On G-Discrepancy and Mixed Monte Carlo and Quasi-Monte Carlo Sequences

Michael Gnewuch

Christian-Albrechts University of Kiel
Germany
email: mig@informatik.uni-kiel.de

Alin V. Roşca

Babeş-Bolyai University of Cluj Napoca
Romania
email: arosca@econ.ubbcluj.ro

December 19, 2008

Abstract

The G-star and the G-discrepancy are generalizations of the well known star and the extreme discrepancy; thereby G denotes a given continuous distribution function on the d-dimensional unit cube $[0,1]^d$. We list and prove some results that describe the behavior of the G-star and the G-discrepancy in terms of the dimension d and the number of sample points N.

Our main focus is on so-called mixed sequences, which are d-dimensional hybrid-Monte Carlo sequences that result from concatenating a d'-dimensional deterministic sequence q with a d''-dimensional sequence X of independently G-distributed random vectors (here $d = d' + d''$). We show that a probabilistic bound on the G-discrepancy of mixed sequence from [11] is unfortunately incorrect, and correct it by proving a new probabilistic bound for the G-discrepancy of mixed sequences. Moreover this new bound exhibits for fixed dimension d a better asymptotical behavior in N, and a similar bound holds also for the G-star discrepancy.

1 Introduction

We consider the d-dimensional unit cube $[0,1]^d$ endowed with the usual σ-algebra of Borel sets. Let λ_G be a probability measure on $[0,1]^d$, and let G be its distribution function, i.e.,

$G(x) = \lambda_G([0,x))$ for all $x \in [0,1]^d$, where $[0,x)$ denotes the d-dimensional subinterval $[0,x_1) \times \cdots \times [0,x_d)$.

Put $\mathcal{C}_d = \{[0,x) \,|\, x \in [0,1]^d\}$ (the set of d-dimensional "cubes") and $\mathcal{R}_d = \{[y,z) \,|\, y,z \in [0,1]^d, y \le z\}$ (the set of d-dimensional "rectangles"); here and in the rest of this paper inequalities like $y \le z$ are meant componentwise.

The *G-star discrepancy* of a point set $P = \{p^1, \ldots, p^N\}$ is defined to be

$$D^*_{G,N}(P) = \sup_{C \in \mathcal{C}_d} \left| \frac{1}{N} \sum_{i=1}^N 1_C(p^i) - \lambda_G(C) \right|,$$

where 1_C is the characteristic function of the set C. Furthermore, the *G-discrepancy* of P is given by

$$D_{G,N}(P) = \sup_{R \in \mathcal{R}_d} \left| \frac{1}{N} \sum_{i=1}^N 1_R(p^i) - \lambda_G(R) \right|.$$

If p is an infinite sequence in $[0,1]^d$, then $D^*_{G,N}(p)$ and $D_{G,N}(p)$ should be understood as the discrepancies of its first N points. Let $D^*_G(N,d)$ be the smallest possible G-star discrepancy of any N-point set in $[0,1]^d$, and define the *inverse of the G-star discrepancy* $N^*_G(\varepsilon,d)$ by

$$N^*_G(\varepsilon,d) = \min\{N \in \mathbb{N} \,|\, D^*_G(N,d) \le \varepsilon\}.$$

Analogously, we define $D_G(N,d)$ and $N_G(\varepsilon,d)$. Observe that we always have $D^*_{G,N}(P) \le D_{G,N}(P)$ for all N-point sets $P \subset [0,1]^d$, and consequently $D^*_G(N,d) \le D_G(N,d)$ and $N^*_G(\varepsilon,d) \le N_G(\varepsilon,d)$.

In the case where λ_G is the ordinary d-dimensional Lebesgue measure λ_d restricted to $[0,1]^d$, we simply omit any reference to the corresponding distribution function. The associated discrepancies $D^*_N(P)$ and $D_N(P)$ are the well-known *star discrepancy* and the *extreme discrepancy* (sometimes also called *unanchored discrepancy* or simply *discrepancy*). (Notice that our denominations G-star and G-discrepancy differ a bit from the denominations, e.g., in [8, 11]. We chose these names, because they are consistent with the classical names "star discrepancy" and "discrepancy".)

G-star and G-discrepancy have, e.g., applications in quasi-Monte Carlo importance sampling, see [1, 8]. Here we are especially interested in the behavior of the G-star and G-discrepancy with respect to the dimension d. The following Theorem is a direct consequence of [4, Theorem 4].

Theorem 1.1. *There exists a universal constant $C > 0$ such that for every d and each probability measure λ_G on $[0,1]^d$ we have for all $N \in \mathbb{N}$*

$$D^*_G(N,d) \le C\, d^{1/2} N^{-1/2} \quad \text{and} \quad D_G(N,d) \le \sqrt{2}\, C\, d^{1/2} N^{-1/2}. \tag{1}$$

This implies

$$N^*_G(\varepsilon,d) \le \lceil C^2 d\, \varepsilon^{-2} \rceil \quad \text{and} \quad N_G(\varepsilon,d) \le \lceil 2\, C^2 d\, \varepsilon^{-2} \rceil. \tag{2}$$

The dependence of the inverses of the G-star and the G-discrepancy on d in (2) can in general not be improved. This was shown by a lower bound on the inverse of the star discrepance in [4], which is, of course, also a lower bound of the extreme discrepancy. An improved lower bound was presented by A. Hinrichs in [5]. To state it in a quite general form, let us quickly recall the notion of covering numbers.

For a class $\mathcal{K}$ of measurable subsets of $[0,1]^d$ we define the pseudo-metric $d_{\lambda_G} = d_G$ by $d_G(K,K') = \lambda_G(K \triangle K')$ for all $K, K' \in \mathcal{K}$; here $K \triangle K'$ denotes the symmetric difference $(K \setminus K') \cup (K' \setminus K)$ of K and K'. The *covering number* $N(G;\mathcal{K},\varepsilon)$ is the smallest number of closed ε-balls $B_{\mathcal{K}}(G;K,\varepsilon) = \{K' \in \mathcal{K} \,|\, d_G(K,K') \le \varepsilon\}$ that cover $\mathcal{K}$.

Theorem 4 from [5] gives us in particular:

Theorem 1.2. *Let λ_G be an arbitrary probability measure on $[0,1]^d$. Assume that there exists a constant $\kappa \ge 1$ such that*

$$N(G;\mathcal{C}_d,\varepsilon) \ge (\kappa\varepsilon)^{-d} \quad \text{for all } \varepsilon \in (0,1]. \tag{3}$$

Then there exist constants $c, \varepsilon_0 > 0$ such that

$$D_G^*(N,d) \ge \min\{\varepsilon_0, cd/N\} \quad \text{for all } N \in \mathbb{N} \tag{4}$$

and

$$N_G^*(\varepsilon,d) \ge cd/\varepsilon \quad \text{for all } \varepsilon \in (0,\varepsilon_0). \tag{5}$$

We want to stress that the constants c, ε_0 do not depend on the dimension d. Indeed, following the proof of [5, Theorem 4] one observes that (4) and (5) hold for the choice $c = \varepsilon_0 = 1/4\kappa\mathrm{e}$.

2 A Result on Covering Numbers

The question that arises here is for which measures λ_G do we have a lower bound on the covering number of the form (3)?

If λ_G is a discrete measure of, e.g., the form $\lambda_G(A) = (1/M)\sum_{i=1}^M 1_A(p^i)$ for $P = \{p^1,\dots,p^M\} \subset [0,1]^d$, then $N(G;\mathcal{C}_d,\varepsilon) \le |\{P \cap C \,|\, C \in \mathcal{C}_d\}| =: \mu_P \le 2^M$. (Due to Sauer's Lemma, see, e.g., [13, Sect.2.6], one gets actually even the better bound $\mu_P \le \sum_{j=0}^d \binom{M}{j}$.) Indeed, if one chooses sets $C_1,\dots C_{\mu_P} \in \mathcal{C}_d$ such that $|\{P \cap C_i \,|\, i = 1,\dots,\mu_P\}| = \mu_P$, then the ε-balls $B_{\mathcal{C}_d}(G;C_i,\varepsilon)$, $i = 1,\dots,\mu_P$, cover $\mathcal{C}_d$. Hence in this case (3) does obviously not hold. On the other hand we get in this case $D_G^*(kM,d) = 0$ for all $k \in \mathbb{N}$, showing that the bound (4) does also not hold. In general, discrete measures do imply less interesting integration problems, since the integral itself is already a cubature formula. Thus approximating the integral via a cubature is trivial (at least if we know the discrete measure explicitly).

So let us confine to another class of distribution functions $\mathcal{G}$, which is interesting for the mixed sequences we want to study in the next section. A distribution function G is in $\mathcal{G}$ if we have for $i = 1,\dots,d$ integrable functions $g_i : [0,1] \to [0,\infty)$ satisfying

$\int_0^1 g_i(t)\,\mathrm{d}t = 1$, and if $G_i(x) = \int_0^x g_i(t)\,\mathrm{d}t$ and $G(x) = G_1(x_1)\dots G_d(x_d)$ for all $x \in [0,1]^d$. In particular, all distribution functions in $\mathcal{G}$ are continuous functions on $[0,1]^d$.

For this restricted class of measures the covering number does not depend on the particular choice of G:

Proposition 2.1. *For all $d \in \mathbb{N}$ and all $G \in \mathcal{G}$ we have*

$$N(G;\mathcal{C}_d,\varepsilon) = N(\mathcal{C}_d,\varepsilon) \quad \textit{for all } \varepsilon \in (0,1]. \tag{6}$$

In particular, we have

$$N(G;\mathcal{C}_d,\varepsilon) \geq (8\mathrm{e}\varepsilon)^{-d} \quad \textit{for all } \varepsilon \in (0,1]. \tag{7}$$

Before starting with the proof, let us define the *generalized inverse functions*

$$\underline{G}_i(t) = \inf\{s \in [0,1] \,|\, G_i(s) = t\} \quad \text{and} \quad \overline{G}_i(t) = \sup\{s \in [0,1] \,|\, G_i(s) = t\}; \tag{8}$$

these functions are well defined since the mapping $G_i : [0,1] \to [0,1]$ is surjective. Put $\underline{G}(x) = (\underline{G}_1(x_1),\dots,\underline{G}_d(x_d))$ and $\overline{G}(x) = (\overline{G}_1(x_1),\dots,\overline{G}_d(x_d))$ for all $x \in [0,1]^d$. Let us consider functions G_i^- with $\underline{G}_i \leq G_i^- \leq \overline{G}_i$. As $\underline{G}_i$ and $\overline{G}_i$ also G_i^- is a right inverse of G_i, i.e., we have $G_i(G_i^-(t)) = t$ for all $t \in [0,1]$. (Observe that any right inverse of G_i has necessarily to be of the form G_i^-.) Notice that if G_i is injective, then there exists an inverse G_i^{-1} and we obviously have $G_i^{-1} = \underline{G}_i = \overline{G}_i = G_i^-$. Put $\widehat{G}(x) = (G_1(x_1),\dots,G_d(x_d))$ and $\widehat{G}^-(x) = (G_1^-(x_1),\dots,G_d^-(x_d))$. We have $\widehat{G}(\widehat{G}^-(x)) = x$ for all $x \in [0,1]^d$.

Proof of Proposition 2.1. For general $\zeta \in [0,1]^d$ and $z = \widehat{G}(\zeta)$ we have

$$\lambda_G([0,\zeta)) = G(\zeta) = G_1(\zeta_1)\dots G_d(\zeta_d) = z_1 \dots z_d = \lambda_d([0,z)). \tag{9}$$

Now let $x^1,\dots,x^m \in [0,1]^d$ be such that the ε-balls $B_{\mathcal{C}_d}([0,x^i),\varepsilon)$, $i = 1,\dots,m$, cover $\mathcal{C}_d$. Then also the ε-balls $B_{\mathcal{C}_d}(G;[0,\widehat{G}^-(x^i)),\varepsilon)$, $i = 1,\dots,m$, cover $\mathcal{C}_d$. To see this, let $\eta \in [0,1]^d$ be arbitrary, and let $y = \widehat{G}(\eta)$. Then we find an $i \in \{1,\dots,m\}$ such that $[0,y) \in B_{\mathcal{C}_d}([0,x^i),\varepsilon)$, which is equivalent to $\lambda_d([0,y)\triangle[0,x^i)) \leq \varepsilon$. Now

$$\begin{aligned}\lambda_d([0,y)\triangle[0,x^i)) =& \lambda_d([0,y)\setminus[0,x^i)) + \lambda_d([0,x^i)\setminus[0,y))\\ =& \lambda_d([0,y)) + \lambda_d([0,x^i)) - 2\lambda_d([0,y\wedge x^i)),\end{aligned}$$

where $(y \wedge x^i)_j = \min\{y_j, x^i_j\}$. Similarly,

$$\lambda_G([0,\eta)\triangle[0,\widehat{G}^-(x^i))) = \lambda_G([0,\eta)) + \lambda_G([0,\widehat{G}^-(x^i))) - 2\lambda_G([0,\eta\wedge\widehat{G}^-(x^i))).$$

Since $y = \widehat{G}(\eta)$, $x^i = \widehat{G}(\widehat{G}^-(x^i))$, and $y \wedge x^i = \widehat{G}(\eta \wedge \widehat{G}^-(x^i))$, we get from (9)

$$\lambda_G([0,\eta)\triangle[0,\widehat{G}^-(x^i))) = \lambda_d([0,y)\triangle[0,x^i)) \leq \varepsilon,$$

implying $[0,\eta) \in B_{\mathcal{C}_d}(G;[0,\widehat{G}^-(x^i)),\varepsilon)$. Thus we have shown that $N(\mathcal{C}_d,\varepsilon) \geq N(G;\mathcal{C}_d,\varepsilon)$. Similarly, if we have $\xi^1,\dots,\xi^k \in [0,1]^d$ such that the ε-balls $B_{\mathcal{C}_d}(G;[0,\xi^i),\varepsilon)$, $i = 1,\dots,k$, cover $\mathcal{C}_d$, then also the ε-balls $B_{\mathcal{C}_d}([0,\widehat{G}(\xi^i)),\varepsilon)$, $i = 1,\dots,k$, cover $\mathcal{C}_d$. This gives $N(\mathcal{C}_d,\varepsilon) \leq N(G;\mathcal{C}_d,\varepsilon)$.

The bound (7) was established by A. Hinrichs for $N(\mathcal{C}_d,\varepsilon)$ in the course of the proof of [5, Theorem 2]. Identity (6) shows that it is true for general $G \in \mathcal{G}$. $\square$

Corollary 2.2. *Let $d \in \mathbb{N}$, and let $G \in \mathcal{G}$. Then the bounds (4) und (5) hold for $c = \varepsilon_0 = 1/32\mathrm{e}^2$.*

3 Mixed Sequences

Let $d, d', d'' \in \mathbb{N}$ with $d = d' + d''$. Let $G \in \mathcal{G}$, and let G', G'' be the uniquely determined distribution functions on $[0,1]^{d'}$ and $[0,1]^{d''}$ respectively that satisfy $G(x) = G'(x')G''(x'')$ for all $x = (x', x'') \in [0,1]^{d'} \times [0,1]^{d''}$. Furthermore, let $q = (q_k)$ be a deterministic sequence in $[0,1)^{d'}$, and let $X = (X_k)$ be a sequence of independent G''-distributed random vectors in $[0,1)^{d''}$. The resulting d-dimensional sequence $m = (m_k) = (q_k, X_k)$ is called a *mixed sequence*. (Other authors call m a mixed sequence if q is a low-discrepancy sequence. Since the latter property is irrelevant for the proof of our main result on mixed sequences, Theorem 3.5, we do not require it here.)

Compared to pure Monte Carlo or quasi-Monte Carlo methods, mixed sequences based on low-discrepancy sequences showed a favorable performance in many numerical experiments. They were introduced by J. Spanier in [12] and studied further in several papers, see, e.g., [3, 6, 7, 9, 11]. Probabilistic bounds for the star discrepancy of mixed sequences were provided in [3, 9], and deterministic bounds were derived in the recent paper [6]. In this section we want to provide probabilistic bounds for the G-star and G-discrepancy of m.

3.1 An Error Correction

Such a result was published in [11, Theorem 8], which states

If q satisfies $D_{G',N}(q) \leq O(\log(N)^{d'}/N)$ for $N \geq 2$, then we have for all $\varepsilon > 0$

$$\mathbb{P}\big(D_{G,N}(m) \leq \varepsilon + D_{G',N}(q)\big) \geq 1 - \frac{1}{\varepsilon^2}\frac{1}{4N}(D_{G',N}(q) + 1). \tag{10}$$

We will show now that this result is unfortunately incorrect, regardless of the distribution function $G \in \mathcal{G}$: Assume (10) is true. Let $\varepsilon \in (0, 1/2]$ be given. Choose $d' = 1$. There exists a $C > 0$ such that $D_{G',N}(q) \leq C\log(N)/N$ for $N \geq 2$. Let us consider N large enough to satisfy $C\log(N)/N < \varepsilon/2$. Recall that $D_{G',N}(q) \leq 1$. Then (10) implies for the mixed sequence m

$$\begin{aligned}\mathbb{P}\big(D_{G,N}(m) \leq \varepsilon\big) \geq &\,\mathbb{P}\big(D_{G,N}(m) \leq (\varepsilon - C\log(N)/N) + D_{G',N}(q)\big)\\ \geq &\,1 - \frac{1}{(\varepsilon - C\log(N)/N)^2}\frac{1}{2N} \geq 1 - \frac{2}{N\varepsilon^2}.\end{aligned}$$

It is easy to see that there exists a constant K independent of d and ε such that for all $N \geq K\varepsilon^{-2}$ the inequalities $C\log(N)/N < \varepsilon/2$ and $P(D_{G,N}(m) \leq \varepsilon) > 0$ hold. Hence, for each $d \in \mathbb{N}$ and each $N \geq K\varepsilon^{-2}$ there exists in particular an N-point set $P \subset [0,1)^d$ such that $D_{G,N}(P) \leq \varepsilon$. Thus $N_G(\varepsilon, d) \leq \lceil K\varepsilon^{-2}\rceil$ for all $d \in \mathbb{N}$, contradicting Corollary 2.2, which says that for $\varepsilon < 1/32\mathrm{e}^2$ we have $N_G(\varepsilon, d) \geq N_G^*(\varepsilon, d) \geq d/32\mathrm{e}^2\varepsilon$.

Since the statement of [11, Theorem 8] is incorrect, the proof has to be incorrect, too. The source of the incorrectness is [11, Lemma 7] whose statement is ambigious (an all-quantifier "$\forall J$" is missing in some place in the if-condition) and which is used for an inadmissable conclusion in the proof of [11, Theorem 8].

Remark 3.1. In [11, Corollary 6] it was shown for a mixed sequence m that

$$\operatorname{Var}\left(\frac{1}{N}\sum_{k=1}^{N} 1_J(m_k) - \lambda_G(J)\right) \le \frac{1}{4N}(D_{G',N}(q)+1),$$

where $J \in \mathcal{R}_d$ and the variance is taken with respect to the random variables $X_1, X_2, \ldots$. In fact a better result holds, namely

$$\operatorname{Var}\left(\frac{1}{N}\sum_{k=1}^{N} 1_J(m_k) - \lambda_G(J)\right) \le \frac{1}{4N}.$$

To realize this, one has to estimate the sum $\sum_{k=1}^{N} 1_{J'}(q_k)/N$ appearing in the course of the proof of [11, Corollary 6] by 1. The same is true for the analogous estimate in [7, Lemma 1].

3.2 New Probabilistic Bounds

Now we derive correct probabilistic bounds for the G-star and G-discrepancy of a mixed sequence, which additionally exhibit a better asymptotical behavior in N than the incorrect bound (10). First we define G-δ-covers and prove two helpful results about them.

Definition 3.2. Let $\delta \in (0,1]$, and let $\mathcal{K}_d \in \{\mathcal{C}_d, \mathcal{R}_d\}$. A finite set $\Gamma \subset \mathcal{K}_d$ is called a *G-δ-cover of $\mathcal{K}_d$* if for all $B \in \mathcal{K}_d$ there are $A, C \in \Gamma \cup \{\emptyset\}$ such that $A \subseteq B \subseteq C$ and $\lambda_G(C) - \lambda_G(A) \le \delta$. We put $\mathcal{N}(G; \mathcal{K}_d, \delta) := \min\{|\Gamma| \,|\, \Gamma \text{ is a } G\text{-}\delta\text{-cover of } \mathcal{K}_d\}$.

In the case where $\mathcal{K}_d = \mathcal{C}_d$, it is convenient to identify the elements $[0,x)$ of $\mathcal{C}_d$ with their upper right edge points x. Following this convention, we view δ-covers of $\mathcal{C}_d$ as finite subsets of $[0,1]^d$.

Proposition 3.3. *For all $d \in \mathbb{N}$ and all $G \in \mathcal{G}$ we have*

$$\mathcal{N}(G; \mathcal{C}_d, \delta) = \mathcal{N}(\mathcal{C}_d, \delta) \quad \textit{and} \quad \mathcal{N}(G; \mathcal{R}_d, \delta) = \mathcal{N}(\mathcal{R}_d, \delta) \quad \textit{for all } \delta \in (0,1]. \tag{11}$$

Proof. Let Γ be a δ-cover of $\mathcal{C}_d$. Put $\Gamma_G := \{\overline{G}(x) \,|\, x \in \Gamma\}$. Then $|\Gamma_G| = |\Gamma|$ (since $\overline{G}$ is injective). Let $\eta \in [0,1)^d$ be given and $y = \widehat{G}(\eta)$. There exist x ,$z \in \Gamma \cup \{0\}$ with $x \le y \le z$ and $\lambda_d([0,z)) - \lambda_d([0,x)) \le \delta$. If $y_i \neq 0$ for all i, then it is easy to see that we can choose x and z in such a way that $x_i < y_i$ for all i. (Consider for ε sufficiently small vectors $y_{-\varepsilon} := (y_1 - \varepsilon, \ldots, y_d - \varepsilon)$. Since Γ is finite, we find $x, z \in \Gamma \cup \{0\}$ with $\lambda_d([0,z)) - \lambda_d([0,x)) \le \delta$ such that for infinitely many $\varepsilon > 0$ we have $y_{-\varepsilon} \in [x,y]$. Hence $y \in [x,y]$ and $x_i < y_i$ for all i.) Since $\widehat{G}$ is non-decreasing with respect to each component, we have $\overline{G}(x) \leq \eta \leq \overline{G}(z)$. Furthermore, $\lambda_G([0,\overline{G}(z))) - \lambda_G([0,\overline{G}(x))) = \lambda_d([0,z)) - \lambda_d([0,x)) \leq \delta$. If $y_i = 0$ for some i, then $\lambda_d([0,x)) = 0$. Thus we get $0 \leq \eta \leq \overline{G}(z)$ and $\lambda_G([0,\overline{G}(z))) - \lambda_G([0,0)) = \lambda_G([0,\overline{G}(z))) = \lambda_d([0,z)) \leq \delta$. Hence Γ_G is a G-δ-cover of $\mathcal{C}_d$. Thus $\mathcal{N}(G;\mathcal{C}_d,\delta) \leq \mathcal{N}(\mathcal{C}_d,\delta)$.

Let now Γ_G be an arbitrary G-δ-cover of $\mathcal{C}_d$. Then it is easily verified that $\Gamma := \{\widehat{G}(\xi) \,|\, \xi \in \Gamma_G\}$ is a δ-cover of $\mathcal{C}_d$ with $|\Gamma| \leq |\Gamma_G|$. Hence $\mathcal{N}(G;\mathcal{C}_d,\delta) \geq \mathcal{N}(\mathcal{C}_d,\delta)$.

Let $\widetilde{\Gamma}$ be a δ-cover of $\mathcal{R}_d$. We may assume $\emptyset \notin \widetilde{\Gamma}$. Put $\widetilde{\Gamma}_G := \{[\underline{G}(x), \overline{G}(z)) \,|\, [x,z) \in \widetilde{\Gamma}\}$. We have $|\widetilde{\Gamma}_G| = |\widetilde{\Gamma}|$. Let $B := [\xi,\eta) \in \mathcal{R}_d \setminus \{\emptyset\}$ be given, and let $\widehat{B} := [\widehat{G}(\xi), \widehat{G}(\eta))$. Then there exist $\widehat{A}, \widehat{C} \in \widetilde{\Gamma} \cup \{\emptyset\}$ with $\widehat{A} \subseteq \widehat{B} \subseteq \widehat{C}$ and $\lambda_d(\widehat{C}) - \lambda_d(\widehat{A}) \leq \delta$. If $\widehat{B} \neq \emptyset$, we may assume that $\widehat{A} = [\underline{x}, \overline{x})$ with $\widehat{G}_i(\xi_i) < \underline{x}_i \leq \overline{x}_i < \widehat{G}_i(\eta_i)$ for all i. (Consider for ε sufficiently small boxes $\widehat{B}_\varepsilon = [\widehat{G}(\xi)_{+\varepsilon}, \widehat{G}(\eta)_{-\varepsilon})$. Since $\widetilde{\Gamma}$ is finite, we find $\widehat{A}, \widehat{C} \in \widetilde{\Gamma} \cup \{\emptyset\}$ with $\lambda_d(\widehat{C}) - \lambda_d(\widehat{A}) \leq \delta$ such that $\widehat{A} \subseteq \widehat{B}_\varepsilon \subseteq \widehat{C}$ for infinitely many $\varepsilon > 0$. These $\widehat{A}$ and $\widehat{C}$ do the job.) But then $[\underline{G}(\underline{x}), \overline{G}(\overline{x})) \subseteq [\xi,\eta)$, since $\widehat{G}$ is non-decreasing with respect to each component. In any case, we have for $\widehat{C} = [\underline{z}, \overline{z})$ that $[\xi,\eta) \subseteq [\underline{G}(\underline{z}), \overline{G}(\overline{z}))$. Furthermore,

$$\begin{aligned}
&\lambda_G([\underline{G}(\underline{z}), \overline{G}(\overline{z}))) - \lambda_G([\underline{G}(\underline{x}), \overline{G}(\overline{x}))) \\
&= \prod_{i=1}^{d} (G_i(\overline{G}_i(\overline{z}_i)) - G_i(\underline{G}_i(\underline{z}_i))) - \prod_{i=1}^{d} (G_i(\overline{G}_i(\overline{x}_i)) - G_i(\underline{G}_i(\underline{x}_i))) \\
&= \prod_{i=1}^{d} (\overline{z}_i - \underline{z}_i) - \prod_{i=1}^{d} (\overline{x}_i - \underline{x}_i) = \lambda_d(\widehat{C}) - \lambda_d(\widehat{A}) \leq \delta.
\end{aligned}$$

If $\widehat{B} = \emptyset$, then $\lambda_d(\widehat{A}) = 0$. If $\widehat{C} = [\underline{z}, \overline{z})$, we thus have $\emptyset \subseteq B \subseteq [\underline{G}(\underline{z}), \overline{G}(\overline{z}))$ and $\lambda_G([\underline{G}(\underline{z}), \overline{G}(\overline{z}))) - \lambda_G(\emptyset) \leq \delta$. This shows that $\widetilde{\Gamma}_G$ is a G-δ-cover of $\mathcal{R}_d$. Thus $\mathcal{N}(G;\mathcal{R}_d,\delta) \leq \mathcal{N}(\mathcal{R}_d,\delta)$.

Again, the second inequality is easier to show; for any G-δ-cover $\widetilde{\Gamma}_G$ of $\mathcal{R}_d$ the set $\widetilde{\Gamma} := \{[\widehat{G}(\xi), \widehat{G}(\zeta)) \,|\, [\xi,\zeta) \in \widetilde{\Gamma}_G\}$ is a δ-cover of $\mathcal{R}_d$ with $|\widehat{\Gamma}| \leq |\widetilde{\Gamma}_G|$. □

From [2, Thm. 1.15] we know that $\mathcal{N}(\mathcal{C}_d,\delta) \leq (2d)^d(\delta^{-1}+1)^d/d! \leq (2\mathrm{e})^d(\delta^{-1}+1)^d$. In [2, Lemma 1.17] the inequality $\mathcal{N}(\mathcal{R}_d,\delta) \leq \mathcal{N}(\mathcal{C}_d,\delta/2)^2$ was proved. Thus we get from Proposition 3.3

$$\mathcal{N}(G;\mathcal{C}_d,\delta) \leq (2\mathrm{e})^d(\delta^{-1}+1)^d \quad \text{and} \quad \mathcal{N}(G;\mathcal{R}_d,\delta) \leq (2e)^{2d}(2\delta^{-1}+1)^{2d}. \tag{12}$$

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

Lemma 3.4. *Let $P = \{p^1, \ldots, p^N\} \subset [0,1]^d$. Let Γ be a G-δ-cover of $\mathcal{C}_d$, then*

$$D^*_{G,N}(P) \leq D^*_{G,N,\Gamma}(P) + \delta, \quad \textit{where} \quad D^*_{G,N,\Gamma}(P) := \max_{x\in\Gamma} \left| \lambda_G([0,x)) - \frac{1}{N}\sum_{j=1}^{N} 1_{[0,x)}(p^j) \right|.$$

Let $\widetilde{\Gamma}$ be a G-δ-cover of $\mathcal{R}_d$, then

$$D_{G,N}(P) \leq D_{G,N,\widetilde{\Gamma}}(P) + \delta, \quad \textit{where} \quad D_{G,N,\widetilde{\Gamma}}(P) := \max_{R\in\widetilde{\Gamma}} \left| \lambda_G(R) - \frac{1}{N}\sum_{j=1}^{N} 1_R(p^j) \right|.$$

Proof. Let $B \in \mathcal{R}_d$. Then we find $A, C \in \widetilde{\Gamma} \cup \{\emptyset\}$ such that $A \subseteq B \subseteq C$ and $\lambda_G(C) - \lambda_G(A) \leq \delta$. Hence we get

$$\lambda_G(A) + \delta - \frac{1}{N}\sum_{j=1}^{N} 1_A(p^j) \geq \lambda_G(B) - \frac{1}{N}\sum_{j=1}^{N} 1_B(p^j) \geq \lambda_G(C) - \delta - \frac{1}{N}\sum_{j=1}^{N} 1_C(p^j).$$

From this the statement for the G-discrepancy follows. Similarly, one shows the statement for the G-star discrepancy. □

Theorem 3.5. *Let $q = (q_k)$ be a deterministic sequence in $[0,1)^{d'}$, $X = (X_k)$ be a sequence of independent and G-distributed random vectors in $[0,1)^{d''}$, and let $m = (m_k) = (q_k, X_k)$ be the resulting d-dimensional mixed sequence. Then we have for all $\varepsilon \in (0,1]$*

$$\mathbb{P}\big(D^*_{G,N}(m) - D^*_{G',N}(q) < \varepsilon\big) > 1 - 2\,\mathcal{N}(G; \mathcal{C}_d, \varepsilon/2) \exp\left(-\frac{\varepsilon^2 N}{2}\right) \tag{13}$$

and

$$\mathbb{P}\big(D_{G,N}(m) - D_{G',N}(q) < \varepsilon\big) > 1 - 2\,\mathcal{N}(G; \mathcal{R}_d, \varepsilon/2) \exp\left(-\frac{\varepsilon^2 N}{2}\right). \tag{14}$$

Let $\theta \in [0,1)$. We have with probability strictly larger than θ

$$D^*_{G,N}(m) < D^*_{G',N}(q) + \sqrt{\frac{2}{N}}\left(d\ln(\rho) + \ln\left(\frac{2}{1-\theta}\right)\right)^{1/2}, \tag{15}$$

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

$$D_{G,N}(m) < D_{G',N}(q) + \sqrt{\frac{2}{N}}\left(2d\ln(\tilde{\rho}) + \ln\left(\frac{2}{1-\theta}\right)\right)^{1/2}, \tag{16}$$

where $\tilde{\rho} = \tilde{\rho}(N,d) := 5\mathrm{e}(\max\{1\,,\, N/(\ln(5\mathrm{e})d)\})^{1/2}$.

Remark 3.6. In front of the exponential function in the bounds (13) and (14) there necessarily has to appear a function depending exponentially on the dimension d: If we have a bound of the form $\mathbb{P}\big(D^*_{G,N}(m) - D^*_{G',N}(q) \leq \varepsilon\big) \geq 1 - f(q; d, \varepsilon)\exp\left(-\frac{\varepsilon^2 N}{2}\right)$ for all ε in some interval $(0, \varepsilon_*]$, all $d > d'$ and all N (or an analoguous bound for the G-discrepancy of mixed sequences), then for a low-discrepancy sequence q and all ε sufficiently small the function $f(q; d, \varepsilon)$ has to increase at least exponentially in d. This was proved in [3, Remark 2.6] for the case of the classical star discrepancy, and the proof can easily be transfered to the case of the G-star and G-discrepancy.

Proof of Theorem 3.5. Let $R \in \mathcal{R}_d$, and let $R' \in \mathcal{R}_{d'}$, $R'' \in \mathcal{R}_{d''}$ such that $R = R' \times R''$. Put $\xi_k = \xi_k(R) := \lambda_G(R) - 1_R(m_k)$ for $k = 1, 2, \ldots$. The ξ_k, $k = 1, 2, \ldots$, are independent random variables with $\mathbb{E}(\xi_k) = \lambda_{G''}(R'')(\lambda_{G'}(R') - 1_{R'}(q_k))$. Thus we have

$$\left|\mathbb{E}\left(\frac{1}{N}\sum_{k=1}^{N}\xi_k\right)\right| = \lambda_{G''}(R'')\left|\lambda_{G'}(R') - \frac{1}{N}\sum_{k=1}^{N} 1_{R'}(q_k)\right| \leq D_{G',N}(q).$$

Let $\delta := \varepsilon/2$. Then Hoeffding's inequality (see, e.g., [10, p.191]) gives us

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{k=1}^{N}\xi_k\right| \geq D_{G',N}(q)+\delta\right) \leq \mathbb{P}\left(\left|\frac{1}{N}\sum_{k=1}^{N}(\xi_k-\mathbb{E}(\xi_k))\right| \geq \delta\right) \leq 2\exp(-2\delta^2 N). \quad (17)$$

Now let $\widetilde{\Gamma}$ be a minimal G-δ-cover of $\mathcal{R}_d$. Then Lemma 3.4 and (17) gives us

$$\begin{aligned}\mathbb{P}\big(D_{G,N}(m)-D_{G',N}(q)<\varepsilon\big) \geq & \mathbb{P}\big(D_{G,N,\widetilde{\Gamma}}(m)-D_{G',N}(q)<\delta\big)\\ = & 1-\mathbb{P}\big(D_{G,N,\widetilde{\Gamma}}(m)\geq D_{G',N}(q)+\delta\big)\\ \geq & 1-\sum_{R\in\widetilde{\Gamma}}\mathbb{P}\left(\left|\frac{1}{N}\sum_{k=1}^{N}\xi_k(R)\right| \geq D_{G',N}(q)+\delta\right)\\ > & 1-2|\widetilde{\Gamma}|\exp\left(-2\delta^2 N\right).\end{aligned}$$

(In the last estimate we obtained "$>$", since we have necessarily $I=[0,1)^d\in\widetilde{\Gamma}$ and $\xi_k(I)=0$ for all k.) This proves (14).

Now $1-2|\widetilde{\Gamma}|\exp(-2\delta^2 N)\geq\theta$ iff

$$\delta^2 \geq \frac{1}{2N}\left(\ln|\widetilde{\Gamma}|+\ln\left(\frac{2}{1-\theta}\right)\right). \quad (18)$$

Due to (12) we have $|\widetilde{\Gamma}|\leq(2e)^{2d}(2\delta^{-1}+1)^{2d}$. Therefore it is easily verified that (18) holds if we choose δ to be

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

This proves that (16) holds with probability $>\theta$.

The verification of (13) and (15) can be done with the same proof sheme and was done in [3, Theorem 3.3] for the classical star discrepancy. □

3.3 An Alternative Approach

One could prove a version of Theorem 3.5 without proving the results on G-δ-covers in Proposition 3.3 and Lemma 3.4, and proving instead the next theorem on the G-star discrepancy and a similar one for the G-discrepancy:

Theorem 3.7. *Let $d\in\mathbb{N}$, and let $G\in\mathcal{G}$. Let $P=\{p^1,\dots,p^N\}\subset[0,1)^d$. Then*

$$D^*_N(\widehat{G}(P)) \leq D^*_{G,N}(P). \quad (19)$$

Under the additional assumption $\widehat{G}(P)\subset[0,1)^d$ we get

$$D^*_N(\widehat{G}(P)) = D^*_{G,N}(P). \quad (20)$$

In particular, for all $P=\{p^1,\dots,p^N\}\subset[0,1)^d$ and a right inverse $\widehat{G}^-$ of $\widehat{G}$ as defined above we get

$$D^*_N(P) = D^*_{G,N}(\widehat{G}^-(P)). \quad (21)$$

Remark 3.8. Without the condition $\widehat{G}(P) \subset [0,1)^d$ the bound $D^*_{G,N}(P) \le D^*_N(\widehat{G}(P))$ does not necessarily hold as the following simple example reveals: Let $d = 2$, $G_1(x) = \min\{1, 2x\}$, and $G_2(x) = x$. Furthermore, let $\varepsilon \in (0, 1/2)$ and $P = \{p^1, p^2\}$ with $p^1 = (1/4, \varepsilon)$, $p^2 = (1/2, \varepsilon)$. Then $\widehat{G}(P) = \{(1/2, \varepsilon), (1, \varepsilon)\}$ and $D^*_N(\widehat{G}(P)) = 1/2 < 1 - \varepsilon = D^*_{G,N}(P)$.

Using Theorem 3.7 and [3, Theorem 3.3] one can conclude (13) and (15) under the additional assumption that $\widehat{G}'(q_k) \in [0,1)^{d'}$ for all k. This assumption seems a little bit artificial to us and is—as Theorem 3.5 shows—in fact unnecessary. Apart from that, even if one proves a result similar to Theorem 3.7 for the G-discrepancy, one still has to establish (14) and (16) for the classical extreme discrepancy.

In the case where the density function g of G is strictly positive, identity (21) is covered by the statement of [8, Lemma 4.1]. Since we only require that the density function g is non-negative and since in the proof of [8, Lemma 4.1] only the case $d = 1$ is treated explicitly, we give here a rigorous proof of Theorem 3.7

Proof of Theorem 3.7. Let us start by showing (19): Let $x \in [0,1]^d$ be given. Then $\lambda_d([0,x)) = \lambda_G([0, \underline{G}(x)))$. Let $p \in [0,1)^d$. Then $\widehat{G}(p) \in [0,x) = [0, \widehat{G}(\underline{G}(x)))$ iff $p \in [0, \underline{G}(x))$, since $\widehat{G}$ is monotone increasing with respect to every component. This gives

$$\left| \frac{1}{N} \sum_{i=1}^{N} 1_{[0,x)}(\widehat{G}(p^i)) - \lambda_d([0,x)) \right| = \left| \frac{1}{N} \sum_{i=1}^{N} 1_{[0,\underline{G}(x))}(p^i) - \lambda_G([0, \underline{G}(x))) \right|.$$

This shows that $D^*_N(\widehat{G}(P)) \le D^*_{G,N}(P)$.

Let us now additionally assume that $\widehat{G}(P) \subset [0,1)^d$ and prove $D^*_{G,N}(P) \le D^*_N(\widehat{G}(P))$: Let $\zeta \in [0,1]^d$ be given and $z = \widehat{G}(\zeta)$. Then $\lambda_G([0,\zeta)) = \lambda_d([0,z))$. Let $p \in [0,1)^d$ with $\widehat{G}(p) \in [0,1)^d$. Since $\widehat{G}(p) \in [0,z)$ implies $p \in [0,\zeta)$, we obtain

$$\lambda_G([0,\zeta)) - \frac{1}{N} \sum_{i=1}^{N} 1_{[0,\zeta)}(p^i) \le \lambda_d([0,z)) - \frac{1}{N} \sum_{i=1}^{N} 1_{[0,z)}(\widehat{G}(p^i)).$$

Since $\widehat{G}$ is not necessarily injective, we cannot conclude from $p \in [0,\zeta)$ that $\widehat{G}(p) \in [0,z)$ holds. So let us define $z_i(\varepsilon) = \min\{z_i + \varepsilon, 1\}$ for $\varepsilon > 0$ and $i = 1, \ldots, d$. For all $i \in \{1, \ldots, d\}$ we get the following: If $z_i = 1$, then $G_i(p_i) \in [0, z_i) = [0, z_i(\varepsilon))$. If $z_i < 1$, then we get from $p_i < \zeta_i$ that $G_i(p_i) \le G_i(\zeta_i) = z_i < z_i(\varepsilon)$. Thus $p \in [0,\zeta)$ always implies $\widehat{G}(p) \in [0, z(\varepsilon))$. With $\delta(\varepsilon) = \lambda_d([0, z(\varepsilon))) - \lambda_d([0,z))$ we get

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N} 1_{[0,\zeta)}(p^i) - \lambda_G([0,\zeta)) \le& \frac{1}{N} \sum_{i=1}^{N} 1_{[0,z(\varepsilon))}(\widehat{G}(p^i)) - \lambda_d([0,z)) \\ =& \frac{1}{N} \sum_{i=1}^{N} 1_{[0,z(\varepsilon))}(\widehat{G}(p^i)) - \lambda_d([0,z(\varepsilon))) + \delta(\varepsilon). \end{aligned}$$

Since we have $\lim_{\varepsilon\to 0}\delta(\varepsilon)=0$, we obtain

$$\frac{1}{N}\sum_{i=1}^{N}1_{[0,\zeta)}(p^i)-\lambda_G([0,\zeta))\le\sup_{\varepsilon>0}\left|\frac{1}{N}\sum_{i=1}^{N}1_{[0,z(\varepsilon))}(\widehat{G}(p^i))-\lambda_d([0,z(\varepsilon)))\right|.$$

These arguments establish $D^*_{G,N}(P)\le D^*_N(\widehat{G}(P))$. Thus identity (20) holds. Then identity (21) follows immediately, since $D^*_N(P)=D^*_N(\widehat{G}(\widehat{G}^-(P)))=D^*_{G,N}(\widehat{G}^-(P))$.

□

References

[1] Chelson, P.: Quasi-random techniques for Monte Carlo methods. Ph.D. Dissertation, The Claremont Graduate School, Claremont, California, 1976.

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

[3] Gnewuch, M.: On probabilistic results for the discrepancy of a hybrid-Monte Carlo sequence. Preprint 2008 (submitted).

[4] 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.

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

[6] Niederreiter, H.: On the discrepancy of some hybrid sequences. Preprint 2008, 20 pp..

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

[8] Ökten, G.: Error reduction techniques in quasi-Monte Carlo integration. Math. Comput. Modelling 30 (1999), 61–69.

[9] 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.

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

[11] 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.

[12] 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.

[13] Van der Vaart, A. W., Wellner, J. A.: Weak Convergence and Empirical Processes. Springer Series in Statistics, Springer, New York, 1996.