorem 3, we see that for each $s \geq (d+1)/2$,

$$\frac{\beta_1(s,d)}{N^{s/d}} \leq \text{wce}(Q[Z_N]; \mathbb{H}^s(\mathbb{S}^d)) \leq c_{s,d}\, D_{\mathbb{L}_\infty}^C(Z_N) \leq \beta_2(s,d) \frac{\sqrt{\log N}}{N^{[(d+1)/2]/d}}.$$

On the other hand, if $\text{wce}(Q[Z_N]; \mathbb{H}^s(\mathbb{S}^d)) < 1$, we have by Lemma 23 of Section 9 (with s replaced by $(d+1)/2$ and s' replaced by s) that, for each $s < (d+1)/2$,

$$(40) \qquad \frac{\beta_1(s,d)}{N^{s/d}} \leq \text{wce}(Q[Z_N]; \mathbb{H}^s(\mathbb{S}^d)) \leq \beta_3(s,d) \frac{(\log N)^{s/(d+1)}}{N^{s/d}}.$$

From (40) we see that for every $s \in (d/2, (d+1)/2)$, low-discrepancy sequences on $\mathbb{S}^d$ have almost optimal order of the worst-case error for $\mathbb{H}^s(\mathbb{S}^d)$, except for a power of $\log N$. This leads to the following natural question: *Do low-discrepancy sequences on $\mathbb{S}^d$ form sequences of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ when $s \in (d/2, (d+1)/2)$?*

We remark that, to the authors' knowledge no explicit constructions of low-discrepancy configurations on $\mathbb{S}^d$ are known, in contrast to the situation for the unit cube. [‖]

7. QMC DESIGNS ARE BETTER THAN AVERAGE

As we show in this section, Theorem 7 is a consequence of a more general result dealing with the average value of the N-point energy defined by a positive definite kernel on the sphere. Given a sequence $(a_\ell)_{\ell \geq 1}$ with $a_\ell \geq 0$ and $\sum_{\ell=1}^\infty a_\ell Z(d, \ell)$ convergent, we set for $\mathbf{x}, \mathbf{y} \in \mathbb{S}^d$

$$(41) \quad \mathcal{A}(\mathbf{x} \cdot \mathbf{y}) := \mathcal{A}(\mathbf{x}, \mathbf{y}) = \sum_{\ell=1}^\infty a_\ell \sum_{k=1}^{Z(d,\ell)} Y_{\ell,k}(\mathbf{x}) Y_{\ell,k}(\mathbf{y}) = \sum_{\ell=1}^\infty a_\ell Z(d, \ell) P_\ell^{(d)}(\mathbf{x} \cdot \mathbf{y}).$$

[¶]This terminology is short for *low spherical cap $\mathbb{L}_\infty$-discrepancy sequence.*

[‖]In [2] it is proved that the spherical cap discrepancy of so-called spherical digital nets and spherical Fibonacci points is bounded by $CN^{-1/2}$ (for some explicit C), the same rate as for random points.

Then $\mathcal{A}$ is a positive definite kernel on $\mathbb{S}^d$ in the sense of Schoenberg [34]; indeed, from the addition theorem, for all N-point configurations $X_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^d$ and all real numbers $\alpha_1, \ldots, \alpha_N$,

$$\sum_{j=1}^{N} \sum_{i=1}^{N} \alpha_j \, \mathcal{A}(\mathbf{x}_j \cdot \mathbf{x}_i) \, \alpha_i = \sum_{\ell=1}^{\infty} a_\ell \sum_{k=1}^{Z(d,\ell)} \left| \sum_{j=1}^{N} \alpha_j Y_{\ell,k}(\mathbf{x}_j) \right|^2 \geq 0.$$

Consequently, the normalized N-point energy of this kernel, defined by

$$(42) \qquad \mathcal{A}[X_N] := \frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} \mathcal{A}(\mathbf{x}_j \cdot \mathbf{x}_i), \qquad X_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^d,$$

satisfies $\mathcal{A}[X_N] \geq 0$.

We are interested in the expected value of the energy (42) for N random points that are chosen independently and identically distributed with respect to the uniform measure on $\mathbb{S}^d$; namely

$$\mathbb{E}\mathcal{A}[X_N] := \int_{\mathbb{S}^d} \cdots \int_{\mathbb{S}^d} \mathcal{A}[X_N] \, \mathrm{d}\,\sigma_d(\mathbf{x}_1) \cdots \mathrm{d}\,\sigma_d(\mathbf{x}_N),$$

which is also known as the *spherical average* of $\mathcal{A}[X_N]$ over all N-point sets $X_N \subset \mathbb{S}^d$. A straightforward computation yields the following result (see Section 9).

Theorem 20. *Given the kernel $\mathcal{A}$ as in* (41),

$$(43) \qquad\qquad\qquad \mathbb{E}\mathcal{A}[X_N] = \frac{\mathcal{A}(1)}{N}.$$

Remark. Theorem 20 generalizes the result [37, Theorem 6].

We now apply this result to the positive definite kernel $\mathcal{K}^{(s)}$ associated with the reproducing kernel $K^{(s)}$ as given in (17) for $\mathbb{H}^s(\mathbb{S}^d)$ with $s > d/2$. By Proposition 11, we have

$$\mathcal{K}^{(s)}[X_N] = \frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} \mathcal{K}^{(s)}(\mathbf{x}_j \cdot \mathbf{x}_i) = \left[\mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) \right]^2,$$

and hence by Theorem 20 we obtain that the expected value of the squared worst-case error is given by $\mathcal{K}^{(s)}(1)/N$, from which Theorem 7 follows. Consequently, sequences of randomly chosen points on the sphere do not generate QMC designs for $s > d/2$.

If, instead of choosing points randomly over the whole sphere, we stratify our approach by requiring that the N points be randomly chosen from N different equal area subsets of $\mathbb{S}^d$ having small diameter, then on average we will obtain a sequence of QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ whenever $s \in (d/2, d/2 + 1)$. In the formal statement of this result we denote by

$$\mathrm{diam}\, A := \sup \left\{ |\mathbf{x} - \mathbf{y}| : \mathbf{x}, \mathbf{y} \in A \right\}$$

the diameter of the set A.

Theorem 21. *Let $(\mathcal{D}_N)$ be a sequence of partitions of $\mathbb{S}^d$ into N equal area subsets $D_{j,N}$, $j = 1, \ldots, N$, such that $\mathrm{diam}\, D_{j,N} \leq c/N^{1/d}$, where c is a positive constant independent of j and N. Let $X_N = \{\mathbf{x}_{1,N}, \ldots, \mathbf{x}_{N,N}\}$, where $\mathbf{x}_{j,N}$ is*

chosen randomly from $D_{j,N}$ with respect to uniform measure on $D_{j,N}$. Then, for $d/2 < s < d/2 + 1$,

$$\frac{\beta'}{N^{s/d}} \leq \sqrt{\mathbb{E}\left[\,\{\mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\}^2\,\right]} \leq \frac{\beta}{N^{s/d}}, \tag{44}$$

where $\beta' > 0$ and $\beta > 0$ depend on the $\mathbb{H}^s(\mathbb{S}^d)$-norm, but are independent of N, and $\beta > 0$ also depends on $(\mathcal{D}_N)$.

We remark that such a sequence of partitions always exists; see [9, 12, 30, 32].

However, as the next result shows, the stratification strategy does not lead, on average, to QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ with $s > d/2 + 1$.

Theorem 22. *Let $(\mathcal{D}_N)$ and X_N be as in Theorem 21. Then for $s > d/2+1$, with $2s - d$ not an even integer,[**]*

$$\sqrt{\mathbb{E}\left[\,\{\mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\}^2\,\right]} \geq \frac{\beta}{N^{(d/2+1)/d}},$$

where $\beta > 0$ depends on the $\mathbb{H}^s(\mathbb{S}^d)$-norm and $(\mathcal{D}_N)$, but is independent of N.

Other concepts of randomness can also be considered. Armentano, Beltrán, and Shub [4] study point configurations on $\mathbb{S}^2$ that are derived from the zeros of random polynomials. This will be a topic of future research.

8. Numerical experiments

8.1. **Point sets.** Many different sequences of point sets on the sphere have been introduced in the literature. We consider the following point sets X_N:

- Pseudo-random points, uniformly distributed on the sphere.
- Equal area points based on an algorithm given in [32].
- Fekete points which maximize the determinant for polynomial interpolation [36].
- Coulomb energy points, which minimize
$$\sum_{j=1}^{N}\sum_{i=1}^{N}\frac{1}{|\mathbf{x}_j - \mathbf{x}_i|}.$$
- Log energy points, which minimize
$$\sum_{j=1}^{N}\sum_{i=1}^{N}\log\frac{1}{|\mathbf{x}_j - \mathbf{x}_i|}.$$
- Generalized spiral points (cf. [7, 32]), with spherical coordinates (θ_j, ϕ_j) given by
$$z_j = 1 - \frac{2j-1}{N}, \quad \theta_j = \cos^{-1}(z_j), \quad \phi_j = 1.8\sqrt{N}\theta_j \mod 2\pi, \quad j = 1, \ldots, N.$$
- Distance points, which maximize
$$\sum_{j=1}^{N}\sum_{i=1}^{N}|\mathbf{x}_j - \mathbf{x}_i|.$$
- Spherical t-designs with $N = \lceil (t+1)^2/2 \rceil + 1$ points.

[**]A similar result for $2s - d$ an even integer can be obtained using an appropriate kernel as discussed in Section 5.

All the point sets that are characterized by optimizing a criterion are faced with the difficulty of many local optima. Thus, for larger values of N, these point sets have objective values near, but not necessarily equal to, the global optimum.

QMC designs for $\mathbb{H}^s\left(\mathbb{S}^2\right)$, for every $s > 1$ (s not an integer), could be calculated by optimizing the expression (33) in terms of the generalized distance $|\mathbf{x}_j - \mathbf{x}_i|^{2s-2}$ (including the low order polynomial $\mathcal{Q}_L$ or by imposing the additional constraints that the point set is an L-design), as discussed in Section 5. We restrict attention to the point sets listed above, which are available from the website [41]. Here the criterion (33) is used to generate points only for the case $s = 3/2$ (maximizing distance sums). In all cases, except the last, the number of points N was taken to be a perfect square.

8.2. **Worst-case error.** The worst-case errors in $\mathbb{H}^s\left(\mathbb{S}^2\right)$ for $s = 3/2$ are illustrated in Figure 1. For all point sets, the worst-case error with $s = 3/2$ is calculated using (29) and the distance kernel. Apart from the random points (which are not QMC designs, see Theorem 7) all the point sets have a worst-case error for $s = 3/2$ that seems to decay like $N^{-3/4}$, implying that they are all QMC designs for $s = 3/2$.
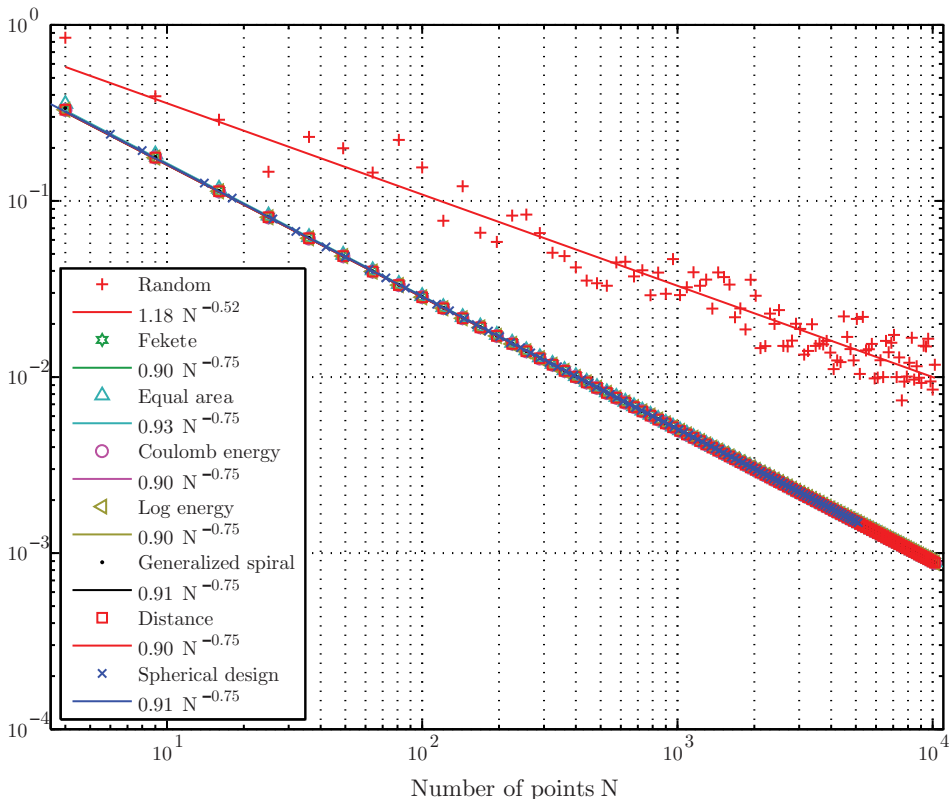

FIGURE 1. Worst-case error for $\mathbb{H}^s(\mathbb{S}^2)$ and $s = 3/2$
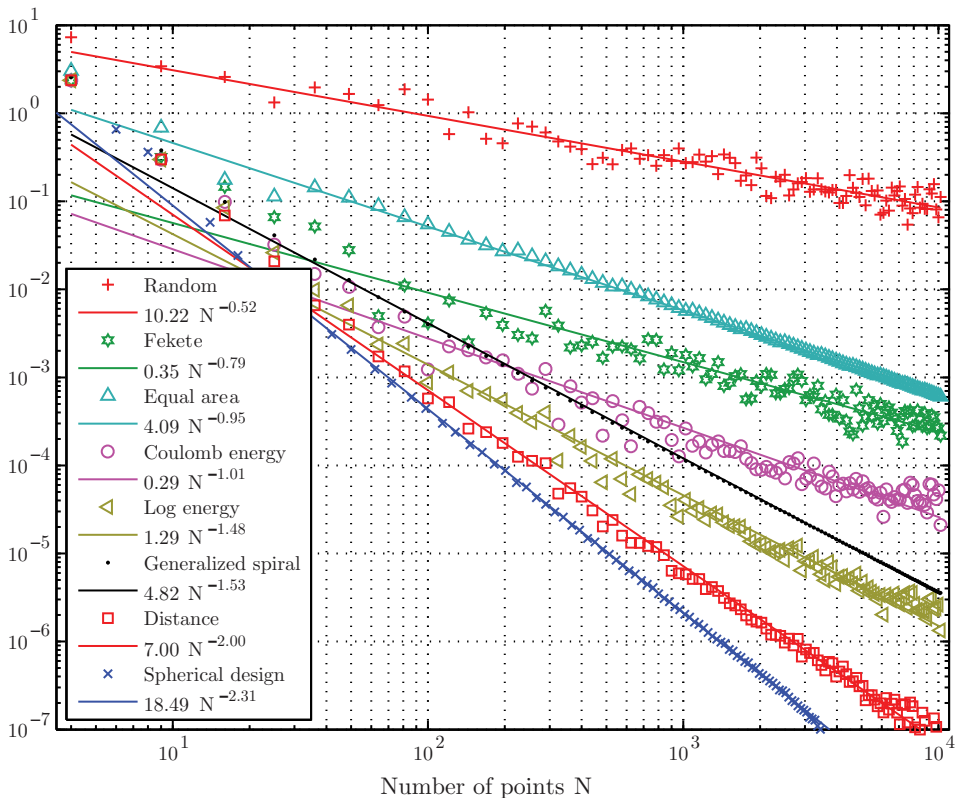
FIGURE 2. Worst-case error for $\mathbb{H}^s(\mathbb{S}^2)$ and $s = 4.5$

Spherical designs with $N = \mathcal{O}(t^2)$ are QMC designs for $\mathbb{H}^s(\mathbb{S}^2)$ and all $s > 1$. From Section 5 and Theorem 14, the distance points are provably QMC designs for $s = 1.5$. The rate of decay for equal area points fits well with Theorem 21, which established that randomized equal area points are also QMC designs for $1 < s < 2$. Moreover, from Theorem 22, randomized equal area points cannot do better than this. Other than the distance points, it has yet to be established rigorously that the non-random point sets are QMC designs for $s = 1.5$. It is rather curious that, for the non-random sequences, the computed worst-case errors in Figure 1 essentially lie on the same curve.

Figure 2 plots the worst-case errors for $s = 4.5$ and estimates the rate of decay by finding a least squares fit of the form $\alpha N^{-\beta}$ for $10 \leq N \leq 10^4$ (except for the spherical designs which use $10 \leq N \leq 5 \times 10^3$). For all point sequences the worst-case error with $s = 4.5$ is calculated using the generalized distance kernel and (33) (for which $L = 3$). As expected from Theorem 7, for random points, the worst-case error still decays like $N^{-1/2}$.

All the point sets, except for the spherical designs, exhibit varying rates of decay slower than $\mathcal{O}(N^{-s/2}) = \mathcal{O}(N^{-2.25})$, indicating that their effectiveness for equal weight numerical integration on $\mathbb{H}^s(\mathbb{S}^2)$ when $s \geq 4.5$ is less than optimal; cf. Theorem 13.
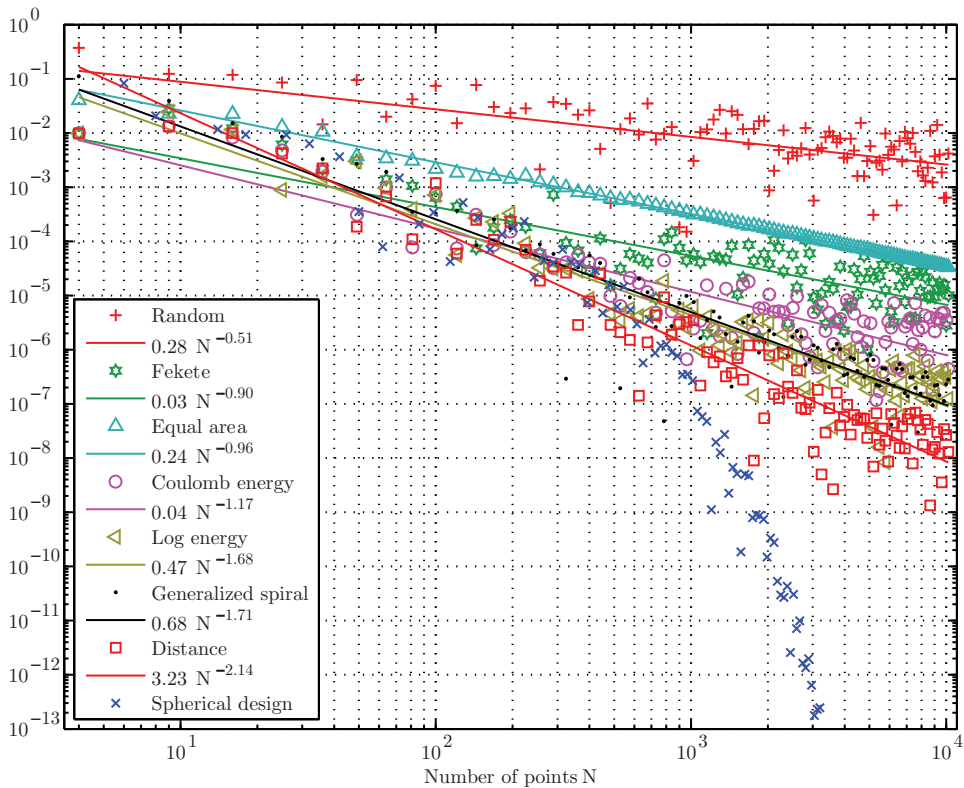
FIGURE 3. Numerical integration errors for the Franke function

8.3. **Integrating a smooth function.** One expects from the error bound (19) that a putative sequence (X_N) of QMC designs for $\mathbb{H}^s(\mathbb{S}^2)$, $s > 1$, will play out its full strength when tested with a smooth test function. Some caution is needed in the choice of this function in order to avoid having an integrand that is accidentally too easy. Our choice is the *Franke* function for the sphere [33] defined by

$$f(x,y,z) := 0.75 \exp(-(9x-2)^2/4 - (9y-2)^2/4 - (9z-2)^2/4)$$
$$+ 0.75 \exp(-(9x+1)^2/49 - (9y+1)/10 - (9z+1)/10)$$
$$+ 0.5 \exp(-(9x-7)^2/4 - (9y-3)^2/4 - (9z-5)^2/4)$$
$$- 0.2 \exp(-(9x-4)^2 - (9y-7)^2 - (9z-5)^2), \qquad (x,y,z)^T \in \mathbb{S}^2,$$

which is in $C^\infty(\mathbb{S}^2)$ and for which

$$\int_{\mathbb{S}^2} f(\mathbf{x})\mathrm{d}\sigma_2(\mathbf{x}) = 0.5328652500843890\ldots.$$

As $f \in \mathbb{H}^s(\mathbb{S}^2)$ for all $s > 1$, the integration error for a particular sequence (X_N) of QMC designs with s^* given by (5), must decay at least as fast as $O\left(N^{-s^*/2+\epsilon}\right)$ for any $\epsilon > 0$. In Figure 3, the faster than algebraic decay of the numerical integration error for spherical designs is apparent.

TABLE 1. Estimates of s^* for $d = 2$

Point set	s^*
Fekete	1.5
Equal area	2
Coulomb energy	2
Log energy	3
Generalized spiral	3
Distance	4
Spherical designs	∞

8.4. **Estimating s^*.** For a given sequence (X_N), the QMC strength s^*, defined by (5), is estimated by calculating the worst-case error for $\mathbb{H}^s(\mathbb{S}^2)$ for some collection of test values $s > d/2$. Any such test value s for which the numerically found convergence rate is not optimal gives an upper bound on s^* (cf. Theorems 4 and 13). Similarly, if the error for the Franke function is decaying approximately like $N^{-\bar{s}/2}$, then that point set can only be a QMC design for $s \leq \bar{s}$.

Some conjectured values of s^* are given in Table 1, based on the results in Figure 2, similar experiments with different values of s, and the results in Figure 3. For example, the equal area points have an estimated rate of decay $N^{-0.95}$ for the worst-case error with $s = 4.5$ in Figure 2, while the error decays like $N^{-0.96}$ for the Franke function in Figure 3, leading us to conjecture that $s^* \approx 2$.

Determining the precise value of s^* is very much an open question.

9. PROOFS

Throughout this proof section we use the shortened notation Q_N for a QMC rule $Q[X_N]$ defined by a node set $X_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^d$.

9.1. **Proofs of Section 1 results.** The proof of Theorem 4 requires the following lemma.

Lemma 23. *Given $s > d/2$, if* $\mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) < 1$, *then*

$$(45) \qquad \mathrm{wce}(Q_N; \mathbb{H}^{s'}(\mathbb{S}^d)) < c(d, s, s') \left[\mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d))\right]^{s'/s}, \qquad d/2 < s' < s,$$

where $c(d, s, s') > 0$ depends on the norms for $\mathbb{H}^s(\mathbb{S}^d)$ and $\mathbb{H}^{s'}(\mathbb{S}^d)$, but is independent of N.

Proof. Relation (45) follows from Theorem 3.1 in Brandolini et al. [13]. For the sake of completeness we give here a proof of (45) along the lines of the proof of Theorem 3.1 tailored to our needs and specifically to the case of spheres $\mathbb{S}^d$. Throughout this proof we use the canonical kernel $K_{\mathrm{can}}^{(s)}$.

Writing $1/(1 + \lambda_\ell)^{s'}$ in terms of a Laplace transform (cf. [1, Eq. 5.9.1]),

$$\frac{1}{(1 + \lambda_\ell)^{s'}} = \frac{1}{\Gamma(s')} \int_0^\infty e^{-(1+\lambda_\ell)\tau} \tau^{s'-1} \, d\tau,$$

and applying Proposition 11, we obtain for all $d/2 < s' \leq s$,

$$(46) \qquad \left[\mathrm{wce}(Q_N; \mathbb{H}^{s'}(\mathbb{S}^d))\right]^2 = \frac{1}{\Gamma(s')} \int_0^\infty \tau^{s'-1} e^{-\tau} g(\tau) \, d\tau.$$

Here

$$g(\tau):=g(\tau; \mathbf{x}_1, \ldots, \mathbf{x}_N):=\sum_{\ell=1}^{\infty} e^{-\lambda_\ell \tau} \sum_{k=1}^{Z(d,\ell)} \left| \frac{1}{N} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) \right|^2$$

$$= \frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} \mathcal{H}(\tau, \mathbf{x}_j \cdot \mathbf{x}_i),$$

where calligraphic $\mathcal{H}$ denotes the *heat kernel* with the constant term removed:

$$H(\tau, \mathbf{x}, \mathbf{y}):= \sum_{\ell=0}^{\infty} e^{-\lambda_\ell \tau} \sum_{k=1}^{Z(d,\ell)} Y_{\ell,k}(\mathbf{x}) Y_{\ell,k}(\mathbf{y})$$

$$= \sum_{\ell=0}^{\infty} e^{-\lambda_\ell \tau} Z(d,\ell) P_\ell^{(d)}(\mathbf{x} \cdot \mathbf{y}), \mathbf{x}, \mathbf{y} \in \mathbb{S}^d,$$

which is the fundamental solution to the heat equation $\partial u / \partial \tau + \Delta_d^* u = 0$ on $\mathbb{R}_+ \times \mathbb{S}^d$. Interchanging integration and summation in (46) is justified, because the heat kernel is uniformly continuous on $[\varepsilon, \infty) \times \mathbb{S}^d \times \mathbb{S}^d$ for $\varepsilon > 0$.

Let $d/2 < s' < s$ and set $\varepsilon := [\mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d))]^{2/s}$. (Then $\varepsilon < 1$ by assumption.) We split the right-hand side of (46) into three parts and use the following estimates:
(i) For "large values" of τ,

$$\frac{1}{\Gamma(s')} \int_1^{\infty} \tau^{s'-1} e^{-\tau} g(\tau) \, \mathrm{d}\tau < \frac{1}{\Gamma(s')} \int_1^{\infty} \tau^{s-1} e^{-\tau} g(\tau) \, \mathrm{d}\tau$$

$$\leq \frac{\Gamma(s)}{\Gamma(s')} \frac{1}{\Gamma(s)} \int_0^{\infty} \tau^{s-1} e^{-\tau} g(\tau) \, \mathrm{d}\tau$$

$$= \frac{\Gamma(s)}{\Gamma(s')} \varepsilon^s < \frac{\Gamma(s)}{\Gamma(s')} \varepsilon^{s'},$$

where we used (46) with s' replaced by s.
(ii) For "medium values" of τ,

$$\frac{1}{\Gamma(s')} \int_{\varepsilon/2}^1 \tau^{s'-1} e^{-\tau} g(\tau) \, \mathrm{d}\tau = \frac{1}{\Gamma(s')} \int_{\varepsilon/2}^1 \tau^{s-1} \tau^{s'-s} e^{-\tau} g(\tau) \, \mathrm{d}\tau$$

$$\leq \frac{(\varepsilon/2)^{s'-s}}{\Gamma(s')} \int_{\varepsilon/2}^1 \tau^{s-1} e^{-\tau} g(\tau) \, \mathrm{d}\tau$$

$$< (\varepsilon/2)^{s'-s} \frac{\Gamma(s)}{\Gamma(s')} \frac{1}{\Gamma(s)} \int_0^{\infty} \tau^{s-1} e^{-\tau} g(\tau) \, \mathrm{d}\tau$$

$$= 2^{s-s'} \frac{\Gamma(s)}{\Gamma(s')} \varepsilon^{s'-s} \varepsilon^s$$

$$= 2^{s-s'} \frac{\Gamma(s)}{\Gamma(s')} \varepsilon^{s'}.$$

(iii) For "small values" of τ, we appeal to the small time Gaussian estimate on the heat kernel (cf. [39]); that is, if $0 < \tau < \varepsilon/2$, then for some $c > 0$,

$$\tau^{d/2} H(\tau, \mathbf{x}, \mathbf{y}) \leq c \, \varepsilon^{d/2} H(\varepsilon, \mathbf{x}, \mathbf{y}),$$

which in turn implies, on the assumption that $g(\tau)$ is uniformly bounded on $[0, 1)$,

$$\tau^{d/2} g(\tau) = \frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} \tau^{d/2} H(\tau, \mathbf{x}_j, \mathbf{x}_i) - \tau^{d/2} \leq c\, \varepsilon^{d/2} \left(g(\varepsilon) + 1 \right) \leq c''' \varepsilon^{d/2}.$$

Postponing the proof of the uniform boundedness, it then follows that

$$\frac{1}{\Gamma(s')} \int_0^{\varepsilon/2} \tau^{s'-1} e^{-\tau} g(\tau) \, \mathrm{d}\tau = \frac{1}{\Gamma(s')} \int_0^{\varepsilon/2} \tau^{s'-d/2-1} e^{-\tau}\, t^{d/2} g(\tau)\, \mathrm{d}\tau$$

$$\leq \frac{c'''}{\Gamma(s')} \varepsilon^{d/2} \int_0^{\varepsilon/2} \tau^{s'-d/2-1} e^{-\tau}\, \mathrm{d}\tau$$

$$\leq \frac{c'''}{\Gamma(s')} \varepsilon^{d/2} \int_0^{\varepsilon/2} \tau^{s'-d/2-1}\, \mathrm{d}\tau = \frac{c'''}{\Gamma(s')} \frac{\varepsilon^{s'}}{s' - d/2}.$$

From (i), (ii) and (iii) we get the required estimate

$$\left[\mathrm{wce}(Q_N; \mathbb{H}^{s'}(\mathbb{S}^d)) \right]^2 \leq \left(\frac{\Gamma(s)}{\Gamma(s')} + 2^{s-s'} \frac{\Gamma(s)}{\Gamma(s')} + \frac{c'''}{\Gamma(s')} \frac{1}{s' - d/2} \right) \varepsilon^{s'}.$$

$$\leq c'''' \left[\mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) \right]^{2s'/s}.$$

It remains to prove uniform boundedness of $g(\tau)$ for $0 \leq \tau < 1$ (we use (19) and the fact that $\int_{\mathbb{S}^d} \mathcal{H}(\tau, \mathbf{x}, \mathbf{x}_k)\, \mathrm{d}\sigma_d(\mathbf{x}) = 0$):

$$0 < g(\tau) = \frac{1}{N} \sum_{j=1}^{N} \frac{1}{N} \sum_{i=1}^{N} \mathcal{H}(\tau, \mathbf{x}_j, \mathbf{x}_i) \leq \mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) \left\| \frac{1}{N} \sum_{j=1}^{N} \mathcal{H}(\tau, \cdot, \mathbf{x}_j) \right\|_{\mathbb{H}^s}$$

$$= \mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) \left\| \sum_{\ell=1}^{\infty} \sum_{k=1}^{Z(d,\ell)} e^{-\lambda_\ell \tau} \left(\frac{1}{N} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) \right) Y_{\ell,k}(\cdot) \right\|_{\mathbb{H}^s}$$

$$= \mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) \left[\sum_{\ell=1}^{\infty} \sum_{k=1}^{Z(d,\ell)} (1 + \lambda_\ell)^s e^{-2\lambda_\ell \tau} \left| \frac{1}{N} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) \right|^2 \right]^{1/2}$$

$$\leq \mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) \sup_{\ell \geq 1} \left\{ (1 + \lambda_\ell)^s e^{-\lambda_\ell \tau} \right\}$$

$$\times \left[\sum_{\ell=1}^{\infty} \sum_{k=1}^{Z(d,\ell)} (1 + \lambda_\ell)^{-s} \left| \frac{1}{N} \sum_{j=1}^{N} Y_{\ell,k}(\mathbf{x}_j) \right|^2 \right]^{1/2}$$

$$= \left[\mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) \right]^2 \sup_{\ell \geq 1} \left\{ (1 + \lambda_\ell)^s e^{-\lambda_\ell \tau} \right\} = \tau^s \sup_{\ell \geq 1} \left\{ (1 + \lambda_\ell)^s e^{-\lambda_\ell \tau} \right\}.$$

The function $(1 + \lambda)^s e^{-\lambda \tau}$ has a unique maximum at $\lambda = s/\tau - 1$ with value $s^s \tau^{-s} e^{\tau - s}$. Hence

$$0 < g(\tau) \leq s^s e^{\tau - s} \leq s^s e^{1-s}. \qquad \square$$

Proof of Theorem 4. If (X_N) is a sequence of N-point QMC designs for $\mathbb{H}^s(\mathbb{S}^d)$ for $s > d/2$, then $\mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) \to 0$ as $N \to \infty$. Thus $\mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) < 1$ for all

$N > N_0$ for some $N_0 > 0$. By (45),

$$\mathrm{wce}(Q_N; \mathbb{H}^{s'}(\mathbb{S}^d)) < c(d, s, s') \left[\mathrm{wce}(Q_N; \mathbb{H}^s(\mathbb{S}^d)) \right]^{s'/s}$$

$$< c(d, s, s') \left[\frac{c(s, d)}{N^{s/d}} \right]^{s'/s} = \frac{c''}{N^{s'/d}}$$

for every $N > N_0$ for each $d/2 < s' < s$. The finitely many exceptions with $N \leq N_0$ satisfy the last inequality with a possibly larger constant c'', depending only on the norms of $\mathbb{H}^s(\mathbb{S}^d)$ and $\mathbb{H}^{s'}(\mathbb{S}^d)$. Consequently, (X_N) is a sequence of QMC designs for $\mathbb{H}^{s'}(\mathbb{S}^d)$ for $d/2 < s' < s$. $\qquad\square$

Proof of Theorem 6. Let $d \geq 2$. By Theorem 8 there exists a sequence (Y_{N_t}) of spherical t-designs Y_{N_t} with $N_t := m_d\, t^d$ points ($t \geq 1$) for some suitably large positive integer m_d. Furthermore, the theorem states that there exist spherical t-designs for every cardinality $\geq m_d\, t^d$. Thus we can fill the gaps in the sequence (Y_{N_t}) by adding spherical t-designs with N points for $N_t < N < N_{t+1}$. If necessary we choose for $N = 2, \ldots, m_d$ spherical 1-designs with N points; that is, configurations with centroid 0. This gives a new sequence $(\widehat{Y}_N)_{N \geq 2}$.

By Theorem 1 there exists a constant $C(s, d) > 0$ such that

$$(47) \qquad \sup_{\substack{f \in \mathbb{H}^s(\mathbb{S}^d), \\ \|f\|_{\mathbb{H}^s} \leq 1}} \left| \frac{1}{N} \sum_{\mathbf{y} \in \widehat{Y}_N} f(\mathbf{y}) - \int_{\mathbb{S}^d} f(\mathbf{y})\, \mathrm{d}\,\sigma_d(\mathbf{y}) \right| \leq \frac{C(s, d)}{t^s} = \frac{C(s, d)}{N^{s/d}} \left(\frac{N^{1/d}}{t} \right)^s$$

for all $N \geq 1$. Since

$$c_d^{1/d} = \frac{N_t^{1/d}}{t} \leq \frac{N^{1/d}}{t} \leq \frac{N_{t+1}^{1/d}}{t} = c_d^{1/d} \frac{t+1}{t} \qquad \text{for all } N_t \leq N \leq N_{t+1},$$

the right-hand side of (47) satisfies for all $N \geq 1$,

$$\frac{C(s, d)}{N^{s/d}} \left(\frac{N^{1/d}}{t} \right)^s \leq \frac{c_d^{s/d} C(s, d)\, (1 + 1/t)^s}{N^{s/d}} \leq \frac{C'(s, d)}{N^{s/d}}, \qquad C'(s, d) := 2^s c_d^{s/d} C(s, d).$$

Consequently, $(\widehat{Y}_N)_{N \geq 2}$ is a sequence of generic QMC designs. $\qquad\square$

9.2. **Proofs of Section 3 results.**

Proof of Theorem 12. For $N \geq 2$, let $X_N^* = \{\mathbf{x}_{1,N}^*, \ldots, \mathbf{x}_{N,N}^*\}$ be as in Theorem 12 and $Y_N = \{\mathbf{y}_{1,N}, \ldots, \mathbf{y}_{N,N}\}$ be as in Theorem 6. By minimality of the X_N^*'s, for every $N \geq 2$,

$$\sum_{j=1}^{N} \sum_{i=1}^{N} \mathcal{K}^{(s)}(\mathbf{x}_{j,N}^* \cdot \mathbf{x}_{i,N}^*) \leq \sum_{j=1}^{N} \sum_{i=1}^{N} \mathcal{K}^{(s)}(\mathbf{y}_{j,N} \cdot \mathbf{y}_{i,N}).$$

Hence, by Proposition 11,

$$\mathrm{wce}(Q[X_N^*]; \mathbb{H}^s(\mathbb{S}^d)) \leq \mathrm{wce}(Q[Y_N]; \mathbb{H}^s(\mathbb{S}^d)) \leq \frac{c(s, d)}{N^{s/d}},$$

where the last inequality follows by Theorem 6. $\qquad\square$

9.3. **Proofs of Section 6 results.**

Proof of Proposition 18. We recall that two different kernels for the same $\mathbb{H}^s(\mathbb{S}^d)$ yield worst-case errors that can differ by at most constant factors, and also that for the particular case of the canonical kernel $K_{\mathrm{can}}^{(s)}$, the worst-case error decreases monotonically with s for $s > d/2$; cf. relation (23). Using these facts and (37), which gives the worst-case error with respect to K_{dist}, we have for $s \geq (d+1)/2$ the following estimates:

$$\mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) \leq \beta_1 \, \mathrm{wce}(Q[X_N]; \mathbb{H}^{(d+1)/2}(\mathbb{S}^d)) \leq \beta_2 D_{\mathbb{L}_2}^C(X_N) \leq \beta_3 D_{\mathbb{L}_\infty}^C(X_N),$$

where the positive constants β_1, β_2 and β_3 depend on the chosen norms. $\square$

9.4. **Proofs of Section 7 results.**

Proof of Theorem 20. Let $\mathcal{A}^L(z)$ be the truncated series

$$\mathcal{A}^L(z) := \sum_{\ell=1}^L a_\ell \, Z(d,\ell) \, P_\ell^{(d)}(z), \qquad z \in [-1,1].$$

On separating the diagonal and off-diagonal terms of the double sum

$$\mathcal{A}^L[X_N] := \frac{1}{N^2} \sum_{j=1}^N \sum_{i=1}^N \mathcal{A}^L(\mathbf{x}_j \cdot \mathbf{x}_i), \qquad X_N = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{S}^d,$$

we obtain

$$\mathbb{E}\mathcal{A}^L[X_N] = \int_{\mathbb{S}^d} \cdots \int_{\mathbb{S}^d} \left[\frac{\mathcal{A}^L(1)}{N} + \frac{1}{N^2} \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{i=1}^N \mathcal{A}^L(\mathbf{x}_j \cdot \mathbf{x}_i) \right] \mathrm{d}\,\sigma_d(\mathbf{x}_1) \cdots \mathrm{d}\,\sigma_d(\mathbf{x}_N)$$

$$= \frac{\mathcal{A}^L(1)}{N} + \frac{1}{N^2} \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{i=1}^N \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \mathcal{A}^L(\mathbf{x}_j \cdot \mathbf{x}_i) \, \mathrm{d}\,\sigma_d(\mathbf{x}_j) \, \mathrm{d}\,\sigma_d(\mathbf{x}_i)$$

$$= \frac{\mathcal{A}^L(1)}{N} + \frac{N(N-1)}{N^2} \sum_{\ell=1}^L a_\ell \sum_{k=1}^{Z(d,\ell)} \left(\int_{\mathbb{S}^d} Y_{\ell,k}(\mathbf{x}) \, \mathrm{d}\,\sigma_d(\mathbf{x}) \right)^2$$

$$= \frac{\mathcal{A}^L(1)}{N}.$$

For constant coefficients $a_1 = a_2 = \cdots = a_L = 1$ this is the result of Theorem 6 in [37].

Note that $\mathcal{A}^L[X_N] - \mathcal{A}^{L-1}[X_N] = a_L \sum_{k=1}^{Z(d,L)} (\frac{1}{N} \sum_{j=1}^N Y_{L,k}(\mathbf{x}_j))^2 \geq 0$. Hence, the quantities $\mathcal{A}^1[X_N]$, $\mathcal{A}^2[X_N]$, ... form a pointwise non-decreasing sequence of non-negative $(\sigma_d \cdots \sigma_d)$-measurable functions with limit function $\mathcal{A}[X_N]$. By the monotone convergence theorem it follows that

$$\mathbb{E}\mathcal{A}[X_N] = \int_{\mathbb{S}^d} \cdots \int_{\mathbb{S}^d} \mathcal{A}[X_N] \, \mathrm{d}\,\sigma_d(\mathbf{x}_1) \cdots \mathrm{d}\,\sigma_d(\mathbf{x}_N)$$

$$= \lim_{L\to\infty} \mathbb{E}\mathcal{A}^L[X_N] = \lim_{L\to\infty} \frac{\mathcal{A}^L(1)}{N} = \frac{\mathcal{A}(1)}{N}.$$

This completes the proof. $\square$

Proof of Theorem 21. We follow the proof idea leading to [32, Theorem 2.2]. Let $\mathcal{D}_N = \{D_{j,N}, \ldots, D_{N,N}\}$ be an equal area partition of $\mathbb{S}^d$ into subsets with small diameter; that is: $\bigcup_{j=1}^N D_{j,N} = \mathbb{S}^d$, where $\sigma_d(D_{j,N} \cap D_{k,N}) = 0$ for all $j, k = 1, \ldots, N$ with $j \neq k$ and $\sigma_d(D_{j,N}) = 1/N$. Furthermore, $\operatorname{diam} D_{j,N} \leq c/N^{1/d}$ for some c not depending on N. Each $D_{j,N}$ is equipped with the probability measure

$$(48) \qquad \mu_{j,N} := \frac{\sigma_d\big|_{D_{j,N}}}{\sigma_d(D_{j,N})}.$$

Let $d/2 < s < d/2 + 1$. Then the expected value of the squared worst-case error $[\operatorname{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))]^2$ for the space $\mathbb{H}^s(\mathbb{S}^d)$ provided with the kernel $K_{\mathrm{gd}}^{(s)}$ given in (27) when the j-th node is chosen randomly from $D_{j,N}$ (with respect to uniform measure on $D_{j,N}$) is (see (29))

$$\mathbb{E}\left[\left\{\operatorname{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\right\}^2\right]$$
$$= \int_{D_{1,N}} \cdots \int_{D_{N,N}} \left[\operatorname{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\right]^2 \mathrm{d}\,\mu_{1,N}(\mathbf{x}_1) \cdots \mathrm{d}\,\mu_{N,N}(\mathbf{x}_N)$$
$$= V_{d-2s}(\mathbb{S}^d) - \frac{1}{N^2} \sum_{\substack{j=1 \\ }}^N \sum_{\substack{i=1 \\ j\neq i}}^N \int_{D_{j,N}} \int_{D_{i,N}} |\mathbf{x}_j - \mathbf{x}_i|^{2s-d} \mathrm{d}\,\mu_{j,N}(\mathbf{x}_j)\, \mathrm{d}\,\mu_{i,N}(\mathbf{x}_i)$$
$$= V_{d-2s}(\mathbb{S}^d) - \Bigg[\int_{\mathbb{S}^d} \int_{\mathbb{S}^d} |\mathbf{x} - \mathbf{y}|^{2s-d} \mathrm{d}\,\sigma_d(\mathbf{x})\, \mathrm{d}\,\sigma_d(\mathbf{y})$$
$$- \frac{1}{N^2} \sum_{j=1}^N \int_{D_{j,N}} \int_{D_{j,N}} |\mathbf{x} - \mathbf{y}|^{2s-d} \mathrm{d}\,\mu_{j,N}(\mathbf{x})\, \mathrm{d}\,\mu_{j,N}(\mathbf{y}) \Bigg].$$

Since the first double integral in brackets equals $V_{d-2s}(\mathbb{S}^d)$ (see (26)), we deduce that

$$\mathbb{E}\left[\left\{\operatorname{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\right\}^2\right] = \frac{1}{N^2} \sum_{j=1}^N \int_{D_{j,N}} \int_{D_{j,N}} |\mathbf{x} - \mathbf{y}|^{2s-d} \mathrm{d}\,\mu_{j,N}(\mathbf{x})\, \mathrm{d}\,\mu_{j,N}(\mathbf{y})$$
$$\leq \frac{1}{N^2} \sum_{j=1}^N [\operatorname{diam} D_{j,N}]^{2s-d} \leq \frac{1}{N^2} \sum_{j=1}^N \left[cN^{-1/d}\right]^{2s-d} = c^{2s-d} N^{-2s/d}.$$

The lower bound in (44) follows from Theorem 3.   □

Proof of Theorem 22. Let $d/2 + L < s < d/2 + L + 1$ for an integer $L \geq 1$. Arguing as in the proof of Theorem 21, but using the kernel in (31), we obtain

$$\mathbb{E}\left[\left\{\operatorname{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\right\}^2\right]$$
$$= \frac{1}{N^2} \sum_{j=1}^N \int_{D_{j,N}} \int_{D_{j,N}} [\mathcal{Q}_L(1) - \mathcal{Q}_L(\mathbf{x} \cdot \mathbf{y})] \mathrm{d}\,\mu_{j,N}(\mathbf{x})\, \mathrm{d}\,\mu_{j,N}(\mathbf{y})$$
$$- \frac{1}{N^2} \sum_{j=1}^N \int_{D_{j,N}} \int_{D_{j,N}} (-1)^{L+1} |\mathbf{x} - \mathbf{y}|^{2s-d} \mathrm{d}\,\mu_{j,N}(\mathbf{x})\, \mathrm{d}\,\mu_{j,N}(\mathbf{y}),$$

where we used the fact that the Laplace-Fourier expansion (32) only contains Legendre polynomials $P_\ell^{(d)}$ with $\ell \geq 1$ and thus

$$\frac{1}{N^2} \sum_{j=1}^{N} \sum_{i=1}^{N} \int_{D_{j,N}} \int_{D_{i,N}} \mathcal{Q}_L(\mathbf{x} \cdot \mathbf{y}) \, d\,\mu_{i,N}(\mathbf{x}) \, d\,\mu_{k,N}(\mathbf{y})$$
$$= \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \mathcal{Q}_L(\mathbf{x} \cdot \mathbf{y}) \, d\,\sigma_d(\mathbf{x}) \, d\,\sigma_d(\mathbf{y}) = 0.$$

For $L = 1$, the definition of $\mathcal{Q}_L$ given in (32) and the fact that $P_1^{(d)}(x) = x$ yields

$$\mathcal{Q}_1(1) - \mathcal{Q}_1(\mathbf{x} \cdot \mathbf{y}) = -\alpha_1^{(s)} Z(d,1)(2 - 2\mathbf{x} \cdot \mathbf{y}) = -\alpha_1^{(s)}(d+1) \, |\mathbf{x} - \mathbf{y}|^2 ,$$

where $\alpha_1^{(s)} < 0$ by (26). More generally, using the following hypergeometric function relation for the polynomials $P_\ell^{(d)}(x)$ (see, e.g, [1, Eq. 18.5.9]),

$$P_\ell^{(d)}(x) = {}_2\mathrm{F}_1\left(\begin{matrix} -\ell, \ell+d-1 \\ d/2 \end{matrix}; \frac{1-x}{2}\right),$$

we have the following formula for $\mathcal{Q}_L(\mathbf{x} \cdot \mathbf{y})$ in terms of even powers of distances:

$$\mathcal{Q}_L(1) - \mathcal{Q}_L(\mathbf{x} \cdot \mathbf{y})$$
$$= \mathcal{Q}_L(1) - \sum_{\ell=1}^{L} \sum_{p=0}^{\ell} \left((-1)^{L+1-\ell} - 1\right) \alpha_\ell^{(s)} Z(d,\ell) \frac{(-\ell)_p(\ell+d-1)_p}{(d/2)_p p!} \left(\frac{1-\mathbf{x} \cdot \mathbf{y}}{2}\right)^p$$
$$= -\sum_{p=1}^{\ell} \left\{ \sum_{\ell=p}^{L} \left((-1)^{L+1-\ell} - 1\right) \alpha_\ell^{(s)} Z(d,\ell)(-\ell)_p(\ell+d-1)_p \right\} \frac{(|\mathbf{x} - \mathbf{y}|/2)^{2p}}{(d/2)_p p!}.$$

For small distances the dominant term above is the square of the distance. Since $(-1)^{L+1-\ell} \alpha_\ell^{(s)}$ is positive by (26), it follows that the coefficient of $(|\mathbf{x} - \mathbf{y}|/2)^2$,

$$\beta_1^{(s)} := \frac{2}{d} \sum_{\ell=1}^{L} \left((-1)^{L+1-\ell} - 1\right) \alpha_\ell^{(s)} Z(d,\ell) \left[-(-\ell)_1\right](\ell+d-1)_1,$$

is positive and therefore

$$\mathcal{Q}_L(1) - \mathcal{Q}_L(\mathbf{x} \cdot \mathbf{y}) = \beta_1^{(s)} \left(\frac{|\mathbf{x} - \mathbf{y}|}{2}\right)^2 + \mathcal{O}(|\mathbf{x} - \mathbf{y}|^4) \qquad \text{as } |\mathbf{x} - \mathbf{y}| \to 0.$$

Hence,

(49)
$$\mathbb{E}\left[\left\{\mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d))\right\}^2\right]$$
$$\geq \frac{\beta_1^{(s)}}{4} \frac{1}{N^2} \sum_{j=1}^{N} \int_{D_{j,N}} \int_{D_{j,N}} |\mathbf{x} - \mathbf{y}|^2 \, d\,\mu_{j,N}(\mathbf{x}) \, d\,\mu_{j,N}(\mathbf{y}) - \mathcal{R}_N,$$

where (as $N \to \infty$)

(50)
$$\mathcal{R}_N = \mathcal{O}\left(\frac{1}{N^2} \sum_{j=1}^{N} [\mathrm{diam}\, D_{j,N}]^4\right) + \mathcal{O}\left(\frac{1}{N^2} \sum_{j=1}^{N} [\mathrm{diam}\, D_{j,N}]^{2s-d}\right)$$
$$= \mathcal{O}(N^{-4/d-1}) + \mathcal{O}(N^{-2s/d}).$$

Next, we observe that for any $c'' > 0$ the following inequalities hold:

$$\int_{D_{j,N}} \int_{D_{j,N}} |\mathbf{x} - \mathbf{y}|^2 \, \mathrm{d}\,\mu_{j,N}(\mathbf{x}) \, \mathrm{d}\,\mu_{j,N}(\mathbf{y})$$

$$\geq \int_{\substack{D_{j,N} \\ |\mathbf{x}-\mathbf{y}|>c''/N^{1/d}}} \int_{D_{j,N}} |\mathbf{x} - \mathbf{y}|^2 \, \mathrm{d}\,\mu_{j,N}(\mathbf{x}) \, \mathrm{d}\,\mu_{j,N}(\mathbf{y})$$

(51)

$$\geq (c'')^2 \, N^{-2/d} \int_{\substack{D_{j,N} \\ |\mathbf{x}-\mathbf{y}|>c''/N^{1/d}}} \int_{D_{j,N}} \mathrm{d}\,\mu_{j,N}(\mathbf{x}) \, \mathrm{d}\,\mu_{j,N}(\mathbf{y})$$

$$= (c'')^2 \, N^{-2/d} \left\{ 1 - \int_{\substack{D_{j,N} \\ |\mathbf{x}-\mathbf{y}|\leq c''/N^{1/d}}} \int_{D_{j,N}} \mathrm{d}\,\mu_{j,N}(\mathbf{x}) \, \mathrm{d}\,\mu_{j,N}(\mathbf{y}) \right\}.$$

Since $\mu_{j,N}$ is a probability measure on $D_{j,N}$ and $\sigma_d(D_{j,N}) = 1/N$, we can bound the above double integral by

(52)

$$\int_{\substack{D_{j,N} \\ |\mathbf{x}-\mathbf{y}|<c''/N^{1/d}}} \int_{D_{j,N}} \mathrm{d}\,\mu_{j,N}(\mathbf{x}) \, \mathrm{d}\,\mu_{j,N}(\mathbf{y}) \leq \int_{D_{j,N}} \int_{\mathcal{C}(\mathbf{x};\theta')} \mathrm{d}\,\mu_{j,N}(\mathbf{y}) \, \mathrm{d}\,\mu_{j,N}(\mathbf{x})$$

$$= \mu_{j,N}(D_{j,N}) \, \mu_{j,N}(\mathcal{C}(\mathbf{x};\theta')) = \frac{\sigma_d(\mathcal{C}(\mathbf{x};\theta'))}{\sigma_d(D_{j,N})},$$

where $2\sin(\theta'/2) = c''/N^{1/d}$. The Funk-Hecke formula gives (cf., e.g., [29])

$$\sigma_d(\mathcal{C}(\mathbf{x};\theta')) = \frac{1}{d} \frac{\omega_{d-1}}{\omega_d} \left[2\sin(\theta'/2) \right]^d \left\{ 1 + \mathcal{O}([2\sin(\theta'/2)]^2) \right\} \quad \text{as } \theta' \to 0.$$

Hence, for N sufficiently large,

$$\frac{\sigma_d(\mathcal{C}(\mathbf{x};\theta')}{\sigma_d(D_{j,N})} = N \frac{1}{d} \frac{\omega_{d-1}}{\omega_d} \left(c''/N^{1/d} \right)^d \left\{ 1 + \mathcal{O}\left(\left(c''/N^{1/d} \right)^2 \right) \right\} \leq 2 \frac{1}{d} \frac{\omega_{d-1}}{\omega_d} (c'')^d.$$

By fixing c'' (which now depends only on d) to be sufficiently small, we can always achieve that the double integral on the left-hand side of (52) is bounded from above by $1/2$. Therefore, for sufficiently large N and $j = 1, \dots, N$, we deduce from (51) that

$$\int_{D_{j,N}} \int_{D_{j,N}} |\mathbf{x} - \mathbf{y}|^2 \, \mathrm{d}\,\mu_{j,N}(\mathbf{x}) \, \mathrm{d}\,\mu_{j,N}(\mathbf{y}) \geq \frac{(c'')^2}{2} \, N^{-2/d}.$$

Combining this estimate with (49) and (50), we obtain for $s > d/2 + 1$ and $2s - d$ a positive integer that is not even,

$$\mathbb{E}\left[\left\{ \mathrm{wce}(Q[X_N]; \mathbb{H}^s(\mathbb{S}^d)) \right\}^2 \right] \geq \frac{\beta_1^{(s)}}{4} \frac{(c'')^2}{2} N^{-2/d-1} + \mathcal{O}(N^{-4/d-1}) + \mathcal{O}(N^{-2s/d})$$

$$\geq \beta \, N^{-2/d-1} = \beta \, N^{-2(d/2+1)/d},$$

where the positive constant β depends on the $\mathbb{H}^s(\mathbb{S}^d)$-norm and the partition sequence $(\mathcal{D}_N)$, but is independent of N. $\qquad\square$

References

[1] Digital Library of Mathematical Functions. 2010-05-07. National Institute of Standards and Technology from http://dlmf.nist.gov/

[2] C. Aistleitner, J. S. Brauchart, and J. Dick, *Point sets on the sphere* $\mathbb{S}^2$ *with small spherical cap discrepancy*, Discrete Comput. Geom. **48** (2012), no. 4, 990–1024, DOI 10.1007/s00454-012-9451-3. MR3000572

[3] Clemens Amstler and Peter Zinterhof, *Uniform distribution, discrepancy, and reproducing kernel Hilbert spaces*, J. Complexity **17** (2001), no. 3, 497–515, DOI 10.1006/jcom.2001.0580. MR1851057 (2002g:65025)

[4] Diego Armentano, Carlos Beltrán, and Michael Shub, *Minimizing the discrete logarithmic energy on the sphere: the role of random polynomials*, Trans. Amer. Math. Soc. **363** (2011), no. 6, 2955–2965, DOI 10.1090/S0002-9947-2011-05243-8. MR2775794 (2012f:31009)

[5] E. Bannai and R. M. Damerell, *Tight spherical designs. I*, J. Math. Soc. Japan **31** (1979), no. 1, 199–207, DOI 10.2969/jmsj/03110199. MR519045 (80b:05014)

[6] E. Bannai and R. M. Damerell, *Tight spherical designs. II*, J. London Math. Soc. (2) **21** (1980), no. 1, 13–30, DOI 10.1112/jlms/s2-21.1.13. MR576179 (81g:05024)

[7] R. Bauer. *Distribution of points on a sphere with application to star catalogs*, J. Guidance Control Dynamics, 23(1):130–137, 2000.

[8] József Beck, *Sums of distances between points on a sphere—an application of the theory of irregularities of distribution to discrete geometry*, Mathematika **31** (1984), no. 1, 33–41, DOI 10.1112/S0025579300010639. MR762175 (86d:52004)

[9] József Beck and William W. L. Chen, *Irregularities of distribution*, Cambridge Tracts in Mathematics, vol. 89, Cambridge University Press, Cambridge, 1987. MR903025 (88m:11061)

[10] Göran Björck, *Distributions of positive mass, which maximize a certain generalized energy integral*, Ark. Mat. **3** (1956), 255–269. MR0078470 (17,1198b)

[11] A. Bondarenko, D. Radchenko, and M. Viazovska, *Optimal asymptotic bounds for spherical designs*, Ann. of Math. (2) **178** (2013), 443–452. MR3071504

[12] J. Bourgain and J. Lindenstrauss, *Distribution of points on spheres and approximation by zonotopes*, Israel J. Math. **64** (1988), no. 1, 25–31, DOI 10.1007/BF02767366. MR981745 (90g:46021)

[13] L. Brandolini, C. Choirat, L. Colzani, G. Gigante, R. Seri, and G. Travaglini, *Quadrature rules and distribution of points on manifolds*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (to appear), arXiv:1012.5409v1 [math.NT], Dec 2010.

[14] Johann S. Brauchart and Josef Dick, *Quasi-Monte Carlo rules for numerical integration over the unit sphere* $\mathbb{S}^2$, Numer. Math. **121** (2012), no. 3, 473–502, DOI 10.1007/s00211-011-0444-6. MR2929076

[15] Johann S. Brauchart and Josef Dick, *A simple proof of Stolarsky's invariance principle*, Proc. Amer. Math. Soc. **141** (2013), no. 6, 2085–2096, DOI 10.1090/S0002-9939-2013-11490-5. MR3034434

[16] Johann S. Brauchart and Kerstin Hesse, *Numerical integration over spheres of arbitrary dimension*, Constr. Approx. **25** (2007), no. 1, 41–71, DOI 10.1007/s00365-006-0629-4. MR2263736 (2007f:41025)

[17] J. S. Brauchart and R. S. Womersley, *Weighted QMC designs: Numerical integration over the unit sphere*, $\mathbb{L}_2$-discrepancy and sum of distances. In preparation.

[18] Xiaojun Chen, Andreas Frommer, and Bruno Lang, *Computational existence proofs for spherical t-designs*, Numer. Math. **117** (2011), no. 2, 289–305, DOI 10.1007/s00211-010-0332-5. MR2754852 (2012h:65048)

[19] Jianjun Cui and Willi Freeden, *Equidistribution on the sphere*, SIAM J. Sci. Comput. **18** (1997), no. 2, 595–609, DOI 10.1137/S1064827595281344. MR1433797 (97k:65024)

[20] P. Delsarte, J. M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geometriae Dedicata **6** (1977), no. 3, 363–388. MR0485471 (58 #5302)

[21] Michael Drmota and Robert F. Tichy, *Sequences, discrepancies and applications*, Lecture Notes in Mathematics, vol. 1651, Springer-Verlag, Berlin, 1997. MR1470456 (98j:11057)

[22] Kerstin Hesse, *A lower bound for the worst-case cubature error on spheres of arbitrary dimension*, Numer. Math. **103** (2006), no. 3, 413–433, DOI 10.1007/s00211-006-0686-x. MR2221056 (2007b:41050)

[23] K. Hesse, H. N. Mhaskar, and I. H. Sloan, *Quadrature in Besov spaces on the Euclidean sphere*, J. Complexity **23** (2007), no. 4-6, 528–552, DOI 10.1016/j.jco.2006.10.004. MR2372012 (2008m:65069)

[24] Kerstin Hesse and Ian H. Sloan, *Optimal lower bounds for cubature error on the sphere S^2*, J. Complexity **21** (2005), no. 6, 790–803, DOI 10.1016/j.jco.2005.07.004. MR2182445 (2006m:65047)

[25] Kerstin Hesse and Ian H. Sloan, *Worst-case errors in a Sobolev space setting for cubature over the sphere S^2*, Bull. Austral. Math. Soc. **71** (2005), no. 1, 81–105, DOI 10.1017/S0004972700038041. MR2127668 (2005k:41090)

[26] Kerstin Hesse and Ian H. Sloan, *Cubature over the sphere S^2 in Sobolev spaces of arbitrary order*, J. Approx. Theory **141** (2006), no. 2, 118–133, DOI 10.1016/j.jat.2006.01.004. MR2252093 (2007k:41073)

[27] Brad J. C. Baxter and Simon Hubbert, *Radial basis functions for the sphere*, Recent progress in multivariate approximation (Witten-Bommerholz, 2000), Internat. Ser. Numer. Math., vol. 137, Birkhäuser, Basel, 2001, pp. 33–47. MR1877496 (2002i:41024)

[28] J. Korevaar and J. L. H. Meyers, *Spherical Faraday cage for the case of equal point charges and Chebyshev-type quadrature on the sphere*, Integral Transform. Spec. Funct. **1** (1993), no. 2, 105–117, DOI 10.1080/10652469308819013. MR1421438 (97g:41046)

[29] A. B. J. Kuijlaars and E. B. Saff, *Asymptotics for minimal discrete energy on the sphere*, Trans. Amer. Math. Soc. **350** (1998), no. 2, 523–538, DOI 10.1090/S0002-9947-98-02119-9. MR1458327 (98e:11092)

[30] Paul Leopardi, *Diameter bounds for equal area partitions of the unit sphere*, Electron. Trans. Numer. Anal. **35** (2009), 1–16. MR2582801 (2010m:51017)

[31] Claus Müller, *Spherical harmonics*, Lecture Notes in Mathematics, vol. 17, Springer-Verlag, Berlin, 1966. MR0199449 (33 #7593)

[32] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou, *Minimal discrete energy on the sphere*, Math. Res. Lett. **1** (1994), no. 6, 647–662. MR1306011 (96e:78011)

[33] Robert J. Renka, *Multivariate interpolation of large sets of scattered data*, ACM Trans. Math. Software **14** (1988), no. 2, 139–148, DOI 10.1145/45054.45055. MR946761 (89d:65009)

[34] I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–108. MR0005922 (3,232c)

[35] P. D. Seymour and Thomas Zaslavsky, *Averaging sets: a generalization of mean values and spherical designs*, Adv. in Math. **52** (1984), no. 3, 213–240, DOI 10.1016/0001-8708(84)90022-7. MR744857 (85m:05031)

[36] Ian H. Sloan and Robert S. Womersley, *Extremal systems of points and numerical integration on the sphere*, Adv. Comput. Math. **21** (2004), no. 1-2, 107–125, DOI 10.1023/B:ACOM.0000016428.25905.da. MR2065291 (2005b:65024)

[37] Ian H. Sloan and Robert S. Womersley, *A variational characterisation of spherical designs*, J. Approx. Theory **159** (2009), no. 2, 308–318, DOI 10.1016/j.jat.2009.02.014. MR2562747 (2011a:05063)

[38] Kenneth B. Stolarsky, *Sums of distances between points on a sphere. II*, Proc. Amer. Math. Soc. **41** (1973), 575–582. MR0333995 (48 #12314)

[39] N. Th. Varopoulos, *Small time Gaussian estimates of heat diffusion kernels. II. The theory of large deviations*, J. Funct. Anal. **93** (1990), no. 1, 1–33, DOI 10.1016/0022-1236(90)90136-9. MR1070036 (91j:58156)

[40] Gerold Wagner, *On means of distances on the surface of a sphere. II. Upper bounds*, Pacific J. Math. **154** (1992), no. 2, 381–396. MR1159518 (93b:52007)

[41] R. S. Womersley, *Numerical integration, approximation and point sets on the sphere.* [http://www.maths.unsw.edu.au/~rsw/Sphere/; accessed August-2012]

School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, 2052, Australia

E-mail address: `j.brauchart@unsw.edu.au`

Center for Constructive Approximation, Department of Mathematics, Vanderbilt University, Nashville, Tennessee 37240

E-mail address: `edward.b.saff@vanderbilt.edu`

School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, 2052, Australia

E-mail address: `i.sloan@unsw.edu.au`

School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, 2052, Australia

E-mail address: `r.womersley@unsw.edu.au`