

# A Strong Law of Large Numbers for Scrambled Net Integration\*

Art B. Owen<sup>†</sup>  
Daniel Rudolf<sup>‡</sup>

**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.

**Key words.** cubature, quadrature, quasi–Monte Carlo, randomized quasi–Monte Carlo, simulation-optimization

**AMS subject classifications.** 65D32, 65C05

**DOI.** 10.1137/20M1320535

**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 [10]. For nonsmooth integrands or higher-dimensional domains, these methods 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 [12, 13, 41] 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 LLN to ensure that the estimate converges to the integral as  $n \rightarrow \infty$ . Such consistency is a minimal requirement of an integration method. For QMC, consistency requires additional assumptions of Riemann integrability or bounded variation, whose

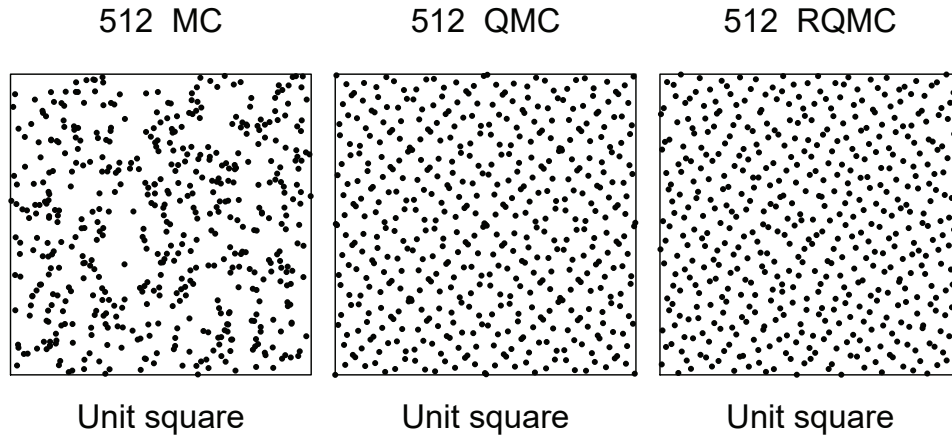
\*Received by the editors February 20, 2020; accepted for publication (in revised form) June 29, 2020; published electronically May 6, 2021.

<https://doi.org/10.1137/20M1320535>

**Funding:** This work was supported by grant IIS-1837931 from the U.S. National Science Foundation.

<sup>†</sup>Department of Statistics, Stanford University, Stanford, CA 94305 USA (owen@stanford.edu, <http://statweb.stanford.edu/~owen/>).

<sup>‡</sup>Felix-Bernstein-Institute for Mathematical Statistics, Universität Göttingen, 37077 Göttingen, Germany (daniel.rudolf@uni-goettingen.de, <http://stochastik.math.uni-goettingen.de/~drudolf/>).



**Fig. 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.

descriptions 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 [18], graphical rendering [30], and solution of PDEs in random environments [34].

In addition to knowing that a method would work as  $n \rightarrow \infty$ , 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 [36], 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 [44, 45], provides estimated integrals that are consistent as  $n \rightarrow \infty$  under weaker conditions than plain QMC requires. It can also reduce the root mean squared error to  $O(n^{-3/2+\epsilon})$  [46, 52] under further conditions on the integrand.

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.

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 the PyTorch [3] tool for Bayesian optimization. A prototypical Bayesian optimization problem is to find  $\min_{\theta \in \Theta} \int_{\mathbb{R}^d} g(\theta, \mathbf{x}) d\mathbf{x}$  for some function  $g(\theta, \mathbf{x})$  and a set  $\Theta$  of allowed values for a parameter  $\theta$ . In a simulation-optimization framework [2] the integral over  $\mathbf{x}$  values may be approximated by a Monte Carlo average. Integration is then a building block in a larger problem. PyTorch has a version using RQMC points instead of plain Monte Carlo. Consistent estimation of the optimal  $\theta$  could be proved assuming a strong LLN for some sample values  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . Such a strong law was available for plain Monte Carlo but not for RQMC, yet RQMC shows much better empirical results in [3].

An outline of this paper is as follows. Section 2 presents the strong and weak LLNs referred to above as well as plain Monte Carlo 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 that is 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 LLN 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  $n \rightarrow \infty$  through the set of values that can be written as  $n = rb^m$  for  $r = 1, \dots, R$  using some integers  $m \geq 0$ ,  $b \geq 2$ , and  $R \geq 1$ . While those are the best sample sizes to use for reasons given in that section, we next extend the result to the ordinary limit as  $n \rightarrow \infty$  through all integer values. Section 5 replaces the assumption that  $f^2$  is integrable by one that  $|f|^p$  has a finite integral for some  $p > 1$ . This result uses the Riesz–Thorin interpolation theorem [6]. 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  $d \geq 1$ . For  $p \geq 1$ , the space  $L^p[0, 1]^d$  consists of all measurable functions  $f$  on  $[0, 1]^d$  for which  $\|f\|_p = (\int_{[0, 1]^d} |f(\mathbf{x})|^p d\mathbf{x})^{1/p} < \infty$ . We consider the problem of computing an estimate  $\hat{\mu}$  of the integral  $\mu = \int_{[0, 1]^d} f(\mathbf{x}) d\mathbf{x}$ . Here  $\mu$  is the expected value of  $f(\mathbf{x})$  when  $\mathbf{x}$  has the uniform distribution on  $[0, 1]^d$ . We write  $\mu = \mathbb{E}(f(\mathbf{x}))$  for  $\mathbf{x} \sim \mathbb{U}[0, 1]^d$  and use  $\Pr(A)$  below to denote the probability of the event  $A$ . Many problems that do not originate as integrals over  $[0, 1]^d$  have such a representation using transformations to generate nonuniformly distributed random variables over the cube and other spaces [11]. 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  $\mathbf{x}_i \sim \mathbb{U}[0, 1]^d$  and estimates  $\mu$  by  $\hat{\mu}_n = \hat{\mu}_n^{\text{MC}} = (1/n) \sum_{i=1}^n f(\mathbf{x}_i)$ . 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  $\epsilon > 0$ ,

$$(2.1) \quad \lim_{n \rightarrow \infty} \Pr(|\hat{\mu}_n^{\text{MC}} - \mu| > \epsilon) = 0.$$

The strong law of large numbers (SLLN) implies that

$$(2.2) \quad \Pr\left(\lim_{n \rightarrow \infty} \hat{\mu}_n^{\text{MC}} = \mu\right) = 1,$$

which we may write as  $\Pr(\limsup_{n \rightarrow \infty} |\hat{\mu}_n^{\text{MC}} - \mu| > \epsilon) = 0$  to parallel the WLLN. Both the WLLN and SLLN hold for independent and identically distributed (IID) random variables  $f(\mathbf{x}_i)$  when  $f \in L^1[0, 1]^d$ . For proofs of these laws, see [15, 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  $\mathbf{x}_i$  are constructed so that the discrete distribution placing probability  $1/n$  on each of  $\mathbf{x}_1, \dots, \mathbf{x}_n$  (with repeated points counted multiple times) is close to the continuous uniform distribution on  $[0, 1]^d$ . There are various ways, called discrepancies [7], to quantify the distance between these discrete and continuous

measures. For a set  $S \subset [0, 1]^d$  define  $\mathbf{1}\{\mathbf{x} \in S\}$  to be 1 if  $\mathbf{x} \in S$ , and 0 otherwise. The most widely used discrepancy is the star discrepancy

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

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

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 [12, 13, 41, 36] among others.

Because QMC is deterministic, it has no analogue of the WLLN (2.1). There is an analogue of the SLLN (2.2), as follows. Let  $\hat{\mu}_n^{\text{QMC}} = (1/n) \sum_{i=1}^n f(\mathbf{x}_i)$ , where now the points  $\mathbf{x}_i$  have been chosen to have small discrepancy. If  $f$  is Riemann integrable and  $D_n^* \rightarrow 0$ , then [32, p. 3]

$$(2.3) \quad \lim_{n \rightarrow \infty} \hat{\mu}_n^{\text{QMC}} = \mu,$$

providing the QMC version of the SLLN (2.2). There is a converse, where if  $|\hat{\mu}_n - \mu| \rightarrow 0$  whenever  $D_n^* \rightarrow 0$ , then  $f$  must be Riemann integrable. See the references and discussion in [39]. 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  $V(f) = \int_0^1 |f'(x)| dx$  and 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 [8]). Of those, the total variation in the sense of Hardy [22] and Krause [31], denoted by  $V_{\text{HK}}(f)$ , is the most useful one for QMC. If  $V_{\text{HK}}(f) < \infty$ , then we write  $f \in \text{BVHK}[0, 1]^d$ . Although we don't need  $f$  to have bounded variation to get the SLLN (2.3) for QMC, bounded variation gives us some information on the rate of convergence, via the Koksma–Hlawka inequality

$$(2.4) \quad |\hat{\mu}_n^{\text{QMC}} - \mu| \leq D_n^*(\mathbf{x}_1, \dots, \mathbf{x}_n) V_{\text{HK}}(f)$$

(see [26]). Typical QMC constructions provide infinite sequences  $\mathbf{x}_i$  whose initial subsequences satisfy

$$D_n^*(\mathbf{x}_1, \dots, \mathbf{x}_n) = O\left(\frac{\log(n)^d}{n}\right).$$

Then  $|\hat{\mu}_n^{\text{QMC}} - \mu| = O(n^{-1+\epsilon})$  by (2.4) for any  $\epsilon > 0$ .

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

$$(2.5) \quad \mathbb{E}((\hat{\mu}_n^{\text{MC}} - \mu)^2)^{1/2} = n^{-1/2} \sigma(f)$$

when, for  $\mathbf{x} \sim \mathbb{U}[0, 1]^d$ , the variance of  $f(\mathbf{x})$  is  $\sigma^2 = \sigma^2(f) = \mathbb{E}((f(\mathbf{x}) - \mu)^2) < \infty$ . Whereas 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 (2.5) is exact, while the QMC version (2.4) is an extremely conservative upper bound, in that it covers even the worst  $f \in \text{BVHK}[0, 1]^d$  for any given  $\mathbf{x}_1, \dots, \mathbf{x}_n$ .

A Riemann integrable function is not necessarily in BVHK. For instance,  $f(\mathbf{x}) = \mathbf{1}\{\sum_{j=1}^d x_j \leq 1\}$  is Riemann integrable but, for  $d \geq 2$ , it is not in BVHK [49]. 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.

LEMMA 2.1. *If  $f \in \text{BVHK}[0, 1]^d$ , then  $f$  is also Riemann integrable.*

*Proof.* If  $f \in \text{BVHK}[0, 1]^d$ , then  $f(\mathbf{x}) = f(\mathbf{0}) + f_+(\mathbf{x}) - f_-(\mathbf{x})$ , where  $f_{\pm}$  are 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 [35].  $\square$

While QMC has a superior convergence rate to MC for  $f \in \text{BVHK}$ , MC has an advantage over QMC in that  $\mathbb{E}((\hat{\mu}^{\text{MC}} - \mu)^2) = \sigma^2/n$  is simple to estimate from independent replicates, while  $D_n^*$  is very expensive to compute [14] and  $V_{\text{HK}}(f)$  is much harder to estimate than  $\mu$ . In a setting where attaining accuracy is important, it must also be important to estimate the attained accuracy. RQMC methods, described next, are hybrids of MC and QMC that support error estimation.

In RQMC [36, 44] one starts with points  $\mathbf{a}_1, \dots, \mathbf{a}_n \in [0, 1]^d$  having a small star discrepancy and randomizes them to produce points  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . These points satisfy the following two conditions: individually  $\mathbf{x}_i \sim \mathbb{U}[0, 1]^d$ , and collectively,  $\mathbf{x}_1, \dots, \mathbf{x}_n$  have small star discrepancy. The RQMC estimate of  $\mu$  is  $\hat{\mu}_n^{\text{RQMC}} = (1/n) \sum_{i=1}^n f(\mathbf{x}_i)$ . From the uniformity of the points  $\mathbf{x}_i$  we find that  $\mathbb{E}(\hat{\mu}_n^{\text{RQMC}}) = \mu$ . Their small star discrepancy means that they are also QMC points, and so they inherit the accuracy properties of QMC. To estimate the error, one takes several independent randomizations of  $\mathbf{a}_i$  producing independent replicates of  $\hat{\mu}_n^{\text{RQMC}}$  whose sample variance can be computed.

**3. Scrambled Nets and Sequences.** In this section, we describe digital nets and sequences and scrambled versions of them. Many authors reserve the term “digital” to mean only points obtained from some certain specific classes of algorithms. Since the overwhelming majority of nets and sequences in use are constructed with such algorithms, we lose little by this simplification.

Let  $b \geq 2$  be an integer base. Let  $\mathbf{k} = (k_1, \dots, k_d)$  for integers  $k_j \geq 0$  and  $\mathbf{c} = (c_1, \dots, c_d)$ , where  $c_j \in \{0, 1, \dots, b^{k_j} - 1\}$ . Then the set

$$(3.1) \quad E(\mathbf{k}, \mathbf{c}) = \prod_{j=1}^d \left[ \frac{c_j}{b^{k_j}}, \frac{c_j + 1}{b^{k_j}} \right)$$

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

DEFINITION 3.1. *For integers  $m \geq t \geq 0$ ,  $b \geq 2$ , and  $d \geq 1$ , the points  $\mathbf{x}_1, \dots, \mathbf{x}_n \in [0, 1]^d$  for  $n = b^m$  are a  $(t, m, d)$ -net in base  $b$  if*

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

*holds for every elementary interval  $E(\mathbf{k}, \mathbf{c})$  from (3.1) with  $|\mathbf{k}| \leq m - t$ .*

An elementary interval of volume  $b^{-|\mathbf{k}|}$  should ideally contain  $nb^{-|\mathbf{k}|} = b^{m-|\mathbf{k}|}$  points from  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . In a  $(t, m, d)$ -net in base  $b$ , 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 3.2.** For integers  $t \geq 0$ ,  $b \geq 2$ , and  $d \geq 1$ , the points  $\mathbf{x}_i \in [0, 1]^d$  for  $i \geq 1$  are a  $(t, d)$ -sequence in base  $b$  if every subsequence of the form  $\mathbf{x}_{(r-1)b^{m+1}+1}, \dots, \mathbf{x}_{rb^m}$  for integers  $m \geq t$  and  $r \geq 1$  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 [55], which also includes lower bounds. The Sobol' sequences of [57] are  $(t, d)$ -sequences in base  $b = 2$ . There are newer versions of Sobol' sequence with improved "direction numbers" in [29, 60]. The Faure sequences [17] have  $t = 0$  but require that the base be a prime number  $b \geq d$ . Faure's construction was generalized to prime powers  $b \geq d$  in [40]. The best presently attainable values of  $t$  for base  $b = 2$  are in the Niederreiter–Xing sequences of [42, 43].

Randomizations of digital nets and sequences operate by applying certain random permutations to their base  $b$  expansions. For details, see the survey in [48]. We will consider the "nested uniform" scramble from [44].

If  $\mathbf{a}_1, \dots, \mathbf{a}_n$  is a  $(t, m, d)$ -net in base  $b$ , then after applying a nested uniform scramble, the resulting points  $\mathbf{x}_1, \dots, \mathbf{x}_n$  are a  $(t, m, d)$ -net in base  $b$  with probability one [44]. If  $\mathbf{a}_i$  for  $i \geq 1$  are a  $(t, d)$ -sequence in base  $b$ , then after applying a nested uniform scramble, the resulting points  $\mathbf{x}_i$  for  $i \geq 1$  are a  $(t, d)$ -sequence in base  $b$  with probability one [44]. In either case, each resulting point satisfies  $\mathbf{x}_i \sim \mathcal{U}[0, 1]^d$ .

If  $f \in L^2[0, 1]^d$  and  $\hat{\mu}_n^{\text{RQMC}}$  is based on a nested uniform scramble of a  $(t, d)$ -sequence in base  $b$  with sample sizes  $n = b^k$  for integers  $k \geq 0$ , then  $\mathbb{E}((\hat{\mu}_n^{\text{RQMC}} - \mu)^2) = o(1/n)$  as  $n \rightarrow \infty$ . It is thus asymptotically better than MC for any  $f$ . For smooth enough  $f$ ,  $\mathbb{E}((\hat{\mu}_n^{\text{RQMC}} - \mu)^2) = O(n^{-3+\epsilon})$  for any  $\epsilon > 0$ . See [46, 52] for sufficient conditions.

The main result that we will use is as follows. Let  $f \in L^2[0, 1]^d$  and write  $\sigma^2$  for the variance of  $f(\mathbf{x})$  when  $\mathbf{x} \sim \mathcal{U}[0, 1]^d$ . Then for a  $(t, m, d)$ -net in base  $b$ , scrambled as in [44], we have

$$(3.2) \quad \mathbb{E}((\hat{\mu}_n^{\text{RQMC}} - \mu)^2) \leq \frac{\Gamma \sigma^2}{n}$$

for some  $\Gamma < \infty$  [47, Theorem 1]. That is, the RQMC estimate for these scrambled nets cannot have more than  $\Gamma$  times the mean squared error that an MC estimate has. The value of  $\Gamma$  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  $\Gamma$  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  $1 \leq r \leq R$  and  $m \geq 0$ , where  $b$  is the base of those nets. The second SLLN extends the first one to include all integer sample sizes.

If  $f \in \text{BVHK}[0, 1]^d$ , then there is an SLLN for RQMC from the Koksma–Hlawka inequality (2.4) when  $\Pr(\lim_{n \rightarrow \infty} D_n^*(\mathbf{x}_1, \dots, \mathbf{x}_n) = 0) = 1$ . More generally, for Riemann integrable  $f$  we get an SLLN for RQMC as an immediate consequence of (2.3).

THEOREM 4.1. Let  $f : [0, 1]^d \rightarrow \mathbb{R}$  be Riemann integrable. For  $i \geq 1$ , let  $\mathbf{x}_i \in [0, 1]^d$  be RQMC points with  $\Pr(\lim_{n \rightarrow \infty} D_n^*(\mathbf{x}_1, \dots, \mathbf{x}_n) = 0) = 1$ . Then

$$\Pr\left(\lim_{n \rightarrow \infty} \hat{\mu}_n^{\text{RQMC}} = \mu\right) = 1.$$

*Proof.* From (2.3),

$$\Pr\left(\lim_{n \rightarrow \infty} \hat{\mu}_n^{\text{RQMC}} = \mu\right) \geq \Pr\left(\lim_{n \rightarrow \infty} D_n^*(\mathbf{x}_1, \dots, \mathbf{x}_n) = 0\right) = 1.$$

Theorem 4.1 is not strong enough for some important applications. It does not cover integration problems where the integrand  $f$  is not in  $\text{BVHK}[0, 1]^d$ , including many where  $f$  is not even Riemann integrable. Integrands with jump discontinuities or kinks (jumps in their gradient) [19, 20, 21, 24] commonly fail to be in  $\text{BVHK}$ , and integrands containing singularities [5, 23, 51, 58] are not even Riemann integrable.

Sobol' [58], after noticing that some of his colleagues were using his QMC points with apparent success on problems with integrable singularities, initiated a theory in which QMC could be consistent provided the points  $\mathbf{x}_i$  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 [5, 50, 51].

In those cases, we can easily get a WLLN if the integrand is in  $L^2$ . The usual results for RQMC show that  $\mathbb{E}((\hat{\mu}_n^{\text{RQMC}} - \mu)^2) \rightarrow 0$  as  $n \rightarrow \infty$  for  $f \in L^2[0, 1]^d$ . From that a WLLN follows by Chebyshev'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 an SLLN for sample sizes equal to  $rb^m$  for  $1 \leq r \leq R$  and  $b \geq 2$  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' [59] recommends using sample sizes in a geometric, rather than an arithmetic, progression such as  $n_\ell = 2^\ell$ , and there is a lengthier discussion of this point in [53]. To see informally how this works, suppose that  $|\hat{\mu}_n - \mu| \leq An^{-1-\delta}$  for  $\delta > 1$  and  $A > 0$ , while  $|\hat{\mu}_n - \hat{\mu}_{n+1}| \geq B/n$ . The first is an instance of better than  $1/n$  error, and the second will be common because  $\hat{\mu}_{n+1} = \hat{\mu}_n(n/(n+1)) + f(\mathbf{x}_{n+1})/(n+1)$ . Then  $|\hat{\mu}_{n+1} - \mu| \geq |\hat{\mu}_{n+1} - \hat{\mu}_n| - |\hat{\mu}_n - \mu|$ , and so for large  $n$ ,  $\hat{\mu}_{n+1}$  will commonly be worse than  $\hat{\mu}_n$ . A rate like  $n^{-1-\delta}$  can be attained only on geometrically spaced sample sizes  $n$  under conditions in [53].

THEOREM 4.2. Let  $\mathbf{x}_1, \mathbf{x}_2, \dots$  be a  $(t, d)$ -sequence in base  $b$ , with gain coefficients no larger than  $\Gamma < \infty$  and randomized as in [44]. Let  $f \in L^2[0, 1]^d$  with  $\int_{[0, 1]^d} f(\mathbf{x}) d\mathbf{x} = \mu$ . For an integer  $R \geq 1$ , let  $\mathcal{N} = \{rb^m \mid 1 \leq r \leq R, m \geq 0\}$ . Then

$$\Pr\left(\lim_{\ell \rightarrow \infty} \hat{\mu}_{n_\ell}^{\text{RQMC}} = \mu\right) = 1,$$

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

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

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

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  $\text{cov}(\hat{\mu}_{\ell,j}, \hat{\mu}_{\ell,j'})$ , but we can bound them by assuming conservatively that the corresponding correlations are 1. Then

$$\text{var}(\hat{\mu}_{n_\ell}^{\text{RQMC}}) = \frac{1}{r^2} \sum_{j=1}^r \sum_{j'=1}^r \text{cov}(\hat{\mu}_{\ell,j}, \hat{\mu}_{\ell,j'}) \leq \text{var}(\hat{\mu}_{\ell,1}) \leq \frac{\Gamma \sigma^2}{n_\ell/r}.$$

Next, by Chebyshev's inequality,  $\Pr(|\hat{\mu}_{n_\ell}^{\text{RQMC}} - \mu| \geq \epsilon) \leq r\Gamma\sigma^2/(n_\ell\epsilon^2)$ . Now

$$\begin{aligned} \sum_{\ell=1}^{\infty} \Pr(|\hat{\mu}_{n_\ell}^{\text{RQMC}} - \mu| \geq \epsilon) &\leq \sum_{m=0}^{\infty} \sum_{r=1}^R \Pr(|\hat{\mu}_{rb^m}^{\text{RQMC}} - \mu| \geq \epsilon) \\ (4.1) \quad &\leq tR + \sum_{m=t}^{\infty} \sum_{r=1}^R \frac{\Gamma\sigma^2}{b^m\epsilon^2}. \end{aligned}$$

The first inequality arises because some sample sizes  $n_\ell$  may have more than one representation of the form  $rb^m$ . Because the sum (4.1) is finite,

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

by the Borel–Cantelli lemma [15, Chapter 2]. Therefore  $\Pr(\lim_{\ell \rightarrow \infty} \hat{\mu}_{n_\ell}^{\text{RQMC}} = \mu) = 1$ .  $\square$

Next we extend this SLLN to a limit as  $n \rightarrow \infty$  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 [16] extends an SLLN for pairwise IID random variables from geometrically spaced sample sizes to all sample sizes.

**THEOREM 4.3.** *Let  $\mathbf{x}_1, \mathbf{x}_2, \dots$  be a  $(t, d)$ -sequence in base  $b$ , with gain coefficients no larger than  $\Gamma < \infty$  and randomized as in [44]. Let  $f \in L^2[0, 1]^d$  with  $\int_{[0,1]^d} f(\mathbf{x}) \, d\mathbf{x} = \mu$ . Then*

$$\Pr\left(\lim_{n \rightarrow \infty} \hat{\mu}_n^{\text{RQMC}} = \mu\right) = 1.$$

*Proof.* First we suppose that  $f(\mathbf{x}) \geq 0$ . There is no loss of generality because  $f(\mathbf{x}) = f_+(\mathbf{x}) - f_-(\mathbf{x})$ , where  $f_+(\mathbf{x}) = \max(f(\mathbf{x}), 0)$  and  $f_-(\mathbf{x}) = \max(-f(\mathbf{x}), 0)$ . 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  $f(\mathbf{x}_i) \geq 0$ , we know that  $T(n) \equiv \sum_{i=1}^n f(\mathbf{x}_i)$  is nondecreasing in  $n$ . Choose  $R = b^k$  for  $k > 1$  and let  $\mathcal{N} = \mathcal{N}(R) = \{rb^m \mid 1 \leq r \leq R, m \geq 0\}$ . For any integer  $n \geq 1$  define  $\bar{n} = \bar{n}(n) = \min\{\nu \in \mathcal{N} \mid \nu \geq n\}$  and  $\underline{n} = \underline{n}(n) = \max\{\nu \in \mathcal{N} \mid \nu \leq n\}$ . Monotonicity of  $T(n)$  combined with  $\hat{\mu}_n^{\text{RQMC}} = T(n)/n$  gives

$$\frac{\underline{n}(n)}{n} \hat{\mu}_{\underline{n}}^{\text{RQMC}} \leq \hat{\mu}_n^{\text{RQMC}} \leq \frac{\bar{n}(n)}{n} \hat{\mu}_{\bar{n}}^{\text{RQMC}}.$$

By Theorem 4.2,  $\Pr(\limsup_{n \rightarrow \infty} \hat{\mu}_{\bar{n}}^{\text{RQMC}} = \mu) = 1$  and  $\Pr(\liminf_{n \rightarrow \infty} \hat{\mu}_{\underline{n}}^{\text{RQMC}} = \mu) = 1$ . 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  $\underline{n} \geq \nu = rb^m$  for  $r = \sum_{s=0}^{L-m} a_{m+s}b^s \leq b^k = R$ . As a result

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

It follows that

$$\Pr\left(\liminf_{n \rightarrow \infty} \hat{\mu}_n^{\text{RQMC}} \geq (1 + b^{1-k})^{-1} \mu\right) = 1,$$

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

**5. An SLLN without Square Integrability.** The SLLN for Monte Carlo requires only 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 some  $p > 1$ .

The proof is based on the Riesz–Thorin interpolation theorem from [6, Chapter 4]. Let  $\mathcal{E}$  be the operator that takes an integrand  $f$  and returns the integration error

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

The integration error is a function of  $\mathbf{x}_1, \dots, \mathbf{x}_n \in [0, 1]^d$ . Together these belong to  $[0, 1]^{dn}$ . Let  $\Omega$  be the set  $[0, 1]^{dn}$  equipped with the distribution induced by the scrambled net randomization producing  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . 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

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

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

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

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

**THEOREM 5.1 (Riesz–Thorin).** *For  $1 \leq q_1 \leq q_2 < \infty$  and  $\theta \in [0, 1]$ , let  $p \geq 1$  satisfy*

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

*For probability spaces  $\Theta_1$  and  $\Theta_2$ , 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*

$$\|\mathcal{T}\|_{L^{q_1}(\Theta_1) \rightarrow L^{q_1}(\Theta_2)} \leq M_1 \quad \text{and} \quad \|\mathcal{T}\|_{L^{q_2}(\Theta_1) \rightarrow L^{q_2}(\Theta_2)} \leq M_2.$$

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

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

*Proof.* This is a special case of Theorem 2.2(b) in [6].  $\square$

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

**COROLLARY 5.2.** *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*

$$\|\mathcal{T}\|_{L^1(\Theta_1) \rightarrow L^1(\Theta_2)} \leq M_1 \quad \text{and} \quad \|\mathcal{T}\|_{L^2(\Theta_1) \rightarrow L^2(\Theta_2)} \leq M_2.$$

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

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

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  $\Theta_1$  will be  $[0, 1]^d$  under the uniform distribution, and the space  $\Theta_2$  will be  $[0, 1]^{nd}$  under the distribution induced by the RQMC points  $\mathbf{x}_1, \dots, \mathbf{x}_n$ .

**THEOREM 5.3.** *Let  $\mathbf{x}_1, \mathbf{x}_2, \dots$  be a  $(t, d)$ -sequence in base  $b$ , with gain coefficients no larger than  $\Gamma < \infty$  and randomized as in [44]. For  $p > 1$ , let  $f \in L^p[0, 1]^d$  with  $\int_{[0, 1]^d} f(\mathbf{x}) \, d\mathbf{x} = \mu$ . Then*

$$\Pr\left(\lim_{n \rightarrow \infty} \hat{\mu}_n^{\text{RQMC}} = \mu\right) = 1.$$

*Proof.* For  $p \geq 2$ , the conclusion follows from Theorem 4.3, and so we assume now that  $1 < p < 2$ . Choose any  $\epsilon > 0$  and suppose that  $n = rb^m$  for  $1 \leq r \leq R < \infty$  and  $m \geq 0$ . The error operator  $\mathcal{E}$  for this  $n$  satisfies  $\|\mathcal{E}\|_{L^1} \leq 2$  and  $\|\mathcal{E}\|_{L^2} \leq (r\Gamma/n)^{1/2}$ . Taking  $\mathcal{T} = \mathcal{E}$  in Corollary 5.2,

$$\sup_{\|f\|_p \leq 1} (\mathbb{E}(|\hat{\mu}_n^{\text{RQMC}} - \mu|^p))^{1/p} \leq 2^{(2-p)/p} \left(\frac{r\Gamma}{n}\right)^{(p-1)/p},$$

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

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

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

$$\Pr\left(\lim_{n \rightarrow \infty} \hat{\mu}_n^{\text{RQMC}} = \mu\right) = 1$$

when the limit is over  $n \in \{rb^m \mid 1 \leq r \leq R, m \geq 0\}$ . We have thus established a version of Theorem 4.2 for  $p > 1$ , and the extension to the unrestricted limit as  $n \rightarrow \infty$  uses the same argument as Theorem 4.3.  $\square$

The Riesz–Thorin theorem has been previously used to bound  $p$ th moments in similar problems. See, for instance, [25, 54, 33].

**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  $\mathbb{E}(|f(\mathbf{x})|^p) < \infty$  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 [38] require less space than the nested uniform scramble. They have the same mean squared discrepancy as the nested uniform scramble [28], and so they might also satisfy an SLLN. A digital shift [36, 48] does not produce the same variance as the nested uniform scramble, and it does not satisfy the critically important bound (3.2) 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 [4, 37].

A second major family of RQMC methods has been constructed from lattice rules [56]. Points  $\mathbf{a}_1, \dots, \mathbf{a}_n$  on a lattice in  $[0, 1]^d$  are randomized into  $\mathbf{x}_i = \mathbf{a}_i + \mathbf{u} \bmod 1$  for  $\mathbf{u} \sim \mathbb{U}[0, 1]^d$ . That is, they are shifted with wraparound in what is known as a Cranley–Patterson rotation [9]. Then the estimate of  $\mu$  is  $\hat{\mu}_n^{\text{RLAT}} = (1/n) \sum_{i=1}^n f(\mathbf{x}_i)$ . For an extensible version of shifted lattice rules, see [27]. The Cranley–Patterson rotation does not provide a  $\Gamma$  bound like (3.2) because there are functions  $f \in L^2[0, 1]^d$  with  $\text{var}(\hat{\mu}_n^{\text{RLAT}}) = \sigma^2(f)$  [36], and so a proof of an SLLN for this form of RQMC would require a different approach. The fact that  $\text{var}(\hat{\mu}_n^{\text{RLAT}}) = \sigma^2(f)$  is possible does not provide a counterexample to an SLLN because this equality might only hold for a finite number of  $n_\ell$  in the infinite sequence. Given a class of functions  $\mathcal{F}$  with  $\text{var}(\hat{\mu}_{n_\ell}^{\text{RLAT}}) \leq B\sigma^2(f)/n_\ell$  for all  $f \in \mathcal{F}$ , all  $\ell \geq 1$ , 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 [36], though they hold for specific  $n_\ell$  and not necessarily an infinite sequence of them.

**Acknowledgments.** We thank Max Balandat for posing the problem of finding an SLLN for randomized QMC. Thanks also to Ektan Bakshy, Wei-Liem Loh, and Fred Hickernell for discussions.

## REFERENCES

- [1] C. AISTLEITNER AND J. DICK, *Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality*, Acta Arithmetica, 167 (2015), pp. 143–171. (Cited on p. 364)
- [2] S. ANDRADÓTTIR, *Simulation optimization*, in Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice, J. Banks, ed., John Wiley & Sons, New York, 1998, pp. 307–333. (Cited on p. 361)
- [3] M. BALANDAT, B. KARRER, D. R. JIANG, S. DAULTON, B. LETHAM, A. G. WILSON, AND E. BAKSHY, *BoTorch: Bayesian Optimization in PyTorch*, preprint, <https://arxiv.org/abs/1910.06403>, 2019. (Cited on p. 361)
- [4] K. BASU AND R. MUKHERJEE, *Asymptotic normality of scrambled geometric net quadrature*, Ann. Statist., 45 (2017), pp. 1759–1788. (Cited on p. 370)
- [5] K. BASU AND A. B. OWEN, *Quasi-Monte Carlo for an integrand with a singularity along a diagonal in the square*, in Contemporary Computational Mathematics: A Celebration of the 80th Birthday of Ian Sloan, J. Dick, F. Y. Kuo, and H. Woźniakowski, eds., Springer, Cham, Switzerland, 2018, pp. 119–130. (Cited on p. 366)
- [6] C. BENNETT AND R. SHARPLEY, *Interpolation of Operators*, Pure Appl. Math. 129, Academic Press, Boston, MA, 1988. (Cited on pp. 362, 368, 369)
- [7] W. CHEN, A. SRIVASTAV, AND G. TRAVAGLINI, EDS., *A Panorama of Discrepancy Theory*, Springer, Cham, Switzerland, 2014. (Cited on p. 362)
- [8] J. A. CLARKSON AND C. R. ADAMS, *On definitions of bounded variation for functions of two variables*, Trans. Amer. Math. Soc., 35 (1933), pp. 824–854. (Cited on p. 363)
- [9] R. CRANLEY AND T. N. L. PATTERSON, *Randomization of number theoretic methods for multiple integration*, SIAM J. Numer. Anal., 13 (1976), pp. 904–914, <https://doi.org/10.1137/0713071>. (Cited on p. 370)
- [10] P. J. DAVIS AND P. RABINOWITZ, *Methods of Numerical Integration*, 2nd ed., Academic Press, San Diego, 1984. (Cited on p. 360)

- [11] L. DEVROYE, *Non-uniform Random Variate Generation*, Springer, Heidelberg, 1986. (Cited on p. 362)
- [12] J. DICK, F. Y. KUO, AND I. H. SLOAN, *High-dimensional integration: The quasi-Monte Carlo way*, *Acta Numer.*, 22 (2013), pp. 133–288. (Cited on pp. 360, 363)
- [13] J. DICK AND F. PILLICHSHAMMER, *Digital Sequences, Discrepancy and Quasi-Monte Carlo Integration*, Cambridge University Press, Cambridge, UK, 2010. (Cited on pp. 360, 363)
- [14] C. DOERR, M. GNEWUCH, AND M. WAHLSTRÖM, *Calculation of discrepancy measures and applications*, in *A Panorama of Discrepancy Theory*, W. Chen, A. Srivastav, and G. Travaglini, eds., Springer, Cham, Switzerland, 2014, pp. 621–678. (Cited on p. 364)
- [15] R. DURRETT, *Probability Theory and Examples*, Cambridge University Press, Cambridge, UK, 2019. (Cited on pp. 362, 367)
- [16] N. ETEMADI, *An elementary proof of the strong law of large numbers*, *Z. Wahrsch. Verw. Gebiete*, 55 (1981), pp. 119–122. (Cited on p. 367)
- [17] H. FAURE, *Discrépance de suites associées à un système de numération (en dimension  $s$ )*, *Acta Arithmetica*, 41 (1982), pp. 337–351. (Cited on p. 365)
- [18] P. GLASSERMAN, *Monte Carlo Methods in Financial Engineering*, Springer Science & Business Media, New York, 2003. (Cited on p. 361)
- [19] M. GRIEBEL, F. Y. KUO, AND I. H. SLOAN, *The smoothing effect of integration in  $\mathbb{R}^d$  and the ANOVA decomposition*, *Math. Comp.*, 82 (2013), pp. 383–400. (Cited on p. 366)
- [20] M. GRIEBEL, F. Y. KUO, AND I. H. SLOAN, *Note on “The smoothing effect of integration in  $\mathbb{R}^d$  and the ANOVA decomposition,”* *Math. Comp.*, 86 (2017), pp. 1847–1854. (Cited on p. 366)
- [21] A. GRIEWANK, F. Y. KUO, H. LEÖVEY, AND I. H. SLOAN, *High dimensional integration of kinks and jumps—Smoothing by preintegration*, *J. Comput. Appl. Math.*, 344 (2018), pp. 259–274. (Cited on p. 366)
- [22] G. H. HARDY, *On double Fourier series, and especially those which represent the double zeta-function with real and incommensurable parameters*, *Quart. J. Math.*, 37 (1905), pp. 53–79. (Cited on p. 363)
- [23] J. HARTINGER AND R. KAINHOFER, *Non-uniform low-discrepancy sequence generation and integration of singular integrands*, in *Monte Carlo and Quasi-Monte Carlo Methods 2004*, H. Niederreiter and D. Talay, eds., Springer, Berlin, 2006, pp. 163–179. (Cited on p. 366)
- [24] Z. HE AND X. WANG, *On the convergence rate of randomized quasi-Monte Carlo for discontinuous functions*, *SIAM J. Numer. Anal.*, 53 (2015), pp. 2488–2503, <https://doi.org/10.1137/15M1007963>. (Cited on p. 366)
- [25] S. HEINRICH, *Random approximation in numerical analysis*, in *Proceedings of the Conference “Functional Analysis” Essen*, Dekker, New York, 1994, pp. 123–171. (Cited on p. 369)
- [26] F. J. HICKERNELL, *Koksma-Hlawka inequality*, *Wiley StatsRef: Statistics Reference Online*, 2014, <https://onlinelibrary.wiley.com/doi/10.1002/0471667196.ess4085.pub2>. (Cited on p. 363)
- [27] F. J. HICKERNELL, H. S. HONG, P. L’ECUYER, AND C. LEMIEUX, *Extensible lattice sequences for quasi-Monte Carlo quadrature*, *SIAM J. Sci. Comput.*, 22 (2000), pp. 1117–1138, <https://doi.org/10.1137/S1064827599356638>. (Cited on p. 370)
- [28] F. J. HICKERNELL AND R. X. YUE, *The mean square discrepancy of scrambled  $(t, s)$ -sequences*, *SIAM J. Numer. Anal.*, 38 (2000), pp. 1089–1112, <https://doi.org/10.1137/S0036142999358019>. (Cited on p. 370)
- [29] S. JOE AND F. Y. KUO, *Constructing Sobol’ sequences with better two-dimensional projections*, *SIAM J. Sci. Comput.*, 30 (2008), pp. 2635–2654, <https://doi.org/10.1137/070709359>. (Cited on p. 365)
- [30] A. KELLER, *A quasi-Monte Carlo algorithm for the global illumination problem in a radiosity setting*, in *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, H. Niederreiter and P. J.-S. Shiue, eds., Springer-Verlag, New York, 1995, pp. 239–251. (Cited on p. 361)
- [31] M. KRAUSE, *Über Fouriersche Reihen mit zwei veränderlichen Größen*, *Leipziger Ber.*, 55 (1903), pp. 164–197. (Cited on p. 363)
- [32] L. KUIPERS AND H. NIEDERREITER, *Uniform Distribution of Sequences*, Wiley, New York, 1974. (Cited on p. 363)
- [33] R. J. KUNSCH, E. NOVAK, AND D. RUDOLF, *Solvable integration problems and optimal sample size selection*, *J. Complexity*, 53 (2019), pp. 40–67. (Cited on p. 369)
- [34] 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*, *Found. Comput. Math.*, 16 (2016), pp. 1631–1696. (Cited on p. 361)

- [35] B. LAVRIČ, *Continuity of monotone functions*, Arch. Math. (Brno), 29 (1993), pp. 1–4. (Cited on p. 364)
- [36] P. L'ECUYER AND C. LEMIEUX, *A survey of randomized quasi-Monte Carlo methods*, in Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications, M. Dror, P. L'Ecuyer, and F. Szidarovszki, eds., Kluwer Academic, Boston, MA, 2002, pp. 419–474. (Cited on pp. 361, 363, 364, 370)
- [37] W.-L. LOH, *On the asymptotic distribution of scrambled net quadrature*, Ann. Statist., 31 (2003), pp. 1282–1324. (Cited on p. 370)
- [38] J. MATOUŠEK, *On the  $L^2$ -discrepancy for anchored boxes*, J. Complexity, 14 (1998), pp. 527–556. (Cited on p. 370)
- [39] H. NIEDERREITER, *Pseudo-random numbers and optimal coefficients*, Adv. Math., 26 (1977), pp. 99–181. (Cited on p. 363)
- [40] H. NIEDERREITER, *Point sets and sequences with small discrepancy*, Monatsh. Math., 104 (1987), pp. 273–337. (Cited on p. 365)
- [41] H. NIEDERREITER, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, PA, 1992, <https://doi.org/10.1137/1.9781611970081>. (Cited on pp. 360, 363)
- [42] H. NIEDERREITER AND C. XING, *Low-discrepancy sequences and global function fields with many rational places*, Finite Fields Appl., 2 (1996), pp. 241–273. (Cited on p. 365)
- [43] H. NIEDERREITER AND C. XING, *Quasirandom points and global function fields*, in Finite Fields and Applications, S. Cohen and H. Niederreiter, eds., London Math. Soc. Lecture Note Ser. 233, Cambridge University Press, Cambridge, UK, 1996, pp. 269–296. (Cited on p. 365)
- [44] A. B. OWEN, *Randomly permuted  $(t, m, s)$ -nets and  $(t, s)$ -sequences*, in Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, H. Niederreiter and P. J.-S. Shiue, eds., Springer-Verlag, New York, 1995, pp. 299–317. (Cited on pp. 361, 364, 365, 366, 367, 369)
- [45] A. B. OWEN, *Monte Carlo variance of scrambled net quadrature*, SIAM J. Numer. Anal., 34 (1997), pp. 1884–1910, <https://doi.org/10.1137/S0036142994277468>. (Cited on p. 361)
- [46] A. B. OWEN, *Scrambled net variance for integrals of smooth functions*, Ann. Statist., 25 (1997), pp. 1541–1562. (Cited on pp. 361, 365)
- [47] A. B. OWEN, *Scrambling Sobol' and Niederreiter-Xing points*, J. Complexity, 14 (1998), pp. 466–489. (Cited on p. 365)
- [48] A. B. OWEN, *Variance with alternative scramblings of digital nets*, ACM Trans. Modeling Comput. Simul., 13 (2003), pp. 363–378. (Cited on pp. 365, 370)
- [49] A. B. OWEN, *Multidimensional variation for quasi-Monte Carlo*, in International Conference on Statistics in Honour of Professor Kai-Tai Fang's 65th Birthday, J. Fan and G. Li, eds., World Scientific, Singapore, 2005. (Cited on p. 364)
- [50] A. B. OWEN, *Halton sequences avoid the origin*, SIAM Rev., 48 (2006), pp. 487–503, <https://doi.org/10.1137/S0036144504441573>. (Cited on p. 366)
- [51] A. B. OWEN, *Randomized QMC and point singularities*, in Monte Carlo and Quasi-Monte Carlo Methods 2004, H. Niederreiter and D. Talay, eds., Springer, Berlin, 2006, pp. 403–418. (Cited on p. 366)
- [52] A. B. OWEN, *Local antithetic sampling with scrambled nets*, Ann. Statist., 36 (2008), pp. 2319–2343. (Cited on pp. 361, 365)
- [53] A. B. OWEN, *A constraint on extensible quadrature rules*, Numer. Math., 132 (2016), pp. 511–518. (Cited on p. 366)
- [54] D. RUDOLF AND N. SCHWEIZER, *Error bounds of MCMC for functions with unbounded stationary variance*, Statist. Probab. Lett., 99 (2015), pp. 6–12. (Cited on p. 369)
- [55] R. SCHÜRER AND W. C. SCHMID, *MinT: New features and new results*, in Monte Carlo and Quasi-Monte Carlo Methods 2008, P. L'Ecuyer and A. B. Owen, eds., Springer, Berlin, 2009, pp. 171–189. (Cited on p. 365)
- [56] I. H. SLOAN AND S. JOE, *Lattice Methods for Multiple Integration*, Oxford Science, Oxford, 1994. (Cited on p. 370)
- [57] I. M. SOBOLOV, *The distribution of points in a cube and the accurate evaluation of integrals*, Zh. Vychisl. Mat. i Mat. Phys., 7 (1967), pp. 784–802 (in Russian). (Cited on p. 365)
- [58] I. M. SOBOLOV, *Calculation of improper integrals using uniformly distributed sequences*, Soviet Math. Dokl. 14 (1973), pp. 734–738. (Cited on p. 366)
- [59] I. M. SOBOLOV, *On quasi-Monte Carlo integrations*, Math. Comput. Simul., 47 (1998), pp. 103–112. (Cited on p. 366)
- [60] I. M. SOBOLOV, D. ASOTSKY, A. KREININ, AND S. KUCHERENKO, *Construction and comparison of high-dimensional Sobol' generators*, Wilmott Mag., 2011 (2011), pp. 64–79. (Cited on p. 365)