Component-by-component construction of low-discrepancy point sets of small size

Benjamin Doerr, Michael Gnewuch, Peter Kritzer and
Friedrich Pillichshammer

Abstract. We investigate the problem of constructing small point sets with low star discrepancy in the s-dimensional unit cube. The size of the point set shall always be polynomial in the dimension s. Our particular focus is on extending the dimension of a given low-discrepancy point set.

This results in a deterministic algorithm that constructs N-point sets with small discrepancy in a component-by-component fashion. The algorithm also provides the exact star discrepancy of the output set. Its run-time considerably improves on the run-times of the currently known deterministic algorithms that generate low-discrepancy point sets of comparable quality.

We also study infinite sequences of points with infinitely many components such that all initial subsegments projected down to all finite dimensions have low discrepancy. To this end, we introduce the inverse of the star discrepancy of such a sequence, and derive upper bounds for it as well as for the star discrepancy of the projections of finite subsequences with explicitly given constants. In particular, we establish the existence of sequences whose inverse of the star discrepancy depends linearly on the dimension.

Keywords. Star discrepancy, multivariate integration, derandomization.

AMS classification. Primary 11K38; Secondary 65D30.

1. Introduction

In numerical integration, point sets with good distribution properties are of great interest. One way of measuring the quality of distribution of a set $\mathcal{P}_s = \{\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{N-1}\}$ of N points in the s-dimensional unit cube $[0, 1)^s$, is to consider its star discrepancy, defined by

$$D_N^*(\mathcal{P}_s) = \sup_{\boldsymbol{x} \in [0,1]^s} |\Delta_s(\boldsymbol{x}, \mathcal{P}_s)| .$$

Here the discrepancy function Δ_s of the point set $\mathcal{P}_s$ is given, for $\boldsymbol{x} = (x_1, \ldots, x_s)$, by

$$\Delta_s(\boldsymbol{x}, \mathcal{P}_s) = \lambda_s([0, \boldsymbol{x})) - \frac{1}{N} \sum_{j=0}^{N-1} 1_{[0, \boldsymbol{x})}(\boldsymbol{p}_j),$$

where λ_s is the s-dimensional Lebesgue measure and $1_{[0,\boldsymbol{x})}$ is the characteristic function of the s-dimensional half-open box $[0, \boldsymbol{x}) = [0, x_1) \times \cdots \times [0, x_s)$.

Another useful quantity in the context of numerical integration is the so-called inverse of the star discrepancy,

$$N^*(\varepsilon, s) = \min\{N \,:\, \exists \mathcal{P}_s \subset [0,1)^s \text{ such that } |\mathcal{P}_s| = N \wedge D_N^*(\mathcal{P}_s) \leq \varepsilon\},$$

where $|\mathcal{P}_s|$ denotes the cardinality of the set $\mathcal{P}_s$. For certain classes of functions, point sets with small star discrepancy yield cubature formulas with a small error of multivariate integration. This is for example illustrated by the well-known Koksma–Hlawka inequality (see [12], but also [15, 17]). If we approximate the integral $\int_{[0,1]^s} f(\boldsymbol{x}) \, d\boldsymbol{x}$ by a quasi-Monte Carlo algorithm $\frac{1}{N} \sum_{i=0}^{N-1} f(\boldsymbol{p}_i)$, where $\boldsymbol{p}_0, \ldots, \boldsymbol{p}_{N-1} \in \mathcal{P}_s$, then the Koksma–Hlawka inequality states that

$$\left| \int_{[0,1]^s} f(\boldsymbol{x}) \, d\boldsymbol{x} - \frac{1}{N} \sum_{i=0}^{N-1} f(\boldsymbol{p}_i) \right| \leq V(f) D_N^*(\mathcal{P}_s),$$

where $V(f)$ is the variation of f in the sense of Hardy and Krause (see, e.g., [17]). This implies that, for any function with bounded variation $V(f)$, a point set with low star discrepancy yields a low integration error for a quasi-Monte Carlo rule using this point set as integration nodes.

However, it is a challenging problem to find point sets with good upper bounds on the star discrepancy. There are many constructions of point sets with low star discrepancy, for example those proposed by Niederreiter (so called (t, m, s)-nets and (t, s)-sequences, see [16]). Most of the known bounds on the star discrepancy of a (t, m, s)-net $\mathcal{P}_s$ are of the form

$$D_N^*(\mathcal{P}_s) \leq C_s \frac{(\log N)^s}{N} \text{ or } D_N^*(\mathcal{P}_s) \leq C_s \frac{(\log N)^{s-1}}{N},$$

where C_s is a constant depending on the dimension, see [14, 16, 17]. These bounds are excellent with respect to the order of magnitude in N. The drawback of such bounds, however, is that the term $(\log N)^s$ becomes large when s is very high and that in general the constants C_s are not small enough to compensate this effect. Therefore, the size of the node set for numerical integration needs to be extremely large if one wants to obtain good bounds on the integration error. In fact it has to be at least exponentially in s, since the function $N \mapsto \log(N)^s/N$ increases for $N \leq e^s$.

Since, during the last years, it has become a major issue to consider numerical integration in particularly high dimensions (e.g., in financial applications, where s might be in the hundreds or thousands), it is of growing interest to find bounds on the star discrepancy displaying a better dependence on the dimension s.

This problem was successfully attacked by Heinrich et al. in [10], who showed the existence of a point set $\mathcal{P}_s$ in $[0,1)^s$ such that

$$D_N^*(\mathcal{P}_s) \leq C\sqrt{s/N}, \tag{1.1}$$

where C is a positive constant not depending on s and N. However, the proof of this result is non-constructive, and, furthermore, good bounds for the constant have not been published yet. In another paper [5] the existence of point sets $\mathcal{P}_s$ satisfying

$$D_N^*(\mathcal{P}_s) \leq C' \sqrt{s/N} \sqrt{\log N}, \tag{1.2}$$

where $C' > 0$ is a known small constant and independent of s and N, was proved. This result is obtained by making use of the concept of so-called δ-covers (see below for a precise definition) and by Hoeffding's inequality. A slightly better result, namely $D_N^*(\mathcal{P}_s) \leq C'' \sqrt{s/N} \sqrt{\log(1 + N/s)}$, was proved in a similar manner in [7]. Like the result of Heinrich et al., the results in [5, 7] are at first based on a probabilistic argument and therefore seem to be non-constructive. On the other hand, in [5] also a deterministic construction of point sets satisfying (1.2) by the means of a derandomized version of Hoeffding's inequality was given. However, the run-time of this algorithm is high, to be more precise, it is exponential in the dimension s.

An algorithm using a different derandomization technique was presented in the recent paper [4]. It also generates an N-point set in dimension s satisfying the bound (1.2), but its run-time improves considerably over the algorithm from [5]. Nevertheless, it is still exponential in s.

The above-mentioned probabilistic bounds were developed further in [1] to allow infinite sequences of points in infinite dimension. For an infinite sequence $\mathcal{P}$ of points in $[0, 1)^{\mathbb{N}}$, let us denote by $\mathcal{P}_s$ the sequence of the s-dimensional projections of the points from $\mathcal{P}$, and by $\mathcal{P}_{N,s}$ the first N points of $\mathcal{P}_s$. Then in [1] the following results were shown:

There exists an unknown constant C such that for every strictly increasing sequence $(N_m)_{m \in \mathbb{N}}$ in $\mathbb{N}$ there is an infinite sequence $\mathcal{P}$ satisfying, for all $m, s \in \mathbb{N}$,

$$D_{N_m}^*(\mathcal{P}_{N_m,s}) \leq C \sqrt{s/N_m} \sqrt{\log(m + 1)}.$$

Furthermore, there exists an explicitly given constant C' such that for every strictly increasing sequence $(N_m)_{m \in \mathbb{N}}$ in $\mathbb{N}$ there is an infinite sequence $\mathcal{P}$ satisfying, for all $m, s \in \mathbb{N}$,

$$D_{N_m}^*(\mathcal{P}_{N_m,s}) \leq C' \sqrt{\left(m + s + s \log\left(1 + \frac{s\sqrt{N_m}}{m + s} \right) \right) / N_m}.$$

Thus, the results in [1] are both extensible in the dimension and in the number of points, which is particularly useful.

Our results

In this paper, we first present another result for infinite sequences $\mathcal{P}$ in $[0, 1)^{\mathbb{N}}$. At the first glance it looks like a modest improvement of [1, Corollary 3], but it establishes

the existence of infinite sequences $\mathcal{P}$ in $[0, 1)^{\mathbb{N}}$ having the following property: To guarantee $D_N^*(\mathcal{P}_{N,s}) \leq \varepsilon$ for a given ε, we only have to take $N \geq c_\varepsilon s$, where c_ε is a constant depending only on ε. Note that this result cannot be deduced directly from the results in [1]. It is known from [10, 11] that we have to take at least $N \geq c_\varepsilon' s$ if ε is sufficiently small. (Here c_ε' depends again only on ε.) In this sense our result shows that the statement "the inverse of the star discrepancy depends linearly on the dimension" (which is the title of the paper [10]) extends to the projections of infinite sequences in $[0, 1)^{\mathbb{N}}$. To be more precise, let us introduce the *inverse of the star discrepancy of an infinite sequence* $\mathcal{P}$,

$$N_{\mathcal{P}}^*(\varepsilon, s) := \min\{N : \forall M \geq N : D_M^*(\mathcal{P}_{M,s}) \leq \varepsilon\}.$$

Then there exist sequences $\mathcal{P}$ such that

$$N_{\mathcal{P}}^*(\varepsilon, s) \leq O(s\varepsilon^{-2} \log(1 + \varepsilon^{-1})). \tag{1.3}$$

In fact, if we endow the set $[0, 1)^{\mathbb{N}}$ with the canonical probability measure $\lambda_{\mathbb{N}} = \otimes_{i=1}^{\infty} \lambda_1$ and allow the implicit constant in the big-O-notation to depend on the particular sequence $\mathcal{P}$, then inequality (1.3) holds almost surely for a random sequence $\mathcal{P}$, see Corollary 2.4. In Theorem 2.3 we provide bounds of the form (1.3) with explicitly given constants and estimates for the measure of the sets of sequences satisfying such bounds.

It would be of great interest to construct infinite sequences whose star discrepancy satisfies bounds as stated above. In this paper we construct, in a component-by-component algorithm (i.e., one component is constructed at a time), finite sequences $\mathcal{P}$ in $[0, 1)^{\mathbb{N}}$ that satisfy similar bounds. For a given N-point set $\mathcal{P}_{s-1} = \{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{N-1}\} \subset [0, 1)^{s-1}$ we deduce from concentration of measure results that we can find $X_0, \ldots, X_{N-1} \in [0, 1)$ such that the extended set

$$\mathcal{P}_s = \{(\boldsymbol{y}_0, X_0), \ldots, (\boldsymbol{y}_{N-1}, X_{N-1})\} \subset [0, 1)^s$$

satisfies

$$D_N^*(\mathcal{P}_s) \leq O\left(\sqrt{\frac{s}{N} \log\left(1 + \frac{N}{s}\right)}\right) + D_N^*(\mathcal{P}_{s-1}).$$

We are able to derandomize the probabilistic argument to generate recursively (component by component) N-point sets $\mathcal{P}_s$ satisfying

$$D_N^*(\mathcal{P}_s) \leq O\left(\sqrt{\frac{s^3}{N} \log\left(1 + \frac{N}{s}\right)}\right).$$

This bound is obviously weaker than the bounds in (1.1) and (1.2), but the run-time of our derandomized algorithm improves considerably on the run-times of the algorithms in [4, 5] generating point sets satisfying (1.2). In Section 5 we compare the run-times of the algorithms to each other and relate the problem of constructing low-discrepancy samples of small size to the problem of approximating the star discrepancy of a given point set.

2. Infinite dimensional infinite sequences

At the beginning of this section we introduce two main tools that will also be useful in the following sections. The first is the concept of so-called δ-covers, as used in [4, 5, 7]. A δ-cover is defined as follows.

Definition 2.1. A finite set $\Gamma \subset [0,1]^s$ is called a δ-cover of $[0,1]^s$ if for every $\boldsymbol{x} = (x_1, \ldots, x_s) \in [0,1]^s$ there are $\boldsymbol{\gamma}_1 = (\gamma_1^{(1)}, \ldots, \gamma_1^{(s)}), \boldsymbol{\gamma}_2 = (\gamma_2^{(1)}, \ldots, \gamma_2^{(s)}) \in \Gamma \cup \{0\}$ with $\lambda_s([0, \boldsymbol{\gamma}_2)) - \lambda_s([0, \boldsymbol{\gamma}_1)) \leq \delta$ and $\gamma_1^{(i)} \leq x_i \leq \gamma_2^{(i)}$ for all $1 \leq i \leq s$.

One can use δ-covers to approximate the star discrepancy of a given set up to some admissible error δ by a quantity whose calculation involves only a finite number of test boxes.

Lemma 2.2. Let Γ be a δ-cover of $[0,1]^s$. Then for any N-point set $\mathcal{P}_s \subset [0,1)^s$ we have
$$D_N^*(\mathcal{P}_s) \leq D_N^{\Gamma}(\mathcal{P}_s) + \delta, \quad \text{where} \quad D_N^{\Gamma}(\mathcal{P}_s) = \max_{\boldsymbol{x} \in \Gamma} |\Delta_s(\boldsymbol{x}, \mathcal{P}_s)|.$$

The proof is straightforward and can, e.g., be found in [5, Lemma 3.1].

Furthermore, we are going to make use of a large deviation bound from probability theory called Hoeffding's inequality (see [13] or [18, p. 58], [19, p. 191]). Let $A_1, \ldots, A_r$ be independent random variables with $\mathbb{E}(A_i) = 0$ and $u_i \leq A_i \leq v_i$ for all $i \in \{1, \ldots, r\}$. Then Hoeffding's inequality states that

$$\mathbb{P}\left[|A_1 + \cdots + A_r| \geq r\eta\right] \leq 2 \cdot e^{-2r^2\eta^2 / \sum_{i=1}^r (v_i - u_i)^2} \tag{2.1}$$

for any $\eta > 0$.

In this section we improve certain aspects of the results in [1]. Let ζ denote the Riemann zeta function, i.e., $\zeta(\gamma) = \sum_{m=1}^{\infty} m^{-\gamma}$. As mentioned in the introduction, we denote, for an infinite sequence $\mathcal{P}$ of points in $[0,1)^{\mathbb{N}}$, by $\mathcal{P}_s$ the sequence of s-dimensional projections of the points from $\mathcal{P}$, and by $\mathcal{P}_{N,s}$ the first N points of $\mathcal{P}_s$.

Theorem 2.3. Let $\mathcal{P} = (p_N)_{N \in \mathbb{N}}$ be a sequence of independent random variables, each of them uniformly distributed with respect to the (infinite) product measure $\lambda_{\mathbb{N}} := \otimes_{N=1}^{\infty} \lambda_1$ on $[0,1)^{\mathbb{N}}$. Let $(N_m)_{m \in \mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$, and let $A \in (1, \infty)$. Let $\gamma > \zeta^{-1}(2)$. Then each of the following two events holds with probability at least $1 - (\zeta(\gamma) - 1)^2$.

(i) For all $m, s \in \mathbb{N}$, with $\rho = \rho(N_m, s) := 6e(\max\{1, N_m/(2\log(6e)s)\})^{1/2}$,

$$D_{N_m}^*(\mathcal{P}_{N_m,s}) \leq \left(\frac{2}{N_m}\left(s\log(\rho) + \gamma(\log(1+m) + \log(1+s)) + \log(2)\right)\right)^{1/2}. \tag{2.2}$$

(ii) *For all $\varepsilon \in (0, 1]$,*

$$N_{\mathcal{P}}^*(\varepsilon, s) \leq \left\lceil 2A^2 \varepsilon^{-2} \Big(s\big(\log(1 + 2A\varepsilon^{-1}) + \log(2e) \big) + \right.$$

$$\left. + \gamma \left(\log\left(2 + \frac{\log(\varepsilon^{-1})}{\log(A)} \right) + \log(1+s) \right) + \log(1 + 2A^2\varepsilon^{-2}) + \log(2) \Big) \right\rceil.$$

$$(2.3)$$

Proof. Let $\theta \in [0, 1)$. Due to Lemma 2.2 we have $\mathbb{P}\left[D_N^*(\mathcal{P}_{N,s}) \leq 2\delta \right] > \theta$ if the inequality $\mathbb{P}\left[D_N^\Gamma(\mathcal{P}_{N,s}) \leq \delta \right] > \theta$ holds for some δ-cover Γ of $[0,1]^s$. For each $s \in \mathbb{N}$ let $p_N^{(s)}$ denote the projection of p_N onto its first s components. If we define for $\boldsymbol{x} \in [0,1]^s$ and $i \in \mathbb{N}$ the random variable $\xi_{\boldsymbol{x}}(p_i^{(s)}) = \lambda_s([0, \boldsymbol{x})) - 1_{[0,\boldsymbol{x})}(p_i^{(s)})$, then the range of $\xi_{\boldsymbol{x}}(p_i^{(s)})$ is contained in an interval of length one and the expectation $\mathbb{E}[\xi_{\boldsymbol{x}}(p_i^{(s)})]$ is zero. Thus Hoeffding's inequality implies

$$\mathbb{P}[|\Delta_s(\boldsymbol{x}, \mathcal{P}_{N,s})| > \delta] = \mathbb{P}\left[\left| \frac{1}{N} \sum_{i=1}^N \xi_{\boldsymbol{x}}(p_i^{(s)}) \right| > \delta \right] \leq 2e^{-2\delta^2 N}.$$

This results in

$$\mathbb{P}\left[D_N^\Gamma(\mathcal{P}_{N,s}) \leq \delta \right] \geq 1 - \sum_{\boldsymbol{x} \in \Gamma} \mathbb{P}\left[\left| \frac{1}{N} \sum_{i=1}^N \xi_{\boldsymbol{x}}(p_i^{(s)}) \right| > \delta \right] > 1 - 2|\Gamma| e^{-2\delta^2 N}.$$

(In the latter estimate we get ">", since we have necessarily $\boldsymbol{1} := (1, \ldots, 1) \in \Gamma$ and $|\Delta_s(\boldsymbol{1}, P)| = 0$ for all finite sets $P \subset [0,1)^s$.) Hence $\mathbb{P}\left[D_N^*(\mathcal{P}_{N,s}) \leq 2\delta \right] > \theta$ is satisfied if

$$2\delta^2 N \geq \log |\Gamma| + \log\left(\frac{2}{1-\theta} \right) \tag{2.4}$$

holds. In [7, Theorem 1.15] a δ-cover Γ was constructed satisfying

$$|\Gamma| \leq 2^s \frac{s^s}{s!} (\delta^{-1} + 1)^s \leq (2e)^s (\delta^{-1} + 1)^s. \tag{2.5}$$

If we choose this δ-cover and restrict ourselves to $\delta \in (0, 1/2]$, then it is easily verified that

$$\delta := \frac{1}{\sqrt{2N}} \left(s \log(\rho) + \log\left(\frac{2}{1-\theta} \right) \right)^{1/2} \tag{2.6}$$

satisfies (2.4). Thus

$$D_N^*(\mathcal{P}_{N,s}) \leq \left(\frac{2}{N} \left(s \log(\rho) + \log\left(\frac{2}{1-\theta} \right) \right) \right)^{1/2}$$

with probability at least θ.

Now let $(N_m)_{m \in \mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$ and define, for $\gamma > 1$,

$$\theta = \theta(m, s, \gamma) = 1 - \left((1 + m)(1 + s)\right)^{-\gamma}. \tag{2.7}$$

The probability that there exists a pair $(m, s) \in \mathbb{N}^2$ such that

$$D^*_{N_m}(\mathcal{P}_{N_m,s}) > \left(\frac{2}{N_m}\left(s\log(\rho) + \log\left(\frac{2}{1 - \theta(m, s, \gamma)}\right)\right)\right)^{1/2}$$

is bounded from above by

$$\sum_{m,s=1}^{\infty} (1 - \theta(m, s, \gamma)) = \left(\sum_{m=1}^{\infty}(1 + m)^{-\gamma}\right)\left(\sum_{s=1}^{\infty}(1 + s)^{-\gamma}\right) = (\zeta(\gamma) - 1)^2.$$

Since the Riemann zeta function is strictly decreasing, the latter expression is strictly less than one if γ is strictly larger than $\zeta^{-1}(2)$. This proves (2.2).

One might be tempted to use inequality (2.2) directly with the sequence $(N_m)_{m \in \mathbb{N}}$ $= (N)_{N \in \mathbb{N}}$ to derive a bound like (2.3) for $N^*_{\mathcal{P}}(\varepsilon, s)$. But this is not easily done, since it is hard to solve the resulting inequality with respect to N. For that reason we again use (2.4) to derive statement *(ii)* of the theorem.

Let us put $\varepsilon := 2\delta$. We know that each N with

$$N \geq 2\varepsilon^{-2}\left(s\left(\log\left(1 + 2\varepsilon^{-1}\right) + \log(2e)\right) + \log\left(\frac{2}{1 - \theta}\right)\right)$$

implies $\mathbb{P}\left[D^*_N(\mathcal{P}_{N,s}) \leq \varepsilon\right] > \theta$, see (2.4) and (2.5). Now put $\varepsilon_m := A^{-m}$ and

$$\theta' = \theta'(m, s, \gamma) = 1 - \frac{((1 + m)(1 + s))^{-\gamma}}{1 + 2\varepsilon_m^{-2}}.$$

For fixed m and s put

$$N_{m,s} := \left\lceil 2\varepsilon_m^{-2}\left(s\left(\log(1 + 2\varepsilon_m^{-1}) + \log(2e)\right) + \log\left(\frac{2}{1 - \theta'(m, s, \gamma)}\right)\right)\right\rceil.$$

For all $N \geq N_{m,s}$ choose $\vartheta(N)$ such that

$$N = 2\varepsilon_m^{-2}\left(s\left(\log(1 + 2\varepsilon_m^{-1}) + \log(2e)\right) + \log\left(\frac{2}{1 - \vartheta(N)}\right)\right),$$

that is,

$$\vartheta(N) = 1 - 2\exp\left(-\frac{1}{2}\varepsilon_m^2 N\right)\left((2e)(1 + 2\varepsilon_m^{-1})\right)^s \geq \theta'(m, s, \gamma).$$

Hence the probability that there exists an $N \geq N_{m,s}$ such that $D_N^*(\mathcal{P}_{N,s}) > \varepsilon_m$ is bounded from above by

$$
\sum_{N=N_{m,s}}^{\infty} (1 - \vartheta(N)) = 2\big((2\mathrm{e})(1 + 2\varepsilon_m^{-1})\big)^s \sum_{N=N_{m,s}}^{\infty} \exp\left(-\frac{1}{2}\varepsilon_m^2 N\right)
$$

$$
\leq 2\big((2\mathrm{e})(1 + 2\varepsilon_m^{-1})\big)^s \left(\exp\left(-\frac{1}{2}\varepsilon_m^2 N_{m,s}\right) + \int_{N_{m,s}}^{\infty} \exp\left(-\frac{1}{2}\varepsilon_m^2 x\right)\, \mathrm{d}x\right)
$$

$$
= 2\big((2\mathrm{e})(1 + 2\varepsilon_m^{-1})\big)^s (1 + 2\varepsilon_m^{-2}) \exp\left(-\frac{1}{2}\varepsilon_m^2 N_{m,s}\right)
$$

$$
\leq (1 + 2\varepsilon_m^{-2})(1 - \theta'(m,s,\gamma)) = ((1+m)(1+s))^{-\gamma}.
$$

Thus the probability that there exists a pair $(m,s) \in \mathbb{N}^2$ such that $N_{\mathcal{P}}^*(\varepsilon_m, s) > N_{m,s}$ is again bounded by $(\zeta(\gamma) - 1)^2$. If $\gamma > \zeta^{-1}(2)$, there exists an infinite sequence $\mathcal{P}$ such that $N_{\mathcal{P}}^*(\varepsilon_m, s)$ is majorized by $N_{m,s}$ for all m, s. Furthermore, we see that for $\varepsilon \in [A^{-m}, A^{1-m})$ the simple estimate $N_{\mathcal{P}}^*(\varepsilon, s) \leq N_{\mathcal{P}}^*(\varepsilon_m, s)$ implies (2.3). This completes the proof of Theorem 2.3.   □

Theorem 2.3 immediately implies the following corollary.

Corollary 2.4. *For an infinite sequence $\mathcal{P}$ in $[0,1)^{\mathbb{N}}$ the following holds with probability one: There exist constants $C_{\mathcal{P}}$, $C_{\mathcal{P}}'$ such that*

$$
D_N^*(\mathcal{P}_{N,s}) \leq C_{\mathcal{P}} \left(\frac{s}{N} \log\left(1 + \frac{N}{s}\right)\right)^{1/2} \quad and \quad N_{\mathcal{P}}^*(\varepsilon, s) \leq C_{\mathcal{P}}' s \varepsilon^{-2} \log\left(1 + \varepsilon^{-1}\right)
$$

for all $N, s \in \mathbb{N}$ and all $\varepsilon \in (0,1]$.

3. Extensions in the dimension

Now we will try to relate the star discrepancies of an arbitrary $(s-1)$-dimensional point set and a suitably chosen s-dimensional extension of it. To be more precise, let $\mathcal{P}_{s-1} = \{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{N-1}\} \subset [0,1)^{s-1}$ be a point set with star discrepancy $D_N^*(\mathcal{P}_{s-1})$. Let $m \geq 2$ be an integer, and let

$$
G = G(m) := \left\{\frac{1}{2m}, \frac{3}{2m}, \ldots, \frac{2m-1}{2m}\right\}.
$$

For $a_0, \ldots, a_{N-1} \in G$ we consider the point set given by

$$
\mathcal{P}_s = \mathcal{P}_s(a_0, \ldots, a_{N-1}) = \{(\boldsymbol{y}_i, a_i) : 0 \leq i < N\}. \tag{3.1}
$$

The following theorem shows that there is always a choice of $a_0, \ldots, a_{N-1} \in G(m)$ such that we can bound the star discrepancy of $\mathcal{P}_s$ as constructed in (3.1) in terms of m, N, and the star discrepancy of $\mathcal{P}_{s-1}$.

Theorem 3.1. *Let $m \geq 2$ be an integer and let $0 \leq \theta < 1$ be a real. Assume that $a_0, \ldots, a_{N-1}$ are chosen independently and uniformly distributed from the set $G(m)$. Then with probability greater than θ we have for all $N \in \mathbb{N}$*

$$D_N^*(\mathcal{P}_s) \leq \frac{\sqrt{2}}{\sqrt{N}} \left(s \log(\rho(N,s)) + \log\left(\frac{2}{1-\theta}\right) \right)^{1/2} + \frac{1}{2m} + D_N^*(\mathcal{P}_{s-1}),$$

where $\rho = \rho(N,s) := 6e(\max\{1, N/(2\log(6e)s)\})^{1/2}$.

Proof. For $\boldsymbol{x} = (x_1, \ldots, x_s) \in [0,1)^s$, $\boldsymbol{y} = (y_1, \ldots, y_{s-1}) \in \mathcal{P}_{s-1}$ and $a \in G$ let

$$\xi_{\boldsymbol{x}}(\boldsymbol{y}, a) = \lambda_s([0, \boldsymbol{x})) - 1_{[0,\boldsymbol{x})}(\boldsymbol{y}, a).$$

Assume that A is chosen according to a uniform distribution from the set G. Then we have

$$\mathbb{E}\left[\xi_{\boldsymbol{x}}((\boldsymbol{y}, A))\right] = \prod_{d=1}^{s} x_d - \prod_{d=1}^{s-1} 1_{[0,x_d)}(y_d)\mathbb{E}\left[1_{[0,x_s)}(A)\right].$$

As

$$\mathbb{E}\left[1_{[0,x_s)}(A)\right] = \frac{1}{m}\sum_{a \in G} 1_{[0,x_s)}(a) = \frac{\lceil mx_s - 1/2 \rceil}{m} = \begin{cases} \leq x_s + 1/(2m), \\ \geq x_s - 1/(2m), \end{cases}$$

it follows that

$$x_s \left(\prod_{d=1}^{s-1} x_d - \prod_{d=1}^{s-1} 1_{[0,x_d)}(y_d)\right) - \frac{1}{2m} \leq \mathbb{E}\left[\xi_{\boldsymbol{x}}(\boldsymbol{y}, A)\right]$$

$$\leq x_s \left(\prod_{d=1}^{s-1} x_d - \prod_{d=1}^{s-1} 1_{[0,x_d)}(y_d)\right) + \frac{1}{2m}.$$

If $A_0, \ldots, A_{N-1}$ are chosen independently and if $\mathcal{P}_s = \mathcal{P}_s(A_0, \ldots, A_{N-1})$, then we obtain

$$x_s\Delta_{s-1}((x_1, \ldots, x_{s-1}), \mathcal{P}_{s-1}) - \frac{1}{2m} \leq \mathbb{E}\left[\Delta_s(\boldsymbol{x}, \mathcal{P}_s)\right]$$

$$\leq x_s\Delta_{s-1}((x_1, \ldots, x_{s-1}), \mathcal{P}_{s-1}) + \frac{1}{2m},$$

where the expectation is taken with respect to $A_0, \ldots, A_{N-1}$. Therefore we get

$$|\mathbb{E}\left[\Delta_s(\boldsymbol{x}, \mathcal{P}_s)\right]| \leq \frac{1}{2m} + |x_s\Delta_{s-1}((x_1, \ldots, x_{s-1}), \mathcal{P}_{s-1})| \leq \frac{1}{2m} + D_N^*(\mathcal{P}_{s-1}). \quad (3.2)$$

Let now $0 < \delta \leq 1/2$. Assume that

$$|\Delta_s(\boldsymbol{x}, \mathcal{P}_s)| \geq D_N^*(\mathcal{P}_{s-1}) + \frac{1}{2m} + \delta.$$

This implies $|\Delta_s(\boldsymbol{x}, \mathcal{P}_s) - \mathbb{E}\left[\Delta_s(\boldsymbol{x}, \mathcal{P}_s)\right]| \geq \delta$ or equivalently

$$\left| \sum_{i=0}^{N-1} \left(\xi_{\boldsymbol{x}}(\boldsymbol{y}_i, A_i) - \mathbb{E}\left[\xi_{\boldsymbol{x}}(\boldsymbol{y}_i, A_i)\right] \right) \right| \geq \delta N.$$

We can apply Hoeffding's inequality to obtain

$$\mathbb{P}\left[|\Delta_s(\boldsymbol{x}, \mathcal{P}_s)| \geq D_N^*(\mathcal{P}_{s-1}) + \frac{1}{2m} + \delta \right]$$

$$\leq \mathbb{P}\left[\left| \sum_{i=0}^{N-1} \left(\xi_{\boldsymbol{x}}(\boldsymbol{y}_i, A_i) - \mathbb{E}[\xi_{\boldsymbol{x}}(\boldsymbol{y}_i, A_i)] \right) \right| \geq \delta N \right] \leq 2\mathrm{e}^{-2\delta^2 N}.$$

Similar to the proof of Theorem 2.3 we get, for a δ-cover Γ of $[0, 1]^s$,

$$\mathbb{P}\left[D_N^*(\mathcal{P}_s) \leq 2\delta + \frac{1}{2m} + D_N^*(\mathcal{P}_{s-1}) \right] \geq \mathbb{P}\left[D_N^\Gamma(\mathcal{P}_s) \leq \delta + \frac{1}{2m} + D_N^*(\mathcal{P}_{s-1}) \right]$$

$$> 1 - 2|\Gamma|\mathrm{e}^{-2\delta^2 N}.$$

The latter term is greater than or equal to θ if

$$2\delta^2 N \geq \log|\Gamma| + \log\left(\frac{2}{1-\theta}\right). \tag{3.3}$$

From [7, Theorem 1.15] we get $|\Gamma| \leq (2\mathrm{e})^s (\delta^{-1} + 1)^s$, hence any solution of

$$\delta^2 \geq \frac{1}{2N}\left(s \log\left(3\mathrm{e}\delta^{-1}\right) + \log\left(\frac{2}{1-\theta}\right) \right) \tag{3.4}$$

satisfies (3.3). Put now

$$\delta = \delta(N, s) := \frac{1}{\sqrt{2N}}\left(s \log\left(\rho\right) + \log\left(\frac{2}{1-\theta}\right) \right)^{1/2}. \tag{3.5}$$

It is easy to check that this δ satisfies (3.4). $\qquad\square$

The result in Theorem 3.1 gives us the opportunity to construct point sets with low star discrepancy in a component-by-component fashion. The following corollary shows an upper bound on the discrepancy of such a point set. Note that this bound displays a worse dependence on the dimension s. The advantage of the construction is that the points lie on a relatively small mesh (with not necessarily the same resolution in each dimension). This allows a cheaper computation of the precise discrepancy (see Equation (4.1) in Section 4) and also a more efficient derandomization (see the remainder of Section 4).

Corollary 3.2. *For any $N \geq 1$ and any sequence $(m_d)_{d \geq 1}$ of positive integers there exists a point set $\mathcal{P} = \{\boldsymbol{x}_0, \dots, \boldsymbol{x}_{N-1}\} \subset [0, 1)^{\mathbb{N}}$ with*

$$\boldsymbol{x}_i = \left(\frac{2i+1}{2N}, \frac{2a_{i,1}+1}{2m_1}, \frac{2a_{i,2}+1}{2m_2}, \dots \right)$$

where $a_{i,d} \in \{0, \dots, m_d - 1\}$ for all $d \geq 1$ and $0 \leq i < N$ such that for all $s \geq 1$ for the point set $\mathcal{P}_s$, consisting of the projections of the points from $\mathcal{P}$ to the first s components, we have

$$D_N^*(\mathcal{P}_s) \leq \frac{\sqrt{2}}{\sqrt{N}} (s-1) \left(s \log \left(\rho(N, s) \right) + \log(2) \right)^{1/2} + \frac{1}{2} \sum_{d=1}^{s-1} \frac{1}{m_d} + \frac{1}{2N}, \quad (3.6)$$

where $\rho(N, s)$ is as in Theorem 3.1.

Proof. It is easy to check that $D_N^*(\mathcal{P}_1) = 1/(2N)$ (see [17, Theorem 2.6]). Due to Theorem 3.1, we can recursively choose sets $\{x_{0,d}, \dots, x_{N-1,d}\}$ such that

$$D_N^*(\mathcal{P}_d) \leq \frac{\sqrt{2}}{\sqrt{N}} \sum_{j=2}^{d} (j \log(\rho(N, j)) + \log(2))^{1/2} + \frac{1}{2} \sum_{j=1}^{d-1} \frac{1}{m_j} + \frac{1}{2N}$$

holds for all $d = 2, \dots, s$. Simple calculus shows that $s \log(\rho(N, s)) \geq d \log(\rho(N, d))$ for $d \leq s$. $\qquad \square$

4. A construction algorithm

One may use Theorem 3.1 and Corollary 3.2 to construct low-discrepancy point sets via a derandomized algorithm similar to what has been done in [5]. Here, we would like to use a slightly different algorithm, since the sets we want to construct are contained in a grid whose cardinality is of the same order as the smallest δ-cover constructed in [7]; instead of only approximating the discrepancy of such a set up to some admissible error δ with the help of a δ-cover, we can calculate its discrepancy exactly with at most the same effort. We make use of the following simple observation:

Let $\mathcal{P}_s = \{\boldsymbol{p}_0, \boldsymbol{p}_1, \dots, \boldsymbol{p}_{N-1}\} \subset [0, 1)^s$ with $\boldsymbol{p}_i = (p_{i,1}, \dots, p_{i,s})$. Define for $d = 1, \dots, s$

$$\Gamma_d(\mathcal{P}_s) := \{p_{0,d}, \dots, p_{N-1,d}\} \quad \text{and} \quad \overline{\Gamma}_d(\mathcal{P}_s) := \Gamma_d(\mathcal{P}_s) \cup \{1\}.$$

Furthermore, set

$$\Gamma(\mathcal{P}_s) := \Gamma_1(\mathcal{P}_s) \times \cdots \times \Gamma_s(\mathcal{P}_s) \quad \text{and} \quad \overline{\Gamma}(\mathcal{P}_s) := \overline{\Gamma}_1(\mathcal{P}_s) \times \cdots \times \overline{\Gamma}_s(\mathcal{P}_s).$$

It is easy to see that

$$
D_N^*(\mathcal{P}_s) = \max \left\{ \max_{\boldsymbol{x} \in \Gamma(\mathcal{P}_s)} \left(\frac{1}{N} \sum_{i=0}^{N-1} 1_{[\boldsymbol{0}, \boldsymbol{x}]}(\boldsymbol{p}_i) - \lambda_s([\boldsymbol{0}, \boldsymbol{x})) \right), \right.
$$
$$
\left. \max_{\boldsymbol{x} \in \overline{\Gamma}(\mathcal{P}_s)} \left(\lambda_s([\boldsymbol{0}, \boldsymbol{x})) - \frac{1}{N} \sum_{i=0}^{N-1} 1_{[\boldsymbol{0}, \boldsymbol{x})}(\boldsymbol{p}_i) \right) \right\}. \tag{4.1}
$$

Observe that for general (more precisely: almost all) N-point sets $\mathcal{P}_s$ the evaluation of the right hand side of (4.1) involves $\Theta((N+1)^s)$ test boxes, while the approximation of their discrepancy up to δ chosen as in (3.5) via a δ-cover as constructed in [7, Theorem 1.15] requires at most

$$
O\left((2\mathrm{e})^s \left(\sqrt{\frac{2N}{s}} \left(\log\left(1 + \frac{N}{s}\right) \right)^{-1/2} + 1 \right)^s \right) \tag{4.2}
$$

test boxes. (In [8] a smaller δ-cover has been constructed in dimension $s = 2$. There it is conjectured that the methods can be extended to arbitrary s and would give δ-covers of size $2\delta^{-s} + O_s(\delta^{-s+1})$, which would be considerably smaller than the δ-covers constructed in [7].) However, in the derandomized construction we have in mind we will confine ourselves to sets $\mathcal{P}_s$ having the property that the cardinality of $\Gamma(\mathcal{P}_s)$ and $\overline{\Gamma}(\mathcal{P}_s)$ is at most of the order of (4.2). To describe our approach in detail, we have to introduce further notation. Let $m_1, m_2, \ldots, m_s \in \mathbb{N}$, and, for $d = 1, \ldots, s$, let

$$
G_d := G(m_d) = \left\{ \frac{1}{2m_d}, \frac{3}{2m_d}, \ldots, \frac{2m_d - 1}{2m_d} \right\} \quad \text{and} \quad \widetilde{G}_d := (G_d \cup \{1\}) \setminus \left\{ \frac{1}{2m_d} \right\}.
$$

Furthermore, put

$$
\mathcal{G}_s := G_1 \times \cdots \times G_s \quad \text{and} \quad \widetilde{\mathcal{G}}_s := \widetilde{G}_1 \times \cdots \times \widetilde{G}_s.
$$

Let $\mathcal{P}_{s-1} = \{\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{N-1}\}$ be a subset of $\mathcal{G}_{s-1}$. For $X_0, \ldots, X_{N-1} \in G_s$ we consider the point set

$$
\mathcal{P}_s = \mathcal{P}_s(X_0, \ldots, X_{N-1}) = \{(\boldsymbol{y}_i, X_i) : 0 \le i < N\}. \tag{4.3}
$$

To avoid having to distinguish between open and closed test boxes and to reduce the number of events we have to control in our random experiment, we reformulate (4.1). It is easy to verify

$$
D_N^*(\mathcal{P}_s) = \max \left\{ \max_{\boldsymbol{x} \in \mathcal{G}_s} \left(\frac{1}{N} \sum_{i=0}^{N-1} 1_{[\boldsymbol{0}, \boldsymbol{x}]}((\boldsymbol{y}_i, X_i)) - \lambda_s([\boldsymbol{0}, \boldsymbol{x})) \right), \right.
$$
$$
\left. \max_{\boldsymbol{x} \in \widetilde{\mathcal{G}}_s} \left(\lambda_s([\boldsymbol{0}, \boldsymbol{x})) - \frac{1}{N} \sum_{i=0}^{N-1} 1_{[\boldsymbol{0}, \boldsymbol{x})}((\boldsymbol{y}_i, X_i)) \right), \frac{1}{2m_s}, D_N^*(\mathcal{P}_{s-1}) \right\}. \tag{4.4}
$$

For $d = 1, \ldots, s$, put

$$G_d^* := \left\{ \frac{1}{m_d}, \frac{2}{m_d}, \ldots, \frac{m_d - 1}{m_d}, 1 \right\}.$$

Additionally, put

$$\mathcal{G}_s^* := G_1^* \times \cdots \times G_s^*.$$

For $\boldsymbol{x} \in \mathcal{G}_s$ we then find a uniquely determined $\boldsymbol{t} \in \mathcal{G}_s^*$ such that

$$|[\boldsymbol{0}, \boldsymbol{x}] \cap S| = |[\boldsymbol{0}, \boldsymbol{t}] \cap S| \quad \text{for all } S \subseteq \mathcal{G}_s;$$

put $\Phi([\boldsymbol{0}, \boldsymbol{x}]) := \boldsymbol{t}$. For $\widetilde{\boldsymbol{x}} \in \widetilde{\mathcal{G}}_s$ we find furthermore a uniquely determined $\tilde{\boldsymbol{t}} \in \mathcal{G}_s^*$ such that

$$|[\boldsymbol{0}, \widetilde{\boldsymbol{x}}) \cap S| = |[\boldsymbol{0}, \tilde{\boldsymbol{t}}) \cap S| \quad \text{for all } S \subseteq \mathcal{G}_s;$$

put $\Phi([\boldsymbol{0}, \widetilde{\boldsymbol{x}})) := \tilde{\boldsymbol{t}}$. This gives us a mapping Φ from the set

$$\mathcal{B} := \{ [\boldsymbol{0}, \boldsymbol{x}] \, : \, \boldsymbol{x} \in \mathcal{G}_s \} \cup \{ [\boldsymbol{0}, \widetilde{\boldsymbol{x}}) \, : \, \widetilde{\boldsymbol{x}} \in \widetilde{\mathcal{G}}_s \}$$

onto $\mathcal{G}_s^*$. Now let $n := |\mathcal{G}_s^*| = \prod_{d=1}^{s} m_d$, and let $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n$ be an enumeration of $\mathcal{G}_s^*$. For $i = 1, \ldots, n$, $\boldsymbol{y} \in \mathcal{P}_{s-1}$, and a randomly chosen $X \in G_s$ let

$$\xi_i(\boldsymbol{y}, X) = 1_{[\boldsymbol{0}, \boldsymbol{t}_i)}(\boldsymbol{y}, X) - \lambda_s([\boldsymbol{0}, \boldsymbol{t}_i)).$$

Then

$$\Xi_i(\boldsymbol{y}, X) := \xi_i(\boldsymbol{y}, X) - \mathbb{E}[\xi_i(\boldsymbol{y}, X)] = 1_{[\boldsymbol{0}, \boldsymbol{t}_i)}(\boldsymbol{y}, X) - t_{i,s} \, 1_{[\boldsymbol{0}, \boldsymbol{t}_i')}(\boldsymbol{y}), \qquad (4.5)$$

where we used the convention to denote the projection of an s-dimensional vector $\boldsymbol{t}$ onto its first $s - 1$ components by $\boldsymbol{t}'$. We use the corresponding convention for subsets of $[0, 1]^s$. Let $B \in \mathcal{B}$ and $\boldsymbol{t}_i = \Phi(B)$. Denote by $B^{(s)}$ the projection of B to the s-th component. Then

$$\frac{1}{N} \sum_{j=0}^{N-1} 1_B((\boldsymbol{y}_j, X_j)) - \lambda_s(B) = \frac{1}{N} \sum_{j=0}^{N-1} 1_{[\boldsymbol{0}, \boldsymbol{t}_i)}((\boldsymbol{y}_j, X_j)) - \lambda_1(B^{(s)}) \lambda_{s-1}(B')$$

$$= \frac{1}{N} \sum_{j=0}^{N-1} \Xi_i(\boldsymbol{y}_j, X_j) + t_{i,s} \left(\frac{1}{N} \sum_{j=0}^{N-1} 1_{[\boldsymbol{0}, \boldsymbol{t}_i')}(\boldsymbol{y}_j) - \lambda_{s-1}(B') \right) +$$

$$+ \lambda_{s-1}(B')(t_{i,s} - \lambda_1(B^{(s)})).$$

Due to $1_{[\boldsymbol{0}, \boldsymbol{t}_i')}(\boldsymbol{y}_j) = 1_{B'}(\boldsymbol{y}_j)$ for $0 \le j < N$ and $|t_{i,s} - \lambda_1(B^{(s)})| \le 1/(2m_s)$ we obtain from (4.4)

$$D_N^*(\mathcal{P}_s) \le \frac{1}{N} \max_{1 \le i \le n} \left| \sum_{j=0}^{N-1} \Xi_i(\boldsymbol{y}_j, X_j) \right| + D_N^*(\mathcal{P}_{s-1}) + \frac{1}{2m_s}. \qquad (4.6)$$

From Hoeffding's inequality we get

$$\mathbb{P}\left[\max_{1\leq i\leq n}\left|\sum_{j=0}^{N-1}\Xi_i(\boldsymbol{y}_j,X_j)\right|\geq\delta N\right]\leq 2n\mathrm{e}^{-2\delta^2 N}. \tag{4.7}$$

This results in

$$\mathbb{P}\left[D_N^*(\mathcal{P}_s)\leq\delta+D_N^*(\mathcal{P}_{s-1})+\frac{1}{2m_s}\right]\geq 1-2n\mathrm{e}^{-2\delta^2 N}.$$

Now the latter term is greater than or equal to θ if

$$2\delta^2 N\geq\log(n)+\log\left(\frac{2}{1-\theta}\right). \tag{4.8}$$

The choice

$$\delta:=\frac{1}{\sqrt{2N}}\left(\sum_{d=1}^{s}\log(m_d)+\log\left(\frac{2}{1-\theta}\right)\right)^{1/2} \tag{4.9}$$

gives us the following theorem.

Theorem 4.1. *Let $\theta\in[0,1)$. Assume that $X_0,\ldots,X_{N-1}$ are chosen independently and uniformly distributed from G_s. Then with probability $\geq\theta$ we have*

$$D_N^*(\mathcal{P}_s)\leq\frac{1}{\sqrt{2N}}\left(\sum_{d=1}^{s}\log(m_d)+\log\left(\frac{2}{1-\theta}\right)\right)^{1/2}+\frac{1}{2m_s}+D_N^*(\mathcal{P}_{s-1}).$$

Let us now choose

$$m_d=m_d(N,\theta)=\left\lceil\frac{\sqrt{N}}{\sqrt{2}}\left(d\log\left(\rho'(N,d)\right)+\log\left(\frac{2}{1-\theta}\right)\right)^{-1/2}\right\rceil \tag{4.10}$$

for $d=1,\ldots,s$, where

$$\rho'(N,d)=2\sqrt{\mathrm{e}}\left(\max\{1,N/(2\log(2\sqrt{\mathrm{e}})d)\}\right)^{1/2}. \tag{4.11}$$

An elementary analysis shows $m_1\geq m_2\geq\cdots\geq m_s$. Now we can estimate n in terms of N,s and θ:

$$n=\prod_{d=1}^{s}m_d\leq\prod_{d=1}^{s}\left(\frac{\sqrt{N}}{\sqrt{2d}}\left(\log\left(\rho'(N,d)\right)+\frac{1}{d}\log\left(\frac{2}{1-\theta}\right)\right)^{-1/2}+1\right)$$

$$=\left(\frac{s^s}{s!}\right)^{1/2}\prod_{d=1}^{s}\left(\frac{\sqrt{N}}{\sqrt{2s}}\left(\log\left(\rho'(N,d)\right)+\frac{1}{d}\log\left(\frac{2}{1-\theta}\right)\right)^{-1/2}+\frac{\sqrt{d}}{\sqrt{s}}\right)$$

$$<(2\pi s)^{-1/4}\mathrm{e}^{s/2}\left(\frac{\sqrt{N}}{\sqrt{2s}}\left(\log\left(\rho'(N,s)\right)+\frac{1}{s}\log\left(\frac{2}{1-\theta}\right)\right)^{-1/2}+1\right)^{s}.$$

Then the particular choice

$$\delta := \frac{1}{\sqrt{2N}} \left(s \log(\rho'(N, s)) + \log \left(\frac{2}{1 - \theta} \right) \right)^{1/2} \tag{4.12}$$

and (4.8) imply the following version of Theorem 4.1.

Theorem 4.2. *Let the conditions of Theorem 4.1 hold. Then, with the choice of* $m_1, \ldots, m_s$ *and* $\rho'(N, s)$ *as in (4.10) and (4.11), we obtain with probability* $> \theta$

$$D_N^*(\mathcal{P}_s) \le \frac{\sqrt{2s}}{\sqrt{N}} \left(\log \left(\rho'(N, s) \right) + \frac{1}{s} \log \left(\frac{2}{1 - \theta} \right) \right)^{1/2} + D_N^*(\mathcal{P}_{s-1}). \tag{4.13}$$

One can prove the following corollary in a similar manner to Corollary 3.2.

Corollary 4.3. *For any* $N \ge 1$ *and the sequence* $(m_d)_{d \ge 1}$ *of positive integers defined as above, there exists a point set* $\mathcal{P} = \{\boldsymbol{p}_0, \ldots, \boldsymbol{p}_{N-1}\} \subset [0, 1)^{\mathbb{N}}$ *with*

$$\boldsymbol{p}_i = \left(\frac{2a_{i,1} + 1}{2m_1}, \frac{2a_{i,2} + 1}{2m_2}, \ldots \right),$$

where $a_{i,d} \in \{0, \ldots, m_d - 1\}$ *for all* $d \ge 1$ *and* $0 \le i < N$ *such that for all* $s \ge 1$ *for the point set* $\mathcal{P}_s$, *consisting of the projections of the points from* $\mathcal{P}$ *to the first* s *components, we have*

$$D_N^*(\mathcal{P}_s) \le \frac{\sqrt{2} \, s^{3/2}}{\sqrt{N}} \left(\log \left(\rho'(N, s) \right) + \frac{1}{s} \log(2) \right)^{1/2}. \tag{4.14}$$

We now derandomize the above construction, that is, we transform it into a deterministic algorithm that computes point sets having a discrepancy similar to the one which the randomized algorithm has with positive probability. In order to be able to use existing derandomizations, we reformulate our problem of "adding one dimension" as a rounding problem with hard constraints.

We briefly recall our problem: Let $\mathcal{P}_{s-1} = \{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{N-1}\} \subseteq \mathcal{G}_{s-1}$ with small star discrepancy $D_N^*(\mathcal{P}_{s-1})$. We aim at finding $X_0, \ldots, X_{N-1} \in G_s$ such that $\mathcal{P}_s = \{(\boldsymbol{y}_j, X_j) : 0 \le j < N\}$ has small discrepancy. For convenience, let us write $m = m_s$.

Our problem becomes a rounding problem as follows. Let $\mathcal{X}$ be the set of all families $(x_{jk})_{j=0,\ldots,N-1;k=1,\ldots,m}$ such that $x_{jk} \in [0, 1]$ for all j and k and $\sum_{k=1}^m x_{jk} = 1$ for all j. Let $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n$ be the enumeration of $\mathcal{G}_s^*$ chosen above. Define a linear function $A : \mathbb{R}^{Nm} \to \mathbb{R}^n$ by

$$A(\boldsymbol{x})_i = \sum_{j=0}^{N-1} \sum_{k=1}^m x_{jk} 1_{[\boldsymbol{0}, \boldsymbol{t}_i)}(\boldsymbol{y}_j, \hat{k}) \quad \text{for } \boldsymbol{x} = (x_{01}, \ldots, x_{(N-1)m}) \text{ and } i = 1, \ldots, n,$$

where we used the shorthand $\hat{k} = \frac{2k-1}{2m}$.

If $\boldsymbol{x} \in \mathcal{X}$ is integral put $X_j = \hat{k}_j$ for $j = 0, 1, \ldots, N - 1$, where k_j is the index for which $x_{jk_j} = 1$. Then $\mathcal{P}(\boldsymbol{x}) := \{(\boldsymbol{y}_j, X_j) : 0 \leq j < N\}$ is an N–point set in $[0, 1]^s$ and $A(\boldsymbol{x})_i$ is $|\mathcal{P}(\boldsymbol{x}) \cap [0, \boldsymbol{t}_i)|$. If in addition $\overline{\boldsymbol{x}} \in \mathcal{X}$ is defined by $\overline{x}_{jk} = \frac{1}{m}$ for all j, k, then we have the following. Let $B \in \mathcal{B}$ and $\boldsymbol{t}_i = \Phi(B)$. Then

$$A(\overline{\boldsymbol{x}})_i = \sum_{j=0}^{N-1} \sum_{k=1}^{m} \frac{1}{m} 1_{[\boldsymbol{0}, \boldsymbol{t}_i)}((\boldsymbol{y}_j, \hat{k})) = \sum_{j=0}^{N-1} 1_{[\boldsymbol{0}, \boldsymbol{t}_i')}(\boldsymbol{y}_j) \left(\sum_{k=1}^{m} \frac{1}{m} 1_{[0, t_{i,s})}(\hat{k}) \right)$$

$$= t_{i,s} \sum_{j=0}^{N-1} 1_{[\boldsymbol{0}, \boldsymbol{t}_i')}(\boldsymbol{y}_j) = t_{i,s} \sum_{j=0}^{N-1} 1_{B'}(\boldsymbol{y}_j).$$

Due to (4.5) we have

$$A(\boldsymbol{x})_i - A(\overline{\boldsymbol{x}})_i = \sum_{j=0}^{N-1} \Xi_i(\boldsymbol{y}_j, X_j), \qquad (4.15)$$

thus (4.6) implies $D_N^*(\mathcal{P}(\boldsymbol{x})) \leq \frac{1}{N} \|A(\boldsymbol{x}) - A(\overline{\boldsymbol{x}})\|_\infty + D_N^*(\mathcal{P}_{s-1}) + \frac{1}{2m}$. Hence low-discrepancy point-sets $\mathcal{P}(\boldsymbol{x})$ correspond to roundings $\boldsymbol{x} \in \mathcal{X}$ of $\overline{\boldsymbol{x}}$ with small rounding error $\|A(\boldsymbol{x}) - A(\overline{\boldsymbol{x}})\|_\infty$.

Generating and derandomizing randomized roundings satisfying certain equalities ("hard constraints"; here $\sum_{k=1}^{m} x_{jk} = 1$ for all j) without violation is highly non-trivial as recent results show, see e.g. [2, 3, 6]. Fortunately, in the case that the hard constraints require variable-disjoint sums to equal one, the classical method of Raghavan can be used. We briefly outline this method.

Let us assume that we have an arbitrary $\overline{\boldsymbol{x}} \in \mathcal{X}$, which we want to round to an integral $\boldsymbol{x} \in \mathcal{X}$ such that $|(A(\boldsymbol{x} - \overline{\boldsymbol{x}}))_i|$ is small for all i. The randomized construction would be to choose for each j independently a $k_j \in [m]$ at random such that $\mathbb{P}[k_j = k] = \overline{x}_{jk}$ for all j, k. Then for all j we define random variables $X_{jk_j} = 1$ and $X_{jk} = 0$, $k \neq k_j$. Note that by construction any outcome of X lies in $\mathcal{X}$. Let $\lambda \in \mathbb{R}$ such that $P := \sum_i \mathbb{P}[|(A(X - \overline{\boldsymbol{x}}))_i| \geq \lambda] < 1$ ("small initial failure probability"). Then there exists an $\boldsymbol{x} \in \mathcal{X}$ such that $|(A(\boldsymbol{x} - \overline{\boldsymbol{x}}))_i| \leq \lambda$ for all i.

We can actually compute such roundings $\boldsymbol{x}$ by derandomizing the probabilistic construction above. Let us confine ourselves to the special case $\overline{x}_{jk} = 1/m$ for all j, k. Due to (4.15) and (4.7) the initial failure probability is at most $P \leq 2n \exp(-2\lambda^2/N)$, which is smaller than one if $\lambda = \delta N$, δ chosen as in (4.12).

For $k = 1, \ldots, m$, let e_k denote the kth m–dimensional unit vector and consider the conditional probability $P_k := \sum_i \mathbb{P}[|(A(X - \overline{\boldsymbol{x}}))_i| \geq \lambda | (X_{01}, \ldots, X_{0m}) = e_k]$. Since $P = \sum_{k=1}^{m} \frac{1}{m} P_k$, there is a $1 \leq k_0^* \leq m$ such that $P_{k_0^*} \leq P < 1$ ("decreasing failure probability"). Next, let $P_{k_0^* k} := \sum_i \mathbb{P}[|(A(X - \overline{\boldsymbol{x}}))_i| \geq \lambda | (X_{01}, \ldots, X_{0m}) = e_{k_0^*}, (X_{11}, \ldots, X_{1m}) = e_k]$. Again, $P_{k_0^*} = \sum_{k=1}^{m} \frac{1}{m} P_{k_0^* k}$, and there is a $1 \leq k_1^* \leq m$ such that $P_{k_0^* k_1^*} \leq P_{k_0^*} < 1$. Proceeding like this we end up with $k_0^*, \ldots, k_{N-1}^*$ such

that $P_{k_0^*,\ldots,k_{N-1}^*} := \mathbb{P}[|(A(X - \overline{x}))_i| \geq \lambda | \forall \, 0 \leq j < N : (X_{j1}, \ldots, X_{jm}) = e_{k_j^*}] < 1$. Since $P_{k_0^*,\ldots,k_{N-1}^*}$ involves no randomness (all variables are bound in the conditional statement), we actually have $P_{k_0^*,\ldots,k_{N-1}^*} = 0$. We define $\boldsymbol{x}$ as follows: For each $1 \leq j < N$, we set $x_{jk_j^*} := 1$ and $x_{jk} := 0$ for all other k. This yields an integral vector $\boldsymbol{x} \in \mathcal{X}$ such that $|(A(\boldsymbol{x} - \overline{\boldsymbol{x}}))_i| \leq \lambda$ for all i.

The only problem with the above derandomization is that we usually cannot compute the conditional probabilities $P_{k_0^* k_1^* \ldots}$ in time polynomially bounded in N, m and n. However, it would suffice if we can compute (in polynomial time) upper bounds $U_{k_0^* k_1^* \ldots}$ for the exact conditional probabilities $P_{k_0^* k_1^* \ldots}$ such that the following key properties are maintained:

- Small initial (estimated) failure probability: $U < 1$.
- Decreasing (estimated) failure probability: For all $0 \leq \ell < N$ and $k_0^*, \ldots, k_{\ell-1}^* \in \{1, \ldots, m\}$ there is a $1 \leq k \leq m$ such that $U_{k_0^* k_1^* \ldots k_{\ell-1}^* k} \leq U_{k_0^* k_1^* \ldots k_{\ell-1}^*}$.

The quantities $U_{k_0^* k_1^* \ldots}$ are called *pessimistic estimators* for the conditional probabilities $P_{k_0^* k_1^* \ldots}$. This notion was introduced by Raghavan [20], who also showed that such pessimistic estimators exist for the conditional probabilities that occur in our derandomization. They are the sum of $2n$ expressions, two for each $1 \leq i \leq n$, estimating the probability that the box $[0, \boldsymbol{t}_i)$ receives too many or too few points. Both cases lead to similar expressions. Hence we sketch only the one for the case of too many points. Note that we do not want give a precise description of how to implement the derandomization, but only prove bounds on the amortized time needed to compute the estimators.

The probability that the box $[0, \boldsymbol{t}_i)$ receives too many points is at most

$$\exp(-c_i) \prod_{j=0}^{N-1} \left(\sum_{k=1}^{m} \widetilde{x}_{jk} \exp(\mathbf{1}_{[0,\boldsymbol{t}_i)}(\boldsymbol{y}_j, \hat{k}) c_i') \right),$$

where c_i and c_i' are suitable constants. Here, $\widetilde{x}_{jk}$ shall always denote the expected value of the random variable X_{jk}, if it has not been rounded, and the outcome of the rounding thereafter. This is the reason why we can compute the pessimistic estimator efficiently using the previous computation: When determining the rounded value for some $(x_{j\cdot})$, we only need to replace the terms involving these variables with the m possible choices for $(x_{j\cdot})$. Hence this can be done quite efficiently in time $O(nm)$. Thus both computing the initial value of the estimator and computing *all* subsequent values take time $O(nNm)$.

Note that for the final rounding $\boldsymbol{x}$, the pessimistic estimator implicitly tells us the number of points in the box $[0, \boldsymbol{t}_i)$: The expression $(\sum_{k=1}^{m} \widetilde{x}_{jk} \exp(\mathbf{1}_{[0,\boldsymbol{t}_i)}(\boldsymbol{y}_j, \hat{k}) c_i')$ is $\exp(c_i')$, if the jth point is in $[0, \boldsymbol{t}_i)$, and one, if not. Hence we can extract the number of points in each box easily from the computations so far, and thus also compute the precise discrepancy easily (by computing for each box the deviation of the actual and the aimed at number of points).

The constants c_i and c_i' depend only on the aimed at rounding error λ (in particular, they do not change during the algorithm). Also, the c_i, c_i' are such that the resulting exponential functions can be computed in the RAM model. As a consequence – since we chose all components of $\overline{x}$ to be rational numbers – we can compute the initial value of the pessimistic estimator in time $O(nNm)$ in the RAM model.

Unfortunately, Raghavan's pessimistic estimators do not admit the Hoeffding bound given in (2.1), but rather one that, in the setting of (2.1) and $u_i = 0$, $v_i = 1$, implies that

$$\mathbb{P}\left[|A_1 + \cdots + A_r| \geq r\eta\right] \leq 2 \cdot e^{-(1/3)r\eta^2}. \tag{4.16}$$

In consequence, to achieve that the initial estimated failure probability U is less than one, we have to choose the δ in Equations (4.7) to (4.9) to be $\sqrt{6}$ times larger than there. Choosing $m_1 \geq \cdots \geq m_s = m$ as in (4.10) and $\theta = 1/2$ in the choice of δ (any small positive constant is fine), this leads to a discrepancy bound analoguous to (4.13) of

$$D_N^*(\mathcal{P}_s) \leq \left(\sqrt{3} + \frac{1}{\sqrt{2}}\right) \frac{\sqrt{s}}{\sqrt{N}} \left(\log\left(\rho'(N, s)\right) + \frac{1}{s}\log(4)\right)^{1/2} + D_N^*(\mathcal{P}_{s-1}). \tag{4.17}$$

Note that this estimate is certainly trivial for $s \geq N/3$ (also, in this case we have $m = m_s = 1$, i.e., our "random experiment" is completely deterministic and $X_0 = \cdots = X_{N-1} = 1/2$). Hence in the following run-time estimate, we may assume $s \leq N/3$. Then, the run-time of our derandomized algorithm is

$$O(nNm) = O\left(\frac{c^s N^{\frac{s+3}{2}}}{s^{\frac{s}{2}+\frac{3}{4}} \log\left(\frac{N}{s}\right)^{\frac{s+1}{2}}}\right),$$

where c is some constant independent of N and s.

We summarize the discussion above in following theorem.

Theorem 4.4. *Let $s, N \in \mathbb{N}$. Let $\theta = 1/2$, and $m_1, \ldots, m_s$ as in (4.10). Let $\mathcal{P}_{s-1} = \{y_0, \ldots, y_{N-1}\}$ be a subset of $\mathcal{G}_{s-1}$. Then there is a deterministic algorithm that computes in time*

$$O(nNM) = O\left(\frac{c^s N^{\frac{s+3}{2}}}{s^{\frac{s}{2}+\frac{3}{4}} \log\left(\frac{N}{s}\right)^{\frac{s+1}{2}}}\right),$$

c a constant independent of N and s, $X_0, \ldots, X_{N-1} \in G_s$ such that the point set $\mathcal{P}_s = \{(y_0, X_0), \ldots, (y_{N-1}, X_{N-1})\}$ satisfies

$$D_N^*(\mathcal{P}_s) \leq \left(\sqrt{3} + \frac{1}{\sqrt{2}}\right) \frac{\sqrt{s}}{\sqrt{N}} \left(\log\left(\rho'(N, s)\right) + \frac{1}{s}\log(4)\right)^{1/2} + D_N^*(\mathcal{P}_{s-1}).$$

Obviously, we can use the derandomized algorithm from Theorem 4.4 to construct point sets $\mathcal{P}_s$ satisfying the discrepancy estimate (4.14) (multiplied by the factor $(\sqrt{3} + \frac{1}{\sqrt{2}})/\sqrt{2}$) component by component.

5. Conclusion

We presented a deterministic algorithm that generates in time

$$
O \left(\frac{c^s N^{\frac{s+3}{2}}}{s^{\frac{s}{2} - \frac{1}{4}} \log \left(\frac{N}{s} \right)^{\frac{s+1}{2}}} \right)
$$

an N-point set $\mathcal{P}_s \subset [0,1)^s$ satisfying

$$
D_N^*(\mathcal{P}_s) \leq O \left(\frac{s^{3/2}}{N^{1/2}} \left(\log \left(1 + \frac{N}{s} \right) \right)^{1/2} \right). \tag{5.1}
$$

In [4, 5] deterministic algorithms were presented that construct N-point sets $\mathcal{P}_s \subset [0,1)^s$ satisfying

$$
D_N^*(\mathcal{P}_s) \leq O \left(\frac{s^{1/2}}{N^{1/2}} \left(\log \left(1 + N \right) \right)^{1/2} \right) \tag{5.2}
$$

in run-time $O(C^s N^{s+2} \log(s)^s \log(N)^{1-s})$, C a constant, and $O(s \log(sN)(\sigma N)^s)$, $\sigma = \sigma(s) = \Theta(\log(s)^2 / s \log \log(s))$, respectively.

The comparison of the algorithms shows that the discrepancy guaranteed by our new algorithm has a worse dependence on s than the discrepancy guaranteed by the other two algorithms. Conversely, the run-time of our new algorithm improves considerably on the other two algorithms, especially with regard to the dependence on the number of points N. Another advantage is that the new algorithm calculates along the way the exact number of points in each box from a distinguished set of half-open boxes. This allows to easily compute the precise discrepancy of the output set $\mathcal{P}_s$.

We close this paper by relating the problem of constructing low-discrepancy sets of small size via derandomization to the problem of approximating the discrepancy of a given set.

Instead of trying to derandomize the random experiment to construct low-discrepancy sets, one may think of a semi-construction by performing a random experiment, calculating the actual discrepancy of the received set, and accept it if bounds like (5.1) or (5.2) are satisfied or start a new random experiment otherwise. Large deviation bounds like Hoeffding's inequality guarantee that with high probability we only have to perform the random experiment a few times to end up with a low-discrepancy point set. Besides the need of (true?) random bits, this overlooks the difficulty of calculating (or approximating) the star discrepancy of a given set.

Indeed, all algorithms that have been presented for this problem so far have a run-time exponential in s or no run-time guarantee at all, see e.g. [7, 21, 22, 23] and the literature mentioned therein. In fact, it has been shown recently in [23] that the decision problem whether an arbitrary point set has discrepancy smaller than ε is (if suitably posed) NP-complete. This indicates that it may be not possible to perform

semi-constructions as described above in polynomial time, as long as the possible output sets do not exhibit a special structure that makes the approximation of their star discrepancy particularly simple.

In the light of these results it is not too surprising that the run-times of the derandomized algorithms in this paper and in [4, 5] are exponentially in s, since we cannot expect to do the (deterministic) derandomized construction with less effort than the (probabilistic) semi-construction.

References

1. J. Dick, *A note on the existence of sequences with small star discrepancy*, J. Complexity 23 (2007), pp. 649–652.

2. B. Doerr, *Generating randomized roundings with cardinality constraints and derandomizations*, in B. Durand and W. Thomas (eds), Proceedings of the 23rd Annual Symposium on Theoretical Aspects of Computer Science (STACS'06), Lecture Notes in Comput. Sci. 3884, pp. 571–583, Springer, Berlin, 2006.

3. B. Doerr, *Randomly rounding rationals with cardinality constraints and derandomizations*, in W. Thomas and P. Weil (eds), Proceedings of the 24rd Annual Symposium on Theoretical Aspects of Computer Science (STACS'07), Lecture Notes in Comput. Sci. 4393, pp. 441–452, Springer, Berlin, 2007.

4. B. Doerr and M. Gnewuch, *Construction of low-discrepancy point sets of small size by bracketing covers and dependent randomized rounding*, in A. Keller, S. Heinrich, H. Niederreiter (eds), Monte Carlo and Quasi-Monte Carlo Methods 2006, pp. 299–312, Springer, Berlin, 2008.

5. B. Doerr, M. Gnewuch and A. Srivastav, *Bounds and constructions for the star discrepancy via δ-covers*, J. Complexity 21 (2005), pp. 691–709.

6. R. Gandhi, S. Khuller, S. Parthasarathy and A. Srinivasan, *Dependent rounding and its applications to approximation algorithms*, J. ACM 53 (2006), pp. 324–360.

7. M. Gnewuch, *Bracketing numbers for axis-parallel boxes and applications to geometric discrepancy*, J. Complexity 24 (2008), pp. 154–172.

8. M. Gnewuch, *Construction of minimal bracketing covers for rectangles*, Berichtsreihe des Mathematischen Seminars der Universität Kiel, Report 07-14, Kiel, 2007.

9. S. Heinrich, *Some open problems concerning the star discrepancy*, J. Complexity 19 (2003), pp. 416–419.

10. S. Heinrich, E. Novak, G. W. Wasilkowski and H. Woźniakowski, *The inverse of the star discrepancy depends linearly on the dimension*, Acta Arith. 96 (2001), pp. 279–302.

11. A. Hinrichs, *Covering numbers, Vapnik–Červonenkis classes and bounds for the star-discrepancy*, J. Complexity 20 (2004), pp. 477–483.

12. E. Hlawka, *Funktionen von beschränkter Variation in der Theorie der Gleichverteilung*, Ann. Mat. Pura Appl. 54 (1961), pp. 325–333.

13. W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. 58 (1963), pp. 13–30.

14. P. Kritzer, *Improved upper bounds on the star disrepancy of (t, m, s)-nets and (t, s)-sequences*, J. Complexity 22 (2006), pp. 336–347.

15. L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.

16. H. Niederreiter, *Point sets and sequences with small discrepancy*, Monatsh. Math. 104 (1987), pp. 273–337.

17. H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Series in Applied Mathematics 63, SIAM, Philadelphia, 1992.

18. V. V. Petrov, *Sums of Independent Random Variables*, Ergebnisse der Mathematik und ihrer Grenzgebiete 82, Springer, Berlin-Heidelberg-New York, 1975.

19. D. Pollard, *Convergence of Stochastic Processes*, Springer, Berlin, 1984.

20. P. Raghavan, *Probabilistic construction of deterministic algorithms: approximating packing integer programs*, J. Comput. System Sci. 37 (1988), 130–143.

21. E. Thiémard, *An algorithm to compute bounds for the star discrepancy*, J. Complexity 17 (2001), 850–880.

22. P. Winker and K. T. Fang, *Application of threshold-accepting to the evaluation of the discrepancy of a set of points*, SIAM J. Numer. Anal 34 (1997), pp. 2028–2042.

23. C. Winzen, *Approximative Berechnung der Sterndiskrepanz*, Diploma Thesis, Department of Computer Science, Christian-Albrechts University Kiel, Kiel, 2007.

Author information

Benjamin Doerr, Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany.
Email: doerr@mpi-sb.mpg.de

Michael Gnewuch, Department of Computer Science, Christian-Albrechts-Universität Kiel, Christian-Albrechts-Platz 4, 24098 Kiel, Germany.
Email: mig@informatik.uni-kiel.de

Peter Kritzer, Department of Mathematics, University of Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria, and School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, Australia.
Email: p.kritzer@unsw.edu.au

Friedrich Pillichshammer, Institut für Finanzmathematik, Universität Linz, Altenbergerstr. 69, 4040 Linz, Austria.
Email: friedrich.pillichshammer@jku.at