



# Spherical Cap Discrepancy of Perturbed Lattices Under the Lambert Projection

Damir Ferizović<sup>1</sup>

Received: 28 February 2022 / Revised: 26 January 2023 / Accepted: 20 February 2023 /  
Published online: 27 August 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

For any given full rank lattice  $\Lambda$  and natural number  $K$ , we consider the point set  $P(K) = \Lambda/K \cap (0, 1)^2$ , with  $N = \#P(K) \approx K^2$ , and bound the spherical cap discrepancy of the projection of these points under the Lambert map to the unit sphere. The bound is of order  $1/\sqrt{N}$ , with leading coefficient given explicitly and depending on  $\Lambda$  only. The proof is established using a lemma that bounds the number of intersections of certain curves with fundamental domains that tile  $\mathbb{R}^2$ , and even allows for local perturbations of  $\Lambda$  without affecting the bound, proving to be suitable for numerical applications. A special case yields the smallest constant for the leading term of the cap discrepancy for deterministic algorithms up to date.

**Keywords** Spherical cap discrepancy · Two sphere · Lambert map · Lattices

**Mathematics Subject Classification** 11K38 · 52C99 · 52A10

## 1 Introduction and Main Result

The problem of distributing  $N$ -many points  $P_N = \{p_1, \dots, p_N\}$  uniformly on a sphere is well known and has applications in numerical integration, approximation, cartography and the applied sciences, see [13, 25, 27] for some reviews surveying over a century of development.

---

Editor in Charge: János Pach

---

The author thankfully acknowledges support by the Methusalem grant METH/21/03—long term structural funding of the Flemish Government.

---

Damir Ferizović  
damir.ferizovic@kuleuven.be

<sup>1</sup> Department of Mathematics, KU Leuven, Celestijnenlaan 200b, Box 2400, 3001 Leuven, Belgium

There are several different notions of uniformity of distribution for a finite point set  $P_N$ : one example is the energy, which is the sum of all pairwise interactions of points under a given lower semi-continuous potential  $F$ . When  $F(x, y) = \|x - y\|^{-s}$  one refers to the corresponding energy as “Riesz  $s$ -energy” and minimizers of this potential have been intensively investigated, see for instance the comprehensive monograph by Borodachov et al. [12]. Another notion of uniformity is that of separation distance, i.e., the minimum of the pairwise distances of distinct elements among  $P_N$ . Our main result uses the notion of spherical cap discrepancy.

**Definition 1.1** Let  $P_N = \{p_1, \dots, p_N\} \subset \mathbb{S}^2$ . The spherical cap discrepancy is

$$\mathcal{D}(P_N) = \sup_{w \in \mathbb{S}^2} \sup_{-1 \leq t \leq 1} \left| \frac{1}{N} \sum_{j=1}^N \chi_{C(w,t)}(p_j) - \sigma(C(w,t)) \right|.$$

Here  $\mathbb{S}^2 \subset \mathbb{R}^3$  is the unit two-sphere, i.e., the set of all unit vectors with norm induced by the standard dot product  $\langle \cdot, \cdot \rangle$ . A spherical cap with center  $w \in \mathbb{S}^2$  and height  $t \in (-1, 1)$  is the set

$$C(w, t) = \{x \in \mathbb{S}^2 \mid \langle w, x \rangle \geq t\}.$$

The boundary of a set  $S$  is denoted by  $\partial S$ , thus

$$\partial C(w, t) = \{x \in \mathbb{S}^2 \mid \langle w, x \rangle = t\},$$

and  $\sigma$  is the normalized surface measure of  $\mathbb{S}^2$ , such that  $\sigma(C(w, t)) = (1 - t)/2$ . The characteristic function of a set  $A$  is  $\chi_A$ , i.e.,  $\chi_A(x) = 1$  iff  $x \in A$ , and 0 otherwise. An application of discrepancy to high-dimensional integration problems is the so-called “Koksma–Hlawka” inequality:

$$\left| \frac{1}{N} \sum_{j=1}^N f(p_j) - \int_{\mathbb{S}^2} f \, d\sigma \right| \leq \hat{\mathcal{D}}(P_N) \mathcal{V}(f),$$

where  $\hat{\mathcal{D}}$  is a notion of discrepancy (not restricted to the spherical cap discrepancy) and  $\mathcal{V}(f)$  is a constant depending on the function  $f$ , see [15, 20].

This paper deals with point sets on the two-sphere which are derived from perturbations of any full rank lattice under a specific equal area map. This adds to the list of known constructions with good distribution properties. For the reader’s convenience, we give below a list of other well-known point sets on  $\mathbb{S}^2$ , including their reference of origin and other works investigating them.

	Origin	Invstigd.
Probabilistic algorithms		
Uniformly at random	[16]	[1, 14]
Ensembles based on determinantal point processes	[22]	[2, 9]
Roots of random polynomials in $\mathbb{R}^2 \rightarrow \mathbb{S}^2$	[11]	[3]
The diamond ensemble (can also be made deterministic)	[8]	[8, 18]
Point sets based on jittered sampling	[5, 7]	[6]
Deterministic algorithms		
Spiral points	[4, 24]	[21]
Hierarchical, equal area and iso-latitude pixelation	[19]	[21]
Point sets using group action	[23]	[23]
Spherical Fibonacci lattice	[1]	[1]
Spherical Fibonacci grid	[26]	[21]

We will prove a bound on the spherical cap discrepancy of our point constructions which is of the same order as the best results obtained so far for deterministic point sets found in the literature. An open (and hard) problem is to improve these bounds to the order obtained in a fundamental result due to Beck in [5, 6], where we find the existence of points  $P_N^*$  and constants independent of  $N$ , such that

$$cN^{-3/4} \leq \mathcal{D}(P_N^*) \leq CN^{-3/4} \sqrt{\log N}.$$

The construction of Beck is probabilistic and an algorithm to generate random point sets with discrepancy matching this upper bound with high probability was obtained for instance in [2].

Bounds for deterministic point sets are usually hard to derive and were first achieved by Lubotzky et al. in [23] with an upper bound of the order  $(\log N)^{2/3} N^{-1/3}$ . Aistleitner et al. showed in [1] that for spherical digital nets and spherical Fibonacci lattices the spherical cap discrepancy is upper bounded by an order of  $N^{-1/2}$ . Further bounds were given for the diamond ensemble by Etayo in [18] and for HEALPix generated points by Hofstadler, Mastrianni, and the author in a preprint, where both point sets are shown to have a spherical cap discrepancy of order  $N^{-1/2}$ .

Etayo opened the contest to find deterministic point sets  $P_N$  which allow for the smallest constant bounding  $\sqrt{N} \mathcal{D}(P_N)$ , where she took the lead in [18] with the Diamond ensemble and a bound of  $4 + 2\sqrt{2}$ . We will identify a family of point sets where  $\sqrt{18}$  suffices, see Sect. 3.1.

We construct point sets based on the method deployed in [1], which originated in [17]. Let  $\text{GL}(\mathbb{R}^2)$  be the group of invertible  $2 \times 2$  matrices acting on  $\mathbb{R}^2$  with identity matrix  $\mathbb{1}$ . The Frobenius norm of the matrix  $Q$  is  $\|Q\|_F = \sqrt{\text{trace } Q^t Q}$ , the Lebesgue measure on  $\mathbb{R}^2$  is  $\lambda$ , and  $e_1, e_2$  denote the standard basis of  $\mathbb{R}^2$ .

**Definition 1.2** A lattice  $\Lambda^Q$  is a set  $Q\mathbb{Z}^2$  for some  $Q \in \text{GL}(\mathbb{R}^2)$ . The associated scaled tiling via a fundamental domain  $\Omega^Q = Q[0, 1)^2 \subset \mathbb{R}^2$  and  $K \in \mathbb{N}$  is

$$\mathbb{R}^2 = \bigcup_{p \in \Lambda^Q} T_K(p), \quad \text{where } T_K(p) = \frac{1}{K}(p + \Omega^Q).$$

Let  $I^2 = [0, 1) \times (0, 1)$  denote the unit square with three boundary sides removed, then for fixed  $K \in \mathbb{N}$  we choose arbitrary  $z_K^p \in T_K(p)$  for  $p \in \Lambda^Q$  and define a point set

$$P^Q(K) = \{z_K^p \mid p \in \Lambda^Q\} \cap I^2.$$

See Fig. 1b. The Lambert map  $L : I^2 \rightarrow \mathbb{S}^2$ , which we introduce later, identifies two sides of  $\partial I^2$  and maps two sides to the poles—and this identification has an effect on the discrepancy, see Fig. 1. To bound it we define a parametrization  $\rho : [0, 4] \rightarrow \partial I^2$  via  $f(x) := \max(\min(x, 1) + \min(2 - x, 0), 0)$ ,

$$x \mapsto \rho(x) = f(x)e_1 + f(4 - x)e_2. \tag{1}$$

Before we state our main result in Theorem 1.3, we need to introduce three terms, which are bounded above by constants depending on  $Q$  only, as we will show:

1) Let  $N^Q(K) = \#P^Q(K)$  be the number of points in  $P^Q(K)$ , and set

$$d^Q(K) = \frac{1}{K} \left| N^Q(K) - \frac{K^2}{|\det(Q)|} \right|.$$

2) Set

$$C_L^Q = \sup_{w \in \mathbb{S}^2} \sup_{-1 \leq t \leq 1} \text{length} \left( Q^{-1}(L^{-1}(\partial C(w, t) \cap \mathbb{S}_p^2)) \right),$$

3) and with the short hand notation  $[a, b]^c := [0, a] \cup [b, 1]$  we define

$$M^Q(K) = \sup_{\substack{S \in \{[a, b], [a, b]^c\} \\ 0 \leq a < b \leq 1}} \frac{K}{|\det(Q)|} \left| \sum_{p \in \text{Ad}(S, K)} \left( \frac{\chi_{P^Q(K)}(z_K^p)}{N^Q(K)} - \lambda(T_K(p) \cap I^2) \right) \right|,$$

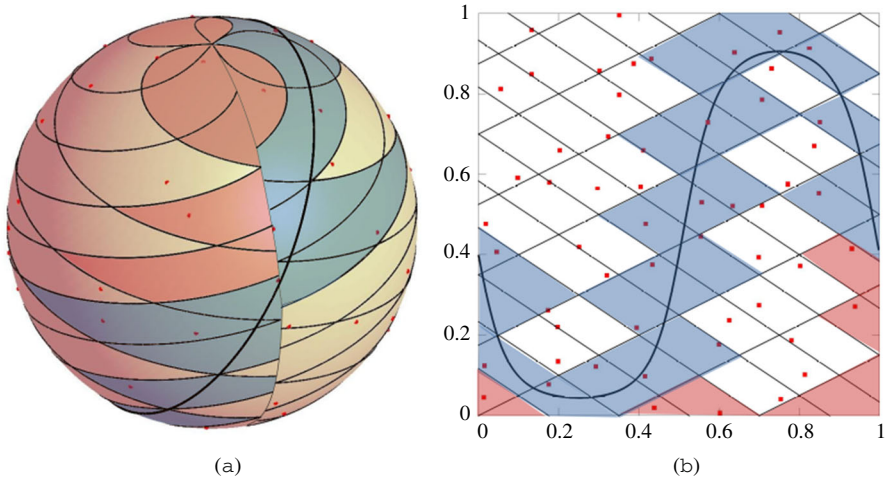
where we sum over all  $p$ , such that the interior of  $T_K(p)$  intersects  $\partial I^2$  at some “height” in  $S$ : We define  $R(S) = \{x \in [0, 4] \mid f(4 - x) \in S\}$  for  $S \subset [0, 1]$  and  $f$  as in (1), to be all  $x$  with  $\langle \rho(x), e_2 \rangle \in S$ ; then

$$\text{Ad}(S, K) = \{p \in \Lambda^Q \mid T_K(p)^\circ \cap \rho(R(S)) \neq \emptyset\}.$$

The term  $M^Q(K)$  bounds the discrepancy that arises around  $L(\partial I^2)$ , while taking the geometry of images of spherical caps under  $L^{-1}$  into account.

**Theorem 1.3** (main result) *Let  $Q \in \text{GL}(\mathbb{R}^2)$ ,  $K \in \mathbb{N}$ , and  $P^Q(K)$  be chosen as in Definition 1.2 with  $N = N^Q(K)$ . Then*

$$\mathcal{D}(L(P^Q(K))) \leq (d^Q(K) + \sqrt{2} \cdot C_L^Q + M^Q(K)) \frac{\sqrt{|\det(Q)|}}{\sqrt{N}} + O(N^{-1}).$$



**Fig. 1** Picture a is the image under the Lambert map of picture b. The dots represent  $P^Q(K)$ , the S-shaped curve is  $L^{-1}(\partial C)$  for a spherical cap  $C$  with the north pole in its interior, see Fig. 2. Contributing factors to the spherical cap discrepancy are fluctuations: of the total number of points, of points in  $T_K(p)$  intersected by  $\partial C$  (shaded in blue) and of points in  $T_K(p)$  intersected by  $\partial I^2$  (shaded in red), which are represented by  $d^Q(K)$ ,  $C^Q(K)$ , and  $M^Q(K)$

**Remark** A possible choice for  $P^Q(K)$  is  $(K^{-1} \Lambda^Q) \cap I^2$ , and Theorem 1.3 shows that small inaccuracies in numerical representation of the positions of these elements do not affect the bound on discrepancy.

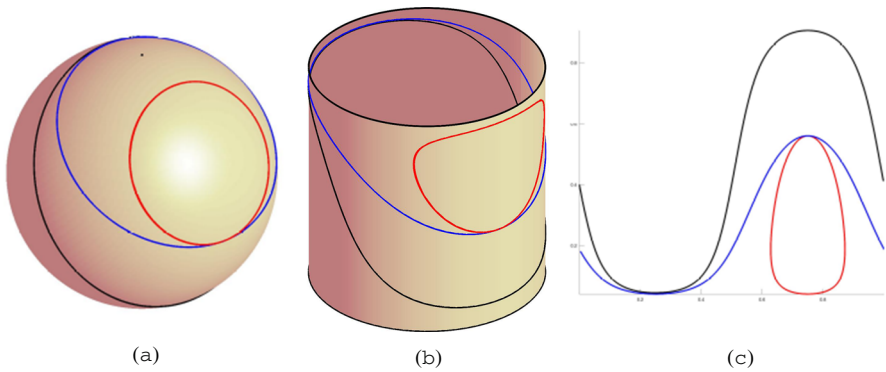
A less precise but more succinct bound is given by the following:

**Corollary 1.4** Let  $Q \in GL(\mathbb{R}^2)$ ,  $K \in \mathbb{N}$ , and  $P^Q(K)$  be chosen as in Definition 1.2 with  $N = N^Q(K)$ . Then

$$\mathcal{D}(L(P^Q(K))) \leq \frac{\|Q\|_F}{\sqrt{|\det(Q)|}} \cdot \frac{8 + 3\sqrt{2}}{\sqrt{N}} + O(N^{-1}).$$

**Proof** This follows from Theorem 1.3 with Lemmas 2.2, 2.3, and 2.8, and the remark after it. □

**Outline of the paper.** In Sect. 2 we first state the definition of the Lambert equal area map and prove a bound on  $C_L^1$ . Next we introduce the notion of an  $n$ -convex curve, and then prove in Lemma 2.7 a bound on how many fundamental domains of  $(1/K)\Lambda^Q$  are intersected by it; this result constitutes the backbone of this work. The section then ends with a bound on  $d^Q(K)$  and a proof of Theorem 1.3. In Sect. 3 we apply Theorem 1.3 to specific cases, thus obtaining many deterministic point sets with the least constant for the leading term up to date.



**Fig. 2** Illustration of the Lambert map. Boundaries of spherical caps on a sphere in a, in b we see their projection to the enveloping cylindrical surface, which is “cut”, “flipped”, and scaled to the unit square in c

## 2 Intermediate Results and Proof of Theorem 1.3

### 2.1 The Lambert Map

By  $\mathbb{S}_p^2$  we denote the punctured unit sphere, i.e.,  $\mathbb{S}^2$  with the poles removed.

**Definition 2.1** The Lambert map  $L$  is a well-known area preserving map, i.e., for open sets  $U \subset I^2$  and the Lebesgue measure  $\lambda$ , we have  $\lambda(U) = \sigma(L(U))$ , where

$$L: I^2 \rightarrow \mathbb{S}_p^2,$$

$$(x, y) \mapsto \begin{pmatrix} 2\sqrt{y - y^2} \cos 2\pi x \\ 2\sqrt{y - y^2} \sin 2\pi x \\ 1 - 2y \end{pmatrix},$$

and the inverse is given in terms of the standard parametrization<sup>1</sup> of  $\mathbb{S}^2$ ,

$$L^{-1}: \mathbb{S}_p^2 \rightarrow I^2,$$

$$\begin{pmatrix} \cos \phi \sin \theta \\ \sin \phi \sin \theta \\ \cos \theta \end{pmatrix} \mapsto \frac{1}{2\pi} \begin{pmatrix} \phi \\ \pi(1 - \cos \theta) \end{pmatrix}. \tag{2}$$

**Lemma 2.2**  $C_L^1 \leq 3$ .

**Proof** Note that the map  $L^{-1}$  projects points from the sphere, parallel to the plane  $\mathbb{R}^2 \times \{0\}$ , to the enveloping cylindrical surface, see Fig. 2. In order to bound the length of the curve  $L^{-1}(\partial C \cap \mathbb{S}_p^2)$ , it will be enough to work with the projection on the enveloping cylinder. We make the following claim, that we consider obviously true, see Fig. 2:

<sup>1</sup> Here  $0 \leq \phi < 2\pi$  and  $0 < \theta < \pi$ .

Given any spherical cap  $C(w, t)$ , then there exists a cap  $C(w', t')$  and  $\varepsilon > 0$ , such that

$$1 - \varepsilon = z_M(w', t') > z_m(w', t') = \varepsilon - 1,$$

where  $z_M(w', t')$  and  $z_m(w', t')$  denote the maximal and minimal  $z$ -value of  $C(w', t')$  in cylindrical coordinates, and

$$C(w', t') \text{ contains exactly one of the poles in its interior,}$$

with the property that

$$\text{length}(\partial C(w, t)) \leq \text{length}(\partial C(w', t')).$$

This inequality then also holds for the length of the projections  $L^{-1}(\partial C(w, t) \cap \mathbb{S}_p^2)$  and  $L^{-1}(\partial C(w', t') \cap \mathbb{S}_p^2)$ .

Thus in order to maximize the length we let  $z_M, z_m$  approach  $\pm 1$  and choose  $w$  so that exactly one of the poles is contained in the cap, i.e., consider  $C(w, 0)$  with

$$w = \left( \sin \left( \frac{\pi}{2} - \varepsilon \right), 0, \cos \left( \frac{\pi}{2} - \varepsilon \right) \right)^t$$

and  $\varepsilon > 0$  small. A parametrization for  $\partial C(w, 0)$  is then given by

$$\begin{aligned} \varphi_{\pm}: [\varepsilon, \pi - \varepsilon] &\rightarrow [0, 2\pi], \\ \theta &\mapsto \pm \arccos \left( -\cot \theta \cot \left( \frac{\pi}{2} - \varepsilon \right) \right), \end{aligned}$$

such that

$$\begin{aligned} \gamma_+ : [\varepsilon, \pi - \varepsilon] &\rightarrow \partial C(w, 0), \\ \theta &\mapsto \begin{pmatrix} \cos \varphi_+(\theta) \sin \theta \\ \sin \varphi_+(\theta) \sin \theta \\ \cos \theta \end{pmatrix}, \end{aligned}$$

defines a curve that runs from  $z_M$  to  $z_m$  along  $\partial C(w, 0)$ , and  $\gamma_-$  defines the other half. Since  $L^{-1} \circ \gamma_+$  and  $L^{-1} \circ \gamma_-$  have the same length, it is enough to consider just one of these curves, bound its length and multiply by 2: thus we consider the curve

$$\theta \mapsto \frac{1}{\pi} \left( \arccos \left( -\cot \theta \cot \left( \frac{\pi}{2} - \varepsilon \right) \right) \right),$$

where we already multiplied by a factor of 2 in (2). Thus we have to bound

$$\text{length} = \int_{\epsilon}^{\pi-\epsilon} \sqrt{\frac{1}{1 - (\cot \theta)^2 (\cot (\pi/2 - \epsilon))^2} \cdot \frac{(\cot (\pi/2 - \epsilon))^2}{\pi^2 (\sin \theta)^4} + (\sin \theta)^2} \, d\theta.$$

We can simplify the expression in the square root, first with the angle sum formula for the cosine (with  $\epsilon < \theta < \pi - \epsilon$ ) to obtain

$$\frac{-1}{\cos (\pi/2 - \epsilon - \theta) \cos (\pi/2 - \epsilon + \theta)} \cdot \frac{(\cos (\pi/2 - \epsilon))^2}{\pi^2 (\sin \theta)^2} + (\sin \theta)^2,$$

and then with  $\cos (\epsilon - \pi/2) = \sin \epsilon$  and a symmetry argument to get

$$\text{length} = 2 \int_{\epsilon}^{\pi/2} \sqrt{\frac{1}{\sin (\epsilon + \theta) \sin (\theta - \epsilon)} \cdot \frac{(\sin \epsilon)^2}{\pi^2 (\sin \theta)^2} + (\sin \theta)^2} \, d\theta.$$

Let

$$\text{sinc } x = \frac{\sin x}{x},$$

then for any choice of  $a \in (0, \pi/2]$  we have

$$x \text{ sinc } a \leq \sin x \leq x$$

for  $x \in [0, a]$ , which we use for small  $\epsilon$  and  $0 < 2\epsilon < a < \pi/2$  to obtain

$$\begin{aligned} \text{length} &\leq \frac{2}{(\text{sinc } a)^2 \pi} \int_{\epsilon}^{a-\epsilon} \sqrt{\frac{1}{\theta^2 - \epsilon^2} \cdot \frac{\epsilon^2}{\theta^2}} \, d\theta + \frac{\epsilon \cdot \pi^2}{4\sqrt{a(a-2\epsilon)^3}} + 2 \int_{\epsilon}^{\pi/2} \sin \theta \, d\theta \\ &= \frac{2}{\pi (\text{sinc } a)^2} \int_1^{a/\epsilon-1} \sqrt{\frac{1}{x^2 - 1}} \cdot \frac{1}{x} \, dx + O(\epsilon) + 2 \stackrel{\lim_{\epsilon \rightarrow 0}}{=} \frac{2}{\pi (\text{sinc } a)^2} \cdot \frac{\pi}{2} + 2, \end{aligned}$$

where we used that the length increases with decreasing  $\epsilon$ , and that

$$\frac{d}{dx} \arctan \sqrt{x^2 - 1} = \sqrt{\frac{1}{x^2 - 1}} \cdot \frac{1}{x}.$$

The inequality above is thus true for any  $0 < a$ , and hence also for  $\text{sinc } 0 = 1$ , which proves the claim. □

The next lemma relates  $C_L^0$  to  $C_L^1$ .



**Lemma 2.3** Given a matrix  $A \in \mathbb{R}^{2 \times 2}$  and a finite length  $C^1$ -curve  $\beta$  in  $\mathbb{R}^2$ , then for  $\ell_\beta = \text{length}(\beta)$ ,  $\ell_{A\beta} = \text{length}(A\beta)$ ,  $m_A = \min \{\|Ae_1\|, \|Ae_2\|\}$ , and  $M_A = \max \{\|Ae_1\|, \|Ae_2\|\}$ , we have

$$\begin{aligned} m_A \cdot \ell_\beta &\leq \ell_{A\beta} \leq M_A \cdot \ell_\beta && \text{if } A \text{ has rank 2 and } A \text{ is orthogonal,} \\ \ell_{A\beta} &\leq \|A\|_F \cdot \ell_\beta && \text{if } A \text{ has rank 2.} \end{aligned}$$

**Proof** Let  $\beta(t) = (x(t), y(t))$  be an injective  $C^1$ -parametrization with  $0 < t < 1$  (thus differing from the geometric curve at most by two endpoints). Let  $s_M = \max_{\|x\|=1} \|Ax\|$  and  $s_m = \min_{\|x\|=1} \|Ax\|$  denote the maximal and minimal singular values of  $A$ , then we have

$$s_m \ell_\beta = s_m \int_0^1 \|\beta'(t)\| dt \leq \int_0^1 \|A\beta'(t)\| dt \leq s_M \int_0^1 \|\beta'(t)\| dt = s_M \ell_\beta.$$

For an orthogonal matrix we have  $s_m = m_A$  and similarly for  $M_A$ ; also it is clearly true that  $s_M \leq \|A\|_F$ .  $\square$

## 2.2 Intersection of Convex Curves with Fundamental Domains

In this section we will introduce the notion of an  $n$ -convex curve, for which we can prove sharp bounds on the number of fundamental domains it intersects. This will then give the boundary contribution in the proof of Theorem 1.3.

**Definition 2.4** A continuous curve  $\beta: [a, b] \rightarrow \mathbb{R}^2$  ( $a, b \in \mathbb{R}$ ,  $a < b$ ) is  $n$ -convex for  $n \in \mathbb{N}$ , if there are points  $a = t_0 < t_1 < \dots < t_{n-1} < t_n = b$  with the property that  $\beta([t_j, t_{j+1}])$  is a convex curve for each  $0 \leq j < n$ , i.e., there exist convex sets  $A_1, \dots, A_n$  with  $\beta([t_j, t_{j+1}]) \subset \partial A_j$ .

A circle is 1-convex, as is a straight line segment. Finite spirals (in length and winding number) are  $n$ -convex. It seems that every (finite) connected part of the boundary of a pseudo-convex set, as introduced in [1], is  $n$ -convex and vice versa. This characterization is not necessary for our purposes, so we do not try to prove it here.

**Corollary 2.5** The images  $L^{-1}(\partial C(w, t))$  for all  $(w, t) \in \mathbb{S}^2 \times (-1, 1)$  are at most three distinct curves, each is at most 7-convex, and they are not self-intersecting.

**Proof** This follows from the analysis of [1, Sect. 6], where the authors show that there is an admissible covering by pseudo-convex sets of at most seven parts (each part has some arc of the curve contained in its boundary).  $\square$

**Definition 2.6** Given a lattice  $\Lambda^Q$ , a continuous curve  $\beta: [a, b] \rightarrow \mathbb{R}^2$  and  $K \in \mathbb{N}$ . Let  $\Omega^Q = Q[0, 1]^2$ , then the intersection number of  $\beta$  with the tiling  $(1/K)(\Lambda^Q + \Omega^Q)$  is

$$I_\beta^Q(K) = \# \left\{ p \in \Lambda^Q \mid \left( \frac{1}{K} \Omega^Q + \frac{1}{K} p \right) \cap \beta([a, b]) \neq \emptyset \right\}.$$

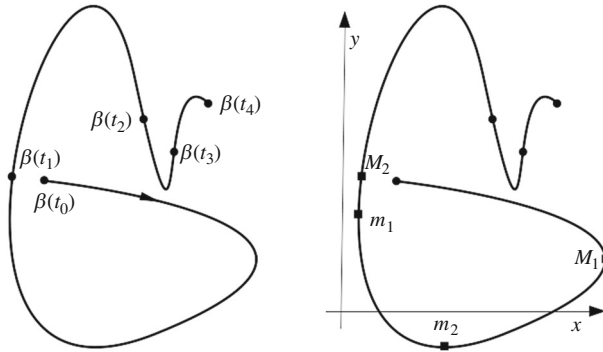


Fig. 3 A 4-convex curve to illustrate the proof of Lemma 2.7

**Lemma 2.7** *Let  $\beta$  be a piece-wise  $C^1$ -curve in  $\mathbb{R}^2$ ,  $n$ -convex with finitely many self-intersections. Let  $\Lambda^Q$  be a full rank lattice, then*

$$I_{\beta}^Q(K) \leq \sqrt{2} \cdot K \cdot \text{length}(Q^{-1}\beta) + 19n + 1.$$

**Proof** Let  $\beta: [0, 1] \rightarrow \mathbb{R}^2$  be a parametrization as in the statement of the lemma, and denote the  $x, y$ -coordinates of  $\beta$  by  $x(t) = \langle \beta(t), e_1 \rangle$  and  $y(t) = \langle \beta(t), e_2 \rangle$ .

We will first show monotonicity of the coordinates for  $\beta$  in certain intervals. Let  $t_1, \dots, t_{n-1}$  be as in Definition 2.4. If  $\beta([0, t_1])$  is a line segment, then  $x(t), y(t)$  are monotonous, otherwise let  $M_j, m_j \in \beta([0, t_1])$  for  $1 \leq j \leq 2$  satisfy

$$\langle M_j, e_j \rangle = \max_{t \in [0, t_1]} \langle \beta(t), e_j \rangle \quad \text{and} \quad \langle m_j, e_j \rangle = \min_{t \in [0, t_1]} \langle \beta(t), e_j \rangle.$$

See Fig. 3. Choose  $o_j^M \in \beta^{-1}(M_j) \cap [0, t_1]$  and  $o_j^m \in \beta^{-1}(m_j) \cap [0, t_1]$  for  $j \in \{1, 2\}$ . Let  $o_1, o_2, o_3, o_4$  denote these  $o_j^m, o_j^M$  in ascending order. In the calculations below the pair  $(\tau_1, \tau_2)$  is  $(o_k, o_{k+1})$  for some choice of  $k \in \{0, 1, 2, 3, 4\}$  such that  $\tau_2 - \tau_1 > 0$ , where  $o_0 = 0$  and  $o_5 = t_1$ .

It follows by the convexity assumption that the  $x, y$ -coordinates of  $\beta(t)$  are monotone for  $t \in [\tau_1, \tau_2]$  as we will show: first, for any  $\kappa \in [0, t_1]$  and  $\tau \in \{s \in [0, t_1] \mid \beta(s) \neq \beta(\kappa)\}$  we either have that the arc segment of  $\beta$  with domain between  $\tau$  and  $\kappa$  is a straight line, or

$$\{(1-s)\beta(\kappa) + s\beta(\tau) \mid 0 < s < 1\} \cap \beta([0, t_1]) = \emptyset.$$

To prove monotonicity of  $x(t)$  and  $y(t)$ , we work out the example  $\beta(\tau_1) = m_2$ . The existence of numbers  $\tau_1 < \epsilon_1 < \epsilon_2 < \tau_2$  such that  $y(\epsilon_1) > y(\epsilon_2)$  will lead to a contradiction in this case (by the assumption on  $\beta(\tau_1)$ ,  $y(t)$  is monotonously increasing). Either  $x(\epsilon_1) \leq x(\epsilon_2)$  or  $x(\epsilon_2) < x(\epsilon_1)$ , and in both cases, one of the

polygons with vertices and edges of the form

$$\begin{aligned} \beta(o_2^M) &\longrightarrow \beta(\tau_1) \longrightarrow \beta(\epsilon_1) \longrightarrow \beta(\epsilon_2), \\ \beta(\tau_1) &\longrightarrow \beta(\epsilon_1) \longrightarrow \beta(\epsilon_2) \longrightarrow \beta(o_2^M), \end{aligned}$$

will contradict convexity, depending if either  $o_2^M < \tau_1$  or  $o_2^M \geq \tau_2$  (by definition  $o_2^M \notin (\tau_1, \tau_2)$ ). For one of the polygons, we can find two points  $A, B$  on it, such that  $\overrightarrow{AB}$  intersects the polygon in three points, and the same will hence also be true for the arc segments  $\beta([o_2^M, \tau_2])$  or  $\beta([\tau_1, o_2^M])$ . Note that  $y(o_2^M) \geq y(\epsilon_1) > y(\epsilon_2) \geq y(\tau_1) = y(o_2^m)$ .

Thus  $y(t)$  is monotonously increasing, and we use this fact to show that  $x(t)$  is monotone. Assume there are numbers  $\tau_1 < \epsilon_1 < \epsilon_2 < \epsilon_3 < \tau_2$  so that  $x(\epsilon_2) < x(\epsilon_1)$  and  $x(\epsilon_2) < x(\epsilon_3)$ , and if either  $o_1^* < \tau_1$  or  $o_1^* \geq \tau_2$  (here  $o_1^*$  is a place holder where  $*$   $\in \{m, M\}$  depending on the polygonal shape), one of the polygons with vertices and edges of the form

$$\begin{aligned} \beta(\tau_1) &\longrightarrow \beta(\epsilon_1) \longrightarrow \beta(\epsilon_2) \longrightarrow \beta(\epsilon_3) \longrightarrow \beta(o_1^*), \\ \beta(o_1^*) &\longrightarrow \beta(\tau_1) \longrightarrow \beta(\epsilon_1) \longrightarrow \beta(\epsilon_2) \longrightarrow \beta(\epsilon_3), \end{aligned}$$

contradicts convexity in a similar fashion as above. The case “ $x(\epsilon_1) < x(\epsilon_2)$  and  $x(\epsilon_3) < x(\epsilon_2)$ ” is similar. Thus  $x(t)$  is monotonous. The other possibilities for  $\beta(\tau_1)$  reduce to the case above by applying rotations to  $\beta([0, t_1])$ .

We define two supporting axes for each non-trivial arc  $\beta([\tau_1, \tau_2])$  as follows:

$$\{(1 - t)\beta(\tau_j) + tZ \mid 0 \leq t \leq 1\} \quad \text{where} \quad Z = \begin{pmatrix} \langle \beta(\tau_1), e_1 \rangle \\ \langle \beta(\tau_2), e_2 \rangle \end{pmatrix} \in \mathbb{R}^2.$$

These line segments above are parallel to the axes, and we denote them accordingly by  $L_x$  and  $L_y$ . Let  $\gamma = \beta([\tau_1, \tau_2])$  and  $Q$  be the identity matrix  $\mathbb{1}$ , then

$$I_\gamma^{\mathbb{1}}(K) \leq I_{L_x}^{\mathbb{1}}(K) + I_{L_y}^{\mathbb{1}}(K)$$

by the following argument: since  $x(t), y(t)$  are monotonous for  $t \in [\tau_1, \tau_2]$ , say both are increasing, then  $\beta$  can only exit a domain  $(1/K)\Omega^{\mathbb{1}} + (1/K)p, p \in \Lambda^{\mathbb{1}} = \mathbb{Z}^2$ , through the top or right side.

- If  $\beta$  leaves through the right side,  $I_\gamma^{\mathbb{1}}(K)$  and  $I_{L_x}^{\mathbb{1}}(K)$  increase by one,
- or else  $I_\gamma^{\mathbb{1}}(K)$  and  $I_{L_y}^{\mathbb{1}}(K)$  increase by one.

Thus, with  $\text{length}(L_x) = c$  and  $\text{length}(L_y) = d$ , we have

$$I_\gamma^{\mathbb{1}}(K) \leq I_{L_x}^{\mathbb{1}}(K) + I_{L_y}^{\mathbb{1}}(K) \leq (c + d)K + 4.$$

We further use the inequality  $c + d \leq \sqrt{2} \cdot \sqrt{c^2 + d^2}$  valid for all  $c, d \in \mathbb{R}$ , to derive

$$I_\gamma^{\mathbb{1}}(K) \leq \sqrt{2} \cdot \sqrt{c^2 + d^2} \cdot K + 4 \leq \sqrt{2} \cdot \text{length}(\gamma) \cdot K + 4,$$

where we used that the shortest path between  $\beta(\tau_1)$  and  $\beta(\tau_2)$  has length  $\sqrt{c^2 + d^2}$ . The same reasoning applies to all sub-intervals (which are at most five), and after summing up we obtain

$$I_\beta^1(K) \leq \sqrt{2} \cdot \text{length}(\beta) \cdot K + 19n + 1.$$

We will reduce the general case to the one above, by regarding the curve  $\gamma = Q^{-1}\beta$ . It is clear that  $I_\gamma^1(K) = I_\beta^Q(K)$ . Regularity of a curve is not affected by invertible matrices, neither is the notion of  $n$ -convexity nor the values  $t_1, \dots, t_n$ .  $\square$

Note that the constant  $\sqrt{2}$  in Lemma 2.7 cannot be improved, as the example of the translated diagonal of  $I^2$  with  $Q = \mathbb{1}$  already shows.

### 2.3 Bound of $d^Q(K)$ and Proof of the Main Result

Before proving Theorem 1.3, we first bound the quantity  $d^Q(K)$  and use this bound to express  $N^Q(K)^{-1/2}$  in terms of  $K$  and  $Q$  in (3).

**Lemma 2.8** *Given a lattice  $\Lambda^Q$  and  $K \in \mathbb{N}$ , then*

$$d^Q(K) \leq \sqrt{2} \cdot 2 \cdot (\|Q^{-1}e_1\| + \|Q^{-1}e_2\|) + \frac{20}{K} \leq \frac{4 \cdot \|Q\|_F}{|\det(Q)|} + \frac{20}{K}.$$

**Proof** Using the notation as introduced in Definition 1.2, we see that  $|\det(Q)| \cdot K^{-2}$  is the area of each  $T_K(p)$ , with  $p \in \Lambda^Q$ . Now the number of points  $\{z_K^p\}$  with  $T_K(p) \subset I^2$  is at most  $\lfloor K^2 \cdot |\det(Q)|^{-1} \rfloor$ , and this differs from  $N^Q(K)$  at most by the number of points  $\{z_K^q\}$  where  $T_K(q)$  intersects  $\partial I^2$ . The number of these points is linear in  $K$  by Lemma 2.7, where the constant can be computed explicitly since  $\text{image}(\beta) = \partial I^2$ . The last part follows by applying the inequality  $a + b \leq \sqrt{2} \cdot \sqrt{a^2 + b^2}$  and the well-known formula for  $Q^{-1}$  in terms of elements of  $Q$  and a factor of  $\det(Q)^{-1}$ .  $\square$

**Remark** The term  $M^Q(K)$  can be bounded in a similar fashion by the same constant  $+ O(K^{-1})$ : one just has to apply the triangle inequality in its definition, bound each term by  $|\det(Q)| \cdot K^{-2}$ , and bound the number of summands as above.

For later use, we note that there is some  $s_K$  with  $|s_K| \leq d^Q(K)/N^Q(K)$ , such that

$$\sqrt{\frac{1}{|\det(Q)| \cdot N^Q(K)}} = \frac{1}{K} \sqrt{1 + \frac{K^2 - |\det(Q)| \cdot N^Q(K)}{|\det(Q)| \cdot N^Q(K)}} = \frac{1}{K} + s_K. \tag{3}$$

**Proof of Theorem 1.3** Let  $P^Q(K) = \{z_1, \dots, z_N\}$  where  $N = N^Q(K)$ . Given a spherical cap  $C(w, t) \subset S^2$ , we regard  $A_{w,t} := L^{-1}(C(w, t) \cap S_p^2) \subset I^2$ . Since  $L$  is area preserving, the proof will be complete once we can bound

$$(\star) := \left| \frac{1}{N} \sum_{j=1}^N \chi_{A_{w,t}}(z_j) - \lambda(A_{w,t}) \right|.$$

An upper bound of  $(\star)$  is established by a classical approach,<sup>2</sup> where  $\Lambda^Q$  is reduced to sets  $V, B$  of volume and boundary contributions in order to apply the triangle inequality: let  $\{p_1, \dots, p_N\} \subset \Lambda^Q$  be such that  $z_j \in T_K(p_j)$ , then

$$(\star) \leq \sum_{p \in V} \left| \frac{1}{N} - \lambda(T_K(p)) \right| + \left| \sum_{p_j \in B} \frac{1}{N} \chi_{A_{w,t}}(z_j) - \sum_{p_j \in B} \lambda(T_K(p_j) \cap A_{w,t}) \right|, \tag{4}$$

where  $V$  and  $B$  are defined below, and where the notation is as in Definition 1.2. Set

$$V = V(w, t, K) = \{p \in \Lambda^Q \mid T_K(p) \subset A_{w,t}\},$$

hence  $z_K^p \in A_{w,t}$  for  $p \in V$ , and we define the boundary term  $B = B_C \cup B_I$  by

$$B_C = B_C(w, t, K) = \{q \in \Lambda^Q \mid T_K(q) \cap (\partial A_{w,t} \setminus \partial I^2) \neq \emptyset\} \quad \text{and}$$

$$B_I = B_I(w, t, K) = \{q \in \Lambda^Q \mid T_K(q) \cap (\partial A_{w,t} \cap \partial I^2) \neq \emptyset\}.$$

(Note that we potentially over-count in  $B_C$  or  $B_I$ , since we do not remove points present in previously defined sets.) For each  $p \in V$ , the difference  $\epsilon_K = N^{-1} - \lambda(T_K(p))$  is constant, thus the contribution to discrepancy coming from the area is given by  $(\#V \leq N)$

$$|\epsilon_K| \cdot \#V \leq \left| \frac{1}{N} - \frac{|\det(Q)|}{K^2} \right| \cdot N = d^Q(K) \cdot \frac{|\det(Q)|}{KN} \cdot N.$$

Next we bound the discrepancy coming from the boundary terms  $B_C$  and  $B_I$ . If for  $q \in B$  we have  $z_K^q \notin A_{w,t}$ , then the difference inside the absolute value of the second sum in (4) has a negative contribution of

$$\lambda(T_K(q) \cap A_{w,t});$$

otherwise the difference has a contribution of

$$\frac{1}{N} - \lambda(T_K(q) \cap A_{w,t}) = \lambda(T_K(q) \setminus A_{w,t}) + \epsilon_K.$$

Both contributions are hence bounded by  $\lambda(T_K(p))$  up to  $\epsilon_K$ , and we obtain

$$\sum_{p \in B_C} \left| \frac{\chi_{A_{w,t}}(z_K^p)}{N} - \lambda(T_K(p) \cap A_{w,t}) \right| \leq \frac{|\det(Q)|}{K} \cdot \sqrt{2} \cdot C_L^Q + O(K^{-2}),$$

where we applied Lemma 2.7 (thanks to Corollary 2.5) and the definition of  $C_L^Q$ . Finally, the discrepancy coming from the contribution of points in  $B_I$  is less than  $|\det(Q)| \cdot K^{-1} M^Q(K)$  by definition, and (3) puts our bounds in terms of  $N$ .  $\square$

<sup>2</sup> See the Gauss circle problem, where Lemma 2.7 could be applied.

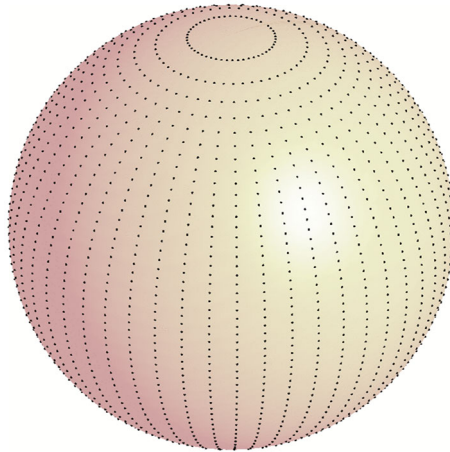


Fig. 4 Image of the points  $P^{\mathbb{1}}(50)$  with  $N = 2500$

### 3 Applications to Specific Cases

In this section we first focus on point sets with special choice  $Q = \mathbb{1}$ , for which we obtain the lowest upper bound on the leading term of the spherical cap discrepancy up to date. This example also shows that in general the order of  $O(N^{-1/2})$  cannot be improved. We further see that a separation distance of matching order cannot be derived in general. The second example serves as visual evidence that some point sets of the type regarded in this article should have a spherical cap discrepancy of much lower order.

#### 3.1 The Standard Lattice

Let

$$P^{\mathbb{1}}(K) = \left( \frac{1}{K} \mathbb{Z}^2 + \frac{1}{2K} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \cap I^2,$$

see Fig. 4. Then  $d^{\mathbb{1}}(K) = 0$  and  $M^{\mathbb{1}}(K) = O(K^{-1})$ , and we obtain with Theorem 1.3 and Lemma 2.2 the constant mentioned in the introduction:

$$\frac{1}{2} \leq \sqrt{N} \cdot \mathcal{D}(L(P^{\mathbb{1}}(K))) \leq \sqrt{18} \approx 4.242641.$$

The lower bound comes from a special choice for a spherical cap: take as center the north pole and let the height  $t \rightarrow ((2K - 1)/(2K))^+$ . The boundary of the aforementioned cap with height  $t = (2K - 1)/(2K)$  has length of order  $K^{-1/2}$ , while there are  $K$  points equi-distributed on it, thus the distance between consecutive points is of order  $K^{-3/2} = N^{-3/4}$ .

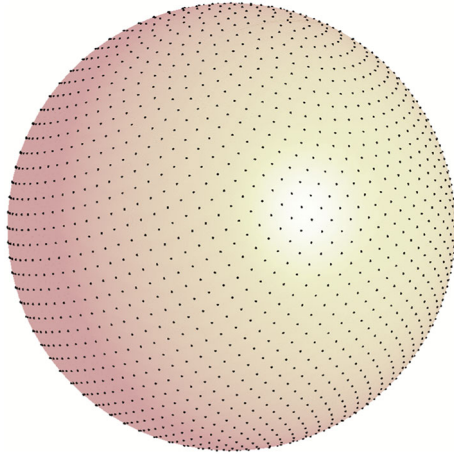


Fig. 5 Image of points related to the lattice  $\Lambda^{Q(\varphi,1)}$  with  $K = 50$  and  $N^Q = 2500$

### 3.2 Orthonormal Lattices

We choose points  $P^{Q(x,y)}(K)$ , where  $Q$  is chosen orthonormal up to a factor, i.e.,

$$Q(x, y) = \frac{1}{y} \begin{pmatrix} x & -1 \\ 1 & x \end{pmatrix}$$

and

$$Q^{-1}(x, y) = \frac{y}{x^2 + 1} \begin{pmatrix} x & 1 \\ -1 & x \end{pmatrix},$$

for  $x \in \mathbb{R}$ ,  $y > 0$ . Then, by Lemma 2.3,

$$C_L^{Q(x,y)} \leq 3 \cdot \|Q^{-1}(x, y)e_1\| = 3 \frac{y}{\sqrt{x^2 + 1}}$$

and also note that  $y \cdot \sqrt{\det(Q(x, y))} = \sqrt{x^2 + 1}$ . We can modify  $P^{Q(x,y)}(K)$  such that  $d^Q(K)$ ,  $M^Q(K)$  are of order  $K^{-1}$ , and hence by Theorem 1.3 we obtain many deterministic point sets with the bound

$$\mathcal{D}(L(P_{\text{mod}}^{Q(x,y)}(K))) \leq \sqrt{\frac{18}{N_m^{Q(x,y)}}} + O\left(\frac{1}{N_m^{Q(x,y)}}\right).$$

For  $\varphi = (1 + \sqrt{5})/2$ , the choice  $Q(\varphi, 1)$  yields the image in Fig. 5, resembling spherical Fibonacci lattices and grids as in [1, 26]. Note that no modification was necessary for  $P^{Q(\varphi,1)}(K)$ —this seems to be related to directional discrepancy as in [10], but we do not pursue this direction.

**Remark** Definition 2.4 and Lemma 2.7 were first developed by this author to prove the result in a work with Julian Hofstadler and Michelle Mastrianni, but the items were improved and streamlined in the current work.

**Remark** Clearly the proof idea of Theorem 1.3 extends beyond the Lambert projection to any area preserving map  $\Gamma$  between a surface and shape  $R \subset \mathbb{R}^2$  with  $\partial R$  being  $n$ -convex—as long as the boundary of the analog of spherical cap under  $\Gamma$  is  $n$ -convex with universally bounded  $n$  and length.

**Acknowledgements** The author thanks the anonymous referees for valuable suggestions and for providing the proof of Lemma 2.3. Further, thanks to Michelle Mastrianni for improving the text of the document, thanks to Dmitriy Bilyk and Arno Kuijlaars for skimming the paper and useful remarks; and thanks to the people involved with GNU Octave, Inkscape, LibreOffice, and TeXstudio, who made this document and many others possible.

**Funding** The work was supported by the Methusalem grant METH/21/03—long term structural funding of the Flemish Government.

**Data Availability** Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

**Code Availability** The code to produce the point distributions in Sect. 3 will be made available on the authors website: [www.damirferizovic.wordpress.com](http://www.damirferizovic.wordpress.com).

## Declarations

**Competing Interests** The author has no relevant financial or non-financial interests to disclose.

## References

1. Aistleitner, C., Brauchart, J.S., Dick, J.: Point sets on the sphere  $\mathbb{S}^2$  with small spherical cap discrepancy. *Discrete Comput. Geom.* **48**(4), 990–1024 (2012)
2. Alishahi, K., Zamani, M.: The spherical ensemble and uniform distribution of points on the sphere. *Electron. J. Probab.* **20**, # 23 (2015)
3. Armentano, D., Beltrán, C., Shub, M.: Minimizing the discrete logarithmic energy on the sphere: the role of random polynomials. *Trans. Am. Math. Soc.* **363**(6), 2955–2965 (2011)
4. Bauer, R.: Distribution of points on a sphere with application to star catalogs. *J. Guid. Cont. Dyn.* **23**(1), 130–137 (2000)
5. Beck, J.: Some upper bounds in the theory of irregularities of distribution. *Acta Arith.* **43**(2), 115–130 (1984)
6. Beck, J.: Sums of distances between points on a sphere—an application of the theory of irregularities of distribution to discrete geometry. *Mathematika* **31**(1), 33–41 (1984)
7. Bellhouse, D.R.: Area estimation by point-counting techniques. *Biometrics* **37**(2), 303–312 (1981)
8. Beltrán, C., Etayo, U.: The diamond ensemble: a constructive set of spherical points with small logarithmic energy. *J. Complexity* **59**, # 101471 (2020)
9. Beltrán, C., Marzo, J., Ortega-Cerda, J.: Energy and discrepancy of rotationally invariant determinantal point processes in high dimensional spheres. *J. Complexity* **37**, 76–109 (2016)
10. Bilyk, D., Ma, X., Pipher, J., Spencer, C.: Directional discrepancy in two dimensions. *Bull. Lond. Math. Soc.* **43**(6), 1151–1166 (2011)
11. Bogomolny, E., Bohigas, O., Lebouf, P.: Distribution of roots of random polynomials. *Phys. Rev. Lett.* **68**(18), 2726–2729 (1992)
12. Borodachov, S.V., Hardin, D.P., Saff, E.B.: *Discrete Energy on Rectifiable Sets*. Springer Monographs in Mathematics. Springer, New York (2019)



13. Brauchart, J.S., Grabner, P.J.: Distributing many points on spheres: minimal energy and designs. *J. Complexity* **31**(3), 293–326 (2015)
14. Brauchart, J.S., Reznikov, A.B., Saff, E.B., Sloan, I.H., Wang, Y.G., Womersley, R.S.: Random point sets on the sphere—hole radii, covering, and separation. *Exp. Math.* **27**(1), 62–81 (2018)
15. Choirat, Ch., Seri, R.: Numerical properties of generalized discrepancies on spheres of arbitrary dimension. *J. Complexity* **29**(2), 216–235 (2013)
16. Cook, J.M.: Rational formulae for the production of a spherically symmetric probability distribution. *Math. Tables Aids Comput.* **11**, 81–82 (1957)
17. Cui, J., Freeden, W.: Equidistribution on the sphere. *SIAM J. Sci. Comput.* **18**(2), 595–609 (1997)
18. Etayo, U.: Spherical cap discrepancy of the diamond ensemble. *Discrete Comput. Geom.* **66**(4), 1218–1238 (2021)
19. Górski, K.M., Hivon, E., Banday, A.J., Wandelt, B.D., Hansen, F.K., Reinecke, M., Bartelmann, M.: HEALPix: a framework for high-resolution discretization and fast analysis of data distributed on the sphere. *Astrophys. J.* **622**(2), 759–771 (2005)
20. Grabner, P.J., Klinger, B., Tichy, R.F.: Discrepancies of point sequences on the sphere and numerical integration. In: 2nd International Conference on Multivariate Approximation (Witten-Bommerholz 1996). *Math. Res.*, vol. 101, pp. 95–112. Akademie Verlag, Berlin (1997)
21. Hardin, D.P., Michaels, T., Saff, E.B.: A comparison of popular point configurations on  $S^2$ . *Dolomites Res. Notes Approx.* **9**, 16–49 (2016)
22. Krishnapur, M.R.: Zeros of Random Analytic Functions. PhD thesis, University of California, Berkeley (2006). [arXiv:math/0607504](https://arxiv.org/abs/math/0607504)
23. Lubotzky, A., Phillips, R., Sarnak, P.: Hecke operators and distributing points on the sphere. I. *Commun. Pure Appl. Math.* **39**(suppl.), S149–S186 (1986)
24. Rakhmanov, E.A., Saff, E.B., Zhou, Y.M.: Minimal discrete energy on the sphere. *Math. Res. Lett.* **1**(6), 647–662 (1994)
25. Saff, E.B., Kuijlaars, A.B.J.: Distributing many points on a sphere. *Math. Intell.* **19**(1), 5–11 (1997)
26. Swinbank, R., Purser, R.J.: Fibonacci grids: A novel approach to global modelling. *Q. J. R. Meteorol. Soc.* **132**(619), 1769–1793 (2006)
27. Whyte, L.L.: Unique arrangements of points on a sphere. *Am. Math. Mon.* **59**, 606–611 (1952)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.