Variance and discrepancy with alternative scramblings

ART B. OWEN
Stanford University

This paper analyzes some schemes for reducing the computational burden of digital scrambling. Some such schemes have been shown not to affect the mean squared L^2 discrepancy. This paper shows that some discrepancy-preserving alternative scrambles can change the variance in scrambled net quadrature. Even the rate of convergence can be adversely affected by alternative scramblings. Finally, some alternatives reduce the computational burden and can also be shown to improve the rate of convergence for the variance, at least in dimension 1.

Categories and Subject Descriptors: G.1.4 [**Mathematics of Computing**]: Numerical Analysis—*Multidimensional (multiple) quadrature*

General Terms: Algorithms

Additional Key Words and Phrases: Derandomization, Randomization, Quasi-Monte Carlo

1. INTRODUCTION

The purpose of this paper is to study some recent proposals for scrambling of digital nets. These proposals can greatly reduce the time and especially the memory required to scramble nets. Most of the alternative proposals produce the same mean squared L^2 discrepancy as the original scrambling proposal. Many have the same variance. We analyze the structure of scrambling methods, show that one computational shortcut adversely affects the rate of convergence of the sampling variance, and present a new shortcut that, at least for one dimensional problems, improves the convergence rate of the variance.

To frame the discussion, suppose that f is a function on the unit cube $[0,1)^d$ in d dimensions, and that $f \in L^2[0,1)^d$. We will compute an approximation to $I = \int_{[0,1)^d} f(x)dx$ of the form $\hat{I} = (1/n)\sum_{i=0}^{n-1} f(x_i)$. We index from 0 and work with the half-open hypercube purely for notational convenience.

In crude Monte Carlo integration, the x_i are sampled independently from the $U[0,1)^d$ distribution. Then $\hat{I}$ has mean I and variance σ^2/n where $\sigma^2 = \int (f(x) - I)^2 dx < \infty$. The root mean square (RMS) error in crude Monte Carlo is $\sigma n^{-1/2}$. In practice pseudo-random numbers are usually substituted for random ones.

Author's address: A. B. Owen, Department of Statistics, Sequoia Hall, Stanford CA 94305
This work was supported by U.S. NSF grant DMS-0072445.

The goal of quasi-Monte Carlo (QMC) methods (see Niederreiter [1992]) is not to simulate randomness. The goal in QMC is to minimize a measure of the distance between the continuous uniform distribution on $[0,1)^d$ and the discrete uniform distribution taking the value x_i with probability $1/n$. Such distance measures are called discrepancies. QMC constructions give deterministic points x_i with $|\hat{I}-I| = O(n^{-1+\epsilon})$ under mild smoothness conditions on f, for any $\epsilon > 0$. The factor n^ϵ is there to hide powers of $\log(n)$.

Randomized versions of quasi-Monte Carlo were introduced in order to obtain sample based error estimates. An early example is Cranley and Patterson [1976]. Randomization can also bring improvements. Roth [1980] introduced randomness into Halton-Hammersley points to obtain asymptotically optimal L_2 discrepancy. The scrambled nets proposed in Owen [1995] were shown in Owen [1997b] to have an RMS error of $O(n^{-3/2+\epsilon})$ under mild smoothness assumptions.

A straightforward implementation of the full scrambling in Owen [1995] requires an amount of memory proportional to that required to store all the x_i. Shortcuts have been proposed by Tan and Boyle [2000], Matousek [1998b], Fox [1999] and Friedel and Keller [2002]. Perhaps greater savings can be achieved by alternative scramblings. Hickernell [1996a] provides sufficient conditions for a scrambling method to have the same mean square L^2 discrepancy as fully scrambled points. Matousek [1998b] provides slightly different sufficient conditions and gives many specific examples of scrambles satisfying them. One key idea is to use a partial derandomization replacing some independent random variables by pairwise independent ones generated from a small set of independent random variables. Tezuka [2002] and **?**] consider a special class of i-binomial scrambles.

For a comprehensive treatment of quasi-Monte Carlo see Niederreiter [1992]. Surveys of randomized quasi-Monte Carlo are presented in [Owen 1998a; 1999] and L'Ecuyer and Lemieux [2002]. Comparative discussions of scrambling methods appear in L'Ecuyer and Lemieux [2002], Matousek [1998b], and in Hong and Hickernell [2000].

The outline of this paper is as follows. Section 2 reviews digital scrambling schemes. The scheme from Owen [1995], called nested uniform scrambling here, is contrasted with alternatives from Matousek [1998b], in which uniform permutations are replaced by random linear ones, and nested scrambling is replaced by positional or matrix scrambling. Work of Hickernell [1996a] and Matousek [1998b] shows that many of these schemes do not alter the mean squared L^2 discrepancy. Section 3 shows by a one dimensional example based on the van der Corput sequence, that replacing nested scrambling by alternatives can radically change the variance of scrambled integration rules. The asymptotic variance rate can be changed for the better or for the worse. Section 4 has some conclusions and raises some open issues.

2. PERMUTING AND SCRAMBLING

A scrambled quadrature rule starts with points $a_0, \ldots, a_{n-1} \in [0,1)^d$. A mapping from $[0,1)^d$ onto $[0,1)^d$ is chosen at random as described below, and applied to the points a_i. The result, denoted x_i is the scrambled version of a_i. Then I is estimated by $\hat{I}$ as in QMC or crude MC.

Most scrambling strategies proposed for $[0,1)^d$ work by independently scrambling

the d components of a_i. We assume such independence throughout, accepting some loss of generality. The strategies are designed with the goal of preserving in $x_0, \dots, x_{n-1}$ some good equidistribution properties built into $a_0, \dots, a_{n-1}$, while making x uniformly distributed on $[0,1)^d$.

We present digital scrambles of $[0,1)$, obtained by randomizing the digits of a in an integer base $b \geq 2$. For $a \in [0,1)$ write $a = \sum_{k=1}^{\infty} a_k b^{-k}$. Scrambling of a results in a point $x = \sum_{k=1}^{\infty} x_k b^{-k}$ where the digits $x_k \in \{0, 1, \dots, b-1\}$ are obtained by permutation schemes described below applied to the digits of a. Some points in $[0,1)$ have two base b expansions, of which one ends in an infinite sequence of $b-1$'s and the other has an infinite sequence of 0's. We adopt the representation with an infinite sequence of 0's for such points a. In some instances we will work formally with infinite sequences of digits distinguishing the two representations of such points. All the schemes below have probability zero of producing any x_i with two distinct representations, when applied to countably many a_i.

The proposal in Owen [1995] scrambles the interval $[0,1)$ by applying random permutations to the digits a_k. Each random permutation is uniformly distributed on the set of $b!$ permutations of $\{0, 1, \dots, b-1\}$. The permutation used for a_1 is $\pi_\bullet$. The permutation used for a_2 depends on the value of a_1, and is written $\pi_{\bullet a_1}$. The permutation used for a_k for $k > 1$ depends on all the values $a_1, \dots, a_{k-1}$. It is written $\pi_{\bullet a_1 a_2 \dots a_{k-1}}$, and there are b^{k-1} such permutations. We use the term *nested uniform* scrambling for this procedure: *nested* describes the dependence of the permutations for digit $k > 1$ on the values of digits $\ell < k$, and *uniform* describes the use of all $b!$ possible permutations. Some alternative methods use non-uniform permutations while others employ non-nested strategies to combine permutations. The non-nested strategies can lead to substantial reductions in the space needed to store the permutations for scrambling.

This article focusses on alternatives to nested uniform scrambling. Here we mention that some steps have been taken to mitigate the costs of nested uniform scrambling. Tan and Boyle [2000] apply nested scrambling to levels up to k' and then take $\pi_{\bullet a_1 a_2 \dots a_k} = \pi_{\bullet a_1 a_2 \dots a_{k'}}$ for $k \geq k'$. Matousek [1998b] describes a method of caching random seeds to reduce storage costs, while requiring some duplicated computation. Friedel and Keller [2002] present a lazy permutation strategy. It may also be possible to implement nested scrambling by using a prodigious number of random seeds, computed by a hash function, though this does not appear to have been tried. The input to the hash function would contain a position $k \leq M$, where $b^M \gg n$, a $k-1$-tuple of base b digits $a_1, \dots, a_{k-1}$, a coordinate $j \in \{1, \dots, d\}$, and possibly a replicate number $r \in \{1, \dots, R\}$.

2.1 Random permutations

The building blocks in scrambling the base b digits of $a \in [0,1)$ are random permutations of the symbols $\mathbb{Z}_b = \{0, 1, \dots, b-1\}$. We write such a permutation as a function π mapping $\mathbb{Z}_b$ onto $\mathbb{Z}_b$. The nonzero elements of $\mathbb{Z}_b$ are denoted by $\mathbb{Z}_b^+ = \{1, \dots, b-1\}$. Here we list some useful random permutations and some key properties of them.

Definition 2.1. In a *uniform random permutation* all $b!$ permutations of $\mathbb{Z}_b$ have probability $1/b!$.

Definition 2.2. Let b be a prime. A *linear random permutation* has the form $\pi(a) = h \times a + g \mod b$ where $h \in \mathbb{Z}_b^+$ and $g \in \mathbb{Z}_b$ are independent and uniformly distributed over their ranges.

Linear random permutations are restricted to prime b because for nonprime b there are $h \neq 0$ for which $h \times a + g$ is not a permutation. As an example take $b = 4$ and $h = 2$. When b is a prime number, then $\mathbb{Z}_b$ is a finite field. For a general finite field $GF(b)$ the number of elements b is a prime number raised to a positive integer power. Only for prime b can arithmetic modulo b be used. To extend the linear permutation we employ invertible functions (bijections) from $GF(b)$ to $\mathbb{Z}_b$. In the definition below, a bijection maps $\mathbb{Z}_b$ into $GF(b)$ where the arithmetic is carried out, and then a second bijection, possibly equal to the first, is used to bring the result back into $\mathbb{Z}_b$:

Definition 2.3. Let b be a prime power. Then a *generalized linear random permutation* has the form $\pi(a) = \Phi^{-1}(h \times \Psi(a) + g)$ where Φ and Ψ are bijections from $\mathbb{Z}_b$ onto $GF(b)$, addition and multiplication are carried out in $GF(b)$, h and g are independent uniformly distributed elements of $GF(b) - \{0\}$ and $GF(b)$ respectively.

Definition 2.4. A *digital shift random permutation* has the form $\pi(a) = a + g \mod b$ where g is uniformly distributed on $\mathbb{Z}_b$. For a prime power b, a *generalized digital shift random permutation* has the form $\pi(a) = \Phi^{-1}(\Psi(a) + g)$ where Φ and Ψ are bijections from $\mathbb{Z}_b$ onto $GF(b)$ and g is uniformly distributed on $GF(b)$.

Linear permutations are described in Matousek [1998b]. Digital shifts are mentioned in L'Ecuyer and Lemieux [2002]. They yield random cyclic permutations. For linear permutations it suffices to store only 2 coefficients per permutation compared to the b coefficients ordinarily used for a uniform random permutation, and digital permutations reduce the storage to one coefficient.

Definition 2.5. The random permutation π has *single-fold uniformity* if

$$\Pr(\pi(a) = x) = \frac{1}{b} \tag{1}$$

for all $a, x \in \mathbb{Z}_b$, *two-fold uniformity* if

$$\Pr(\pi(a_1) = x_1,\ \pi(a_2) = x_2) = \frac{1}{b(b-1)} \tag{2}$$

whenever $a_1, a_2, x_1, x_2 \in \mathbb{Z}_b$ with $a_1 \neq a_2$ and $x_1 \neq x_2$, and it has *k-fold uniformity*, for $1 \leq k \leq b$ if

$$\Pr(\pi(a_1) = x_1, \ldots, \pi(a_k) = x_k) = \frac{1}{b(b-1)\cdots(b-k+1)} \tag{3}$$

whenever $a_1, \ldots, a_k$ and $x_1, \ldots, x_k$ are both lists of k distinct elements from $\mathbb{Z}_b$.

When π has k-fold uniformity, the random vector $(\pi(0), \ldots, \pi(b-1)) \in {\mathbb{Z}_b}^b$ has k dimensional margins matching those of uniform random permutations. It is easy to see that $(b-1)$-fold uniformity implies b-fold uniformity, which in turn is equivalent to uniformity. The definition of k-fold uniformity is equivalent to $\pi(a_1), \ldots, \pi(a_k)$ being a simple random sample (without replacement) from $\mathbb{Z}_b$ whenever $a_1, \ldots, a_k$

are distinct elements of $\mathbb{Z}_b$. See Cochran [1977] for background on simple random sampling.

Conditions (1) and (2) were used by Hickernell [1996a]. Matousek [1998b] shows that linear random permutations satisfy them, as do generalized linear random permutations. Digital shifts and generalized digital shifts have single-fold uniformity (1). When $b = 2$ there are only 2 permutations, $\pi(a) = a$ and $\pi(a) = 1 - a$. Then uniform, generalized linear and generalized digital shift permutations coincide. When $b = 3$ then (generalized) linear permutations have 3-fold uniformity.

Multiplicative permutations such as $\pi(a) = ha \mod b$ for prime b and h uniformly distributed on nonzero values, do not satisfy (1). For instance $\pi(0) = 0$ with probability 1 under multiplicative permutation. Similarly the random permutation that has $\pi(a) = a$ or $\pi(a) = b - 1 - a$ each with probability 1/2 does not satisfy (1) when $b > 2$.

2.2 Random scrambles

A scramble is a method of randomizing the digits a_k in the base b expansion of $a \in [0, 1)$. A scramble applies random permutations which may be of the various types described above.

Definition 2.6. In a *nested scramble* $x_1 = \pi_\bullet(a_k)$, and $x_k = \pi_{\bullet a_1 a_2 \dots a_{k-1}}(a_k)$, for $k \geq 2$ for independent random permutations $\pi_\bullet$ and $\pi_{\bullet a_1 a_2 \dots a_{k-1}}(a_k)$ for $k \geq 2$ and $a_1, \dots, a_{k-1} \in \mathbb{Z}_b$.

Nested scrambling uses b^{k-1} permutations to randomize the k'th digit. Useful scrambles can be constructed with many fewer permutations. One proposal is to use a single permutation $x_k = \pi_k(a_k)$ at the k'th position in the base b expansion of a.

Definition 2.7. In a *positional scramble* $x_k = \pi_k(a_k)$, where π_k for $k \geq 1$ are independent random permutations.

The nomenclature for scrambling strategies is not standardized. To distinguish between alternatives to nesting and alternatives to uniformity, we suggest the names *positional uniform*, *nested linear*, and *positional linear* to describe some of the alternatives to nested uniform scrambling. The first term describes the scrambling framework and the second describes the type of permutations used in that framework. Positional linear scrambling corresponds to one proposal in Hickernell and Yue [2000]. There are also positional and nested digital shifts.

Matousek [1998b] proposed random linear scrambling that for prime b can be written

$$x_k = \sum_{j=1}^{k} M_{kj} a_j + C_k, \tag{4}$$

where M_{kj} and C_k are elements of $\mathbb{Z}_b$. We'll adopt a convention that elements M_{kj} and C_k are independent and uniformly distributed over their ranges, unless otherwise specified. One version has $C_k = 0$ and another has $C_k \in \mathbb{Z}_b$. Choosing $C_k \in \mathbb{Z}_b$ versus $C_k = 0$ has the effect of applying a positional digital shift scramble.

Matousek took $M_{kk} \in \mathbb{Z}_b^+$ and $M_{jk} \in \mathbb{Z}_b$ for $0 < j < k$. Matousek's scrambling can be represented by a lower triangular matrix M. The structure of the scrambling

can be described by the structure of the matrix M. We use the term matrix scrambling to describe Matousek's random linear scrambling as well as some other scrambles. In linear matrix scrambling $C_k = 0$ while for affine matrix scrambling C_k is a random element of $\mathbb{Z}_b$.

Matousek [1998b; 1998a] takes the generalized Faure construction of Tezuka [1995] as a starting point for his linear scrambles, and remarks that the affine version might have some additional value. Generalized Faure sequences have digits obtained by multiplying the generator matrices of the Faure [1982] nets on the left by some invertible matrices. Faure and Tezuka [2002a] describe a randomization obtained by multiplying the generator matrices on the right. Such randomizations reorder the integration points. Hong and Hickernell [2000] discuss and implement both types of randomization. **?**] consider left and right versions of i-binomial scrambling.

Definition 2.8. A *linear matrix scramble* of $a = \sum_{k=1}^{\infty} a_k b^{-k}$ takes the form $a \to x = \sum_{k=1}^{\infty} x_k b^{-k}$, where each x_k is given by (4), for random elements M_{kj} with $C_k = 0$. An *affine matrix scramble* of $a = \sum_{k=1}^{\infty} a_k b^{-k}$ takes the form $a \to x = \sum_{k=1}^{\infty} x_k b^{-k}$, where each x_k is given by (4), for random elements M_{kj} with C_k uniformly distributed in $\mathbb{Z}_b$ independently of each other and independent of M_{kj}.

As k increases for fixed b, matrix scrambling requires $O(k^2)$ storage to permute k digits compared to $O(k)$ for positional scrambling and $O(b^k)$ for nested scrambling. Tezuka's [2002] i-binomial scrambling described below reduces the storage to $O(k)$.

The permutation applied to a_k is a linear one with $h = M_{kk}$ and $g = C_k + \sum_{j=1}^{k-1} M_{kj} a_j$. As in nested scrambling the permutation applied to a_k depends on the digits $a_1, \ldots, a_{k-1}$. We will consider some linear matrix scrambles where the matrix has a distribution different from that proposed by Matousek [1998b]. A matrix element uniformly distributed on $\mathbb{Z}_b - \{0\}$ is represented by the letter h, an element that is uniformly distributed on $\mathbb{Z}_b$ is represented by a g, and one that must be zero is denoted 0. Elements that are constrained to be equal to each other have the same indices, otherwise all elements are independent. Scrambling by matrices

$$\begin{pmatrix} h_{11} & 0 & 0 & 0 & \cdots \\ g_{21} & h_{22} & 0 & 0 & \cdots \\ g_{31} & g_{32} & h_{33} & 0 & \cdots \\ g_{41} & g_{41} & g_{43} & h_{44} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \begin{pmatrix} h_1 & 0 & 0 & 0 & \cdots \\ g_2 & h_1 & 0 & 0 & \cdots \\ g_3 & g_2 & h_1 & 0 & \cdots \\ g_4 & g_3 & g_2 & h_1 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}, \text{ and, } \begin{pmatrix} h_1 & 0 & 0 & 0 & \cdots \\ h_1 & h_2 & 0 & 0 & \cdots \\ h_1 & h_2 & h_3 & 0 & \cdots \\ h_1 & h_2 & h_3 & h_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (5)$$

corresponds respectively, to Matousek's random linear scrambling (4), Tezuka's [2002] i-binomial scrambling, and a new matrix scramble proposed here. The third matrix has constant vertical stripes below the diagonal, and is called striped matrix scrambling below. Notice that for $b = 2$ striped matrix and i-binomial scrambling both have 1's on the diagonal, but they differ because i-binomial scrambling allows 0's below the diagonal. Faure [2001] considered discrepancy of generalized Faure sequences in which the generators are premultiplied by a lower triangular matrix with ones on and below the diagonal.

In all these examples the diagonal has nonzero elements. It is important for the upper m by m submatrix of M to be invertible, for otherwise two distinct

digit m-tuples $(a_1, \ldots, a_m)$ and $(a'_1, \ldots, a'_m)$ can both give the same output digits $(x_1, \ldots, x_m)$.

It is clear that matrix scrambles can be generalized to prime power bases b, using bijections and finite field arithmetic. The details are straightforward and are omitted.

The structure of matrix scrambling can be mimicked, using a matrix of uniform random permutations, replacing addition by composition. For $1 \le j \le k < \infty$ let π_{jk} be a uniformly distributed random permutation. A uniform matrix scramble has

$$x_k = \pi_{kk}\bigl(\pi_{k\,k-1}\bigl(\cdots\bigl(\pi_{k1}(a_k)\bigr)\bigr)\bigr).$$

A closer analogue to linear matrix scrambling would require that the digits a_j for $j < k$ play a role in π_{kj}. An example is

$$x_k = \pi_{kk}\bigl(a_k + \pi_{k\,k-1}\bigl(a_{k-1} + \pi_{k\,k-2}\bigl(a_{k-2} + \cdots \pi_{k1}\bigl(a_1\bigr)\bigr)\bigr)\bigr),$$

with addition interpreted modulo b.

3. VARIANCE AND DISCREPANCY EFFECTS

The primary application of scrambled nets is as an alternative to crude Monte Carlo for integration problems. First we note that the expected value of $\hat{I}$ remains I when crude Monte Carlo is replaced by various scrambles of digital nets.

PROPOSITION 3.1. *For $a \in [0,1)$ let x be the scrambled version of a in base b, under affine matrix scrambling, or under nested or positional scrambling with permutations that satisfy* (1)*. Assume that all of the permutations employed in scrambling a are independent. Then $x \sim U[0,1)$.*

Proof: Write $a = \sum_{k=1}^{\infty} a_k b^{-k}$. Under positional scrambling all of the $x_k = \pi_k(a_k)$ are independent and by (1) uniformly distributed on $\{0, 1, \ldots, b-1\}$. Therefore $x_k \sim U[0,1)$ under positional scrambling. The result is proved for nested scrambles in Owen [1995]. The result follows for affine matrix scrambling because it is a linear matrix scramble followed by a positional digital shift. □

Proposition 3.1 applies to an individual point $a \in [0,1)$, and so it applies to $a \in [0,1)^d$ when independent scrambles are applied to each coordinate. Thus all points a_i in a scrambled quadrature rule are uniformly distributed in $[0,1)^d$ under the scramblings in Proposition 3.1. Unbiasedness under these scrambling rules does not require that the a_i are points of a net.

Remark 3.2. Unlike affine matrix scrambling, linear matrix scrambling does not necessarily give a uniform distribution for x. For example x_1 equals $M_{11}a_1$, and therefore $x_1 = 0$ if and only if $a_1 = 0$, in the usual setting where $M_{11} \neq 0$. Thus x cannot have the $U[0,1)$ distribution. The estimate $\hat{I}$ may still be unbiased if n is a multiple of b. But, because affine matrix scrambling is unbiased and requires so little extra time and space compared to linear matrix scrambling, we prefer affine matrix scrambling to linear matrix scrambling.

Variance and discrepancy provide ways of quantifying the quality of alternative unbiased scrambling schemes. For the scrambling schemes here with $E(\hat{I}) = I$, the

variance of $\hat{I}$ is

$$E((\hat{I}-I)^2) = \frac{1}{n^2}\sum_{i=0}^{n-1}\sum_{i'=0}^{n-1} E\Big((f(x_i)-I)(f(x_{i'})-I)\Big). \tag{6}$$

While variance is specific to an integrand f, discrepancy measures can describe integration error over a large class of integrands. Discrepancy bounds take the form $|\hat{I}-I| \le D(x_0,\dots,x_{n-1})\|f\|$ where D is a discrepancy, and $\|f\|$ is a compatible norm or semi-norm on functions. The Zaremba bound [Zaremba 1968] and the Koksma-Hlawka bound [Hlawka 1961] are two well known examples. These have been generalized and extended in a series of papers by Fred Hickernell [1996b; 1997; 1998].

The squared L^2 discrepancy, $D_2^{*2} = D_2^{*2}(x_0,\dots,x_{n-1})$ is

$$D_2^{*2} = \Big(\frac{4}{3}\Big)^d - \frac{2}{n}\sum_{i=0}^{n-1}\prod_{j=1}^{d}\frac{3-x_{ij}^2}{2} + \frac{1}{n^2}\sum_{i=0}^{n-1}\sum_{i'=0}^{n-1}\prod_{j=1}^{d}\Big(2-\max(x_{ij},x_{i'j})\Big). \tag{7}$$

The expectation of (7) over random x_i is the mean squared L^2 discrepancy, studied in Hickernell [1996a] and Matousek [1998b]. The Zaremba bound uses D_2^*.

Both the mean square discrepancy and the variance depend on the joint distribution of $x_0,\dots,x_{n-1}$ only through pairwise joint distributions of x_i and $x_{i'}$. Many alternative scrambling schemes have been shown to leave the expected value of (7) unchanged [Hickernell 1996b; Matousek 1998b].

3.1 Example

We show by an example that the variance (6) can change if nested scrambling is replaced by positional scrambling or by certain matrix scrambles. The example uses a scrambled $(0,m,1)$–net in base 2, described below in concrete terms. For now we note that such a scrambled $(0,m,1)$–net consists of 2^m points, of which each interval $[i2^{-m},(i+1)2^{-m})$ for integer $0 \le i < 2^m$ contains exactly one point. That point is uniformly distributed in the interval. Where the scrambling methods differ is in the joint distribution of pairs of points.

We study the scrambling variance with a scrambled version of the van der Corput [1935a; 1935b] points in $[0,1)$. Write the integer $i \ge 0$ in base 2 as $i = \sum_{k=1}^{\infty} a_{ik}2^{k-1}$ for digits $a_{ik} \in \{0,1\}$ of which only finitely many are nonzero for each i. The i'th van der Corput point is $a_i = \sum_{k=1}^{\infty} a_{ik}2^{-k}$.

There are only two permutations of $\{0,1\}$, the identity permutation $a \to a$ and $a \to 1-a$. Thus for $b=2$, the uniform, linear, and digital shift permutations coincide. We compare scrambling strategies using the integrand $f(x) = x = \sum_{k=1}^{\infty} x_k 2^{-k}$. Trivially $I = 1/2$ and ordinary Monte Carlo sampling has $V(\hat{I}) = 1/(12n)$.

Nested uniform scrambling of $n = 2^m$ points of the van der Corput sequence is known [Owen 1997b] to be equivalent to stratified sampling in which one point is chosen uniformly and independently from within each of n intervals $[i/n,(i+1)/n)$ for $i = 0,\dots,n-1$. Accordingly $V(\hat{I}) = 1/(12n^3)$ for nested scrambling of 2^m van der Corput points with $f(x) = x$.

Now suppose that positional scrambling is used. Then $x_i = \sum_{k=1}^{\infty} \pi_k(a_{ik})2^{-k}$

$(i)_2$	a_i	z_i	x_i
0	$0.0000\dot{0}$	$0_\circ 0000\dot{0}$	$0.g_1g_2g_3g_4\cdots$
1	$0.1000\dot{0}$	$0_\circ 111\dot{1}$	$0.\bar{g}_1\bar{g}_2\bar{g}_3\bar{g}_4\cdots$
10	$0.0100\dot{0}$	$0_\circ 011\dot{1}$	$0.g_1\bar{g}_2\bar{g}_3\bar{g}_4\cdots$
11	$0.1100\dot{0}$	$0_\circ 100\dot{0}$	$0.\bar{g}_1g_2g_3g_4\cdots$
100	$0.0010\dot{0}$	$0_\circ 001\dot{1}$	$0.g_1g_2\bar{g}_3\bar{g}_4\cdots$
101	$0.1010\dot{0}$	$0_\circ 110\dot{0}$	$0.\bar{g}_1\bar{g}_2g_3g_4\cdots$
110	$0.0110\dot{0}$	$0_\circ 010\dot{0}$	$0.g_1\bar{g}_2g_3g_4\cdots$
111	$0.1110\dot{0}$	$0_\circ 101\dot{1}$	$0.\bar{g}_1g_2\bar{g}_3\bar{g}_4\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table I. This table shows the results of affine striped matrix scrambling of the van der Corput sequence. The first column has i running from 0 to 7 in base 2. The second column shows the van der Corput points obtained by reflecting the digits of i about the base 2 point. The third column shows cumulative sums of digits of a_i modulo 2. The open decimal point is a reminder that z_i is written in a formal representation of the points, in which, for example, $0_\circ 0\dot{1}$ is distinct from $0_\circ 1\dot{0}$. The fourth column shows the scrambled points x_i.

and $\hat{I}$ becomes

$$\begin{aligned}\frac{1}{n}\sum_{i=0}^{n-1}\sum_{k=1}^{\infty}\pi_k(a_{ik})2^{-k} &= \frac{1}{n}\sum_{k=1}^{m-1}\bigl[\pi_k(0)+\pi_k(1)\bigr]2^{-k+m-1}+\sum_{k=m}^{\infty}\pi_k(0)2^{-k}\\ &= \frac{1}{2}\Bigl(1-2^{1-m}\Bigr)+\sum_{k=m}^{\infty}\pi_k(0)2^{-k}.\end{aligned}$$

Elementary arguments then give that $E(\hat{I}) = I = 1/2$ and

$$V(\hat{I}) = \frac{1}{4}\sum_{k=m}^{\infty}4^{-k} = \frac{1}{12n^2}.$$

Nested scrambling yields $V(\hat{I}) = 1/(12n^3)$ in this example, so positional scrambling has adversely affected the asymptotic rate of variance. In this instance positional scrambling is equivalent to a randomized shifted lattice rule [Cranley and Patterson 1976] in which $x_i = i/n+U \mod 1$ where $U \sim U[0,1)$ and $z \mod 1$ means $z-\lfloor z\rfloor$.

Next consider affine striped matrix (ASM) scrambling corresponding to the third matrix pattern in (5). There is only one nonzero element in $\{0,1\}$ and so $M_{kj} = 1$ for $1 \le j \le k < \infty$. Each C_k is 0 or 1 independently with probability 1/2. We take the digit sequence $a_i = 0.a_{i1}a_{i2}a_{i3}\cdots$ in base 2 and apply first the M_{kj} producing z_i with formal digits $z_{ik} = \sum_{j=1}^{k} a_{ij} \mod 2$. The digits of z_i are cumulative sums of those of a_i and because every a_i ends in an infinite sequence of zeros, every z_i ends in an infinite sequence of either zeros or ones, depending on what the cumulative sum was when the tail of zeros began. Finally $x_i = \sum_{k=1}^{\infty} x_{ik}2^{-k}$ where $x_{ik} = z_{ik} + g_k \mod 2$ with $g_k = C_k$.

Table I illustrates ASM scrambling of van der Corput sequence. The table shows, and it is easy to prove directly, that $x_{2\ell} + x_{2\ell+1} = 1$ for $\ell \ge 0$. It follows that if n is even then ASM scrambling of the van der Corput sequence takes observations in antithetic pairs and so $V(\hat{I}) = 0$ for $f(x) = x$ or any other f with $f(x)+f(1-x) = 2I$. To summarize:

PROPOSITION 3.3. *Let $f(x) = x$ on $[0,1)$ and put $n = 2^m$ for an integer $m \geq 1$. Let $a_0, \ldots, a_{n-1}$ be the first n points of the van der Corput sequence, suppose that x_i is a base b scrambled version of a_i, and let $\hat{I} = (1/n)\sum_{i=0}^{n-1} f(x_i)$. Then*

$$V(\hat{I}) = \begin{cases} \frac{1}{12n^3} & \text{for nested scrambling} \\ \frac{1}{12n^2} & \text{for positional scrambling} \\ 0 & \text{for affine striped matrix scrambling} \end{cases}$$

Proof: The first three results follow from the preceeding discussion. If $m = 0$ then all three methods have variance $1/12$. The result for ASM scrambling extends to any even n. □

3.2 Local antithetic property

The previous example shows that the variance differs according to the type of scrambling used. The function $f(x) = x$ from that example is clearly special. But some superiority of ASM scrambling of the van der Corput sequence extends to more general functions on $[0,1)$.

The scrambled points are more than just antithetic with respect to the interval $[0,1)$. The $n/2$ points in $[0,1/2)$ are arranged in pairs centered on $1/4$ and the $n/2$ points in $[1/2,1)$ are arranged in pairs centered on $3/4$. More generally the $n/2^{\bar{m}}$ points in the interval $[i2^{-\bar{m}}, (i+1)2^{-\bar{m}})$ for $0 \leq i < 2^{\bar{m}}$ and $0 \leq \bar{m} < m$ are arranged in pairs centered on the point $(i+1/2)2^{-\bar{m}}$.

For $i = 0, \ldots, 2^{m-1} - 1$ the points x_i and $x_{i+2^{m-1}}$ are antithetic complements in a subinterval of width 2^{-m+1}. We refer to this as a local antithetic property. Local antithetic sampling was studied by Haber [1967]. ASM scrambled points are locally and globally antithetic.

All the randomness in ASM scrambling of the van der Corput sequence is contained in the M_{k0} which are the digits of x_0. Accordingly the whole sequence can be simulated to K bits of accuracy using just K random bits, and various reflections to produce x_i from x_0. Such extreme multiple use of a single random value x_0 complicate expressions for the variance of $\hat{I}$, but for integrands with bounded second derivatives we can get an upper bound on the variance. We use a Lemma from De Boor[1978] on piecewise linear interpolation:

LEMMA 3.4. *Let $f(x)$ be a function with $\sup_{0\leq x\leq 1} |f''(x)| \leq B < \infty$ on $[0,1)$. Let $0 = t_0 < t_1 < \cdots < t_J = 1$ and set $\Delta = \max_{0\leq j<J} t_{j+1} - t_j$. Let $\widetilde{f}(x)$ be the unique continuous function, linear over $[t_j, t_{j+1}]$ for $0 \leq j < J$, with $\widetilde{f}(t_j) = f(t_j)$ for $0 \leq j \leq J$. Then*

$$\sup_{0\leq x\leq 1} |f(x) - \widetilde{f}(x)| \leq \frac{B\Delta^2}{8}, \tag{8}$$

and when $t_j = j/J$, then $\sup_{0\leq x\leq 1} |f(x) - \widetilde{f}(x)| \leq B/(8J^2)$.

Proof: Equation (8) is obtained from Chapter III (equation (2)) of De Boor [1978]. The second conclusion follows because $\Delta = 1/J$ for the $J+1$ equispaced points. □

PROPOSITION 3.5. *Suppose that* $\sup_{0\le x\le 1}|f''(x)|\le B<\infty$ *on* $[0,1)$. *Then for* $n=2^m$ *where* $m\ge 1$,

$$V(\hat{I})\le\frac{B^2}{n^4}=O(n^{-4})$$

under ASM matrix scrambling in base 2 *of the first* n *points of the van der Corput sequence.*

Proof: Assume that $m\ge 1$, so $n\ge 2$. Let $\widetilde{f}_{m-1}(x)$ be a piecewise linear function that interpolates $f(x)$ at $x=i/2^{m-1}$ for $i=0,1,\dots,2^{m-1}$. Then $\sup_{0\le x\le 1}|f(x)-\widetilde{f}_{m-1}(x)|\le B/(2n^2)$ by Lemma 3.4 with $J=n/2$. The function $\widetilde{f}_{m-1}$ is integrated without error by ASM matrix scrambling because of the local antithetic property. Now

$$\begin{aligned}|\hat{I}-I|&\le\left|\frac{1}{n}\sum_{i=0}^{n-1}\widetilde{f}_{m-1}(x)-\int_{[0,1)}\widetilde{f}_{m-1}(x)dx\right|\\&\quad+\left|\frac{1}{n}\sum_{i=0}^{n-1}\Bigl(\widetilde{f}_{m-1}(x)-f(x)\Bigr)-\int_{[0,1)}\widetilde{f}_{m-1}(x)-f(x)dx\right|\\&\le B/n^2.\end{aligned}$$

Finally $V(\hat{I})=E((\hat{I}-I)^2)\le B^2n^{-4}=O(n^{-4})$. □

The proof of Proposition 3.5 applies to any unbiased locally antithetic method. Local antithetic sampling applied independently within each of $n/2$ strata attains a variance that is $O(n^{-5})$ in this case as shown in Haber [1967]. Proposition 3.5 would still be true if the variance were $o(n^{-4})$, but simple numerical experiments with $f(x)=x^2$ on $[0,1)$ show that the rate is no better than n^{-4}. Independent local stratification yields a beneficial error cancellation not seen in ASM scrambling.

While the variance of ASM scrambling is complicated, one generally useful technique is to analyze it in two steps, the first generating the M_{kj} and producing a formal base b expansion, and the second generating C_k conditionally on M_{kj}.

PROPOSITION 3.6. *Let* $a_0,\dots,a_{n-1}\in[0,1)^d$. *Suppose that the* d *components of the* a_i *are scrambled independently using affine matrix scrambling. Then* $V(\hat{I})=E(V(\hat{I}\mid M))$.

Proof: We may write $V(\hat{I})=E(V(\hat{I}\mid M))+V(E(\hat{I}\mid M))$. The inner means and variances are with respect to C_k with M_{kj} fixed for $0<j\le k$, while the outer ones are with respect to M_{kj}. Given M_{kj} adding C_k introduces a positional digital shift scramble satisfying (1). Therefore $E(\hat{I}\mid M)=I$ by Proposition 3.1 and so $V(E(\hat{I}\mid M))=0$. □

Using Proposition 3.6 and a notion in which $b\ge 2$ points can be locally antithetic, it is possible to extend Proposition 3.5 to radical inverse schemes in integer bases $b\ge 2$. To define the radical inverse points let $i=\sum_{k=1}^{\infty}a_{ik}b^{k-1}$ where $a_{ik}\in\mathbb{Z}_b$, and put $a_i=\sum_{k=1}^{\infty}a_{ik}b^{-k}$.

PROPOSITION 3.7. *Suppose that* $\sup_{0\le x\le 1}|f''(x)|\le B<\infty$ *on* $[0,1)$. *Let* b *be*

a prime number and let $n = b^m$ for $m \geq 1$. Then

$$V(\hat{I}) \leq \frac{B^2 b^4}{16 n^4} = O(n^{-4})$$

under ASM scrambling in base b of the first n points of the radical inverse sequence in base b.

Proof: We will show that conditionally on the values of M_{kj} for $1 \leq j \leq k < \infty$, the scrambled points integrate without error a piecewise linear continuous interpolation of f at $b^{m-1}+1$ points tb^{-m+1} for $0 \leq t \leq b^{m-1}$. Then from Lemma 3.4 we can conclude that $|\widetilde{f} - f|$ is uniformly smaller than $B/(8b^{2m-2}) = Bb^2/(8n^2)$. Then $|\hat{I} - I| \leq Bb^2/(4n^2)$, so $V(\hat{I} \mid M) \leq B^2 b^4/(16n^4)$, and the result will follow by Proposition 3.6.

Let Z be the infinite matrix with elements

$$z_{ik} = \sum_{j=1}^{k} M_{kj} a_{ij} = \sum_{j=1}^{k} h_j a_{ij} \pmod b$$

for indices $i \geq 0$ and $k \geq 1$ and independent h_j uniformly distributed on $\mathbb{Z}_b^+$. Because $M_{kk} = h_k \neq 0$ the upper left m by m submatrix of M is invertible. Thus the upper left b^m by m submatrix of Z contains each of the b^m points of ${\mathbb{Z}_b}^m$ exactly once. Because $a_{ik} = 0$ for $k > m$ and $i < b^m$ it follows that $z_{ik} = z_{im}$ for $k > m$ and $i < b^m$. Thus every row $i < b^m$ of Z has an infinite tail of repeated values, from element m onwards.

Each distinct $m-1$ tuple in ${\mathbb{Z}_b}^{m-1}$ appears exactly b times as the first $m-1$ rows of Z. Among every such set of b rows there is one row with an infinite tail of 0's, one with an infinite tail of 1's, and, $\cdots$ one with an infinite tail of $b-1$'s.

Let X be the infinite matrix with $x_{ik} = z_{ik} + g_k \mod b$ where $g_k = C_k$ are independent uniform random elements of $\mathbb{Z}_b$. Each distinct $m-1$ tuple in ${\mathbb{Z}_b}^{m-1}$ appears exactly b times as the first $m-1$ rows of X.

It happens with probability one that g_k for $k \geq m$ are not all equal to a common value. Because we are working out the variance of a bounded function we can simply assume that the g_k for $k \geq m$ are not all equal. Then for each non-negative integer $t < b^{m-1}$ the interval $[tb^{-m+1}, (t+1)b^{-m+1})$ contains exactly b of the points $x_i = \sum_{k=1}^{\infty} x_{ik} b^{-k}$ for $0 \leq i < b^m$. Let $\mathcal{I}_t$ be the set of indices $i < b^{m-1}$ with $\lfloor b^{m-1} x_i \rfloor = t$. For each $c \in \mathbb{Z}_b$ there is one $i \in \mathcal{I}_t$ such that $x_{ik} = c + g_k \mod b$ for $k \geq m$.

The average value of x_i for $i \in \mathcal{I}_t$ is

$$tb^{-m+1} + \sum_{k=m}^{\infty} b^{-k} \left(\frac{1}{b} \sum_{c=0}^{b-1} c \right) = tb^{-m+1} + \left(\frac{b^{-m}}{1 - b^{-1}} \right) \left(\frac{b-1}{2} \right) = \left(t + \frac{1}{2} \right) b^{-m+1}.$$

It follows that a piece-wise linear function, continuous on $[b^{-m+1}t, b^{-m+1}(t+1))$, for $t \in \mathbb{Z}_{b^{m-1}}$ is integrated without error by averaging over x_i for $i \in \mathbb{Z}_{b^m}$. □

Scrambling schemes may be applied very generally without assuming that a_i are a net. This example shows that even in the special case of a scrambled $(0, m, s)$-net, positional, ASM, and nested scrambling are not variance equivalent, even to the extent of having different rates of convergence.

3.3 Worst case comparisons

The three methods of scrambling the van der Corput sequence differ in their worst case performance relative to crude Monte Carlo. Let $\sigma^2 = \int_{[0,1)}(f(x)-I)^2dx$ be the variance of f. In Monte Carlo sampling $V(\hat{I}) = \sigma^2/n$. For nested scrambling of the van der Corput sequence and $n = 2^m$ it is known [Owen 1997a] that $V(\hat{I}) \le \sigma^2/n$. For positional or ASM scrambling we can construct a function (depending on m) for which $f(x_i)$ is constant in i with probability 1 leading to $V(\hat{I}) = \sigma^2$. Such unfavorable functions for positional scrambling are periodic with period $1/n$. Unfavorable functions for ASM scrambling are periodic with period $2/n$ and are symmetric about $1/n$ within the interval $[0, 2/n)$.

3.4 Discrepancy

To see why the scrambling strategies from Section 3.1 don't affect mean squared L^2 discrepancy, note that for $d = 1$ formula (7) reduces to

$$D_2^{*2} = \frac{4}{3} - \frac{1}{n}\sum_{i=0}^{n-1}\bigl(3 - x_i^2\bigr) + \frac{1}{n^2}\sum_{i=0}^{n-1}\sum_{i'=0}^{n-1}\Bigl(2 - \max(x_i, x_{i'})\Bigr). \tag{9}$$

The expected value of x_i^2 is common to all scrambling methods because in all cases x_i has the $U[0,1)$ distribution.

The double sum in (9) appears to depend on pairs of x_i but the dependence is not essential. For each $0 \le i < n$, there is exactly one $x_{i'}$ in $[i/n, (i+1)/n)$. Suppose the x_i are written in increasing order $x_{(0)} \le x_{(1)} \le \cdots \le x_{(n-1)}$. Then $x_{(i)} = (i+U_i)/n$ where the U_i are uniformly distributed on $[0,1)$ but are not necessarily independent. Then

$$\begin{aligned}\sum_{i=0}^{n-1}\sum_{i'=0}^{n-1}\max(x_i, x_{i'}) &= \sum_{i=0}^{n-1}\sum_{i'=0}^{n-1}\max(x_{(i)}, x_{(i')})\\ &= \frac{1}{n}\sum_{i=0}^{n-1}\sum_{i'=0}^{n-1}\max(i+U_i, i'+U_{i'})\\ &= \frac{1}{n}\sum_{i=0}^{n-1}(i+1)(i+U_i),\end{aligned}$$

which has an expectation unaffected by the joint behavior of the U_i. Dependence among the U_i does not play a role because the points $x_{(i)}$ cannot cross each other no matter what values the U_i take.

3.5 Discussion

If we follow through the development in Owen [1997a] of the variance under nested uniform scrambling, we find that one argument in Lemma 4 can fail to hold if nested scrambling is replaced by another scrambling, such as positional or ASM scrambling. The issue that arises for non-nested scramblings is that the k'th digits of a_i and $a_{i'}$ can get a statistically dependent permutation even if one or more of the j'th digits of a_i and $a_{i'}$ differ for $1 \le j < k$. With nested scrambling these digits are randomized independently. For example any nested scrambling with permutations

that satisfy one and two-fold uniformity, (1) and (2), has the same variance as nested uniform scrambling. In particular, nested random linear scrambling has the same variance as nested uniform scrambling.

A quadrature rule with small discrepancy has a correspondingly small error for a large class of integrands. In a sequence of such rules indexed by the number n of points, the relative error bound holds even if a different function f_n is chosen for each rule. Hickernell [1996a] shows that the mean squared L^2 discrepancy for scrambled $(0, m, d)$–nets (and some alternatives scramblings) decreases as $O(n^{-2}(\log n)^{d-1})$. As a consequence positional scrambling should yield errors that are $O_p(n^{-1}(\log n)^{(d-1)/2})$.

Under mild smoothness conditions on f, the variance of nested scrambled net quadrature decreases as $O(n^{-3}(\log n)^{d-1})$, corresponding to RMS errors that are $O(n^{-3/2}(\log n)^{(d-1)/2})$. Hickernell [1996a] explains the better rate as a consequence of having an integrand that does not change with n.

The rate at which $V(\hat{I})$ converges to zero for nested uniform scrambling depends on the decay rate of coefficients in an expansion of f into products of base b Haar-like functions. Rates of this type apply equally to nested uniform and nested pair-wise uniform scrambling. The variance rate $O(n^{-3}(\log n)^{d-1})$ has been obtained under varying assumptions on f and the scrambled quadrature rule, in Owen [1997b], Owen [1998b] and Yue and Hickernell [2002].

4. CONCLUSIONS

Scrambling techniques vary greatly in the costs of storage and execution. The mean squared discrepancy is common for a great many scrambling techniques. But integration variance can be strongly affected by the scrambling method. Positional scrambling can greatly increase variance over that of nested scrambling while ASM scrambling can greatly reduce it. Even the rate of convergence for variance can be changed by scrambling methods that leave the discrepancy unchanged.

Haber [1966] shows that stratified sampling of $[0, 1)^d$ with congruent cubical strata, has variance $O(n^{-1-2/d})$ which deterioriates with dimension. In one dimension, nested scrambling generalizes stratified sampling and has variance $O(n^{-3})$ for smooth f. Nested scrambling in d dimensions attains a variance of $O(n^{-3}(\log n)^{d-1})$ in which the dimension effect is less than for stratification. For $d = 1$ ASM scrambling of radical inverse sequences has a local antithetic property that leads to a variance that is $O(n^{-4})$. It is remains to be seen how ASM scrambling compares to other methods for $d > 1$.

Haber [1967] shows that for $d \geq 1$ local antithetic sampling (within congruent subcubes) attains a variance that is $O(n^{-1-4/d})$ for f with uniformly bounded second derivatives. There remains the possibility of incorporating additional randomness into ASM scrambling to preserve its good properties while getting an additional error cancellation like that attained by local antithetic sampling.

Schemes that replace uniformity by pairwise uniformity can leave discrepancies and variances unchanged, but they would be expected to change third and higher moments. Accordingly a central limit theorem like the one of Loh [2002] for nested scrambling might not hold for ASM or other alternative scrambling methods.

Finally, it takes only a small number of random digits to implement ASM scram-

bling, and it emerges that for unfavorable integrands ASM scrambling of nets can be worse than ordinary Monte Carlo sampling. It remains to be seen whether schemes that consume a small number of random digits must necessarily have bad worst case performance.

ACKNOWLEDGMENTS

I thank Fred Hickernell, Christiane Lemieux and Shu Tezuka for helpful comments.

REFERENCES

Cochran, W. G. 1977. *Sampling Techniques (3rd Ed)*. John Wiley & Sons.

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

de Boor, C. 1978. *A Practical Guide to Splines*. Springer, New York.

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

Faure, H. 2001. Variations on (0,s)-sequences. *Journal of Complexity 17*, 741–753.

Faure, H. and Tezuka, S. 2002a. Another random scrambling of digital (t,s)–sequences. In *Monte Carlo and Quasi-Monte Carlo Methods 2000*, K. T. Fang, F. J. Hickernell, and H. Niederreiter, Eds. Springer-Verlag, Berlin, 242–256.

Faure, H. and Tezuka, S. 2002b. I-binomial scrambling of digital nets and sequences. Tech. rep., IBM Tokyo Research Laboratory.

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

Friedel, I. and Keller, A. 2002. Fast generation of randomized low-discrepancy point sets. In *Monte Carlo and Quasi-Monte Carlo Methods 2000*, K. T. Fang, F. J. Hickernell, and H. Niederreiter, Eds. Springer-Verlag, Berlin, 257–273.

Haber, S. 1966. A modified Monte Carlo quadrature. *Mathematics of Computation 20*, 361–368.

Haber, S. 1967. A modified Monte Carlo quadrature, II. *Mathematics of Computation 21*, 388–397.

Hickernell, F. J. 1996a. The mean square discrepancy of randomized nets. *ACM Trans. Model. Comput. Simul. 6*, 274–296.

Hickernell, F. J. 1996b. Quadrature error bounds and figures of merit for quasi-random points. *SIAM Journal of Numerical Analysis 33*, 1995–2016. corrected printing of Sections 3-6 in ibid., **34** (1997), 853–866.

Hickernell, F. J. 1997. A generalized discrepancy and quadrature error bound. *Mathematics of Computation*. To appear.

Hickernell, F. J. 1998. Lattice rules: how well do they measure up? In *Random and Quasi-Random Point Sets*, P. Hellekalek and G. Larcher, Eds. Springer, New York, 109–168.

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

Hlawka, E. 1961. Funktionen von beschränkter Variation in der Theorie der Gleichverteilung. *Annali di Matematica Pura ed Applicata 54*, 325–333.

Hong, H. S. and Hickernell, F. J. 2000. Implementing scrambled digital sequences. Tech. rep., Hong Kong Baptist University.

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

Loh, W.-L. 2002. On the asymptotic distribution of scrambled net quadrature. Tech. rep., National University of Singapore, Dept of Statistics and Applied Probability.

Matousek, J. 1998a. *Geometric Discrepancy : An Illustrated Guide*. Springer-Verlag, Heidelberg.

Matousek, J. 1998b. On the L^2–discrepancy for anchored boxes. *Journal of Complexity 14*, 527–556.

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

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*, H. Niederreiter and P. J.-S. Shiue, Eds. Springer-Verlag, New York, 299–317.

Owen, A. B. 1997a. Monte Carlo variance of scrambled equidistribution quadrature. *SIAM Journal of Numerical Analysis 34,* 5, 1884–1910.

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

Owen, A. B. 1998a. Latin supercube sampling for very high dimensional simulations. *ACM Transactions on Modeling and Computer Simulation 8,* 2, 71–102.

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

Owen, A. B. 1999. Monte Carlo quasi-Monte Carlo and randomized quasi-Monte Carlo. In *Monte Carlo and quasi-Monte Carlo Methods 1998*, H. Niederreiter and J. Spanier, Eds. 86–97.

Roth, K. F. 1980. On irregularities of distribution IV. *Acta Arithmetica 37*, 67–75.

Tan, K. S. and Boyle, P. P. 2000. Applications of randomized low discrepancy sequences to the valuation of complex securities. *Journal of Economic Dynamics and Control 24*, 1747–1782.

Tezuka, S. 1995. *Uniform random numbers: theory and practice*. Kluwer Academic Publishers, Boston.

Tezuka, S. 2002. On randomization of generalized faure sequences. Tech. Rep. RT0494, IBM Tokyo Research Laboratory.

van der Corput, J. G. 1935a. Verteilungsfunktionen I. *Nederl. Akad. Wetensch. Proc. 38*, 813–821.

van der Corput, J. G. 1935b. Verteilungsfunktionen II. *Nederl. Akad. Wetensch. Proc. 38*, 1058–1066.

Yue, R. X. and Hickernell, F. J. 2002. The discrepancy and gain coefficients of scrambled digital nets. *Journal of Complexity 18*, 135–151.

Zaremba, S. K. 1968. Some applications of multidimensional integration by parts. *Annales Polonici Mathematici 21*, 85–96.