Probabilistic error bounds for the discrepancy of mixed sequences

Christoph Aistleitner[*], Markus Hofer[†]

Abstract

In many applications Monte Carlo (MC) sequences or Quasi-Monte Carlo (QMC) sequences are used for numerical integration. In moderate dimensions the QMC method typically yield better results, but its performance significantly falls off in quality if the dimension increases. One class of randomized QMC sequences, which try to combine the advantages of MC and QMC, are so-called mixed sequences, which are constructed by concatenating a d-dimensional QMC sequence and an $s-d$-dimensional MC sequence to obtain a sequence in dimension s. Ökten, Tuffin and Burago proved probabilistic asymptotic bounds for the discrepancy of mixed sequences, which were refined by Gnewuch. In this paper we use an interval partitioning technique to obtain improved probabilistic bounds for the discrepancy of mixed sequences. By comparing them with lower bounds we show that our results are almost optimal.

1 Introduction and statement of results

A common notion to measure the regularity of point distributions is the so-called star discrepancy. Roughly speaking, the star discrepancy compares the relative number of elements of a point set, which are contained in an axis-parallel box, to the volume of this box, and finally takes the maximal deviation over all possible boxes. The Quasi-Monte Carlo method for numerical integration is based on the fact that the difference of the integral of a function and the arithmetic mean of the function values at certain sampling points can be estimated by the product of the variation of this function and the discrepancy of the set of sampling points. Therefore point sets having a small star discrepancy can serve as a tool for numerical integration, a method which is frequently used for the evaluation of high-dimensional integrals in applied mathematics. Many constructions of low-discrepancy point sets only provide good bounds for the discrepancy if the number of points is large (in comparison with the dimension). This led to the development of the so-called randomized Quasi-Monte Carlo method, which tries to combine the advantages of the (deterministic) Quasi-Monte Carlo method and the advantages of the (random) Monte Carlo method. For an introduction to discrepancy theory and its applications in numerical mathematics we refer the reader to the books of Dick

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

[†]Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: markus.hofer@tugraz.at. Research supported by the Austrian Research Foundation (FWF), Project S9603-N23.

Mathematics Subject Classification: 11K38, 65C05, 65D32

Keywords: Monte Carlo, Quasi-Monte Carlo, discrepancy, hybrid sequences, mixed sequences

and Pillichshammer [4], Drmota and Tichy [5], Kuipers and Niederreiter [13] and Glasserman [6].

To formulate our results in a precise way we need some notation. We write $(x^{(1)}, \dots, x^{(s)})$ for the coordinates of a point $x \in [0,1]^s$. We write $x \leq y$ if $x^{(i)} \leq y^{(i)}$ for $1 \leq i \leq s$. We write 0 and 1 for the points $(0, \dots, 0)$ and $(1, \dots, 1)$ in $[0,1]^s$. For $a \in [0,1]^s$ we define an s-dimensional interval $[0,a]$ as the set $\{x \in [0,1]^s : 0 \leq x \leq a\}$ (which is an s-dimensional axis-parallel box).
Let $(x_1, \dots, x_N)$ be a sequence of points in the s-dimensional unit cube. The star discrepancy D_N^* of $(x_1, \dots, x_N)$ is defined as

$$D_N^*(x_1, \dots, x_N) = \sup_{a \in [0,1]^s} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[0,a]}(x_n) - \lambda([0,a]) \right|.$$

Here and in the sequel λ denotes the Lebesgue measure. For simplicity we write $D_N^*(x_n)$ instead of $D_N^*(x_1, \dots, x_N)$. If $(x_n)_{n \geq 1}$ is an infinite sequence, we write $D_N^*(x_n)$ for the discrepancy of the first N elements of $(x_n)_{n \geq 1}$.

The importance of discrepancy theory in numerical mathematics is based on the Koksma-Hlawka inequality, which states that for a sequence $(x_1, \dots, x_N)$ of points in $[0,1]^s$ and a function f having total variation $\operatorname{Var} f$ on $[0,1]^s$ (in the sense of Hardy and Krause)

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_0^1 f(x) \, dx \right| \leq D_N^*(x_n) \cdot \operatorname{Var} f.$$

There exist many constructions of so-called low-discrepancy sequences, i.e. sequences $(x_n)_{n \geq 1}$ for which

$$D_N^*(x_n) \ll (\log N)^s N^{-1} \qquad \text{as} \qquad N \to \infty \tag{1}$$

(this should be compared with a result of Roth [21] which states that every infinite sequence of points from $[0,1]^s$ has discrepancy $\gg (\log N)^{s/2} N^{-1}$ for infinitely many N; this has been slightly improved by Beck [2] and Bilyk, Lacey and Vagharshakyan [3], but the precise minimal asymptotic order of the discrepancy is still an open problem). Sequences of this type are only of practical use if the number of sampling points N is "large" in comparison with the dimension s; in particular the right-hand side of (1) is increasing for $N \leq e^s$. On the other hand, the so-called Monte Carlo method (which uses i.i.d. randomly generated points instead of deterministic points) gives an probabilistic bound of asymptotical order $N^{-1/2}$, independently of the dimension. This led to the development of randomized QMC integration schemes, which try to combine the advantages of (random) MC and (deterministic) QMC. There exist several methods for "randomizing" QMC rules; see for example Hickernell [11], Matoušek [15], Owen [19] and L'Ecuyer and Lemieux [14]. In this paper we consider s-dimensional sequences which are constructed by concatenating the coordinates of a d-dimensional QMC sequence and an $s-d$-dimensional MC sequence. Sequences of this type are called "mixed" sequences, and have been investigated e.g. by Spanier [22], Ökten [16][17] and Roşca [20].

Let $(q_n)_{n \geq 1}$ be a d-dimensional QMC sequence, and let $(X_n)_{n \geq 1}$ be a sequence of i.i.d. random variables having uniform distribution on $[0,1]^{s-d}$. We write $(x_n)_{n \geq 1}$ for the sequence which consists of the points $x_n = (q_n, X_n)$, i.e. $x_n = (q_n^{(1)}, \dots, q_n^{(d)}, X_n^{(1)}, \dots, X_n^{(s-d)})$ for $n \geq 1$. Ökten, Tuffin and Burago [18] showed that for such a sequence, under the additional assumption $D_N^*(q_n) \to 0$, for arbitrary $\varepsilon > 0$

$$\mathbb{P}\left(D_N^*(x_n) \leq D_N^*(q_n) + \varepsilon\right) \geq 1 - 2e^{-\varepsilon^2 N/2} \tag{2}$$

for sufficiently large N (in [18, Theorem 5] the exponent $-2\varepsilon^2 N$ appears, but as Gnewuch [9] remarks, the proof only gives $-\varepsilon^2 N/2$). Their paper contains no information on the size of the values of N for which (2) holds. Gnewuch [9] showed that

$$\mathbb{P}\left(D_N^*(x_n) \leq D_N^*(q_n) + \varepsilon\right) \geq 1 - 2\mathcal{N}(s, \varepsilon/2) e^{-\varepsilon^2 N/2}, \tag{3}$$

where $\mathcal{N}(d, \delta)$ is defined as the smallest number M for which there exists a set Γ of M points in $[0,1]^s$ such that for all $y \in [0,1]^s$ there exist $x, z \in \Gamma \cup \{0\}$ such that $x \leq y \leq z$ and $\lambda([0,z]) - \lambda([0,x]) \leq \delta$ (the set Γ is called a δ-*cover* of $[0,1]^s$, and the number $\mathcal{N}$ the *covering number*). By [7, Theorem 1.15]

$$\mathcal{N}(s, \delta) \leq (2e)^s (\delta^{-1} + 1)^s,$$

and therefore (3) implies

$$\mathbb{P}\left(D_N^*(x_n) \leq D_N^*(q_n) + \varepsilon\right) > 1 - 2(2e)^s (2/\varepsilon + 1)^s e^{-\varepsilon^2 N/2}. \tag{4}$$

In dimension $s = 2$ Gnewuch [8] proved a stronger upper bound for covering numbers, and conjectured that in all dimensions

$$\mathcal{N}(s, \delta) \leq 2\delta^{-s} + o_s(\delta^{-s}).$$

(where o_s means that the implied constant may depend on s). This would lead to an improvement of (4).

We will also need the notion of δ-*bracketing covers*: Let $\delta \in (0,1]$. A finite set Δ of pairs of points from $[0,1]^s$ is called a δ-bracketing cover of $[0,1]^s$, if for every pair $(x,z) \in \Delta$ the estimate $\lambda([0,z]) - \lambda([0,x]) \leq \delta$ holds, and if for every $y \in [0,1]^s$ there exists a pair (x,z) from Δ such that $x \leq y \leq z$. The number $\mathcal{N}_{[\,]}(s, \delta)$, which is called the *bracketing number*, denotes the smallest cardinality of a δ-bracketing cover of $[0,1]^s$. By [7, Theorem 1.15]

$$\mathcal{N}_{[\,]}(s, \delta) \leq 2^{s-1} e^s (\delta^{-1} + 1)^s.$$

Gnewuch's result (3) has the advantage of being valid for all $N \geq 1$. However, (2) is asymptotically stronger than (3) (as N increases, for fixed ε). On the one hand, the purpose of this paper is to show an improved version of (2), which is almost optimal. On the other hand, we want to show that the factor ε^{-s} in (3) and (4), which essentially comes from the necessity to discretize the discrepancy with respect to a grid of precision ε, is not necessary and can be replaced by γ^s for an appropriate constant γ. This might be surprising at first sight: the impact of the necessity to discretize the discrepancy with respect to a certain (possibly extremely close-meshed) grid does not depend on the accuracy of this grid.

More precisely, we will prove the following theorem:

Theorem 1 *Let $(q_n)_{n\geq 1}$ be a d-dimensional sequence, and let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables having uniform distribution on $[0,1]^{s-d}$. Let $(x_n)_{n\geq 1}$ denote the mixed sequence which consists of the points $x_n = (q_n, X_n)$. Then for every $\eta > 0$ there exists a constant $\gamma = \gamma(\eta)$ such that for every $\varepsilon > 0$*

$$\mathbb{P}\left(D_N^*(x_n) \leq 2D_N^*(q_n) + \varepsilon\right) \geq 1 - \gamma^s e^{-2(1-\eta)\varepsilon^2 N}. \tag{5}$$

In (5) we can choose

$$\gamma = e^{2\cdot\lceil 4\log_2(3/\eta)+2\log_2 7\rceil}. \tag{6}$$

As a direct consequence of Theorem 1 we obtain the following corollary, which is an improvement of the result of Ökten, Tuffin and Burago (2).

Corollary 1 *Assume that $D_N^*(q_n) \to 0$, and let $\eta > 0$ be given. Then for arbitrary $\varepsilon > 0$*

$$\mathbb{P}\left(D_N^*(x_n) \leq \varepsilon\right) \geq 1 - e^{-2(1-\eta)\varepsilon^2 N}$$

for sufficiently large N.

Proof of Corollary 1: Let $\eta > 0$ be given, and let $\hat{\eta}$ be so small that $(1-\hat{\eta})^3 > 1-\eta$. Since $D_N^*(q_n) \to 0$ we have $D_N^*(q_n) \leq \hat{\eta}\varepsilon/2$ for sufficiently large N. Thus by Theorem 1

$$\begin{aligned}
\mathbb{P}\left(D_N^*(x_n) \leq \varepsilon\right) &\geq \mathbb{P}\left(D_N^*(x_n) \leq 2D_N^*(q_n) + (1-\hat{\eta})\varepsilon\right) \\
&\geq 1 - \gamma(\hat{\eta})^s e^{-2(1-\hat{\eta})((1-\hat{\eta})\varepsilon)^2 N} \\
&\geq 1 - e^{-2(1-\eta)\varepsilon^2 N}
\end{aligned}$$

for sufficiently large N. This proves the corollary. □

Remark 1: Theorem 1 and Gnewuch's result (3) both give probability zero for $\varepsilon \leq s^{1/2}N^{-1/2}$. It is clear that a result like Theorem 1 can not give a positive probability for all possible $d \geq 1,\ s > d$ and $\varepsilon > 0$, since this would imply (by choosing $d = 1$ and $(q_1, \dots, q_N)$ such that $D_N^*(q_n) = 1/N$) the existence of an s-dimensional sequence $(x_1, \dots, x_N)$ with discrepancy $\leq 2/N + \varepsilon$ for arbitrary s and N, which is in conflict with Roth's result. In fact the bound $s^{1/2}N^{-1/2}$ might be crucial: it is know that for all $N \geq 1$ and $s \geq 1$ there exists an N-element sequence having discrepancy $\leq 10s^{1/2}N^{-1/2}$, but it is unknown how far this upper bound is from optimality. For more information we refer to [1], [10] and [12].

Remark 2: Gnewuch [9, Remark 3.4] showed that in every bound of the form

$$\mathbb{P}\left(D_N^*(x_n) \leq D_N^*(q_n) + \varepsilon\right) \geq 1 - f(s,\varepsilon)e^{-\varepsilon^2 N/2}$$

the function $f(s,\varepsilon)$ has to grow at least exponentially in s (this follows from a general result of Heinrich, Novak, Wasilkowski and Woźniakowski [10]). Using exactly the same argument it can be easily shown that every function $f(s)$ replacing the factor γ^s in our Theorem 1 (for some fixed η) has to grow at least exponentially in s. Thus the only possible improvement of Theorem 1 with respect to s is a reduction of the base γ of the term γ^s.

Remark 3: For any dimensions $d \geq 1$ and $s > d$ it is impossible to find constants $\eta > 0$ and $\gamma > 0$ such that for arbitrary $\varepsilon > 0$

$$\mathbb{P}\left(D_N^*(x_n) \leq 2D_N^*(q_n) + \varepsilon\right) \geq 1 - \gamma^s e^{-2(1+\eta)\varepsilon^2 N}$$

for sufficiently large N. Thus the exponent $2(1-\eta)\varepsilon^2 N$ in Theorem 1 can not be improved to $2(1+\eta)\varepsilon^2 N$ (a proof of this remark will be given at the end of this paper).

Remark 4: Our corollary shows that it is possible to obtain an asymptotic order of $e^{-2(1-\eta)\varepsilon^2 N}$ (for ε fixed, as $N \to \infty$) for arbitrarily small $\eta > 0$. However, as η gets smaller the necessary value of the constant γ in (5) and (6) increases , and in particular $\gamma \to \infty$ as $\eta \to 0$. We are not able to decide whether it is possible to improve Theorem 1 to

$$\mathbb{P}\left(D_N^*(x_n) \leq 2D_N^*(q_n) + \varepsilon\right) \geq 1 - \gamma^s e^{-2\varepsilon^2 N}$$

for some constant γ. Summarizing these results, we know for every $\eta > 0$ that an asymptotic order of $e^{-2(1-\eta)\varepsilon^2 N}$ is possible and $e^{-2(1+\eta)\varepsilon^2 N}$ is impossible, while the "critical" case $e^{-2\varepsilon^2 N}$ remains open.

Remark 5: There are two differences between Theorem 1 and Gnewuch's result (3). On the one hand, our bound for the discrepancy is $2D_N^*(q_n)+\varepsilon$ instead of $D_N^*(q_n)+\varepsilon$. The additional term $D_N^*(q_n)$ comes from the interval partitioning method which is used in our proof, and it seems that this extra term can not be avoided. In applications this should not cause problems, since the deterministic sequence $(q_n)_{1 \leq n \leq N}$ is chosen in such a way that $D_N^*(q_n)$ is very small, whereas ε can not be arbitrarily small (see Remark 1). On the other hand, we can avoid the factor ε^{-s} from Gnewuch's result, which can have a significant contribution particularly for large values of s.

2 Preliminaries

We will use Hoeffding's inequality and Bernstein's inequality, two classical inequalities from probability theory.

Hoeffding's inequality: For $Z_1, \dots, Z_N$ being independent random variables, satisfying $a \leq |Z_n| \leq b$ a.s. for some $a < b,\ b - a \leq 1$,

$$\mathbb{P}\left(\left|\sum_{n=1}^N (Z_n - \mathbb{E}Z_n)\right| > t\right) \leq 2e^{-2t^2}.$$

Bernstein's inequality: For $Z_1, \dots, Z_N$ being independent random variables, satisfying $|Z_n - \mathbb{E}Z_n| \leq 1$ a.s.,

$$\mathbb{P}\left(\left|\sum_{n=1}^N (Z_n - \mathbb{E}Z_n)\right| > t\right) \leq 2\exp\left(-\frac{t^2}{2\left(\sum_{n=1}^N \mathbb{E}Z_n^2\right) + 2t/3}\right).$$

The following lemma will be needed for the proof of Remark 3:

Lemma 1 *Let $(Z_n)_{n\geq 1}$ be independent, fair Bernoulli random variables. Let $\eta > 0$ be given. Then there exists an $\varepsilon_0 = \varepsilon_0(\eta)$ such that for every fixed $\varepsilon \in (0, \varepsilon_0)$ for all sufficiently large N*

$$\mathbb{P}\left(\sum_{n=1}^{N} Z_n \geq N/2 + \varepsilon N\right) \geq e^{-2\varepsilon^2(1+\eta)N}.$$

Proof: To simplify notations we assume w.l.o.g. that εN is an integer. Let η be given, and set

$$p = \mathbb{P}\left(\sum_{n=1}^{N} Z_n \geq N/2 + \varepsilon N\right).$$

By Taylor's formula we have for sufficiently small ε

$$\log(1/2 + \varepsilon) \geq \log 1/2 + 2\varepsilon - (1+\eta)\varepsilon^2$$

and

$$\log(1/2 - \varepsilon) \geq \log 1/2 - 2\varepsilon - (1+\eta)\varepsilon^2.$$

By Stirling's formula for sufficiently large N

$$\begin{aligned}
\binom{N}{N/2+\varepsilon N} &\geq \frac{1}{2}\frac{N^N}{\sqrt{2\pi N}(N/2+\varepsilon N)^{(N/2+\varepsilon N)}(N/2-\varepsilon N)^{(N/2-\varepsilon N)}} \\
&\geq \left(e^{-\eta\varepsilon^2}\right)^N \left(\frac{N^N}{(N/2+\varepsilon N)^{(N/2+\varepsilon N)}(N/2-\varepsilon N)^{(N/2-\varepsilon N)}}\right)
\end{aligned}$$

and therefore, also for sufficiently large N,

$$\begin{aligned}
p^{1/N} &= \left(\sum_{k=N/2+\varepsilon N}^{N} \binom{N}{k}\frac{1}{2^N}\right)^{1/N} \\
&\geq \left(\binom{N}{N/2+\varepsilon N}\frac{1}{2^N}\right)^{1/N} \\
&\geq e^{-\eta\varepsilon^2}\left(\frac{N^N}{(N/2+\varepsilon N)^{(N/2+\varepsilon N)}(N/2-\varepsilon N)^{(N/2-\varepsilon N)}2^N}\right)^{1/N} \\
&= e^{-\eta\varepsilon^2}\left(\frac{1}{(1/2+\varepsilon)^{(1/2+\varepsilon)}(1/2-\varepsilon)^{(1/2-\varepsilon)}2}\right) \\
&\geq e^{-\eta\varepsilon^2}\exp\Big(-(1/2+\varepsilon)\log(1/2+\varepsilon) - (1/2-\varepsilon)\log(1/2-\varepsilon) - \log 2\Big) \\
&\geq e^{-\eta\varepsilon^2}\exp\Big(-(1/2+\varepsilon)(\log 1/2 + 2\varepsilon - (1+\eta)\varepsilon^2) \\
&\qquad -(1/2-\varepsilon)(\log 1/2 - 2\varepsilon - (1+\eta)\varepsilon^2) - \log 2\Big) \\
&= e^{-\eta\varepsilon^2}\exp\Big(-(2+\eta)\varepsilon^2\Big).
\end{aligned}$$

Thus for sufficiently large N

$$p \geq \exp\Big(-2(1+\eta)\varepsilon^2 N\Big). \qquad \square$$

3 Proof of Theorem 1

We use a refined version of the dyadic partitioning technique in [1]. Let $N \geq 1, \varepsilon > 0, \eta > 0$ and a parameter $\mu \geq 10$ be given (μ will be chosen as a function of η, see equation (26) below). For simplicity we assume that μ is an integer.
Let $(q_n)_{n\geq 1}$ be a d-dimensional sequence, and write D for the (d-dimensional) star discrepancy of $(q_n)_{1\leq n\leq N}$. Let $X_1, \ldots, X_N$ be i.i.d. random variables defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, having uniform distribution on $[0,1]^{s-d}$, and write $(x_n)_{1\leq n\leq N}$ for the mixed sequence which consists of the s-dimensional points $x_n = (q_n, X_n)$. We will use the estimate

$$(2e)^s (2^k + 1)^s \leq e^{(k-1)s}, \tag{7}$$

which holds for all $k \geq \mu$ (since we assumed $\mu \geq 10$, and since of course $s \geq 2$).

Assume now that

$$\varepsilon \geq 2^{-\mu}, \tag{8}$$

and let Γ be a $2^{-2\mu}$-cover of $[0,1]^s$ for which

$$\#\Gamma \leq (2e)^s (2^{2\mu} + 1)^s \leq e^{(2\mu-1)s} \leq \frac{e^{2\mu s}}{2}.$$

Then, using Gnewuch's method from [9] and Hoeffding's inequality we can easily show that

$$\begin{aligned} \mathbb{P}\left(D_N^*(x_n) \leq D + \varepsilon + 2^{-2\mu}\right) &\geq 1 - 2e^{-2\varepsilon^2 N}(\#\Gamma) \\ &\geq 1 - e^{2\mu s} e^{-2\varepsilon^2 N}, \end{aligned}$$

which by (8) implies

$$\mathbb{P}\left(D_N^*(x_n) \leq D + \varepsilon + \varepsilon 2^{-\mu}\right) \geq 1 - e^{2\mu s} e^{-2\varepsilon^2 N}. \tag{9}$$

For the rest of the proof we assume that instead of (8)

$$\varepsilon \leq 2^{-\mu} \tag{10}$$

holds (which is the much more difficult case). Additionally we assume that

$$\varepsilon > \frac{\sqrt{\mu s}}{\sqrt{2N}}. \tag{11}$$

(this additional assumption will be dropped later). Let

$$K = K(\varepsilon) := \min\left\{k \geq 1 : \ 2^{-k/2} k^{-1/2} \leq \varepsilon\right\}.$$

Then

$$2^{-K/2} K^{-1/2} \leq \varepsilon \leq 2 \cdot 2^{-K/2} K^{-1/2}, \tag{12}$$

and $\mu \geq 10$ implies

$$K \geq \mu + 15 \geq 25. \tag{13}$$

By (12) and (13) we have $2^{-K} \geq \varepsilon^2$. Thus by (10) and (11)

$$K \leq \log_2(\varepsilon^{-2}) \leq \log_2(2N/s\mu) \leq \log_2(N/10) \leq 2N^{1/4} \leq \varepsilon^{3/2} N \leq 2^{-\mu/2} \varepsilon N. \tag{14}$$

For $\mu \leq k \leq K-1$ let Γ_k be a 2^{-k}-cover of $[0,1]^s$, for which

$$\#\Gamma_k \leq (2e)^s(2^k+1)^s. \tag{15}$$

Let Δ_K denote a 2^{-K}-bracketing cover of $[0,1]^s$ for which

$$\#\Delta_K \leq (2e)^s(2^k+1)^s. \tag{16}$$

Such sets Γ_k and Δ_K exist by a result of Gnewuch [7, Theorem 1.15]. For notational convenience we also define

$$\Gamma_K = \{x \in [0,1]^s : \ (x,y) \in \Delta_K \text{ for some } y \in [0,1]^s\}$$

and

$$\Gamma_{K+1} = \{y \in [0,1]^s : \ (x,y) \in \Delta_K \text{ for some } x \in [0,1]^s\}$$

For every $x \in [0,1]^s$ there exists a pair $(p_K, p_{K+1}) = (p_K(x), p_{K+1}(x))$ for which $(p_K, p_{K+1}) \in \Delta_K$ such that $p_K \leq x \leq p_{K+1}$ and

$$\lambda([0,p_{K+1}]) - \lambda([0,p_K]) \leq \frac{1}{2^K}. \tag{17}$$

For every $x \in [0,1]^s$ and $k = K, K-1, \dots, \mu+1$ we can recursively determine points $p_{k-1} = p_{k-1}(x) \in \Gamma_{k-1} \cup \{0\}$, such that $p_{k-1}(x) \leq p_k(x)$ and

$$\lambda([0,p_k]) - \lambda([0,p_{k-1}]) \leq \frac{1}{2^{k-1}}.$$

For notational convenience we also define

$$p_{\mu-1} = 0.$$

We define for $x, y \in [0,1]^s$

$$\overline{[x,y]} := \begin{cases} [0,y]\backslash[0,x] & \text{if } x \neq 0 \\ [0,y] & \text{if } x = 0, \ y \neq 0. \\ \emptyset & \text{if } x = y = 0. \end{cases}$$

Then the sets

$$I_k(x) := \overline{[p_k(x), p_{k+1}(x)]}, \quad \mu - 1 \leq k \leq K,$$

are disjoint, we have

$$\bigcup_{k=\mu-1}^{K-1} I_k(x) \subset [0,x] \subset \bigcup_{k=\mu-1}^{K} I_k(x),$$

and for all $x, y \in [0,1]^s$

$$\sum_{k=\mu-1}^{K-1} \mathbb{1}_{I_k(x)}(y) \leq \mathbb{1}_{[0,x]}(y) \leq \sum_{k=\mu-1}^{K} \mathbb{1}_{I_k(x)}(y). \tag{18}$$

For $\mu - 1 \leq k \leq K$ we write A_k for the set of all sets of the form $I_k(x)$, where x can take any possible value from $[0,1]^s$. Then by (7), (15) and (16), A_k contains at most

$$\#\Gamma_{k+1} \leq e^{ks} \tag{19}$$

elements. All elements of A_k, where $\mu \le k \le K$, have Lebesgue measure bounded by 2^{-k}. The elements of $A_{\mu-1}$ can have Lebesgue measure between 0 and 1.

For any $k \in \{\mu, \dots, K+1\}$ we will represent the numbers $p_k \in \Gamma_k$ in the form (u_k, v_k), where $u_k \in [0,1]^d$ and $v_k \in [0,1]^{s-d}$, such that p_k has the coordinates $(u_k^{(1)}, \dots, u_k^{(d)}, v_k^{(1)}, \dots, v_k^{(s-d)})$. We write U_k and V_k for the intervals $[0, u_k]$ and $[0, v_k]$, and (U_k, V_k) for the sets $U_k \times V_k = [0, p_k]$. Every $x \in [0,1]^s$ uniquely determines points $p_k \in \Gamma_k,\ \mu \le k \le K+1$, and hence the according values of I_k, u_k, v_k, U_k, V_k are also uniquely defined.

For two sets $I_{k-1} \in A_{k-1}$ and $I_k \in A_k$ we write $I_{k-1} \prec I_k$ if there exists an $x \in [0,1]^s$ such that $I_k = I_k(x)$ and $I_{k-1} = I_{k-1}(x)$. For every $I_k \in A_k,\ \mu \le k \le K$ there exists exactly one element I_{k-1} of A_{k-1} for which $I_{k-1} \prec I_k$. Every $I_k \in A_k,\ \mu \le k \le K$ uniquely determines sets $I_{\mu-1}, \dots, I_{k-1}$ such that $I_{\mu-1} \prec \cdots \prec I_{k-1} \prec I_k$. Whenever I_k is fixed we will write $I_{\mu-1} \dots, I_{k-1}$ for these sets, which are uniquely determined, and $p_l, u_l, v_l, U_l, V_l,\ \mu \le l \le k-1$ for the according values, which are also uniquely determined.

Every $I_k \in A_k,\ \mu \le k \le K$, is of the form

$$(U_{k+1}, V_{k+1}) \backslash (U_k, V_k) = ((U_{k+1} \backslash U_k) \times V_{k+1}) \cup (U_k \times (V_{k+1} \backslash V_k)).$$

Every $I_{\mu-1} \in A_{\mu-1}$ is of the form $[0, p_\mu] = (U_\mu, V_\mu)$.

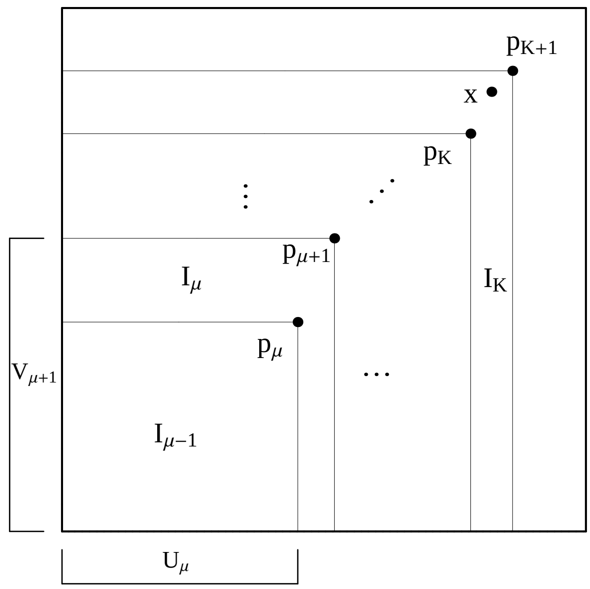

Figure 1: An illustration of our construction in the case $d = 1,\ s = 2$. A point $x \in [0,1]^2$ is given and determines points $p_\mu, p_{\mu+1}, \dots, p_{K+1}$ and sets $I_{\mu-1} \prec I_\mu \prec \cdots \prec I_K$. For exemplification we have also marked the sets U_μ and $V_{\mu+1}$. Every set $I_k,\ \mu \le k \le K$, is of the form $(U_{k+1}, V_{k+1}) \backslash (U_k, V_k) = ((U_{k+1} \backslash U_k) \times V_{k+1}) \cup (U_k \times (V_{k+1} \backslash V_k))$.

Step by step we construct a function $S(I)$ for intervals I from $A_{\mu-1},\dots,A_K$, such that for every I the function value $S(I)$ is a subset of $\{1,\dots,N\}$ (we explain the necessity of this function S in the footnote[1]).
Firstly, let $I_{\mu-1}\in A_{\mu-1}$. Then $I_{\mu-1}$ is of the form (U_μ,V_μ), and we can find $\lceil N\lambda(U_\mu)-ND\rceil$ indices n from $\{1,\dots,N\}$ for which $q_n\in U_\mu$. This is possible since the discrepancy of $(q_n)_{1\le n\le N}$ is bounded by D, and hence the interval U_μ of Lebesgue measure $\lambda(U_\mu)$ contains at least $\lceil N\lambda(U_\mu)-ND\rceil$ points of $(q_n)_{1\le n\le N}$. Denote this set of indices by $S(I_\mu)$.
In the next step let I_μ denote an element of A_μ. Then I_μ is of the form $(U_\mu,V_\mu)\backslash I_{\mu-1}$, where $I_{\mu-1}\in A_{\mu-1}$ and $I_{\mu-1}\prec I_\mu$. We can find $\lfloor N\lambda(U_{\mu+1}\backslash U_\mu)\rfloor$ indices n which are not contained in $S(I_{\mu-1})$ but for which $q_n\in U_{\mu+1}$. We write $S(I_{\mu+1})$ for this set of indices.
Generally, assume that the function S is defined for all intervals in A_k for $k=\mu-1,\mu,\dots,m$ for some m. Let I_{m+1} denote an element of A_{m+1}. Then I_{m+1} is of the form $(U_{m+2},V_{m+2})\backslash(I_{\mu-1}\cup I_\mu\cup\dots\cup I_m)$, where $I_k\in A_k$ for $k=\mu-1,\dots,m$ and $I_{\mu-1}\prec\dots\prec I_m\prec I_{m+1}$. We can find $\lfloor N\lambda(U_{m+2}\backslash U_{m+1})\rfloor$ indices n which are not contained in $\bigcup_{k=\mu-1}^m S(I_k)$, but for which $q_n\in U_{m+2}$. We write $S(I_{k+1})$ for this set of indices.
Proceeding in this way we define the function S for all elements of $A_{\mu-1},\dots,A_K$.

Additionally we define for every $I_k\in A_k,\ \mu\le k\le K$,

$$R(I_k)=S(I_{\mu-1})\cup\dots\cup S(I_{k-1}),$$

where $I_{\mu-1}\prec\dots\prec I_k$.

Then

$$\begin{aligned}
\#\bigcup_{k=\mu-1}^{K-1}S(I_k) &\ge \lceil N\lambda(U_\mu)-ND\rceil+\sum_{k=\mu}^{K-1}\lfloor N\lambda(U_{k+1}\backslash U_k)\rfloor\\
&\ge N\lambda\left(\bigcup_{k=\mu}^{K-1}U_{k+1}\backslash U_k\right)-ND-(K-\mu)\\
&= N\lambda(U_K)-ND-(K-\mu)\\
&\ge \sum_{n=1}^N \mathbb{1}_{U_K}(q_n)-2ND-(K-\mu),
\end{aligned}$$

and accordingly

$$\#\bigcup_{k=\mu-1}^{K}S(I_k)\ \ge\ \sum_{n=1}^N \mathbb{1}_{U_{K+1}}(q_n)-2ND-(K+1-\mu).$$

[1] Our proof is based on the decomposition of the unit cube into parts, and the fact that an arbitrary interval can be written as an union of sets of quickly decreasing Lebesgue measure. However, in our situation this method can only be directly applied if the number of elements of $(q_n)_{1\le n\le N}$ in a subset U of $[0,1]^d$ is $\approx\lambda(U)N$. Unfortunately, this is not necessarily the case: the sets U we consider can be written in the form $U^+\backslash U^-$ for some axis-parallel boxes U^+ and U^-. Thus, if the discrepancy D of $(q_n)_{1\le n\le N}$ is large in comparison with $\lambda(U)$, the number of elements of $(q_n)_{1\le n\le N}$, which are contained in U (which can be any number from $[N\lambda(U)-2ND,N\lambda(U)+2ND]$) can be much larger than $N\lambda(U)$ (and this may hold not only for one, but for several of the sets which we need in our decomposition!). To solve this problem, we distribute the indices $\{1,\dots,N\}$ to the sets in our decomposition in an appropriate regular way, instead of assigning them directly to the sets to which they actually belong.

Thus

$$
\begin{aligned}
& \sum_{n=1}^{N} \mathbb{1}_{[0,x]}(x_n) \\
\geq \quad & \sum_{n=1}^{N} \mathbb{1}_{[0,p_K]}(x_n) \\
= \quad & \sum_{n=1}^{N} \mathbb{1}_{U_K}(q_n) \cdot \mathbb{1}_{V_K}(X_n) \\
= \quad & \sum_{n=1}^{N} \mathbb{1}_{U_\mu}(q_n) \cdot \mathbb{1}_{V_\mu}(X_n) + \sum_{k=\mu}^{K-1} \sum_{n=1}^{N} \left(\mathbb{1}_{U_{k+1} \setminus U_k}(q_n) \cdot \mathbb{1}_{V_{k+1}}(X_n) + \mathbb{1}_{U_k}(q_n) \cdot \mathbb{1}_{V_{k+1} \setminus V_k}(X_n) \right) \\
\geq \quad & \sum_{n \in S(I_{\mu-1})} \mathbb{1}_{V_\mu}(X_n) + \sum_{k=\mu}^{K-1} \left(\sum_{n \in S(I_k)} \mathbb{1}_{V_{k+1}}(X_n) + \sum_{n \in R(I_k)} \mathbb{1}_{V_{k+1} \setminus V_k}(X_n) \right) \qquad (20)
\end{aligned}
$$

and

$$
\begin{aligned}
& \sum_{n=1}^{N} \mathbb{1}_{[0,x]}(x_n) \\
\leq \quad & \sum_{n=1}^{N} \mathbb{1}_{U_{K+1}}(q_n) \cdot \mathbb{1}_{V_{K+1}}(X_n) \\
\leq \quad & \left(\sum_{n \in \bigcup_{k=\mu-1}^{K} S(I_k)} \mathbb{1}_{[0,p_{K+1}]}(x_n) \right) + 2ND + (K+1-\mu) \\
\leq \quad & \sum_{n \in S(I_{\mu-1})} \mathbb{1}_{V_\mu}(X_n) + \sum_{k=\mu}^{K} \left(\sum_{n \in S(I_k)} \mathbb{1}_{V_{k+1}}(X_n) + \sum_{n \in R(I_k)} \mathbb{1}_{V_{k+1} \setminus V_k}(X_n) \right) \\
& \quad +2ND + (K+1-\mu). \qquad (21)
\end{aligned}
$$

Let $I_{\mu-1} \in A_{\mu-1}$ and define

$$Z = Z(I_{\mu-1}) = \sum_{n \in S(I_{\mu-1})} \mathbb{1}_{V_\mu}(X_n).$$

Then by Hoeffding's inequality

$$\mathbb{P}\left(|Z - \mathbb{E}Z| > \varepsilon N\right) \leq 2e^{-2\varepsilon^2 N}. \qquad (22)$$

Now assume that $I_k \in A_k$ for some $k,\ \mu \leq k \leq K$. Then the random variable

$$Z = Z(I_k) \quad = \quad \sum_{n \in S(I_k)} \mathbb{1}_{V_{k+1}}(X_n) + \sum_{n \in R(I_k)} \mathbb{1}_{V_{k+1} \setminus V_k}(X_n)$$

(which is a sum of independent random variables) has expected value

$$\sum_{n \in S(I_k)} \lambda(V_{k+1}) + \sum_{n \in R(I_k)} \lambda(V_{k+1} \setminus V_k)$$

and variance

$$\begin{aligned}
& \sum_{n\in S(I_k)} \lambda(V_{k+1})(1-\lambda(V_{k+1})) + \sum_{n\in R(I_k)} \lambda(V_{k+1}\backslash V_k)(1-\lambda(V_{k+1}\backslash V_k)) \\
\leq \; & \sum_{n\in S(I_k)} \lambda(V_{k+1}) + \sum_{n\in R(I_k)} \lambda(V_{k+1}\backslash V_k) \\
\leq \; & \lambda(V_{k+1})\cdot \# S(I_k) + \lambda(V_{k+1}\backslash V_k)\cdot \# R(I_k) \\
\leq \; & N\lambda(U_{k+1}\backslash U_k)\lambda(V_{k+1}) + N\lambda(U_k)\lambda(V_{k+1}\backslash V_k) \\
= \; & N\lambda(I_k) \\
\leq \; & N 2^{-k}.
\end{aligned}$$

We apply Bernstein's inequality and obtain for $t>0$

$$\mathbb{P}\left(|Z-\mathbb{E}Z|>t\right) \leq 2\exp\left(-\frac{t^2}{2^{-k+1}N+2t/3}\right). \tag{23}$$

If we let

$$t=\frac{6k^{1/2}\varepsilon N}{5\cdot 2^{k/2}},$$

then by (12) we have

$$2t/3 \leq \frac{24N}{15\cdot 2^k},$$

and therefore

$$\begin{aligned}
\mathbb{P}\left(|Z-\mathbb{E}Z|>t\right) & \leq 2\exp\left(-\frac{36k\varepsilon^2 N}{25\,(2+24/15)}\right) \\
& = 2e^{-2k\varepsilon^2 N/5}.
\end{aligned} \tag{24}$$

Let

$$B_{\mu-1}=\bigcup_{I\in A_{\mu-1}} \left(|Z(I_{\mu-1})-\mathbb{E}Z(I_{\mu-1}|>\varepsilon N\right)$$

Then by (19) and (22) we have

$$\mathbb{P}(B_{\mu-1}) \leq 2e^{-2\varepsilon^2 N/5}e^{\mu s}.$$

For $\mu\leq k\leq K$ define

$$B_k=\bigcup_{I_k\in A_k}\left(|Z(I_k)-\mathbb{E}Z(I_k)|>\frac{6k^{1/2}\varepsilon N}{5\cdot 2^{k/2}}\right).$$

Then by (19) and (24), and since $\varepsilon^2 N>\mu s/2\geq 5s$,

$$\sum_{k=\mu}^{K}\mathbb{P}(B_k)\leq \sum_{k=\mu}^{K} 2e^{-2k\varepsilon^2 N/5}e^{ks} \leq \sum_{k=\mu}^{K} 2e^{-k\varepsilon^2 N/5} \leq 3e^{-\mu\varepsilon^2 N/5} \leq 3e^{-2\varepsilon^2 N}.$$

Overall we have

$$\mathbb{P}\left(\bigcup_{k=\mu-1}^{K} B_k\right) \leq 3e^{-2\varepsilon^2 N}+2e^{-2\varepsilon^2 N}e^{3\mu s} \leq 3e^{-2\varepsilon^2 N}e^{\mu s}.$$

Thus by (12), (14), (17) and (21) we have with probability at least $1-3e^{-2\varepsilon^2 N}e^{\mu s}$ for all $x \in [0,1]^s$

$$
\begin{aligned}
& \sum_{n=1}^{N} \mathbb{1}_{[0,x]}(x_n) \\
\leq \quad & \sum_{n \in S(I_{\mu-1})} \mathbb{1}_{V_\mu}(X_n) + \sum_{k=\mu}^{K} \left(\sum_{n \in S(I_k)} \mathbb{1}_{V_{k+1}}(X_n) + \sum_{n \in R(I_k)} \mathbb{1}_{V_{k+1} \setminus V_k}(X_n) \right) \\
& \quad +2ND + (K+1-\mu) \\
\leq \quad & \mathbb{E} \left(\sum_{n \in S(I_{\mu-1})} \mathbb{1}_{V_\mu}(X_n) + \sum_{k=\mu}^{K} \left(\sum_{n \in S(I_k)} \mathbb{1}_{V_{k+1}}(X_n) + \sum_{n \in R(I_k)} \mathbb{1}_{V_{k+1} \setminus V_k}(X_n) \right) \right) \\
& \quad + \varepsilon N \left(1 + \sum_{k=\mu}^{K} \frac{6k^{1/2}}{5 \cdot 2^{k/2}} \right) + 2ND + 2^{-\mu/2} \varepsilon N \\
= \quad & \mathbb{E} \left(\sum_{n \in \bigcup_{k=\mu-1}^{K} S(I_k)} \mathbb{1}_{V_K}(X_n) \right) + \varepsilon N \left(1 + 2^{-\mu/2} + \sum_{k=\mu}^{K} 4k^{1/2} 2^{-k/2} \right) + 2ND \\
\leq \quad & N\lambda([0, p_{K+1}]) + \varepsilon N \left(1 + 2^{-\mu/2} + \sum_{k=\mu}^{K} \frac{6k^{1/2}}{5 \cdot 2^{k/2}} \right) + 2ND + (K+1-\mu) \\
\leq \quad & N\lambda([0,x]) + N2^{-K} + \varepsilon N \left(1 + 2^{-\mu/2} + \sum_{k=\mu}^{K} \frac{6k^{1/2}}{5 \cdot 2^{k/2}} \right) + 2ND) \\
\leq \quad & N\lambda([0,x]) + \varepsilon N \left(1 + K^{1/2} 2^{-K/2} + 2^{-\mu/2} + \underbrace{\sum_{k=\mu}^{K} \frac{6k^{1/2}}{5 \cdot 2^{k/2}}}_{\leq 5\mu^{1/2} 2^{-\mu/2}} \right) + 2ND \\
\leq \quad & N\lambda([0,x]) + \varepsilon N \left(1 + 7\mu^{1/2} 2^{-\mu/2} \right) + 2ND
\end{aligned}
$$

Similarly by (12), (14), (17) and (20) we have with probability at least $1-3e^{-2\varepsilon^2 N}e^{\mu s}$ for all $x \in [0,1]$

$$
\sum_{n=1}^{N} \mathbb{1}_{[0,x]}(x_n) \quad \geq \quad N\lambda([0,x]) - \varepsilon N \left(1 + 7\mu^{1/2} 2^{-\mu/2} \right) - 2ND.
$$

Therefore we have, with probability at least $1-3e^{-2\varepsilon^2 N}e^{\mu s}$,

$$
D_N^*(x_n) \quad \leq \quad 2D + \varepsilon \left(1 + 7\mu^{1/2} 2^{-\mu/2} \right). \tag{25}
$$

This holds under assumptions (10) and (11). Now it is easy to see that (25) also holds without assuming (11), since in this case $1-3e^{-2\varepsilon^2 N}e^{\mu s} \leq 0$ (cf. Remark 1). Comparing this result with (9), which holds under assumption (8) we see that (25) holds with probability greater than or equal to

$$1 - e^{-2\varepsilon^2 N} e^{2\mu s}.$$

Now let η be given. Set

$$\mu = \lceil 4 \log_2(3/\eta) + 2 \log_2 7 \rceil \tag{26}$$

and

$$\gamma = \gamma(\eta) = e^{2\mu} = e^{2 \cdot \lceil 4 \log_2(3/\eta) + 2 \log_2 7 \rceil}.$$

Then $\mu \geq 10$. Some calculations show that for $y \in (0, 1]$

$$\sqrt{4 \log_2(3/y) + 2 \log_2 7} \leq \frac{4}{y}$$

and consequently

$$\begin{aligned}
& \left(1 + 7\sqrt{4 \log_2(3/y) + 2 \log_2 7} \cdot 2^{-(4 \log_2(3/y) + 2 \log_2 7)/2}\right)^{-2} \\
\geq\ & \left(1 + 7 \cdot 4 \cdot y \cdot y^2/9 \cdot 7^{-1}\right)^{-2} \\
\geq\ & (1 + y^2/2)^{-2} \\
\geq\ & 1 - y.
\end{aligned} \tag{27}$$

Thus by (25) and (27) for $\varepsilon > 0$

$$\begin{aligned}
\mathbb{P}\left(D_N^*(x_n) > D + \varepsilon\right) & \leq e^{2\mu s} \exp\left(-2\varepsilon^2 N \left(1 + 7\mu^{1/2} 2^{-\mu/2}\right)^{-2}\right) \\
& \leq \gamma^s e^{-2(1-\eta)\varepsilon^2 N},
\end{aligned}$$

which proves the theorem. $\square$

In conclusion we prove Remark 3 on the asymptotic optimality of the probability estimate

$$1 - \gamma(\eta)^s e^{-2(1-\eta)\varepsilon^2 N}.$$

We show that this lower bound can not be replaced by

$$1 - \gamma(\eta)^s e^{-2(1+\eta)\varepsilon^2 N}$$

for any positive η, no matter how large the constant $\gamma(\eta)$ is chosen. More precisely, let $d \geq 1$, $s > d$ and $\eta > 0$ be given, and assume that it is possible to find a constant γ such that for every sequence $(q_n)_{n \geq 1}$ and every $\varepsilon > 0$ for sufficiently large N

$$\mathbb{P}\left(D_N^*(x_n) \leq 2D_N^*(q_n) + \varepsilon\right) \geq 1 - \gamma^s e^{-2\varepsilon^2 N(1+\eta)}. \tag{28}$$

Chose $\hat{\eta}$ so small that

$$(1 + \hat{\eta})^3 \leq (1 + \eta), \tag{29}$$

and let $(q_n)_{n \geq 1}$ be a d-dimensional sequence for which $D_N^*(q_n) \to 0$. Write I for the indicator of the s-dimensional box of the form $[0,1]^d \times [0, 2^{1/(s-d)}]$. Then I has Lebesgue measure $1/2$. Let $X_n,\ n \geq 1$ be i.i.d. random variables having uniform distribution on $[0,1]^{s-d}$, and write

$(x_n)_{n\geq 1}$ for the mixed sequence. Then $x_n \in I$ if and only if $X_n \in [0, 2^{1/(s-d)}]$. The random variables

$$\mathbb{1}_I(x_n) = \mathbb{1}_{[0,2^{1/(s-d)}]}(X_n)$$

are independent, fair Bernoulli random variables. Thus, if ε is chosen appropriately small, we have by Lemma 1

$$\mathbb{P}\left(\sum_{k=1}^{N} \mathbb{1}_I(x_n) \geq \frac{N}{2} + (1+\hat{\eta})\varepsilon N\right) \geq e^{-2\varepsilon^2 N(1+\hat{\eta})^3},$$

for sufficiently large N. Since $D_N^*(q_n) \to 0$, this implies

$$\mathbb{P}\left(D_N^*(x_n) \geq 2D_N^*(q_n) + \varepsilon\right) \quad \geq \quad e^{-2\varepsilon^2 N(1+\hat{\eta})^3} \tag{30}$$

for sufficiently large N. By (29)

$$\frac{e^{-2\varepsilon^2 N(1+\hat{\eta})^3}}{e^{-2\varepsilon^2 N(1+\eta)}} \to 0 \qquad \text{as} \qquad N \to \infty,$$

and hence (30) implies

$$\mathbb{P}\left(D_N^*(x_n) \leq 2D_N^*(q_n) + \varepsilon\right) < 1 - \gamma^s e^{-2\varepsilon^2 N(1+\eta)}$$

for sufficiently large N, which is a contradiction to (28).

References

[1] C. Aistleitner. Covering numbers, dyadic chaining and discrepancy. *J. Complexity.* to appear.

[2] J. Beck. A two-dimensional van Aardenne-Ehrenfest theorem in irregularities of distribution. *Compositio Math.*, 72(3):269–339, 1989.

[3] D. Bilyk, M. T. Lacey, and A. Vagharshakyan. On the small ball inequality in all dimensions. *J. Funct. Anal.*, 254(9):2470–2502, 2008.

[4] J. Dick and F. Pillichshammer. *Digital nets and sequences. Discrepancy theory and quasi-Monte Carlo integration.* Cambridge: Cambridge University Press. xvii, 600 p., 2010.

[5] M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications*, volume 1651 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1997.

[6] P. Glasserman. *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York).* Springer-Verlag, New York, 2004. Stochastic Modelling and Applied Probability.

[7] M. Gnewuch. Bracketing numbers for axis-parallel boxes and applications to geometric discrepancy. *J. Complexity*, 24(2):154–172, 2008.

[8] M. Gnewuch. Construction of minimal bracketing covers for rectangles. *Electron. J. Combin.*, 15(1):Research Paper 95, 20, 2008.

[9] M. Gnewuch. On probabilistic results for the discrepancy of a hybrid-Monte Carlo sequence. *J. Complexity*, 25(4):312–317, 2009.

[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(3):279–302, 2001.

[11] F. J. Hickernell. The mean square discrepancy of randomized nets. *ACM Trans. Model. Comput. Simul.*, 6:274–296, October 1996.

[12] A. Hinrichs. Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy. *J. Complexity*, 20(4):477–483, 2004.

[13] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics.

[14] C. Lemieux and P. L'Ecuyer. Variance reduction via lattice rules. *Management Science*, 46:1214–1235, 2000.

[15] J. Matoušek. On the L_2-discrepancy for anchored boxes. *J. Complexity*, 14(4):527–556, 1998.

[16] G. Ökten. A probabilistic result on the discrepancy of a hybrid-Monte Carlo sequence and applications. *Monte Carlo Methods Appl.*, 2(4):255–270, 1996.

[17] G. Ökten. High dimensional simulation. *Math. Comput. Simulation*, 55(1-3):215–222, 2001. The Second IMACS Seminar on Monte Carlo Methods (Varna, 1999).

[18] G. Ökten, B. Tuffin, and V. Burago. A central limit theorem and improved error bounds for a hybrid-Monte Carlo sequence with applications in computational finance. *J. Complexity*, 22(4):435–458, 2006.

[19] A. B. Owen. Monte Carlo, quasi-Monte Carlo, and randomized quasi-Monte Carlo. In *Monte Carlo and quasi-Monte Carlo methods 1998 (Claremont, CA)*, pages 86–97. Springer, Berlin, 2000.

[20] A. V. Roşca. A mixed Monte Carlo and quasi-Monte Carlo sequence for multidimensional integral estimation. *Acta Univ. Apulensis Math. Inform.*, (14):141–160, 2007.

[21] K. F. Roth. On irregularities of distribution. *Mathematika*, 1:73–79, 1954.

[22] J. Spanier. Quasi-Monte Carlo methods for particle transport problems. In *Monte Carlo and quasi-Monte Carlo methods in scientific computing (Las Vegas, NV, 1994)*, volume 106 of *Lecture Notes in Statist.*, pages 121–148. Springer, New York, 1995.