Construction of Low-Discrepancy Point Sets of Small Size by Bracketing Covers and Dependent Randomized Rounding

Benjamin Doerr[1] and Michael Gnewuch[2]

[1] Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany
[2] Institut für Informatik, Christian-Albrechts-Universität zu Kiel, Christian-Albrechts-Platz 4, 24098 Kiel, Germany

In memory of our friend, colleague and former fellow student Manfred Schocker

Summary. We provide a deterministic algorithm that constructs small point sets exhibiting a low star discrepancy. The algorithm is based on bracketing and on recent results on randomized roundings respecting hard constraints. It is structurally much simpler than the previous algorithm presented for this problem in [B. Doerr, M. Gnewuch, A. Srivastav. Bounds and constructions for the star discrepancy via δ-covers. J. Complexity, 21:691–709, 2005]. Besides leading to better theoretical run time bounds, our approach also can be implemented with reasonable effort.

1 Introduction

The L^∞*-star discrepancy* of an n-point set T in the d-dimensional unit cube $[0,1]^d$ is given by

$$d_\infty^*(T) := \sup_{x\in[0,1]^d} \left| \frac{1}{n} |T \cap [0,x[| - \mathrm{vol}([0,x[) \right| ,$$

where $[0,x[$ is the d-dimensional anchored half-open box $[0,x_1[\times \ldots \times [0,x_d[$. Here, as in the whole article, the cardinality of a finite set S is denoted by $|S|$. The smallest possible discrepancy of any n-point configuration in $[0,1]^d$ is

$$d_\infty^*(n,d) := \inf_{T\subset[0,1]^d\,;\,|T|=n} d_\infty^*(T) .$$

The *inverse of the star discrepancy* is given by

$$n_\infty^*(\varepsilon,d) := \min\{n \in \mathbb{N} \,|\, d_\infty^*(n,d) \le \varepsilon\} .$$

The star discrepancy is related to the worst case error of multivariate integration of a certain class of functions by the Koksma-Hlawka inequality (see, e.g., [DT97, HSW04, Nie92]). The inequality shows that points with small discrepancy induce quasi-Monte Carlo algorithms with small worst case errors. Since the number of sample points is roughly proportional to the costs of those algorithms, it is of interest to find n-point configurations with small discrepancy and n not too large. In particular, n should not depend exponentially on d.

For fixed dimension d the asymptotically best upper bounds for $d^*_\infty(n,d)$ that have been proved so far are of the form

$$d^*_\infty(n,d) \le C_d \ln(n)^{d-1} n^{-1}\,, \quad n \ge 2\,. \tag{1}$$

These bounds give us no helpful information for moderate values of n, since $\ln(n)^{d-1}n^{-1}$ is an increasing function for $n \le e^{d-1}$. Additionally, point configurations satisfying (1) will in general lead to constants C_d that depend critically on d. If we take, e.g., the Halton-Hammersley points, then C_d grows superexponentially in d (see, e.g., [Nie92]).

A bound more suitable for high-dimensional integration was established by Heinrich, Novak, Wasilkowski and Woźniakowski [HNWW01], who proved

$$d^*_\infty(n,d) \le c d^{1/2} n^{-1/2} \quad \text{and} \quad n^*_\infty(d,\varepsilon) \le \lceil c^2 d \varepsilon^{-2} \rceil\,, \tag{2}$$

where c does not depend on d, n or ε. Here the dependence of the inverse of the star discrepancy on d is optimal. This was also established in [HNWW01] by a lower bound for $n^*_\infty(d,\varepsilon)$, which was later improved by Hinrichs [Hin04] to $n^*_\infty(d,\varepsilon) \ge c_0 d \varepsilon^{-1}$ for $0 < \varepsilon < \varepsilon_0$, where $c_0, \varepsilon_0 > 0$ are constants. The proof of (2) is not constructive but probabilistic, and the proof approach does not provide an estimate for the value of c. (A. Hinrichs presented a more direct approach to prove (2) with $c = 10$ at the Dagstuhl Seminar 04401 "Algorithms and Complexity for Continuous Problems" in 2004.)

In the same paper the authors proved a slightly weaker bound with an explicitly known small constant k:

$$d^*_\infty(n,d) \le k d^{1/2} n^{-1/2} \big(\ln(d) + \ln(n)\big)^{1/2}\,. \tag{3}$$

The proof is again probabilistic and uses Hoeffding's inequality. For the sake of explicit constants the proof technique has been adapted in subsequent papers on high-dimensional integration of certain function classes [HSW04, Mha04]. In [DGS05] Srivastav and the authors were able to improve (3) to

$$d^*_\infty(n,d) \le k' d^{1/2} n^{-1/2} \ln(n)^{1/2}\,, \tag{4}$$

where k' is smaller than k. Of course the estimate (4) is asymptotically not as good as (2). But the constant k' is small—essentially we have $k' = \sqrt{2}$. If, e.g., $c = 10$, then (4) is superior to (2) for all n that are roughly smaller than e^{50}, i.e., for all values of n of practical interest. By derandomizing the probabilistic argument used in the proof of the inequality (4), the authors and Srivastav additionally gave a deterministic algorithm constructing point sets satisfying (4). The algorithm is based on a quite general derandomization approach of Srivastav and Stangier [SS96] and essentially a point-by-point construction using the method of conditional probabilities and so-called pessimistic estimators.

Our Results

In this paper, we use a novel approach to randomized rounding presented in [Doe06]. Contrary to the classical one, it allows to generate randomized roundings that respect certain hard constraints. This enables us to use a construction that needs significantly fewer random variables, which in turn speeds up the randomized construction.

A second speed-up and considerable simplification from the implementation point of view stems from the fact that the general approach in [Doe06] may be derandomized via the more restricted approach of Raghavan [Rag88]. This runs in time $O(mn)$, where n is the number of (random) variables and m the number of constraints.

It thus avoids the general, but more costly solution by Srivastav and Stangier [SS96]. The latter was a break-through from the theoretical point of view as it showed that randomized rounding for arbitrary linear constraints can be derandomized. From the practical point of view, it suffers from a higher run-time of $O(mn^2 \log(mn))$ and its extremely high technical demands. To the best of our knowledge, the algorithm implicit in the 30 pages proof has never been implemented.

We show the following result. For a given $n \in \mathbb{N}$ the algorithm computes an n-point set T with discrepancy

$$d_\infty^*(T) \leq (4 + \sqrt{3})\sqrt{n^{-1}\left(\tfrac{1}{2} d \ln(\sigma n) + \ln 2\right)} + 2^{-d\ln(dn)-1} n^{-1}$$

in time $O(d(\sigma n)^d \log(dn))$. Here $\sigma = \sigma(d)$ is less than one and converges to zero as d tends to infinity. In [DGS05] the running time for constructing an n-point set with the same discrepancy order was $O(C^d n^{d+2} \log(d)^d / \log(n)^{d-1})$, C some constant. That the running times of our deterministic algorithms are exponential in d may not be too surprising. Already any deterministic algorithm known so far that approximates the L^∞-star discrepancy of arbitrary given n-point sets has running time exponential in d (see [Thi96], the literature mentioned therein, and the discussion in [Gne06]). A comparison with other deterministic algorithms for constructing low-discrepancy sets of small size can be found in [DGS05].

Let us stress that the main advance in this paper is providing a simple solution (even though we are also faster than the previous one). We feel that our solution can be implemented with reasonable effort (and our future research will include this implementation project). Since often the quality of computed solutions is much better than what is guaranteed by theoretical worst-case error bounds, our solution presented in this paper opens an interesting line of research.

2 Randomized Construction

We start by introducing some useful notation: For arbitrary $n \in \mathbb{N}$ put $[n] := \{1, \ldots, n\}$. If $x, y \in [0,1]^d$, we write $x \le y$ if $x_i \le y_i$ holds for all $i \in [d]$. We write $[x, y] = \prod_{i \in [d]} [x_i, y_i]$ and use corresponding notation for open and half-open intervals. For a point $x \in [0,1]^d$ we denote by V_x the volume of the box $[0, x]$. Similarly, we denote the volume of a Lebesgue measurable subset S of $[0,1]^d$ by V_S.

2.1 Grids and Covers

Let $0 = q_0 < q_1 < \ldots < q_k = 1$ and $G := \{q_i \,|\, 1 \le i \le k\}^d$. G is a (not necessarily equidistant) grid in the d-dimensional unit cube $[0,1]^d$. Let $\delta = \delta(G)$ be the smallest real number such that for all $y \in [0,1]^d$ there are $x, z \in G \cup \{0\}$ with $x \le y \le z$ and $V_z - V_x \le \delta$. In the language of [DGS05] δ is minimal such that G is a δ–cover. Let us restate the definition from [DGS05]:

A finite set $\Gamma \subset [0,1]^d$ is a *δ-cover* of $[0,1]^d$ if for every $y \in [0,1]^d$ there exist $x, z \in \Gamma \cup \{0\}$ with $V_z - V_x \le \delta$ and $x \le y \le z$. Essentially the same concept is known in the literature of empirical processes as *bracketing*, see also [Gne06].

The helpfulness of δ-covers in discrepancy theory lies in the fact that one can use them to discretize discrepancy while controlling the discretization error:

Lemma 1. *Let Γ be a δ-cover of $[0,1]^d$. Then for all n-point sets $T \subset [0,1]^d$*

$$d^*_\infty(T) \le d^*_\Gamma(T) + \delta\,, \quad \textit{where} \quad d^*_\Gamma(T) := \max_{x \in \Gamma} \left| \frac{1}{n} |T \cap [0, x[| - V_x \right| . \qquad (5)$$

The proof is straightforward and can, e.g., be found in [DGS05].

Let now $\mathcal{I} := \{[q_{i-1}, q_i[\,|\, 1 \le i \le k\}$ and $\mathcal{B} := \{\prod_{i=1}^d I_i \,|\, I_1, \ldots, I_d \in \mathcal{I}\}$. Note that $\mathcal{B}$ is a partition of $[0,1[^d$ into axis-parallel boxes with upper right corners in G. Let $\mathcal{C}_0 := \{[0, g[\,|\, g \in G\}$. $\mathcal{C}_0$ is a subset of the set $\mathcal{C}$ of all axis-parallel boxes that are anchored in 0 (these boxes are sometimes called *corners*). If $g \in G$, let $B(g)$ be the uniquely determined $B \in \mathcal{B}$ and $C(g)$ the uniquely determined $C \in \mathcal{C}_0$ whose upper right corners are g. Furthermore, let $\mathcal{B}(g) := \{B \in \mathcal{B} \,|\, B \subseteq [0, g[\}$. If $B = B(g)$ or $C = C(g)$, we denote $\mathcal{B}(g)$ also by $\mathcal{B}(B)$ or $\mathcal{B}(C)$ respectively. Note that $(G, \le)$ is a partially ordered set and that via the identification of elements from $\mathcal{B}$ and $\mathcal{C}_0$ with their upper right corners we get induced partial orderings on $\mathcal{B}$ and $\mathcal{C}_0$ respectively. Let us denote all these orderings simply by $\le$. Finally, denote for a given $B \in \mathcal{B}$ the smallest set in $\mathcal{C}_0$ that contains B by $C(B)$. That is, B and $C(B)$ share the same upper right corner.

2.2 Reducing the Binary Length

Our aim is to construct n points in the unit cube exhibiting a fairly good discrepancy. We proceed as follows. For $B \in \mathcal{B}$, let $x_B := n \operatorname{vol}(B)$ be the fair number of points to lie in B.

We first round the x_B, $B \in \mathcal{B}$, to non-negative numbers having a finite binary expansion. Having numbers with finite binary expansion is necessary in the subsequent rounding step, but also carries the advantages of allowing (from this point on) efficient and exact computations.

Using a simple rounding approach, we could obtain $(\tilde{x}_B)$ such that $2^\ell \tilde{x}_B \in \mathbb{Z}$, $|x_B - \tilde{x}_B| < 2^{-\ell}$ and $\sum_{B \in \mathcal{B}} \tilde{x}_B = \sum_{B \in \mathcal{B}} x_B = n$. This yields a rounding error of $|\sum_{B \in \mathcal{B}(C)} (x_B - \tilde{x}_B)| < 2^{-\ell} |\mathcal{B}(C)|$ in all corners $C \in \mathcal{C}_0$.

However, we can achieve much smaller rounding errors by using a very recent result of Güntürk, Yılmaz and the first author [DGY06]. Let us remark first that there are higher-dimensional matrices $A \in [0,1]^{[k]^d}$ such that for any roundings $B \in \{0,1\}^{[k]^d}$ there is an $x \in [k]^d$ such that the rounding error

$$\left| \sum_{i_1=1}^{x_1} \cdots \sum_{i_d=1}^{x_d} (a_{i_1,\ldots,i_d} - b_{i_1,\ldots,i_d}) \right|$$

is of order $\Omega((\log k)^{(d-1)/2})$. Hence if we round the x_B to multiples of $2^{-\ell}$, we may get rounding errors $|\sum_{B \in \mathcal{B}(C)} (x_B - \tilde{x}_B)|$ of order $\Omega(2^{-\ell} (\log k)^{(d-1)/2})$. This follows from a result of Beck [Bec81], which in turn relies heavily on lower bounds for geometric discrepancies (Roth [Rot64], Schmidt [Sch72]). We refer to [Doe07] for a more extensive discussion of these connections.

The surprising result of [DGY06] is that by allowing larger deviations in the variables we can guarantee much better errors in the subarrays. In particular, we are able to remove any dependence on the grid size k. To ease reading, we reformulate and prove their result in our language.

Lemma 2. *Let $\ell \in \mathbb{N}$. There is a simple $O(|\mathcal{B}|)$ time algorithm computing $(\tilde{x}_B)$ such that*

(i) $2^\ell \tilde{x}_B \in \mathbb{Z}$ *for all* $B \in \mathcal{B}$*;*
(ii) $|x_B - \tilde{x}_B| \le 2^{-\ell-1+d}$ *for all* $B \in \mathcal{B}$*;*
(iii) $\sum_{B \in \mathcal{B}} \tilde{x}_B = \sum_{B \in \mathcal{B}} x_B = n$*;*
(iv) $|\sum_{B \in \mathcal{B}(C)} (x_B - \tilde{x}_B)| \le 2^{-\ell-1}$ *for all* $C \in \mathcal{C}_0$*.*

To assure that all quantities $\tilde{x}_B$, $B \in \mathcal{B}$, are non-negative, we will later choose $\ell \ge d - 1 + \log_2(\max_{B \in \mathcal{B}} x_B^{-1})$.

Proof. We may sort $\mathcal{B}$ in a way that B_1 is prior to B_2 if $B_1 \le B_2$. In this order, we traverse (x_B) and choose the uniquely determined $\tilde{x}_B$ satisfying $2^\ell \tilde{x}_B \in \mathbb{Z}$ and

$$-2^{-\ell-1} < \sum_{B' \in \mathcal{B}(B)} (\tilde{x}_{B'} - x_{B'}) \le 2^{-\ell-1} \,. \tag{6}$$

We may express any B by unions and differences of at most 2^d corners $C(\tilde{B})$. Hence we can write $\tilde{x}_B - x_B$ as sum and difference of at most 2^d terms $\sum_{B' \in \mathcal{B}(\tilde{B})} (\tilde{x}_{B'} - x_{B'})$, resulting in $|\tilde{x}_B - x_B| \le 2^{-\ell-1+d}$.

Finally, from $\sum_{B \in \mathcal{B}} \tilde{x}_B \in 2^{-\ell}\mathbb{Z}$, $\sum_{B \in \mathcal{B}} x_B = n \in \mathbb{Z}$ and $|\sum_{B \in \mathcal{B}} (x_B - \tilde{x}_B)| \le 2^{-\ell-1}$, we conclude $\sum_{B \in \mathcal{B}} \tilde{x}_B = \sum_{B \in \mathcal{B}} x_B = n$.

2.3 Randomized Rounding with Cardinality Constraint

We now randomly round $(\tilde{x}_B)$ to integers (y_B) and then choose our point set in a way that is has exactly y_B points in the box B. This rounding is done via a recent extension of the classical randomized rounding method due to Raghavan [Rag88]. We briefly review the basics.

Randomized Rounding

For a number r we write $\lfloor r \rfloor = \max\{z \in \mathbb{Z} \,|\, z \le r\}$, $\lceil r \rceil = \min\{z \in \mathbb{Z} \,|\, z \ge r\}$ and $\{r\} = r - \lfloor r \rfloor$. Let $\xi \in \mathbb{R}$. An integer-valued random variable y is called *randomized rounding of* ξ if

$$\Pr(y = \lfloor \xi \rfloor + 1) = \{\xi\},$$
$$\Pr(y = \lfloor \xi \rfloor) = 1 - \{\xi\}.$$

Since only the fractional part of ξ is relevant, we often may ignore the integer part and then have $\xi \in [0,1]$. In this case, a randomized rounding y of ξ satisfies

$$\Pr(y = 1) = \xi,$$
$$\Pr(y = 0) = 1 - \xi.$$

For $\xi \in \mathbb{R}^n$, we call $y = (y_1, \ldots, y_n)$ *randomized rounding of* ξ if y_j is a randomized rounding of ξ_j for all $j \in [n]$. We call y *independent randomized rounding* of ξ, if the y_i are mutually independent random variables.

Independent randomized rounding was introduced by Raghavan [Rag88] and since has found numerous application. It takes its strength from the fact that sums of independent random variables are strongly concentrated around their mean. This allows to bound the deviation of a weighted sum of the ξ_i from the corresponding sum of the y_i (this is done via so-called Chernoff bounds).

Independent randomized rounding can be derandomized. That is, one can transform the above sketched approach into a deterministic rounding algorithm (at the price of a slightly higher run-time) that guarantees large deviation bounds comparable to those that randomized rounding satisfy with high probability.

For our purposes, independent randomized rounding is not fully satisfactory since we would like to construct exactly n points. In other words, we prefer to have $\sum_{B\in\mathcal{B}} y_B = \sum_{B\in\mathcal{B}} \tilde{x}_B = n$ without any deviation. Fortunately, this can be achieved relatively easy with the randomized rounding method proposed by the first author in [Doe06]. It allows to generate randomized roundings that always fulfill constraints like $\sum_{B\in\mathcal{B}} y_B = \sum_{B\in\mathcal{B}} \tilde{x}_B = n$. In consequence, this is not independent randomized rounding. However, though not being independent, these roundings still satisfy Chernoff bounds and can be derandomized. The following makes this precise.

Theorem 1 ([Doe06]). *Let $\xi \in \mathbb{R}^N$ such that all ξ_i have binary length at most ℓ and $\sum_{i=1}^n \xi_i \in \mathbb{N}$. Then in time $O(\ell N)$ a randomized rounding y of ξ can be generated such that* $\Pr(\sum_{i=1}^N y_i = \sum_{i=1}^N \xi_i) = 1$ *and for all $a \in [0,1]^N$, $Y := \sum_{i=1}^N a_i y_i$, $\mu := E(Y) = \sum_{i=1}^N a_i \xi_i$ and all $\delta \in [0,1]$,*

$$\Pr(Y \geq (1+\delta)\mu) \leq \exp(-\tfrac{1}{3}\mu\delta^2),$$
$$\Pr(Y \leq (1-\delta)\mu) \leq \exp(-\tfrac{1}{2}\mu\delta^2).$$

The Chernoff bounds given above are not strongest possible. Bounds like $\Pr(Y \geq (1+\delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu$ would also hold, but are often not practical to work with. In fact, we will use an even simpler bound.

Lemma 3. *In the setting of Theorem 1, assume further that ξ is non-negative and $n := \sum_{i=1}^N \xi_i$. Then for all $\lambda \geq 0$, we have*

$$\Pr(|Y - \mu| \geq \lambda) \leq 2\exp(-\tfrac{1}{3}\lambda^2/n).$$

Proof. We may assume $\lambda \leq n$, as Y never exceeds n by non-negativity of ξ. Let $\xi_{N+1} = \lceil n - \mu \rceil$ and $a_{N+1} = (n-\mu)/\lceil n-\mu\rceil$. Note that since ξ_{N+1} is integral, any randomized rounding of $\xi_1, \ldots, \xi_{N+1}$ as in Theorem 1 yields a randomized rounding for $\xi_1, \ldots, \xi_N$ as in Theorem 1 (by just forgetting the $(N+1)$-st variable) and vice versa (by taking $y_{N+1} = \xi_{N+1}$ with probability one). Hence we need not to distinguish between the two.

Let $\tilde{Y} = \sum_{i=1}^{N+1} a_i y_i$ and $\tilde{\mu} = E(\tilde{Y})$. Note that by construction, $\tilde{\mu} = n$. Hence with $\delta = \lambda/n$, the first bound of Theorem 1 yields

$$\begin{aligned}\Pr(Y - \mu \geq \lambda) &= \Pr(\tilde{Y} - \tilde{\mu} \geq \lambda)\\ &= \Pr(\tilde{Y} \geq (1+\delta)\tilde{\mu})\\ &\leq \exp(-\tfrac{1}{3}n\delta^2) \;=\; \exp(-\tfrac{1}{3}\lambda^2/n).\end{aligned}$$

The second bound of Theorem 1 analogously yields $\Pr(-(Y-\mu) \geq \lambda) \leq \exp(-\frac{1}{3}\lambda^2/n)$. Both estimates give this lemma.

Construction of the Point Set

We use the theorem above to generate random variables (y_B) as randomized roundings of $(\tilde{x}_B)$. Since we required the $\tilde{x}_B$, $B \in \mathcal{B}$, to be non-negative, the y_B, $B \in \mathcal{B}$, are non-negative integers. Let T be an n-point set in the unit cube such that for all $B \in \mathcal{B}$, T contains exactly y_B points in B.

Lemma 4. *Let $C \in \mathcal{C}_0$. Then for all non-negative λ we have*

$$\Pr\left(\big||C \cap T| - nV_C\big| > \lambda + 2^{-\ell-1}\right) \leq 2\exp\left(-\frac{\lambda^2}{3n}\right).$$

Proof. By construction, we have $|C \cap T| = \sum_{B \in \mathcal{B}(C)} y_B$ and $nV_C = \sum_{B \in \mathcal{B}(C)} x_B$. From Lemma 2 we get

$$\big||C \cap T| - nV_C\big| = \left|\sum_{B \in \mathcal{B}(C)} (y_B - x_B)\right| \leq \left|\sum_{B \in \mathcal{B}(C)} (y_B - \tilde{x}_B)\right| + 2^{-\ell-1}.$$

Put $Y := \sum_{B \in \mathcal{B}(C)} y_B$. Since the $\tilde{x}_B$ are non-negative and $n = \sum_{B \in \mathcal{B}} \tilde{x}_B$, we get from Lemma 3

$$\Pr\left(\left|\sum_{B \in \mathcal{B}(C)} (y_B - \tilde{x}_B)\right| \geq \lambda\right) \leq 2\exp\left(-\frac{\lambda^2}{3n}\right).$$

Theorem 2. *Let T be as above. Let $\delta \in]0,1]$. For all $\theta \in [1, \infty[$ we have*

$$\Pr\left(d_\infty^*(T) > \sqrt{3n^{-1}\ln(2\theta|\mathcal{B}|)} + \delta + 2^{-\ell-1}n^{-1}\right) \leq \theta^{-1}. \tag{7}$$

Proof. By Lemma 1, we have $d_\infty^*(T) \leq d_G^*(T) + \delta$. (Of course, G should be a grid as in Subsection 2.1.) Choosing $\lambda = \sqrt{3n\ln(2\theta|\mathcal{B}|)}$, we deduce from Lemma 4 that

$$\Pr(\big||C \cap T| - nV_C\big| > \lambda + 2^{-\ell-1}) \leq (\theta|\mathcal{B}|)^{-1} \quad \text{for all } C \in \mathcal{C}_0.$$

Hence, since $|\mathcal{C}_0| = |\mathcal{B}|$,

$$\Pr(d_G^*(T) > (\lambda + 2^{-\ell-1})/n) \leq \sum_{C \in \mathcal{C}_0} \Pr\left(\big||C \cap T| - nV_C\big| > \lambda + 2^{-\ell-1}\right) \leq \theta^{-1}.$$

Choice of Parameters

Note that inequality (7) depends on the parameters θ, ℓ and δ (in particular, $|\mathcal{B}|$ depends on δ). In the following we make some reasonable choices for these parameters to get a version of inequality (7) that only depends on d and n.

Let $d \geq 2$. In [DGS05, Thm.2.3] a δ-cover in form of a non-equidistant grid $G = \{q_1, \ldots, q_k\}^d$ was constructed satisfying

$$k = \left\lceil \frac{d}{d-1} \frac{\ln(1-(1-\delta)^{1/d}) - \ln\delta}{\ln(1-\delta)} \right\rceil + 1 \leq \left\lceil \frac{d}{d-1} \frac{\ln d}{\delta} \right\rceil + 1.$$

The explicit construction goes as follows: Put $p_0 := 1$ and $p_1 := (1-\delta)^{1/d}$. If $p_i > \delta$, then define $p_{i+1} := (p_i - \delta)p_1^{1-d}$. If $p_{i+1} \leq \delta$, then put $\kappa(\delta, d) := i + 1$, otherwise proceed by calculating p_{i+2}. Then $k = \kappa(\delta, d) + 1$ and $q_{k-i} = p_i$.

For this grid G and $\delta \le 1/2$ we get

$$\ln|\mathcal{B}| = \ln|G| = d\ln k \le d(\ln\delta^{-1} + \ln\ln d + \ln 4)\,.$$

Choosing

$$\delta = \left(3n^{-1}(d\,(\ln\ln d + \ln 8) + \ln 4)\right)^{1/2} \tag{8}$$

leads to

$$\ln|\mathcal{B}| \le \frac{d}{2}\ln(\sigma n)\,, \quad \text{where} \quad \sigma = \sigma(d) := \frac{16(\ln d)^2}{3(d(\ln\ln d + \ln 8) + \ln 4)}\,. \tag{9}$$

An elementary analysis shows that σ takes its maximum in $d = 6$ and therefore $\max_{d\ge 2}\sigma(d) < 0.9862$. In the table below we listed some values of σ.

d	$\sigma(d)$	d	$\sigma(d)$
2	0.5324886424	20	0.7528969387
3	0.8141209699	30	0.6139386902
4	0.9308908286	40	0.5306094834
5	0.9754256341	50	0.4702720050
6	0.9861774970	60	0.4242704800
7	0.9802221264	70	0.3878425250
8	0.9657904472	80	0.3581541672
9	0.9471133088	90	0.3334088723
10	0.9264689299	100	0.3124089382
11	0.9051224430	360	0.1331152560
12	0.8837875877	1000	0.0634092061

Let us assume that for a given dimension $d \ge 2$ the number of points n is large enough to imply $\delta \le 1/2$. Then, for $\theta = 2$,

$$\sqrt{3n^{-1}\ln(2\theta|\mathcal{B}|)} + \delta \le 2\sqrt{3n^{-1}\left(\frac{d}{2}\ln(\sigma n) + \ln 4\right)}\,.$$

Now let us specify the choice of ℓ. Due to our choice of the δ-cover G we may assume that

$$\min_{B\in\mathcal{B}} V_B > \frac{\delta^d}{d^d}\,. \tag{10}$$

(It is easy to see that $q_k - q_{k-1} > \delta/d$ and that $q_i - q_{i-1}$, $i = 2,\dots,k$, is a strictly monotonic decreasing sequence. Nevertheless, (10) does not hold in the situation where q_1 is almost zero. In this case we substitute $q_1 := q_2/2$. It is easy to see that the new grid G is still a δ-cover. Then $q_1 > \delta/2$ and therefore $q_i - q_{i-1} > \delta/d$ for all $i = 1,\dots,k$.)

To guarantee that our quantities $\tilde{x}$, $B \in \mathcal{B}$, from Lemma 2 are non-negative, we choose ℓ such that $2^{-\ell+d-1} \le \delta^d/d^d$. According to our choice of δ in (8), the last inequality holds if

$$\ell \geq \left\lceil \frac{d}{2\ln 2} \ln\left(\frac{2}{3}\, d\, n\right)\right\rceil - 1\,.$$

For simplicity we choose

$$\ell = \lceil d \ln(d\, n) \rceil\,.$$

Our choices of G, δ, θ and ℓ result in the following corollary.

Corollary 1. *Let G, δ, θ and ℓ be chosen as above, and let $\sigma = \sigma(d) = \frac{16(\ln d)^2}{3(d(\ln\ln d + \ln 8) + \ln 4)}$ be as above. Then*

$$\Pr\left(d^*_\infty(T) > 2\sqrt{3n^{-1}\left(\tfrac{1}{2} d \ln(\sigma\, n) + \ln 4\right)} + 2^{-d\ln(dn)-1} n^{-1}\right) \leq \frac{1}{2}. \qquad (11)$$

Remark 1. Note that above we were using a non-equidistant grid as δ-cover. In [DGS05, Gne06], also δ-covers were constructed that had no grid structure (by a grid, we shall always mean a point set G in $[0,1]^d$ that can be written as $G = (G_0)^d$ for some $G_0 \subset [0,1]$). These δ-covers were superior in the sense that they needed fewer points. For the approach we use in this paper, however, they cannot be applied. The reason is that in Lemma 2 and 4, we heavily use the fact that corners (elements from $\mathcal{C}_0$) are the union of all boxes (elements from $\mathcal{B}$) which they have a non-trivial intersection with.

3 Derandomized Construction

The randomized roundings of Theorem 1 and hence the whole construction above can be derandomized. Combining Theorem 4 of [Doe06] with the simple derandomization without pessimistic estimators (this is derandomization (i) in Section 3.2 of [Doe06]) yields the following.

Theorem 3. *Let $A \in \{0,1\}^{m\times n}$. Let $\xi \in \mathbb{R}^n$ such that $\sum_{i=1}^n \xi_i \in \mathbb{Z}$ and $2^\ell \xi \in \mathbb{Z}^n$. Then a rounding y of ξ such that $\sum_{i=1}^n y_i = \sum_{i=1}^n \xi_i$ and*

$$\forall i \in [m] : |(A\xi)_i - (Ay)_i| \leq 13\sqrt{\max\{(A\xi)_i, \ln(4m)\}\ln(4m)}$$

can be computed in time $O(mn\ell)$.

The rounding errors we are interested in are all of the kind $\sum_{B\in\mathcal{B}(C)}(\tilde{x}_B - y_B)$ for some $C \in \mathcal{C}_0$. Hence the matrix encoding all these errors is an $|\mathcal{C}_0| \times |\mathcal{B}|$ matrix having entries 0 and 1 only. More precisely, we consider the matrix $A = (a_{C,B})_{C\in\mathcal{C}_0, B\in\mathcal{B}}$, where $a_{C,B} = 1$ if $B \subseteq C$ and $a_{C,B} = 0$ else. For each $C \in \mathcal{C}_0$ we have

$$(A\tilde{x})_C = \sum_{B\in\mathcal{B}(C)} \tilde{x}_B \leq \sum_{B\in\mathcal{B}} \tilde{x}_B = n\,.$$

Thus, if $n \geq \ln(4|\mathcal{C}_0|)$, we get from Theorem 3 the bound

$$|(A\tilde{x})_C - (Ay)_C| \le 13\sqrt{n \ln(4|\mathcal{C}_0|)}\,.$$

If $n \le \ln(4|\mathcal{C}_0|)$, this bound holds trivially, since always $|(A\tilde{x})_C - (Ay)_C| \le n$. Altogether we get the following theorem.

Theorem 4. *Let $n \in \mathbb{N}$ be given. There is a deterministic algorithm that*

(i) computes a point set $T \subseteq [0,1]^d$ that has exactly n points;
(ii) $d^*_\infty(T) \le 13\sqrt{n^{-1}\ln(4|\mathcal{C}_0|)} + \delta + 2^{-\ell-1}n^{-1}$;
(iii) has run time $O(\ell|\mathcal{B}||\mathcal{C}_0|)$.

We get the following corollary.

Corollary 2. *Let $n \in \mathbb{N}$ be given. Let G, δ and ℓ be as chosen in the last section. Furthermore, let σ be as defined in (9). There is a deterministic algorithm that*

(i) computes a point set $T \subseteq [0,1]^d$ that has exactly n points;
(ii) $d^*_\infty(T) \le (13+\sqrt{3})\sqrt{n^{-1}\left(\frac{d}{2}\ln(\sigma\, n) + \ln 4\right)} + 2^{-d\ln(d\,n)-1}n^{-1}$;
(iii) has run time $O(d\,\ln(d\,n)(\sigma\,n)^d)$.

3.1 Improvements on the Constants

In [Doe06], having good estimates for the constant in the error term (which e.g. yield the 13 in Theorem 3) was not too important. Here, this constant has roughly a quadratic influence on the size of the point set having a fixed discrepancy. We therefore discuss a simple improvement of the result in [Doe06].

We note that the simple derandomization for $\{0, \frac{1}{2}\}$ vectors and $A \in \{0,1\}^{m\times n}$ (this is derandomization (i) in Section 3.2 of [Doe06]) works as well for $A \in \{-1,0,1\}^{m\times n}$. This saves us from separating positive and negative entries of A as in the proof of Lemma 4 in [Doe06]. Consequently, the $\ln(4m)$ terms become $\ln(2m)$ and the constant $13 \ge f(2\sqrt{\frac{1}{2}})$ becomes $f(\sqrt{\frac{1}{2}}) \le 4$ (where f is defined as in [Doe06]).

Further improvements, in particular, for certain values of the variables involved, are definitely possible (cf. also the note after Theorem 1). We feel, however, that in this paper further technicalities would rather hide the main ideas. We thus decided not to follows such lines of research in this paper.

Future Work

We provided a deterministic algorithm to construct low-discrepancy sets of small size. This algorithm can be implemented with reasonable effort, and we will concentrate on this task in the near future. It would be interesting to test the quality of the resulting point sets T. Further "fine tuning" may improve the discrepancy of T. Notice, e.g., that so far we only distributed n points in boxes $B \in \mathcal{B}$, but we have not specified where to place them inside these boxes. Indeed, this has no influence on our given analysis. But in practise it may, e.g., lead to better results if one places the points in a more sophisticated way inside the box instead of just putting them into the lower left corner. A further investigation of this topic seems to be interesting.

References

[Bec81] Beck, J.: Balanced two-colorings of finite sets in the square. Combinatorica, **1**, 327–335 (1981)

[DGS05] Doerr, B., Gnewuch, M., Srivastav, A.: Bounds and constructions for the star discrepancy via δ-covers. J. Complexity, **21**, 691–709 (2005)

[DGY06] Doerr, B., Güntürk, S., Yılmaz, O.: Matrix Quasi-Rounding. Preprint (2006)

[Doe06] Doerr, B.: Generating randomized roundings with cardinality constraints and derandomizations. In: Durand, B., Thomas, W. (eds) Proceedings of the 23rd Annual Symposium on Theoretical Aspects of Computer Science (STACS'06), Lecture Notes in Comput. Sci., **3884**, 571–583 (2006)

[Doe07] Doerr, B.: Matrix approximation and Tusnády's problem. European J. Combin., **28**, 990–995 (2007)

[DT97] Drmota, M., Tichy, R.F.: Sequences, Discrepancies and Applications. Lecture Notes in Math., **1651**, Springer, Berlin Heidelberg New York (1997)

[Gne06] Gnewuch, M.: Bracketing numbers for d-dimensional boxes and applications to geometric discrepancy. Preprint (2006)

[Hin04] Hinrichs, A.: Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy. J. Complexity, **20**, 477–483 (2004)

[HNWW01] Heinrich, S., Novak, E., Wasilkowski, G.W., Woźniakowski, H.: The inverse of the star-discrepancy depends linearly on the dimension. Acta Arith., **96**, 279–302 (2001)

[HSW04] Hickernell, F.J., Sloan, I.H., Wasilkowski, G.W.: On tractability of weighted integration over bounded and unbounded regions in $\mathbb{R}^s$. Math. Comp., **73**, 1885–1901 (2004)

[Mha04] Mhaskar, H.N.: On the tractability of multivariate integration and approximation by neural networks. J. Complexity, **20**, 561–590 (2004)

[Nie92] Niederreiter, H.: Pseudo Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia (1992)

[Rag88] Raghavan, P.: Probabilistic construction of deterministic algorithms: Approximating packing integer programs. J. Comput. Syst. Sci., **37**, 130–143 (1988)

[Rot64] Roth, K.F.: Remark concerning integer sequences. Acta Arith., **9**, 257–260 (1964)

[Sch72] Schmidt, W.M.: On irregularities of distribution VII. Acta Arith., **21**, 45–50 (1972)

[SS96] Srivastav, A., Stangier, P.: Algorithmic Chernoff-Hoeffding inequalities in integer programming. Random Structures Algorithms, **8**, 27–58 (1996)

[Thi96] Thiémard, E.: An algorithm to compute bounds for the star discrepancy. J. Complexity, **17**, 850-880 (2001)