Proceedings of the 2003 Winter Simulation Conference
S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds.

QUASI-MONTE CARLO METHODS FOR SIMULATION

Pierre L'Ecuyer

Département d'Informatique et de Recherche Opérationnelle
Université de Montréal, C.P. 6128, Succ. Centre-Ville
Montréal (Québec), H3C 3J7, CANADA

ABSTRACT

Quasi-Monte Carlo (QMC) methods are numerical techniques for estimating large-dimensional integrals, usually over the unit hypercube. They can be applied, at least in principle, to any simulation whose aim is to estimate a mathematical expectation. This covers a very wide range of applications.

In this paper, we review some of the key ideas of quasi-Monte Carlo methods from a broad perspective, with emphasis on some recent results. We visit lattice rules in different types of spaces and make the connections between these rules and digital nets, thus covering the two most widely used QMC methods.

1 INTRODUCTION

When running a stochastic simulation on a computer, the required (pseudo)randomness is usually produced by a random number generator (RNG), whose output is a sequence (or *stream*) of real numbers between 0 and 1. This sequence is supposed to imitate a typical realization of a sequence of independent and identically distributed (i.i.d.) random variables uniformly distributed over the interval $(0, 1)$. The simulation program can then be viewed as a complicated function f that transforms this stream of real numbers $\mathbf{u} = (u_0, u_1, u_2, \dots)$ into an output value $f(\mathbf{u})$. Frequently, the goal of the simulation is to estimate a mathematical expectation that can be written as

$$\mu = \int_{[0,1)^s} f(\mathbf{u})d\mathbf{u} \tag{1}$$

where s is an integer that represents the number of calls to the RNG required by the simulation. In the case where this number of calls is random and unbounded, we can simply view s as infinite, and assume that the number of uniforms that are actually *used* by the simulation is finite with probability one.

If n independent simulation runs are performed, with run i using the random stream $\mathbf{u}_i \in [0, 1)^s$, for $i = 0, \dots, n-1$, the *Monte Carlo* (MC) estimator of μ is

$$Q_n = \frac{1}{n}\sum_{i=0}^{n-1} f(\mathbf{u}_i). \tag{2}$$

This estimator is unbiased, has variance σ^2/n where

$$\sigma^2 = \int_{[0,1)^s} f^2(\mathbf{u})d\mathbf{u} - \mu^2 \tag{3}$$

is assumed finite throughout this paper, and obeys the central-limit theorem $\sqrt{n}(Q_n - \mu)/\sigma \Rightarrow N(0, 1)$. The error $Q_n - \mu$ thus converges at rate $O_p(\sigma/\sqrt{n})$.

The idea of *Quasi-Monte Carlo* (QMC) methods is to replace the random points $\mathbf{u}_i$ by a set of points $P_n = \{\mathbf{u}_0, \dots, \mathbf{u}_{n-1}\} \subset [0, 1)^s$ that cover the unit hypercube $[0, 1)^s$ *more uniformly* than typical random points. The two main classes of methods for constructing such point sets are *digital nets* and *integration lattices* (Niederreiter 1992b, Sloan and Joe 1994, L'Ecuyer and Lemieux 2002). We will explain how they work in Section 3.

Can these methods beat the $O_p(1/\sqrt{n})$ convergence rate? The short theoretical answer is *yes*. A standard way to bound the integration error and obtain its convergence rate is via the Koksma-Hlawka inequality and its generalizations (Niederreiter 1992b, Hickernell 1998a). The idea is to consider a Banach space $\mathcal{F}$ of functions with norm $\|\cdot\|$, where $\|f-\mu\|$ measures the *variability* of f, and a measure $D(P_n)$ of the *discrepancy* (or non-uniformity) of P_n, chosen in a way that the *worst-case* deterministic error bound $|Q_n-\mu| \le \|f - \mu\| D(P_n)$ holds for all $f \in \mathcal{F}$. Then, for functions f with bounded variability the error is guaranteed to converge at least as fast (asymptotically) as $D(P_n)$. It is known that there are point sets P_n (constructed via lattice rules and digital nets) for which $O(D(P_n)) = O(n^{-1}(\ln n)^{s-1})$ (Niederreiter 1992b). If we impose the additional condition that $P_m \subseteq P_n$ whenever $m < n$, so that $\lim_{n\to\infty} P_n$ represents an infinite sequence of points whose first n points are P_n for each n, then the best known rate becomes $O(D(P_n)) = O(n^{-1}(\ln n)^s)$. In both cases, this rate beats $O(n^{-1/2})$ asymptotically. But for practical values of n (say, $n \le 10^9$), $O(n^{-1}(\ln n)^{s-1})$ wins only if the dimension s does not exceed 7 or 8. QMC methods have been shown to beat standard MC for certain problems in up to 1000 dimensions or more. However, the $O(n^{-1}(\ln n)^{s-1})$ convergence rate implied by the Koksma-Hlawka inequality does not suffice to explain this success. A key additional explanation will be given in Section 2: roughly, QMC can still work nicely if f can be approximated by a sum of low-dimensional functions.

In classical QMC methods, P_n is a purely deterministic point set, so the estimator Q_n has zero variance and the error (or *bias*) $Q_n - \mu$ is hard to estimate. In *randomized* QMC methods, P_n is randomized in a way that it retains its high uniformity over $[0,1)^s$ when taken as a set, while each of its points has the uniform distribution over $[0,1)^s$ when taken individually. Then, Q_n becomes an unbiased estimator of μ, hopefully with smaller variance than the standard MC estimator. To estimate the variance and perhaps compute a confidence interval on μ, one can apply m independent randomizations to the same P_n, and compute $\bar{X}$ and S_x^2, the sample mean and sample variance of the m corresponding (independent) values of Q_n. Then, $\mathrm{E}[\bar{X}] = \mu$ and $\mathrm{E}[S_x^2] = \mathrm{Var}[Q_n] = m\mathrm{Var}[\bar{X}]$ (L'Ecuyer and Lemieux 2000).

One simple example of such a randomization is a *random shift modulo 1*, proposed by Cranley and Patterson (1976): generate a *single* point $\mathbf{u}$ uniformly distributed over $[0,1)^s$ and add it to each point of P_n, coordinatewise, modulo 1. Since all points of P_n are shifted by the same amount, the set retains most of its structure and uniformity. Another example is a *random digital shift in base b*: generate again a single $\mathbf{u} = (u_1, \ldots, u_s)$ uniformly over $[0,1)^s$, write the digital expansion in base b of each of its coordinates, say $u_j = \sum_{\ell=1}^{\infty} d_{j,\ell} b^{-\ell}$, then add $d_{j,\ell}$ modulo b to the ℓth digit of the digital expansion in base b of the jth coordinate of each point $\mathbf{u}_i \in P_n$. For $b = 2$, the digitwise addition modulo b becomes a bitwise exclusive-or, which is fast to perform on a computer. An interesting property of this randomization is that if the hypercube $[0,1)^s$ is partitioned into $b^{q_1+\cdots+q_s}$ rectangular boxes of the same size by partitioning the jth axis into b^{q_j} equal parts for each j, for some integers $q_j \ge 0$ (such a partition is called a $\mathbf{q}$*-equidissection in base b* of the unit hypercube, where $\mathbf{q} = (q_1, \ldots, q_s)$), then the number of boxes that contain m points, for each integer m, is unchanged by the randomization. In particular, if each box contains the same number of point of P_n before the randomization, then it also does after the randomization. In this case, we say that P_n is $\mathbf{q}$*-equidistributed in base b*. Several other randomization methods exist and most are adapted to special types of point sets; see, e.g., L'Ecuyer and Lemieux (2002) and Owen (2003).

For randomized QMC point sets, the convergence rate of the variance $E[(Q_n - \mu)^2]$ can easily beat that of standard MC, especially if the function f is smooth. For example, if $\mathcal{F}$ be the Sobolev class of functions on $[0,1)^s$ whose mixed partial derivatives $D^i f$ of order $|i| \le k$ all have Euclidean norm $\|D^i f\|_2 \le 1$, then $\inf_{P_n} \sup_{f \in \mathcal{F}} (\mathrm{E}[(Q_n - \mu)^2])^{1/2} = O(n^{-k/s-1/2})$ where the infimum is taken over all randomized point sets P_n (Bakhvalov 1962, Heinrich and Nowak 2002). When k/s is large, this is much better than $O(n^{-1/2})$. On the other hand, concrete constructions giving this convergence rate for any k and s are not available, and the hidden constant could be large.

The remainder of this paper is organized as follows. In Section 2, we recall the functional ANOVA decomposition of a function f and discuss the importance of looking at the lower-dimensional projections when studying the uniformity of a point set P_n. In section 3, we give the definitions and outline some basic properties of lattice rules and digital nets. Randomized versions of these point sets, and corresponding variance expressions and bounds, are also examined. A short conclusion completes the paper.

2 ANOVA DECOMPOSITION

The *functional ANOVA decomposition* (Hoeffding 1948, Owen 1998, Liu and Owen 2003) writes f as $f(\mathbf{u}) = \mu + \sum_{I \subseteq \{1,\ldots,s\},\, I \neq \phi} f_I(\mathbf{u})$ where each f_I depends only on $\{u_i,\ i \in I\}$, the f_I's integrate to zero and are orthogonal, and the variance decomposes as $\sigma^2 = \sum_{I \subseteq \{1,\ldots,s\}} \sigma_I^2$ where $\sigma_I^2 = \mathrm{Var}[f_I(\mathbf{U})]$ for $\mathbf{U}$ uniformly distributed over $[0,1)^s$. See the references for explicit definitions of these f_I and additional properties.

For each set of coordinates I, let $P_n(I)$ denote the projection of P_n over the subspace determined by I. If there is a set $\mathcal{J}$ of subsets of $\{1,\ldots,s\}$ of cardinality much smaller than 2^s and such that $\sum_{I \in \mathcal{J}} \sigma_I^2 \approx \sigma^2$, then it suffices to construct P_n so that the projections $P_n(I)$ are highly uniform for all $I \in \mathcal{J}$, in order to reduce the important variance terms σ_I^2. This is generally easier to achieve than having *all* projections $P_n(I)$ highly uniform. The set $\mathcal{J}$ of important projections depends of course on the function f.

In this context, a function f is said to have *effective dimension d in proportion ρ in the superposition sense* if $\sum_{|I| \le d} \sigma_I^2 \ge \rho\sigma^2$ (Owen 1998). If ρ is close to 1, this means that f is well approximated by a sum of d-dimensional (or less) functions. For example, a multivariate polynomial of degree d has effective dimension d in proportion 1 in the superposition sense ($d = 1$ for a linear function, $d = 2$ for a quadratic function, etc.). Real-life simulations often involve high-dimensional functions with low effective dimension in proportion ρ close to 1. Special techniques can also be used to change f in order to reduce the effective dimension, without changing μ (Spanier and Maize 1994, Morokoff 1998, Owen 1998, Fox 1999). Nevertheless, for large s and moderate d, the number of projections $P_n(I)$ for which $|I| \le d$, which equals $\sum_{k=1}^{d} \binom{s}{k}$, becomes so large that it may be too hard or impossible to construct point sets for which all these projections are very uniform.

Sometimes, $\sum_{I \in \mathcal{J}} \sigma_I^2$ is close to σ^2 if $\mathcal{J}$ contains all the sets I formed by indices that are not too far apart, and it suffices to have good uniformity for the corresponding projections. For example, if one wishes to estimate the average waiting time per customer in a queueing system, the result will typically depend strongly on the interaction between the interarrival and service times of customers that are close to each other in time, and very little on the interaction between customers that are far from each other in time. This leads to the following definition: f has effective dimension d in proportion ρ in the *successive-dimensions sense* if $\sum_{I \subseteq \{i,\dots,i+d-1\},\, 0 \le i \le s-d} \sigma_I^2 \ge \rho\sigma^2$ (L'Ecuyer and Lemieux 2000). The following third definition further reduces the number of projections considered: f has effective dimension d in proportion ρ in the *truncation sense* (Caflisch, Morokoff, and Owen 1997) if $\sum_{I \subseteq \{1,\dots,d\}} \sigma_I^2 \ge \rho\sigma^2$. Low effective dimension in the truncation sense can sometimes be achieved by setting the simulation experiment (or program) in a way that the first few random variables that are generated account for most of the variance in f (Caflisch, Morokoff, and Owen 1997, Fox 1999, L'Ecuyer and Lemieux 2000).

Point sets should thus be constructed by considering the uniformity of certain sets of projections. It is natural to ask that all projections contain as many distinct points as the original point set, i.e., points should not be superposed in projections. Adopting constructions for which several projections are identical can also make the analysis easier. A point set P_n in $[0,1)^s$ is called *fully projection-regular* (Sloan and Joe 1994, L'Ecuyer and Lemieux 2000) if for each non-empty $I \subseteq \{1,\dots,s\}$, $P_n(I)$ has n distinct points. It is called *dimension-stationary* (Lemieux and L'Ecuyer 2001) if whenever $1 \le i_1 < \dots < i_\eta < s$ and $1 \le j \le s - i_\eta$, $P_n(\{i_1,\dots,i_\eta\}) = P_n(\{i_1+j,\dots,i_\eta+j\})$. This means that $P_n(I)$ depends only on the *spacings* between the indices in I. Note that naïve rectangular grids in $s \ge 2$ are *not* projection-regular, because their projections have several points superposed on each other. In this sense, they are bad QMC point sets.

3 LATTICE RULES AND DIGITAL NETS

3.1 Ordinary Lattice Rules

We now summarize the main types of construction methods for QMC point sets, and some of their basic properties. An *integration lattice* is a vector space of the form

$$L_s = \left\{ \mathbf{v} = \sum_{j=1}^{s} h_j \mathbf{v}_j \text{ such that each } h_j \in \mathbb{Z} \right\},$$

where $\mathbf{v}_1,\dots,\mathbf{v}_s \in \mathbb{R}^s$ are linearly independent over $\mathbb{R}$ and $\mathbb{Z}^s \subseteq L_s$. The approximation of μ by Q_n with the node set $P_n = L_s \cap [0,1)^s$ is a called a *lattice rule* (Korobov 1959, Sloan and Joe 1994). The condition $\mathbb{Z}^s \subseteq L_s$ implies that L_s is periodic with period 1 along each of the s coordinates.

Let $\mathbf{V}$ be the matrix whose rows are the basis vectors $\mathbf{v}_1,\cdots,\mathbf{v}_s$ and $\mathbf{V}^{-1}$ its inverse. The columns $\mathbf{h}_1^T,\dots,\mathbf{h}_s^T$ of $\mathbf{V}^{-1}$ form a basis of the *dual lattice*, defined as $L_s^* = \{\mathbf{h} \in \mathbb{R}^s : \mathbf{h}\cdot\mathbf{v} \in \mathbb{Z} \text{ for all } \mathbf{v} \in L_s\}$, where $\cdot$ denotes the scalar product. One has $\mathbb{Z}^s \subseteq L_s$ iff (if and only if) $L_s^* \subseteq \mathbb{Z}^s$ iff all entries of $\mathbf{V}^{-1}$ are integer. When this holds, $n = \det(\mathbf{V}^{-1})$ and all entries of $\mathbf{V}$ are multiples of $1/n$.

The *rank* of the lattice is the smallest r such that one can find a basis of the form $\mathbf{v}_1,\dots,\mathbf{v}_r,\mathbf{e}_{r+1},\cdots,\mathbf{e}_s$, where $\mathbf{e}_j$ is the jth unit vector in s-dimensions. In particular, a lattice rule of *rank 1* has a basis of the form $\mathbf{v}_1 = (a_1,\dots,a_s)/n$ and $\mathbf{v}_j = \mathbf{e}_j$ for $j > 1$, where $a_j \in \mathbb{Z}_n$ for each j. It is a *Korobov* rule if $\mathbf{v}_1$ has the special form $\mathbf{v}_1 = (1,\ a,\ a^2 \bmod n,\ \dots,\ a^{s-1} \bmod n)/n$ for some $a \in \mathbb{Z}_n$. The point set P_n of a Korobov lattice rule can also be written as $P_n = \{(x_1,\dots,x_s)/n$ such that $x_1 \in \mathbb{Z}_n$ and $x_j = a x_{j-1} \bmod n$ for all $j > 1\}$. This is the set of all vectors of successive values produced by a linear congruential generator (LCG) with modulus n and multiplier a, from all possible initial states (including 0). In this case, the points are easy to enumerate by using the recurrence.

The *projection* $L_s(I)$ of L_s over the subspace determined by $I = \{i_1,\dots,i_\eta\}$ is also a lattice, with point set $P_n(I)$. A rule of rank 1 is fully projection-regular iff $\gcd(n, a_j) = 1$ for all j, and a Korobov rule is fully projection-regular and dimension-stationary iff $\gcd(n, a) = 1$ (L'Ecuyer and Lemieux 2000).

Figure 1 illustrates the point set P_n in $s = 2$ dimensions for a Korobov lattice rule with $n = 1021$ and $a = 90$. The vectors $\mathbf{v}_1 = (1/1021, 90/1021)$ and $\mathbf{v}_2 = (0,1)$ are a basis of the lattice. This rule is both fully projection-regular and dimension-stationary. The high regularity and uniformity of the points over the unit square is apparent. The projection of P_n on each of the two coordinates gives the set of equidistant points $P_n(\{1\}) = P_n(\{2\}) = \{0, 1/n, \dots, (n-1)/n\}$.

It is possible to construct sequences of lattices $L_s^1 \subset L_s^2 \subset L_s^3 \subset \dots$, so that each lattice contains the previous one (Cranley and Patterson 1976, Joe and Sloan 1992, Hickernell, Hong, L'Ecuyer, and Lemieux 2001). Such sequences permit one to increase the cardinality of P_n sequentially, without throwing away the points already considered. If the point set $L_s^\xi \cap [0,1)^s$ contains n_ξ points, then $n_{\xi-1}$ must divide n_ξ, for each ξ. For example, if the ξth rule is a Korobov rule with $n_\xi = 2^\xi$ points and multiplier a_ξ, then one must have $a_\xi = a_{\xi-1}$ or $a_\xi = a_{\xi-1} + n_{\xi-1}$ for each ξ. That is, when doubling the number of points, there are only two possibilities for the new lattice in this case.

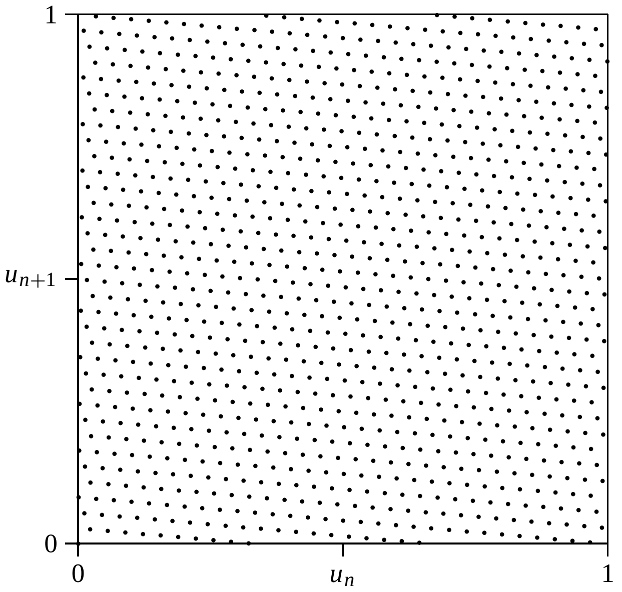

Figure 1: The 1021 Points of the Korobov Lattice Rule with $n = 1021$ and $a = 90$ in Two Dimensions

The lattice structure of P_n and each of its projections $P_n(I)$ implies that their points belong to a limited number of equidistant parallel hyperplanes. In two dimensions, these hyperplanes are lines, as illustrated in Figure 1. We want the distance between these hyperplanes to be small, in order to avoid large slices of space that contain no point. This distance for $P_n(I)$ happens to equal one over the Euclidean length of a shortest nonzero vector in the dual lattice $L_s^*(I)$. Computing it is often called the *spectral test* (Knuth 1998). It is commonly used to assess the quality of LCGs, and can be used in exactly the same way for integration lattices. For example, for each I in some arbitrarily selected class $\mathcal{J}$ of subsets of $\{0, \ldots, s-1\}$, one could compute the length $\ell(I)$ of the shortest nonzero vector in $L_s^*(I)$, divide it by an upper bound on the best possible value that can be achieved for an arbitrary lattice with *n* points per unit of volume in $|I|$ dimensions, in order to obtain a normalized value between 0 and 1 for any I, say, $S(I)$, and then take the worst case $\min_{I \in \mathcal{J}} S(I)$ as a figure of merit for the lattice. Ideally, this figure of merit should be as close to 1 as possible. For the lattice of Figure 1, one has $S(\{1,2\}) = 0.958$ and $S(\{1\}) = S(\{2\}) = 1$. L'Ecuyer and Lemieux (2000) provide further details and explicit Korobov rules selected via this type of criterion in more than two dimensions. Another possibility could be to take a weighted average of the $S(I)$. Several other measures of uniformity have been proposed and used for selecting integration lattices; see, e.g., (Sloan and Joe 1994, Hellekalek 1998, Hickernell 1998b, Lemieux and L'Ecuyer 2001).

3.2 Fourier Expansion of f and Variance for Randomly-Shifted Lattice Rules

A randomly-shifted lattice retains its lattice structure. For this reason, applying a random shift modulo 1 to an integration lattice provides a randomized point set with a nice structure that facilitates mathematical analysis of the error and variance. More precisely, let us write the *Fourier expansion* of f as

$$f(\mathbf{u}) = \sum_{\mathbf{h} \in \mathbb{Z}^s} \hat{f}(\mathbf{h}) \exp(2\pi\sqrt{-1}\,\mathbf{h} \cdot \mathbf{u}), \qquad (4)$$

with *Fourier coefficients*

$$\hat{f}(\mathbf{h}) = \int_{[0,1)^s} f(\mathbf{u}) \exp(-2\pi\sqrt{-1}\,\mathbf{h} \cdot \mathbf{u}) d\mathbf{u}.$$

Then, for the Monte Carlo method,

$$n\text{Var}[Q_n] = \sigma^2 = \sum_{\mathbf{0} \neq \mathbf{h} \in \mathbb{Z}^s} |\hat{f}(\mathbf{h})|^2, \qquad (5)$$

whereas for a randomly-shifted lattice rule,

$$\text{Var}[Q_n] = \sum_{\mathbf{0} \neq \mathbf{h} \in L_s^*} |\hat{f}(\mathbf{h})|^2 \qquad (6)$$

(Tuffin 1998, L'Ecuyer and Lemieux 2000). Note that the terms in (4) that corresponds to small vectors **h** represent the main trends (low-frequency components) of the function f and are usually more important than the high-frequency ones (large **h**). Note that the decomposition (5) is finer than the one given in Section 2, in the sense that each σ_I^2 corresponds to the sum of terms in (5) for the **h** whose nonzero coordinates are those with indices in I. Expression (6) suggests measures of discrepancy of the form $\sum_{\mathbf{0} \neq \mathbf{h} \in L_s^*} w(\mathbf{h})$ or $\sup_{\mathbf{0} \neq \mathbf{h} \in L_s^*} w(\mathbf{h})$ for the integration lattice L_s, where the weights $w(\mathbf{h})$ could (ideally) be chosen to decrease with the norm of **h** proportionally with the anticipated values of $|\hat{f}(\mathbf{h})|^2$. In practice, these weights are chosen somewhat arbitrarily, because the Fourier coefficients are hard to anticipate, but perhaps more work is needed in that direction. If we take $w(\mathbf{h})$ equal to $1/\|\mathbf{h}\|_2$ multiplied by an appropriate normalization constant that depends on the set I of nonzero coordinates of **h**, then $\sum_{\mathbf{0} \neq \mathbf{h} \in L_s^*} w(\mathbf{h})$ becomes equivalent to (is the inverse of) the spectral test figure of merit $\min_{I \in \mathcal{J}} S(I)$ discussed earlier. This type of figure of merit is thus justified both by a geometric argument (distance between hyperplanes) and by a variance expression in terms of Fourier coefficients for randomly-shifted lattice rules.

3.3 Lattice Rules in Formal Series

In the lattice L_s defined above, the vector coordinates are in $\mathbb{R}$ and the linear combinations of the basis vectors are over $\mathbb{Z}$. But integration lattices can also be defined in different spaces. For example, we can replace $\mathbb{R}$ and $\mathbb{Z}$ by the ring $\mathbb{L}_b$ of formal Laurent series with coefficients in $\mathbb{Z}_b$ and by the ring $\mathbb{Z}_b[z]$ of polynomials with coefficients in $\mathbb{Z}_b$, respectively, where b is an arbitrary integer larger than 1 and $\mathbb{Z}_b$ is the residue ring of integers modulo b (Niederreiter 1992b, L'Ecuyer 2003, Lemieux and L'Ecuyer 2003). The lattice

$$\mathcal{L}_s = \left\{ \mathbf{v}(z) = \sum_{j=1}^{s} q_j(z)\mathbf{v}_j(z) \; : \; \text{each } q_j(z) \in \mathbb{Z}_b[z] \right\}, \tag{7}$$

thus obtained, where $\mathbf{v}_1(z), \ldots, \mathbf{v}_s(z)$ are in $(\mathbb{L}_b)^s$ is called a *polynomial integration lattice* under the additional condition that $(\mathbb{Z}_b[z])^s \subseteq \mathcal{L}_s$.

An output mapping $\varphi : \mathbb{L}_b \to \mathbb{R}$ can be defined by

$$\varphi\left(\sum_{\ell=\omega}^{\infty} x_\ell z^{-\ell}\right) = \sum_{\ell=\omega}^{\infty} x_\ell b^{-\ell}.$$

The *polynomial lattice rule* (*PLR*) uses the node set $P_n = \varphi(\mathcal{L}_s) \cap [0,1)^s = \varphi(\mathcal{L}_s \cap \mathbb{L}_{b,0})$, where $\mathbb{L}_{b,0} = \mathbb{L}_b \bmod \mathbb{Z}_b[z]$. Most of the properties of ordinary lattice rules have counterparts in the context of PLRs.

The basis vectors form a matrix $\mathbf{V}$ with rows $\mathbf{v}_1(z), \ldots, \mathbf{v}_s(z)$, whose inverse $\mathbf{V}^{-1}$ has columns $\mathbf{h}_1(z)^T, \ldots, \mathbf{h}_s(z)^T$ that form a basis of the *dual lattice*

$$\mathcal{L}_s^* = \{\mathbf{h}(z) \in (\mathbb{L}_b)^s : \mathbf{h}(z) \cdot \mathbf{v}(z) \in \mathbb{Z}_b[z] \text{ for all } \mathbf{v}(z) \in \mathcal{L}_s\},$$

where $\mathbf{h}(z) \cdot \mathbf{v}(z) = \sum_{j=1}^{s} h_j(z) v_j(z)$. The *determinants* $\det(\mathcal{L}_s) = \det(\mathbf{V})$ and $\det(\mathcal{L}_s^*) = \det(\mathbf{V}^{-1}) = 1/\det(\mathcal{L}_s)$ do not depend on the choice of basis. The condition $(\mathbb{Z}_b[z])^s \subseteq \mathcal{L}_s$ is crucial to guarantee that all the inverses defined above do exist even when b is not a prime (i.e., when $\mathbb{Z}_b$ is not a field). This condition holds iff $\mathbf{V}^{-1}$ exist and all its entries are polynomials. Then, $\det(\mathcal{L}_s^*)$ is a polynomial $P(z)$, the number of points is $n = b^k$ where k is the degree of $P(z)$, and each coordinate of any vector $\mathbf{v}(z) \in \mathcal{L}_s$ has the form $v(z) = p(z)/P(z)$ for some polynomial $p(z)$.

The *rank* of the rule is the smallest r such that there is a basis of the form $\mathbf{v}_1(z), \ldots, \mathbf{v}_r(z), \mathbf{e}_{r+1}, \cdots, \mathbf{e}_s$. A *Korobov PLR* is a rule of rank 1 with $P(z)\mathbf{v}_1(z) = (1, \; a(z), \; a^2(z) \bmod P(z), \; \ldots, \; a^{s-1}(z) \bmod P(z))$, where $P(z)$ is a polynomial of degree k over $\mathbb{Z}_b$, having a multiplicative inverse $1/P(z)$ in $\mathbb{L}_b$, and $a(z) \in \mathbb{Z}_b[z]/(P)$ (the polynomials with degree less than k). Such a rule is equivalent to using the point set $P_n = \{\varphi((p_0(z), \ldots, p_{s-1}(z))/P(z))$ such that $p_0(z) \in \mathbb{Z}_b[z]/(P)\}$ where $p_j(z) = a(z)p_{j-1}(z) \bmod P(z)$ for all j, i.e., the image by φ of all vectors of successive values produced by a polynomial LCG with modulus $P(z)$ and multiplier $a(z)$, from all initial states $p_0(z)$. For the special case where $b = 2$ (the most interesting case), this corresponds to using all cycles of a *linear feedback shift register* (LFSR) or *Tausworthe* generator with characteristic polynomial $P(z)$ (Tezuka 1995, Lemieux and L'Ecuyer 2003).

The projection of $\mathcal{L}_s$ over the subspace determined by $I = \{i_1, \ldots, i_\eta\} \subset \{1, \ldots, s\}$ is a polynomial integration lattice $\mathcal{L}_s(I)$ with dual lattice $\mathcal{L}_s^*(I)$ and point set $P_n(I)$. For prime b, one can show (Lemieux and L'Ecuyer 2003) that a rule of rank 1 with $\mathbf{v}_1(z) = (g_1(z), g_2(z), \ldots, g_s(z))/P(z)$ is fully projection-regular iff $\gcd(g_j(z), P(z)) = 1$ for all j, and that a Korobov rule, with $g_j(z) = a^{j-1}(z) \bmod P(z)$, is fully projection-regular and dimension-stationary iff $\gcd(a(z), P(z)) = 1$.

PLRs of rank 1 were introduced by Niederreiter (1992a), and by Tezuka (1990) for $b = 2$ and the special case of an irreducible $P(z)$; see also Niederreiter (1992b), Section 4.4. They were generalized to PLRs of arbitrary rank over a finite field by Lemieux and L'Ecuyer (2003) and over the ring $\mathbb{Z}_b$ by L'Ecuyer (2003).

3.4 Equidistribution and Measures of Uniformity for Polynomial Lattice Rules

Recall that P_n is called $\mathbf{q}$-equidistributed in base b, for an integer vector $\mathbf{q} = (q_1, \ldots, q_s) \geq \mathbf{0}$, if each box of the $\mathbf{q}$-equidissection of $[0,1)^s$ contains the same number of points from P_n, namely b^t points where $t = k - q_1 - \cdots - q_s$ if $n = b^k$. If this holds for $q_1 = \cdots = q_s = \ell$ for some $\ell \geq 1$, we have *s-distribution with* ℓ *digits of accuracy* (Tezuka 1995), and the largest such ℓ is called the *s-dimensional resolution* of P_n. This value cannot exceed $\lfloor k/s \rfloor$.

These definitions also apply to projections: For $I = \{i_1, \ldots, i_\eta\} \subset \{1, \ldots, s\}$, the set $P_n(I)$ is $(q_{i_1}, \ldots, q_{i_\eta})$-*equidistributed* if each box of the $(q_{i_1}, \ldots, q_{i_\eta})$-equidissection of $[0,1)^\eta$ contains $2^{t(I)}$ points of $P_n(I)$, where $k - t(I) = q_{i_1} + \ldots + q_{i_\eta}$. The *resolution gap* of $P_n(I)$ is $\delta_I = \lfloor k/\eta \rfloor - \ell_I$, where ℓ_I is the η-dimensional resolution of $P_n(I)$.

For $n = b^k$, P_n is called a (t,k,s)-*net in base* b if it is $(q_1, \ldots, q_s)$-equidistributed for all non-negative integers $q_1, \ldots, q_s$ summing to $k - t$ (Niederreiter 1992b). We call the smallest such t the t-*value* of the net.

For ordinary lattice rules, measures of uniformity such as the distance between hyperplanes are obtained by computing a shortest vector in the dual lattice. Interestingly, this is also true for PLRs: The equidistribution and (t,k,s)-net properties can be verified by computing the length of a shortest nonzero vector in the dual lattice $\mathcal{L}_s^*$, with an appropriate choice of norm.

For each integer vector $\mathbf{q} = (q_1, \dots, q_s)$, define a length (or *distance function*) $\| \cdot \|_{-\mathbf{q}}$ on $(\mathbb{Z}_b[z])^s$ by

$$\log_b \|\mathbf{h}(z)\|_{-\mathbf{q}} = \max_{1\le j\le s} (\deg(h_j) - q_j), \tag{8}$$

for $\mathbf{h}(z) = (h_1(z), \dots, h_s(z)) \in \mathbb{Z}_b[z]$, where $\deg(h_j)$ is the degree of the polynomial $h_j(z)$ and $\deg(0) = -\infty$ by convention. Let $\sigma_1^* = \min_{\mathbf{0}\ne\mathbf{h}(z)\in\mathcal{L}_s^*} \|\mathbf{h}(z)\|_{-\mathbf{q}}$, the length of the shortest nonzero vector in the dual lattice. Under the assumption that b is prime, it is proved in L'Ecuyer (2003) that P_n is $\mathbf{q}$-equidistributed iff $\sigma_1^* \ge 1$. In particular, the s-dimensional resolution of P_n is equal to $\log_b \min_{\mathbf{0}\ne\mathbf{h}(z)\in\mathcal{L}_s^*} \|\mathbf{h}(z)\|_{\mathbf{0}}$. This shortest vector with respect to the distance function $\|\mathbf{h}(z)\|_{\mathbf{0}}$ is relatively easy to compute (Tezuka 1995, Couture and L'Ecuyer 2000).

The t-value of P_n can also be obtained by computing the length of a shortest nonzero vector in the dual lattice, with a different definition of length. For $\mathbf{h}(z) \in \mathbb{Z}_b[z]$, define $\|\mathbf{h}(z)\|_\pi$ by

$$\log_b \|\mathbf{h}(z)\|_\pi = \sum_{j=1}^{s} \deg(h_j)$$

and let $\tau_1^* = \min_{\mathbf{0}\ne\mathbf{h}(z)\in\mathcal{L}_s^*} \|\mathbf{h}(z)\|_\pi$. Then the t-value of P_n is equal to $k - s + 1 - \log_b \tau_1^*$ (Niederreiter and Pirsic 2001, L'Ecuyer 2003).

Following the discussion in Section 2 and similar to what we suggested for ordinary lattice rules, one can consider for PLRs uniformity criteria of the form $\Delta_{\mathcal{J}} = \max_{I\in\mathcal{J}} \delta_I$ (worst-case resolution gap) or $\max_{I\in\mathcal{J}} t^*_{|I|}/t_I$ or $\max_{I\in\mathcal{J}} (t_I - t^*_{|I|})$, where $\mathcal{J}$ is a selected class of sets I, t_I is the t-value for $P_n(I)$, $t^*_{|I|}$ a lower bound on the best possible t-value in $|I|$ dimensions, and with the convention that $0/0 = 1$. The choice of $\mathcal{J}$ is again arbitrary and a matter of compromise. If $\mathcal{J}$ contains too many sets, not only the selection criterion will be more costly to compute, but the best value that it can achieve will be larger, and therefore the criterion will become less demanding for the equidistribution of the more important projections. Other types of uniformity criteria are discussed, e.g., in L'Ecuyer and Lemieux (2002).

3.5 Lattice Rules in Formal Series Over $\mathbb{Z}_b$

We now consider a lattice of the form

$$\mathcal{C}_s = \left\{ \mathbf{v}(z) = \sum_{i=1}^{k} y_i \mathbf{c}_i(z) \text{ such that } y_i \in \mathbb{Z}_b \text{ for each } i \right\}, \tag{9}$$

where $\mathbf{c}_1(z), \dots, \mathbf{c}_k(z)$ are k vectors of $\mathbb{L}^s_{b,0}$ independent over $\mathbb{Z}_b$, and let $P_n = \varphi(\mathcal{C}_s) \subset [0,1)^s$, where φ is defined as before. Here, the lattice is defined over $\mathbb{Z}_b$ instead of over $\mathbb{Z}_b[z]$ as in (7). The set P_n contains exactly $n = b^k$ distinct points, because $\mathcal{C}_s \subset \mathbb{L}^s_{b,0}$.

To define the dual lattice, we first define a (noncommutative) product $\odot$ in $\mathbb{L}_b$ by

$$\left(\sum_{\ell=-\infty}^{w_2} x_\ell z^\ell\right) \odot \left(\sum_{\ell=w_1}^{\infty} y_\ell z^{-\ell}\right) = \sum_{\ell=w_1-1}^{w_2} x_\ell y_{\ell+1}$$

where the last sum is in $\mathbb{Z}_b$. For $\mathbf{x}(z) = (x_1(z), \dots, x_s(z))$ and $\mathbf{y}(z) = (y_1(z), \dots, y_s(s))$ in $\mathbb{L}_b^s$, we define $\mathbf{x}(z) \odot \mathbf{y}(z) = \sum_{j=1}^{s} x_j(z) \odot y_j(z)$. The *dual lattice* is then defined as

$$\begin{aligned} \mathcal{C}_s^\perp \;=\; & \{\mathbf{h}(z) \in (\mathbb{Z}_b[z])^s \text{ such that} \\ & \quad \mathbf{h}(z) \odot \mathbf{v}(z) = 0 \text{ for all } \mathbf{v}(z) \in \mathcal{C}_s\}. \end{aligned}$$

This is a lattice over $\mathbb{Z}_b$, i.e., can be written as $\mathcal{C}_s^\perp = \{\mathbf{h}(z) = \sum_{j=1}^{\nu} x_i \mathbf{h}_j(z)$ such that $x_i \in \mathbb{Z}_b$ for each $i\}$ for some basis $\mathbf{h}_1(z), \dots, \mathbf{h}_\nu(z)$, where ν is the dimension of $\mathcal{C}_s^\perp$ over $\mathbb{Z}_b$.

Equidistribution properties can be determined by computing the lengths of shortest vectors in this dual lattice, just as for PLRs (L'Ecuyer 2003). Specifically, at least for prime b, P_n is $\mathbf{q}$-equidistributed iff $\min_{\mathbf{0}\ne\mathbf{h}\in\mathcal{C}_s^\perp} \|\mathbf{h}\|_{-\mathbf{q}} \ge 1$, the resolution of P_n is equal to $\log_b \min_{\mathbf{0}\ne\mathbf{h}\in\mathcal{C}_s^\perp} \|\mathbf{h}\|_{\mathbf{0}}$, and its t-value is equal to $k - s + 1 - \log_b \min_{\mathbf{0}\ne\mathbf{h}\in\mathcal{C}_s^\perp} \|\mathbf{h}\|_\pi$.

3.6 Digital Nets and Sequences

The lattice rules over $\mathbb{Z}_b$ defined in the previous section turn out to be equivalent to another very well-known class of QMC methods: the digital nets, introduced by Sobol' (1967) in base 2, later generalized by Faure (1982), Niederreiter (1987), and Tezuka (1995), and defined as follows (Niederreiter 1992b). Let $\mathbf{C}^{(1)}, \dots, \mathbf{C}^{(s)}$ be matrices of dimension $\infty \times k$ with elements in $\mathbb{Z}_b$, for some integer $k \ge 1$. They are the *generating matrices* of the net. For $i = 0, \dots, b^k - 1$, write $i = \sum_{\ell=0}^{k-1} a_{i,\ell} b^\ell$ and define $\mathbf{u}_i = (u_{i,1}, \dots, u_{i,s})$ where $u_{i,j} = \sum_{\ell=1}^{\infty} u_{i,j,\ell} b^{-\ell}$ and $(u_{i,j,1}, u_{i,j,2}, \dots)^T = \mathbf{C}^{(j)}(a_{i,0}, a_{i,1}, \dots, a_{i,k-1})^T$. The point set $P_n = \{\mathbf{u}_0, \dots, \mathbf{u}_{n-1}\}$ thus obtained, with $n = b^k$, is a *digital net* over $\mathbb{Z}_b$. These n points are distinct in their first ℓ digits iff the $\ell s \times k$ matrix formed by taking the first ℓ rows of each $\mathbf{C}^{(j)}$ has rank k. The matrices $\mathbf{C}^{(j)}$ can also be defined with an infinite number of columns: we then have an infinite sequence of points, called a *digital sequence*, whose first b^k points form a digital net for each integer k. In concrete implementations, it is worth considering only a finite number of rows of each $\mathbf{C}^{(j)}$, because of the finite precision of computers.

Digital nets and sequences can in fact be defined over an arbitrary commutative ring R of cardinality b, with an identity element. It suffices to define bijections between R and $\mathbb{Z}_b$ to map the digits of the b-ary expansion of i to elements of R and to recover the b-ary digits of $u_{i,j}$ from elements of R (Niederreiter 1992b, L'Ecuyer and Lemieux 2002). A similar generalization also applies to lattices rules in formal series by incorporating the bijections from R to $\mathbb{Z}_b$ into φ. However, the result is no longer a lattice over $\mathbb{Z}_b$ or $\mathbb{Z}_b[z]$. Here, we assume that $R = \mathbb{Z}_b$ and that all bijections are the identity, which is usually the case in practice.

For the lattice $\mathcal{C}_s$ defined in (9), if we write the basis vectors $\mathbf{c}_i(z) = (c_{i,1}(z), \dots, c_{i,s}(z))$ where $c_{i,j}(z) = \sum_{\ell=1}^{\infty} c_{\ell,i}^{(j)} z^{-\ell}$, and let $\mathbf{C}^{(j)}$ be the $\infty \times k$ matrix with elements $c_{\ell,i}^{(j)}$ then it turns out that this methods yields exactly the same point set as the digital net in base b with generating matrices $\mathbf{C}^{(1)}, \dots, \mathbf{C}^{(s)}$ (L'Ecuyer and Lemieux 2002, L'Ecuyer 2003). In other words, a lattice rule in formal series over $\mathbb{Z}_b$ is just an alternative definition of a digital net over $\mathbb{Z}_b$, with identity bijections. These digital nets are thus lattice rules in an appropriate space.

Several special cases of digital sequences (from which digital nets can be extracted by taking the first $n = b^k$ points for any k) have been proposed over the years. The matrices $\mathbf{C}^{(j)}$ are normally chosen on the basis of some uniformity criterion, which is often the t-value.

In the original construction of Sobol' (1967) each matrix $\mathbf{C}^{(j)}$ is filled up using a recurrence with primitive characteristic polynomial f_j over the finite field $\mathbb{F}_2$, where the f_j are all distinct and have small degree. The initial states of these recurrences are called the *direction numbers* and their choice may have a significant impact on the quality of the point set. Specific values are suggested by Sobol' and Levitan (1976) and used in the implementation of Bratley and Fox (1988). These values have been chosen so that P_n has s-distribution with one bit of accuracy when $n = 2^s$ and two bits of accuracy when $n = 4^s$. Equidistribution for other equidissections was not examined.

In the construction of Faure (1982) and its generalizations, the basis b is the first prime larger or equal to the dimension s and $\mathbf{C}^{(j)} = \mathbf{A}_j \mathbf{P}^{j-1}$ where P is the transposed Pascal matrix, with element (i, j) equal to $\binom{j-1}{i-1}$, and $\mathbf{A}_j$ is an arbitrary non-singular lower-triangular matrix. The resulting point set has the remarkable property of being a $(0, k, s)$-net (i.e., has the best possible t-value) when $n = b^k$ for any k. Noticing that Faure's construction is not practical for large s because it would require too many points, Niederreiter (1988) has proposed a construction where b is a prime power, but can be smaller than s, and where the t-value is reasonably small when $n = b^k$. More recently, Niederreiter and Xing (1997) proposed a new class of digital net sequences with optimal asymptotic t-value as a function of s and n, for a fixed b. These sequences improve on the t-value of Sobol's sequence for $b = 2$.

The *Salzburg tables* (Pirsic and Schmid 2001) list the best parameters found for the special case where the digital net is a Korobov PLR, in an attempt to optimize its t-value. Other sets of parameters for PLRs, chosen via a criterion of the form $\Delta_{\mathcal{J}}$ defined earlier, can be found in Lemieux and L'Ecuyer (2001) and L'Ecuyer and Panneton (2002).

3.7 Variance Expressions and Bounds for Randomized Nets and Lattice Rules in Formal Series

Randomly shifting a set in $\mathbb{L}_b$ corresponds to adding a random formal series to all series in the set, by adding the corresponding coefficients in $\mathbb{Z}_b$. In the point set P_n, this transposes to a random digital shift in base b. As mentioned in the introduction, this type of shift preserves the equidistribution of every $\mathbf{q}$-equidissection in base b.

A variance expression similar to (6) is available for (randomly) digitally-shifted lattice rules in formal series (Lemieux and L'Ecuyer 2003, L'Ecuyer and Lemieux 2002). The Fourier expansion and coefficients are replaced by Walsh expansion and coefficients. Just as for ordinary lattice rules, these expressions suggest that the integration lattices should be selected so that their dual lattice does not contain short vectors.

Randomly shifting a point set provides an unbiased estimator of μ with a minimal amount of randomization and "perturbation" of the point set. However, more randomization can in some cases reduce the variance. Owen (1995) has proposed a randomization method called *nested uniform scrambling*, for digital nets, which randomly permutes the values $\{0, \dots, b-1\}$ used for the digits $u_{i,j,\ell}$, independently for each coordinate j, as follows. One uses a first permutation for $\ell = 1$. For $\ell > 1$, one uses a different permutation for each possible value of the preceding $\ell - 1$ digits $u_{i,j,1} \cdots u_{i,j,\ell-1}$. To scramble the first ℓ digits thus requires $(1 + b + \cdots + b^{\ell-1})s$ permutations, and all these permutations are independent. Owen (1997) has shown that for smooth enough functions (whose mixed partial derivatives satisfy a Lipschitz condition) the variance is in $O(n^{-3}(\log n)^s)$. With a random digital shift, the bound is $O(n^{-2}(\log n)^s)$ instead (L'Ecuyer and Lemieux 2002). However, nested uniform scrambling is much more expensive to apply than the digital shift. Several other less expensive scramblings have been proposed whose amount of randomization lie somewhere in between these two; see Owen (2003) for an overview and a discussion.

4 CONCLUSION

We have reviewed the most common QMC methods and their randomizations, in the framework of lattice rules in different spaces. These methods can be used to improve the efficiency of simulations. Numerical illustrations of their application and effectiveness can be found in several of the references given below. On-going work on these methods includes, among other things, making computer searches for good parameters in terms of various selection criteria, developing extensive general-purpose software tools for QMC, studying the effectiveness of QMC methods and comparing them for specific classes of applications, developing QMC rules that may adapt to the integrand, and studying how the methods can be made more effective for high-dimensional problems.

ACKNOWLEDGMENTS

This work has been supported by NSERC-Canada grant No. ODGP0110050, NATEQ-Québec grant No. 02ER3218, and a Killam Research Fellowship to the author.

REFERENCES

Bakhvalov, N. S. 1962. On the rate of convergence of indeterministic integration processes within the functional classes $w_p^{(l)}$. *Theory of Probability and its Applications* 7:227.

Bratley, P., and B. L. Fox. 1988. Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software* 14 (1): 88–100.

Caflisch, R. E., W. Morokoff, and A. Owen. 1997. Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension. *The Journal of Computational Finance* 1 (1): 27–46.

Couture, R., and P. L'Ecuyer. 2000. Lattice computations for random numbers. *Mathematics of Computation* 69 (230): 757–765.

Cranley, R., and T. N. L. Patterson. 1976. Randomization of number theoretic methods for multiple integration. *SIAM Journal on Numerical Analysis* 13 (6): 904–914.

Faure, H. 1982. Discrépance des suites associées à un système de numération. *Acta Arithmetica* 61:337–351.

Fox, B. L. 1999. *Strategies for quasi-Monte Carlo*. Boston, MA: Kluwer Academic.

Heinrich, S., and E. Nowak. 2002. Optimal summation and integration by deterministic, randomized, and quantum algorithms. In *Monte Carlo and Quasi-Monte Carlo Methods 2000*, ed. K.-T. Fang, F. J. Hickernell, and H. Niederreiter, 50–62. Berlin: Springer-Verlag.

Hellekalek, P. 1998. On the assessment of random and quasi-random point sets. In *Random and Quasi-Random Point Sets*, ed. P. Hellekalek and G. Larcher, Volume 138 of *Lecture Notes in Statistics*, 49–108. New York: Springer.

Hickernell, F. J. 1998a. A generalized discrepancy and quadrature error bound. *Mathematics of Computation* 67:299–322.

Hickernell, F. J. 1998b. Lattice rules: How well do they measure up? In *Random and Quasi-Random Point Sets*, ed. P. Hellekalek and G. Larcher, Volume 138 of *Lecture Notes in Statistics*, 109–166. New York: Springer.

Hickernell, F. J., H. S. Hong, P. L'Ecuyer, and C. Lemieux. 2001. Extensible lattice sequences for quasi-Monte Carlo quadrature. *SIAM Journal on Scientific Computing* 22 (3): 1117–1138.

Hoeffding, W. 1948. A class of statistics with asymptotically normal distributions. *Annals of Mathematical Statistics* 19:293–325.

Joe, S., and I. H. Sloan. 1992. Embedded lattice rules for multidimensional integration. *SIAM Journal on Numerical Analysis* 29:1119–1135.

Knuth, D. E. 1998. *The art of computer programming, volume 2: Seminumerical algorithms*. Third ed. Reading, Mass.: Addison-Wesley.

Korobov, N. M. 1959. The approximate computation of multiple integrals. *Dokl. Akad. Nauk SSSR* 124:1207–1210. in Russian.

L'Ecuyer, P. 2003. Polynomial lattice rules. In *Monte Carlo and Quasi-Monte Carlo Methods 2002*, ed. H. Niederreiter. Berlin: Springer-Verlag. to appear.

L'Ecuyer, P., and C. Lemieux. 2000. Variance reduction via lattice rules. *Management Science* 46 (9): 1214–1235.

L'Ecuyer, P., and C. Lemieux. 2002. Recent advances in randomized quasi-Monte Carlo methods. In *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, ed. M. Dror, P. L'Ecuyer, and F. Szidarovszki, 419–474. Boston: Kluwer Academic Publishers.

L'Ecuyer, P., and F. Panneton. 2002. Construction of equidistributed generators based on linear recurrences modulo 2. In *Monte Carlo and Quasi-Monte Carlo Methods 2000*, ed. K.-T. Fang, F. J. Hickernell, and H. Niederreiter, 318–330. Berlin: Springer-Verlag.

Lemieux, C., and P. L'Ecuyer. 2001. Selection criteria for lattice rules and other low-discrepancy point sets. *Mathematics and Computers in Simulation* 55 (1–3): 139–148.

Lemieux, C., and P. L'Ecuyer. 2003. Randomized polynomial lattice rules for multivariate integration and simulation. *SIAM Journal on Scientific Computing* 24 (5): 1768–1789.

Liu, R., and A. B. Owen. 2003. Estimating mean dimensionality. manuscript.

Morokoff, W. J. 1998. Generating quasi-random paths for stochastic processes. *SIAM Review* 40 (4): 765–788.

Niederreiter, H. 1987. Point sets and sequences with small discrepancy. *Monatshefte für Mathematik* 104:273–337.

Niederreiter, H. 1988. Low-discrepancy and low-dispersion sequences. *Journal of Number Theory* 30:51–70.

Niederreiter, H. 1992a. Low-discrepancy point sets obtained by digital constructions over finite fields. *Czechoslovak Math. Journal* 42:143–166.

Niederreiter, H. 1992b. *Random number generation and quasi-Monte Carlo methods*, Volume 63 of *SIAM CBMS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia: SIAM.

Niederreiter, H., and G. Pirsic. 2001. Duality for digital nets and its applications. *Acta Arithmetica* 97:173–182.

Niederreiter, H., and C. Xing. 1997. The algebraic-geometry approach to low-discrepancy sequences. In *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, ed. P. Hellekalek, G. Larcher, H. Niederreiter, and P. Zinterhof, Volume 127 of *Lecture Notes in Statistics*, 139–160. New York: Springer-Verlag.

Owen, A. B. 1995. Randomly permuted (t,m,s)-nets and (t,s)-sequences. In *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, ed. H. Niederreiter and P. J.-S. Shiue, Number 106 in Lecture Notes in Statistics, 299–317. Springer-Verlag.

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

Owen, A. B. 1998. Latin supercube sampling for very high-dimensional simulations. *ACM Transactions of Modeling and Computer Simulation* 8 (1): 71–102.

Owen, A. B. 2003. Variance and discrepancy with alternative scramblings. *ACM Transactions of Modeling and Computer Simulation* 13 (4). To appear.

Pirsic, G., and W. C. Schmid. 2001. Calculation of the quality parameter of digital nets and application to their construction. *Journal of Complexity* 17 (4): 827–839.

Sloan, I. H., and S. Joe. 1994. *Lattice methods for multiple integration*. Oxford: Clarendon Press.

Sobol', I. M. 1967. The distribution of points in a cube and the approximate evaluation of integrals. *U.S.S.R. Comput. Math. and Math. Phys.* 7:86–112.

Sobol', I. M., and Y. L. Levitan. 1976. The production of points uniformly distributed in a multidimensional. Technical Report Preprint 40, Institute of Applied Mathematics, USSR Academy of Sciences. In Russian.

Spanier, J., and E. H. Maize. 1994. Quasi-random methods for estimating integrals using relatively small samples. *SIAM Review* 36:18–44.

Tezuka, S. 1990, Jan.. A new family of low-discrepancy point sets. Technical Report RT-0031, IBM Research, Tokyo Research Laboratory.

Tezuka, S. 1995. *Uniform random numbers: Theory and practice*. Norwell, Mass.: Kluwer Academic Publishers.

Tuffin, B. 1998. Variance reduction order using good lattice points in Monte Carlo methods. *Computing* 61:371–378.

AUTHOR BIOGRAPHY

PIERRE L'ECUYER is Professor in the Département d'Informatique et de Recherche Opérationnelle, at the Université de Montréal, Canada. His main research interests are random number generation, quasi-Monte Carlo methods, efficiency improvement via variance reduction, sensitivity analysis and optimization of discrete-event stochastic systems, and discrete-event simulation in general. He obtained the prestigious *E. W. R. Steacie* fellowship in 1995-97 and a *Killam* fellowship in 2001-03. His recent research articles are available on-line at `http://www.iro.umontreal.ca/~lecuyer`.