On Probabilistic Results for the Discrepancy of a Hybrid-Monte Carlo Sequence

Michael Gnewuch

Department of Computer Science, University of Kiel,
Christian-Albrechts-Platz 4, 24098 Kiel, Germany
email: mig@informatik.uni-kiel.de

December 19, 2008

Abstract

In many applications it has been observed that hybrid-Monte Carlo sequences perform better than Monte Carlo and quasi-Monte Carlo sequences, especially in difficult problems. For a mixed s-dimensional sequence m, whose elements are vectors obtained by concatenating d-dimensional vectors from a low-discrepancy sequence q with $(s-d)$-dimensional random vectors, probabilistic upper bounds for its star discrepancy have been provided. In a paper of G. Ökten, B. Tuffin and V. Burago [J. Complexity 22 (2006), 435–458] it was shown that for arbitrary $\varepsilon > 0$ the difference of the star discrepancies of the first N points of m and q is bounded by ε with probability at least $1 - 2\exp(-\varepsilon^2 N/2)$ for N sufficiently large. The authors did not study how large N actually has to be and if and how this actually depends on the parameters s and ε. In this note we derive a lower bound for N, which significantly depends on s and ε. Furthermore, we provide a probabilistic bound for the difference of the star discrepancies of the first N points of m and q, which holds without any restrictions on N. In this sense it improves on the bound of Ökten, Tuffin and Burago and is more helpful in practice, especially for small sample sizes N. We compare this bound to other known bounds.

1 Introduction

A commonly used measure for the uniformity of point distributions is the well-known star discrepancy. Let λ_s denote the s-dimensional Lebesgue measure. The *star discrepancy* of an N-point set $P = \{p_1, \ldots, p_N\} \subset [0,1]^s$ is defined by

$$D_N^*(P) := \sup_{\alpha \in [0,1]^s} \left| \lambda_s([0,\alpha)) - \frac{1}{N} \sum_{k=1}^{N} 1_{[0,\alpha)}(p_k) \right| ;$$

here $[0,\alpha)$ denotes the s-dimensional axis-parallel box $[0,\alpha_1)\times\cdots\times[0,\alpha_s)$ and $1_{[0,\alpha)}$ its characteristic function. If p is an infinite sequence in $[0,1]^s$, then $D_N^*(p)$ should be understood as the discrepancy of its first N points. Let $D^*(N,s)$ denote the smallest possible star discrepancy of any N-point set, i.e.,

$$D^*(N,s) = \inf_{P\subset[0,1]^s,\,|P|=N} D_N^*(P).$$

Furthermore, let

$$N^*(\varepsilon,s) := \inf\{N\in\mathbb{N}\,|\,D^*(N,s)\le\varepsilon\}$$

be the *inverse of the star discrepancy.* Apart from the "classical" asymptotic bounds for the star discrepancy (see, e.g., [9]), there are bounds known which describe its behavior in the number of points N and in the dimension s. In [7] Heinrich, Novak, Wasilkowski, and Woźniakowski proved

$$D^*(N,s)\le C\sqrt{s/N} \quad\text{and}\quad N^*(\varepsilon,s)\le\lceil C^2 s\varepsilon^{-2}\rceil, \tag{1}$$

where the constant C does not depend on N, s or ε. The dependence of the inverse of the star discrepancy on s is optimal here; this was proved by a lower bound in [7], which was improved by A. Hinrichs in [8]. He proved the existence of constants $c,\varepsilon_0>0$ such that

$$D^*(N,s)\ge\min\{\varepsilon_0, cs/N\} \quad\text{and}\quad N^*(\varepsilon,s)\ge cs\varepsilon^{-1} \quad\text{for all } \varepsilon\in(0,\varepsilon_0). \tag{2}$$

Studying the proofs, one sees that (2) holds in particular for $c=\varepsilon_0=1/32\mathrm{e}^2$.

The star discrepancy is closely related via the Koksma-Hlawka inequality to the problem of numerical integration of certain function classes, see, e.g., [9]. The essence is that the smaller the discrepancy of a certain point set P, the better the worst-case error guarantee of the corresponding quasi-Monte Carlo cubature $\frac{1}{|P|}\sum_{p\in P} f(p)$.

A rule of thumb is that in most problems low-discrepancy point sets outperform (pseudo-)random sets in moderate dimensions, but lose their effectiveness in high dimensions. For this reason several researchers studied hybrid methods which try to use advantages of both Monte Carlo and quasi-Monte Carlo methods. One example are so-called mixed sequences used by J. Spanier [16] and studied further by G. Ökten and his collaborators [10, 11, 13], and other researchers as, e.g., A. V. Roşca [15]. Mixed sequences showed a favorable performance in many numerical experiments.

2 Previous Work

To obtain a somehow "more objective" measure for the quality of mixed sequences than "just performing" numerical tests, some probabilistic bounds on their star discrepancy have been derived. Let $q=(q_k)$ be a low-discrepancy sequence in $[0,1)^d$, and let $X=(X_k)$ be a sequence of independent and uniformly distributed random variables in $[0,1)^{s-d}$. The resulting s-dimensional sequence $m=(m_k)=(q_k,X_k)$ is called a *mixed sequence.*

The main result in [13], Theorem 5, reads as follows:

Theorem 2.1. *For $s \in \mathbb{N}$, $d \in \{1, \dots, s\}$, and $\varepsilon > 0$*

$$\mathbb{P}\big(D_N^*(m) - D_N^*(q) < \varepsilon\big) \geq 1 - 2\exp\left(-\frac{\varepsilon^2 N}{2}\right) \quad \textit{for } N \textit{ sufficiently large.} \tag{3}$$

(In the actual formulation of [13, Thm.5] one finds the term $\exp(-\varepsilon^2 N/2)$ replaced by $\exp(-2\varepsilon^2 N)$; looking at the proof one sees that [13, Lemma 4] has to be employed with $\varepsilon/2$ instead of ε, leading to the correct result (3).) In [13] Ökten, Tuffin, and Burago did not investigate how large N actually has to be or on which parameters the required size of N really depends. So let $N(q; s, \varepsilon)$ be the smallest number such that (3) holds for all $N \geq N(q; s, \varepsilon)$. We derive now a lower bound on $N(q; s, \varepsilon)$ via the discrepancy bound (2); for simplicity we work with $c = \varepsilon_0 = 1/32\mathrm{e}^2$. So let q (and in particular d) be fixed, let $\varepsilon < 1/64\mathrm{e}^2$, and put $N_q = N_q(\varepsilon) := \min\{N \mid \forall M \geq N : D_M^*(q) \leq \varepsilon\}$. If $N = \max\{N(q; s, \varepsilon), N_q, \lceil 4\ln(2)\varepsilon^{-2}\rceil\}$, then (3) implies in particular the existence of N-point sets P with $D_N^*(P) < 2\varepsilon < 1/32\mathrm{e}^2$. Due to (2) we obtain

$$(1/32\mathrm{e}^2)\min\{1, s/N\} < 2\varepsilon < 1/32\mathrm{e}^2 \quad \text{for all } s \in \mathbb{N}.$$

This leads first to $1 > s/N$, implying $N > (1/64\mathrm{e}^2)(s/\varepsilon)$ for all s. Since N_q and $\lceil 4\ln(2)\varepsilon^{-2}\rceil$ are constant with respect to s, this gives us

$$N(q; s, \varepsilon) > \frac{1}{64\mathrm{e}^2}\frac{s}{\varepsilon} \quad \text{for all but finitely many } s. \tag{4}$$

To be more precise, (4) holds for all $s \geq 64\mathrm{e}^2\varepsilon \max\{N_q, \lceil 4\ln(2)\varepsilon^{-2}\rceil\}$. We can simplify this condition, if, e.g., q satisfies $D_N^*(q) \leq C_q(\ln N)^d/N$ for some $C_q > 0$ and all $N > 2$ (cf. [9]): This implies $N_q \leq C_q'(\ln(\varepsilon^{-1}))^d\varepsilon^{-1}$ for a suitable constant C_q'. Hence there exists some $C_q'' > 0$ such that (4) holds for all $s \geq C_q''\varepsilon^{-1}$. The lower bound (4) is certainly not sharp, but it serves to show that $N(q; s, \varepsilon)$ depends indeed significantly on s and ε.

In the literature one can find bounds similar to (3) without any restrictions on N. Examples are [10, Corollary 1] and [15, Theorem 8] (the latter result is not about the star discrepancy, but about a generalization of the extreme discrepancy). As discussed in [12] and [6], these two results are unfortunately incorrect.

3 An Improved Probabilistic Bound

In this section we want to derive a valid version of the bound (3) that imposes no restrictions on the size of N. Let us first restate a definition from [3].

Definition 3.1. Let $\delta \in (0, 1]$. A finite set $\Gamma \subset [0, 1]^s$ is called a *δ-cover of* $[0, 1]^s$ if for all $y \in [0, 1]^s$ there are $x, z \in \Gamma \cup \{0\}$ such that $x_k \leq y_k \leq z_k$ for $k = 1, \dots, s$ and $\lambda_s([0, z)) - \lambda_s([0, x)) \leq \delta$. We put $\mathcal{N}(s, \delta) := \min\{|\Gamma| \mid \Gamma \text{ is a } \delta\text{-cover of } [0, 1]^s\}$.

From [4, Thm. 1.15] we know that

$$\mathcal{N}(s, \delta) \leq (2s)^s(\delta^{-1} + 1)^s/s! \leq (2e)^s(\delta^{-1} + 1)^s. \tag{5}$$

In [5] a better bound was provided constructively in dimension $s = 2$, and it was conjectured that the construction method can be extended to arbitrary $s \geq 3$ and would lead to $\mathcal{N}(s,\delta) \leq 2\delta^{-s} + O_s(\delta^{-s+1})$.

One may use δ-covers to discretize the star discrepancy at the cost of a discretization error at most δ.

Lemma 3.2. *Let Γ be a δ-cover of $[0,1]^s$, and let $P = \{p_1,\dots,p_N\} \subset [0,1]^s$. Then*

$$D_N^*(P) \leq D_N^\Gamma(P) + \delta, \qquad \textit{where} \qquad D_N^\Gamma(P) := \max_{\alpha\in\Gamma} \left| \lambda_s([0,\alpha)) - \frac{1}{N}\sum_{j=1}^N 1_{[0,\alpha)}(p_j) \right|.$$

The proof is straightforward, see, e.g., [3, Lemma 3.1].

Theorem 3.3. *Let $q = (q_k)$ be a sequence in $[0,1)^d$, $X = (X_k)$ be a sequence of independent and uniformly distributed random variables in $[0,1)^{s-d}$, and let $m = (m_k) = (q_k, X_k)$ be the resulting s-dimensional mixed sequence. Then we have for all $\varepsilon \in (0,1]$*

$$\mathbb{P}\big(D_N^*(m) - D_N^*(q) < \varepsilon\big) > 1 - 2\mathcal{N}(s,\varepsilon/2)\exp\left(-\frac{\varepsilon^2 N}{2}\right). \tag{6}$$

Let $\theta \in [0,1)$. Using the upper bound $\mathcal{N}(s,\varepsilon/2) \leq (2\mathrm{e})^s(2\varepsilon^{-1}+1)^s$, we have with probability strictly larger than θ

$$D_N^*(m) < D_N^*(q) + \sqrt{\frac{2}{N}}\left(s\ln(\rho) + \ln\left(\frac{2}{1-\theta}\right)\right)^{1/2}, \tag{7}$$

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

Remark 3.4. Let us assume that a bound of the form

$$\mathbb{P}\big(D_N^*(m) - D_N^*(q) \leq \varepsilon\big) \geq 1 - f(q;s,\varepsilon)\exp\left(-\frac{\varepsilon^2 N}{2}\right) \tag{8}$$

holds for all ε in some interval $(0,\varepsilon_*]$, all $s > d$ and all N. If $D_N^*(q) < 1/64\mathrm{e}^2$ for all N sufficiently large, then for all ε sufficiently small the function $f(q;s,\varepsilon)$ has to increase at least exponentially in s. Indeed, put $N_q := \min\{N \,|\, \forall M \geq N : D_M^*(q) < 1/64\mathrm{e}^2\}$. Let $\varepsilon < 1/64\mathrm{e}^2$, and put $N := \max\{N_q, \lceil 2(\ln f(q;s,\varepsilon) + \ln 2)\varepsilon^{-2}\rceil\}$. Then (8) implies in particular the existence of N-point sets $P \subset [0,1]^s$ with $D_N^*(P) \leq \varepsilon + D_N^*(q) < 1/32\mathrm{e}^2$. Hence (2) gives us for each $s > d$

$$s < \frac{1}{32\mathrm{e}^2}\,\frac{s}{\varepsilon + D_N^*(q)} \leq N^*(\varepsilon + D_N^*(q), s) \leq \max\{N_q, \lceil 2(\ln f(q;s,\varepsilon) + \ln 2)\varepsilon^{-2}\rceil\},$$

showing that $\ln f(q;s,\varepsilon)$ has to grow for large s at least linearly in s.

This shows that the factor $\mathcal{N}(s,\varepsilon/2)$ on the right hand side of (6) is not an indication of the coarseness of our estimate, but a factor growing exponentially in s has to appear there necessarily.

Proof of Theorem 3.3. Let $\alpha = (\alpha', \alpha'')$, where $\alpha' \in [0,1]^d$, $\alpha'' \in [0,1]^{s-d}$. Let $\xi_k = \xi_k(\alpha) := \lambda_s([0,\alpha)) - 1_{[0,\alpha)}(m_k)$ for $k = 1, 2, \ldots$. The ξ_k, $k = 1, 2, \ldots$, are independent random variables with $\mathbb{E}(\xi_k) = \lambda_{s-d}([0,\alpha''))(\lambda_d([0,\alpha')) - 1_{[0,\alpha')}(q_k))$. Thus we have

$$\left|\mathbb{E}\left(\frac{1}{N}\sum_{k=1}^{N}\xi_k\right)\right| = \lambda_{s-d}([0,\alpha''))\left|\lambda_d([0,\alpha')) - \frac{1}{N}\sum_{k=1}^{N}1_{[0,\alpha')}(q_k)\right| \le D_N^*(q).$$

Let $\delta := \varepsilon/2$. Then the inequality

$$\left|\frac{1}{N}\sum_{k=1}^{N}\xi_k\right| \ge D_N^*(q) + \delta \quad \text{implies} \quad \left|\frac{1}{N}\sum_{k=1}^{N}(\xi_k - \mathbb{E}(\xi_k))\right| \ge \delta.$$

Thus Hoeffding's large deviation bound for sums of independent random variables (see, e.g., [14, p.191]) gives us

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{k=1}^{N}\xi_k\right| \ge D_N^*(q) + \delta\right) \le \mathbb{P}\left(\left|\frac{1}{N}\sum_{k=1}^{N}(\xi_k - \mathbb{E}(\xi_k))\right| \ge \delta\right) \le 2\exp(-2\delta^2 N). \tag{9}$$

To prove (6) we discretize the star discrepancy with the help of a δ-cover Γ and employ a union bound over Γ. So let Γ be a δ-cover of $[0,1]^s$ of minimal size. Then, with the help of Lemma 3.2 and (9), we obtain

$$\begin{aligned}
\mathbb{P}\big(D_N^*(m) - D_N^*(q) < \varepsilon\big) \ge& \mathbb{P}\big(D_N^\Gamma(m) - D_N^*(q) < \delta\big) = 1 - \mathbb{P}\big(D_N^\Gamma(m) \ge D_N^*(q) + \delta\big) \\
=& 1 - \mathbb{P}\left(\max_{\alpha\in\Gamma}\left|\frac{1}{N}\sum_{k=1}^{N}\xi_k(\alpha)\right| \ge D_N^*(q) + \delta\right) \\
=& 1 - \mathbb{P}\left(\bigcup_{\alpha\in\Gamma}\left\{\left|\frac{1}{N}\sum_{k=1}^{N}\xi_k(\alpha)\right| \ge D_N^*(q) + \delta\right\}\right) \\
\ge& 1 - \sum_{\alpha\in\Gamma}\mathbb{P}\left(\left|\frac{1}{N}\sum_{k=1}^{N}\xi_k(\alpha)\right| \ge D_N^*(q) + \delta\right) \\
>& 1 - 2|\Gamma|\exp\left(-2\delta^2 N\right).
\end{aligned}$$

(In the last estimate we obtained "$>$", since we have necessarily $z := (1, 1, \ldots, 1) \in \Gamma$ and $\xi_k(z) = 0$ for all k.) This proves (6).

Now $1 - 2|\Gamma|\exp(-2\delta^2 N) \ge \theta$ iff

$$\delta^2 \ge \frac{1}{2N}\left(\ln|\Gamma| + \ln\left(\frac{2}{1-\theta}\right)\right). \tag{10}$$

Inequality (5) gives us $|\Gamma| \le (2e)^s(\delta^{-1} + 1)^s$. Therefore it is easily verified that (10) holds if we choose δ to be

$$\delta = \sqrt{\frac{1}{2N}}\left(s\ln(\rho) + \ln\left(\frac{2}{1-\theta}\right)\right)^{1/2}.$$

This proves that (7) holds with probability $> \theta$. □

The technique to prove Theorem 3.3 is similar to the one that was used to prove [2, Thm. 3.1]. There the authors wanted to extend a given $(s-1)$-dimensional N-point set $P_{s-1} = \{y_0, \ldots, y_{N-1}\}$ to an s-dimensional set $P_s = \{(y_0, a_0), \ldots, (y_{N-1}, a_{N-1})\}$ with a relatively small star discrepancy by choosing $a_0, \ldots, a_{N-1}$ randomly from a grid in $[0, 1)$ with step size $1/k$. The probabilistic bound on $D^*_N(P_s)$ is essentially the bound (7) (with m replaced by P_s and q replaced by P_{s-1}), but one has to add the term $1/2k$ (which may be viewed as the prize of discretizing). In this way one can generate randomly "component-by-component" (CBC) point sets in arbitrarily high dimension s. Notice that such random CBC-constructions were not considered in [2] because their resulting discrepancy is extraordinarily small; in fact the behavior of the upper bounds proved in [2] with respect to the dimension s is worse than that of the discrepancy bound (1) and the bounds proved in [1, 3, 4, 7] for other random constructions. The aim of [2] was to provide a fast derandomized algorithm (based on the CBC-idea) that generates sets with relatively small star discrepancy. In fact its running time is reasonably faster than that of the preceding algorithms presented in [1, 3], at the prize of a not too much worse theoretical discrepancy bound (which hopefully may even not be observed in practice).

References

[1] Doerr, B., Gnewuch, M.: 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, 299–312, Springer, Berlin Heidelberg, 2008.

[2] Doerr, B., Gnewuch, M., Kritzer, P., Pillichshammer, F.: Component-by-component construction of low-discrepancy point sets of small size. Monte Carlo Methods Appl. 14 (2008), 129–149.

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

[4] Gnewuch, M.: Bracketing numbers for d-dimensional boxes and applications to geometric discrepancy. J. Complexity 24 (2008), 154–172.

[5] Gnewuch, M.: Construction of Minimal Bracketing Covers for Rectangles, Electron. J. Combin., Vol. 15(1) (2008), Research Paper 95, 20 pp..

[6] Gnewuch, M., Roşca, A. V.: On G-discrepancy and mixed Monte Carlo and quasi-Monte Carlo sequences. Submitted.

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

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

[9] Niederreiter, H.: Random Number Generation and Quasi-Monte Carlo Methods. CBMS-NSF Series in Applied Mathematics 63, SIAM, Philadelphia, 1992.

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

[11] Ökten, G.: High dimensional simulation. Math. Comput. Simul. 55 (2001), 215–222.

[12] Ökten, G., Gnewuch, M.: Correction of a proof in "A probabilistic result on the discrepancy of a hybrid-Monte Carlo sequence and applications". To appear in Monte Carlo Methods Appl..

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

[14] Pollard, D.: Convergence of Stochastic Processes. Springer, Berlin, 1984.

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

[16] Spanier, J.: Quasi-Monte Carlo methods for particle transport problems. In: Niederreiter, H., Shiue, P. J.-S. (Eds.), Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, 121–148, Lecture Notes in Statistics 106, Springer, Berlin, 1995.