

Variance Bounds and Existence Results for Randomly Shifted Lattice Rules

Vasile Sinescu and Pierre L'Ecuyer

DIRO, University of Montreal, Canada

Abstract

We study the convergence of the variance for randomly shifted lattice rules for numerical multiple integration over the unit hypercube in an arbitrary number of dimensions. We consider integrands that are square integrable but whose Fourier series are not necessarily absolutely convergent. For such integrands, a bound on the variance is expressed through a certain type of weighted discrepancy. We prove existence and construction results for randomly shifted lattice rules such that the variance bounds are almost $O(n^{-\alpha})$, where n is the number of function evaluations and $\alpha > 1$ depends on our assumptions on the convergence speed of the Fourier coefficients. These results hold for general weights, arbitrary n , and any dimension. With additional conditions on the weights, we obtain a convergence that holds uniformly in the dimension, and this provides sufficient conditions for strong tractability of the integration problem. We also show that lattice rules that satisfy these bounds are not difficult to construct explicitly and we provide numerical illustrations of the behaviour of construction algorithms.

Keywords: Numerical multiple integration, quasi-Monte Carlo, lattice rules, discrepancy, random shift, variance.

2000 MSC: 65D30, 65D32, 11K38, 62J10.

Email address: vsinescu@yahoo.co.nz, lecuyer@iro.umontreal.ca (Vasile Sinescu and Pierre L'Ecuyer)

1. Introduction and summary

Integrals over the d -dimensional unit cube given by

$$I_d(f) = \int_{(0,1)^d} f(\mathbf{x}) \, d\mathbf{x}$$

can be approximated by quadrature rules of the form

$$Q_{n,d}(f) = \frac{1}{n} \sum_{k=0}^{n-1} f(\mathbf{t}_k),$$

which average function evaluations over the set of quadrature points $P_n := \{\mathbf{t}_0, \mathbf{t}_1, \dots, \mathbf{t}_{n-1}\}$. In standard Monte Carlo (MC), these points are independent and have the uniform distribution over the unit cube $(0, 1)^d$. In classical quasi-Monte Carlo (QMC) methods, the points are deterministic and they are selected to cover the unit cube very evenly, that is, so that a given (pre-specified) measure of discrepancy between their empirical distribution and the uniform distribution is smaller than for independent random points. In randomised QMC (RQMC), the points are randomised in a way that each point \mathbf{t}_k has the uniform distribution over the unit cube while the points keep their low discrepancy when taken together. The performance of QMC methods is often studied by bounding the convergence rate of the worst-case integration error as a function of n , for given classes of integrands. With RQMC, we have a noisy but unbiased estimator of $I_d(f)$, so it makes sense to assess the performance of this estimator via the convergence rate of its variance as a function of n , instead of the worst-case error. This will be our viewpoint in this paper.

Lattice rules (with a shift) are a class of QMC constructions for which the set of quadrature points is

$$P_n = (L + \mathbf{\Delta}) \cap [0, 1)^d$$

where $\mathbf{\Delta} \in [0, 1)^d$ is called the *shift*, and L is an integration lattice of density n in \mathbb{R}^d , defined as a discrete subset of \mathbb{R}^d which is closed under addition and subtraction, which contains \mathbb{Z}^d as a subset, and has n points per unit of volume in \mathbb{R}^d . When $\mathbf{\Delta} = \mathbf{0}$, we have a “plain” or “unshifted” lattice rule, which is a QMC method. If $\mathbf{\Delta}$ is random with the uniform distribution over the unit cube $(0, 1)^d$, we have a RQMC method known as a *randomly-shifted lattice rule* (RSLR). This is the type of method considered here.

In fact, in this paper, we restrict ourselves to a subclass of shifted *rank-1 lattice rules* for which P_n can be written as

$$P_n := \left\{ \left\{ \frac{k\mathbf{z}}{n} + \mathbf{\Delta} \right\}, 0 \leq k \leq n-1 \right\},$$

with *generating vector* $\mathbf{z} \in \mathcal{Z}_n^d$, where $\mathcal{Z}_n := \{z \in \{1, 2, \dots, n-1\} : \gcd(z, n) = 1\}$, and where the braces around a vector indicate that we take only the fractional part of each component (this is the “modulo 1” operator). In this case, the *dual lattice* to L is given by

$$L^\perp = \{\mathbf{h} \in \mathbb{Z}^d : \mathbf{h} \cdot \mathbf{z} \equiv 0 \pmod{n}\}.$$

More details on lattice rules can be found in [14] and [21].

Suppose that f has the Fourier series representation

$$f(\mathbf{x}) = \sum_{\mathbf{h} \in \mathbb{Z}^d} \hat{f}(\mathbf{h}) e^{2\pi i \mathbf{h} \cdot \mathbf{x}},$$

with Fourier coefficients

$$\hat{f}(\mathbf{h}) = \int_{(0,1)^d} f(\mathbf{x}) e^{-2\pi i \mathbf{h} \cdot \mathbf{x}} d\mathbf{x}.$$

The following proposition, proved in [11], tells us that the RSLR yields an unbiased estimator regardless of the choice of lattice, and it provides an explicit expression for the variance in terms of the dual lattice and the Fourier coefficients of f . As pointed out in [11], the Fourier series does not have to be absolutely convergent.

Proposition 1. *Suppose that f is square integrable. With either MC or a RSLR, we have $\mathbb{E}[Q_{n,d}(f)] = I_d(f)$. With MC, the variance is*

$$\text{Var}[Q_{n,d}(f)] = \frac{\sigma^2}{n} = \frac{1}{n} \sum'_{\mathbf{h} \in \mathbb{Z}^d} |\hat{f}(\mathbf{h})|^2,$$

where $\sigma^2 = \int_{(0,1)^d} f^2(\mathbf{x}) d\mathbf{x} - I_d^2(f)$ and the ' in the sum indicates that we omit the $\mathbf{h} = \mathbf{0}$ term, whereas with a RSLR with integration lattice L , it is

$$\text{Var}[Q_{n,d}(f)] = \sum'_{\mathbf{h} \in L^\perp} |\hat{f}(\mathbf{h})|^2. \quad (1)$$

Ideally, for any given function f , we would like to find a lattice L that minimises the variance expression (1). This suggests figures of merit, or measures of “discrepancy” (for P_n), of the form

$$\sum'_{\mathbf{h} \in L^\perp} w(\mathbf{h}), \quad (2)$$

where the *weights* $w(\mathbf{h})$ are chosen in correspondence with the class of functions f that we want to consider. (Here the quotes around “discrepancy” reflect the fact that (2) may not represent a natural measure of discrepancy from the uniform distribution.) As noted in [10, 11], this discrepancy (2) provides an obvious bound on the RSLR variance for all functions f whose Fourier coefficients satisfy $|\hat{f}(\mathbf{h})|^2 \leq w(\mathbf{h})$. Giving an arbitrary weight $w(\mathbf{h})$ to each vector \mathbf{h} as in (2), seems to be the most general way to assign those weights. However, finding optimal weights at that level of generality is impractical, because it would require knowledge of all the Fourier coefficients of f and there are infinitely many. Moreover, given a selection of weights $w(\mathbf{h})$ and a parameter $\alpha > 1$ that controls the decay of the weights, a key question of interest is whether we can construct sequences of lattices indexed by n so that the corresponding discrepancy (2) converges as $O(n^{-\alpha})$.

This last question is easier to study for a slightly more restrictive class of weights defined as follows. For each subset of coordinates (or projection) $\mathbf{u} \subseteq \mathcal{D} = \{1, \dots, d\}$, we select a *projection-dependent* weight $\gamma_{\mathbf{u}} \geq 0$. Such weights have been introduced in [7, 22] together with the concept of “weighted spaces of functions”. More precisely, [22] considered only the special case of *product weights*, defined by $\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \gamma_j$ for all \mathbf{u} , for some positive constants $\gamma_1, \dots, \gamma_d$, whereas [7] introduced the more general weights considered here. Note that these projection-dependent weights are not the same as the weights $w(\mathbf{h})$ in (2) and that for the rest of the paper, by “weights”, we mean the projection-dependent weights $\gamma_{\mathbf{u}}$. Then for each $\mathbf{h} = (h_1, \dots, h_d) \in \mathbb{Z}^d$, we take

$$w(\mathbf{h}) = \gamma_{\mathbf{u}(\mathbf{h})} \prod_{j \in \mathbf{u}(\mathbf{h})} |h_j|^{-\alpha}$$

where $\mathbf{u}(\mathbf{h}) = \mathbf{u}(h_1, \dots, h_d)$ is the set of indices j for which $h_j \neq 0$, and $\alpha > 1$ is a given constant. These types of projection-dependent weights have also been adopted earlier by several authors, under the name “general weights”

[4, 18]. With these weights, the discrepancy (2) becomes

$$D_{n,d,\gamma}(\mathbf{z}, \alpha) := \sum_{\mathbf{u} \subseteq \mathcal{D}} \sum_{\{\mathbf{h} \in L^\perp : \mathbf{u}(\mathbf{h}) = \mathbf{u}\}} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} |h_j|^{-\alpha}. \quad (3)$$

This weighted discrepancy, which turns out to be a square worst-case error as explained later, provides a bound on the RSLR variance for the class of functions f whose Fourier coefficients satisfy

$$|\hat{f}(\mathbf{h})| \leq \gamma_{\mathbf{u}(\mathbf{h})}^{1/2} \prod_{j \in \mathbf{u}(\mathbf{h})} |h_j|^{-\alpha/2}. \quad (4)$$

One motivation for adopting projection-dependent weights $\gamma_{\mathbf{u}}$ is that these weights can be selected by “matching” the variance components $\sigma_{\mathbf{u}}^2$ in a functional ANOVA decomposition of f . This decomposition writes f as

$$f(\mathbf{x}) = \sum_{\mathbf{u} \subseteq \mathcal{D}} f_{\mathbf{u}}(\mathbf{x}),$$

where $f_{\emptyset} = I_d(f)$, the other $f_{\mathbf{u}}$ are orthogonal and have mean zero, and if we denote $\sigma_{\mathbf{u}}^2 = \int_{[0,1]^{|\mathbf{u}|}} f_{\mathbf{u}}^2(\mathbf{x}) \, d\mathbf{x}$, the MC variance of $f_{\mathbf{u}}$, then the total MC variance has the corresponding decomposition $\sigma^2 = \sum_{\mathbf{u} \subseteq \mathcal{D}} \sigma_{\mathbf{u}}^2$ [13, 23]. The variance components $\sigma_{\mathbf{u}}^2$ can be estimated by MC techniques described in [23] and the weights $\gamma_{\mathbf{u}}$ can be selected as increasing functions of these $\sigma_{\mathbf{u}}^2$, as suggested in [24] and [12], for example.

We prove that for any $\alpha > 1$, any β satisfying $1 \leq \beta < \alpha$, and any fixed dimension d , regardless of the choice of weights $\gamma_{\mathbf{u}}$, for each $n \geq 3$ there exists a generating vector $\mathbf{z}^* = \mathbf{z}^*(n)$ such that

$$D_{n,d,\gamma}(\mathbf{z}^*, \alpha) \leq \kappa^\beta C(\alpha, \beta, d, \gamma) n^{-\beta} (\log \log n)^\beta, \quad (5)$$

where κ is an absolute constant and

$$C(\alpha, \beta, d, \gamma) := \left(\sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|\mathbf{u}|} \right)^\beta, \quad (6)$$

where $\zeta(\alpha) := \sum_{h=1}^{\infty} h^{-\alpha}$ denotes the Riemann zeta function. The constant $C(\alpha, \beta, d, \gamma)$ does not depend on n but it may be unbounded in d , depending on the choice of weights. Under the additional condition that

$$C(\alpha, \beta, d, \gamma) \leq C(\alpha, \beta, \gamma) \quad \text{for all } d \geq 1, \quad (7)$$

for some constant $C(\alpha, \beta, \gamma)$ that does not depend on d , the bound in (5) becomes uniform in the dimension d .

We also provide algorithms that provably construct (either in a deterministic or in a probabilistic sense) vectors \mathbf{z} that satisfy the above conditions for any given d , α , $1 \leq \beta < \alpha$, weights $\gamma_{\mathbf{u}}$, and n , by evaluating the expression (3) for only a small number of vectors \mathbf{z} . These construction methods include the well-known component-by-component (CBC) technique used in [4, 9, 18, 19, 20] and a few randomised versions similar to those used in [19, 24]. Under the assumption that (3) can be evaluated efficiently, we show that finding vectors \mathbf{z} that satisfy the bound (5) is quite easy.

Our assumptions on the integrand are not much stronger than square integrability, which is a minimal smoothness assumption even for standard MC. The variance expression in Proposition 1, proved in [11], also holds in such generality. However [11] gives no result on the existence and construction of good randomly shifted lattice rules. The purpose of our paper is to provide such results. It turns out that (3) has the same expression as the square worst-case error in weighted Korobov spaces considered in [4], with the difference that we assume $\alpha > 1$ in (4), while $\alpha \geq 2$ was assumed in [4] and in other references that proved existence results. The class of integrands considered here is much larger than in [3, 4, 8, 9], where integrands are assumed to be in certain reproducing kernel Hilbert spaces such as Korobov spaces of periodic functions, or in Sobolev spaces with square integrable partial mixed derivatives. We also relax the assumptions on the functions considered in [18, 20] where the integrands were assumed to have integrable partial mixed first derivatives.

In particular, our results cover integrands that may have discontinuities and we make no assumption on their derivatives. These relaxations are important from a practical viewpoint, since many integrands encountered in practice are not smooth. For example, the expected payoff of a barrier options in finance [13], or the probability that the completion of a project exceeds a given time limit when its components have random durations [11], or the probability that more than 20% of the received calls in one month of operation of a call centre wait more than 30 seconds [2], are all integrals of discontinuous functions. Since our bounds on the variance are expressed via the worst-case error in certain Korobov spaces, the results already proved in [3, 4, 8, 9] also hold here, but we add new knowledge to those results. Our results cover the combination of arbitrary (non-prime) number of points n , general projection-dependent weights, and random shift, which has not been

considered before. For instance, the results of [4] are for general projection-dependent weights, but only for prime n and unshifted lattice rules. However, since the space of functions considered in [4] is the weighted Korobov space with shift-invariant kernel, adding a shift does not affect the discrepancy (3). Other results in [3, 8, 9] were developed only for the product weights mentioned earlier. The results in these papers also used a second sequence of weights and their derivation involve the number of prime factors of n , which is not needed here. We will show that the results in those previous papers are particular cases of ours, under no more restrictive assumptions. Our results are also presented here in a form that is easier to follow. Our discrepancy bounds also differ from previous ones and are significantly smaller in some situations, for reasonable values of α and n . It is known from [3] that in a weighted Korobov space of functions whose Fourier coefficients satisfy (4) for some $\alpha > 2$, there exist rank-1 lattice rules for which the square worst-case error converges as $O(n^{-\alpha}(\log n)^{d\alpha})$ as a function of n . In this paper, by bounding the variance instead of the worst-case error, we can extend the known results in Korobov spaces for $\alpha \geq 2$ by covering the case where $1 < \alpha \leq 2$, that is, situations where the Fourier series associated with the integrand is not absolutely convergent. Our bounds replace the $O(n^{-\alpha}(\log n)^{d\alpha})$ expression by $O(n^{-\beta}(\log \log n)^\beta)$ for any $1 \leq \beta < \alpha$, with β arbitrarily close to α .

The remainder of the paper is organised as follows. The main theoretical results are presented and proved in Sections 2 and 3. These results concern the existence of good lattice rules, the analysis of the convergence of the figure of merit and the construction of lattice rules that are good with respect to the figure of merit. We then illustrate the empirical performance of the construction methods on a typical example.

2. Existence and convergence results

Existence results for good lattice rules with respect to a certain figure of merit are usually proved by an averaging argument; see for instance [4, 9, 18, 20]. We will use this type of argument, and for that purpose we consider the average of the quantities (3) over all possible generating vectors \mathbf{z} . Of course, there has to be a generating vector \mathbf{z} that produces a discrepancy not bigger than the average. Then, to analyse the convergence of the quadrature error for a better than average \mathbf{z} , we will prove that the average is bounded by the right-hand side of (5).

We now proceed to bound the average discrepancy. We first expand the inside sum in the right side of (3) as

$$\sum_{\{\mathbf{h} \in L^\perp : \mathbf{u}(\mathbf{h}) = \mathbf{u}\}} \prod_{j \in \mathbf{u}} |h_j|^{-\alpha} = \frac{1}{n} \sum_{k=0}^{n-1} \prod_{j \in \mathbf{u}} \left(\sum'_{h \in \mathbb{Z}} \frac{e^{2\pi i h k z_j / n}}{|h|^\alpha} \right), \quad (8)$$

which follows from [21, Theorem 2.8] applied to the function

$$g_{\mathbf{u}}(\mathbf{x}) := \prod_{j \in \mathbf{u}} \left(\sum'_{h \in \mathbb{Z}} \frac{e^{2\pi i h x_j}}{|h|^\alpha} \right).$$

By using (8) in (3), we obtain

$$D_{n,d,\gamma}(\mathbf{z}, \alpha) = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} \left(\sum'_{h \in \mathbb{Z}} \frac{e^{2\pi i h k z_j / n}}{|h|^\alpha} \right). \quad (9)$$

The average of this discrepancy (9) over all admissible vectors \mathbf{z} is

$$M_{n,d,\gamma}(\alpha) := \frac{1}{\varphi(n)^d} \sum_{\mathbf{z} \in \mathbb{Z}_n^d} D_{n,d,\gamma}(\mathbf{z}, \alpha), \quad (10)$$

where $\varphi(n)$ denotes the Euler totient function of n . For prime n , we have an exact formula for this average, also established in [4]:

$$M_{n,d,\gamma}(\alpha) = \frac{1}{n} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}} (2\zeta(\alpha))^{\|\mathbf{u}\|} + \frac{n-1}{n} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}} (W(\alpha))^{\|\mathbf{u}\|},$$

where

$$W(\alpha) = -\frac{2\zeta(\alpha)(1 - n^{1-\alpha})}{n-1}.$$

If the weights have a product form, that is, $\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \gamma_j$, where $\gamma_j \geq 0$ is a weight associated with coordinate j for each j , then the average for prime n is given by

$$M_{n,d,\gamma}(\alpha) = \frac{1}{n} \prod_{j=1}^d (1 + 2\gamma_j \zeta(\alpha)) + \frac{n-1}{n} \prod_{j=1}^d (1 + \gamma_j W(\alpha)) - 1.$$

For non-prime n , no closed form formula for the average is available, but we establish an upper bound for it.

Theorem 1. For any $n \geq 2$, any dimension $d \geq 1$, and any given $\alpha > 1$, we have

$$M_{n,d,\gamma}(\alpha) \leq \overline{M}_{n,d,\gamma}(\alpha) := \frac{1}{\varphi(n)} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}(2\zeta(\alpha))^{|\mathbf{u}|}. \quad (11)$$

Proof. Expanding the average (10) using (9), as in [4] but with the difference that here n is an arbitrary positive integer, we obtain

$$M_{n,d,\gamma}(\alpha) = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}(T_{n,\alpha}(k))^{|\mathbf{u}|}, \quad (12)$$

where

$$T_{n,\alpha}(k) := \frac{1}{\varphi(n)} \sum_{z \in \mathcal{Z}_n} \sum_{h \in \mathbb{Z}}' \frac{e^{2\pi i h k z / n}}{|h|^\alpha}. \quad (13)$$

From [5] and [9, Lemmas 2.1 and 2.2], we obtain

$$\sum_{k=0}^{n-1} (T_{n,\alpha}(k))^{|\mathbf{u}|} \leq \frac{n}{\varphi(n)} (2\zeta(\alpha))^{|\mathbf{u}|}, \quad (14)$$

for any subset $\mathbf{u} \subseteq \mathcal{D}$. By using (14) in (12), we obtain

$$M_{n,d,\gamma}(\alpha) \leq \frac{1}{\varphi(n)} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}(2\zeta(\alpha))^{|\mathbf{u}|},$$

which is the desired (11). □

When n is prime, this gives the same bound as in [4], namely

$$M_{n,d,\gamma}(\alpha) \leq \frac{1}{n-1} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}(2\zeta(\alpha))^{|\mathbf{u}|}.$$

If $d = 1$, then for any $z \in \mathcal{Z}_n$, it follows easily from (3) that

$$D_{n,1,\gamma}(z, \alpha) = \frac{2\gamma_{\{1\}}\zeta(\alpha)}{n^\alpha}. \quad (15)$$

Of course, there must be at least one vector \mathbf{z} as good as the average, and therefore at least one generating vector \mathbf{z} whose discrepancy does not exceed the bound given by Theorem 1.

It is known from [3] that for any fixed $\alpha > 1$, there is at least one vector \mathbf{z} for which the expression (3) converges as $O(n^{-\alpha}(\log n)^{d\alpha})$ when $n \rightarrow \infty$, and that the exponent $-\alpha$ in $n^{-\alpha}$ is optimal. We can apply this result here and it provides a bound on the convergence rate of the discrepancy for the best $\mathbf{z} = \mathbf{z}(n)$, as a function of n . In the next theorem, we follow a different path and obtain a different bound.

Theorem 2. *Let $\alpha > 1$ be fixed. For any dimension $d \geq 1$ and integer $n \geq 3$, there exists a vector $\mathbf{z}^* \in \mathcal{Z}_n^d$ such that for any β satisfying $1 \leq \beta < \alpha$, we have*

$$D_{n,d,\gamma}(\mathbf{z}^*, \alpha) \leq C(\alpha, \beta, d, \gamma) \left(\frac{\kappa \log \log n}{n} \right)^\beta, \quad (16)$$

where $\kappa > 0$ is an absolute constant and

$$C(\alpha, \beta, d, \gamma) = \left(\sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|u|} \right)^\beta$$

is as given in (6). Moreover, if the weights are chosen so that condition (7) holds, that is, $C(\alpha, \beta, d, \gamma) \leq C(\alpha, \beta, \gamma)$, then the bound (16) is also uniform in d .

Proof. We will use Jensen's inequality [6, Theorem 19, p. 28], which states that for arbitrary non-negative numbers a_i with $i = 1, 2, \dots$ and $0 < t < s$, we have

$$\left(\sum_i a_i^s \right)^{1/s} \leq \left(\sum_i a_i^t \right)^{1/t}. \quad (17)$$

By taking a β such that $1 \leq \beta < \alpha$ and applying Jensen's inequality (17) in (3), we obtain

$$D_{n,d,\gamma}(\mathbf{z}, \alpha) \leq \left(\sum_{\mathbf{u} \subseteq \mathcal{D}} \sum'_{\{\mathbf{h} \in L^\perp : u(\mathbf{h}) = \mathbf{u}\}} \gamma_{\mathbf{u}}^{1/\beta} \prod_{j \in \mathbf{u}} |h_j|^{-\alpha/\beta} \right)^\beta.$$

Consider now a vector \mathbf{z}^* such that $D_{n,d,\gamma}(\mathbf{z}^*, \alpha) \leq D_{n,d,\gamma}(\mathbf{z}, \alpha)$ for all $\mathbf{z} \in \mathcal{Z}_n^d$. From the previous inequality, we have

$$D_{n,d,\gamma}(\mathbf{z}^*, \alpha) \leq D_{n,d,\gamma}(\mathbf{z}, \alpha) \leq (D_{n,d,\gamma^{1/\beta}}(\mathbf{z}, \alpha/\beta))^\beta, \quad \text{for all } \mathbf{z} \in \mathcal{Z}_n^d. \quad (18)$$

From Theorem 1, there exists a vector $\mathbf{z} \in \mathcal{Z}_n^d$ such that

$$D_{n,d,\gamma^{1/\beta}}(\mathbf{z}, \alpha/\beta) \leq M_{n,d,\gamma^{1/\beta}}(\alpha/\beta) \leq \frac{1}{\varphi(n)} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|\mathbf{u}|},$$

Combining this with (18) leads to

$$D_{n,d,\gamma}(\mathbf{z}^*, \alpha) \leq \left(\frac{1}{\varphi(n)} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|\mathbf{u}|} \right)^\beta = \frac{C(\alpha, \beta, d, \gamma)}{(\varphi(n))^\beta}, \quad (19)$$

where $C(\alpha, \beta, d, \gamma)$ is given by (6). We now use an inequality from [17], which states that

$$\frac{n}{\varphi(n) \log \log n} < e^\omega + \frac{2.50637}{(\log \log n)^2},$$

for any $n \geq 3$, where ω is the Euler-Mascheroni constant. This leads to

$$\frac{1}{\varphi(n)} \leq \left(e^\omega + \frac{2.50637}{(\log \log n)^2} \right) \frac{\log \log n}{n}.$$

Clearly, the expression in parentheses decreases as n increases and therefore there exists an absolute constant $\kappa > 0$ such that

$$\frac{1}{\varphi(n)} \leq \frac{\kappa \log \log n}{n}. \quad (20)$$

For instance, we can take $\kappa = e^\omega + 2.5$ for any $n > 15$. Replacing this in (19), we obtain (16). If the weights are chosen so that (7) holds, then obviously the bound in (16) does not depend on d , and this completes the proof. \square

Theorem 2 shows that there exists a generating vector \mathbf{z} whose discrepancy is $O(n^{-\beta}(\log \log n)^\beta)$, where the dimension d appears only in the constant $C(\alpha, \beta, d, \gamma)\kappa^\beta$ and not in the function of n . Similar convergence results for the figure of merit have been obtained previously [3, 4, 14], based on the asymptotically optimal bound of $O(n^{-\alpha}(\log n)^{d\alpha})$ (which is asymptotically slightly stronger because $\beta < \alpha$), by writing $(\log n)^{d\alpha} \leq C_1(\alpha, \delta, d)n^\delta$, where $\delta > 0$ can be taken arbitrarily small. However, there are many situations where our bound is much smaller than the bound provided in those references, for reasonable values of α and n . To see this, let us compare our bound with the bound in [3, Theorem 6], for the case of product weights,

where $\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \gamma_j$ for any $\mathbf{u} \in \mathcal{D}$. That theorem states that there exists a generating vector \mathbf{z} for which

$$D_{n,d,\gamma}(\mathbf{z}, \alpha) \leq \frac{1}{\varphi(n)^\alpha} \prod_{j=1}^d \left(1 + 2\gamma_j^{1/\alpha} (1 - \log 2 + \zeta(\alpha)^{1/\alpha} + \log n)\right)^\alpha. \quad (21)$$

For the same product weights, the constant $C(\alpha, \beta, d, \gamma)$ in (6) can be written as

$$C(\alpha, \beta, d, \gamma) = \left(\prod_{j=1}^d \left(1 + 2\gamma_j^{1/\beta} \zeta(\alpha/\beta)\right) - 1 \right)^\beta.$$

If we take for instance $\alpha = 2$, $d = 10$, and weights $\gamma_j = 1/j^2$, then for $n = 16384 = 2^{14}$, the bound given by (21) is approximately 9.5046×10^7 , whereas the bound in (16) reaches a minimal value (as a function of β) of 0.0022 when $\beta = 1$. Thus, our upper bound is smaller by a factor of about 4×10^{10} . As another example, for $\alpha = 2$, $d = 5$, weights $\gamma_j = 1/j^2$, and $n = 1048576 = 2^{20}$, the bound in (21) is 0.5027, while that in (16) reaches a minimum value of 2.5873×10^{-5} at $\beta = 1$. Also for $\alpha = 2$ and weights $\gamma_j = 1/j^2$, if we now take $n = 2097152 = 2^{21}$ and $d = 10$, the bound in (21) is 2.3005×10^6 , while the bound in (16) reaches its minimum value of 1.0146×10^{-5} when $\beta \approx 1.19$. These numerical examples show that by minimising (16) over β , one can obtain a much tighter bound on the discrepancy than the bound given by (21).

Given that $1 \leq \beta < \alpha$ and because the variance is bounded by this discrepancy, it follows that with a RSLR we can obtain a variance that converges at a faster rate than the usual $O(n^{-1})$ achieved by MC methods.

3. Construction results and algorithms

In this section we present algorithms to construct generating vectors of rank-1 lattice rules and prove that a component-by-component (CBC) construction method (see [15, 16]) returns a generating vector whose weighted discrepancy (9) satisfies the bound given in Theorem 2. We also compare the performance of the CBC algorithm with simpler (and more naive) random search methods. We suppose that d , α and the weights are fixed. We also assume that $n \geq 3$, and that the discrepancy (9) can be computed in constant time for any vector \mathbf{z} . The latter assumption is not always true in

practice, but it holds when α is an even integer and was used in [4, 8, 9, 21] and perhaps not in other cases, as we explain in the final section.

The CBC algorithm constructs the generating vector $\mathbf{z} = (z_1, z_2, \dots, z_d)$ as follows.

CBC construction algorithm:

Let $z_1 := 1$;

For $s = 2, 3, \dots, d$, find $z_s \in \mathcal{Z}_n$ that minimises $D_{n,s,\gamma}((z_1, z_2, \dots, z_s), \alpha)$, defined in (9), while z_1, \dots, z_{s-1} remain unchanged.

We now discuss the total computing time required by this algorithm for the situation where α is an even integer. Then, it is known (see [21]) that the infinite sum

$$C_k(z, \alpha) := \sum'_{h \in \mathbb{Z}} \frac{e^{2\pi i h k z / n}}{|h|^\alpha} \quad (22)$$

can be expressed as the finite sum

$$\sum'_{h \in \mathbb{Z}} \frac{e^{2\pi i h x}}{|h|^\alpha} = \frac{(-1)^{\frac{\alpha}{2}+1} (2\pi)^\alpha}{\alpha!} B_\alpha(x) \quad (23)$$

for all $x \in [0, 1]$, where B_α is the Bernoulli polynomial of degree α . In this case, since there are 2^d weights, the computation of (9) for fixed \mathbf{z} , n , and d requires $O(nd2^d)$ operations, and the full CBC construction requires $O(n^2 d 2^d)$ time, plus additional storage. This is explained in [18] for a different type of discrepancy, but the argument remains the same. Note that if we would recompute the product over j in (8) for each value of s in the CBC algorithm, we would get $O(n^2 d^2 2^d)$ instead of $O(n^2 d 2^d)$ for the total time; we obtain the latter by storing the partial products from previous iterations, for each k . This requires $O(n)$ storage. Of course, having 2^d weights is unpractical, but the cost of the CBC algorithm can be dramatically reduced for particular classes of weights. See [1, 3, 4, 15, 16, 18, 20] for further details.

There is also a fast version of this CBC construction algorithm which returns a generating vector \mathbf{z} in $O(nd \log n)$ time for certain types of weights, such as product weights [15, 16]. The fast CBC construction of [1, 15, 16] replaces the n^2 factor with a much smaller one of $n \log n$ by using a fast Fourier transform, while particular classes of weights such as product weights allow a further reduction on the dimension dependence of the construction algorithm, to yield the mentioned $O(nd \log n)$. Although the $O(nd \log n)$ time

required by the full CBC construction is optimal, we think it is nevertheless interesting to compare this algorithm with simpler (more naive) random search methods, because they are easier to implement and can easily provide good point sets, as we will see later.

The following randomised CBC construction algorithm simplifies the previous one by examining only a small number of integers $z_s \in \mathcal{Z}_n$ (chosen at random) at each step. It has been already used in [19], where the discrepancy measure to minimise was the classical weighted star discrepancy of [18, 20]. A similar algorithm was considered in [24], with the additional feature that for any given s , new integers z_s are examined until $D_{n,s,\gamma}((z_1, z_2, \dots, z_s), \alpha) \leq M_{n,s,\gamma}(\alpha)$.

Randomised CBC construction algorithm (R-CBC):

Let $z_1 := 1$;

For $s = 2, 3, \dots, d$,

choose r integers z_s at random in \mathcal{Z}_n , and select the one that minimises $D_{n,s,\gamma}((z_1, z_2, \dots, z_s), \alpha)$, while z_1, \dots, z_{s-1} remain unchanged.

An even simpler (and more naive) algorithm is a uniform random search in $(\mathcal{Z}_n)^d$, as follows:

Uniform random search algorithm (R-search):

Choose r vectors \mathbf{z} at random in $(\mathcal{Z}_n)^d$, and select the one that minimises $D_{n,s,\gamma}((z_1, z_2, \dots, z_s), \alpha)$.

Both R-CBC and R-search algorithms yield a random output, while the full CBC construction gives a deterministic output. Following the same idea as for the full CBC construction, one can see that for product weights, the R-CBC algorithm takes $O(ndr)$ time, where r is the number of random selections at each step. If $r \approx \log n$ and if we assume that the hidden constants in the two $O(\cdot)$ expressions are the same, this is comparable to the $O(nd \log n)$ time required by the fast implementation of the CBC algorithm. Those hidden constants depend on the specific implementations and comparing them is beyond the scope of this paper. But to give a rough idea of what they can be, with an implementation of fast CBC and R-CBC currently developed by D. Munger and P. L'Ecuyer, for a prime n near 2^{20} , fast CBC needs about twice the CPU time of R-CBC with $r = 14 \approx \log n$.

As pointed out in [15], the hidden constant in the $O(nd \log n)$ expression depends on the number of divisors of n and the number of prime factors of n , since the number of fast Fourier transforms on which the fast CBC construction is based, is dependent on these numbers. The hidden constant in $O(ndr)$ does not depend either on the number of divisors or the number of prime factors of n . In a similar way, the R-search algorithm also takes $O(ndr)$ time, where r is this time the total number of random trials. See also [19] for further discussion on the R-CBC and R-search algorithms. The randomised algorithms can be attractive from a practical viewpoint if we find that they return vectors \mathbf{z} with figures of merit comparable to those returned by the CBC method even for small r . This is basically what we will find for R-CBC in our numerical experiments. Similar results, for a different figure of merit, were reported in [19].

Another special category of weights, used for example in [4, 18, 20], are *order-dependent* weights, where $\gamma_{\mathbf{u}}$ is assumed to depend only on $|\mathbf{u}|$, the cardinality of \mathbf{u} , so we can write $\gamma_{\mathbf{u}} = \Gamma_{\ell}$ when $|\mathbf{u}| = \ell$, where $\Gamma_1, \dots, \Gamma_d$ are non-negative constants. It was shown in [1] that the cost of the fast CBC construction for these weights is also $O(nd \log n)$, but with $O(nd)$ storage. Our R-CBC algorithm for these weights will also require $O(ndr)$ plus $O(nd)$ storage by using the same arguments as in [1].

We now prove that the (full) CBC algorithm constructs a generating vector \mathbf{z} whose corresponding discrepancy satisfies the bound of Theorem 2. The proof is by induction on d and a similar idea was used in [4, 9, 18, 20] under the specific assumptions made in these papers. It basically shows that we can construct good lattice rules that are extensible in d , if we assume that we can compute the discrepancy for any given \mathbf{z} .

Theorem 3. *For any integer $n \geq 3$ and any dimension $d \geq 1$, there exists a $\mathbf{z} \in \mathcal{Z}_n^d$ such that*

$$D_{n,d,\gamma}(\mathbf{z}, \alpha) \leq \frac{1}{(\varphi(n))^\beta} \left(\sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|u|} \right)^\beta. \quad (24)$$

and this vector can be obtained by using the CBC technique, that is, we can set $z_1 = 1$ and then, every component z_d , with $d \geq 1$ can be obtained by minimising $D_{n,d,\gamma}(\mathbf{z}, \alpha)$ with respect to $z_d \in \mathcal{Z}_n$ without altering the previous $d - 1$ components.

Proof. For $d = 1$, the result follows easily from (15) together with $\beta < \alpha$, $\zeta(\alpha) \leq \zeta(\alpha/\beta)$ and $\varphi(n) < n$, which show that

$$D_{n,1,\gamma}(z, \alpha) \leq \frac{2\gamma_{\{1\}}\zeta(\alpha/\beta)}{\varphi(n)^\beta}.$$

For $d \geq 1$, let us assume now that (24) holds. We want to prove that there exists an integer $z_{d+1} \in \mathcal{Z}_n$ such that

$$D_{n,d+1,\gamma}((z, z_{d+1}), \alpha) \leq \frac{1}{(\varphi(n))^\beta} \left(\sum_{u \subseteq \mathcal{D}_1} \gamma_u^{1/\beta} (2\zeta(\alpha/\beta))^{|u|} \right)^\beta, \quad (25)$$

where $\mathcal{D}_1 := \mathcal{D} \cup \{d+1\}$. For any $d \geq 1$, consider

$$D_{n,d+1,\gamma}((z_1, z_2, \dots, z_d, z_{d+1}), \alpha) = \frac{1}{n} \sum_{u \subseteq \mathcal{D}_1} \gamma_u \sum_{k=0}^{n-1} \prod_{j \in u} C_k(z_j, \alpha)$$

where the $C_k(z_j, \alpha)$ are as given by (22). Then we separate out the discrepancy in dimension d from this $(d+1)$ -dimensional discrepancy to obtain

$$D_{n,d+1,\gamma}((z, z_{d+1}), \alpha) = D_{n,d,\gamma}(z, \alpha) + L_{n,d+1,\gamma}((z, z_{d+1}), \alpha), \quad (26)$$

where

$$L_{n,d+1,\gamma}((z, z_{d+1}), \alpha) = \frac{1}{n} \sum_{\substack{u \subseteq \mathcal{D}_1 \\ d+1 \in u}} \gamma_u \sum_{k=0}^{n-1} C_k(z_{d+1}, \alpha) \prod_{j \in u \setminus \{d+1\}} C_k(z_j, \alpha).$$

Following a similar idea as in [4, 18, 20], we then average over all possible integers $z_{d+1} \in \mathcal{Z}_n$ and focus on the last term in the above, because it is the only one depending on z_{d+1} . Using (13) and (14), we see that we have

$$\frac{1}{\varphi(n)} \sum_{z_{d+1} \in \mathcal{Z}_n} C_k(z_{d+1}, \alpha) = T_{n,\alpha}(k),$$

and

$$\sum_{k=0}^{n-1} T_{n,\alpha}(k) \leq \frac{n}{\varphi(n)} 2\zeta(\alpha).$$

We also have $|C_k(z, \alpha)| \leq 2\zeta(\alpha)$. From all these inequalities, it follows that the average on $L_{n,d+1,\gamma}$ over z_{d+1} satisfies

$$\begin{aligned} \text{Avg}(L_{n,d+1,\gamma}((\mathbf{z}, z_{d+1}), \alpha)) &:= \frac{1}{\varphi(n)} \sum_{z_{d+1} \in \mathcal{Z}_n} L_{n,d+1,\gamma}((\mathbf{z}, z_{d+1}), \alpha) \\ &\leq \frac{1}{n} \sum_{\substack{\mathbf{u} \subseteq \mathcal{D}_1 \\ d+1 \in \mathbf{u}}} \gamma_{\mathbf{u}} (2\zeta(\alpha))^{|\mathbf{u}|-1} \frac{n}{\varphi(n)} 2\zeta(\alpha). \end{aligned}$$

Consequently, there exists a $z_{d+1} \in \mathcal{Z}_n$ such that

$$L_{n,d+1,\gamma}((\mathbf{z}, z_{d+1}), \alpha) \leq \frac{1}{\varphi(n)} \sum_{\substack{\mathbf{u} \subseteq \mathcal{D}_1 \\ d+1 \in \mathbf{u}}} \gamma_{\mathbf{u}} (2\zeta(\alpha))^{|\mathbf{u}|}. \quad (27)$$

On the other hand, using (8), we see that we can write

$$\begin{aligned} L_{n,d+1,\gamma}((\mathbf{z}, z_{d+1}), \alpha) &= \frac{1}{n} \sum_{\substack{\mathbf{u} \subseteq \mathcal{D}_1 \\ d+1 \in \mathbf{u}}} \gamma_{\mathbf{u}} \sum_{k=0}^{n-1} \prod_{j \in \mathbf{u}} C_k(z_j, \alpha) \\ &= \sum_{\substack{\mathbf{u} \subseteq \mathcal{D}_1 \\ d+1 \in \mathbf{u}}} \gamma_{\mathbf{u}} \sum_{\{\mathbf{h} \in L^\perp : \mathbf{u}(\mathbf{h}) = \mathbf{u}\}} \prod_{j \in \mathbf{u}} |h_j|^{-\alpha}. \end{aligned}$$

Using then Jensen's inequality (17), it follows that (see also the proof of Theorem 2)

$$\sum_{\substack{\mathbf{u} \subseteq \mathcal{D}_1 \\ d+1 \in \mathbf{u}}} \sum_{\{\mathbf{h} \in L^\perp : \mathbf{u}(\mathbf{h}) = \mathbf{u}\}} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} |h_j|^{-\alpha} \leq \left(\sum_{\substack{\mathbf{u} \subseteq \mathcal{D}_1 \\ d+1 \in \mathbf{u}}} \sum_{\{\mathbf{h} \in L^\perp : \mathbf{u}(\mathbf{h}) = \mathbf{u}\}} \gamma_{\mathbf{u}}^{1/\beta} \prod_{j \in \mathbf{u}} |h_j|^{-\alpha/\beta} \right)^\beta,$$

which shows that

$$L_{n,d+1,\gamma}((\mathbf{z}, z_{d+1}), \alpha) \leq (L_{n,d+1,\gamma^{1/\beta}}((\mathbf{z}, z_{d+1}), \alpha/\beta))^\beta.$$

This inequality together with (27) shows that the chosen $z_{d+1} \in \mathcal{Z}_n$ satisfies

$$L_{n,d+1,\gamma}((\mathbf{z}, z_{d+1}), \alpha) \leq \frac{1}{(\varphi(n))^\beta} \left(\sum_{\substack{\mathbf{u} \subseteq \mathcal{D}_1 \\ d+1 \in \mathbf{u}}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|\mathbf{u}|} \right)^\beta.$$

From this inequality and (26) and by using the induction hypothesis (24), it follows that the chosen z_{d+1} satisfies

$$D_{n,d+1,\gamma}(\mathbf{z}, z_{d+1}, \alpha) \leq \frac{1}{(\varphi(n))^\beta} \left(\sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|\mathbf{u}|} \right)^\beta + \frac{1}{(\varphi(n))^\beta} \left(\sum_{\substack{\mathbf{u} \subseteq \mathcal{D}_1 \\ d+1 \in \mathbf{u}}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|\mathbf{u}|} \right)^\beta.$$

The desired result (25) follows from this inequality by applying again Jensen's inequality (17). \square

Theorem 3 implies that the vector \mathbf{z} constructed by the CBC algorithm satisfies the optimal bound given by Theorem 2. This is stated in the next corollary.

Corollary 4. *For any $\alpha > 1$, any dimension $d \geq 1$, and any $n \geq 3$, the vector \mathbf{z} constructed by the CBC algorithm satisfies (16) for any β satisfying $1 \leq \beta < \alpha$.*

Markov's inequality allows us to state a probabilistic bound on the figure of merit of the vector \mathbf{z} returned by the R-search algorithm. Similar results in the prime case have been proved in [4, Theorem 1 and Theorem 2].

Theorem 5. *If $\check{\mathbf{z}}$ is the vector returned by the R-search algorithm, r is the number of independent trials of the algorithm, and $t > 1$, then*

$$\mathbb{P} \left[D_{n,d,\gamma}(\check{\mathbf{z}}, \alpha) \leq t \left(\frac{\kappa \log \log n}{n} \right)^\beta \left(\sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|\mathbf{u}|} \right)^\beta \right] \geq 1 - t^{-r/\beta}.$$

Proof. Since the average of $D_{n,d,\gamma}(\mathbf{z}, \alpha)$ over all vectors $\mathbf{z} \in \mathcal{Z}_n^d$ is bounded by $\overline{M}_{n,d,\gamma}(\alpha)$ (see Theorem 1), we have from Markov's inequality that for all $a > 1$ and for \mathbf{z} drawn at random uniformly from \mathcal{Z}_n^d ,

$$\mathbb{P} \left[D_{n,d,\gamma^{1/\beta}}(\mathbf{z}, \alpha/\beta) \geq a \overline{M}_{n,d,\gamma^{1/\beta}}(\alpha/\beta) \right] \leq 1/a,$$

where $\overline{M}_{n,d,\gamma^{1/\beta}}(\alpha/\beta)$ is defined in (11). We also have from (18) that

$$D_{n,d,\gamma}(\mathbf{z}, \alpha) \leq (D_{n,d,\gamma^{1/\beta}}(\mathbf{z}, \alpha/\beta))^\beta$$

for all \mathbf{z} . Combining these two inequalities with (20) gives that

$$\mathbb{P} \left[D_{n,d,\gamma}(\mathbf{z}, \alpha) \geq t \left(\frac{\kappa \log \log n}{n} \right)^\beta \left(\sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}}^{1/\beta} (2\zeta(\alpha/\beta))^{|\mathbf{u}|} \right)^\beta \right] \leq t^{-1/\beta}.$$

Since r is the number of independent trials, the result follows by a geometric argument. \square

This last theorem shows that if we choose $t > 1$ and r large enough so that $t^{-r/\beta}$ is very small, we have a very high probability that a vector \mathbf{z} leads to a discrepancy satisfying the desired error bound.

4. Numerical experiments with the construction algorithms

We will now compare empirically the performance of the three construction algorithms given in Section 3. We made numerical investigations with various values of d , n , even integer values of α and different choices of weights. We report the results for a small representative subset of those experiments. Similar experiments were reported in [19] for a different type of figure of merit, namely the weighted star discrepancy, which provides worst-case deterministic error bounds for certain classes of functions. Other experiments with the CBC construction method for the same measure of discrepancy as considered here, with $\alpha = 2$ and specific types of weights, are also reported in [9] (for product weights) and [4] (for order-dependent weights, defined below).

When α is not an even integer, there is no closed-form formula for $C_k(z, \alpha)$. Then, in the construction algorithms presented in the previous section, one could truncate the infinite sum that defines $C_k(z, \alpha)$. However it is unclear what would be the impact of such a truncation over the figure of merit.

For the reported experiments, we take $\alpha = 2$, which means that we consider functions f whose Fourier coefficients satisfy

$$|\hat{f}(\mathbf{h})| \leq \gamma_{\mathbf{u}(\mathbf{h})} \prod_{j \in \mathbf{u}} |h_j|^{-1}$$

for all \mathbf{h} . These functions do not have absolutely convergent Fourier series, as typically assumed in other papers [3, 4, 8, 9]. In that particular case, the discrepancy (9) which bounds the variance is equivalent to the figure of merit used in [4], that is

$$\begin{aligned} D_{n,d,\gamma}(\mathbf{z}, 2) &= \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} \left(\sum'_{h \in \mathbb{Z}} \frac{e^{2\pi i h k z_j / n}}{|h|^2} \right) \\ &= \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} \left(2\pi^2 B_2 \left(\left\{ \frac{k z_j}{n} \right\} \right) \right), \end{aligned} \quad (28)$$

where B_2 is the Bernoulli polynomial of degree 2, given by $B_2(x) = x^2 - x + 1/6$. This expression is easy to compute.

If the weights have the product form $\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \gamma_j$, the expression of $D_{n,d,\gamma}(\mathbf{z}, 2)$ given by (28) can be rewritten as

$$D_{n,d,\gamma}(\mathbf{z}, 2) = \frac{1}{n} \sum_{k=0}^{n-1} \prod_{j=1}^d \left(1 + 2\pi^2 \gamma_j B_2 \left(\left\{ \frac{k z_j}{n} \right\} \right) \right) - 1.$$

For order-dependent weights, that is $\gamma_{\mathbf{u}} = \Gamma_{\ell}$ when $|\mathbf{u}| = \ell$, we can write

$$D_{n,d,\gamma}(\mathbf{z}, 2) = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\ell=1}^d \Gamma_{\ell} \sum_{\substack{\mathbf{u} \subseteq \mathcal{D} \\ |\mathbf{u}|=\ell}} \prod_{j \in \mathbf{u}} \left(2\pi^2 B_2 \left(\left\{ \frac{k z_j}{n} \right\} \right) \right). \quad (29)$$

For $\alpha = 2$, the average is bounded by (see also (11)):

$$M_{n,d,\gamma}(2) \leq \frac{1}{\varphi(n)} \sum_{\mathbf{u} \subseteq \mathcal{D}} \gamma_{\mathbf{u}} \left(\frac{\pi^2}{3} \right)^{|\mathbf{u}|}.$$

For the experiments with $D_{n,d,\gamma}(\mathbf{z}, 2)$ reported here, we selected dimensions ranging from 5 to 80, and values of n ranging from 2^{10} to 2^{20} . We consider product weights with $\gamma_j = 1/j^2$ for all j . For those weights, we have $\sum_{j=1}^{\infty} \gamma_j < \infty$ and this is sufficient to ensure that condition (7) holds, which implies that that variance is bounded independently of the dimension d . We also consider order-dependent weights given by $\Gamma_{\ell} = (d(d-1) \cdots (d-\ell+1))^{-1}$ for all $\mathbf{u} \subseteq \mathcal{D}$ with $|\mathbf{u}| = \ell$, where $\ell = 1, \dots, d$.

Replacing these weights in (29) and by noting that $|B_2(x)| \leq 1/6$ for any $x \in [0, 1]$, we obtain

$$D_{n,d,\gamma}(\mathbf{z}, 2) \leq \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\ell=1}^d \frac{1}{\ell!} \left(\frac{\pi^2}{3}\right)^\ell \leq \exp\left(\frac{\pi^2}{3}\right),$$

which clearly shows that this discrepancy is bounded independently of d .

Table 1 summarises the values of the discrepancy (28) obtained for different values of n (powers of 2), product weights in $d = 20$ dimensions and order-dependent weights in $d = 10$ dimensions, with the weights defined as indicated above, with the (full) CBC method, randomised CBC with $r = 10$ (this is smaller than $\log n$), and uniform random search. For the latter method, we used a sample size of $r = 10^5$ for $n = 2^{14}, 2^{15}, 2^{16}$, and $r = 10^4$ for $n = 2^{17}, 2^{18}$. These large sample sizes were used for the purpose of estimating the probability distribution of the figure of merit, as illustrated in Figure 1. When applying the R-search algorithm to quickly find a good lattice, one would normally use a smaller r than this, and the best returned figure of merit will typically be a bit larger.

Table 1: Values of the figure of merit obtained by the CBC algorithm (CBC), the randomised CBC algorithm (R-CBC) with $r = 10$, the uniform random search algorithm with a large r (R-search), and the value of the bound for the mean $\overline{M}_{n,d,\gamma}(2)$

Weights	d	n	CBC	R-CBC	R-search	$\overline{M}_{n,d,\gamma}(2)$
product	20	2^{14}	4.65e-5	7.94e-5	7.51e-5	2.60e-3
		2^{15}	1.81e-5	2.19e-5	3.03e-5	1.30e-3
		2^{16}	6.76e-6	8.69e-6	1.23e-5	6.50e-4
		2^{17}	2.56e-6	3.48e-6	5.14e-6	3.24e-4
		2^{18}	9.73e-7	1.39e-6	2.09e-6	1.62e-4
order-dep.	10	2^{14}	5.20e-4	5.63e-4	5.71e-4	3.16e-3
		2^{15}	2.25e-4	2.37e-4	2.64e-4	1.58e-4
		2^{16}	9.80e-5	1.07e-4	1.13e-4	7.88e-4
		2^{17}	4.26e-5	4.71e-5	5.19e-5	3.94e-4
		2^{18}	1.86e-5	2.02e-5	2.24e-5	1.97e-4

We observed that the R-CBC algorithm always returned vectors \mathbf{z} with figure of merit smaller than the bound $\overline{M}_{n,d,\gamma}(2)$, even for small values of r

(such as $r = 10$, as reported in the table). This means that finding a good generating vector (in the sense of doing better than the bound) is easy and does not require the full CBC construction. As expected, the figures of merit for R-CBC are not as good as for full CBC. In other experiments, it was observed that when the weights are fixed, the gap between CBC and R-CBC generally decreases with the dimension. Note that despite the large value of r used in the R-search algorithm here, the R-search does not perform as well as the R-CBC. Nevertheless, except in one case, it still returns a value much smaller than the mean.

In our exploration of the uniform random search algorithm method, we computed an empirical version \hat{F} of the cumulative distribution function F of $D_{n,d,\gamma}(\mathbf{z}, 2)$, defined by $F(x) = \mathbb{P}[D_{n,d,\gamma}(\mathbf{z}, 2) \leq x]$, for a purely random \mathbf{z} drawn uniformly from \mathcal{Z}_n^d . The empirical distribution was computed with $r = 10^5$ for $n \leq 2^{16}$, and $r = 10^4$ otherwise. We observed that the shapes of the corresponding distributions were quite similar for different values of d , n , and weights (with proper scaling). We found that typically, this distribution is positively skewed, and the median is smaller than the average $M_{n,d,\gamma}(2)$, or the bound on this average, so the probability $q_{n,d,\gamma}$ of getting a value smaller than the average with a random \mathbf{z} is more than $1/2$ (often around 0.9). Note that these empirical observations may change with the figure of merit. For instance in the case of the weighted star discrepancy used in [19], we observed that the probability of finding a vector better than the mean is often around 0.75. This implies that a vector \mathbf{z} whose corresponding discrepancy is smaller than the average (and thus satisfies the bound in Theorem 2) is easy to find even by simple uniform random search. By applying this algorithm with r trials, the probability of finding such a vector is $1 - (1 - q_{n,d,\gamma})^r$. With $q_{n,d,\gamma} = 0.9$, this probability is $1 - 10^{-r}$, which is very close to 1 even for moderate values of r . Note that the bound on the probability given (and proved) in Theorem 5 is only valid for a discrepancy value larger than the mean and is usually very conservative. This easyness of finding a vector better than the mean just by random search may suggest that we should be more ambitious than only beating the average if the goal is to really find one of the best rules. On the other hand, the discrepancy bounds and convergence rate results are only in terms of the average, and there is no proof that one can always do significantly better.

An example of an empirical distribution is given in Figure 1. The minimum value returned by the figure of merit out of 10000 random tries was 2.24×10^{-5} , while the maximum was 2.41×10^{-2} . In the figure, the green (left-

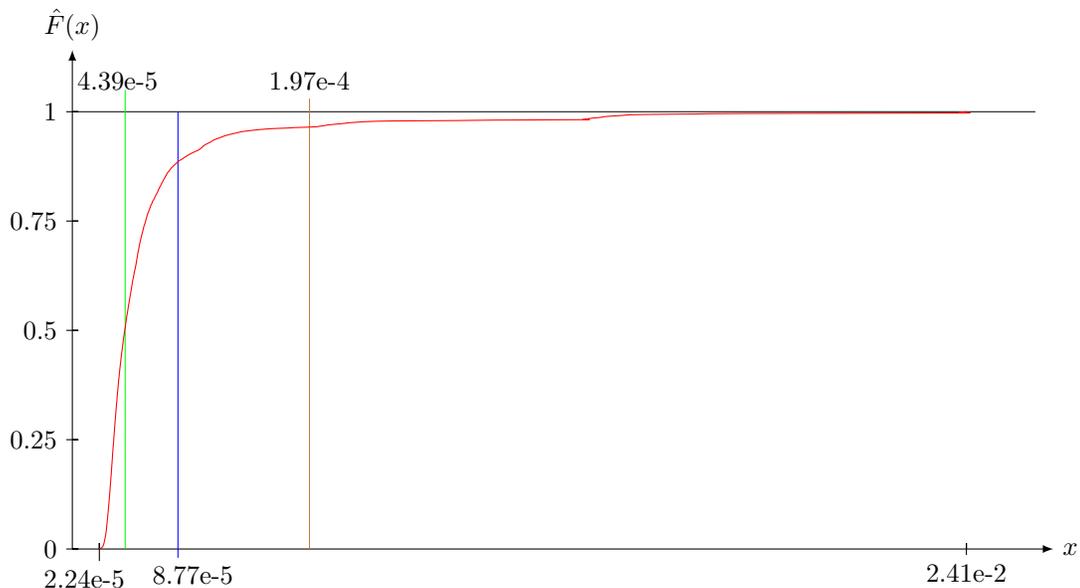


Figure 1: Empirical distribution of $D_{n,d,\gamma}(\mathbf{z}, 2)$ for 10^4 random vectors when $n = 2^{18} = 262144$, $d = 10$ and order-dependent weights $\Gamma_\ell = (d(d-1)\cdots(d-\ell+1))^{-1}$, for all $\ell = 1, \dots, d$.

most) vertical line indicates the median (4.39×10^{-5}), the blue line (middle one) indicates the empirical average (8.77×10^{-5}), while the brown (right-most) line indicates the value of the bound (11) on the mean (1.97×10^{-4}).

Acknowledgements

This work has been supported by an NSERC-Canada Discovery Grant and a Canada Research Chair to the second author. The first draft of the paper was written when the first author was a postdoctoral fellow at the University of Montreal. The first author acknowledges support from Professor Ronald Cools, via the Research Programme of the Research Foundation-Flanders (FWO-Vlaanderen) G.0597.06. David Munger helped us with some comments and experiments with the CBC method.

References

- [1] R. Cools, F.Y. Kuo and D. Nuyens. (2006). *Constructing embedded lattice rules for multivariate integration*, SIAM J. Sci. Comput. **28**, pp. 2162–2168.
- [2] M.T. Çezik and P. L’Ecuyer. (2008). *Staffing multiskill call centers via linear programming and simulation*, Management Science **54**, pp. 310–323.
- [3] J. Dick. (2004). *On the convergence rate of the component-by-component construction of good lattice rules*, J. Complexity **20**, pp. 493–522.
- [4] J. Dick, I.H. Sloan, X. Wang and H. Woźniakowski. (2006). *Good lattice rules in weighted Korobov spaces with general weights*, Numer. Math. **103**, pp. 63–97.
- [5] S. Disney. (1990). *Error bounds for rank 1 lattice quadrature rules modulo composites*, Monatsh. Math. **110**, pp. 89–100.
- [6] G.H. Hardy, J.E. Littlewood and G. Pólya. (1964). *Inequalities*, 2nd ed., Cambridge University Press.
- [7] F.J. Hickernell. (1998). *Lattice rules: how well do they measure up?*, Random and quasi-random point sets, Lecture Notes in Statist. **138**, Springer, New York, pp. 109–166.
- [8] F.Y. Kuo. (2003). *Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces*, J. Complexity **19**, pp. 301–320.
- [9] F.Y. Kuo and S. Joe. (2002). *Component-by-component construction of good lattice rules with a composite number of points*, J. Complexity **18**, pp. 943–976.
- [10] P. L’Ecuyer. (2009). *Quasi-Monte Carlo methods with applications in finance*, Finance and Stochastics **13**, pp. 307–349.
- [11] P. L’Ecuyer and C. Lemieux. (2000). *Variance reduction via lattice rules*, Management Science **46**, pp. 1214–1235.

- [12] P. L'Ecuyer and D. Munger. (2011). *On Figures of Merit for Randomly-Shifted Lattice Rules*, in Monte Carlo and Quasi-Monte Carlo Methods 2010, H. Wozniakowski and L. Plaskota (Eds.), Springer, Berlin, 2011.
- [13] R. Liu and A.B. Owen. (2006). *Estimating mean dimensionality of analysis of variance decompositions*, J. Amer. Statist. Assoc. **101**, pp. 712–721.
- [14] H. Niederreiter. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia.
- [15] D. Nuyens and R. Cools. (2006). *Fast component-by-component construction of rank-1 lattice rules with a non-prime number of points*, J. Complexity **22**, pp. 4–28.
- [16] D. Nuyens and R. Cools. (2006). *Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces*, Math. Comp. **75**, pp. 903–920.
- [17] J.B. Rosser and L. Schoenfeld. (1962). *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6**, pp. 64–94.
- [18] V. Sinescu and S. Joe. (2007). *Good lattice rules based on the general weighted star discrepancy*, Math. Comp. **76**, pp. 989–1004.
- [19] V. Sinescu and P. L'Ecuyer. (2009). *On the behavior of the weighted star discrepancy bounds for shifted lattice rules*, in Monte Carlo and Quasi-Monte Carlo Methods 2008, P. L'Ecuyer and A.B. Owen (Eds.), Springer, Berlin, pp. 603–616.
- [20] V. Sinescu and P. L'Ecuyer. (2011). *Existence and construction of shifted lattice rules with an arbitrary number of points and bounded weighted star discrepancy for general decreasing weights*, J. Complexity **27**, pp. 449–465.
- [21] I.H. Sloan and S. Joe. (1994). *Lattice Methods for Multiple Integration*, Clarendon Press, Oxford.
- [22] I.H. Sloan and H. Woźniakowski. (1998). *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?*, J. Complexity **14**, pp. 1–33.

- [23] I.M. Sobol' and E.E. Myshetskaya. (2007). *Monte Carlo estimators for small sensitivity indices*, Monte Carlo Methods and Applications **13**, pp. 455–465.
- [24] X. Wang and I.H. Sloan. (2006). *Efficient weighted lattice rules with applications to finance*, SIAM J. Sci. Comput. **28**, pp. 728–750.