Discrete Comput Geom (2012) 48:990–1024
DOI 10.1007/s00454-012-9451-3

Point Sets on the Sphere $\mathbb{S}^2$ with Small Spherical Cap Discrepancy

C. Aistleitner · J.S. Brauchart · J. Dick

Received: 14 September 2011 / Revised: 31 July 2012 / Accepted: 7 August 2012 /
Published online: 28 August 2012

Abstract In this paper we study the geometric discrepancy of explicit constructions of uniformly distributed points on the two-dimensional unit sphere. We show that the spherical cap discrepancy of random point sets, of spherical digital nets and of spherical Fibonacci lattices converges with order $N^{-1/2}$. Such point sets are therefore useful for numerical integration and other computational simulations. The proof uses an area-preserving Lambert map. A detailed analysis of the level curves and sets of the pre-images of spherical caps under this map is given.

Keywords Discrepancy · Isotropic discrepancy · Lambert map · Level curve · Level set · Numerical integration · Quasi-Monte Carlo · Spherical cap discrepancy

1 Introduction

Let $\mathbb{S}^2 = \{\boldsymbol{z} \in \mathbb{R}^3 : \|\boldsymbol{z}\| = 1\}$ be the unit sphere in the Euclidean space $\mathbb{R}^3$ provided with the norm $\|\cdot\|$ induced by the usual inner product $\boldsymbol{x} \cdot \boldsymbol{y}$. On this sphere we consider the Lebesgue surface area measure σ normalised to a probability measure ($\int_{\mathbb{S}^2} \mathrm{d}\sigma = 1$).

This paper is concerned with uniformly distributed sequences of points on $\mathbb{S}^2$. Informally speaking, a sequence of points is called uniformly distributed if every

C. Aistleitner
Institute of Mathematics A, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria
e-mail: aistleitner@math.tugraz.at

J.S. Brauchart · J. Dick (✉)
School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, 2052, Australia
e-mail: josef.dick@unsw.edu.au

J.S. Brauchart
e-mail: j.brauchart@unsw.edu.au

reasonably defined (clopen) $A \subseteq \mathbb{S}^2$ gets a fair share of points as their number N grows. Given a triangular scheme $\{z_{1,N}, \dots, z_{N,N}\}$, $N \geq 1$, of points on $\mathbb{S}^2$ in such a case one has

$$\lim_{N\to\infty} \frac{\operatorname{card}(\{j : z_{j,n} \in A\})}{N} = \sigma(A) \tag{1a}$$

(where card denotes the cardinality of the set) or, equivalently (defined in terms of numerical integration),

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} f(z_{j,N}) = \int_{\mathbb{S}^2} f \, \mathrm{d}\sigma \quad \text{for every } f \text{ continuous on } \mathbb{S}^2. \tag{1b}$$

The degree of uniformity is quantified by the so-called *spherical cap discrepancy*.

A spherical cap $C = C(\boldsymbol{w}, t)$ centred at $\boldsymbol{w} \in \mathbb{S}^2$ with height $t \in [-1, 1]$ is given by the set

$$C(\boldsymbol{w}, t) = \{\boldsymbol{y} \in \mathbb{S}^2 : \boldsymbol{w} \cdot \boldsymbol{y} > t\}.$$

(We assume that spherical caps are open subsets of $\mathbb{S}^2$.) The boundary of $C(\boldsymbol{w}, t)$ then is

$$\partial C(\boldsymbol{w}, t) = \{\boldsymbol{y} \in \mathbb{S}^2 : \boldsymbol{w} \cdot \boldsymbol{y} = t\}.$$

Let $Z_N = \{z_0, \dots, z_{N-1}\} \subseteq \mathbb{S}^2$ be an N-point set on the sphere $\mathbb{S}^2$. The local discrepancy with respect to a spherical cap C measures the difference between the proportion of points in C (the empirical measure of C) and the normalised surface area of C. The spherical cap discrepancy is then the supremum of the local discrepancy over all spherical caps, as stated in the following definition.

Definition 1 The *spherical cap discrepancy* of an N-point set $Z_N = \{z_0, \dots, z_{N-1}\} \subseteq \mathbb{S}^2$ is

$$D(Z_N) = \sup_{\boldsymbol{w} \in \mathbb{S}^2} \delta(Z_N; \boldsymbol{w}), \quad \delta(Z_N; \boldsymbol{w}) = \sup_{-1 \leq t \leq 1} \left| \frac{1}{N} \sum_{n=0}^{N-1} 1_{C(\boldsymbol{w},t)}(z_n) - \sigma\big(C(z, t)\big) \right|.$$

If the point set Z_N is well-distributed, then this discrepancy is small. In fact, a sequence of N-point systems $(Z_N)_{N\geq 1}$ satisfying

$$\lim_{N\to\infty} D(Z_N) = 0, \tag{2}$$

is called *asymptotically uniformly distributed*. Using, for example, the classical *Erdös–Turán type inequality* (cf. Grabner [25], also cf. Li and Vaaler [38]) or *Le-Veque type inequalities* (Narcowich et al. [41]) and the fact that the set of polynomials is dense in the set of continuous functions, one can show that (2) is equivalent to (1b).

It is known from [6] that there are constants $c, C > 0$, independent of N, such that a *low-discrepancy* scheme $\{Z_N^*\}_{N\geq 2}$ satisfies

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

The lower bound holds for all N-point sets Z_N on $\mathbb{S}^2$ and there always exists an N-point set $Z_N \subseteq \mathbb{S}^2$ such that the upper bound holds. The proof of the upper bound is probabilistic in nature and is thus non-constructive. To our best knowledge, explicit constructions of low-discrepancy schemes are not known. (In this paper we restrict ourselves to the sphere $\mathbb{S}^2$, though some of the results are known for spheres of dimension $d \geq 2$.)

An explicit construction of points Z_N with small spherical cap discrepancy has been given in [39, 40]. For instance, in [39] it was shown that

$$D(Z_N) \leq C(\log N)^{2/3} N^{-1/3}. \tag{4}$$

The numerical experiments in [39] indicate a convergence rate of $\mathcal{O}(N^{-1/2})$.

In this paper we give explicit constructions of point sets Z_N for which we have

$$D(Z_N) \leq 44\sqrt{2} N^{-1/2}.$$

Our numerical results indicate a convergence rate of $\mathcal{O}((\log N)^c N^{-3/4})$ for some $1/2 \leq c \leq 1$, see Tables 1, 2 below.

The *spherical cap* $\mathbb{L}_2$*-discrepancy*

$$D_{\mathbb{L}_2}(Z_N) := \left\{ \int_{-1}^{1} \int_{\mathbb{S}^2} \left| \frac{1}{N} \sum_{n=0}^{N-1} 1_{C(\boldsymbol{w},t)}(z_n) - \sigma\big(C(z,t)\big) \right|^2 \mathrm{d}\sigma(\boldsymbol{w})\, \mathrm{d}t \right\}^{1/2},$$

which averages the local discrepancy for a spherical cap over all caps, provides a lower bound for the spherical cap discrepancy. It is closely related to the *sum of distances* and its continuous counterpart the *distance integral* by means of *Stolarsky's Invariance Principle* [50] for the Euclidean distance and the 2-sphere:

$$\begin{aligned}\frac{1}{N^2} \sum_{j=1}^{N} \sum_{k=1}^{N} |z_j - z_k| + 4\big[D_{\mathbb{L}_2}^C(z_1, \ldots, z_N)\big]^2 &= \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} |z - \boldsymbol{w}|\, \mathrm{d}\sigma(z)\, \mathrm{d}\sigma(\boldsymbol{w}) \\ &=: V_{-1}\big(\mathbb{S}^2\big) = \frac{4}{3}.\end{aligned}$$

This gives a simple way of computing the spherical cap $\mathbb{L}_2$-discrepancy of point sets on $\mathbb{S}^2$. In [14] it is shown that the spherical cap $\mathbb{L}_2$-discrepancy of Z_N can be interpreted as the worst-case error of an equal-weight numerical integration rule with node set Z_N for functions in the unit ball of a certain Sobolev space over $\mathbb{S}^2$. It is shown in [13] that on average (i.e., for randomly chosen points independently identically uniformly distributed over the sphere), the expected squared worst-case error is of the form $(4/3)N^{-1}$. Thus the expected value of the squared spherical cap discrepancy satisfies

$$8\mathbb{E}\big[D(Z_N)\big]^2 \geq 4\mathbb{E}\big[D_{\mathbb{L}_2}(Z_N)\big]^2 = \frac{4}{3} N^{-1}. \tag{5}$$

We study the expected value and the *typical* asymptotic order of the spherical cap discrepancy of random point sets in detail in Sect. 4. Among other results, we show that there is also a constant $C > 0$ such that $\mathbb{E}[D(Z_N)] \leq C N^{-1/2}$.

Point configurations maximising the sum of distances, by Stolarsky's Invariance Principle, have low spherical cap $\mathbb{L}_2$-discrepancy. It is known from [6] that low spherical cap $\mathbb{L}_2$-discrepancy point sets satisfy relations similar to (3) except for the logarithmic term introduced by the probabilistic approach. The upper bound for the spherical cap discrepancy of maximum sum of distances points $\hat{Z}_N^*$ obtained in [41] is much weaker but still better than (4): For some positive constant $c > 0$, not depending on N,

$$D\bigl(Z_N^*\bigr) \le c N^{-3/8}.$$

For point configurations Z_N^* emulating electrons restricted to move on $\mathbb{S}^2$ in the most stable equilibrium, i.e. minimising their *Coulomb potential energy* essentially given by

$$\sum_{\substack{j=1}}^{N} \sum_{\substack{k=1 \\ j \neq k}}^{N} \frac{1}{|z_j - z_k|},$$

one can show the bound

$$D\bigl(Z_N^*\bigr) \le C N^{-1/2} \log N.$$

The estimate $D(Z_N^*) = \mathcal{O}(N^{-1/2})$ was conjectured by Korevaar [32] and later proved (up to the logarithmic factor) by Götz [24]. When allowing so-called K*-regular* test sets[1] introduced by Sjögren [46], the estimate above is sharp in the following sense: The upper bound holds for any K-regular test set, whereas there are some numbers K_0 and c such that to any N points $z_1, \dots, z_N \in \mathbb{S}^2$ there is a K_0-regular test set B with [24, Corollary 2]

$$c\, K_0\, N^{-1/2} \le \left| \frac{1}{N} \sum_{n=1}^{N} 1_B(z_n) - \sigma(B) \right|.$$

(The lower bound also applies to the explicit constructions given in this paper.) Bounds for the spherical cap discrepancy of so-called minimal Riesz energy configurations (for the concept of Riesz energy, see, e.g., Saff and Kuijlaars [45] and Hardin and Saff [28]) can be found in [11] (for the logarithmic energy), Damelin and Grabner [17, 18] (the first hyper-singular case), and [41] (sums of generalised distances). Wagner [54] estimates the spherical cap discrepancy in terms of the Riesz energy. It should be mentioned that there are very few known explicit constructions of point configurations with optimal Riesz energy. In general, one has to rely on numerical optimisation to generate such point sets. The underlying (constrained) optimisation problem is highly nonlinear. Moreover, numerical results indicate that the number of local minima increases exponentially with the number of points. (For the computational complexity see, e.g., Bendito et al. [9].)

[1] Roughly speaking, σ-measurable subsets B of $\mathbb{S}^2$ whose δ-neighbourhoods $\partial_\delta B$ relative to $\mathbb{S}^2$ satisfy $\sigma(\partial_\delta B) \le K\delta$. For example, spherical caps are K_1-regular for some fixed K_1.

Spherical n-designs introduced by Delsarte, Goethals and Seidel in the landmark paper [19] are node sets for equal-weight numerical integration rules such that all spherical polynomials of degree $\leq n$ are integrated exactly. Grabner and Tichy [26] give the following upper bound of the spherical cap discrepancy of a spherical n-design with $N(n)$ points:

$$D\big(Z^*_{N(n)}\big) \leq C n^{-1}, \tag{6}$$

which immediately follows from the aforementioned *Erdös–Turán type inequality*. (See also Andrievskii et al. [4] for a similar form for K-regular test sets.)

A spherical n-design is the solution of a system of polynomial equations (one for every spherical harmonic of the real orthonormal basis of the space of spherical polynomials of degree $\leq n$). Hence, a natural lower bound for the number of points of a spherical n-design is given by the dimension of the involved polynomial space; that is, one needs at least $\geq n^2/4$ points. The famous conjecture that $C\,n^2$ points (for some universal $C > 0$) are sufficient for a spherical n-design seems to have been settled by Bondarenko et al. [10]. The proposed proof is non-constructive. Hardin and Sloane [29] propose a construction of so-called putative spherical n-designs with $(1/2)n^2 + o(n^2)$ points. The variational characterisation of spherical designs introduced in [48] (also cf. [26]) leads to a minimisation problem for a certain energy functional (changing with n) whose minimiser is a spherical n-design if and only if the functional becomes zero. Numerical results also suggest a coefficient $1/2$. When allowing more points, $N(n) = (n+1)^2$, interval-based methods yield, in principle, the existence of a spherical n-design near so-called extremal (maximum determinant points, cf. [47]). Due to the computational cost this approach was carried out only for $n \leq 20$. Very recently, Chen et al. [16] devised a computational algorithm based on interval arithmetic that, upon successful completion, verifies the existence of a spherical n-design with $(n+1)^2$ points and provides narrow interval enclosures which are known to contain these nodes with mathematical certainty. The spherical cap discrepancy of all such obtained spherical n-design with $\mathcal{O}(n^2)$ points can then be bounded by $C' N^{-1/2}$ by (6). For the sake of completeness it should be mentioned that the tensor product rules used by Korevaar and Meyers [33] to prove the existence of spherical n-designs of $N(n) = \mathcal{O}(n^3)$ points give rise to $N(n)$-point configurations whose spherical cap discrepancy can be bounded by $C''[N(n)]^{-1/3}$ by (6).

From [14] it follows that the spherical cap discrepancy of a point set $Z_N = \{z_0, \dots, z_{N-1}\} \subseteq \mathbb{S}^2$ yields an upper bound on the integration error in certain Sobolev spaces of functions defined on $\mathbb{S}^2$ using a quadrature rule $Q_N(f) = \frac{1}{N}\sum_{n=0}^{N-1} f(z_n)$. Thus, our results here provide an explicit mean of finding quadrature points for numerical integration of functions defined on $\mathbb{S}^2$. Our result here improves the bound on the integration error in [12] by a factor of $\sqrt{\log N}$.

The construction of the points on $\mathbb{S}^2$ is obtained by mapping low-discrepancy points on $[0,1]^2$ to $\mathbb{S}^2$ using an equal area transformation $\boldsymbol{\Phi} : [0,1]^2 \to \mathbb{S}^2$. The same approach has previously been used in [12] and [27], in both cases in the context of numerical integration. The low-discrepancy points in $[0,1]^2$ are obtained from digital nets and Fibonacci lattices, see [20, 42]. These point sets are well-distributed with respect to rectangles anchored at the origin $(0,0)$. However, the set

$$\boldsymbol{\Phi}^{-1}\big(C(\boldsymbol{w}, t)\big) = \big\{\boldsymbol{x} \in [0,1]^2 : \boldsymbol{\Phi}(\boldsymbol{x}) \in C(\boldsymbol{w}, t)\big\}$$

is, in general, not a rectangle. In fact, it is not even a convex set, although the boundary

$$\boldsymbol{\Phi}^{-1}\big(\partial C(\boldsymbol{w},t)\big)=\big\{\boldsymbol{x}\in[0,1]^2:\boldsymbol{\Phi}(\boldsymbol{x})\in\partial C(\boldsymbol{w},t)\big\}$$

is a continuous curve.

Hence, in order to prove bounds on the spherical cap discrepancy of digital nets and Fibonacci lattices lifted to the sphere using $\boldsymbol{\Phi}$ (we call those point sets spherical digital nets and spherical Fibonacci lattices), we need to prove bounds on a general notion of discrepancy in $[0,1]^2$. To this end we study discrepancy in $[0,1]^2$ with respect to convex sets, the corresponding discrepancy being known as *isotropic discrepancy* [8]. We show that digital nets and Fibonacci lattices have isotropic discrepancy of order $\mathcal{O}(N^{-1/2})$. Using these result and some properties of the function $\boldsymbol{\Phi}$, we can show that spherical digital nets and spherical Fibonacci lattices have spherical cap discrepancy at most $CN^{-1/2}$ for an explicitly given constant C, see Corollaries 16 and 18. Note that the best possible rate of convergence of the isotropic discrepancy is $N^{-2/3}(\log N)^c$ for some $0\le c\le 4$, see [7] and [8, p. 107]. Hence the approach via the isotropic discrepancy cannot give the optimal rate of convergence for the spherical cap discrepancy.

In the following we define the equal area Lambert map $\boldsymbol{\Phi}$ and show some of its properties.

2 The Equal-Area Lambert Transform and Some Properties

The points on the sphere are obtained by using the *Lambert cylindrical equal-area projection*

$$\boldsymbol{\Phi}(\alpha,\tau)=\big(2\sqrt{\tau-\tau^2}\cos(2\pi\alpha),2\sqrt{\tau-\tau^2}\sin(2\pi\alpha),1-2\tau\big),\quad \alpha,\tau\in[0,1]. \tag{7}$$

The area-preserving Lambert map can be illustrated in the following way. The unit square $[0,1]^2$ is linearly stretched to the rectangle $[0,2\pi]\times[-1,1]$, rolled into a cylinder of radius 1 and height 2 and fitted around the unit sphere such that the polar axis is the main z-axis. This way a point (α,τ) in $[0,1]^2$ is mapped to a point on the cylinder which is radially projected along a ray orthogonal to the polar axis onto the sphere giving the point $\boldsymbol{\Phi}(\alpha,\tau)$.

Axis-parallel rectangles in the unit square are mapped to spherical "rectangles" of equal area, see Fig. 1.

The pre-image of a spherical cap centred at $\boldsymbol{w}$ with height t under the Lambert map is the set

$$B(\boldsymbol{w},t)=\boldsymbol{\Phi}^{-1}\big(C(\boldsymbol{w},t)\big)=\big\{(\alpha,\tau)\in[0,1)\times[0,1]:\boldsymbol{\Phi}(\alpha,\tau)\in C(\boldsymbol{w},t)\big\}$$

and the pre-image of the boundary of this spherical cap is

$$\partial B(\boldsymbol{w},t)=\boldsymbol{\Phi}^{-1}\big(\partial C(\boldsymbol{w},t)\big)=\big\{(\alpha,\tau)\in[0,1)\times[0,1]:\boldsymbol{\Phi}(\alpha,\tau)\in\partial C(\boldsymbol{w},t)\big\}.$$

The sets $B(\boldsymbol{w},t)$ are not convex, in general. Thus, we consider a more general class of sets which we call pseudo-convex. A definition is given in the following.

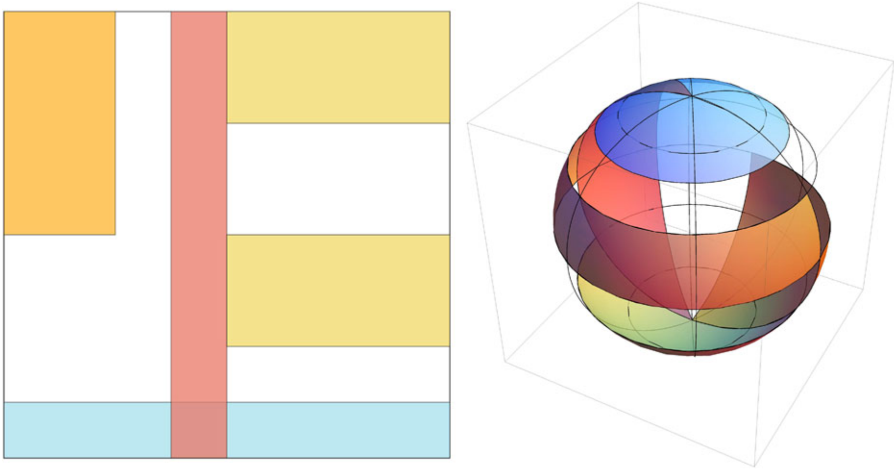


Fig. 1 Axis-parallel rectangles in the square and their images under $\boldsymbol{\Phi}$ on $\mathbb{S}^2$

Definition 2 Let A be an open subset of $[0,1]^2$ such that there exists a collection of p convex subsets $A_1,\dots,A_p$ of $[0,1]^2$ with the following properties: (a) $A_j \cap A_k$ is empty for $j \neq k$; (b) $A \subseteq A_1 \cup \cdots \cup A_p$; (c) either A_j is a convex part of A ($A_j \subseteq A$) or the complement of A with respect to A_j, $A_j' = A_j \setminus A$, is convex. Then A is called a *pseudo-convex* set and $A_1,\dots,A_p$ is an admissible convex covering for A with p parts (with q convex parts of A).

Lemma 3 *For every $\boldsymbol{w} \in \mathbb{S}^2$ and all $-1 \leq t \leq 1$, the pre-image $B(\boldsymbol{w},t)$ of the spherical cap $C(\boldsymbol{w},t)$ centred at $\boldsymbol{w}$ with height t under the Lambert map is pseudo-convex with an admissible convex covering with at most 7 parts. More precisely, taking into account the number of convex parts of the pre-image, among the convex coverings with p parts and q of which are convex, the worst case has $p=7$ and $q=3$ which implies the constant $2p-q=11$.*

The proof of Lemma 3 in Sect. 7 gives details how to construct admissible coverings.

3 Isotropic- and Spherical Cap Discrepancy

We introduce the isotropic discrepancy of a point set and a sequence as follows. Let λ be the Lebesgue area measure in the unit square.

Definition 4 The *isotropic discrepancy* J_N of an N-point set $P_N = \{\boldsymbol{x}_0,\dots,\boldsymbol{x}_{N-1}\}$ in $[0,1]^2$ is defined as

$$J_N(P_N) = \sup_{A \in \mathcal{A}} \left| \frac{1}{N} \sum_{n=0}^{N-1} 1_A(\boldsymbol{x}_n) - \lambda(A) \right|,$$

where $\mathcal{A}$ is the family of all convex subsets of $[0,1]^2$.

For an infinite sequence $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots \in [0,1]^2$ the isotropic discrepancy is defined as the isotropic discrepancy of the initial N points of the sequence.

Lemma 5 *Let A be a pseudo-convex subset of $[0,1]^2$ with an admissible convex covering of p parts with q convex parts of A. Then for any N-point set $P_N = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\} \subseteq [0,1]^2$,*

$$\left| \frac{1}{N} \sum_{n=0}^{N-1} 1_A(\boldsymbol{x}_n) - \lambda(A) \right| \leq (2p-q) J_N(P_N).$$

Proof Let $A_1, \ldots, A_p$ be an admissible convex covering of A with p parts. Without loss of generality, let $A_1, \ldots, A_q$ be the convex parts of A and $A_{q+1}, \ldots, A_p$ those for which $A_j' = A_j \setminus A$ $(q+1 \leq j \leq p)$ is convex. Clearly,

$$A = \bigcup_{j=1}^{q} A_j \cup \bigcup_{j=q+1}^{p} \bigl(A_j \setminus A_j'\bigr).$$

Thus,

$$\begin{aligned}
&\frac{1}{N} \sum_{n=0}^{N-1} 1_A(\boldsymbol{x}_n) - \lambda(A) \\
&\quad = \sum_{j=1}^{q} \left[\frac{1}{N} \sum_{n=0}^{N-1} 1_{A_j}(\boldsymbol{x}_n) - \lambda(A_j) \right] + \sum_{j=q+1}^{p} \left[\frac{1}{N} \sum_{n=0}^{N-1} 1_{A_j \setminus A_j'}(\boldsymbol{x}_n) - \lambda\bigl(A_j \setminus A_j'\bigr) \right] \\
&\quad = \sum_{j=1}^{q} \left[\frac{1}{N} \sum_{n=0}^{N-1} 1_{A_j}(\boldsymbol{x}_n) - \lambda(A_j) \right] + \sum_{j=q+1}^{p} \left[\frac{1}{N} \sum_{n=0}^{N-1} 1_{A_j}(\boldsymbol{x}_n) - \lambda(A_j) \right] \\
&\qquad - \sum_{j=q+1}^{p} \left[\frac{1}{N} \sum_{n=0}^{N-1} 1_{A_j'}(\boldsymbol{x}_n) - \lambda\bigl(A_j'\bigr) \right].
\end{aligned}$$

In the last line all sets are convex and we can use the isotropic discrepancy in the estimation

$$\left| \frac{1}{N} \sum_{n=0}^{N-1} 1_A(\boldsymbol{x}_n) - \lambda(A) \right| \leq \bigl[q + (p-q) + (p-q)\bigr] J_N(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}).$$ □

Theorem 6 *Let $P_N = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\} \subseteq [0,1]^2$ and let $Z_N = \{\boldsymbol{\Phi}(\boldsymbol{x}_0), \ldots, \boldsymbol{\Phi}(\boldsymbol{x}_{N-1})\} \subseteq \mathbb{S}^2$. Then*

$$D(Z_N) \leq 11 J_N(P_N).$$

Proof Let $\boldsymbol{w} \in \mathbb{S}^2$ and $-1 \le t \le 1$. A point $\boldsymbol{\Phi}(\boldsymbol{x}_n) \in C(\boldsymbol{w}, t)$ if and only if $\boldsymbol{x}_n \in B(\boldsymbol{w}, t)$. Thus,

$$\sum_{n=0}^{N-1} 1_{C(\boldsymbol{w},t)}\big(\boldsymbol{\Phi}(\boldsymbol{x}_n)\big) = \sum_{n=0}^{N-1} 1_{B(\boldsymbol{w},t)}(\boldsymbol{x}_n).$$

Further, since the transformation $\boldsymbol{\Phi}$ preserves areas, we have

$$\sigma\big(C(\boldsymbol{w}, t)\big) = \lambda\big(B(\boldsymbol{w}, t)\big).$$

Hence,

$$\left|\frac{1}{N}\sum_{n=0}^{N-1} 1_{C(\boldsymbol{w},t)}\big(\boldsymbol{\Phi}(\boldsymbol{x}_n)\big) - \sigma\big(C(\boldsymbol{w}, t)\big)\right| = \left|\frac{1}{N}\sum_{n=0}^{N-1} 1_{B(\boldsymbol{w},t)}(\boldsymbol{x}_n) - \lambda\big(B(\boldsymbol{w}, t)\big)\right|.$$

The pre-images are pseudo-convex in the sense of Definition 2 by Lemma 3. Applying Lemma 5 with the constant $2p - q = 11$ from Lemma 3 we arrive at the result. □

We have now reduced the problem of proving bounds on the spherical cap discrepancy to prove bounds on the isotropic discrepancy of points in the square $[0, 1]^2$. We will study this problem in Sect. 5.

4 Spherical Cap Discrepancy of Random Points Sets

Let $(M, \mathcal{M})$ be a measurable space, and let P be a probability on it. Let further X_n, $n \ge 0$, denote a sequence of independent, identically distributed (i.i.d.) random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in M, and let $\mathcal{C} \subseteq \mathcal{M}$ denote a class of subsets of M. To avoid measurability problems we will assume throughout the rest of this section that the class $\mathcal{C}$ is countable. Let $A \subseteq M$ be an arbitrary set. Then $\mathcal{C}$ is said to *shatter* A if to every possible subset B of A there exists a set $C \in \mathcal{C}$ such that

$$C \cap A = B.$$

For $k \ge 1$ the kth *shattering coefficient* $S_{\mathcal{C}}(k)$ of $\mathcal{C}$ is defined as

$$S_{\mathcal{C}}(k) := \max_{x_1,\dots,x_k \in M} \operatorname{card}\big\{\{x_1, \dots, x_k\} \cap C : C \in \mathcal{C}\big\}.$$

The *Vapnik–Červonenkis dimension* (VC-dimension) of $\mathcal{C}$ is defined as

$$v(\mathcal{C}) := \min_k \big\{k : S_{\mathcal{C}} < 2^k\big\}.$$

(Here we use the convention that the minimum of the empty set is ∞.) A class $\mathcal{C}$ with finite VC-dimension is called a *Vapnik–Červonenkis class* (VC class). The theory of VC classes is of extraordinary importance in the theory of empirical processes indexed by classes of functions. For example, a class $\mathcal{C}$ is *uniformly Glivenko–Cantelli* if and only if it is a VC class, see [53]. We will use the following theorem, which is a combination of results of Talagrand [51, Theorem 6.6] and Haussler [30, Corollary 1], and has already been used by Heinrich et al. [31] in the context of probabilistic discrepancy theory.

Theorem 7 (See [31, Theorem 2]) *There exists a positive number K such that, for each VC class $\mathcal{C}$ and each probability P and sequence X_n, $n \geq 0$, as above, the following holds: For all $s \geq K\sqrt{v(\mathcal{C})}$ we have*

$$\mathbb{P}\left\{\sup_{C\in\mathcal{C}}\left|\frac{1}{N}\sum_{n=0}^{N-1} 1_C(X_n) - P(C)\right| \geq \frac{s}{\sqrt{N}}\right\} \leq \frac{1}{s}\left(\frac{Ks^2}{v(\mathcal{C})}\right)^{v(\mathcal{C})} e^{-2s^2}.$$

In our setting we will have $M = \mathbb{S}^2$, $\mathcal{M}$ will denote the sigma-field generated by the class of spherical caps, P will stand for the normalised Lebesgue surface area measure σ, and $\mathcal{C}$ will denote the class of all spherical caps for which the centre $\boldsymbol{w}$ is a vector of rational numbers and the height t is also a rational number (this restriction is necessary to assure that the class $\mathcal{C}$ is countable; of course, the spherical cap discrepancy with respect to this class is the same as the discrepancy with respect to the class of *all* spherical caps). In the sequel we assume that the i.i.d. random variables X_n, $n \geq 0$, are uniformly distributed on $\mathbb{S}^2$. We will write $Z_N = Z_N(\omega)$ for the (random) point set $\{X_0, \dots, X_{N-1}\} = \{X_0(\omega), \dots, X_{N-1}(\omega)\}$.

The following proposition asserts that the class $\mathcal{C}$ is a VC class (the proof of this and the subsequent results of this section can be found in Sect. 7).

Proposition 8 *The class $\mathcal{C}$ has VC dimension* 5.

Using Theorem 7 and Proposition 8 we can prove the following results:

Theorem 9 *There exist constants C_1, C_2 such that for $N \geq 1$,*

$$\frac{C_1}{\sqrt{N}} \leq \mathbb{E}\big[D(Z_N)\big] \leq \frac{C_2}{\sqrt{N}}.$$

Remark The existence of such a constant C_1 for the lower bound follows directly from (5); we can choose $C_1 = 6^{-1/2}$.

Theorem 10 *For any $\varepsilon > 0$ there exist positive constants $C_3(\varepsilon), C_4(\varepsilon)$ such that for sufficiently large N,*

$$\mathbb{P}\big\{C_3 \leq \sqrt{N} D(Z_N) \leq C_4\big\} \geq 1 - \varepsilon.$$

Theorem 10 shows that the *typical* discrepancy of a random set of N points is of order $N^{-1/2}$. However, actually much more is true, since by classical results any VC class $\hat{\mathcal{C}}$ on a measurable space $(\hat{M}, \hat{\mathcal{M}})$ is a so-called *Donsker class*, which essentially means that for every probability measure $\hat{P}$ and every sequence V_n, $n \geq 0$, of i.i.d.

random variables having law $\hat{P}$ the *empirical process indexed by sets*

$$\alpha_N(C) = \sqrt{N}\left|\frac{1}{N}\sum_{n=0}^{N-1} 1_C(V_n) - P(C)\right|, \quad C \in \hat{\mathcal{C}},$$

converges weakly to a centred, bounded Gaussian process $B(C)$, which has covariance structure

$$\mathbb{E}B(C_1)B(C_2) = P(C_1 \cap C_2) - P(C_1)P(C_2), \quad C_1, C_2 \in \hat{\mathcal{C}}.$$

This weak convergence could, for example, be used to prove the existence of a limit distribution of $\sqrt{N}D(Z_N)$ as $N \to \infty$; however, to keep this presentation short and self-contained we will not pursue this method any further, and refer the interested reader to [2, 21, 22, 52] and the references therein.

Remark The upper bounds in Theorems 9 and 10 follow from Theorem 7. However, since no concrete value for the constant K in Theorem 7 is known, the value of the constants C_2 and C_4 in Theorem 9 and Theorem 10, respectively, is also unknown. It is possible that the decomposition technique from [1] can be used to achieve a version of Theorems 9 and 10 with explicitly known constants in the upper bound.

Finally, the following theorem describes the asymptotic order of a *typical* infinite sequence of random points.

Theorem 11 *We have*

$$D(Z_N) = \mathcal{O}\left(\frac{\sqrt{\log\log N}}{\sqrt{N}}\right) \quad \text{as } N \to \infty, \text{ almost surely}.$$

Theorem 11 is a so-called *bounded law of the iterated logarithm*, and follows easily from Theorem 6 and Philipp's law of the iterated logarithm (LIL) for the isotropic discrepancy of random point sets in the plane. More precisely, Philipp [44] proved that for a sequence of i.i.d. uniformly distributed random variables Y_n, $n \geq 0$, on the unit square (writing P_N for the (random) point set $\{Y_0, \dots, Y_{N-1}\}$), the law of the iterated logarithm

$$\limsup_{N\to\infty} \frac{N J_N(P_N)}{\sqrt{2N\log\log N}} = \frac{1}{2} \quad \text{a.s.}$$

holds. Together with Theorem 6 this implies for $Z_N = \{\boldsymbol{\Phi}(Y_0), \dots, \boldsymbol{\Phi}(Y_{N-1})\} \subseteq \mathbb{S}^2$ that

$$\limsup_{N\to\infty} \frac{N D(Z_N)}{\sqrt{2N\log\log N}} \leq \frac{11}{2} \quad \text{a.s.},$$

which proves Theorem 11 (it is necessary to observe that the image of a sequence of i.i.d. uniformly distributed random variables on the unit square under the area-preserving Lambert map is a sequence of i.i.d. uniformly distributed random variables on the sphere). It is easy to see that Theorem 11 is optimal, except for the value of the implied constant. More precisely, let C^* denote a fixed spherical cap with area 2π (which means that C^* is a hemisphere, and has normalised surface area measure $\sigma(C^*) = 1/2$). Then clearly the random variables

$$1_{C^*}\big(\boldsymbol{\Phi}(Y_n)\big) - \sigma\big(C^*\big), \quad n \geq 0,$$

have expected value 0 and variance 1/4. Thus, by the classical law of the iterated logarithm for sequences of i.i.d. random variables,

$$\limsup_{N\to\infty} \frac{|\frac{1}{N}\sum_{n=0}^{N-1} 1_{C^*}(\boldsymbol{\Phi}(Y_n)) - 1/2|}{\sqrt{2N\log\log N}} = \frac{1}{2} \quad \text{a.s.},$$

and since

$$D(Z_N) \geq \left| \frac{1}{N} \sum_{n=0}^{N-1} 1_{C^*}\big(\boldsymbol{\Phi}(Y_n)\big) - \sigma\big(C^*\big) \right|,$$

we finally arrive at

$$\limsup_{N\to\infty} \frac{ND(Z_N)}{\sqrt{2N\log\log N}} \geq \frac{1}{2} \quad \text{a.s.},$$

which proves the optimality of Theorem 11. We remark that it should also be possible to prove Theorem 11 without using Theorem 6 and Philipp's LIL for the isotropic discrepancy, by deducing it directly from the bounded LIL for empirical processes on VC classes of Alexander and Talagrand [3]. We conjecture that Theorem 11 can be improved to

$$\limsup_{N\to\infty} \frac{ND(Z_N)}{\sqrt{2N\log\log N}} = \frac{1}{2} \quad \text{a.s.},$$

but this seems to be very difficult to prove.

5 Point Sets with Small Isotropic Discrepancy

In this section we investigate the isotropic discrepancy of $(0, m, 2)$-nets and Fibonacci lattices. In particular we show that the isotropic discrepancy of those point sets converges with order $\mathcal{O}(N^{-1/2})$. Note that the best possible rate of convergence of the isotropic discrepancy is $N^{-2/3}(\log N)^c$ for some $0 \leq c \leq 4$, see [7] and [8, p. 107]. Whether $(0, m, 2)$-nets and/or Fibonacci lattices achieve the optimal rate of convergence for the isotropic discrepancy is an open question.

5.1 Nets and Sequences

We give the definition of $(0, m, 2)$-nets in base b in the following.

Definition 12 Let $b \geq 2$ and $m \geq 1$ be integers. A point set $P_{b^m} \subseteq [0, 1)^2$ consisting of b^m points is called a $(0, m, 2)$-*net in base* b, if for all nonnegative integers d_1, d_2

with $d_1 + d_2 = m$, each of the elementary intervals

$$\prod_{i=1}^{2}\left[\frac{a_i}{b^{d_i}}, \frac{a_i+1}{b^{d_i}}\right), \quad 0 \le a_i < b^{d_i} \ (a_i \text{ an integer}),$$

contains exactly 1 point of P_{b^m}.

It is also possible to construct nested $(0, m, 2)$-nets, thereby obtaining an infinite sequence of points.

Definition 13 Let $b \ge 2$ be an integer. A sequence $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots \in [0,1)^2$ is called a $(0,2)$-*sequence in base* b, if for all $m > 0$ and for all $k \ge 0$, the point set $\boldsymbol{x}_{kb^m}, \boldsymbol{x}_{kb^m+1}, \ldots, \boldsymbol{x}_{(k+1)b^m-1}$ is a $(0, m, 2)$-net in base b.

Explicit constructions of $(0, m, 2)$-nets and $(0, 2)$-sequences are due to Sobol' [49] and Faure [23], see also [20, Chap. 8].

The following is a special case of an unpublished result due to Gerhard Larcher. For completeness we include a proof here.

Theorem 14 *For the isotropic discrepancy J_N of a $(0, m, 2)$-net P_N in base b ($N = b^m$) we have*

$$J_N(P_N) \le 4\sqrt{2}b^{-\lfloor m/2\rfloor} \le \frac{4\sqrt{2b}}{\sqrt{N}}.$$

Proof Let $P_N = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{b^m-1}\}$. Let $k = \lfloor m/2\rfloor$ and consider a subcube W of $[0,1)^s$ of the form

$$W = \left[\frac{c_1}{b^k}, \frac{c_1+1}{b^k}\right) \times \left[\frac{c_2}{b^k}, \frac{c_2+1}{b^k}\right)$$

with $0 \le c_i < b^k$ (c_i an integer) for $i = 1, 2$. The cube W has volume b^{-2k} and is the union of b^{m-2k} elementary intervals of order m. Indeed,

$$W = \bigcup_{v=0}^{b^{m-2k}-1}\left(\left[\frac{c_1}{b^k}, \frac{c_1+1}{b^k}\right) \times \left[\frac{c_2}{b^k} + \frac{v}{b^{m-k}}, \frac{c_2}{b^k} + \frac{v+1}{b^{m-k}}\right)\right).$$

So W contains exactly b^{m-2k} points of the net. The diagonal of W has length $\sqrt{2}/b^k$.

Let now A be an arbitrary convex subset of $[0,1]^2$. Let W° denote the union of cubes W fully contained in A and let $\overline{W}$ denote the union of cubes W having nonempty intersection with A or its boundary. The sets $\overline{W}$ and W° are fair with respect to the net, that is,

$$\frac{1}{N}\sum_{n=0}^{N-1} 1_{\overline{W}}(\boldsymbol{x}_n) = \lambda(\overline{W}) \quad \text{and} \quad \frac{1}{N}\sum_{n=0}^{N-1} 1_{W^\circ}(\boldsymbol{x}_n) = \lambda\bigl(W^\circ\bigr).$$

We have

$$\frac{1}{N}\sum_{n=0}^{N-1} 1_A(\boldsymbol{x}_n) - \lambda(A) \le \frac{1}{N}\sum_{n=0}^{N-1} 1_{\overline{W}}(\boldsymbol{x}_n) - \lambda(\overline{W}) + \lambda(\overline{W}\setminus A) = \lambda(\overline{W}\setminus A)$$

and

$$\frac{1}{N}\sum_{n=0}^{N-1} 1_A(\boldsymbol{x}_n) - \lambda(A) \ge \frac{1}{N}\sum_{n=0}^{N-1} 1_{W^\circ}(\boldsymbol{x}_n) - \lambda\bigl(W^\circ\bigr) - \lambda\bigl(A\setminus W^\circ\bigr) = -\lambda\bigl(A\setminus W^\circ\bigr).$$

Since the set A is convex, the length of the boundary of A is at most the circumference of the unit square, which is 4. Further we have

$$\overline{W}\setminus A \subseteq \bigl\{\boldsymbol{x}\in[0,1]^2\setminus A : \|\boldsymbol{x}-\boldsymbol{y}\| \le \sqrt{2}b^{-k} \text{ for some } \boldsymbol{y}\in A\bigr\}$$

and therefore

$$\lambda(\overline{W}\setminus A) \le \lambda\bigl(\bigl\{\boldsymbol{x}\in[0,1]^2\setminus A : \|\boldsymbol{x}-\boldsymbol{y}\| \le \sqrt{2}b^{-k} \text{ for some } \boldsymbol{y}\in A\bigr\}\bigr) \le 4\sqrt{2}b^{-k},$$

where the last inequality follows from the fact that the outer boundary of the enclosing set has length at most 4 (which is the circumference of the square $[0,1]^2$). Moreover,

$$A\setminus W^\circ \subseteq \bigl\{\boldsymbol{x}\in A : \|\boldsymbol{x}-\boldsymbol{y}\| \le \sqrt{2}b^{-k} \text{ for some } \boldsymbol{y}\in[0,1]^2\setminus A\bigr\}$$

and therefore

$$\lambda\bigl(A\setminus W^\circ\bigr) \le \lambda\bigl(\bigl\{\boldsymbol{x}\in A : \|\boldsymbol{x}-\boldsymbol{y}\| \le \sqrt{2}b^{-k} \text{ for some } \boldsymbol{y}\in[0,1]^2\setminus A\bigr\}\bigr) \le 4\sqrt{2}b^{-k},$$

since, by the convexity of A, the boundary of A has length at most 4 (which is the circumference of the square $[0,1]^2$).

Thus we obtain

$$\left|\frac{1}{N}\sum_{n=0}^{N-1} 1_A(\boldsymbol{x}_n) - \lambda(A)\right| \le 4\sqrt{2}b^{-k}$$

and hence the result follows. □

Note that the above result only applies when the number of points N is of the form $N=b^m$ (notice that choosing $m=1$ only yields a trivial result, hence one usually chooses a small base b and a 'large' value of m). In the following we give an extension where the number of points can take on arbitrary positive integers.

Theorem 15 *For the isotropic discrepancy J_N of the first N points $P_N = \{\boldsymbol{x}_0,\dots,\boldsymbol{x}_{N-1}\}$ of a $(0,2)$-sequence in base b, we have*

$$J_N(P_N) \le 4\sqrt{2}\bigl(b^2+b^{3/2}\bigr)\frac{1}{\sqrt{N}}.$$

Proof Let $N \in \mathbb{N}$ have base b expansion $N = N_0 + N_1 b + \cdots + N_m b^m$. Let $P_N = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$ denote the first N points of a $(0,2)$-sequence in base b. For $0 \le k \le m$ with $N_k > 0$ and $0 \le \ell < N_k$ the point set

$$Q_{k,\ell} = \{\boldsymbol{x}_{N_m b^m + \cdots + N_{k+1} b^{k+1} + \ell b^k}, \ldots, \boldsymbol{x}_{N_m b^m + \cdots + N_{k+1} b^{k+1} + (\ell+1) b^k - 1}\}$$

is a $(0,k,2)$-net in base b by Definition 13. Thus P_N is a disjoint union of such $(0,k,2)$-nets:

$$P_N = \bigcup_{\substack{0 \le k \le m \\ N_k > 0}} \bigcup_{0 \le \ell < N_k} Q_{k,\ell}.$$

We have the following triangle inequality for the isotropic discrepancy (which is an analogue to the triangle inequality for the star-discrepancy [34, p. 115, Theorem 2.6]):

$$J_N(P_N) \le \sum_{\substack{k=0 \\ N_k > 0}}^{m} \sum_{\ell=0}^{N_k - 1} \frac{b^k}{N} J_{b^k}(Q_{k,\ell}).$$

This inequality holds, since for a spherical cap C we have

$$\frac{1}{N} \sum_{n=0}^{N-1} 1_C(\boldsymbol{x}_n) - \lambda(C) = \sum_{\substack{k=0 \\ N_k > 0}}^{m} \sum_{\ell=0}^{N_k - 1} \frac{b^k}{N} \left(\frac{1}{b^k} \sum_{\boldsymbol{x} \in Q_{k,\ell}} 1_C(\boldsymbol{x}) - \lambda(C) \right).$$

Thus we can use Theorem 14 to obtain

$$\begin{aligned} J_N(P_N) &\le \sum_{\substack{k=0 \\ N_k > 0}}^{m} \sum_{\ell=0}^{N_k - 1} \frac{b^k}{N} 4\sqrt{2} b^{-\lfloor k/2 \rfloor} = 4\sqrt{2} \sum_{\substack{k=0 \\ N_k > 0}}^{m} \sum_{\ell=0}^{N_k - 1} \frac{b^{\lceil k/2 \rceil}}{N} = 4\sqrt{2} \frac{\sum_{k=0}^m N_k b^{\lceil k/2 \rceil}}{\sum_{k=0}^m N_k b^k} \\ &\le 4\sqrt{2}(b-1) \sum_{k=0}^{m} \frac{b^{\lceil k/2 \rceil}}{b^m} \le 4\sqrt{2} \frac{b(b-1)}{\sqrt{b}-1} b^{-m/2} \le 4\sqrt{2} \frac{b^{3/2}(b-1)}{\sqrt{b}-1} \frac{1}{\sqrt{N}}. \end{aligned}$$

The estimate follows from the identity $(a-1)(a+1) = a^2 - 1$. □

Corollary 16

(1) *Let* P_N *be a* $(0,m,2)$*-net in base* b *and let* $Z_N = \boldsymbol{\Phi}(P_N) \subseteq \mathbb{S}^2$*. Then the spherical cap discrepancy* $D(Z_N)$ *is bounded by*

$$D(Z_N) \le 44\sqrt{2} b^{-\lfloor m/2 \rfloor}.$$

(2) *Let* P_N *be the first* N *points of a* $(0,2)$*-sequence in base* b *and let* $Z_N = \boldsymbol{\Phi}(P_N) \subseteq \mathbb{S}^2$*. Then the spherical cap discrepancy* $D(Z_N)$ *is bounded by*

$$D(Z_N) \le 44\sqrt{2}\big(b^2 + b^{3/2}\big) \frac{1}{\sqrt{N}} \quad \textit{for all } N.$$

Table 1 Comparison of the effect of two different normalisations of the value $\widetilde{D}(Z_N)$ computing $\max_{\boldsymbol{w}\in Z_N} \delta(Z_N;\boldsymbol{w})$, cf. Definition 1, for a digital net Z_N based on a two-dimensional Sobol' point set

m	6	7	8	9	10	11	12	13
$N=2^m$	64	128	256	512	1024	2048	4096	8192
$\frac{N^{3/4}\widetilde{D}(Z_N)}{\sqrt{\log N}}$	0.8829	0.8436	0.8279	0.8632	0.8518	1.2128	1.2285	0.9546
$\frac{N^{3/4}\widetilde{D}(Z_N)}{\log N}$	0.4329	0.3829	0.3515	0.3456	0.3235	0.4392	0.4259	0.3180
m	14	15	16	17	18	19	20	21
$N=2^m$	16384	32768	65536	131072	262144	524288	1048576	2097152
$\frac{N^{3/4}\widetilde{D}(Z_N)}{\sqrt{\log N}}$	0.7925	0.8862	1.0331	0.8337	0.8562	0.9854	1.1167	1.1463
$\frac{N^{3/4}\widetilde{D}(Z_N)}{\log N}$	0.2544	0.2748	0.3102	0.2428	0.2424	0.2715	0.2999	0.3004

Note that item (2) improves upon Theorem 11 by a factor of $\sqrt{\log\log N}$ and hence, asymptotically, spherical digital sequences are better than random sequences almost surely.

The numerical experiments shown in Table 1 seem to suggest that the correct order of the spherical cap discrepancy of spherical digital nets is

$$\frac{(\log N)^c}{N^{3/4}} \quad \text{for some } 1/2 \le c \le 1.$$

In those experiments we calculated an approximation from below of the spherical cap discrepancy of the members of a sequence of spherical digital nets Z_N based on two-dimensional Sobol' point sets by explicit numerical computation of

$$\widetilde{D}(Z_N) = \max_{\boldsymbol{w}\in Z_N} \delta(Z_N;\boldsymbol{w}).$$

5.2 Fibonacci Lattices

The Fibonacci numbers F_m are given by $F_1=1$, $F_2=1$ and $F_m=F_{m-1}+F_{m-2}$ for all $m>2$. A Fibonacci lattice is a point set of F_m points in $[0,1)^2$ given by

$$\boldsymbol{f}_m := \left(\frac{n}{F_m}, \left\{\frac{nF_{m-1}}{F_m}\right\}\right), \quad 0\le n<F_m,$$

where $\{x\}=x-\lfloor x\rfloor$ denotes the fractional part for nonnegative real numbers x. The set

$$\mathcal{F}_m := \{\boldsymbol{f}_0,\dots,\boldsymbol{f}_{F_m-1}\}$$

is called a Fibonacci lattice point set.

The spherical Fibonacci lattice points are then given by

$$z_n = \boldsymbol{\Phi}(\boldsymbol{f}_n), \quad 0\le n<F_m,$$

and the point set

$$Z_{F_m} = \{z_0, \ldots, z_{F_m-1}\}$$

is the spherical Fibonacci lattice point set.

In the following we prove a bound on the isotropic discrepancy of Fibonacci lattices, see also [35–37].

Lemma 17 *For the isotropic discrepancy J_{F_m} of a Fibonacci lattice $\mathcal{F}_m$ we have*

$$J_{F_m}(\mathcal{F}_m) \le \begin{cases} 4\sqrt{2/F_m} & \text{if } m \text{ is odd,} \\ 4\sqrt{8/F_m} & \text{if } m \text{ is even.} \end{cases}$$

Proof Consider the case of odd integers m first. From [43, Theorem 3] it follows that for $m \in \mathbb{N}$ the Fibonacci lattice $\mathcal{F}_{2m+1}$ can be generated by the vectors

$$\begin{aligned} \boldsymbol{a}_{2m+1} &= \big(F_m/F_{2m+1}, (-1)^{m-1} F_{m+1}/F_{2m+1}\big), \\ \boldsymbol{b}_{2m+1} &= \big(F_{m+1}/F_{2m+1}, (-1)^{m} F_m/F_{2m+1}\big). \end{aligned}$$

This means that

$$\mathcal{F}_{2m+1} = \{u\boldsymbol{a}_{2m+1} + v\boldsymbol{b}_{2m+1} : u, v \in \mathbb{Z}\} \cap [0,1)^2.$$

Let

$$U(\boldsymbol{y}) = \big\{\boldsymbol{y} + \boldsymbol{x} \in [0,1]^2 : \boldsymbol{x} = s\boldsymbol{a}_{2m+1} + t\boldsymbol{b}_{2m+1}, 0 \le s, t < 1\big\}.$$

We call $U(\boldsymbol{f}_n)$ a unit cell (belonging to the point $\boldsymbol{f}_n$). Note that the area of a unit cell is $1/F_{2m+1}$ and each unit cell contains exactly one point of the lattice, see [43].

Since $\boldsymbol{a}_{2m+1} \perp \boldsymbol{b}_{2m+1}$, it follows that the minimum distance between points of the Fibonacci lattice is

$$d_{\min}(\mathcal{F}_{2m+1}) = \min\big\{\|\boldsymbol{a}_{2m+1}\|, \|\boldsymbol{b}_{2m+1}\|\big\} = \frac{\sqrt{F_m^2 + F_{m+1}^2}}{F_{2m+1}} = \frac{1}{\sqrt{F_{2m+1}}}.$$

Thus the diameter of a unit cell is $\sqrt{2/F_{2m+1}}$.

Let now A be an arbitrary convex subset of $[0,1]^2$. Let W° denote the union of all unit cells fully contained in A and let $\overline{W}$ denote the union of all unit cells with nonempty intersection with A or its boundary.

We have

$$\begin{aligned} \frac{1}{F_{2m+1}} \sum_{n=0}^{F_{2m+1}-1} 1_A(\boldsymbol{f}_n) - \lambda(A) &\le \frac{1}{F_{2m+1}} \sum_{n=0}^{F_{2m+1}-1} 1_{\overline{W}}(\boldsymbol{f}_n) - \lambda(\overline{W}) + \lambda(\overline{W} \setminus A) \\ &= \lambda(\overline{W} \setminus A) \end{aligned}$$

and

$$\frac{1}{F_{2m+1}}\sum_{n=0}^{F_{2m+1}-1} 1_A(\boldsymbol{f}_n)-\lambda(A)\ge \frac{1}{F_{2m+1}}\sum_{n=0}^{F_{2m+1}-1} 1_{W^\circ}(\boldsymbol{f}_n)-\lambda\bigl(W^\circ\bigr)-\lambda\bigl(A\setminus W^\circ\bigr)$$
$$=-\lambda\bigl(A\setminus W^\circ\bigr).$$

Since the set A is convex, the length of the boundary of A is at most the circumference of the unit square, which is 4. Further we have

$$\overline{W}\setminus A\subseteq\Bigl\{\boldsymbol{x}\in[0,1]^2\setminus A:\|\boldsymbol{x}-\boldsymbol{y}\|\le\sqrt{2/F_{2m+1}}\text{ for some }\boldsymbol{y}\in A\Bigr\}$$

and therefore

$$\begin{aligned}\lambda(\overline{W}\setminus A)&\le\lambda\Bigl(\Bigl\{\boldsymbol{x}\in[0,1]^2\setminus A:\|\boldsymbol{x}-\boldsymbol{y}\|\le\sqrt{2/F_{2m+1}}\text{ for some }\boldsymbol{y}\in A\Bigr\}\Bigr)\\&\le 4\sqrt{2/F_{2m+1}},\end{aligned}$$

where the last inequality follows since the outer boundary of enclosing set has length at most 4 (which is the circumference of the square $[0,1]^2$).

On the other hand,

$$A\setminus W^\circ\subseteq\Bigl\{\boldsymbol{x}\in A:\|\boldsymbol{x}-\boldsymbol{y}\|\le\sqrt{2/F_{2m+1}}\text{ for some }\boldsymbol{y}\in[0,1]^2\setminus A\Bigr\}$$

and therefore

$$\begin{aligned}\lambda\bigl(A\setminus W^\circ\bigr)&\le\lambda\Bigl(\Bigl\{\boldsymbol{x}\in A:\|\boldsymbol{x}-\boldsymbol{y}\|\le\sqrt{2/F_{2m+1}}\text{ for some }\boldsymbol{y}\in[0,1]^2\setminus A\Bigr\}\Bigr)\\&\le 4\sqrt{2/F_{2m+1}},\end{aligned}$$

since, by the convexity of A, the boundary of A has length at most 4 (which is the circumference of the square $[0,1]^2$).

Thus we obtain

$$\left|\frac{1}{F_{2m+1}}\sum_{n=0}^{F_{2m+1}-1} 1_A(\boldsymbol{f}_n)-\lambda(A)\right|\le 4\sqrt{\frac{2}{F_{2m+1}}}.$$

Now, we consider even integers $n=2m$ with $m\ge 2$. Using the identity $F_mF_{2m-1}-F_{m-1}F_{2m}=(-1)^{m-1}F_m$, we obtain $F_mF_{2m-1}\equiv(-1)^{m-1}\times F_m\pmod{F_{2m}}$. Consequently,

$$\left(\frac{k}{F_{2m}},\left\{\frac{kF_{2m-1}}{F_{2m}}\right\}\right)=\left(\frac{F_m}{F_{2m}},\left\{\frac{(-1)^{m-1}F_m}{F_{2m}}\right\}\right)\quad\text{for }k=F_m.$$

Thus the Fibonacci lattice has the equivalent generating vector

$$\boldsymbol{a}_{2m}=\left(\frac{F_m}{F_{2m}},\frac{(-1)^{m-1}F_m}{F_{2m}}\right).$$

Analogously, using the equality $F_{m+1}F_{2m-1} - F_m F_{2m} = (-1)^m F_{m-1}$, we obtain the generating vector

$$\boldsymbol{b}_{2m} = \left(\frac{F_{m+1}}{F_{2m}}, \frac{(-1)^m F_{m-1}}{F_{2m}}\right).$$

The area of the parallelogram spanned by $\boldsymbol{a}_{2m}$ and $\boldsymbol{b}_{2m}$ is

$$\left|\det\begin{pmatrix} F_m/F_{2m} & (-1)^{m-1}F_m/F_{2m} \\ F_{m+1}/F_{2m} & (-1)^m F_{m-1}/F_{2m}\end{pmatrix}\right| = \frac{F_m F_{m-1} + F_m F_{m+1}}{F_{2m}^2} = \frac{1}{F_{2m}}.$$

Thus the parallelogram spanned by $\boldsymbol{a}_{2m}$ and $\boldsymbol{b}_{2m}$ does not contain any point of the Fibonacci lattice in its interior (i.e. is a unit cell of the Fibonacci lattice, see [15, 43]). Thus, we have

$$\mathcal{F}_{2m} = \{u\boldsymbol{a}_{2m} + v\boldsymbol{b}_{2m} : u, v \in \mathbb{Z}\} \cap [0, 1)^2.$$

Let

$$U(\boldsymbol{y}) = \left\{\boldsymbol{y} + \boldsymbol{x} \in [0, 1]^2 : \boldsymbol{x} = s\boldsymbol{a}_{2m+1} + t\boldsymbol{b}_{2m+1}, 0 \le s, t < 1\right\}.$$

We call $U(\boldsymbol{f}_n)$ a unit cell (belonging to the point $\boldsymbol{f}_n$). Note that the area of a unit cell is $1/F_{2m+1}$ and each unit cell contains exactly one point of the lattice.

Now,

$$\|\boldsymbol{a}_{2m}\|^2 = \frac{2F_m^2}{F_{2m}^2}$$

and

$$\|\boldsymbol{b}_{2m}\|^2 = \frac{F_{m+1}^2 + F_{m-1}^2}{F_{2m}^2} > \frac{F_m^2 + F_m(2F_{m-1})}{F_{2m}^2} > \frac{2F_m^2}{F_{2m}^2} = \|\boldsymbol{a}_{2m}\|^2.$$

Further, it can be checked that $\|\boldsymbol{a}_{2m} + \boldsymbol{b}_{2m}\|, \|\boldsymbol{a}_{2m} - \boldsymbol{b}_{2m}\| > \|\boldsymbol{a}_{2m}\|$. Thus the minimum distance between points of the Fibonacci lattice is

$$d_{\min}(\mathcal{F}_{2m}) = \|\boldsymbol{a}_{2m}\| = \frac{\sqrt{2}F_m}{F_{2m}}.$$

The diameter of a unit cell is given by $\|\boldsymbol{a}_{2m} + \boldsymbol{b}_{2m}\|$. Using the relations $F_m = F_{m-1} + F_{m-2}$, $F_m^2 + F_{m-1}^2 = F_{2m-1}$ and $F_{2m} = (2F_{m-1} + F_m)F_m$, we obtain

$$\begin{aligned} F_{2m}^2\|\boldsymbol{a}_{2m} + \boldsymbol{b}_{2m}\|^2 &= (F_m + F_{m+1})^2 + (F_m - F_{m-1})^2 \\ &= 2F_{2m} + 4F_{2m-1} + 2F_m^2 < 8F_{2m}. \end{aligned}$$

Thus the diameter of a unit cell is bounded by

$$\|\boldsymbol{a}_{2m} + \boldsymbol{b}_{2m}\| \le \sqrt{\frac{8}{F_{2m}}}.$$

The result now follows by using the same arguments as in the previous case. □

Table 2 Comparison of the effect of two different normalisations of the value $\widetilde{D}(Z_{F_m})$ computing $\max_{\boldsymbol{w}\in Z_{F_m}} \delta(Z_{F_m};\boldsymbol{w})$, cf. Definition 1, for spherical Fibonacci points Z_{F_m}

m	17	18	19	20	21	22	23	24
F_m	1597	2584	4181	6765	10946	17711	28657	46368
$\frac{F_m^{3/4}\widetilde{D}(Z_{F_m})}{\sqrt{\log F_m}}$	0.6729	0.6373	0.6228	0.6661	0.6953	0.6890	0.7427	0.6900
$\frac{F_m^{3/4}\widetilde{D}(Z_{F_m})}{\log F_m}$	0.2477	0.2273	0.2156	0.2243	0.2279	0.2203	0.2318	0.2105
m	25	26	27	28	29	30	31	32
F_m	75025	121393	196418	317811	514229	832040	1346269	2178309
$\frac{F_m^{3/4}\widetilde{D}(Z_{F_m})}{\sqrt{\log F_m}}$	0.6957	0.7249	0.7531	0.7205	0.8562	0.7455	0.7862	0.8082
$\frac{F_m^{3/4}\widetilde{D}(Z_{F_m})}{\log F_m}$	0.2076	0.2118	0.2157	0.2024	0.2361	0.2019	0.2092	0.2115

Corollary 18 *Let $\mathcal{F}_m$ be a Fibonacci lattice and let $Z_{F_m} = \boldsymbol{\Phi}(\mathcal{F}_m) \subseteq \mathbb{S}^2$. Then the spherical cap discrepancy $D(Z_{F_m})$ is bounded by*

$$D(Z_{F_m}) \le \begin{cases} 44\sqrt{2/F_m} & \text{if } m \text{ is odd},\\ 44\sqrt{8/F_m} & \text{if } m \text{ is even}. \end{cases}$$

The numerical experiments shown in Table 2 seem to suggest that the correct order of the spherical cap discrepancy of spherical Fibonacci lattice points is

$$\frac{(\log F_m)^c}{F_m^{3/4}} \qquad \text{for some } 1/2 \le c \le 1.$$

In those experiments we calculated an approximation from below of the spherical cap discrepancy of the members of a sequence of spherical Fibonacci point configurations Z_{F_m} by explicit numerical computation of

$$\widetilde{D}(Z_{F_m}) = \max_{\boldsymbol{w}\in Z_{F_m}} \delta(Z_{F_m};\boldsymbol{w}).$$

6 Level Curves of the Distance Function and Their Properties

The Euclidean distance of two points on $\mathbb{S}^2$, given by $\boldsymbol{w} = \boldsymbol{\Phi}(u,v)$ and $\boldsymbol{z} = \boldsymbol{\Phi}(\alpha,\tau)$, can be written as

$$\begin{aligned}\|\boldsymbol{w}-\boldsymbol{z}\|^2 &= 2(1-\boldsymbol{w}\cdot\boldsymbol{z})\\ &= 2\Big[\underbrace{1-(1-2v)(1-2\tau)}_{2(v+\tau-2v\tau)} - 4\sqrt{(1-v)v(1-\tau)\tau}\,\cos\big(2\pi(u-\alpha)\big)\Big].\end{aligned}$$

The well-defined boundary curve of a spherical cap $C(\boldsymbol{w},t)$ has the implicit representation

$$F(\alpha,\tau):=\|\boldsymbol{w}-\boldsymbol{z}\|^2-2(1-t)=0, \quad \text{where } \boldsymbol{z}=\boldsymbol{\Phi}(\alpha,\tau) \text{ moves on } \partial C(\boldsymbol{w},t). \tag{8}$$

(Here, $\sqrt{2(1-t)}$ is the distance from the centre $\boldsymbol{w}$ to a point on the boundary of $C(\boldsymbol{w},t)$.) Relation (8) describes a level curve $\mathcal{C}_{\boldsymbol{w}}$ of the distance function $\|\boldsymbol{w}-\boldsymbol{\Phi}(\alpha,\tau)\|$ (for $\boldsymbol{w}$ fixed) in the parameter space which is the unit square, cf. Fig. 2. For further references we record that for each $\boldsymbol{w}$ there are, in general, two exceptional levels,

$$r_{\boldsymbol{w}}^2=2(1-t_{\boldsymbol{w}})=\|\boldsymbol{w}-\boldsymbol{p}\|^2=4v,$$
$$\rho_{\boldsymbol{w}}^2=2\bigl(1-t_{\boldsymbol{w}}'\bigr)=2(1+t_{\boldsymbol{w}})=\|\boldsymbol{w}+\boldsymbol{p}\|^2=4(1-v),$$

where the boundary of the spherical cap centred at $\boldsymbol{w}$ passes through the North Pole ($\boldsymbol{p}$) and the South Pole ($-\boldsymbol{p}$), respectively, which may coincide if $\boldsymbol{w}$ is on the equator. For these level curves the singular behaviour at the poles imposed by the parameterisation $\boldsymbol{\Phi}$ plays a role.

Suppose $u=1/2$. Because the sign of the difference $u-\alpha$ is absorbed by the cosine function in the distance function, a level curve (a level set) is symmetric with respect to the vertical line $\alpha=u$. The shape of the curves (sets) does not change when the point $\boldsymbol{w}$ is rotated about the polar axis except a part moving outside the left side of the unit square enters at the right side ("wrap around"). This "modulo 1" behaviour complicates considerations regarding convexity of the pre-image of a spherical cap centred at $\boldsymbol{w}$ under the distance function. Similarly and in the same sense, the level curves (sets) are symmetric with respect to the vertical line $\alpha=u\pm 1/2 \mod 1$ which passes through the parameter point of the antipodal point $-\boldsymbol{w}=\boldsymbol{\Phi}(u\pm 1/2 \mod 1, 1-v)$. When identifying the left and right sides of the unit square $[0,1]^2$, we get the "cylindrical view". Clearly, all level curves except the critical ones associated with the distance to one of the poles are closed on the open cylinder. Thus, the critical level curves separate the open cylinder into three parts corresponding to the cases when neither pole is contained in the spherical cap centred at $\boldsymbol{w}$, only one pole is contained in the spherical cap, and both poles are contained in the spherical cap. In both the first and the last case the level curve cannot escape the level set bounded by a critical curve (and the boundary of the cylinder). It is closed even in the unit square provided the level set is contained in its interior. In the middle case the level curves wrap around the cylinder; that is, start at the left side of the unit square and end at its right side at the same height; cf. Fig. 2.

Let $\boldsymbol{w}$ be the North or the South Pole Then the level curves are horizontal lines in the unit square which are smooth curves. Moreover, the pre-image of a spherical cap centred at one of these poles is a convex set (a rectangle).

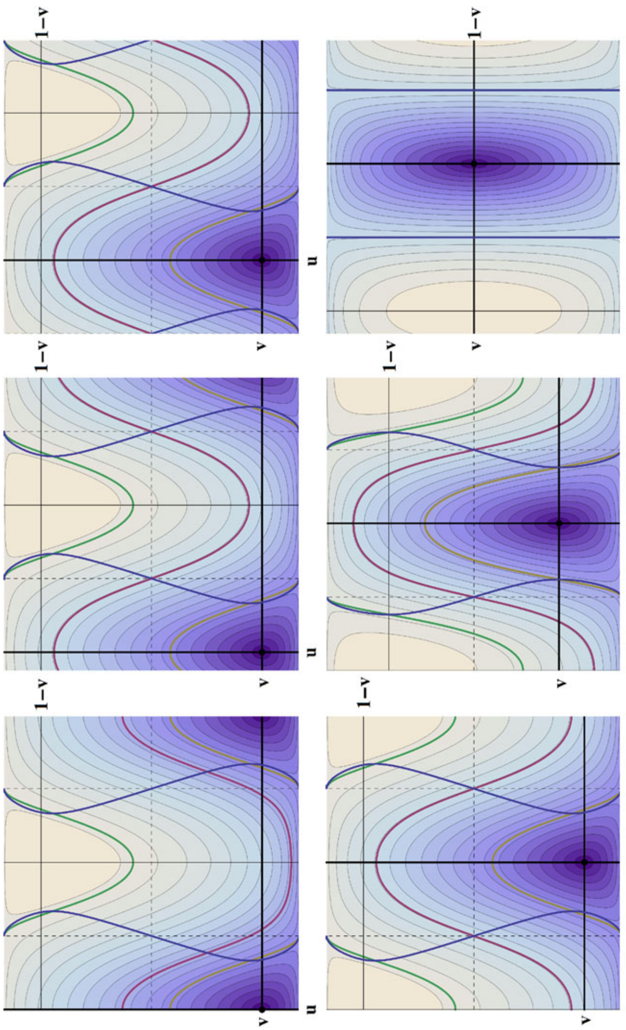


Fig. 2 Contour plot of the distance function $\|\boldsymbol{w} - \boldsymbol{\Phi}(\alpha, \tau)\|$ for various $\boldsymbol{w}$. Level curves for distances $\sqrt{2}$ and $2\sqrt{v}$, $2\sqrt{1-v}$ are emphasised. The *solid curve* shows vanishing curvature

Let $\boldsymbol{w}$ be Different from Either Pole (that is, $0 < v < 1$*)* We use the signed curvature of the implicitly given level curve to determine the segments where it is convex (concave). First, we collect the partial derivatives up to second order:

$$F_\alpha(\alpha,\tau) = -16\pi\sqrt{(1-v)v(1-\tau)\tau}\,\sin\bigl(2\pi(u-\alpha)\bigr), \tag{9a}$$

$$F_\tau(\alpha,\tau) = 4\left(1-2v-(1-2\tau)\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}\cos\bigl(2\pi(u-\alpha)\bigr)\right), \tag{9b}$$

$$F_{\alpha\alpha}(\alpha,\tau) = 32\pi^2\sqrt{(1-v)v(1-\tau)\tau}\,\cos\bigl(2\pi(u-\alpha)\bigr), \tag{9c}$$

$$F_{\alpha\tau}(\alpha,\tau) = -8\pi(1-2\tau)\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}\sin\bigl(2\pi(u-\alpha)\bigr), \tag{9d}$$

$$F_{\tau\tau}(\alpha,\tau) = \frac{2}{(1-\tau)\tau}\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}\cos\bigl(2\pi(u-\alpha)\bigr). \tag{9e}$$

We observe that the partial derivatives involving differentiation with respect to τ become singular as τ approaches 0 (North Pole) or 1 (South Pole).

The signed curvature at a point (α,τ) of $\mathcal{C}_{\boldsymbol{w}}$ is given by

$$\kappa = \kappa(\alpha,\tau) = \frac{F_{\alpha\alpha}F_\tau^2 - 2F_{\alpha\tau}F_\alpha F_\tau + F_{\tau\tau}F_\alpha^2}{(F_\alpha^2+F_\tau^2)^{3/2}}. \tag{10}$$

First, we discuss the denominator. Substituting the relations (9a) and (9b), we obtain

$$\begin{aligned} F_\alpha^2+F_\tau^2 &= 16\left(1-2v-(1-2\tau)\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}\cos\bigl(2\pi(u-\alpha)\bigr)\right)^2 \\ &\quad + 16^2\pi^2(1-v)v(1-\tau)\tau\bigl[\sin\bigl(2\pi(u-\alpha)\bigr)\bigr]^2. \end{aligned}$$

A necessary condition for the vanishing of $F_\alpha^2+F_\tau^2$ is $\sin(2\pi(u-\alpha))=0$; that is, either $\alpha=u$ or $\alpha = u\pm 1/2 \mod 1$. In the first case one has $\cos(2\pi(u-\alpha))=1$ and, therefore, $F_\alpha^2+F_\tau^2=0$ if and only if $\tau=v$. In the second case one has $\cos(2\pi(u-\alpha))=-1$ and, therefore, $F_\alpha^2+F_\tau^2=0$ if and only if $\tau=1-v$. It follows that the denominator of the curvature formula (10) vanishes if and only if $\boldsymbol{z}=\boldsymbol{\Phi}(\alpha,\tau)$ (on the boundary of the spherical cap centred at $w=\boldsymbol{\Phi}(u,v)$) coincides with $\boldsymbol{w}$ (that is, the spherical cap degenerates to the point $\boldsymbol{w}$) or $\boldsymbol{z}$ coincides with the antipodal point of $\boldsymbol{w}$ (that is, the closed spherical cap is the whole sphere). In either of these cases the pre-image of the spherical cap under $\boldsymbol{\Phi}$ is convex.

Suppose that the *spherical cap is neither a point nor the whole sphere*. Then $F_\alpha^2+F_\tau^2\neq 0$. For the numerator in (10) we obtain after some simplifications:

$$\begin{aligned} &F_{\alpha\alpha}F_\tau^2 - 2F_{\alpha\tau}F_\alpha F_\tau + F_{\tau\tau}F_\alpha^2 \\ &\quad = 512\pi^2\sqrt{(1-v)v(1-\tau)\tau}\left(1-2v-(1-2\tau)\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}\cos\bigl(2\pi(u-\alpha)\bigr)\right)^2 \\ &\qquad \times\cos\bigl(2\pi(u-\alpha)\bigr) \end{aligned}$$

$$+ 512\pi^2(1-v)v\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}\cos\bigl(2\pi(u-\alpha)\bigr)\bigl[\sin\bigl(2\pi(u-\alpha)\bigr)\bigr]^2$$
$$- 1024\pi^2(1-v)v(1-2\tau)\left(1-2v-(1-2\tau)\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}\cos\bigl(2\pi(u-\alpha)\bigr)\right)$$
$$\times \bigl[\sin\bigl(2\pi(u-\alpha)\bigr)\bigr]^2.$$

The right-hand side above can be written as a polynomial in $x=\cos(2\pi(u-\alpha))$ as follows:

$$F_{\alpha\alpha}F_\tau^2 - 2F_{\alpha\tau}F_\alpha F_\tau + F_{\tau\tau}F_\alpha^2 = Ax(1-2v-BHx)^2 + \frac{A^2Bx(1-x^2)}{512\pi^2(1-\tau)\tau} - 2ABH(1-2v-BHx)\bigl(1-x^2\bigr),$$

where

$$A = 512\pi^2\sqrt{(1-v)v(1-\tau)\tau}, \qquad B=\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}, \qquad H = 1-2\tau.$$

Reordering with respect to falling powers of x, we observe that the coefficient of x^2 vanishes, and after simplifications we arrive at

$$F_{\alpha\alpha}F_\tau^2 - 2F_{\alpha\tau}F_\alpha F_\tau + F_{\tau\tau}F_\alpha^2 = -AB\left(BH^2+\frac{A}{512\pi^2(1-\tau)\tau}\right)x^3 + A\left(2B^2H+(1-2v)^2+\frac{AB}{512\pi^2(1-\tau)\tau}\right)x - 2ABH(1-2v).$$

The coefficient of x^3 does not vanish for $0<v<1$ and $0<\tau<1$. Hence, we divide and get

$$-\frac{F_{\alpha\alpha}F_\tau^2 - 2F_{\alpha\tau}F_\alpha F_\tau + F_{\tau\tau}F_\alpha^2}{AB(BH^2+\frac{A}{512\pi^2(1-\tau)\tau})} = x^3+px+q =: Q(\tau;x) =: Q(x),$$

where (using the definitions of A, B, and H)

$$p=p(\tau) = -\frac{(1-2v)^2+B^2(1+2H^2)}{B^2(1+H^2)}, \qquad q=q(\tau)=\frac{2(1-2v)(1-2\tau)}{B(1+H^2)}.$$

We observe that $p(1-\tau)=p(\tau)$ and $q(1-\tau)=-q(\tau)$. Hence, $Q(\tau;x)=-Q(1-\tau;-x)$ for all x. In particular, if ξ is a zero of $Q(\tau;\cdot)$, then so is $-\xi$ a zero of $Q(1-\tau;\cdot)$ and vice versa. The monic polynomial Q of degree 3 with real coefficients has either one or three real solutions (counting multiplicity). With the help of Mathematica we find that the discriminant of the polynomial Q is positive:

$$\operatorname{discr}(Q) = -4p^3-27q^2>0;$$

that is, the polynomial Q has three distinct real roots.

Table 3 The number of real zeros of Q in the intervals $(-1,0)$ and $(0,1)$ as it follows from Sturm's theorem

Range of v	Range of τ	$\sigma(-1)$	$\sigma(0)$	$\sigma(1)$	$(-1,0)$	$(0,1)$
$0<v<1/2$	$0<\tau<1/2$	2	2	1	0	1
$0<v<1/2$	$1/2<\tau<1$	2	1	1	1	0
$1/2<v<1$	$0<\tau<1/2$	2	1	1	1	0
$1/2<v<1$	$1/2<\tau<1$	2	2	1	0	1

For $v=1/2$ the polynomial Q reduces to

$$Q(x)=x^3+px, \quad \text{where } p=-\frac{1+2H^2}{1+H^2}=-\frac{1+2(1-2\tau)^2}{1+(1-2\tau)^2}.$$

The solutions ± 1 (if $\tau=1/2$) correspond to $\boldsymbol{\Phi}(\alpha,\tau)=\pm\boldsymbol{w}$ and can be discarded, since we assumed that the spherical cap is neither a point nor the whole sphere. The solution zero yields that $\cos(2\pi(u-\alpha))=0$, which in turn shows that the zeros of the curvature (10) form the vertical lines at $\alpha=u\pm 1/4 \mod 1$ if $v=1/2$.

Let $v\neq 1/2$. Suppose Q has a zero at ± 1. Then

$$0=Q(1)=1+p+q=\pm\frac{(1-2v+BH)^2}{B^2(1+H^2)}=\pm\frac{(1-2v+\frac{\sqrt{(1-v)v}}{\sqrt{(1-\tau)\tau}}(1-2\tau))^2}{B^2(1+H^2)},$$

which can only happen when $\tau=v$. This implies that $\boldsymbol{\Phi}(\alpha,\tau)=\boldsymbol{w}$, which is excluded by our assumptions. Suppose Q has a zero at 0. Since $v\neq 1/2$, this can only happen when $\tau=1/2$.

Having established that $-1,0$ (except when $\tau=1/2$), and 1 cannot be zeros of the polynomial Q, we use Sturm's theorem to show that the polynomial Q has precisely one solution either in the interval $(-1,0)$ or in the interval $(0,1)$ if $\tau\neq 1/2$, cf. Table 3. First, we generate the canonical Sturm chain by applying Euclid's algorithm to Q and its derivative:

$$\begin{aligned}
p_0(x)&=Q(x)=x^3+px+q,\\
p_1(x)&=Q'(x)=3x^2+p,\\
p_2(x)&=p_1(x)q_0(x)-p_0(x)=-\frac{2p}{3}x-q,\\
p_3(x)&=p_2(x)q_1(x)-p_1(x)=-p-\frac{27q^2}{4p^2}=\frac{-4p^3-27q^2}{4p^2}=\frac{\operatorname{discr}(Q)}{4p^2}>0,\\
p_4(x)&=0.
\end{aligned}$$

Let $\sigma(x)$ denote the number of sign changes (not counting a zero) in the sequence

$$\{p_0(x),p_1(x),p_2(x),p_3(x)\}.$$

For $x=0$ we obtain the canonical Sturm chain $\{q,p,-q,\operatorname{discr}(Q)/(4p^2)\}$ and we conclude that $\sigma(0)=2$ for $(1/2-v)(1/2-\tau)>0$ and $\sigma(0)=1$ otherwise. For $x=1$

we have

$$p_0(1) = 1 + p + q = -\frac{(1-2v+BH)^2}{B^2(1+H^2)} < 0,$$

$$p_1(1) = 3 + p = \frac{2+H^2}{1+H^2} - \frac{(1-2v)^2}{B^2(1+H^2)} > 0 \quad \text{iff } (1-\tau)\tau < 3(1-v)v,$$

$$p_2(1) = -\frac{2p}{3} - q = \frac{2[B^2(1+2H^2) + (1-2v)^2 - 3BH(1-2v)]}{3B^2(1+H^2)} > 0,$$

$$p_3(1) = \frac{\operatorname{discr}(Q)}{4p^2} > 0.$$

(The positivity of $p_2(1)$ has been verified using Mathematica.) Hence, in all three cases $p_1(1) < 0$, $p_1(1) = 0$, and $p_1(1) > 0$, one gets $\sigma(1) = 1$. For $x = -1$ we have

$$p_0(-1) = -1 - p + q = \frac{(1-2v+BH)^2}{B^2(1+H^2)} > 0,$$

$$p_1(-1) = 3 + p = \frac{2+H^2}{1+H^2} - \frac{(1-2v)^2}{B^2(1+H^2)} > 0 \quad \text{iff } (1-\tau)\tau < 3(1-v)v,$$

$$p_2(-1) = \frac{2p}{3} - q = -\frac{2[B^2(1+2H^2) + (1-2v)^2 + 3BH(1-2v)]}{3B^2(1+H^2)} < 0,$$

$$p_3(-1) = \frac{\operatorname{discr}(Q)}{4p^2} > 0.$$

Here, we obtain $\sigma(-1) = 2$. Thus, by Sturm's theorem, the difference $\sigma(-1) - \sigma(0)$ gives the number of real zeros of Q in the interval $(-1, 0]$ and $\sigma(0) - \sigma(1)$ is the number of zeros in $(0, 1]$, see Table 3.

Having established that the polynomial Q has to every $0 < \tau < 1$ precisely one zero in the interval $(-1, 1)$ (cf. Table 3 and previous considerations), it follows that to each such zero $x = x(\tau)$ there correspond two values of α by means of the trigonometric equation

$$\cos\bigl(2\pi(u-\alpha)\bigr) = x(\tau). \tag{11}$$

Because of the continuity of the coefficients in the polynomial Q (if $0 < \tau < 1$), the zero $x(\tau)$ is also changing continuously and so are the solutions α_1 and α_2. A jump can happen when they are taken modulo 1. We further record that along the vertical lines $\alpha = u \pm 1/4 \bmod 1$ one has

$$\kappa(u \pm 1/4, \tau) = -16\pi^2 \frac{(1-v)v(1-2v)(1-2\tau)}{[1 - 4(1-v)v(1-4\pi^2(1-\tau)\tau)]^{3/2}}$$

which vanishes at $\tau = 1/2$, and along the vertical lines $\alpha = u \pm 1/2 \bmod 1$ and $\alpha = u$ one has

$$\kappa(u, \tau) = \kappa(u \pm 1/2, \tau) = -8\pi^2 \frac{\sqrt{(1-v)v}(1-\tau)\tau}{|(1-2v)\sqrt{(1-\tau)\tau} + (1-2\tau)\sqrt{(1-v)v}|}$$

which vanishes only as $\tau \to 0$ or $\tau \to 1$ (if $0 < v < 1$). When identifying the left and right sides of the unit square, these two lines separate the two solutions of (11) in such a way that in each part the points at which $\kappa(\alpha, \tau)$ vanishes form a connected curve varying about the "base lines" $\alpha = u \pm 1/4 \bmod 1$. It follows that these curves (together with the boundary of the cylinder) divide the cylinder into two parts in each of which the curvature $\kappa(\alpha, \tau)$ has the same sign. The shapes of these curves do not change when $\boldsymbol{w}$ is rotated about the polar axis. We may fix $u = 1/2$ and because of the symmetries (including relation $Q(\tau; x) = -Q(1-\tau; -x)$) it suffices to consider the curve of the zeros of $\kappa(\alpha, \tau)$ for $0 < v < 1/2$ (recall that these curves are vertical lines for $v = 1/2$) and $0 < \tau < 1/2$ which lies in the strip $0 < \alpha < 1/2$. We know that the zero $x = x(\tau)$ of Q (we are interested in) in the given setting is in $(0, 1)$ (cf. Table 3). Using $\dot{x}$ to denote the derivative of x with respect to τ, implicit differentiation gives

$$Q'\big(x(\tau)\big)\,\dot{x}(\tau) = -\dot{p}(\tau)\,x(\tau) - \dot{q}(\tau), \tag{12}$$

where it can be easily seen that $Q'(x(\tau)) < 0$, since $x(\tau)$ is simple and $Q(x)$ has a negative global minimum for positive x. For τ's in $(0, 1/2)$ at which $\dot{x}(\tau)$ vanishes, one has

$$x(\tau) = -\dot{q}(\tau)/\dot{p}(\tau) = \frac{1-6(1-\tau)\tau}{1-6(1-v)v}\,\frac{1-2v}{1-2\tau}\sqrt{\frac{(1-v)v}{(1-\tau)\tau}}, \tag{13}$$

which follows by substituting

$$\dot{p}(\tau) = -\frac{(1-6(1-v)v)(1-2\tau)}{2(1-v)v(1-2(1-\tau)\tau)^2},$$
$$\dot{q}(\tau) = \frac{(1-6(1-\tau)\tau)(1-2v)}{2(1-\tau)\tau(1-2(1-\tau)\tau)^2}\sqrt{\frac{(1-\tau)\tau}{(1-v)v}}.$$

For the second derivative of x at such τ's we get

$$\begin{aligned} Q'\big(x(\tau)\big)\ddot{x}(\tau) &= -\ddot{p}(\tau)x(\tau) - \ddot{q}(\tau) \\ &= \frac{1-2v}{4((1-\tau)\tau)^{3/2}\sqrt{(1-v)v}[1-2\tau(2-\tau(3-2\tau))]}. \end{aligned}$$

The square-bracketed expression is strictly monotonically decreasing on $(0, 1/2)$ and evaluates to zero at $\tau = 1/2$. Thus, the left-hand side has to be positive for all critical τ in $(0, 1/2)$ which in turn implies that $\ddot{x}(\tau) < 0$ at such τ's. We conclude that $x(\tau)$ has a single maximum in $(0, 1/2)$, since it cannot be constant as can be seen from (13) by evaluating $x(\tau)$ at that τ' at which $1 - 6(1-\tau')\tau' = 0$ and $x(\tau) \to \infty$ as $\tau \to 1/2$. By (11) (recall $u = 1/2$)

$$-\cos\big(2\pi\alpha(\tau)\big) = x(\tau)$$

and it follows that $\alpha(\tau)$ has a single maximum in $(0, 1/2)$.

Proposition 19 *The set of zeros of $\kappa(\alpha, \tau)$ and the horizontal sides of the unit square divide the unit square either into three parts of equal sign of $\kappa(\alpha, \tau)$ or in four parts, cf. Fig. 2.*

Using the trigonometric method, we obtain an explicit expression of the zero of Q in $(-1, 1)$. The change of variable $x = 2\sqrt{-p/3}\cos\theta$ gives the equivalence

$$Q(x) = 0 \quad \text{if and only if} \quad \cos(3\theta) = \frac{3q}{2p}\sqrt{\frac{3}{-p}}$$

and therefore

$$x(\tau) = 2\sqrt{-p(\tau)/3}\cos\left(\frac{1}{3}\arccos\left(\frac{3q(\tau)}{2p(\tau)}\sqrt{\frac{3}{-p(\tau)}}\right) - \frac{2\pi}{3}\right).$$

(The discarded solutions are either smaller or larger than the given one. By our reasoning, they have to lie outside the interval $(-1, 1)$.) A limit process shows that $x(\tau) \to 0$ as $\tau \to 0$ or $\tau \to 1$. Hence, when moving towards the upper or lower side of the square along the curve of zeros of the curvature, one approaches the corresponding "base line" $\alpha = u \pm 1/4 \bmod 1$.

Eliminating the trigonometric term, along a level curve with parameter t (cf. (8)) we have

$$\begin{aligned} F_{\alpha\alpha}F_\tau^2 - 2F_{\alpha\tau}F_\alpha F_\tau + F_{\tau\tau}F_\alpha^2 = -16\pi^2\Bigg[&\left(\frac{1}{(1-\tau)^2} + \frac{1}{\tau^2}\right)X^3 \\ &- 8\left(1 + 3\frac{(1-v)v}{(1-\tau)\tau}(1-2\tau)^2\right)X + 64(1-v)v(1-2v)(1-2\tau)\Bigg], \end{aligned}$$

where $X = t - (1-2v)(1-2\tau)$. Reordering the terms and using the substitution $G = t - (1-2v)$, we arrive at

$$\begin{aligned} F_{\alpha\alpha}F_\tau^2 - 2F_{\alpha\tau}F_\alpha F_\tau + F_{\tau\tau}F_\alpha^2 = -\frac{16}{(1-\tau)^2\tau^2}\big[&G^3 - 2G\big(3 + G^2 + 3Gt - 3t^2\big)\tau \\ &+ 2\big(2t + 6G^2t - 2t^3 + 3G\big(3 - t^2\big)\big)\tau^2 - 4\big(5t - 3t^3 + 3G\big(1 + t^2\big)\big)\tau^3 + 16t\tau^4\big]. \end{aligned} \tag{14}$$

The zeros of the numerator of the curvature (10) are determined by a polynomial in τ of degree 4. Thus, there can be at most four pairs (symmetry with respect to $\alpha = u$) of points on the level set at which the curvature vanishes. For the sake of completeness, in a similar way one obtains

$$\begin{aligned} \big(F_\alpha^2 + F_\tau^2\big)^{3/2} = \frac{1}{(1-\tau)^3\tau^3}\big[&(G - 2t\tau)^2 - 16\pi^2(1-\tau)^2\tau^2 \\ &\times \big(G^2 - 4Gt\tau - 4\tau\big(1 - t^2 - \tau\big)\big)\big]^{3/2}. \end{aligned}$$

The critical curves when the boundary of the spherical cap passes through a pole are of particular interest. In the case of the North Pole (that is, $t = 1 - 2v$, or equivalently $G = 0$), the curvature along the corresponding level curve reduces to

$$\kappa(\tau) = -16\pi^2 \frac{(1-2v)(1-\tau)[2(1-v)v - (1+6(1-v)v)\tau + 2\tau^2]}{[(1-2v)^2 - 16\pi^2(1-\tau)^2\tau(\tau - 4(1-v)v)]^{3/2}}. \tag{15}$$

For given $0 < \tau < 1$, the corresponding value(s) of α can be obtained from the relation

$$\cos\bigl(2\pi(u-\alpha)\bigr) = \frac{(1-2v)\tau}{2\sqrt{(1-v)v(1-\tau)\tau}}. \tag{16}$$

From (8) it follows that the range for τ is $(0, \tau_1]$ with $\tau_1 = 4v(1-v)$. For future reference we record that for $0 < v < 1$ with $v \neq 1/2$ the curvature (15) vanishes only for

$$\tau_v = \frac{1}{4}\Bigl(1 + 6(1-v)v - \sqrt{1 - 4(1-v)v\bigl(1 - 9(1-v)v\bigr)}\Bigr). \tag{17}$$

(The other solution lies outside the interval $[-1, 1]$ as one can verify with Mathematica.)

Proposition 20 *The curves of zeros of the curvature* (10) (*as functions of* τ) *assume their extrema at* τ_v *and* $1 - \tau_v$ *with* τ_v *given in* (17).

Proof Suppose that $u = 1/2$ and $0 < v < 1/2$. Then $0 < \tau_v < 1/2$. On observing that the right-hand side of (16) for $\tau = \tau_v$ is also the zero $x(\tau_v)$ in the interval $(0, 1)$ of the polynomial Q, it can be verified with the help of Mathematica that the right-hand side of (12) vanishes and therefore $\dot{x}(\tau_v) = 0$; that is, the zero $x(\tau)$ is extremal at $\tau = \tau_v$. Using the symmetry relation $Q(\tau; x) = -Q(1-\tau; -x)$, the zero $x(\tau)$ is also extremal at $\tau = 1 - \tau_v$. By means of (16) this translates into extrema of the curve of zeros of the curvature (10). A shift in u (rotation of $\boldsymbol{w}$ about the polar axis) does not change the shape of the level curves and the general result follows. □

Substituting (17) into the right-hand side of (16) gives the extremal value a zero of Q in $(-1, 1)$ can assume, also cf. (13):

$$x(\tau_v) = \frac{1-2v}{\sqrt{1 + 2(1-v)v + \sqrt{1 - (1-v)v(9(1-2v)^2 - 5)}}}.$$

It can be shown that $x(\tau_v)$ is a strictly monotonically decreasing function in v which is symmetric with respect to $v = 1/2$. Hence $|x(\tau_v)| \leq x(0^+) = 1/\sqrt{2}$. Using this bound in (16) yields that $|\alpha - (u \pm 1/4)| \leq 1/8$ (when wrapping around).

When moving along the critical level curve towards the lower side of the unit square, which is associated with the North Pole, we have

$$\lim_{\tau \to 0} \kappa(\tau) = -32\pi^2 \frac{1-2v}{|1-2v|^3}(1-v)v, \quad 0 < v < 1,\ v \neq 1/2.$$

We conclude that the critical level curve with $t = 1 - 2v$ (associated with the North Pole) has precisely one symmetric (in the cylindrical view) pair of intersection points with the curves of zeros of the curvature function (10) at $\tau = \tau_v$ in the strip $0 < \tau < 1$. A similar result holds for the level curve associated with the South Pole ($t = -(1 - 2v)$).

Proposition 21 *Let* $0 < v < 1/2$. *Then the level curves with* t *in the range* $-(1 - 2v) \le t \le 1 - 2v$ *have precisely one symmetric* (*in the cylindrical view*) *pair of intersection points with the curve of zeros of the curvature function* (10). *For* t *in the ranges* $-1 < t < -(1 - 2v)$ *or* $1 - 2v < t < 1$ *there are either no intersection points*, *one pair of tangential points*, *or two pairs*.

The analogous result holds for $1/2 < t < 1$. (For $v = 1/2$ the level curves for distance $\sqrt{2}$ are the verticals at $\alpha = u \pm 1/4 \mod 1$ and coincide with the curve of zeros of $\kappa(\alpha, \tau)$ and also coincide with the critical curves. The other level curves have no intersections.)

Proof Without loss of generality assume that $u = 1/2$. We have already established that either critical level curve has precisely one pair of symmetric (with respect to $\alpha = u$) intersection point with the two curves of zero curvature about the base lines $\alpha = u \pm 1/4$ at the values $\tau = \tau_v$ and $\tau = 1 - \tau_v$. These parameter values also give the position of the extrema of the zero curves, cf. Fig. 2. The left zero curve $\mathcal{Z}$ is increasing for τ in $(0, \tau_v)$, decreasing for τ in $(\tau_v, 1 - \tau_v)$ and increasing again for τ in $(1 - \tau_v, 1)$.

Let $-(1 - 2v) < t < 1 - 2v$ and Γ_t denote the left half of the corresponding level curve starting at the left side at some point $(0, \tau_1)$ and ending at some point (u, τ_2). (The other half is symmetric.) We note that the part where the zero curve is increasing is contained in the regions separated off by the critical level curves. Thus, an intersection between $\mathcal{Z}$ and Γ_t can only occur for τ in the interval $[t_v, 1 - t_v]$. The curvature along Γ_t changes continuously (cf. (14) and subsequent formula) from negative to positive value. Hence, there is an intersection point of $\mathcal{Z}$ and Γ_t and the Γ_t cannot change abruptly. In particular, both $\mathcal{Z}$ and Γ_0[2] pass through $(1/4, 1/2)$, which is their only intersection point because for $\tau \ne 1/2$ the vertical line $\alpha = 1/4$ separates both curves. A Γ_t with $0 < t < 1 - 2v$ ($-(1 - 2v) < t < 0$) has to intersect $\mathcal{Z}$ in the strip $1/4 < \alpha < 1/2$ ($0 < \alpha < 1/4$). Inspecting the partial derivatives of F (cf. (9a)–(9e)) (9c), (9d), (9e)) it follows that the gradient of F at the intersection point, which is the outward normal at the level curve Γ_t, points into the upper left part; that is, the tangent vector at Γ_t at the intersection point shows to the right whereas the tangent vector at $\mathcal{Z}$ at this point shows to the left. Moreover, if $0 < t < 1 - 2v$, then the curve Γ_0 separates $\mathcal{Z}$ and Γ_t for $\tau \ge 1/2$ and in the remaining part both curves $\mathcal{Z}$ and Γ_t bend away from each other because $\mathcal{Z}$ is decreasing with growing τ and the curvature along Γ_t becomes positive. Consequently, there is only one intersection point of Γ_t and $\mathcal{Z}$. A similar argument holds for $-(1 - 2v) < t < 0$.

[2]The spherical cap is the half-sphere.

Let $1-2v<t<1$. Let Γ_t denote the left half of the level curve. For t sufficiently close to 1 there is no intersection with $\mathcal{Z}$ and the level curve is convex. If Γ_t and $\mathcal{Z}$ intersect in only one point, then the level curve is still convex, since the curvature function $\kappa(\alpha,\tau)$ has positive sign in the section between $\mathcal{Z}$ and the vertical line $\alpha=u$. In this case $\mathcal{Z}$ and Γ_t share a common tangent at the intersection point. If the curvature along Γ_t changes its sign to negative, then it has to become positive again, since it is positive when crossing the vertical $\alpha=u$. But after changing back to positive curvature, both curves are bending away from each other. So, there can be no other intersection point.

By symmetry with respect to the line $\alpha=u$ one has pairs of symmetric intersection points.

Shifting u does not change the form of the curves and their relative positions. This completes the proof. □

7 Proofs

Proof of Lemma 3 For $t=1$ the spherical cap is a point and for $t=-1$ it is the whole sphere. Their pre-images (a point and the whole unit square) are convex. So, we may assume that $-1<t<1$.

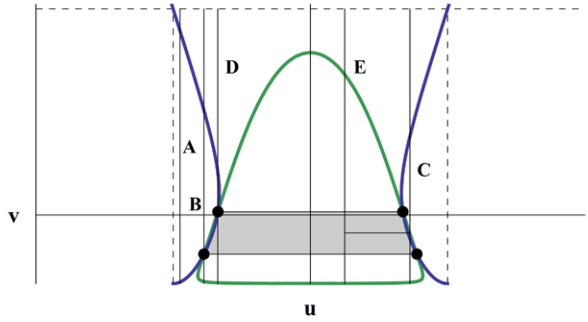
Case (i): Let $\boldsymbol{w}$ be either the North or the South Pole. Then the pre-images of the boundary of spherical caps centred at $\boldsymbol{w}$ are horizontal lines in the unit square. Hence, the pre-image of such a spherical cap is convex.

Case (ii): Let $\boldsymbol{w}$ be on the equator (that is, $v=1/2$). We know that the curvature (10) vanishes along the lines $\alpha=u\pm 1/4 \mod 1$. First, suppose that $u=1/4$. Then the pre-image of any spherical cap centred at $\boldsymbol{w}$ with boundary points at most Euclidean distance $\sqrt{2}$ away from $\boldsymbol{w}$ is convex. For a larger spherical cap C it follows that its complement $\overline{C}$ with respect to the sphere (centred at the antipodal point $-\boldsymbol{w}$) has the property that points on the boundary have distance $\leq\sqrt{2}$ from $-\boldsymbol{w}$. Hence, the pre-image of $\overline{C}$ is convex. When rotating $\boldsymbol{w}$ about the polar axis (that is, shifting u), the vertical boundaries of the square cut these convex sets into two parts. We conclude that the pre-image of a spherical cap centred at $\boldsymbol{w}$ or its complement with respect to the sphere is the union of at most two convex sets.

Case (iii): Let $\boldsymbol{w}$ be neither the poles nor located at the equator. Without loss of generality we may assume that $\boldsymbol{w}$ is in the upper half of the sphere; that is, $0<v<1/2$. (Otherwise we can use reflection with respect to the equator.) First, let us consider the canonical position $u=1/2$. Let $-(1-2v)\leq t\leq 1-2v$. Then, by Proposition 21, there are precisely two (symmetric) points along the level curve at which the curvature vanishes, say at (α_1,τ_t) and (α_2,τ_t). This yields a decomposition of the unit square into three vertical rectangles such that either the part above or below the level curve is convex. Let $1-2v<t<1$. By Proposition 21 the level curve is already convex or there are two pairs of symmetric points at which the curvature along the level curve vanishes and a sign change occurs. Hence, there are numbers $\tau_1<\tau_2$ such that the level curve is convex for $\tau\leq\tau_1$ and convex for $\tau\geq t_2$. The remaining middle part can be covered by a convex isosceles trapezoid which in turn can be split by some vertical line contained in the level set associated with the level curve. Thus,

Table 4 Worst-case admissible convex covering with p part and q of which are convex. The *vertical lines* show canonical positions of the *vertical borders* of $[0, 1]^2$

	p	q	$2p - q$
A	4	2	6
B	5	3	7
C	7	3	11
D	6	3	9
E	6	4	8



one has again two convex polygons which are divided into a convex and non-convex part by the level curve. A similar argument holds for $-1 < t < -(1 - 2v)$.

A shift of u does not increase the number of vertical rectangles needed for $0 \le t \le 1 - 2v$. (In fact, one may even reduce the number of elements of the partition.) In the case $1 - 2v < t < 1$ one may need to use a covering of the pre-image of the spherical cap with up to 7 pieces. A more precise analysis is listed in Table 4. □

Proof of Proposition 8 Radon's theorem (see e.g. [5, Theorem 4.1]) states that any set of $d + 2$ points from $\mathbb{R}^d$ can be partitioned into two disjoint subsets whose convex hulls intersect. Particularly, let A denote a set of 5 points on the sphere. Then by Radon's theorem there exists a partitioning of A into disjoint subsets whose convex hulls intersect. Thus the set A cannot be shattered by the class of half-spaces. Since every spherical cap is the intersection of the sphere with an appropriate half-space, the set A can also not be shattered by the class of spherical caps. Thus, the VC dimension of the class of spherical caps is at most 5.

On the other hand, let the set $\hat{A}$ consist of the points of a regular simplex, which lie on the sphere. Then some simple considerations show that the set $\hat{A}$ is shattered by the class of spherical caps. Thus, the VC dimension of the class of spherical caps (and therefore of course also the VC dimension of the class $\mathcal{C}$ of spherical caps for which the centre $\mathbf{w}$ and the height t are rational numbers, which was used in Sect. 4) equals 5. □

Proof of Theorem 9 As mentioned directly after the statement of Theorem 9, the lower bound in the theorem follows directly from (5). To prove the upper bound we use Theorem 7. By Proposition 8 the VC dimension of the class $\mathcal{C}$ in Sect. 4 is 5. For simplicity we assume that the constant in Theorem 7 is an integer. Then, for any $s \ge 2K$,

$$\mathbb{P}\left\{D(Z_N) \ge \frac{s}{\sqrt{N}}\right\} \le \frac{1}{s}\left(\frac{Ks^2}{5}\right)^5 e^{-2s^2},$$

and consequently we have

$$\begin{aligned}\mathbb{E}\big(D(Z_N)\big) &\le \frac{2K}{\sqrt{N}} + \sum_{s=2K}^{\infty}\left(\frac{(s+1)}{\sqrt{N}}\cdot \mathbb{P}\left\{D(Z_N)\ge \frac{s}{\sqrt{N}}\right\}\right)\\ &\le \frac{2K}{\sqrt{N}} + \sum_{s=2K}^{\infty}\frac{s+1}{s\sqrt{N}}\left(\frac{Ks^2}{5}\right)^5 e^{-2s^2}\\ &\le \frac{\hat{K}}{\sqrt{N}}\end{aligned}$$

for some appropriate constant $\hat{K}$. This proves the theorem. □

Proof of Theorem 10 Let $C^* \subseteq \mathcal{C}$ denote a hemisphere, i.e. a spherical cap whose normalised surface area measure is $\sigma(C^*) = 1/2$. By the central limit theorem, for any $t \ge 0$,

$$\mathbb{P}\left\{\left|\frac{1}{N}\sum_{n=0}^{N-1} 1_{C^*}(X_n) - \sigma\big(C^*\big)\right| \le tN^{-1/2}\right\} \to \frac{\sqrt{2}}{\sqrt{\pi}}\int_{-t}^{t} e^{-2u^2}\,\mathrm{d}u \quad \text{as } N\to\infty,$$

and consequently, for any given $\varepsilon > 0$ and sufficiently small $C_3(\varepsilon) > 0$,

$$\mathbb{P}\left\{\left|\frac{1}{N}\sum_{n=0}^{N-1} 1_{C^*}(X_n) - \sigma\big(C^*\big)\right| \le C_3 N^{-1/2}\right\} \le \varepsilon/2$$

for sufficiently large N. Since $D(Z_N) \ge |\frac{1}{N}\sum_{n=0}^{N-1} 1_{C^*}(X_n) - \sigma(C^*)|$, this implies

$$\mathbb{P}\big\{D(Z_N) \le C_3 N^{-1/2}\big\} \le \varepsilon/2 \tag{18}$$

for sufficiently large N.

On the other hand, by Theorem 7 for any given $\varepsilon > 0$ and sufficiently large $C_4(\varepsilon)$,

$$\mathbb{P}\left\{D(Z_N) \ge \frac{C_4}{\sqrt{N}}\right\} \le \varepsilon/2 \tag{19}$$

for sufficiently large N. Combining (18) and (19) we obtain

$$\mathbb{P}\left\{\frac{C_3}{\sqrt{N}} \le D(Z_N) \le \frac{C_4}{\sqrt{N}}\right\} \ge 1-\varepsilon$$

for sufficiently large N, which proves Theorem 11. □

Acknowledgements The authors are grateful to Peter Kritzer and Friedrich Pillichshammer for inspiring suggestions. In particular, F. Pillichshammer pointed out Theorem 14.

The research of C.A. was supported by the Austrian Research Foundation (FWF), Project S9603-N23. The research of J.S.B. was supported by an Australian Research Council Discovery Project. The research of J.D. was supported by an Australian Research Council Queen Elizabeth 2 Fellowship.

References

1. Aistleitner, C.: Covering numbers, dyadic chaining and discrepancy. J. Complex. **27**(6), 531–540 (2011)
2. Alexander, K.S.: The central limit theorem for empirical processes on Vapnik–Červonenkis classes. Ann. Probab. **15**(1), 178–203 (1987)
3. Alexander, K.S., Talagrand, M.: The law of the iterated logarithm for empirical processes on Vapnik–Červonenkis classes. J. Multivar. Anal. **30**(1), 155–166 (1989)
4. Andrievskii, V.V., Blatt, H.-P., Götz, M.: Discrepancy estimates on the sphere. Monatshefte Math. **128**(3), 179–188 (1999)
5. Barvinok, A.: A Course in Convexity. Graduate Studies in Mathematics, vol. 54. American Mathematical Society, Providence (2002)
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. Beck, J.: On the discrepancy of convex plane sets. Monatshefte Math. **105**(2), 91–106 (1988)
8. Beck, J., Chen, W.W.L.: Irregularities of Distribution. Cambridge Tracts in Mathematics, vol. 89. Cambridge University Press, Cambridge (2008). Reprint of the 1987 original. [MR0903025]
9. Bendito, E., Carmona, A., Encinas, A.M., Gesto, J.M., Gómez, A., Mouriño, C., Sánchez, M.T.: Computational cost of the Fekete problem. I. The forces method on the 2-sphere. J. Comput. Phys. **228**(9), 3288–3306 (2009)
10. Bondarenko, A., Radchenko, D., Viazovska, M.: Optimal asymptotic bounds for spherical designs (2011). arXiv:1009.4407v3 [math.MG]
11. Brauchart, J.S.: Optimal logarithmic energy points on the unit sphere. Math. Comput. **77**(263), 1599–1613 (2008)
12. Brauchart, J.S., Dick, J.: Quasi-Monte Carlo rules for numerical integration over the unit sphere $\mathbb{S}^2$. Numer. Math. **121**(3), 473–502 (2012)
13. Brauchart, J.S., Saff, E.B., Sloan, I.H., Womersley, R.S.: Sequences of approximate spherical designs. Manuscript, 62 pages
14. Brauchart, J.S., Womersley, R.S.: Numerical integration over the unit sphere, $\mathbb{L}_2$-discrepancy and sum of distances. Manuscript, 26 pages
15. Cassels, J.W.S.: An Introduction to the Geometry of Numbers. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, vol. 99. Springer, Berlin (1959)
16. Chen, X., Frommer, A., Lang, B.: Computational existence proofs for spherical t-designs. Numer. Math. **117**(2), 289–305 (2011)
17. Damelin, S.B., Grabner, P.J.: Energy functionals, numerical integration and asymptotic equidistribution on the sphere. J. Complex. **19**(3), 231–246 (2003). Numerical integration and its complexity (Oberwolfach, 2001)
18. Damelin, S.B., Grabner, P.J.: Corrigendum to: "Energy functionals, numerical integration and asymptotic equidistribution on the sphere" [J. Complexity **19** (2003), no. 3, 231–246; mr1984111]. J. Complex. **20**(6), 883–884 (2004)
19. Delsarte, P., Goethals, J.M., Seidel, J.J.: Spherical codes and designs. Geom. Dedic. **6**(3), 363–388 (1977)
20. Dick, J., Pillichshammer, F.: Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration. Cambridge University Press, Cambridge (2010)
21. Dudley, R.M.: Central limit theorems for empirical measures. Ann. Probab. **6**(6), 899–929 (1979). 1978
22. Durst, M., Dudley, R.M.: Empirical processes, Vapnik–Chervonenkis classes and Poisson processes. Probab. Math. Stat. **1**(2), 109–115 (1981). 1980
23. Faure, H.: Discrépance de suites associées à un système de numération (en dimension s). Acta Arith. **41**(4), 337–351 (1982)
24. Götz, M.: On the distribution of weighted extremal points on a surface in $\mathbf{R}^d$, $d \geq 3$. Potential Anal. **13**(4), 345–359 (2000)
25. Grabner, P.J.: Erdős–Turán type discrepancy bounds. Monatshefte Math. **111**(2), 127–135 (1991)
26. Grabner, P.J., Tichy, R.F.: Spherical designs, discrepancy and numerical integration. Math. Comput. **60**(201), 327–336 (1993)
27. Hannay, J.H., Nye, J.F.: Fibonacci numerical integration on a sphere. J. Phys. A **37**(48), 11591–11601 (2004)

28. Hardin, D.P., Saff, E.B.: Discretizing manifolds via minimum energy points. Not. Am. Math. Soc. **51**(10), 1186–1194 (2004)
29. Hardin, R.H., Sloane, N.J.A.: McLaren's improved snub cube and other new spherical designs in three dimensions. Discrete Comput. Geom. **15**(4), 429–441 (1996)
30. Haussler, D.: Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik–Chervonenkis dimension. J. Comb. Theory, Ser. A **69**(2), 217–232 (1995)
31. Heinrich, S., Novak, E., Wasilkowski, G.W., Woźniakowski, H.: The inverse of the star-discrepancy depends linearly on the dimension. Acta Arith. **96**(3), 279–302 (2001)
32. Korevaar, J.: Fekete extreme points and related problems. In: Approximation Theory and Function Series, Budapest, 1995. Bolyai Soc. Math. Stud., vol. 5, pp. 35–62. János Bolyai Math. Soc., Budapest (1996)
33. Korevaar, J., Meyers, J.L.H.: Spherical Faraday cage for the case of equal point charges and Chebyshev-type quadrature on the sphere. Integral Transforms Spec. Funct. **1**(2), 105–117 (1993)
34. Kuipers, L., Niederreiter, H.: Uniform Distribution of Sequences. Pure and Applied Mathematics. Wiley-Interscience, New York (1974)
35. Larcher, G.: Optimale Koeffizienten bezüglich zusammengesetzter Zahlen. Monatshefte Math. **100**(2), 127–135 (1985)
36. Larcher, G.: On the distribution of s-dimensional Kronecker-sequences. Acta Arith. **51**(4), 335–347 (1988)
37. Larcher, G.: Corrigendum: "On the distribution of s-dimensional Kronecker-sequences" [Acta Arith. **51** (1988), no. 4, 335–347; MR0971085 (90f:11065)]. Acta Arith. **60**(1), 93–95 (1991)
38. Li, X.-J., Vaaler, J.D.: Some trigonometric extremal functions and the Erdős–Turán type inequalities. Indiana Univ. Math. J. **48**(1), 183–236 (1999)
39. Lubotzky, A., Phillips, R., Sarnak, P.: Hecke operators and distributing points on the sphere. I. Commun. Pure Appl. Math. **39**(S, suppl.), S149–S186 (1986). Frontiers of the Mathematical Sciences: 1985 (New York, 1985)
40. Lubotzky, A., Phillips, R., Sarnak, P.: Hecke operators and distributing points on S^2. II. Commun. Pure Appl. Math. **40**(4), 401–420 (1987)
41. Narcowich, F.J., Sun, X., Ward, J.D., Wu, Z.: LeVeque type inequalities and discrepancy estimates for minimal energy configurations on spheres. J. Approx. Theory **162**(6), 1256–1278 (2010)
42. Niederreiter, H.: Random Number Generation and Quasi-Monte Carlo Methods. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 63. SIAM, Philadelphia (1992)
43. Niederreiter, H., Sloan, I.H.: Integration of nonperiodic functions of two variables by Fibonacci lattice rules. J. Comput. Appl. Math. **51**(1), 57–70 (1994)
44. Philipp, W.: Empirical distribution functions and uniform distribution mod 1. In: Proc. Conf. Diophantine Approximation and Its Application, Washington, DC, 1972, pp. 211–234. Academic Press, New York (1973)
45. Saff, E.B., Kuijlaars, A.B.J.: Distributing many points on a sphere. Math. Intell. **19**(1), 5–11 (1997)
46. Sjögren, P.: Estimates of mass distributions from their potentials and energies. Ark. Mat. **10**, 59–77 (1972)
47. Sloan, I.H., Womersley, R.S.: Extremal systems of points and numerical integration on the sphere. Adv. Comput. Math. **21**(1–2), 107–125 (2004)
48. Sloan, I.H., Womersley, R.S.: A variational characterisation of spherical designs. J. Approx. Theory **159**(2), 308–318 (2009)
49. Sobol', I.M.: Distribution of points in a cube and approximate evaluation of integrals. Ž. Vyčisl. Mat. i Mat. Fiz. **7**, 784–802 (1967)
50. Stolarsky, K.B.: Sums of distances between points on a sphere. II. Proc. Am. Math. Soc. **41**, 575–582 (1973)
51. Talagrand, M.: Sharper bounds for Gaussian and empirical processes. Ann. Probab. **22**(1), 28–76 (1994)
52. Talagrand, M.: Vapnik–Chervonenkis type conditions and uniform Donsker classes of functions. Ann. Probab. **31**(3), 1565–1582 (2003)
53. Vapnik, V.N., Červonenkis, A.J.: The uniform convergence of frequencies of the appearance of events to their probabilities. Teor. Verojatnost. i Primenen. **16**, 264–279 (1971)
54. Wagner, G.: Erdős–Turán inequalities for distance functions on spheres. Mich. Math. J. **39**(1), 17–34 (1992)