A strong law of large numbers for scrambled net integration

Art B. Owen
Stanford University

Daniel Rudolf
University of Goettingen

May 2020

Abstract

This article provides a strong law of large numbers for integration on digital nets randomized by a nested uniform scramble. The motivating problem is optimization over some variables of an integral over others, arising in Bayesian optimization. This strong law requires that the integrand have a finite moment of order p for some $p > 1$. Previously known results implied a strong law only for Riemann integrable functions. Previous general weak laws of large numbers for scrambled nets require a square integrable integrand. We generalize from L^2 to L^p for $p > 1$ via the Riesz-Thorin interpolation theorem.

1 Introduction

Numerical integration is a fundamental building block in many applied mathematics problems. When the integrand is a smooth function of a low dimensional input, then classical methods such as tensor products of Simpson's rule are very effective [8]. For non-smooth integrands or higher dimensional domains, these method may perform poorly. One then turns to Monte Carlo methods, where the integrand is expressed as the expected value of a random variable which is then sampled in a simulation and averaged. Sample averages converge to population averages by the law of large numbers (LLN), providing a justification for the Monte Carlo method.

The Monte Carlo method converges very slowly to the true answer as the number n of sampled values increases. The root mean squared error is $O(n^{-1/2})$. Quasi-Monte Carlo (QMC) methods [10, 11, 39] replace random sampling by deterministic sampling methods. These may be heuristically described as space filling samplers using n points constructed to reduce the unwanted gaps and clusters that would arise among randomly chosen inputs. Because the inputs are not random, we cannot use the law of large numbers to ensure that the estimate converges to the integral as $n \to \infty$. Such consistency is a minimal requirement of an integration method. For QMC, consistency requires additional assumptions of Riemann integrability or bounded variation, whose description we defer. Under the latter condition, the integration error is $O(n^{-1+\epsilon})$ for any $\epsilon > 0$. QMC has proved valuable in financial valuation [16], graphical rendering [28] and solving PDEs in random environments [32].

In addition to knowing that a method would work as $n \to \infty$, users also need to have some estimate of how well it has worked for a given sample size n. Monte Carlo methods make it easy to quantify uncertainty by using the central limit theorem in conjunction with a sample variance estimate. Plain QMC lacks such a convenient error estimate. Randomized QMC (RQMC) methods, surveyed in [34], produce random points with QMC properties. Then a few statistically independent repeats of the whole RQMC process support uncertainty quantification. One of these methods, scrambled nets [42, 43], provides estimated integrals that are consistent as $n \to \infty$ under weaker conditions than plain QMC requires. It can also reduce the root mean squared error to $O(n^{-3/2+\epsilon})$ [44, 50] under further conditions on the integrand.

Up to this point, we have considered the LLN as just one result. There are in fact strong and weak forms of the LLN that we discuss below. The distinction does not come up for plain Monte Carlo sampling because both laws hold at once. For RQMC, mostly weak laws of large numbers have been proved. Our contribution here is to establish strong laws. The motivation to do this comes from a personal communication by Max Balandat. He and co-authors at Facebook Research are developing a Bayesian optimization tool. In those problems one must optimize over variables θ the integral over variables $\boldsymbol{x}$ of some function $g(\theta, \boldsymbol{x})$. Integration is thus a building block in a larger problem. Consistent estimation of the optimal θ could be proved assuming a strong LLN for samples $\boldsymbol{x}$. Such a strong law was available for plain Monte Carlo but not for RQMC, yet RQMC has much better empirical results in their work in progress.

An outline of this paper is as follows. Section 2 presents the strong and weak laws of large numbers referred to above as well as MC and QMC and RQMC sampling, making more precise some of the conditions stated in this introduction. It includes a lemma to show that functions of bounded variation in the sense of Hardy and Krause (the usual regularity assumption in QMC) must also be Riemann integrable. That is either a new result or one hard to find in the literature. Section 3 defines the QMC method known as digital nets whose RQMC counterparts are called scrambled nets. Section 4 has the main result. It is a strong law of large numbers for scrambled net sampling. The integrand is assumed to be square integrable. The first new strong law is a form of consistency for scrambled net integration as $n \to \infty$ through the set of values that can be written rb^m for $r = 1, \ldots, R$ and $m \geqslant 0$ and an integer $b \geqslant 2$ instead of through all $n \geqslant 1$ as in plain MC. While those are the best sample sizes to use for reasons given in that section, we next extend the result to the plain limit as $n \to \infty$. Section 5 replaces the assumption that f^2 be integrable by one that $|f|^p$ have a finite integral for some $p > 1$. This result uses the Riesz-Thorin interpolation theorem [4]. Section 6 provides some additional context and discussion, including randomly shifted lattice versions of RQMC.

2 Background on LLNs, QMC and RQMC

We begin with the unit cube $[0,1]^d$ in dimension $d \geqslant 1$. For $p \geqslant 1$, the space $L^p[0,1]^d$ consists of all measurable functions f on $[0,1]^d$ for which $\|f\|_p = \bigl(\int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\bigr)^{1/p} < \infty$. We consider the problem of computing an estimate $\hat\mu$ of the integral $\mu = \int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. Here μ is the expected value of $f(\boldsymbol{x})$ for $\boldsymbol{x} \sim \mathbb{U}([0,1]^d)$. Many problems that do not originate as integrals over $[0,1]^d$ have such a representation using transformations to generate non-uniformly distributed random variables over the cube and other spaces [9]. We suppose that those transformations are subsumed into f. Also, while our theory works for genuinely random numbers, in practice one ordinarily uses deterministic output of a random number generator that simulates randomness.

The plain Monte Carlo (MC) method takes independent $\boldsymbol{x}_i \sim \mathbb{U}[0,1]^d$ and estimates μ by $\hat\mu_n = \hat\mu_n^{\mathrm{MC}} = (1/n)\sum_{i=1}^n f(\boldsymbol{x}_i)$. There are many more sophisticated Monte Carlo methods but when we refer to Monte Carlo below we mean this simple one.

The weak law of large numbers (WLLN) implies that for any $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr\bigl(|\hat\mu_n^{\mathrm{MC}} - \mu| > \epsilon\bigr) = 0. \tag{1}$$

The strong law of large numbers (SLLN) implies that

$$\Pr\Bigl(\lim_{n\to\infty} \hat\mu_n^{\mathrm{MC}} = \mu\Bigr) = 1 \tag{2}$$

which we may write as $\Pr(\limsup_{n\to\infty} |\hat\mu_n^{\mathrm{MC}} - \mu| > \epsilon) = 0$ to parallel the WLLN. Both the WLLN and SLLN hold for independent and identically distributed (IID) random variables $f(\boldsymbol{x}_i)$ when $f \in L^1[0,1]^d$. For proofs of these laws, see [13, Chapter 2]. For an example of a sequence of independent random variables that satisfies the WLLN but not the SLLN, let $\hat\mu_n = \mu$ with probability $1 - 1/n$ and $\hat\mu_n = \mu + 1$ otherwise.

In QMC sampling, the $\boldsymbol{x}_i$ are constructed so that the discrete distribution placing probability $1/n$ on each of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ (with repeated points counted multiple times) is close to the continuous uniform distribution on $[0,1]^d$. There are various ways, called discrepancies [5], to quantify the distance between these discrete and continuous measures. For a set $S \subset [0,1]^d$ define $\mathbf{1}\{\boldsymbol{x} \in S\}$, to be 1 if $\boldsymbol{x} \in S$ and 0 otherwise. The most widely used discrepancy is the star discrepancy

$$D_n^* = D_n^*(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \sup_{\boldsymbol{a}\in[0,1]^d} \biggl| \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{\boldsymbol{x}_i \in [\mathbf{0}, \boldsymbol{a})\} - \prod_{j=1}^d a_j \biggr|$$

where $[\mathbf{0}, \boldsymbol{a}) = \{\boldsymbol{x} \in [0,1]^d \mid 0 \leqslant x_j < a_j,\ j = 1, \dots, d\}$.

To keep this paper at a manageable length, the relevant properties of QMC and RQMC methods are presented but the details of their constructions are omitted. For the latter, see [10, 11, 39, 34] among others.

Because QMC is deterministic, it has no analogue of the WLLN (1). There is an analogue of the SLLN (1), as follows. Let $\hat{\mu}_n^{\mathrm{QMC}} = (1/n)\sum_{i=1}^n f(\boldsymbol{x}_i)$ where now the points $\boldsymbol{x}_i$ have been chosen to have small discrepancy. If f is Riemann integrable and $D_n^* \to 0$ then [30, p. 3]

$$\lim_{n\to\infty} \hat{\mu}_n^{\mathrm{QMC}} = \mu \tag{3}$$

providing the QMC version of the SLLN (2). There is a converse, where if $|\hat{\mu}_n - \mu| \to 0$ whenever $D_n^* \to 0$, then f must be Riemann integrable. See the references and discussion in [37]. That is, QMC could fail to be consistent when f is not Riemann integrable. Riemann integrable f must also be bounded.

A better known result about QMC is the Koksma-Hlawka inequality below which uses the notion of bounded variation. Recall that a differentiable function f on $[0,1]$ has total variation $V(f) = \int_0^1 |f'(x)|\,\mathrm{d}x$ and it is of bounded variation for $V(f) < \infty$. There are numerous generalizations of the total variation for functions on the unit cube $[0,1]^d$ when $d > 1$ (see [6]). Of those, the total variation in the sense of Hardy and Krause [20, 29], denoted by $V_{\mathrm{HK}}(f)$, is the most useful one for QMC. If $V_{\mathrm{HK}}(f) < \infty$, then we write $f \in \mathrm{BVHK}[0,1]^d$. Although we don't need f to have bounded variation to get the SLLN (3) for QMC, bounded variation gives us some information on the rate of convergence, via the Koksma-Hlawka inequality

$$|\hat{\mu}_n^{\mathrm{QMC}} - \mu| \leqslant D_n^* \times V_{\mathrm{HK}}(f) \tag{4}$$

(see [24]). Typical QMC constructions provide infinite sequences $\boldsymbol{x}_i$ whose initial subsequences satisfy

$$D_n^*(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = O\Bigl(\frac{\log(n)^d}{n}\Bigr).$$

Then $|\hat{\mu}_n^{\mathrm{QMC}} - \mu| = O(n^{-1+\epsilon})$ by (4) for any $\epsilon > 0$.

The counterpart in MC to the Koksma-Hlawka inequality is that

$$\mathbb{E}((\hat{\mu}_n^{\mathrm{MC}} - \mu)^2)^{1/2} = n^{-1/2}\sigma(f) \tag{5}$$

when, for $\boldsymbol{x} \sim \mathbb{U}[0,1]^d$ we have $\sigma^2 = \sigma^2(f) = \mathbb{E}((f(\boldsymbol{x}) - \mu)^2) < \infty$. Where the rate for QMC comes after strengthening the regularity requirement on f from Riemann integrability to bounded variation, the rate for MC comes about after strengthening the requirement from $f \in L^1[0,1]^d$ to $f \in L^2[0,1]^d$. The MC counterpart (5) is exact while the QMC version (4) is an extremely conservative upper bound, in that it covers even the worst $f \in \mathrm{BVHK}[0,1]^d$ for any given $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$.

A Riemann integrable function is not necessarily in BVHK. For instance $f(\boldsymbol{x}) = \mathbf{1}\{\sum_{j=1}^d x_j \leqslant 1\}$ is Riemann integrable but, for $d \geqslant 2$, it is not in BVHK [47]. A function in BVHK is necessarily Riemann integrable. This result is hard to find in the literature. It must almost certainly have been known to Hardy, Krause, Hobson and others over a century ago, at least for $d = 2$, which earlier work emphasized. Here is a short proof based on some recent results.

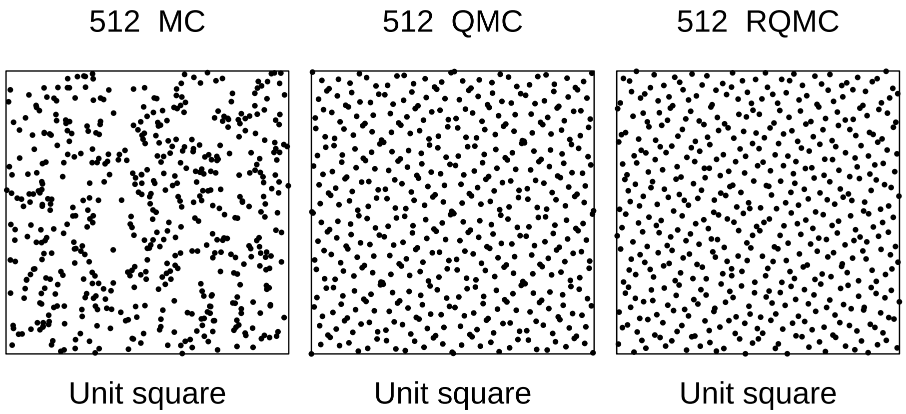

Figure 1: Each panel depicts 512 points in the unit square $[0,1]^2$. From left to right: plain Monte Carlo points, Sobol' points, scrambled Sobol' points.

Lemma 1. *If $f \in \mathrm{BVHK}[0,1]^d$, then f is also Riemann integrable.*

Proof. If $f \in \mathrm{BVHK}[0,1]^d$ then $f(\boldsymbol{x}) = f(\mathbf{0}) + f_+(\boldsymbol{x}) - f_-(\boldsymbol{x})$ where $f_\pm$ are uniquely determined completely monotone functions on $[0,1]^d$ with $f_\pm(\mathbf{0}) = 0$ [1, Theorem 2]. Completely monotone functions are, *a fortiori*, monotone. Now both $f_\pm$ are bounded monotone functions on $[0,1]^d$. They are then Riemann integrable by the corollary in [33]. $\square$

While QMC has a superior convergence rate to MC for $f \in \mathrm{BVHK}$, MC has an advantage over QMC in that $\mathbb{E}((\hat\mu^{\mathrm{MC}} - \mu)^2) = \sigma^2/n$ is simple to estimate from independent replicates, while D_n^* is very expensive to compute [12] and $V_{\mathrm{HK}}(f)$ is much harder to estimate than μ. In a setting where attaining accuracy is important, it will also be important to estimate the attained accuracy. RQMC methods, described next, are hybrids of MC and QMC that support error estimation.

In RQMC [34, 42] one starts with points $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n \in [0,1]^d$ having a small star discrepancy and randomizes them to produce points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$. These points satisfy the following two conditions: individually $\boldsymbol{x}_i \sim \mathbb{U}[0,1]^d$, and collectively, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ have small star discrepancy. The RQMC estimate of μ is $\hat\mu_n^{\mathrm{RQMC}} = (1/n)\sum_{i=1}^n f(\boldsymbol{x}_i)$. From the uniformity of the points $\boldsymbol{x}_i$ we find that $\mathbb{E}(\hat\mu_m^{\mathrm{RQMC}}) = \mu$. Their small star discrepancy means that they are also QMC points and so they inherit the accuracy properties of QMC. To estimate error, one takes several independent randomizations of $\boldsymbol{a}_i$ producing independent replicates of $\hat\mu_n^{\mathrm{RQMC}}$ whose sample variance can be computed.

The first panel in Figure 1 shows 512 MC points in the unit square $[0,1]^2$. We see clear gaps and clumps among those points. The second panel shows 512 QMC points from a Sobol' sequence described in Section 3. The points are very structured and fill the space quite evenly. The third panel shows a scrambled version of those 512 points also described in Section 3.

3 Scrambled nets and sequences

In this section, we describe digital nets and sequences and scrambled versions of them. Let $b \geqslant 2$ be an integer base. Let $\boldsymbol{k} = (k_1, \dots, k_d) \in \mathbb{N}^d$ and $\boldsymbol{c} = (c_1, \dots, c_d)$ where $c_j \in \{0, 1, \dots, b^{k_j} - 1\}$. Then the set

$$E(\boldsymbol{k}, \boldsymbol{c}) = \prod_{j=1}^{d} \Bigl[\frac{c_j}{b_j^{k_j}}, \frac{c_j + 1}{b_j^{k_j}}\Bigr) \tag{6}$$

is called an elementary interval in base b. It has volume $b^{-|\boldsymbol{k}|}$ where $|\boldsymbol{k}| = \sum_{j=1}^{d} k_j$.

Definition 1. For integers $m \geqslant t \geqslant 0$, $b \geqslant 2$ and $d \geqslant 1$, the points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in [0, 1)^d$ for $n = b^m$ are a (t, m, d)-net in base b if

$$\sum_{i=1}^{n} \mathbf{1}\{\boldsymbol{x}_i \in E(\boldsymbol{k}, \boldsymbol{c})\} = b^{m-|\boldsymbol{k}|}$$

holds for every elementary interval $E(\boldsymbol{k}, \boldsymbol{c})$ from (6) with $|\boldsymbol{k}| \leqslant m - t$.

An elementary interval of volume $b^{-|\boldsymbol{k}|}$ should ideally contain $nb^{-|\boldsymbol{k}|} = b^{m-|\boldsymbol{k}|}$ points from $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$. In a digital net, every elementary interval that should ideally contain b^t of the points does so. For any given b, m and d, smaller t imply finer equidistribution. It is not always possible to attain $t = 0$.

Definition 2. For integers $t \geqslant 0$, $b \geqslant 2$ and $d \geqslant 1$, the points $\boldsymbol{x}_i \in [0, 1)^d$ for $i \geqslant 1$ are a (t, d)-sequence in base b if every subsequence of the form $\boldsymbol{x}_{(r-1)b^m+1}, \dots, \boldsymbol{x}_{rb^m}$ for integers $m \geqslant t$ and $r \geqslant 1$ is a (t, m, d)-net in base b.

The best available values of t for nets and sequences are recorded in the online resource MinT described in [53], which also includes lower bounds. The Sobol' sequences of [55] are (t, d)-sequences in base $b = 2$. There are newer versions of Sobol's sequence with improved 'direction numbers' in [27, 58]. The Faure sequences [15] have $t = 0$ but require that the base be a prime number $b \geqslant d$. Faure's construction was generalized to prime powers $b \geqslant d$ in [38]. The best presently attainable values of t for base $b = 2$ are in the Niederreiter-Xing sequences of [40, 41].

Randomizations of digital nets and sequences operate by applying certain random permutations to their base b expansions. For details, see the survey in [46]. We will consider the 'nested uniform' scramble from [42].

If $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ is a (t, m, d)-net in base b then after applying a nested uniform scramble, the resulting points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ are a (t, m, d)-net in base b with probability one [42]. If $\boldsymbol{a}_i$ for $i \geqslant 1$ are a (t, d)-sequence in base b then after applying a nested uniform scramble, the resulting points $\boldsymbol{x}_i$ for $i \geqslant 1$ are a (t,d)-sequence in base b with probability one [42]. In either case, each resulting $\boldsymbol{x}_i \sim \mathbb{U}([0,1]^d)$.

If $f \in L^2[0,1]^d$ and $\hat\mu_n^{\mathrm{RQMC}}$ is based on a nested uniform scramble of a (t,d)-sequence in base b with sample sizes $n = b^k$ for integers $k \geqslant m$, then $\mathbb{E}((\hat\mu_n^{\mathrm{RQMC}} - \mu)^2) = o(1/n)$ as $n \to \infty$. It is thus asymptotically better than MC for any f. For smooth enough f, $\mathbb{E}((\hat\mu_n^{\mathrm{RQMC}} - \mu)^2) = O(n^{-3+\epsilon})$ for any $\epsilon > 0$. See [44, 50] for sufficient conditions.

The main result that we will use is as follows. Let $f \in L^2[0,1]^d$ and write σ^2 for the variance of $f(\boldsymbol{x})$ when $\boldsymbol{x} \sim \mathbb{U}[0,1]^d$. Then for a (t,m,d)-net in base b, scrambled as in [42], we have

$$\mathbb{E}((\hat\mu_n^{\mathrm{RQMC}} - \mu)^2) \leqslant \frac{\Gamma\sigma^2}{n} \tag{7}$$

for some $\Gamma < \infty$ [45, Theorem 1]. That is, the RQMC estimate for these scrambled nets cannot have more than Γ times the mean squared error that an MC estimate has. The value of Γ is found using some conservative upper bounds. We can use $\Gamma = b^t[(b+1)/(b-1)]^d$. If $t = 0$, then we can take $\Gamma = [b/(b-1)]^d$, and for $d = 1$ we can take $\Gamma = b^t$. The quantity Γ arises as an upper bound on an infinite set of 'gain coefficients' relating the RQMC variance to the MC variance for parts of a basis expansion of f. The worst case bound $\sigma\sqrt{\Gamma/n}$ for the RQMC root mean squared error does not contain the factor $\log(n)^d$ that makes the QMC worst case error so large for large d and n of practical interest.

4 RQMC laws of large numbers

This section outlines some very simple LLNs for RQMC before going on to prove two SLLN results for scrambled net integration when $f \in L^2[0,1]^d$. The first SLLN requires sample sizes to be of the form rb^m for $1 \leqslant r \leqslant R$ and $m \geqslant 0$ where b is the base of those nets. The second SLLN extends the first one to include all integer sample sizes.

If $f \in \mathrm{BVHK}[0,1]^d$, then there is an SLLN for RQMC from the Koksma-Hlawka inequality (4) when $\Pr(\lim_{n\to\infty} D_n^*(\boldsymbol{x}_1,\dots,\boldsymbol{x}_n) = 0) = 1$. More generally, for Riemann integrable f we get an SLLN for RQMC as an immediate consequence of equation (3).

Theorem 1. *Let* $f : [0,1]^d \to \mathbb{R}$ *be Riemann integrable. For* $i \geqslant 1$*, let* $\boldsymbol{x}_i \in [0,1]^d$ *be RQMC points with* $\Pr(\lim_{n\to\infty} D_n^*(\boldsymbol{x}_1,\dots,\boldsymbol{x}_n) = 0) = 1$*. Then*

$$\Pr\Bigl(\lim_{n\to\infty} \hat\mu_n^{\mathrm{RQMC}} = \mu\Bigr) = 1.$$

Proof. From equation (3),

$$\Pr\Bigl(\lim_{n\to\infty} \hat\mu_n^{\mathrm{RQMC}} = \mu\Bigr) \geqslant \Pr\Bigl(\lim_{n\to\infty} D_n^*(\boldsymbol{x}_1,\dots,\boldsymbol{x}_n) = 0\Bigr) = 1. \quad \square$$

Theorem 1 is not strong enough for some important applications. It does not cover integration problems where the integrand f is not in $\mathrm{BVHK}[0,1]^d$

including many where f is not even Riemann integrable. Integrands with jump discontinuities or kinks (jumps in their gradient) [17, 18, 19, 22] commonly fail to be in BVHK and integrands containing singularities [3, 21, 49, 56] are not even Riemann integrable.

Sobol' [56] noticed that some of his colleagues were using his QMC points with apparent success on problems with integrable singularities and then he initiated a theory in which QMC could be consistent provided the points $\boldsymbol{x}_i$ avoided the singularities in a suitable and problem specific way. Uniform random points show no preference for the region near a singularity no matter where it is and this is enough to get consistent integral estimates on some problems with integrable singularities [3, 48, 49].

In those cases, we can easily get a WLLN, if the integrand is in L^2. The usual results for RQMC show that $\mathbb{E}((\hat{\mu}_n^{\mathrm{RQMC}} - \mu)^2) \to 0$ as $n \to \infty$ for $f \in L^2[0,1]^d$. From that a WLLN follows by Chebychev's inequality. A WLLN proves to be not quite enough for some problems, so we seek an SLLN for scrambled net quadrature.

First we prove a strong law of large numbers for sample sizes equal to rb^m for $1 \leqslant r \leqslant R$ and $b \geqslant 0$ and $f \in L^2[0,1]^d$. These are the best sample sizes to use in a (t,d)-net with values $n = b^m$ being the best of those because they are the smallest sample sizes to properly balance elementary intervals of size b^{t-m}.

Sobol' [57] recommends using sample sizes in a geometric progression such as $n_\ell = 2^\ell$, not an arithmetic one and there is a lengthier discussion of this point in [51]. To see informally how this works, suppose that $|\hat{\mu}_n - \mu| \leqslant An^{-1-\delta}$ for $\delta > 1$ and $A > 0$ while $|\hat{\mu}_n - \hat{\mu}_{n+1}| \geqslant B/n$. The first is an instance of better than $1/n$ error and the second will be common because $\hat{\mu}_{n+1} = \hat{\mu}_n(n/(n+1)) + f(\boldsymbol{x}_{n+1})/(n+1)$. Then $|\hat{\mu}_{n+1} - \mu| \geqslant |\hat{\mu}_{n+1} - \hat{\mu}_n| - |\hat{\mu}_n - \mu|$ and so for large n, $\hat{\mu}_{n+1}$ will commonly be worse than $\hat{\mu}_n$. A rate like $n^{-1-\delta}$ can only be attained on geometrically spaced sample sizes n under conditions in [51].

Theorem 2. *Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots$ be a (t,d)-sequence in base b, with gain coefficients no larger than $\Gamma < \infty$ and randomized as in [42]. Let $f \in L^2[0,1]^d$ with $\int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \mu$. For an integer $R \geqslant 1$, let $\mathcal{N} = \{rb^m \mid 1 \leqslant r \leqslant R, m \geqslant 0\}$. Then*

$$\Pr\Bigl(\lim_{\ell\to\infty} \hat{\mu}_{n_\ell}^{\mathrm{RQMC}} = \mu\Bigr) = 1$$

where n_ℓ for $\ell \geqslant 1$ are the unique elements of $\mathcal{N}$ arranged in increasing order.

Proof. Pick any $\epsilon > 0$. Let $\sigma^2 < \infty$ be the variance of $f(\boldsymbol{x})$ for $\boldsymbol{x} \sim \mathbb{U}[0,1]^d$. First we consider $n_\ell = rb^m$ for $m \geqslant t$ and $1 \leqslant r \leqslant R$. Because $m \geqslant t$, the definition of a (t,d)-sequence implies that

$$\hat{\mu}_{n_\ell}^{\mathrm{RQMC}} = \frac{1}{r}\sum_{j=1}^{r} \hat{\mu}_{\ell,j}$$

where each $\hat{\mu}_{\ell,j}$ is the average of f over a scrambled (t,m,d)-net in base b. We don't know the covariances $\mathrm{cov}(\hat{\mu}_{\ell,j}, \hat{\mu}_{\ell,j'})$ but we can bound them by assuming conservatively that the corresponding correlations are 1. Then

$$\mathrm{var}(\hat\mu_{n_\ell}^{\mathrm{RQMC}}) = \frac{1}{r^2}\sum_{j=1}^{r}\sum_{j'=1}^{r}\mathrm{cov}(\hat\mu_{\ell,j},\hat\mu_{\ell,j'}) \leqslant \mathrm{var}(\hat\mu_{\ell,1}) \leqslant \frac{\Gamma\sigma^2}{n_\ell/r}.$$

Next, by Chebychev's inequality, $\Pr(|\hat\mu_{n_\ell}^{\mathrm{RQMC}}-\mu|\geqslant\epsilon)\leqslant r\Gamma\sigma^2/(n_\ell\epsilon^2)$. Now

$$\begin{aligned}\sum_{\ell=1}^{\infty}\Pr(|\hat\mu_{n_\ell}^{\mathrm{RQMC}}-\mu|\geqslant\epsilon) &\leqslant \sum_{m=0}^{\infty}\sum_{r=1}^{R}\Pr(|\hat\mu_{rb^m}^{\mathrm{RQMC}}-\mu|\geqslant\epsilon)\\ &\leqslant tR+\sum_{m=t}^{\infty}\sum_{r=1}^{R}\frac{\Gamma\sigma^2}{b^m\epsilon^2}. \qquad (8)\end{aligned}$$

The first inequality arises because some samples sizes n_ℓ may have more than one representation of the form rb^m. Because the sum (8) is finite,

$$\Pr(|\hat\mu_{n_\ell}^{\mathrm{RQMC}}-\mu|\geqslant\epsilon \text{ for infinitely many } \ell)=0$$

by the Borel-Cantelli lemma [13, Chapter 2]. Therefore $\Pr\bigl(\lim_{\ell\to\infty}\hat\mu_{n_\ell}^{\mathrm{RQMC}} = \mu\bigr)=1$. ☐

Next we extend this SLLN to a limit as $n\to\infty$ without a restriction to geometrically spaced sample sizes. While geometrically spaced sample sizes should be used, it is interesting to verify this limit as well. The proof method is adapted from the way that Etemadi [14] extends an SLLN for pairwise independent and identically distributed random variables from geometrically spaced sample sizes to all sample sizes.

Theorem 3. *Let $\boldsymbol{x}_1,\boldsymbol{x}_2,\dots$ be a (t,d)-sequence in base b, with gain coefficients no larger than $\Gamma<\infty$ and randomized as in [42]. Let $f\in L^2[0,1]^d$ with $\int_{[0,1]^d}f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=\mu$. Then*

$$\Pr\Bigl(\lim_{n\to\infty}\hat\mu_n^{\mathrm{RQMC}}=\mu\Bigr)=1.$$

Proof. First we suppose that $f(\boldsymbol{x})\geqslant 0$. This is no loss of generality because $f(\boldsymbol{x})=f_+(\boldsymbol{x})-f_-(\boldsymbol{x})$ where $f_+(\boldsymbol{x})=\max(f(\boldsymbol{x}),0)$ and $f_-(\boldsymbol{x})=\max(-f(\boldsymbol{x}),0)$. If $f\in L^2[0,1]^d$ then both $f_\pm\in L^2[0,1]^d$ and an SLLN for $f_\pm$ would imply one for f.

Because $f(\boldsymbol{x}_i)\geqslant 0$, we know that $T(n)\equiv\sum_{i=1}^n f(\boldsymbol{x}_i)$ is nondecreasing in n. Choose $R=b^k$ for $k>1$ and let $\mathcal{N}=\mathcal{N}(R)=\{rb^m \mid 1\leqslant r\leqslant R, m\geqslant 0\}$. For any integer $n\geqslant 1$ define $\bar n=\bar n(n)=\min\{\nu\in\mathcal{N}\mid \nu\geqslant n\}$ and $\underline{n}=\underline{n}(n)=\max\{\nu\in\mathcal{N}\mid\nu\leqslant n\}$. Monotonicity of $T(n)$ combined with $\hat\mu_n^{\mathrm{RQMC}}=T(n)/n$ gives

$$\frac{\underline{n}(n)}{n}\times\hat\mu_{\underline{n}}^{\mathrm{RQMC}}\leqslant\hat\mu_n^{\mathrm{RQMC}}\leqslant\frac{\bar n(n)}{n}\times\hat\mu_{\bar n}^{\mathrm{RQMC}}$$

By Theorem 2, $\Pr(\limsup_{n\to\infty}\hat\mu_{\bar n}^{\mathrm{RQMC}}=\mu)=1$ and $\Pr(\liminf_{n\to\infty}\hat\mu_{\underline{n}}^{\mathrm{RQMC}}=\mu)=1$. What remains is to bound $\bar n/n$ and $\underline{n}/n$.

We can suppose that $n > b^k$. The base b expansion of n is $\sum_{\ell=0}^{L} a_\ell b^\ell$ where $a_\ell = a_\ell(n) \in \{0, 1, \dots, b-1\}$ and $L = L(n) = 1 + \lfloor \log_b(n) \rfloor$ is the smallest number of base b digits required to write n. Choosing $m = L - k + 1$ we know that $\underline{n} \geqslant \nu = b^m \times r$ for $r = \sum_{s=0}^{L-m} a_{m+s} b^s \leqslant b^k = R$. As a result

$$\frac{\underline{n}(n)}{n} \geqslant \frac{\sum_{\ell=L-k+1}^{L} a_\ell b^\ell}{\sum_{\ell=0}^{L} a_\ell b^\ell} \geqslant \frac{\sum_{\ell=L-k+1}^{L} a_\ell b^\ell}{b^{L-k+1} + \sum_{\ell=L-k+1}^{L} a_\ell b^\ell} \geqslant \frac{b^L}{b^{L-k+1} + b^L}.$$

It follows that

$$\Pr\Bigl(\liminf_{n\to\infty} \hat{\mu}_n^{\mathrm{RQMC}} \geqslant (1 + b^{1-k})^{-1} \mu\Bigr) = 1$$

and since we may choose k as large as we like, $\Pr(\liminf_{n\to\infty} \hat{\mu}_n^{\mathrm{RQMC}} \geqslant \mu) = 1$. Similarly, if $n = A_L b^L$ then $n \in \mathcal{N}$ and we may take $\bar{n} = n$. Otherwise, $\bar{n} \leqslant \nu + b^m = (r+1)b^m$ with $r+1 \leqslant R$ and then $\Pr(\liminf_{n\to\infty} \hat{\mu}_n^{\mathrm{RQMC}} \leqslant \mu) = 1$. $\square$

5 An SLLN without square integrability

The SLLN for Monte Carlo only requires that $f \in L^1[0,1]^d$. The results in Section 4 for RQMC require the much stronger condition that $f \in L^2[0,1]^d$. In this section, we narrow the gap by proving an SLLN for scrambled nets when $f \in L^p[0,1]^d$ for any $p > 1$.

The proof is based on the Riesz-Thorin interpolation theorem from [4, Chapter 4]. Let $\mathcal{E}$ be the operator that takes an integrand f and returns the integration error

$$\hat{\mu}_n^{\mathrm{RQMC}} - \mu = \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{x}_i) - \mu.$$

The integration error is a function of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in [0,1]^d$. Together these belong to $[0,1]^{dn}$. Let Ω be the set $[0,1]^{dn}$ equipped with the distribution induced by the scrambled net randomization producing $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$. If $n = b^m$, then $\mathcal{E}$ is a bounded linear operator from $L^2[0,1]^d$ to $L^2(\Omega)$. The norm of $\mathcal{E}$ is

$$\|\mathcal{E}\|_{L^2[0,1]^d \to L^2(\Omega)} = \sup_{\|f\|_2 \leqslant 1} \bigl(\mathbb{E}(\hat{\mu}_n^{\mathrm{RQMC}} - \mu)^2\bigr)^{1/2} \leqslant \sqrt{\Gamma/n}.$$

The operator $\mathcal{E}$ is also a bounded linear operator from $L^1[0,1]^d$ to $L^1(\Omega)$. Here the norm is

$$\|\mathcal{E}\|_{L^1[0,1]^d \to L^1(\Omega)} = \sup_{\|f\|_1 \leqslant 1} \mathbb{E}(|\hat{\mu}_n^{\mathrm{RQMC}} - \mu|) \leqslant \sup_{\|f\|_1 \leqslant 1} |\mu(f)| + \int_{[0,1]^d} |f(\boldsymbol{x})| \,\mathrm{d}\boldsymbol{x} \leqslant 2.$$

By the Riesz-Thorin theorem below, $\mathcal{E}$ is also a bounded linear operator from $L^p[0,1]$ to $L^p(\Omega)$ for any p with $1 \leqslant p \leqslant 2$.

Theorem 4 (Riesz-Thorin)**.** *For $1 \leqslant q_1 \leqslant q_2 < \infty$ and $\theta \in [0,1]$, let $p \geqslant 1$ satisfy*

$$\frac{1}{p} = \frac{1-\theta}{q_1} + \frac{\theta}{q_2}.$$

For probability spaces Θ_1 and Θ_2, let $\mathcal{T}$ be a linear operator from $L^{q_1}(\Theta_1)$ to $L^{q_1}(\Theta_2)$ and at the same time a linear operator from $L^{q_2}(\Theta_1)$ to $L^{q_2}(\Theta_2)$ satisfying

$$\|\mathcal{T}\|_{L^{q_1}(\Theta_1)\to L^{q_1}(\Theta_2)} \leqslant M_1 \quad \textit{and} \quad \|\mathcal{T}\|_{L^{q_2}(\Theta_1)\to L^{q_2}(\Theta_2)} \leqslant M_2.$$

Then $\mathcal{T}$ is a linear operator from $L^p(\Theta_1)$ to $L^p(\Theta_2)$ satisfying

$$\|\mathcal{T}\|_{L^p(\Theta_1)\to L^p(\Theta_2)} \leqslant M_1^{1-\theta} M_2^{\theta}.$$

Proof. This is a special case of Theorem 2.2(b) in [4]. □

Because $1/p$ is a convex combination of $1/q_1$ and $1/q_2$ we must have $q_1 \leqslant p \leqslant q_2$. Our interest is in $q_1 = 1$ and $q_2 = 2$ and $1 \leqslant p \leqslant 2$. The following corollary handles that case.

Corollary 1. *Let $\mathcal{T}$ be a linear operator from $L^1(\Theta_1)$ to $L^1(\Theta_2)$ and at the same time from $L^2(\Theta_1)$ to $L^2(\Theta_2)$ with*

$$\|\mathcal{T}\|_{L^1(\Theta_1)\to L^1(\Theta_2)} \leqslant M_1 \quad \textit{and} \quad \|\mathcal{T}\|_{L^2(\Theta_1)\to L^2(\Theta_2)} \leqslant M_2.$$

Then for $1 \leqslant p \leqslant 2$,

$$\|\mathcal{T}\|_{L^p(\Theta_1)\to L^p(\Theta_2)} \leqslant M_1^{(2-p)/p} M_2^{2(p-1)/p}.$$

Now we are ready to use the Riesz-Thorin theorem to get an SLLN. The operator $\mathcal{T}$ will be the RQMC error $\mathcal{E}$, the space Θ_1 will be $[0,1]^d$ under the uniform distribution and the space Θ_2 will be $[0,1]^{nd}$ under the distribution induced by the RQMC points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$.

Theorem 5. *Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots$ be a (t,d)-sequence in base b, with gain coefficients no larger than $\Gamma < \infty$ and randomized as in [42]. For $p > 1$, let $f \in L^p[0,1]^d$ with $\int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \mu$. Then*

$$\Pr\Bigl(\lim_{n\to\infty} \hat{\mu}_n^{\mathrm{RQMC}} = \mu\Bigr) = 1.$$

Proof. For $p \geqslant 2$, the conclusion follows from Theorem 3 and so we assume now that $1 < p < 2$. Choose any $\epsilon > 0$ and suppose that $n = rb^m$ for $1 \leqslant r \leqslant R < \infty$ and $m \geqslant 0$. The error operator $\mathcal{E}$ for this n satisfies $\|\mathcal{E}\|_{L^1} \leqslant 2$ and $\|\mathcal{E}\|_{L^2} \leqslant (r\Gamma/n)^{1/2}$. Taking $\mathcal{T} = \mathcal{E}$ in Corollary 1,

$$\sup_{\|f\|_p\leqslant 1} \bigl(\mathbb{E}(|\hat{\mu}_n^{\mathrm{RQMC}} - \mu|^p)\bigr)^{1/p} \leqslant 2^{(2-p)/p}\Bigl(\frac{r\Gamma}{n}\Bigr)^{(p-1)/p}$$

from which $\mathbb{E}(|\hat{\mu}_n^{\mathrm{RQMC}} - \mu|^p) \leqslant 2^{2-p}(r\Gamma/n)^{p-1}$ and then

$$\Pr(|\hat{\mu}_n^{\mathrm{RQMC}} - \mu| > \epsilon) \leqslant 2^{2-p}\epsilon^{-p}(r\Gamma)^{p-1}\|f\|_p^p n^{1-p}.$$

This probability has a finite sum over $r = 1, \dots, R$ and $m \geqslant 0$ and so

$$\Pr\Bigl(\lim_{n\to\infty} \hat{\mu}_n^{\mathrm{RQMC}} = \mu\Bigr) = 1$$

when the limit is over $n \in \{rb^m \mid 1 \leqslant r \leqslant R, m \geqslant 0\}$. We have thus established a version of Theorem 2 for $p > 1$ and the extension to the unrestricted limit as $n \to \infty$ uses the same argument as Theorem 3. □

The Riesz-Thorin theorem has been previously used to bound p'th moments in similar problems. See for instance [23, 52, 31].

6 Discussion

We have proved a strong law of large numbers for scrambled digital net integration, first for geometrically spaced sample sizes and a square integrable integrand, then removing the geometric spacing assumption and finally, reducing the squared integrability condition to $\mathbb{E}(|f(\boldsymbol{x})|^p) < \infty$ for some $p > 1$. It is interesting that this strong law for $p > 1$ is obtained before an equally general weak law was found.

There are other ways to scramble digital nets and sequences. The linear scrambles of [36] require less space than the nested uniform scramble. They have the same mean squared discrepancy as the nested uniform scramble [26] and so they might also satisfy an SLLN. A digital shift [34, 46] does not produce the same variance as the nested uniform scramble and it does not satisfy the critically important bound (7) on gain coefficients, so the methods used here would not provide an SLLN for it. The nested uniform scramble is the only one for which central limit theorems have been proved [2, 35].

A second major family of RQMC methods has been constructed from lattice rules [54]. Points $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ on a lattice in $[0,1]^d$ are randomized into $\boldsymbol{x}_i = \boldsymbol{a}_i + \boldsymbol{u} \bmod 1$, for $\boldsymbol{u} \sim \mathbb{U}[0,1]^d$. That is, they are shifted with wraparound in what is known as a Cranley-Patterson rotation [7]. For an extensible version of shifted lattice rules, see [25]. The Cranley-Patterson rotation does not provide a Γ bound like (7) because there are functions $f \in L^2[0,1]^d$ with $\mathrm{var}(\hat{\mu}_n^{\mathrm{RQMC}}) = \sigma^2(f)$ [34], and so a proof of an SLLN for this form of RQMC would require a different approach. The fact that $\mathrm{var}(\hat{\mu}_n^{\mathrm{RQMC}}) = \sigma^2(f)$ is possible does not provide a counter-example to an SLLN because this equality might only hold for a finite number of n_ℓ in the infinite sequence. Given a class of functions $\mathcal{F}$ with $\mathrm{var}(\hat{\mu}_{n_\ell}^{\mathrm{RQMC}}) \leqslant B\sigma^2(f)/n_\ell$ for all $f \in \mathcal{F}$, all $\ell \geqslant 1$, and some $B < \infty$, we get an SLLN for $f \in \mathcal{F}$ if $\sum_{\ell=1}^{\infty} 1/n_\ell < \infty$. Some such bounds B for randomly shifted lattices appear in [34] though they hold for specific n_ℓ not necessarily an infinite sequence of them.

Acknowledgments

We thank Max Balandat for posing the problem of finding and SLLN for randomized QMC. Thanks also to Ektan Bakshy, Wei-Liem Loh and Fred Hickernell for discussions. This work was supported by grant IIS-1837931 from the U.S. National Science Foundation.

References

[1] Ch. Aistleitner and J. Dick. Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality. *Acta Arithmetica*, 167(2):143–171, 2015.

[2] K. Basu and R. Mukherjee. Asymptotic normality of scrambled geometric net quadrature. *The Annals of Statistics*, 45(4):1759–1788, 2017.

[3] K. Basu and A. B. Owen. Quasi-Monte Carlo for an integrand with a singularity along a diagonal in the square. In J. Dick, F. Y. Kuo, and H. Woźniakowski, editors, *Contemporary Computational Mathematics-A Celebration of the 80th Birthday of Ian Sloan*, pages 119–130. Springer, 2018.

[4] C. Bennett and R. Sharpley. *Interpolation of Operators*, volume 129 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1988.

[5] W. Chen, A. Srivastav, and G. Travaglini, editors. *A Panorama of Discrepancy Theory*. Springer, Cham, Switzerland, 2014.

[6] J. A. Clarkson and C. R. Adams. On definitions of bounded variation for functions of two variables. *Transactions of the American Mathematical Society*, 35(4):824–854, 1933.

[7] R. Cranley and T.N.L. Patterson. Randomization of number theoretic methods for multiple integration. *SIAM Journal of Numerical Analysis*, 13:904–914, 1976.

[8] P. J. Davis and P. Rabinowitz. *Methods of Numerical Integration (2nd Ed.)*. Academic Press, San Diego, 1984.

[9] Luc Devroye. *Non-uniform Random Variate Generation*. Springer, 1986.

[10] J. Dick, F. Y. Kuo, and I. H. Sloan. High-dimensional integration: the quasi-Monte Carlo way. *Acta Numerica*, 22:133–288, 2013.

[11] J. Dick and F. Pillichshammer. *Digital sequences, discrepancy and quasi-Monte Carlo integration*. Cambridge University Press, Cambridge, 2010.

[12] C. Doerr, M. Gnewuch, and M. Wahlström. Calculation of discrepancy measures and applications. In Chen W., Srivastav A., and Travaglini G., editors, *A Panorama of Discrepancy Theory*, pages 621–678. Springer, 2014.

[13] R. Durrett. *Probability Theory and Examples*. Cambridge University Press, Cambridge, 2019.

[14] N. Etemadi. An elementary proof of the strong law of large numbers. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 55(1):119–122, 1981.

[15] H. Faure. Discrépance de suites associées à un système de numération (en dimension s). *Acta Arithmetica*, 41:337–351, 1982.

[16] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer Science & Business Media, New York, 2003.

[17] M. Griebel, F. Y. Kuo, and I. H. Sloan. The smoothing effect of integration in $\mathbb{R}^d$ and the ANOVA decomposition. *Mathematics of Computation*, 82(281):383–400, 2013.

[18] M. Griebel, F. Y. Kuo, and I. H. Sloan. Note on "The smoothing effect of integration in $\mathbb{R}^d$ and the ANOVA decomposition". *Mathematics of Computation*, 86(306):1847–1854, 2017.

[19] A. Griewank, F. Y. Kuo, H. Leövey, and I. H. Sloan. High dimensional integration of kinks and jumps–Smoothing by preintegration. *Journal of Computational and Applied Mathematics*, 344:259–274, 2018.

[20] G. H. Hardy. On double Fourier series, and especially those which represent the double zeta-function with real and incommensurable parameters. *Quarterly Journal of Mathematics*, 37:53–79, 1905.

[21] J. Hartinger and R. Kainhofer. Non-uniform low-discrepancy sequence generation and integration of singular integrands. In H. Niederreiter and D. Talay, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pages 163–179. Springer, 2006.

[22] Z. He and X. Wang. On the convergence rate of randomized quasi–Monte Carlo for discontinuous functions. *SIAM Journal on Numerical Analysis*, 53(5):2488–2503, 2015.

[23] S. Heinrich. Random approximation in numerical analysis. *Proceedings of the Conference "Functional Analysis" Essen*, pages 123–171, 1994.

[24] F. J. Hickernell. Koksma-Hlawka inequality. *Wiley StatsRef: Statistics Reference Online*, 2014.

[25] F. J. Hickernell, H. S. Hong, P. L'Ecuyer, and C. Lemieux. Extensible lattice sequences for quasi-Monte Carlo quadrature. *SIAM Journal on Scientific Computing*, 22(3):1117–1138, 2000.

[26] F. J. Hickernell and R. X. Yue. The mean square discrepancy of scrambled (t, s)-sequences. *SIAM Journal of Numerical Analysis*, 38:1089–1112, 2000.

[27] S. Joe and F. Y. Kuo. Constructing Sobol' sequences with better two-dimensional projections. *SIAM Journal on Scientific Computing*, 30(5):2635–2654, 2008.

[28] Alexander Keller. A quasi-Monte Carlo algorithm for the global illumination problem in a radiosity setting. In H. Niederreiter and P. Jau-Shyong Shiue, editors, *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, pages 239–251, New York, 1995. Springer-Verlag.

[29] M. Krause. Über Fouriersche Reihen mit zwei veränderlichen Größen. *Leipziger Ber.*, 55:164–197, 1903.

[30] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences.* Wiley, New York, 1974.

[31] R. J. Kunsch, E. Novak, and D. Rudolf. Solvable integration problems and optimal sample size selection. *Journal of Complexity*, 53:40–67, 2019.

[32] F. Y. Kuo and D. Nuyens. Application of quasi-Monte Carlo methods to elliptic PDEs with random diffusion coefficients: a survey of analysis and implementation. *Foundations of Computational Mathematics*, 16(6):1631–1696, 2016.

[33] Boris Lavrič. Continuity of monotone functions. *Archivum Mathematicum*, 29(1):1–4, 1993.

[34] P. L'Ecuyer and C. Lemieux. A survey of randomized quasi-Monte Carlo methods. In M. Dror, P. L'Ecuyer, and F. Szidarovszki, editors, *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, pages 419–474. Kluwer Academic Publishers, 2002.

[35] W.-L. Loh. On the asymptotic distribution of scrambled net quadrature. *Annals of Statistics*, 31(4):1282–1324, 2003.

[36] J. Matoušek. On the L^2–discrepancy for anchored boxes. *Journal of Complexity*, 14:527–556, 1998.

[37] H. Niederreiter. Pseudo-random numbers and optimal coefficients. *Advances in Mathematics*, 26:99–181, 1977.

[38] H. Niederreiter. Point sets and sequences with small discrepancy. *Monatshefte fur mathematik*, 104:273–337, 1987.

[39] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods.* S.I.A.M., Philadelphia, PA, 1992.

[40] H. Niederreiter and C. Xing. Low-discrepancy sequences and global function fields with many rational places. *Finite Fields and Their Applications*, 2:241–273, 1996.

[41] H. Niederreiter and C. Xing. Quasirandom points and global function fields. In S. Cohen and H. Niederreiter, editors, *Finite Fields and Applications*, volume 233, pages 269–296, Cambridge, 1996. Cambridge University Press.

[42] A. B. Owen. Randomly permuted (t, m, s)-nets and (t, s)-sequences. In H. Niederreiter and P. Jau-Shyong Shiue, editors, *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, pages 299–317, New York, 1995. Springer-Verlag.

[43] A. B. Owen. Monte Carlo variance of scrambled equidistribution quadrature. *SIAM Journal of Numerical Analysis*, 34(5):1884–1910, 1997.

[44] A. B. Owen. Scrambled net variance for integrals of smooth functions. *Annals of Statistics*, 25(4):1541–1562, 1997.

[45] A. B. Owen. Scrambling Sobol' and Niederreiter-Xing points. *Journal of Complexity*, 14(4):466–489, December 1998.

[46] A. B. Owen. Variance with alternative scramblings of digital nets. *ACM Transactions on Modeling and Computer Simulation*, 13(4):363–378, 2003.

[47] A. B. Owen. Multidimensional variation for quasi-Monte Carlo. In J. Fan and G. Li, editors, *International Conference on Statistics in honour of Professor Kai-Tai Fang's 65th birthday*, 2005.

[48] A. B. Owen. Halton sequences avoid the origin. *SIAM Review*, 48:487–583, 2006.

[49] A. B. Owen. Randomized QMC and point singularities. In H. Niederreiter and D. Talay, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pages 403–418. Springer, 2006.

[50] A. B. Owen. Local antithetic sampling with scrambled nets. *Annals of Statistics*, 36(5):2319–2343, 2008.

[51] A. B. Owen. A constraint on extensible quadrature rules. *Numerische Mathematik*, 132(3):511–518, 2016.

[52] D. Rudolf and N. Schweizer. Error bounds of MCMC for functions with unbounded stationary variance. *Statistics & Probability Letters*, 99:6–12, 2015.

[53] R. Schürer and W. Ch. Schmid. MinT: new features and new results. In P. L'Ecuyer and A. B. Owen, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2008*. Springer, 2009.

[54] I. H. Sloan and S. Joe. *Lattice Methods for Multiple Integration.* Oxford Science Publications, Oxford, 1994.

[55] I. M. Sobol'. The distribution of points in a cube and the accurate evaluation of integrals (in Russian). *Zh. Vychisl. Mat. i Mat. Phys.*, 7:784–802, 1967.

[56] I. M. Sobol'. Calculation of improper integrals using uniformly distributed sequences. *Soviet Math Dokl*, 14(3):734–738, 1973.

[57] I. M. Sobol'. On quasi-Monte Carlo integrations. *Mathematics and Computers in Simulation*, 47:103–112, 1998.

[58] I. M. Sobol', D. Asotsky, A. Kreinin, and S. Kucherenko. Construction and comparison of high-dimensional Sobol' generators. *Wilmott magazine*, 2011(56):64–79, 2011.