Discrete & Computational Geometry (2022) 68:540–565
https://doi.org/10.1007/s00454-022-00396-7

Efficient Quantisation and Weak Covering of High Dimensional Cubes

Jack Noonan[1] · **Anatoly Zhigljavsky**[1]

Received: 22 May 2020 / Revised: 22 January 2022 / Accepted: 2 February 2022 /
Published online: 3 June 2022

Abstract
Let $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ be a design; that is, a collection of n points $Z_j \in [-1, 1]^d$. We study the quality of quantisation of $[-1, 1]^d$ by the points of $\mathbb{Z}_n$ and the problem of quality of coverage of $[-1, 1]^d$ by $\mathcal{B}_d(\mathbb{Z}_n, r)$, the union of balls centred at $Z_j \in \mathbb{Z}_n$. We concentrate on the cases where the dimension d is not small, $d \geq 5$, and n is not too large, $n \leq 2^d$. We define the design $\mathbb{D}_{n,\delta}$ as a 2^{d-1} design defined on vertices of the cube $[-\delta, \delta]^d$, $0 \leq \delta \leq 1$. For this design, we derive a closed-form expression for the quantisation error and very accurate approximations for the coverage area $\mathrm{vol}\,([-1, 1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))$. We provide results of a large-scale numerical investigation confirming the accuracy of the developed approximations and the efficiency of the designs $\mathbb{D}_{n,\delta}$.

Keywords Covering · Quantisation · Facility location · Space-filling · Computer experiments · High dimension · Voronoi set

Mathematics Subject Classification 11H31 · 05B40 · 52C17

1 Introduction

1.1 Main Notation

- $\|\cdot\|$: the Euclidean norm;
- $\mathcal{B}_d(Z, r) = \{Y \in \mathbb{R}^d : \|Y - Z\| \leq r\}$: d-dimensional ball of radius r centered at $Z \in \mathbb{R}^d$;

Editor in Charge: Kenneth Clarkson

Jack Noonan
Noonanj1@cardiff.ac.uk

Anatoly Zhigljavsky
ZhigljavskyAA@cardiff.ac.uk

[1] School of Mathematics, Cardiff University, Cardiff, CF244AG, UK

- $\mathbb{Z}_n = \{Z_1, \ldots, Z_n\}$: a design; that is, a collection of n points $Z_j \in \mathbb{R}^d$;
- $\mathcal{B}_d(\mathbb{Z}_n, r) = \bigcup_{j=1}^n \mathcal{B}_d(Z_j, r)$;
- $C_d(\mathbb{Z}_n, r) = \mathrm{vol}\,([-1,1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))/2^d$: the proportion of the cube $[-1,1]^d$ covered by $\mathcal{B}_d(\mathbb{Z}_n, r)$;
- vectors in $\mathbb{R}^d$ are row-vectors;
- for any $a \in \mathbb{R}$, $\boldsymbol{a} = (a, a, \ldots, a) \in \mathbb{R}^d$.

1.2 Main Problems of Interest

We will study the following two main characteristics of designs $\mathbb{Z}_n = \{Z_1, \ldots, Z_n\} \subset \mathbb{R}^d$.

1. Quantization error. Let $X = (x_1, \ldots, x_d)$ be uniform random vector on $[-1,1]^d$. The mean squared quantisation error for a design $\mathbb{Z}_n$ is defined by

$$\theta(\mathbb{Z}_n) = \mathbb{E}_X \varrho^2(X, \mathbb{Z}_n), \quad \text{where} \quad \varrho^2(X, \mathbb{Z}_n) = \min_{Z_i \in \mathbb{Z}_n} \|X - Z_i\|^2. \tag{1}$$

2. Weak covering. Denote the proportion of the cube $[-1,1]^d$ covered by the union of n balls $\mathcal{B}_d(\mathbb{Z}_n, r) = \bigcup_{j=1}^n \mathcal{B}_d(Z_j, r)$ by

$$C_d(\mathbb{Z}_n, r) := \frac{\mathrm{vol}\,([-1,1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))}{2^d}.$$

For given radius $r > 0$, the union of n balls $\mathcal{B}_d(\mathbb{Z}_n, r)$ makes the $(1-\gamma)$-coverage of the cube $[-1,1]^d$ if

$$C_d(\mathbb{Z}_n, r) = 1 - \gamma. \tag{2}$$

Complete coverage corresponds to $\gamma = 0$. In this paper, the complete coverage of $[-1,1]^d$ will not be enforced and we will mostly be interested in *weak covering*, that is, achieving (2) with some small $\gamma > 0$.

Two n-point designs $\mathbb{Z}_n$ and $\mathbb{Z}'_n$ will be differentiated in terms of performance as follows: (a) $\mathbb{Z}_n$ dominates $\mathbb{Z}'_n$ for quantisation if $\theta(\mathbb{Z}_n) < \theta(\mathbb{Z}'_n)$; (b) if for a given $\gamma \geq 0$, $C_d(\mathbb{Z}_n, r_1) = C_d(\mathbb{Z}'_n, r_2) = 1 - \gamma$ and $r_1 < r_2$, then the design $\mathbb{Z}_n$ provides a more efficient $(1-\gamma)$-coverage than $\mathbb{Z}'_n$ and is therefore preferable. In Sect. 1.4 we extend these definitions by allowing the two designs to have different number of points and, moreover, to have different dimensions.

Numerical construction of n-point designs with moderate values of n with good quantisation and coverage properties has recently attracted much attention in view of diverse applications in several fields including computer experiments [7, 8, 11], global optimization [15], function approximation [12, 13], and numerical integration [9]. Such designs are often referred to as *space-filling designs*. Readers can find many additional references in the citations above. Unlike the exiting literature on space-filling, we concentrate on theoretical properties of a family of very efficient designs and derivation of accurate approximations for the characteristics of interest.

1.3 Relation Between Quantisation and Weak Coverage

The two characteristics, $C_d(\mathbb{Z}_n, r)$ and $\theta(\mathbb{Z}_n)$, are related: $C_d(\mathbb{Z}_n, r)$, as a function of $r \geq 0$, is the c.d.f. of the r.v. $\varrho(X, \mathbb{Z}_n)$ while $\theta(\mathbb{Z}_n)$ is the second moment of the distribution with this c.d.f.:

$$\theta(\mathbb{Z}_n) = \int_{r \geq 0} r^2 dC_d(\mathbb{Z}_n, r). \tag{3}$$

In particular, this yields that if an n-point design $\mathbb{Z}_n^*$ maximises, in the set of all n-point designs, $C_d(\mathbb{Z}_n, r)$ for all $r > 0$, then it also minimises $\theta(\mathbb{Z}_n)$. Moreover, if r.v. $\varrho(X, \mathbb{Z}_n)$ stochastically dominates $\varrho(X, \mathbb{Z}_n')$, so that $C_d(\mathbb{Z}_n', r) \leq C_d(\mathbb{Z}_n, r)$ for all $r \geq 0$ and the inequality is strict for at least one r, then $\theta(\mathbb{Z}_n) < \theta(\mathbb{Z}_n')$.

The relation (3) can alternatively be written as

$$\theta(\mathbb{Z}_n) = \int_{r \geq 0} r\, dC_d(\mathbb{Z}_n, \sqrt{r}), \tag{4}$$

where $C_d(\mathbb{Z}_n, \sqrt{r})$, considered as a function of r, is the c.d.f. of the r.v. $\varrho^2(X, \mathbb{Z}_n)$ and hence $\theta(\mathbb{Z}_n)$ is the mean of this r.v. Relation (4) is simply another form of (1).

1.4 Re-Normalised Versions and Formulation of Optimal Design Problems

In view of (13), the naturally defined re-normalised version of $\theta(\mathbb{Z}_n)$ is $Q_d(\mathbb{Z}_n) = n^{2/d}\theta(\mathbb{Z}_n)/(4d)$. From (4) and (3), $Q_d(\mathbb{Z}_n)$ is the expectation of $n^{2/d}\varrho^2(X, \mathbb{Z}_n)/(4d)$ and the second moment of the r.v. $n^{1/d}\varrho(X, \mathbb{Z}_n)/(2\sqrt{d})$ respectively. This suggests the following re-normalization of the radius r with respect to n and d:

$$R = \frac{n^{1/d} r}{2\sqrt{d}}. \tag{5}$$

We can then define optimal designs as follows. Let d be fixed, $\mathcal{Z}_n = \{\mathbb{Z}_n\}$ be the set of all n-point designs and $\mathcal{Z} = \bigcup_{n=1}^{\infty} \mathcal{Z}_n$ be the set of all designs.

Definition 1.1 The design $\mathbb{Z}_m^*$ with some m is optimal for quantisation in $[-1, 1]^d$, if

$$Q_d(\mathbb{Z}_m^*) = \min_n \min_{\mathbb{Z}_n \in \mathcal{Z}_n} Q_d(\mathbb{Z}_n) = \min_{\mathbb{Z} \in \mathcal{Z}} Q_d(\mathbb{Z}). \tag{6}$$

Definition 1.2 The design $\mathbb{Z}_m^*$ with some m is optimal for $(1-\gamma)$-coverage of $[-1, 1]^d$, if

$$R_{1-\gamma}(\mathbb{Z}_m^*) = \min_n \min_{\mathbb{Z}_n \in \mathcal{Z}_n} R_{1-\gamma}(\mathbb{Z}_n) = \min_{\mathbb{Z} \in \mathcal{Z}} R_{1-\gamma}(\mathbb{Z}). \tag{7}$$

Here $0 \leq \gamma \leq 1$ and for a given design $\mathbb{Z}_n \in \mathcal{Z}_n$,

$$R_{1-\gamma}(\mathbb{Z}_n) = \frac{n^{1/d} r_{1-\gamma}(\mathbb{Z}_n)}{2\sqrt{d}}, \tag{8}$$

where $r_{1-\gamma}(\mathbb{Z}_n)$ is defined as the smallest r such that $C_d(\mathbb{Z}_n, r) = 1-\gamma$.

Importance of the factor $\sqrt{d}$ in (5) will be seen in Sect. 3.5 where we shall study the asymptotical behaviour of $(1-\gamma)$-coverings for large d.

1.5 Thickness of Covering

Let $\gamma = 0$ in Definition 1.2. Then $r_1(\mathbb{Z}_n)$ is the covering radius associated with $\mathbb{Z}_n$ so that the union of the balls $\mathcal{B}_d(\mathbb{Z}_n, r)$ with $r = r_1(\mathbb{Z}_n)$ makes a coverage of $[-1,1]^d$. Let us tile up the whole space $\mathbb{R}^d$ with the translations of the cube $[-1,1]^d$ and corresponding translations of the balls $\mathcal{B}_d(\mathbb{Z}_n, r)$. This would make a full coverage of the whole space; denote this space coverage by $\mathcal{B}_d(\mathbb{Z}_{(n)}, r)$. The thickness Θ of any space covering is defined, see [1, (1), Chap. 2], as the average number of balls containing a point of the whole space. In our case of $\mathcal{B}_d(\mathbb{Z}_{(n)}, r)$, the thickness is

$$\Theta(\mathcal{B}_d(\mathbb{Z}_{(n)}, r)) = \frac{n \operatorname{vol}(\mathcal{B}_d(0,r))}{\operatorname{vol}([-1,1]^d)} = \frac{nr^d \operatorname{vol}(\mathcal{B}_d(0,1))}{2^d}.$$

The normalised thickness, θ, is the thickness Θ divided by $\operatorname{vol}(\mathcal{B}_d(0,1))$, the volume of the unit ball, see [1, (2), Chap. 2]. In the case of $\mathcal{B}_d(\mathbb{Z}_{(n)}, r)$, the normalised thickness is

$$\theta(\mathcal{B}_d(\mathbb{Z}_{(n)}, r)) = \frac{nr^d}{2^d} = d^{d/2}[R_1(\mathbb{Z}_{(n)})]^d,$$

where we have recalled that $r = r_1(\mathbb{Z}_n)$ and $R_{1-\gamma}(\mathbb{Z}_n) = n^{1/d} r_{1-\gamma}(\mathbb{Z}_n)/(2\sqrt{d})$ for any $0 \le \gamma \le 1$. We can thus define the normalised thickness of the covering of the cube by the same formula and extend it to any $0 \le \gamma \le 1$:

Definition 1.3 Let $\mathcal{B}_d(\mathbb{Z}_n, r)$ be a $(1-\gamma)$-coverage of the cube $[-1,1]^d$ with $0 \le \gamma \le 1$. Its normalised thickness is defined by

$$\theta(\mathcal{B}_d(\mathbb{Z}_n, r)) = (\sqrt{d}\, R)^d, \tag{9}$$

where $R = n^{1/d} r/(2\sqrt{d})$, see (5).

In view of (9), we can reformulate the definition (7) of the $(1-\gamma)$-covering optimal design by saying that this design minimises (normalised) thickness in the set of all $(1-\gamma)$-covering designs.

1.6 The Design of the Main Interest

We will be mostly interested in the following n-point design $\mathbb{Z}_n = \mathbb{D}_{n,\delta}$ defined only for $n = 2^{d-1}$:

Design $\mathbb{D}_{n,\delta}$: *a* 2^{d-1} *design defined on vertices of the cube* $[-\delta,\delta]^d$, $0 \le \delta \le 1$.

For theoretical comparison with design $\mathbb{D}_{n,\delta}$, we shall consider the following simple design, which extends to the integer point lattice Z^d (shifted by $\mathbf{1/2}$) in the whole space $\mathbb{R}^d$:

Design $\mathbb{D}_n^{(0)}$: *the collection of* 2^d *points* $(\pm 1/2, \ldots, \pm 1/2)$, *all vertices of the cube* $[-1/2, 1/2]^d$.

Without loss of generality, while considering the design $\mathbb{D}_{n,\delta}$ we assume that the point $Z_1 \in \mathbb{D}_{n,\delta} = \{Z_1, \ldots, Z_n\}$ is $Z_1 = \boldsymbol{\delta} = (\delta, \ldots, \delta)$. Similarly, the first point in $\mathbb{D}_n^{(0)}$ is $Z_1 = \mathbf{1/2} = (1/2, \ldots, 1/2)$. Note also that for numerical comparisons, in Sect. 4 we shall introduce one more design.

The design $\mathbb{D}_{n,1/2}$ extends to the lattice D_d (shifted by $\mathbf{1/2}$) containing points $X = (x_1, \ldots, x_d)$ with integer components satisfying $x_1 + \ldots + x_d = 0 \pmod 2$, see [1, Sect. 7.1, Chap. 4]; this lattice is sometimes called "checkerboard lattice". The motivation to theoretically study the design $\mathbb{D}_{n,\delta}$ is a consequence of numerical results reported in [14] and [4], where the present authors have considered n-point designs in d-dimensional cubes providing good coverage and quantisation and have shown that for all dimensions $d \geq 7$, the design $\mathbb{D}_{n,\delta}$ with suitable δ provides the best quantisation and coverage per point among all other designs considered. Aiming at practical applications mentioned in Sect. 1.2, our aim was to consider the designs with n which is not too large and in any case does not exceed 2^d.

If the number of points n in a design is much larger than 2^d, then we may use the following scheme of construction of efficient quantisers in the cube $[-1, 1]^d$: (a) construct one of the very efficient lattice space quantisers, see [1, Sect. 3, Chap. 2], (b) take the lattice points belonging to a very large cube, and (c) scale the chosen large cube to $[-1, 1]^d$. In view of [2, Thm. 8.9], as $n \to \infty$, the normalised quantisation error $Q_d(\mathbb{Z}_n)$ of the sequence of resulting designs $\mathbb{Z}_n$ converges to the respective quantisation error of the lattice space quantiser. However, for any given n the study of quantisation error of such designs is difficult (both, numerically and theoretically) as there could be several non-congruent types of Voronoi cells due to boundary conditions. Note also that the boundary conditions make significant difference in relative efficiencies of the resulting designs. In particular, the checkerboard lattice D_d is better than the integer-point lattice $\mathbb{Z}^d$ for all $d \geq 3$ as a space quantiser and becomes the best lattice space quantiser for $d = 4$ but in the case of cube $[-1, 1]^d$, the design $\mathbb{D}_{n,\delta}$ (with optimal δ) makes a better quantiser than $\mathbb{D}_n^{(0)}$ for $d \geq 7$ only; see Sect. 2.4 for theoretical and numerical comparison of the two designs.

1.7 Structure of the Rest of the Paper and the Main Results

In Sect. 2 we study $Q_d(\mathbb{D}_{n,\delta})$, the normalised mean squared quantisation error for the design $\mathbb{D}_{n,\delta}$. There are two important results, Theorems 2.2 and 2.3. In Theorem 2.2, we derive the explicit form for the Voronoi cells for the points of the design $\mathbb{D}_{n,\delta}$ and in Theorem 2.3 we derive a closed-form expression for $Q_d(\mathbb{D}_{n,\delta})$ for any $\delta > 0$. As a consequence, in Corollary 2.4 we determine the optimal value of δ.

The main result of Sect. 3 is Theorem 3.2, where we derive closed-form expressions (in terms of $C_{d,Z,r}$, the fraction of the cube $[-1, 1]^d$ covered by a ball $\mathcal{B}_d(Z, r)$) for the coverage area with vol $([-1,1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))$. Then, using accurate approximations for $C_{d,Z,r}$, we derive approximations for vol $([-1,1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))$. In Theorem 3.4 we derive asymptotic expressions for the $(1-\gamma)$-coverage radius for the design $\mathbb{D}_{d,1/2}$ and show that for any $\gamma > 0$, the ratio of the $(1-\gamma)$-coverage radius to the 1-coverage radius tends to $1/\sqrt{3}$ as $d \to \infty$. Numerical results of Sect. 3.5 confirm that even for rather small d, the 0.999-coverage radius is much smaller than the 1-coverage radius providing the full coverage.

In Sect. 4 we demonstrate that the approximations developed in Sect. 3 are very accurate and make a comparative study of selected designs used for quantisation and covering.

In Appendices A–C, we provide proofs of the most technical results. In Appendix D, for completeness, we briefly derive an approximation for $C_{d,Z,r}$ with arbitrary d, Z, and r.

The two most important contributions of this paper are: a) derivation of the closed-form expression for the quantisation error for the design $\mathbb{D}_{n,\delta}$, and b) derivation of accurate approximations for the coverage area vol $([-1,1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))$ for the design $\mathbb{D}_{n,\delta}$.

2 Quantisation

2.1 Reformulation in Terms of the Voronoi Cells

Consider any n-point design $\mathbb{Z}_n = \{Z_1, \ldots, Z_n\}$. The Voronoi cell $V(Z_i)$ for $Z_i \in \mathbb{Z}_n$ is defined as

$$V(Z_i) = \{x \in [-1,1]^d : \|Z_i - x\| \le \|Z_j - x\| \text{ for } j \ne i\}.$$

The mean squared quantisation error $\theta(\mathbb{Z}_n)$ introduced in (1) can be written in terms of the Voronoi cells as follows:

$$\theta(\mathbb{Z}_n) = \mathbb{E}_X \min_{i=1,\ldots,n} \|X - Z_i\|^2 = \frac{1}{\mathrm{vol}([-1,1]^d)} \sum_{i=1}^{n} \int_{V(Z_i)} \|X - Z_i\|^2 dX, \quad (10)$$

where $X = (x_1, \ldots, x_d)$ and $dX = dx_1 dx_2 \ldots dx_d$.

This reformulation has significant benefit when the design $\mathbb{Z}_n$ has certain structure. In particular, if all of the Voronoi cells $V(Z_i)$, $i = 1, \ldots, n$, are congruent, then we can simplify (10) to

$$\theta(\mathbb{Z}_n) = \frac{1}{\mathrm{vol}(V(Z_1))} \int_{V(Z_1)} \|X - Z_1\|^2 dX. \quad (11)$$

In Sect. 2.4, this formula will be the starting point for derivation of the closed-form expression for $\theta(\mathbb{Z}_n)$ for the design $\mathbb{D}_{n,\delta}$.

2.2 Re-Normalisation of the Quantisation Error

To compare efficiency of n-point designs $\mathbb{Z}_n$ with different values of n, one must suitably normalise $\theta(\mathbb{Z}_n)$ with respect to n. Specialising a classical characteristic for quantisation in space, as formulated in [1, (86), Chap. 2], we obtain

$$Q_d(\mathbb{Z}_n) = \frac{1}{d} \cdot \frac{(1/n)\sum_{i=1}^{n} \int_{V(Z_i)} \|X - Z_i\|^2 \, dX}{\left[(1/n)\sum_{i=1}^{n} \mathrm{vol}(V(Z_i))\right]^{1+2/d}}. \tag{12}$$

Note that $Q_d(\mathbb{Z}_n)$ is re-normalised with respect to dimension d too, not only with respect to n. Normalization $1/d$ with respect to d is very natural in view of the definition of the Euclidean norm. Using (10), for the cube $[-1, 1]^d$, (12) can be expressed as

$$Q_d(\mathbb{Z}_n) = \frac{n^{2/d}\theta(\mathbb{Z}_n)}{d\left[\sum_{i=1}^{n} \mathrm{vol}(V(Z_i))\right]^{2/d}} = \frac{n^{2/d}\theta(\mathbb{Z}_n)}{d\,\mathrm{vol}([-1, 1]^d)^{2/d}} = \frac{n^{2/d}}{4d}\theta(\mathbb{Z}_n). \tag{13}$$

2.3 Voronoi Cells for $\mathbb{D}_{n,\delta}$

Proposition 2.1 *Consider the design $\mathbb{D}_{n,\delta}^{(0)}$, the collection of $n = 2^d$ points $(\pm\delta, \dots, \pm\delta)$, $0 < \delta < 1$. The Voronoi cells for this design are all congruent. The Voronoi cell for the point $\boldsymbol{\delta} = (\delta, \delta, \dots, \delta)$ is the cube*

$$C_0 = \{X = (x_1, \dots, x_d) \in \mathbb{R}^d : 0 \le x_i \le 1,\ i = 1, 2, \dots, d\}. \tag{14}$$

Proof Consider the Voronoi cells created by the design $\mathbb{D}_{n,\delta}^{(0)}$ in the whole space $\mathbb{R}^d$. For the point $\boldsymbol{\delta} = (\delta, \delta, \dots, \delta)$, the Voronoi cell is clearly $\{X = (x_1, \dots, x_d) : x_i \ge 0\}$. By intersecting this set with the cube $[-1, 1]^d$ we obtain (14). □

Theorem 2.2 *The Voronoi cells of the design $\mathbb{D}_{n,\delta} = \{Z_1, \dots, Z_n\}$ are all congruent. The Voronoi cell for the point $Z_1 = \boldsymbol{\delta} = (\delta, \delta, \dots, \delta) \in \mathbb{R}^d$ is*

$$V(Z_1) = C_0 \cup \bigcup_{j=1}^{d} U_j, \tag{15}$$

where C_0 is the cube (14) and

$$U_j = \{X = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d : -1 \le x_j \le 0,\ |x_j| \le x_k \le 1 \ \text{for all}\ k \ne j\}. \tag{16}$$

The volume of $V(Z_1)$ is $\mathrm{vol}(V(Z_1)) = 2$.

Proof The design $\mathbb{D}_{n,\delta}$ is symmetric with respect to all components implying that all $n = 2^{d-1}$ Voronoi cells are congruent immediately yielding that their volumes equal 2.

Consider $V(Z_1)$ with $Z_1 = \boldsymbol{\delta}$. Since $\mathbb{D}_{n,\delta} \subset \mathbb{D}_{n,\delta}^{(0)}$, where design $\mathbb{D}_{n,\delta}^{(0)}$ is introduced in Proposition 2.1, and C_0 is the Voronoi set of $\boldsymbol{\delta}$ for design $\mathbb{D}_{n,\delta}^{(0)}$, $C_0 \subset V(\boldsymbol{\delta})$ for design $\mathbb{D}_{n,\delta}$ too. Consider the d cubes adjacent to C_0:

$$C_j = \{X = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d : -1 \le x_j \le 0,\ 0 \le x_i \le 1 \text{ for all } i \ne j\}; \tag{17}$$

$j = 1, \ldots, d$. A part of each cube C_j belongs to $V(Z_1)$. This part is exactly the set U_j defined by (16). This can be seen as follows. A part of C_j also belongs to the Voronoi set of the point $X_{jk} = \boldsymbol{\delta} - 2\delta e_j - 2\delta e_k$, where $e_l = (0, \ldots, 0, 1, 0, \ldots, 0)$ with 1 placed at l-th place; all components of X_{jk} are δ except j-th and k-th components which are $-\delta$. We have to have $|x_j| \le x_k$, for a point $X \in C_j$ to be closer to Z_1 than to X_{jk}. Joining all constraints for $X = (x_1, x_2, \ldots, x_d) \in C_j$ $(k = 1, \ldots, d, k \ne j)$ we obtain (16) and hence (15). □

2.4 Explicit Formulae for the Quantisation Error

Theorem 2.3 *For the design $\mathbb{D}_{n,\delta}$ with $0 \le \delta \le 1$, we obtain*

$$\theta(\mathbb{D}_{n,\delta}) = d\left(\delta^2 - \delta + \frac{1}{3}\right) + \frac{2\delta}{d+1}, \tag{18}$$

$$Q_d(\mathbb{D}_{n,\delta}) = 2^{-2/d}\left(\delta^2 - \delta + \frac{1}{3} + \frac{2\delta}{d(d+1)}\right). \tag{19}$$

Proof To compute $\theta(\mathbb{D}_{n,\delta})$, we use (11), where, in view of Theorem 2.2, $\mathrm{vol}(V(Z_1)) = 2$. Using the expression (15) for $V(Z_1)$ with $Z_1 = \boldsymbol{\delta}$, we obtain

$$\begin{aligned} \theta(\mathbb{Z}_n) &= \frac{1}{2}\int_{V(Z_1)} \|X - Z_1\|^2 \, dX \\ &= \frac{1}{2}\left[\int_{C_0} \|X - Z_1\|^2 \, dX + d\int_{U_1} \|X - Z_1\|^2 \, dX\right]. \end{aligned} \tag{20}$$

Consider the two terms in (20) separately. The first term is easy:

$$\begin{aligned} \int_{C_0} \|X - Z_1\|^2 \, dX &= \int_{C_0} \sum_{i=1}^{d} (x_i - \delta)^2 dx_1 \ldots dx_d \\ &= d\int_0^1 (x - \delta)^2 dx = d\left(\delta^2 - \delta + \frac{1}{3}\right). \end{aligned} \tag{21}$$

For the second term we have

$$\begin{aligned}\int_{U_1} \|X - Z_1\|^2 \, dX &= \int_{-1}^{0} \left[\int_{|x_1|}^{1} \cdots \int_{|x_1|}^{1} \sum_{i=1}^{d} (x_i - \delta)^2 \, dx_2 \ldots dx_d \right] dx_1 \\ &= \int_{-1}^{0} (x_1 - \delta)^2 (1 + x_1)^{d-1} \, dx_1 \\ &\quad + (d-1) \int_{-1}^{0} (1 + x_1)^{d-2} \int_{|x_1|}^{1} (x_2 - \delta)^2 \, dx_2 \, dx_1 \\ &= \delta^2 - \delta + \frac{1}{3} + \frac{4\delta}{d(d+1)}. \end{aligned} \tag{22}$$

Inserting the obtained expressions into (20) we obtain (18). The expression (19) is a consequence of (13), (18), and $n = 2^{d-1}$. □

A simple consequence of Theorem 2.3 is the following corollary.

Corollary 2.4 *The optimal value of δ minimising $\theta(\mathbb{D}_{n,\delta})$ and $Q_d(\mathbb{D}_{n,\delta})$ is*

$$\delta^* = \frac{1}{2} - \frac{1}{d(d+1)}; \tag{23}$$

for this value,

$$Q_d(\mathbb{D}_{n,\delta^*}) = \min_{\delta} Q_d(\mathbb{D}_{n,\delta}) = 2^{-2/d} \left[\frac{1}{12} + \frac{d^2 + d - 1}{(d+1)^2 d^2} \right]. \tag{24}$$

Let us make several remarks.

1. The value δ^* can be alternatively characterised by the well-known optimality condition of a general design saying that each design point of an optimal quantiser must be a centroid of the related Voronoi cell; see e.g. [10]. Specifically, each design point $Z_i \in \mathbb{D}_{n,\delta}$ is the centroid of $V(Z_i)$ if and only if $\delta = \delta^*$.
2. From (19), for the design $\mathbb{D}_{n,1/2}$ we get

$$Q_d(\mathbb{D}_{n,1/2}) = 2^{-2/d} \left[\frac{1}{12} + \frac{1}{(d+1)d} \right]; \tag{25}$$

   this value is always slightly larger than (24).
3. For the one-point design $\mathbb{D}^{(0)} = \{0\}$ with the single point 0 and the design $\mathbb{D}_n^{(0)}$ with $n = 2^d$ points $(\pm 1/2, \ldots, \pm 1/2)$ we have $Q_d(\mathbb{D}^{(0)}) = Q_d(\mathbb{D}_n^{(0)}) = 1/12$, which coincides with the value of Q_d in the case of space quantisation by the integer-point lattice Z^d, see [1, Chaps. 2 and 21].
4. The quantisation error (25) for the design $\mathbb{D}_{n,1/2}$ have almost exactly the same form as the quantisation error for the 'checkerboard lattice' D_d in $\mathbb{R}^d$; the difference is in the factor 1/2 in the last term in (25), see [1, (27), Chap. 21]. Naturally,

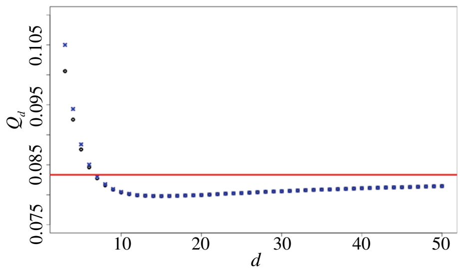

Fig. 1 $Q_d(\mathbb{D}_{n,\delta^*})$ and $Q_d(\mathbb{D}_{n,1/2})$ as functions of d and $Q_d(\mathbb{D}_n^{(0)}) = 1/12$; $d = 3, \ldots, 50$

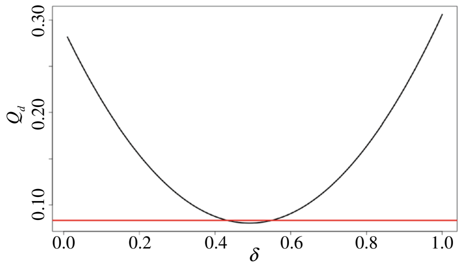

Fig. 2 $Q_d(\mathbb{D}_{n,\delta})$ as a function of δ and $Q_d(\mathbb{D}_n^{(0)}) = 1/12$; $d = 10$

the quantisation error Q_d for D_d in $\mathbb{R}^d$ is slightly smaller than Q_d for $\mathbb{D}_{n,1/2}$ in $[-1, 1]^d$.

5. The optimal value of δ in (23) is smaller than 1/2. This is caused by a non-symmetrical shape of the Voronoi cells $V(Z_j)$ for designs $\mathbb{D}_{n,\delta}$, which is clearly visible in (15).
6. The minimal value of $Q_d(\mathbb{D}_{n,\delta^*})$ with respect to d is attained at $d = 15$.
7. Formulae (23) and (24) are in agreement with numerical results presented in Table 4 of [14] and Table 5 of [4].

Let us now briefly illustrate the results above. In Fig. 1, the black circles depict the quantity $Q_d(\mathbb{D}_{n,\delta^*})$ as a function of d. The quantity $Q_d(\mathbb{D}_n^{(0)}) = 1/12$ is shown with the solid red line. We conclude that from dimension seven onwards, the design $\mathbb{D}_{n,\delta^*}$ provides better quantisation per points than the design $\mathbb{D}_n^{(0)}$. Moreover, for $d > 15$, the quantity $Q_d(\mathbb{D}_{n,\delta^*})$ slowly increases and converges to 1/12. Typical behaviour of $Q_d(\mathbb{D}_{n,\delta})$ as a function of δ is shown in Fig. 2. This figure demonstrates the significance of choosing δ optimally.

3 Closed-Form Expressions for the Coverage Area with $\mathbb{D}_{n,\delta}$ and Approximations

In this section, we will derive explicit expressions for the coverage area of the cube $[-1, 1]^d$ by the union of the balls $\mathcal{B}_d(\mathbb{D}_{n,\delta}, r)$ associated with the design $\mathbb{D}_{n,\delta}$ introduced in Sect. 1.2. That is, we will derive expressions for the quantity $C_d(\mathbb{D}_{n,\delta}, r)$ for all values of r. Then, in Sect. 3.3, we shall obtain approximations for $C_d(\mathbb{D}_{n,\delta}, r)$. The accuracy of the approximations will be assessed in Sect. 4.2.

3.1 Reduction to Voronoi Cells

For an n-point design $\mathbb{Z}_n = \{Z_1, \ldots, Z_n\}$, denote the proportion of the Voronoi cell around Z_i covered by the ball $\mathcal{B}_d(Z_i, r)$ as

$$V_{d,Z_i,r} := \frac{\mathrm{vol}\,(V(Z_i) \cap \mathcal{B}_d(Z_i, r))}{\mathrm{vol}\,(V(Z_i))}.$$

The following lemma is straightforward.

Lemma 3.1 *Consider a design $\mathbb{Z}_n = \{Z_1, \ldots, Z_n\}$ such that all Voronoi cells $V(Z_i)$ are congruent. Then for any $Z_i \in \mathbb{Z}_n$, $C_d(\mathbb{Z}_n, r) = V_{d,Z_i,r}$.*

In view of Theorem 2.2, for design $\mathbb{D}_{n,\delta}$ all Voronoi cells $V(Z_i)$ are congruent and $\mathrm{vol}\,(V(Z_i)) = 2$; recall that $n = 2^{d-1}$. By then applying Lemma 3.1 and without loss of generality we have chosen $Z_1 = \boldsymbol{\delta} = (\delta, \delta, \ldots, \delta) \in \mathbb{R}^d$, we have for any $r > 0$

$$V_{d,\boldsymbol{\delta},r} = \frac{\mathrm{vol}\,(V(\boldsymbol{\delta}) \cap \mathcal{B}_d(\boldsymbol{\delta}, r))}{2} = C_d(\mathbb{D}_{n,\delta}, r). \tag{26}$$

In order to formulate explicit expressions for $V_{d,\boldsymbol{\delta},r}$, we need the important quantity, proportion of intersection of $[-1, 1]^d$ with one ball. Take the cube $[-1, 1]^d$ and a ball $\mathcal{B}_d(Z, r) = \{Y \in \mathbb{R}^d : \|Y - Z\| \le r\}$ centred at a point $Z = (z_1, \ldots, z_d) \in \mathbb{R}^d$; this point Z could be outside $[-1, 1]^d$. The fraction of the cube $[-1, 1]^d$ covered by the ball $\mathcal{B}_d(Z, r)$ is denoted by

$$C_{d,Z,r} = \frac{\mathrm{vol}\,([-1, 1]^d \cap \mathcal{B}_d(Z, r))}{2^d}.$$

3.2 Expressing $C_d(\mathbb{D}_{n,\delta}, r)$ Through $C_{d,Z,r}$

Theorem 3.2 *Depending on the values of r and δ, the quantity $C_d(\mathbb{D}_{n,\delta}, r)$ can be expressed through $C_{d,Z,r}$ for suitable Z as follows.*

- *For $r \le \delta$:*

$$C_d(\mathbb{D}_{n,\delta}, r) = \frac{C_{d,\mathbf{2\delta-1},2r}}{2}. \tag{27}$$

- *For $\delta \le r \le 1 + \delta$:*

$$C_d(\mathbb{D}_{n,\delta}, r) = \frac{1}{2}\left[C_{d,\mathbf{2\delta-1},2r} + d\int_0^{r-\delta} C_{d-1,\frac{\mathbf{2\delta-1-t}}{\mathbf{1-t}},\frac{2\sqrt{r^2-(t+\delta)^2}}{1-t}}(1-t)^{d-1}\,dt\right]. \tag{28}$$

– *For* $r \geq 1+\delta$*:*

$$C_d(\mathbb{D}_{n,\delta}, r) = \frac{1}{2}\left[C_{d,\mathbf{2\delta-1},2r} + d\int_0^1 C_{d-1,\frac{\mathbf{2\delta-1-t}}{\mathbf{1-t}},\frac{2\sqrt{r^2-(t+\delta)^2}}{1-t}}(1-t)^{d-1}\,dt\right]. \tag{29}$$

The proof of Theorem 3.2 is given in Appendix A.

3.3 Approximation for $C_d(\mathbb{D}_{n,\delta}, r)$

Accurate approximations for $C_{d,Z,r}$ for arbitrary d, Z, and r were developed in [14]. By using the general expansion in the central limit theorem for sums of independent non-identical r.v., the following approximation was developed:

$$C_{d,Z,r} \cong \Phi(t) + \frac{\|Z\|^2 + d/63}{5\sqrt{3}(\|Z\|^2 + d/15)^{3/2}}(1-t^2)\varphi(t), \tag{30}$$

where

$$t = \frac{\sqrt{3}(r^2 - \|Z\|^2 - d/3)}{2\sqrt{\|Z\|^2 + d/15}}.$$

A short derivation of this approximation is included in Appendix D. Using (30), we formulate the following approximation for $C_d(\mathbb{D}_{n,\delta}, r)$.

Approximations for $C_d(\mathbb{D}_{n,\delta}, r)$. *Approximate the values* $C_{\cdot,\cdot,\cdot}$ *in formulae* (27), (28), *and* (29) *with corresponding approximations* (30).

3.4 Simple Bounds for $C_d(\mathbb{D}_{n,\delta}, r)$

Lemma 3.3 *For any* $r \geq 0$, $0 < \delta < 1$, *and* $\boldsymbol{\delta} = (\delta, \delta, \dots, \delta) \in \mathbb{R}^d$, *the quantity* $C_d(\mathbb{D}_{n,\delta}, r)$ *can be bounded as follows:*

$$\frac{C_{d,\mathbf{2\delta-1},2r} + C_{d,A,2r}}{2} \leq C_d(\mathbb{D}_{n,\delta}, r) \leq C_{d,\mathbf{2\delta-1},2r}, \tag{31}$$

where $A = (2\delta+1, 2\delta-1, \dots, 2\delta-1) \in \mathbb{R}^d$.

The proof of Lemma 3.3 is given in Appendix B. In Figs. 3 and 4, using the approximation given in (30) we study the tightness of the bounds given in (31). In these figures, the dashed red line, dashed blue line and solid black line depict the upper bound, the lower bound and the approximation for $C_d(\mathbb{D}_{n,\delta}, r)$ respectively. We see that the upper bound is very sharp across r and d; this behaviour is not seen with the lower bound.

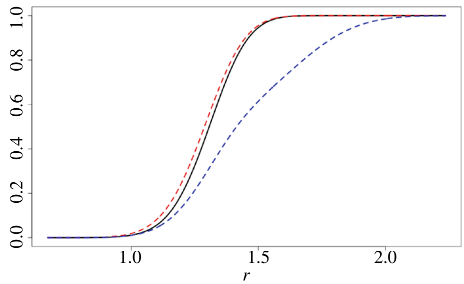

Fig. 3 $C_d(\mathbb{D}_{n,\delta}, r)$ with upper and lower bounds: $d = 20$

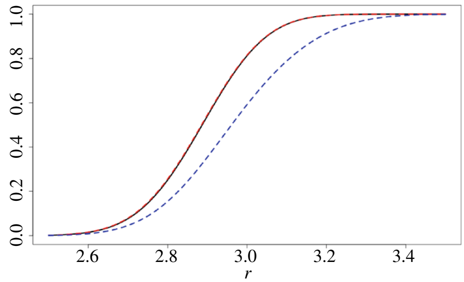

Fig. 4 $C_d(\mathbb{D}_{n,\delta}, r)$ with upper and lower bounds: $d = 100$

3.5 'Do Not Try to Cover the Vertices'

In this section, we theoretically support the recommendation 'do not try to cover the vertices' which was first stated in [14] and supported in [4] on the basis of numerical evidence. In other words, we will show on the example of the design $\mathbb{D}_{n,1/2}$ that in large dimensions the attempt to cover the whole cube rather than 0.999 of it leads to a dramatic increase of the required radius of the balls.

Theorem 3.4 *Let γ be fixed, $0 \le \gamma \le 1$. Consider $(1-\gamma)$-coverings of $[-1, 1]^d$ generated by the designs $\mathbb{D}_{n,\delta}$ and the associated normalised radii $R_{1-\gamma}(\mathbb{D}_{n,\delta})$, see (8). For any $0 < \gamma < 1$ and $0 \le \delta \le 1$, the limit of $R_{1-\gamma}(\mathbb{D}_{n,\delta})$, as $d \to \infty$, exists and achieves minimal value for $\delta = 1/2$. Moreover, $R_{1-\gamma}(\mathbb{D}_{n,1/2})/R_1(\mathbb{D}_{n,1/2}) \to 1/\sqrt{3}$ as $d \to \infty$, for any $0 < \gamma < 1$.*

Proof is given in Appendix C.

In Figs. 5 and 6 using a solid red line we depict the approximation of $C_d(\mathbb{D}_{n,1/2}, r)$ as a function of $R = n^{1/d} r/(2\sqrt{d})$, see (5). The vertical green line illustrates the value of $R_{0.999}$ and the vertical blue line depicts $R_1 = n^{1/d}\sqrt{d+8}/(4\sqrt{d})$. These figures illustrate that as d increases, for all γ we have $R_{1-\gamma}/R_1$ slowly tending to $1/\sqrt{3}$. From the proof of Theorem 3.4, it transpires that $C_d(\mathbb{D}_{n,\delta}, r)$ as a function of R converges to the jump function with the jump at $1/(2\sqrt{3})$.

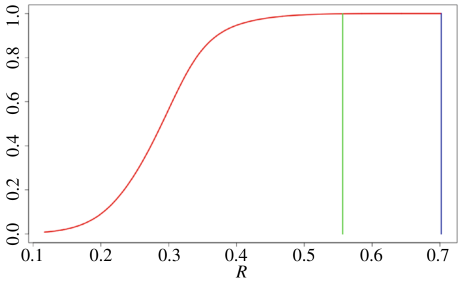

Fig. 5 $C_d(\mathbb{D}_{n,1/2}, r)$ with $R_{0.999}$ and R_1: $d = 5$

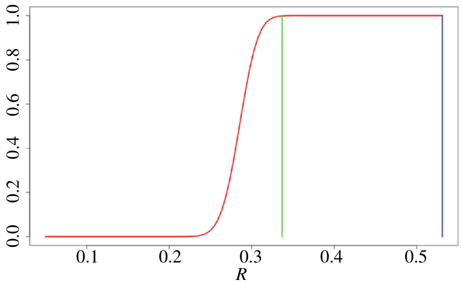

Fig. 6 $C_d(\mathbb{D}_{n,1/2}, r)$ with $R_{0.999}$ and R_1: $d = 50$

4 Numerical Studies

For comparative purposes, we introduce another design which is one of the most popular designs (both, for quantisation and covering) considered in applications.

Design $\mathbb{S}_n$: $Z_1, \dots, Z_n$ *are taken from a low-discrepancy Sobol's sequence on the cube* $[-1, 1]^d$.

For constructing the design $\mathbb{S}_n$, we use the R-implementation provided in the well-known 'SobolSequence' package [3]. For $\mathbb{S}_n$, we have set $n = 1024$ and $F2 = 10$ (an input parameter for the Sobol sequence function). Sobol sequences $\mathbb{S}_n$ attain their best space-filling properties when n is a power of 2; that is, when $n = 2^\ell$ for some integer ℓ. We have chosen $\ell = 10$. As we study renormalised characteristics $Q_d(\cdot)$ and $R_{1-\gamma}(\cdot)$ of designs, exact value of ℓ for $\mathbb{S}_n$ with $n = 2^\ell$ is almost irrelevant: in particular, numerically computed values $Q_d(\mathbb{S}_{2^\ell})$ and $R_{1-\gamma}(\mathbb{S}_{2^\ell})$ for $\ell = 8, 9, 11, 12$ are almost indistinguishable from the corresponding values for $\ell = 10$ provided below in Tables 1 and 2. By varying values of ℓ, we are not improving space-filling properties of $\mathbb{S}_{2^\ell}$. In fact, increase of ℓ generally leads to a slight deterioration of normalised space-filling characteristics (including $Q_d(\cdot)$ and $R_{1-\gamma}(\cdot)$) of Sobol sequences.

4.1 Quantisation and Weak Covering Comparisons

In Table 1, we compare the normalised mean squared quantisation error $Q_d(\mathbb{Z}_n)$ defined in (13) across three designs: $\mathbb{D}_{n,\delta^*}$ with δ^* given in (23), $\mathbb{D}_n^{(0)}$ and $\mathbb{S}_n$. In Table 2, we compare the normalised statistic $R_{1-\gamma}$ introduced in (7), where we have fixed $\gamma = 0.01$. For designs $\mathbb{D}_{n,\delta}$ (with the optimal value of δ), $\mathbb{D}_{n,1/2}$ and $\mathbb{D}_n^{(0)}$ we have also included R_1, the smallest normalised radius that ensures the full coverage. Let us make some remarks concerning Tables 1 and 2.

- In conjunction with Fig. 1, Table 1 shows that for $d \geq 7$, the quantisation for design $\mathbb{D}_{n,\delta^*}$ is superior over all other designs considered.
- For the weak coverage statistic $R_{1-\gamma}$, the superiority of $\mathbb{D}_{n,\delta}$ with optimal δ over all other designs considered is seen for $d \geq 10$.
- For the designs $\mathbb{D}_{n,\delta}$, the optimal value of δ minimising $R_{1-\gamma}$ depends on γ.
- From remark 6 of Sect. 2.4, the minimal value of $Q_d(\mathbb{D}_{n,\delta^*})$ with respect to d is attained at $d = 15$. For $d > 15$, the quantity $Q_d(\mathbb{D}_{n,\delta^*})$ increases with d, slowly converging to $Q_d(\mathbb{D}_n^{(0)}) = 1/12$. This non-monotonic behaviour can be seen in Table 1.
- Unlike the case of $Q_d(\mathbb{D}_{n,\delta^*})$, such non-monotonic behaviour is not seen for the quantity $R_{1-\gamma}$ and $R_{1-\gamma}(\mathbb{D}_{n,\delta})$ monotonically decreases as d increases. Also, Theorem 3.4 implies that for any $\gamma \in (0,1)$, $R_{1-\gamma}(\mathbb{D}_{n,\delta}) \to 1/(2\sqrt{3}) \cong 0.289$ as $d \to \infty$.

Table 1 Normalised mean squared quantisation error Q_d for three designs and different d

	$d=5$	$d=7$	$d=10$	$d=15$	$d=20$
$Q_d(\mathbb{D}_{n,\delta^*})$	0.0876	0.0827	0.0804	0.0798	0.0800
$Q_d(\mathbb{D}_n^{(0)})$	0.0833	0.0833	0.0833	0.0833	0.0833
$Q_d(\mathbb{S}_n)$	0.0988	0.1003	0.1022	0.1060	0.1086

Table 2 Normalised statistic $R_{1-\gamma}$ across d with $\gamma = 0.01$ (value in brackets corresponds to optimal δ)

	$d=5$	$d=7$	$d=10$	$d=15$	$d=20$
$R_{1-\gamma}(\mathbb{D}_{n,\delta})$	0.4750 (0.54)	0.3992 (0.53)	0.3635 (0.52)	0.3483 (0.51)	0.3417 (0.50)
$R_{1-\gamma}(\mathbb{D}_{n,1/2})$	0.4765	0.4039	0.3649	0.3484	0.3417
$R_{1-\gamma}(\mathbb{D}_n^{(0)})$	0.4092	0.3923	0.3766	0.3612	0.3522
$R_{1-\gamma}(\mathbb{S}_n)$	0.4714	0.4528	0.4256	0.4074	0.3967
$R_1(\mathbb{D}_{n,\delta})$	0.6984 (0.54)	0.6555 (0.53)	0.6178 (0.52)	0.5856 (0.51)	0.5714 (0.50)
$R_1(\mathbb{D}_{n,1/2})$	0.7019	0.6629	0.6259	0.5912	0.5714
$R_1(\mathbb{D}_n^{(0)})$	0.5000	0.5000	0.5000	0.5000	0.5000

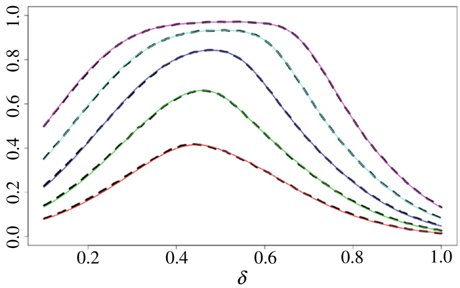

Fig. 7 $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 5$, r from 0.7 to 1.1 increasing by 0.1

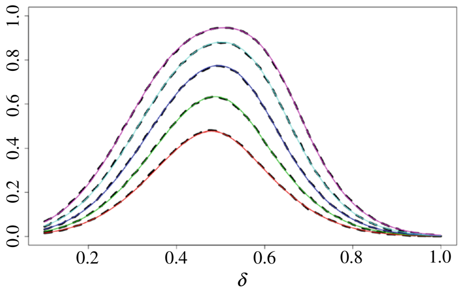

Fig. 8 $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 10$, r from 0.95 to 1.15 increasing by 0.05

4.2 Accuracy of Covering Approximation and Dependence on δ

In this section, we assess the accuracy of the approximation of $C_d(\mathbb{D}_{n,\delta}, r)$ developed in Sect. 3.3 and the behaviour of $C_d(\mathbb{D}_{n,\delta}, r)$ as a function of δ. In Figs. 7, 8, 9, and 10, the thick dashed black lines depict $C_d(\mathbb{D}_{n,\delta}, r)$ for several different choices of r; these values are obtained via Monte Carlo simulations. The thinner solid lines depict its approximation of Sect. 3.3. These figures show that the approximation is extremely accurate for all r, δ, and d; we emphasise that the approximation remains accurate even for very small dimensions like $d = 3$. These figures also clearly demonstrate the δ-effect saying that a significantly more efficient weak coverage can be achieved with a suitable choice of δ. This is particularly evident in higher dimensions, see Figs. 9 and 10.

Figs. 11 and 12 illustrate Theorem 3.4 and show the rate of convergence of the covering radii as d increases. Let the probability density function $f(r)$ be defined by $dC_d(\mathbb{D}_{n,\delta}, r) = f(r)\, dr$, where $C_d(\mathbb{D}_{n,\delta}, r)$ as a function of r is viewed as the c.d.f. of the r.v. $r = \varrho(X, \mathbb{Z}_n)$, see Sect. 1.3. Trivial calculations yield that the density for the normalised radius R expressed by (5) is $p_d(R) := 2\sqrt{d}\, n^{-1/d} f(2\sqrt{d}\, n^{-1/d} R)$. In Fig. 11, we depict the density $p_d(\cdot)$ for $d = 5, 10, 20$ with blue, red, and black lines respectively. The respective c.d.f.'s $\int_0^R p_d(\tau)\, d\tau$ are shown in Fig. 12 under the same colouring scheme.

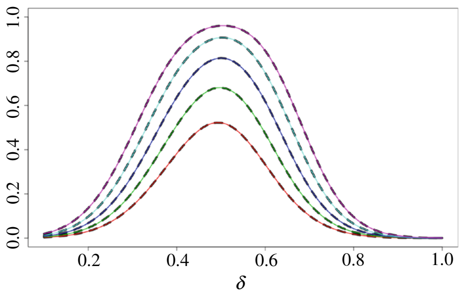

Fig. 9 $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 15$, r from 1.15 to 1.35 increasing by 0.05

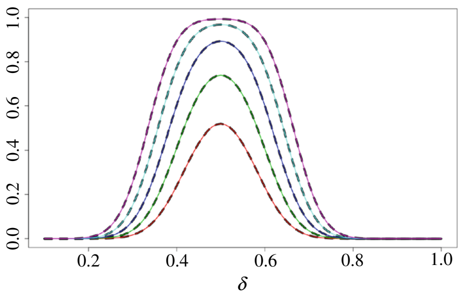

Fig. 10 $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 50$, r from 2.05 to 2.35 increasing by 0.075

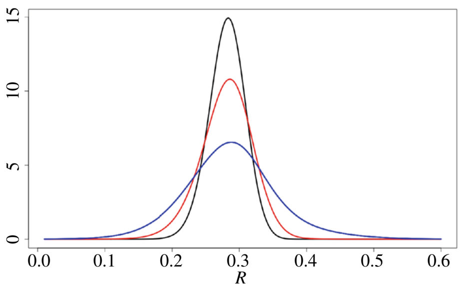

Fig. 11 Densities $f_d(R)$ for the design $\mathbb{D}_{n,\delta^*}$; $d = 5, 10, 20$

4.3 Stochastic Dominance

In Figs. 13 and 14, we depict the c.d.f.'s for the normalised distance $n^{1/d}\varrho(X, \mathbb{Z}_n)/(2\sqrt{d})$ for two designs: $\mathbb{D}_{n,\delta^*}$ in red, and $\mathbb{D}_n^{(0)}$ in black. We can see that the design $\mathbb{D}_{n,\delta^*}$ stochastically dominates the design $\mathbb{D}_n^{(0)}$ for $d = 10$ but for $d = 5$ the design $\mathbb{D}_n^{(0)}$ is preferable to the design $\mathbb{D}_{n,\delta^*}$ although there is no clear domination; this is in line with findings from Sects. 2.4 and 4.1, see e.g. Fig. 1, Tables 1 and 2.

In Fig. 15, we depict the c.d.f.'s for the normalised distance $n^{1/d}\varrho(X, \mathbb{Z}_n)/(2\sqrt{d})$ foresign $\mathbb{D}_n^{(0)}$ (in red) and design $\mathbb{S}_n$ (in black). We can see that for $d = 5$, the design $\mathbb{D}_n^{(0)}$ stochastically dominates the design $\mathbb{S}_n$. The style of Fig. 16 is the same as Fig. 15,

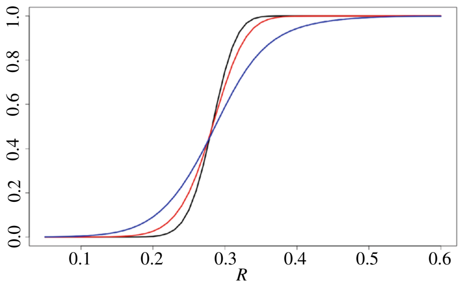

Fig. 12 c.d.f.'s of R for the design $\mathbb{D}_{n,\delta*}$; $d = 5, 10, 20$

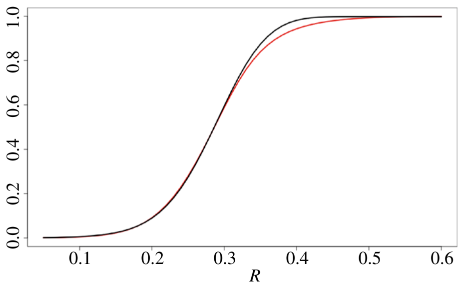

Fig. 13 $d = 5$: design $\mathbb{D}_n^{(0)}$ is preferable to design $\mathbb{D}_{n,\delta*}$

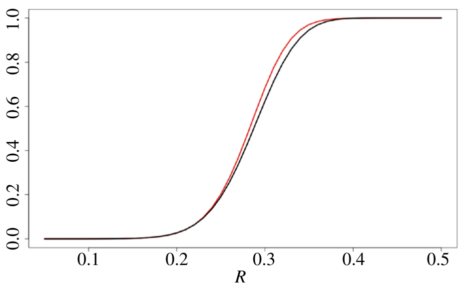

Fig. 14 $d = 10$: design $\mathbb{D}_{n,\delta*}$ stochastically dominates design $\mathbb{D}_n^{(0)}$

however we set $d = 10$ and the design $\mathbb{D}_n^{(0)}$ is replaced with the design $\mathbb{D}_{n,\delta*}$. Here we see a very clear stochastic dominance of the design $\mathbb{D}_{n,\delta*}$ over the design $\mathbb{S}_n$. All findings are consistent with findings from Sect. 4.1, see Tables 1 and 2.

Acknowledgements The authors are grateful to our colleague Iskander Aliev for numerous intelligent discussions. The authors are very grateful to the referee for their comments that greatly improved the presentation of the paper. Furthermore, we are grateful for the concise proof of Theorem 3.2 suggested by the reviewer which has been included in this manuscript.

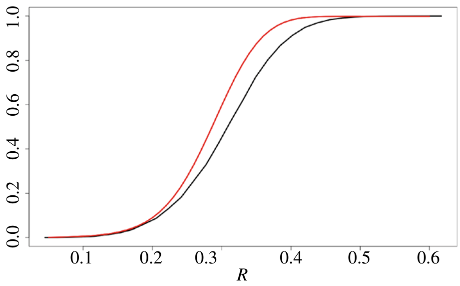

Fig. 15 $d = 5$: design $\mathbb{D}_n^{(0)}$ stochastically dominates design $\mathbb{S}_n$

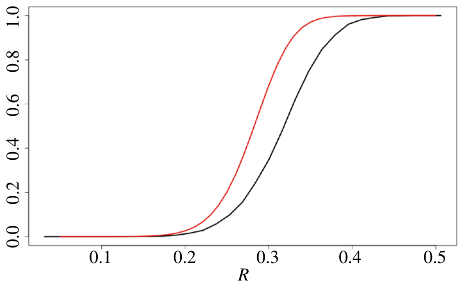

Fig. 16 $d = 10$: design $\mathbb{D}_{n,\delta^*}$ stochastically dominates design $\mathbb{S}_n$

Appendix A: Proof of Theorem 3.2

In view of (26), $C_d(\mathbb{D}_{n,\delta}, r) = V_{d,\boldsymbol{\delta},r}$ for all $0 \le \delta \le 1$ and $r \ge 0$ and we shall derive expressions for $V_{d,\boldsymbol{\delta},r}$ rather than $C_d(\mathbb{D}_{n,\delta}, r)$.

Case (a): $r \le \delta$. To prove this case, we observe i) for this range of r, $\mathcal{B}_d(\boldsymbol{\delta}, r) \subset [0,1]^d$; ii) the fraction of a cube covered by a ball is preserved under invertible affine transformations; iii) the affine transformation $x \mapsto 2x - \mathbf{1}$ maps the ball $\mathcal{B}_d(\boldsymbol{\delta}, r)$ and cube $[0,1]^d$ to $\mathcal{B}_d(\mathbf{2\delta - 1}, 2r)$ and $[-1,1]^d$, respectively. This leads to

$$V_{d,\boldsymbol{\delta},r} = \frac{\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r))}{2\,\mathrm{vol}\,([0,1]^d)} = \frac{\mathrm{vol}\,(\mathcal{B}_d(\mathbf{2\delta - 1}, 2r))}{2\,\mathrm{vol}\,([-1,1]^d)} = \frac{C_{d,\mathbf{2\delta-1},2r}}{2}.$$

Case (b): $\delta \le r \le 1 + \delta$. Using (15) we obtain

$$V_{d,\boldsymbol{\delta},r} = \frac{\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + d\,\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap U_1)}{2}.$$

The first quantity in the numerator has been considered in case (a) and it is simply $C_{d,\mathbf{2\delta-1},2r}$. Therefore we aim to reformulate the second quantity within the brackets, $\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta},r)\cap U_1)$. Denote by $\mathcal{P}(t)=\{(x_1,x_2,\dots,x_d):x_1=t\}$, the $(d-1)$-dimensional hyperplane. Then

$$\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta},r)\cap U_1)=\int_{\delta-r}^{0}\mathrm{vol}_{d-1}(\mathcal{P}(t)\cap\mathcal{B}_d(\boldsymbol{\delta},r)\cap U_1)\,dt.$$

Notice further that

$$\begin{aligned} U_1\cap\mathcal{P}(t)&=\{t\}\times[|t|,1]^{d-1}, && \text{for } -1\le t\le 0,\\ \mathcal{B}_d(\boldsymbol{\delta},r)\cap\mathcal{P}(t)&=\{t\}\times\mathcal{B}_{d-1}\Big(\boldsymbol{\delta},\sqrt{r^2-(t-\delta)^2}\Big) && \text{for } \delta-r\le t\le 0,\ r\ge\delta, \end{aligned}\tag{32}$$

where $\boldsymbol{\delta}=(\delta,\dots,\delta)\in\mathbb{R}^{d-1}$ and the natural identification of $\mathcal{P}(t)$ with $\mathbb{R}^{d-1}$ is used. The r.h.s. in (32) are a $(d-1)$-dimensional cube and ball respectively. Since covered fraction is preserved under affine transformations in $\mathbb{R}^{d-1}$, it suffices to construct one, denote by ϕ, for which $\phi([|t|,1]^{d-1})=[-1,1]^d$. In $\mathcal{P}(t)$, such ϕ maps the cube from (32) to the standard cube $[-1,1]^d$. Clearly, ϕ can be taken as

$$\phi\colon x\mapsto\frac{x-(\mathbf{1}+|\boldsymbol{t}|)/2}{(1-|t|)/2}=\frac{2x-(\mathbf{1}+|\boldsymbol{t}|)}{1-|t|},$$

where $\mathbf{1}=(1,\dots,1)$ and $|\boldsymbol{t}|=(|t|,\dots,|t|)$ are constant vectors in $\mathbb{R}^{d-1}$. Note that

$$\phi\Big(\mathcal{B}_{d-1}\Big(\boldsymbol{\delta},\sqrt{r^2-(t-\delta)^2}\Big)\Big)=\mathcal{B}_{d-1}\bigg(\frac{\mathbf{2\delta-(1+|t|)}}{\mathbf{1-|t|}},\frac{2\sqrt{r^2-(t-\delta)^2}}{1-|t|}\bigg).$$

Finally, by the preservation of covered fraction, we obtain

$$\mathrm{vol}_{d-1}(\mathcal{P}(t)\cap\mathcal{B}_d(\boldsymbol{\delta},r)\cap U_1)=\mathrm{vol}_{d-1}([|t|,1]^{d-1})\cdot C_{d-1,\frac{\mathbf{2\delta-1-|t|}}{\mathbf{1-|t|}},\frac{2\sqrt{r^2-(t-\delta)^2}}{1-|t|}}.$$

As a result,

$$\begin{aligned} V_{d,\boldsymbol{\delta},r}&=\frac12\bigg[C_{d,\mathbf{2\delta-1},2r}+d\int_{\delta-r}^{0}C_{d-1,\frac{\mathbf{2\delta-1-|t|}}{\mathbf{1-|t|}},\frac{2\sqrt{r^2-(t-\delta)^2}}{1-|t|}}(1-|t|)^{d-1}\,dt\bigg]\\ &=\frac12\bigg[C_{d,\mathbf{2\delta-1},2r}+d\int_{0}^{r-\delta}C_{d-1,\frac{\mathbf{2\delta-1-t}}{\mathbf{1-t}},\frac{2\sqrt{r^2-(t+\delta)^2}}{1-t}}(1-t)^{d-1}\,dt\bigg]. \end{aligned}\tag{33}$$

Case (c): $r\ge 1+\delta$. Case (c) is almost identical to case (b), with the only change occurring within the lower limit of integration in (33); the lower limit of the integral remains at -1 for all $r\ge 1+\delta$. Since the steps are almost identical to case (b), they are omitted and we simply conclude:

$$V_{d,\boldsymbol{\delta},r}=\frac12\bigg[C_{d,\mathbf{2\delta-1},2r}+d\int_{0}^{1}C_{d-1,\frac{\mathbf{2\delta-1-t}}{\mathbf{1-t}},\frac{2\sqrt{r^2-(t+\delta)^2}}{1-t}}(1-t)^{d-1}\,dt\bigg].$$

Appendix B: Proof of Lemma 3.3

(a) Let us first prove the upper bound in (31). Consider the set U_j defined in (16) and the associated set

$$U'_j = \left\{X = (x_1, x_2, \dots, x_d) \in [0,1]^d : |x_j| \le x_k \le 1 \text{ for all } k \ne j\right\} \subset C_0.$$

We have $\mathrm{vol}(U_j) = \mathrm{vol}(U'_j) = 1/d$ and

$$V(\boldsymbol{\delta}) = C_0 \cup \bigcup_{j=1}^{d} U_j, \qquad \bigcup_{j=1}^{d} U'_j = C_0 = [0,1]^d.$$

Let us prove that for any $r \ge 0$ we have

$$\mathrm{vol}\,(U_j \cap \mathcal{B}_d(\boldsymbol{\delta}, r)) \le \mathrm{vol}\,(U'_j \cap \mathcal{B}_d(\boldsymbol{\delta}, r)). \tag{34}$$

With any $X = (x_1, x_2, \dots, x_d) \in U'_1$, we associate the point $X' = (-x_1, x_2, \dots, x_d) \in U_1$ by simply changing the sign in the first component. For these two points, we have

$$\|\boldsymbol{\delta} - X\|^2 = (x_1 - \delta)^2 + \sum_{k=2}^{d} (x_k - \delta)^2 < (-x_1 - \delta)^2 + \sum_{k=2}^{d} (x_k - \delta)^2 = \|\boldsymbol{\delta} - X'\|^2.$$

Therefore, $\|\boldsymbol{\delta} - X\|^2 \le r$ implies $\|\boldsymbol{\delta} - X'\|^2 \le r$ yielding (34).

To prove the upper bound in (31) for all r we must consider two cases: $r \le \delta$ and $r \ge \delta$. For $r \le \delta$, we clearly have

$$V_{d,\boldsymbol{\delta},r} = \frac{C_{d,\mathbf{2\delta-1},2r}}{2} \le C_{d,\mathbf{2\delta-1},2r}.$$

For $r \ge \delta$, using (34) we have

$$\begin{aligned} V_{d,\boldsymbol{\delta},r} &= \frac{\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + d\,\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap U_1)}{2} \\ &\le \frac{\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + d\,\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap U'_1)}{2} \\ &= \mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) = C_{d,\mathbf{2\delta-1},2r} \end{aligned}$$

and hence the upper bound in (31).

(b) Consider now the lower bound in (31). For $j \ge 2$, with the set U_j we now associate the set

$$V_j = \Big\{\tilde{X} = (x_1, \dots, x_d) : -1 \le x_1 \le 0,\ 0 \le x_m \le 1 \quad (\text{for } m > 1),\ |x_j| \le |x_k| \le 1 \ \text{ for } \ k \ne j\Big\}.$$

With any point $X = (x_1, x_2, \ldots, x_d) \in U_j$ (here x_j is negative and $|x_j| \le |x_k| \le 1$ for $k \ne j$) we associate point $\tilde{X} = (-x_1, x_2, \ldots, x_{j-1}, -x_j, x_{j+1}, \ldots, x_d) \in V_j$ by changing sign in the first and j-the component of $X \in U_j$. Setting without loss of generality $j = 2$, we have for these two points:

$$\|\boldsymbol{\delta} - X\|^2 = (x_1 - \delta)^2 + (x_2 - \delta)^2 + \sum_{k=3}^{d} (x_k - \delta)^2$$
$$\le (-x_1 - \delta)^2 + (-x_2 - \delta)^2 + \sum_{k=3}^{d} (x_k - \delta)^2 = \|\boldsymbol{\delta} - \tilde{X}\|^2,$$

where the inequality follows from the inequalities $x_1 \ge 0$, $x_2 < 0$, and $|x_2| < x_1$ containing in the definition of U_2. Therefore, $\|\boldsymbol{\delta} - \tilde{X}\|^2 \le r$ implies $\|\boldsymbol{\delta} - X\|^2 \le r$, where from

$$\mathrm{vol}\,(U_j \cap \mathcal{B}_d(\boldsymbol{\delta}, r)) \ge \mathrm{vol}\,(V_j \cap \mathcal{B}_d(\boldsymbol{\delta}, r)). \tag{35}$$

To prove the lower bound in (31) for all r we must consider two cases: $r \le \delta$ and $r \ge \delta$. For $r \ge \delta$, using (35) we have

$$\begin{aligned} V_{d,\boldsymbol{\delta},r} &= \frac{1}{2}\left[\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + \sum_{i=1}^{d} \mathrm{vol}\,((\mathcal{B}_d(\boldsymbol{\delta}, r) \cap U_i))\right] \\ &\ge \frac{1}{2}\left[\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + \mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap U_1) + \sum_{i=2}^{d} \mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap V_i)\right] \\ &= \frac{\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + \mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1)}{2}, \end{aligned}$$

where C_1 is given in (17). To compute $\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1)$, we shall use a similar technique to the proof of Theorem 3.2. The affine transformation

$$x \mapsto 2x + (1, -1, -1, \ldots, -1)$$

maps the ball $\mathcal{B}_d(\boldsymbol{\delta}, r)$ and the cube C_1 to $\mathcal{B}_d(A, 2r)$ and $[-1, 1]^d$ respectively, where $A = (2\delta + 1, 2\delta - 1, \ldots, 2\delta - 1)$. Since the fraction of covered volume is preserved under invertible affine transformations, one has

$$\frac{\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1)}{\mathrm{vol}\,(C_1)} = C_{d,A,2r}$$

and hence we can conclude:

$$V_{d,\boldsymbol{\delta},r} \ge \frac{\mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + \mathrm{vol}\,(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1)}{2} = \frac{C_{d,\mathbf{2\delta-1},2r} + C_{d,A,2r}}{2}.$$

For $r \le \delta$, since vol $(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1) = C_{d,A,2r} = 0$, we have

$$V_{d,\boldsymbol{\delta},r} = \frac{C_{d,\mathbf{2\delta-1},2r} + C_{d,A,2r}}{2}$$

and hence the lower bound in (31).

Appendix C: Proof of Theorem 3.4

Before proving Theorem 3.4, we prove three auxiliary lemmas.

Lemma C.1 *Let $r = r_{\alpha,d} = \alpha\sqrt{d}$ with $\alpha \ge 0$ and $Z_{a,b;d} = (a, b, b, \dots, b) \in \mathbb{R}^d$. Then the limit $\lim_{d\to\infty} C_{d,Z_{a,b;d},2r}$ exists and*

$$\lim_{d\to\infty} C_{d,Z_{a,b;d},2r} = \begin{cases} 0 & \text{if } \alpha < \sqrt{1/3 + b^2}/2, \\ 1/2 & \text{if } \alpha = \sqrt{1/3 + b^2}/2, \\ 1 & \text{if } \alpha > \sqrt{1/3 + b^2}. \end{cases}$$

Proof Define

$$t_\alpha = \frac{\sqrt{3}(d(4\alpha^2 - b^2 - 1/3) + b^2 - a^2)}{2\sqrt{a^2 + (d-1)b^2 + d/15}}.$$

As the r.v. η_z introduced in Appendix D are concentrated on a finite interval, for finite a and b the quantities of $\rho_a := \mathbb{E}(|\eta_a - a^2 - 1/3|^3)$ and $\rho_b := \mathbb{E}(|\eta_b - b^2 - 1/3|^3)$ are bounded. By applying the Berry–Esseen theorem (see [5, Sect. 2, Chap. 5]) to $C_{d,Z_{a,b},2r}$, there exists some constant C such that

$$-\frac{C \max\{\rho_a/\sigma_a^2, \rho_b/\sigma_b^2\}}{(\sigma_a^2 + (d-1)\sigma_b^2)^{1/2}} + \Phi(t_\alpha) \le C_{d,Z_{a,b},2r} \le \Phi(t_\alpha) + \frac{C \max\{\rho_a/\sigma_a^2, \rho_b/\sigma_b^2\}}{(\sigma_a^2 + (d-1)\sigma_b^2)^{1/2}},$$

where $\sigma_a^2 = \mathrm{var}(\eta_a)$ and $\sigma_b^2 = \mathrm{var}(\eta_b)$. By the squeeze theorem, it is clear that if $4\alpha^2 - b^2 - 1/3 > 0$ and hence $\alpha > \sqrt{1/3 + b^2}/2$, then $C_{d,Z_{a,b},2r} \to 1$ as $d \to \infty$. If $\alpha < \sqrt{1/3 + b^2}/2$, then $C_{d,Z_{a,b},2r} \to 0$ as $d \to \infty$. If $\alpha = \sqrt{1/3 + b^2}/2$, then $C_{d,Z_{a,b},2r} \to 1/2$ as $d \to \infty$. □

Lemma C.2 *Let $r = \alpha\sqrt{d}$. Then for $\boldsymbol{\delta} = (\delta, \delta, \dots, \delta)$, we have*

$$\lim_{d\to\infty} V_{d,\boldsymbol{\delta},r} = \lim_{d\to\infty} C_{d,\mathbf{2\delta-1},2r} = \begin{cases} 0 & \text{if } \alpha < \sqrt{1/3 + (2\delta-1)^2}/2, \\ 1/2 & \text{if } \alpha = \sqrt{1/3 + (2\delta-1)^2}/2, \\ 1 & \text{if } \alpha > \sqrt{1/3 + (2\delta-1)^2}/2. \end{cases}$$

Proof Using Lemma C.1 with $Z_{a,b} = A = (2\delta + 1, 2\delta - 1, \dots, 2\delta - 1)$, we obtain

$$\lim_{d\to\infty} C_{d,A,2r} = \lim_{d\to\infty} C_{d,\mathbf{2\delta-1},2r} = \begin{cases} 0 & \text{if } \alpha < \sqrt{1/3 + (2\delta-1)^2}/2, \\ 1/2 & \text{if } \alpha = \sqrt{1/3 + (2\delta-1)^2}/2, \\ 1 & \text{if } \alpha > \sqrt{1/3 + (2\delta-1)^2}/2. \end{cases}$$

By then applying the squeeze theorem to the bounds in Lemma 3.3 using the fact from Lemma 3.1 we have $V_{d,\boldsymbol{\delta},r} = C_d(\mathbb{Z}_n, r)$, we obtain the result. □

To determine the value of r that leads to the full coverage, we utilise the following simple lemma.

Lemma C.3 *For design $\mathbb{D}_{n,\delta}$, the smallest value of r that ensures a complete coverage of $[-1, 1]^d$ satisfies*

$$\lim_{d\to\infty} \frac{r_1}{\sqrt{d}} = \begin{cases} 1-\delta & \textit{if } \delta \le 1/2, \\ \delta & \textit{if } \delta > 1/2. \end{cases}$$

Proof of Theorem 3.4 From Lemma C.2, it is clear that the smallest α and hence r is attained with $\delta = 1/2$. Moreover, Lemma C.2 provides

$$\lim_{d\to\infty} V_{d,\mathbf{1/2},r} = \lim_{d\to\infty} C_{d,\mathbf{0},2r} = \begin{cases} 0 & \text{if } \alpha < 1/(2\sqrt{3}), \\ 1/2 & \text{if } \alpha = 1/(2\sqrt{3}), \\ 1 & \text{if } \alpha > 1/(2\sqrt{3}), \end{cases}$$

meaning for any $0 < \gamma < 1$, $r_{1-\gamma} = \sqrt{d}/(2\sqrt{3})$. By then applying Lemma C.3 with $\delta = 1/2$, we obtain $r_1 = \sqrt{d}/2$ and hence $r_{1-\gamma}/r_1 \to 1/\sqrt{3}$ as $d \to \infty$. □

Appendix D: Derivation of Approximation (30)

Let $U = (u_1, \dots, u_d)$ be a random vector with uniform distribution on $[-1, 1]^d$ so that $u_1, \dots, u_d$ are i.i.d.r.v. uniformly distributed on $[-1, 1]$. Then for given $Z = (z_1, \dots, z_d) \in \mathbb{R}^d$ and any $r > 0$,

$$C_{d,Z,r} = \mathbb{P}\{\|U - Z\| \le r\} = \mathbb{P}\{\|U - Z\|^2 \le r^2\} = \mathbb{P}\left\{\sum_{j=1}^d (u_j - z_j)^2 \le r^2\right\}.$$

That is, $C_{d,Z,r}$, as a function of r, is the c.d.f. of the r.v. $\|U - Z\|$.

Let u have the uniform distribution on $[-1, 1]$ and $z \in \mathbb{R}$. The first three central moments of the r.v. $\eta_z = (u - z)^2$ can be easily computed:

$$\mathbb{E}\eta_z = z^2 + \frac{1}{3}, \quad \mathrm{var}(\eta_z) = \frac{4}{3}\left(z^2 + \frac{1}{15}\right),$$

$$\mu_z^{(3)} = E[\eta_z - E\eta_z]^3 = \frac{16}{15}\left(z^2 + \frac{1}{63}\right). \tag{36}$$

Consider the r.v. $\|U - Z\|^2 = \sum_{i=1}^d \eta_{z_j} = \sum_{j=1}^d (u_j - z_j)^2$. From (36) and independence of $u_1, \ldots, u_d$, we obtain

$$\mu_{d,Z} = \mathbb{E}\|U - Z\|^2 = \|Z\|^2 + \frac{d}{3},$$
$$\sigma_{d,Z}^2 = \mathrm{var}(\|U - Z\|^2) = \frac{4}{3}\left(\|Z\|^2 + \frac{d}{15}\right), \qquad \text{and}$$
$$\mu_{d,Z}^{(3)} = \mathbb{E}[\|U - Z\|^2 - \mu_{d,Z}]^3 = \sum_{j=1}^d \mu_{z_j}^{(3)} = \frac{16}{15}\left(\|Z\|^2 + \frac{d}{63}\right).$$

If d is large enough then the conditions of the CLT for $\|U - Z\|^2$ are approximately met and the distribution of $\|U - Z\|^2$ is approximately normal with mean $\mu_{d,Z}$ and variance $\sigma_{d,Z}^2$. That is, we can approximate $C_{d,Z,r}$ by

$$C_{d,Z,r} \cong \Phi\left(\frac{r^2 - \mu_{d,Z}}{\sigma_{d,Z}}\right), \tag{37}$$

where $\Phi(\cdot)$ is the c.d.f. of the standard normal distribution:

$$\Phi(t) = \int_{-\infty}^t \varphi(v)\,dv \qquad \text{with} \quad \varphi(v) = \frac{e^{-v^2/2}}{\sqrt{2\pi}}.$$

The approximation (37) can be improved by using an Edgeworth-type expansion in the CLT for sums of independent non-identically distributed r.v.

General expansion in the central limit theorem for sums of independent non-identical r.v. has been derived by Petrov, see [5, Thm. 7, Chap. 6]; the first three terms of this expansion have been specialised in [6, Sect. 5.6]. By using only the first term in this expansion, we obtain the following approximation for the distribution function of $\|U - Z\|^2$:

$$P\left(\frac{\|U - Z\|^2 - \mu_{d,Z}}{\sigma_{d,Z}} \le x\right) \cong \Phi(x) + \frac{\mu_{d,Z}^{(3)}}{6(\sigma_{d,Z}^2)^{3/2}}(1 - x^2)\varphi(x),$$

leading to the following improved form of (37):

$$C_{d,Z,r} \cong \Phi(t) + \frac{\|Z\|^2 + d/63}{5\sqrt{3}(\|Z\|^2 + d/15)^{3/2}}(1 - t^2)\varphi(t),$$

where

$$t = t_{d,\|Z\|,r} = \frac{r^2 - \mu_{d,Z}}{\sigma_{d,Z}} = \frac{\sqrt{3}(r^2 - \|Z\|^2 - d/3)}{2\sqrt{\|Z\|^2 + d/15}}.$$

References

1. Conway, J.H., Sloane, N.J.A.: Sphere Packings, Lattices and Groups. Grundlehren der Mathematischen Wissenschaften, vol. 290. Springer, New York (1999)
2. Graf, S., Luschgy, H.: Foundations of Quantization for Probability Distributions. Lecture Notes in Mathematics, vol. 1730. Springer, Berlin (2000)
3. Kuo, F., Joe, S., Matsumoto, M., Mori, S., Saito, M.: Sobol sequences with better two-dimensional projections (2017). https://CRAN.R-project.org/package=SobolSequence
4. Noonan, J., Zhigljavsky, A.: Non-lattice covering and quantization of high dimensional sets. In: Black Box Optimization, Machine Learning, and No-Free Lunch Theorems, vol. 170, pp. 273–318. Springer, Cham (2021)
5. Petrov, V.V.: Sums of Independent Random Variables. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 82. Springer, New York (1975)
6. Petrov, V.V.: Limit Theorems of Probability Theory. Sequences of Independent Random Variables. Oxford Studies in Probability, vol. 4. Oxford Science Publications. Oxford University Press, New York (1995)
7. Pronzato, L.: Minimax and maximin space-filling designs: some properties and methods for construction. J. Société Française de Stat. **158**(1), 7–36 (2017)
8. Pronzato, L., Müller, W.G.: Design of computer experiments: space filling and beyond. Stat. Comput. **22**(3), 681–701 (2012)
9. Pronzato, L., Zhigljavsky, A.: Bayesian quadrature, energy minimization, and space-filling design. SIAM/ASA J. Uncertain. Quantif. **8**(3), 959–1011 (2020)
10. Saka, Y., Gunzburger, M., Burkardt, J.: Latinized, improved LHS, and CVT point sets in hypercubes. Int. J. Numer. Anal. Model. **4**(3–4), 729–743 (2007)
11. Santner, T.J., Williams, B.J., Notz, W.I.: The Design and Analysis of Computer Experiments. Springer Series in Statistics. Springer, New York (2003)
12. Schaback, R., Wendland, H.: Kernel techniques: from machine learning to meshless methods. Acta Numer. **15**, 543–639 (2006)
13. Wendland, H.: Scattered Data Approximation. Cambridge Monographs on Applied and Computational Mathematics, vol. 17. Cambridge University Press, Cambridge (2005)
14. Zhigljavsky, A., Noonan, J.: Covering of high-dimensional cubes and quantization. Oper. Res. Forum **1**(3), # 18 (2020)
15. Zhigljavsky, A., Žilinskas, A.: Bayesian and High-Dimensional Global Optimization. SpringerBriefs in Optimization. Springer, Cham (2021)