Proceedings of the 1998 Winter Simulation Conference
D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, eds.

MONTE CARLO EXTENSION OF QUASI-MONTE CARLO

Art B. Owen

Department of Statistics
Stanford University
Stanford CA 94305, U.S.A.

ABSTRACT

This paper surveys recent research on using Monte Carlo techniques to improve quasi-Monte Carlo techniques. Randomized quasi-Monte Carlo methods provide a basis for error estimation. They have, in the special case of scrambled nets, also been observed to improve accuracy. Finally through Latin supercube sampling it is possible to use Monte Carlo methods to extend quasi-Monte Carlo methods to higher dimensional problems.

1 INTRODUCTION

The problem we consider is the estimation of an integral

$$I = \int_{[0,1]^d} f(x)dx. \tag{1}$$

Standard manipulations can be applied to express integrals over domains other than the unit cube or with respect to nonuniform measures in the form (1). Similarly, the integrand f in (1) subsumes weighting functions from importance sampling or periodization. We are especially interested in cases where the dimension d is large, and some of the methods considered here apply to the case $d = \infty$.

The focus of this article is on ways of combining Monte Carlo and quasi-Monte Carlo solutions to this integration problem. Our goal is to provide readers with enough information to see what can be done and decide whether the approach is worthy of further investigation for their problems. For those readers who want to implement these constructions or to gain a full understanding of when and why the methods can work, there are references to the literature.

Section 2 describes the effect of the dimension d on the problem of computing (1). This section also presents an ANOVA decomposition of the integrand and some notions of the effective dimension of the integrand. Section 3 describes some simulation methods, Monte Carlo, quasi-Monte Carlo and hybrids thereof, that can be used on high dimensional integration problems. For very high dimensional problems, some of these methods lose effectiveness. Section 4 describes methods of using lower dimensional integration methods on higher dimensional problems. Brief conclusions are given in Section 5.

2 THE PROBLEM OF DIMENSION

This section presents working definitions of high and very high dimensional problems, taken from Owen (1998). The ANOVA decomposition is based on Owen (1992) and other references cited there. The definitions of effective dimension are from on Caflisch, Morokoff and Owen (1997).

2.1 High and very high dimensions

When $d = 1$, there are standard integration techniques that have very good accuracy when f is smooth. See Davis and Rabinowitz (1984). For small $d > 1$ iterated versions of such rules, based on Fubini's theorem, can be very effective. But for a rule with errors $O(n^{-r})$ in one dimension, the errors become $O(n^{-r/d})$ in d dimensions. A working definition of a high dimensional problem is one where iterated integrals are computationally infeasible or insufficiently accurate.

High dimensional problems are best handled by simulation methods, including Monte Carlo and quasi-Monte Carlo (equidistribution). These are reviewed in Section 3.

In sufficiently large dimensions it becomes difficult to even construct quasi-Monte Carlo point sets with meaningful equidistribution properties. For example, some constructions are not especially equidistributed unless $n = O(d^2)$ which can be too large. A working definition

of a very high dimensional problem is one where iterated integrals are computationally infeasible or insufficiently accurate. Monte Carlo is still available for such problems, as is Latin hypercube sampling (Section 4.1). Latin Supercube sampling (Section 4.4) is directed at extending quasi-Monte Carlo into very high dimensional problems.

2.2 ANOVA Decomposition

Let $\mathcal{A} = \{1, 2, \ldots, d\}$ denote the set of input variables to the function f on $[0,1]^d$. We can write f as a sum of 2^d functions, one for each subset of $\mathcal{A}$, with that function only depending on the variables in its subset. That is

$$f(x) = \sum_{u \subseteq \mathcal{A}} f_u(x), \tag{2}$$

where $f_u(x)$ only depends on those components of x whose indices are in u.

For any choices of f_u with $u \neq \mathcal{A}$, we can make (2) hold by choosing $f_{\mathcal{A}}$ by subtraction. For example, suppose

$$f(x) = 30 + 20x^1 + 10x^2 - 16x^1x^2$$

where $x = (x^1, x^2) \in [0,1]^2$. This can be rewritten as

$$f(x) = f_\emptyset(x) + f_{\{1\}}(x) + f_{\{2\}}(x) + f_{\{1,2\}}(x)$$

where $f_\emptyset(x) = 41$, $f_{\{1\}}(x) = 12(x^1 - 0.5)$, $f_{\{2\}}(x) = 2(x^2 - 0.5)$, and $f_{\{1,2\}}(x) = -16(x^1 - 0.5)(x^2 - 0.5)$. Notice that $\int_0^1 f_{\{1\}}(x^1)dx^1 = 0$. This is reasonable; had $f_{\{1\}}$ integrated to some other constant we could have added that constant to $f_\emptyset$ and subtracted it from $f_{\{1\}}$. More generally, when some structure can be attributed to either f_u or f_v with $v \subset u$ we prefer on grounds of parsimony to attribute it to f_v.

A particularly useful choice for the f_u is based on the analysis of variance (ANOVA) decomposition from statistical experimental design. (Montgomery (1984) is a standard reference on design.) In concept, one simply embeds an equispaced q^d grid in $[0,1]^d$, defines the main effects and interactions on this grid, and then lets $q \to \infty$ replacing sums by integrals.

We employ the following notation: $|u|$ is the cardinality of u, x^u denotes the $|u|$-tuple consisting of components x^j with $j \in u$, and $-u$ is the complement of u in $\mathcal{A}$. In the function setting we let

$$f_u(x) = \int_{z : z^u = x^u} \left(f(z) - \sum_{v \subset u} f_v(z) \right) dz^{-u} \tag{3}$$

where the sum in (3) is over strict subsets $v \neq u$. Equation (3) defines f_u by subtracting what can be attributed to subsets of u, and then averaging over all components not in u. Using a natural convention $f_\emptyset(x) = I$, and another convention gives $f_{\mathcal{A}}(x) = f(x) - \sum_{|u|<d} f_u(x)$.

The function $f_u(x)$ only depends on x^u. When it is desired to emphasize this point, we write $f_u(x^u)$. Formally, $f_u(x^u) = f_u(z)$ at any point z for which $z^u = x^u$. The value of z^{-u} does not enter. For $u = \emptyset$ we may write $f_\emptyset$ without an argument x, since the function is constant.

Let $\sigma^2 = \int (f(x) - I)^2 dx$ and suppose that $\sigma^2 < \infty$. Now let $\sigma_u^2 = \int f_u(x)^2 dx$ for $|u| > 0$ and $\sigma_\emptyset^2 = 0$. Then

$$\sigma^2 = \sum_{u \subseteq \mathcal{A}} \sigma_u^2. \tag{4}$$

Equation (4) partitions the variance of f into parts corresponding to each subset $u \subseteq \mathcal{A}$. The f_u enjoy some other easily verified properties: if $j \in u$, then the line integral $\int_0^1 f_u(x)dx^j = 0$, for any values of x^k with $k \neq j$, and if $u \neq v$ then $\int f_u(x) f_v(x) dx = 0$.

2.3 Effective Dimensions

The ANOVA decomposition can be used to consider notions of the "effective" dimension of an integrand. For example, because an additive integrand

$$f(x) = f_\emptyset + f_{\{1\}}(x^1) + \cdots + f_{\{d\}}(x^d) \tag{5}$$

is a sum of one dimensional integrands it can be much easier to integrate than a general d dimensional integrand. In many application areas, additive integrands are very unlikely.

Nearly additive integrands may however be common in some application areas. Caflisch, Morokoff and Owen (1997) found that a 360 dimensional integrand motivated by a problem in computational finance was very nearly additive. They then defined two notions of effective dimension using the ANOVA decomposition.

Definition 1 *The effective dimension of f, in the superposition sense, is the smallest integer d_S such that $\sum_{0<|u|\le d_S} \sigma_u^2 \ge 0.99\sigma^2$.*

Definition 2 *The effective dimension of f, in the truncation sense, is the smallest integer d_T such that $\sum_{u \subseteq \{1,2,\ldots,d_T\}} \sigma_u^2 \ge 0.99\sigma^2$.*

The truncation definition reflects that for some integrands, only a small number of the inputs might really matter. The superposition definition reflects that for some integrands, the inputs might only influence the outcome through their joint action within small groups. For example, an additive function has superposition dimension 1 and a quadratic function has superposition dimension at most 2, but either could have truncation definition d.

Clearly the threshold 0.99 is an arbitrary choice and other values could be used. It is immediate that $d_S \le d_T \le d$. The value of d_T depends on the order in which input variables are indexed. If one has subject matter knowledge about which variables are most important, then one would first order the variables in decreasing order of importance before applying Definition 2.

3 SIMULATION METHODS

The simulation methods we consider here are all of the form

$$\hat{I} = \frac{1}{n}\sum_{i=1}^{n} f(X_i) \tag{6}$$

where $X_i \in [0,1]^d$. Using the ANOVA decomposition (2) and noting that $f_\emptyset = I$, we find that

$$\hat{I} - I = \sum_{|u|>0} \hat{I}_u \tag{7}$$

where for $|u| > 0$,

$$\hat{I}_u = \frac{1}{n}\sum_{i=1}^{n} f_u(X_i) = \frac{1}{n}\sum_{i=1}^{n} f_u(X_i^u) \tag{8}$$

is the error in the estimate of $I_u = \int f_u(x)dx = 0$. From (8) we see that the contribution of $\hat{I}_u$ to the error comes from the $|u|$-dimensional projected quadrature rule X_i^u, $i = 1, \ldots, n$.

3.1 Monte Carlo

Simple Monte Carlo samples X_i independently from the $U[0,1]^d$ distribution and then applies the estimate (6). It is well known that $\hat{I} - I$ has mean zero and variance σ^2/n, so that the Monte Carlo errors are of order $n^{-1/2}$ in probability. Variance reduction techniques like stratification and control variates (with finitely many strata or variates) and importance sampling, do not affect this rate, though they may improve the constant.

The dimension d does not appear in this rate. This means that the effectiveness of Monte Carlo is independent of the dimension, unless one is considering a dimension indexed sequence of functions for which σ^2 has a dimension effect.

A third important feature of Monte Carlo sampling is that error estimation is comparatively easy. An unbiased estimate of $\mathrm{var}(\hat{I})$ is s^2/n where $s^2 = (1/(n-1))\sum_{i=1}^{n}(f(X_i) - \hat{I})^2$. This estimate is available from the same data used to construct $\hat{I}$.

3.2 Quasi-Monte Carlo

The aim in quasi-Monte Carlo integration is to choose X_i without the usual clusters and gaps seen in Monte Carlo sampling. The reader unfamiliar with quasi-Monte Carlo may consult Niederreiter (1992) for more information about this topic, including background on the results cited here. In particular, we refer below to (t, m, s)-nets and (t, s)-sequences. These are quasi-Monte Carlo point sets defined and discussed in Niederreiter (1992).

If the integrand has bounded variation in the sense of Hardy and Krause, then it is possible to find a deterministic sequence X_i, $i \ge 1$ along which

$$|\hat{I} - I| = O(n^{-1}(\log n)^d). \tag{9}$$

If we do not require the n point integration rule to include the points of the $n-1$ point integration rule, then it is possible to reduce the exponent of $\log n$ to $d-1$.

The rate in (9) is asymptotically superior to the rate $n^{-1/2}$ that characterizes Monte Carlo. In high dimensions, the rate (9) does not set in until n is large. One simple observation is that the error bound increases with n until n equals $\exp(d)$. Thus the smallest n for which (9) could be relevant is likely to be at least $\exp(d)$.

Morokoff and Caflisch (1995) have reported that QMC methods usually beat MC in practice although the advantage usually disappears by about $d = 8$. Paskov and Traub (1995) by contrast found that QMC was very effective on some integrands with $d = 360$. Caflisch, Morokoff and Owen (1997) suggested that QMC was superior to MC if the effective dimension of the integrand was not large.

Accuracy considerations favor QMC over MC. QMC has superior asymptotic accuracy, and in examples it usually has better small sample accuracy. The main practical disadvantage of QMC with respect to MC is that there is no way to estimate the accuracy achieved from the sample values. The constant implicit in (9) is the total variation of f in the sense of Hardy and Krause. There appear to be no good ways to estimate that quantity, and in any case, it can be a gross upper bound on the error.

3.3 Randomized Quasi-Monte Carlo

Randomized QMC methods have long been used to provide a basis for error estimation in QMC methods. Owen (1995) surveys the use of such methods.

Here is a generic recipe for randomizing QMC methods. Let $A_1, \ldots, A_n$ be a QMC point set. Let X_i be a randomized version of A_i. The randomization should have the following properties:

RQMC-1 $X_i \sim U[0,1]^d$, $i = 1, \ldots, n$,

RQMC-2 $X_1, \ldots, X_n$ is a QMC point set with probability 1.

Property RQMC-1 makes the estimator (6) an unbiased estimate of I. Property RQMC-2 simply means that the randomization has preserved whatever special properties the underlying QMC point set had. The examples below illustrate this.

Space does not allow a detailed description of QMC points and their randomizations, but we discuss two examples briefly. Further details can be found in Owen (1998).

Cranley and Patterson (1976) describe a form of random rotation in which

$$X_i^j = A_i^j + U^j \bmod 1. \tag{10}$$

Here U^j are independent $U[0,1]$ random variables, $j = 1,\dots,d$ and $z \bmod 1$ means $z - \lfloor z \rfloor$ where $\lfloor z \rfloor$ is the greatest integer less than or equal to z. Certain lattice rules (Sloan and Joe (1994)) have a structure that makes them very accurate for periodic functions whose Fourier coefficients decay rapidly. The rules are still accurate after rotation.

Some QMC methods, known as (t,m,s)-nets and (t,s)-sequences, construct points A_i so that certain hyperrectangles obtain a number of sample points proportional to their volumes. The hyperrectangles involved have coordinates that are integers divided by powers of an integer base $b \geq 2$. For such A_i it is possible to apply random permutations to their digits in base b in a way that preserves their net properties and renders the resulting X_i uniformly distributed. See Owen (1995,1997a) for details. Owen (1995) surveys earlier work on randomizing digits.

In one special case, the randomization of a QMC point set can be shown to enhance the accuracy of the integration rule. The explanation is that randomization leads to cancellation of some error components. Owen (1997b) shows that scrambled nets can lead to errors of size $n^{-3/2}(\log n)^{(d-1)/2}$ in probability. The integrand must have greater smoothness than bounded variation: a sufficient condition is that $\partial^d f(X)/\prod_{j=1}^d \partial X^j$ obey a Lipschitz condition (Owen 1997b). Hickernell (1996) shows that this randomization can improve the equidistribution of nets.

3.4 Using randomized QMC rules

In practice one can take a small number r of independent replicates of QMC points. The corresponding estimates $\hat{I}_1,\dots,\hat{I}_r$ are unbiased estimates of I with common variance σ^2_{RQMC}. The pooled estimate $\hat{I} = (1/r)\sum_{k=1}^r \hat{I}_k$ has variance $\sigma^2_{\mathrm{RQMC}}/r$, and an unbiased estimate of this variance is

$$\frac{1}{r(r-1)}\sum_{k=1}^r (\hat{I}_k - \hat{I})^2.$$

Taking a large value of r makes for a more accurate variance estimate. But for a given number of function evaluations $N = nr$ one can usually expect that a larger value of n with a smaller value of r should give better accuracy in $\hat{I}$. At the extreme, taking $n = 1$ and $r = N$ simply reproduces Monte Carlo estimation.

For scrambled nets, Owen (1997a) describes a form of internal replication in which consecutive blocks of observations can be treated as replicates. The cost of this, compared to genuine replication, is usually an upward bias in the estimated variance, while the gain is usually greater accuracy in $\hat{I}$.

4 VERY HIGH DIMENSIONS

For high enough dimensions it can be difficult to construct QMC point sets with meaningful QMC properties and reasonably small values of n. To illustrate the difficulties, consider $(0,d)$-sequences in base b. These only exist for $b \geq d$. To have all hyperrectangles of size $1/b \times 1/b \times 1 \cdots \times 1$ get n/b^2 points each, takes at least $b^2 \geq d^2$ points. If one takes such a large n then every one of the $d(d-1)/2$ bivariate projections of the X_i will be equidistributed. But taking such a large n is costly if d is large.

4.1 Latin Hypercube Sampling

Latin hypercube sampling (LHS) is a form of simultaneous stratification on all d dimensions. McKay, Beckman and Conover (1979) introduced a version of LHS for computer experiments. Let

$$X_i^j = \frac{\pi_j(i) - U_i^j}{n} \tag{11}$$

where π_j are uniform random permutations of the integers $1,\dots,n$, the U_i^j are $U[0,1]$ random variables, and all of the π_j and U_i^j are independent.

An older version of LHS, due to Patterson (1954) has

$$X_i^j = \frac{\pi_j(i) - 0.5}{n}. \tag{12}$$

In either (11) or (12), for each input $j = 1,\dots,d$ and every interval of the form $((m-1)/n, m/n)$ for $m = 1,\dots,n$, there is one observation X_i^j in that interval.

Latin hypercube sampling can be used in any dimension d, even $d > n$. Because LHS stratifies each input variable individually, it is able to integrate near additive functions with great accuracy. Stein (1987) shows that

$$\mathrm{var}_{\mathrm{LHS}}(\hat{I}) = \frac{1}{n}\sum_{|u|>1}\sigma_u^2 + o\left(\frac{1}{n}\right). \tag{13}$$

The additive portion of the integrand f does not contribute to the asymptotic variance under LHS.

In finite samples, LHS is never much worse than MC. Owen (1997a) shows that

$$\mathrm{var}_{\mathrm{LHS}}(\hat{I}) \le \frac{1}{n-1}\sigma^2. \tag{14}$$

Owen (1998) shows that LHS can also work for $d = \infty$. As in the finite d setting, the additive part of f does not contribute to the asymptotic variance.

4.2 Padding

If an integrand f has nominal dimension d but is nearly of dimension $s \ll d$ then one can employ an s dimensional QMC or RQMC rule on the important input variables and use something else for the others.

If one thinks of the n by d matrix of X_i^j values as the input to the simulation, then the leftmost s columns can be filled with (R)QMC points. Something has to go in the other $d-s$ columns. One might simply replace all of those values by the central value 0.5. But, this does not make for an unbiased estimate of I and so it can be hard to estimate the error in the resulting estimate.

As an alternative, one can simply pad out the matrix by filling in a lower quality rule. For particle transport problems, Spanier (1995) and Okten (1996) suggest filling out the remaining columns with simple Monte Carlo points. Owen (1994) considers filling out the remaining columns with a Latin hypercube sample. If one has used an RQMC rule on the first s dimensions and MC or LHS padding, then it becomes possible to estimate the variance of $\hat{I}$. A further benefit is that if one has guessed incorrectly, so that some of the variables thought to be unimportant really are important, then the padding procedure can lose much less accuracy than one would lose by filling in 0.5's.

Suppose that $\mathcal{A}_1 = \{1, 2, \ldots, s\} \subset \mathcal{A}$ is the set of all input variables handled by an RQMC rule. It follows from Theorem 1 of Owen (1998) that the resulting variance of $\hat{I}$ is, under mild conditions,

$$\frac{1}{n}\left(\sigma^2 - \sum_{u \subseteq \mathcal{A}_1} \sigma_u^2\right) + o\left(\frac{1}{n}\right)$$

under padding by Monte Carlo, and

$$\frac{1}{n}\left(\sigma^2 - \sum_{u \subseteq \mathcal{A}_1} \sigma_u^2 - \sum_{j=s+1}^{d} \sigma_{\{j\}}^2\right) + o\left(\frac{1}{n}\right)$$

under padding by LHS. For these results the ANOVA components of the integrand must enjoy some extra smoothness that the underlying RQMC method requires. These results say that one gets what one pays for: the RQMC rule eventually balances out the errors in f_u for $u \subseteq \mathcal{A}_1$, padding by LHS balances out some additional additive components.

4.3 Engineering the Integrand

It is often possible to re-arrange a simulation so as to reduce the effective dimension. For example, simulations driven by Brownian motion sampled at d time points, may be generated in any order whatsoever, not just in order of time sequence. Caflisch, Morokoff and Owen (1997) generate the end point first then fill in the midpoint, quarter-points and so forth of the Brownian motion, in each case sampling the new point from its conditional distribution given the existing points. This process replaces the integrand f by another one with the same value of I, the same variance σ^2 and the same nominal dimension d. The new integrand was more strongly dominated by the first few steps and this reduced the effective dimension. They were able to employ LHS padding with scrambled nets, and also to use Sobol' sequences to good effect.

Similarly, Acworth, Broadie and Glasserman (1997) used the principal components of Brownian motion to reduce effective dimension. Fox (1996) discusses several ways to reduce effective dimension in discrete event simulation.

4.4 Latin Supercube Sampling

Given a well engineered integrand, RQMC with padding is able to reduce the variance substantially. But it only reaps the benefits of RQMC for ANOVA effects within the set $\mathcal{A}_1$ of variables balanced by the simulation.

In some cases, one would like to be able to obtain the benefits of RQMC balance within several groups of variables. Of course there are restrictions on what one can obtain. A good fully d dimensional set of QMC points may not exist for reasonable n.

The idea of Latin Supercube Sampling (LSS) is to use (R)QMC within multiple groups of input variables. In a simulation driven by several Brownian motions, there might be one group of variables for each Brownian motion. In a simulation that follows a sequence of particle collisions, there might be one group of variables for each collision.

Suppose for example that $d = ks$ and that $\mathcal{A}_1 = \{1, 2, \ldots, s\}$, $\mathcal{A}_2 = \{s+1, s+2, \ldots, 2s\}$, and so on until $\mathcal{A}_k = \{(k-1)s+1, (k-1)s+2, \ldots, ks\}$.

Suppose further that $\mathcal{X}_i^j \in [0,1]^s$, $i = 1, \ldots, n$ is an s dimensional (R)QMC point set for each $j = 1, \ldots, k$. Then LSS takes points

$$X_i = (\mathcal{X}^1_{\pi_1(i)}, \mathcal{X}^2_{\pi_2(i)}, \ldots, \mathcal{X}^k_{\pi_k(i)}) \in [0,1]^d, i = 1, \ldots, n,$$

where the π_j are independent uniform random permutations of $1, 2, \cdots, n$. More generally, there is no need for the $\mathcal{A}_j$

to all contain equal numbers of input variables. It must be true that $\mathcal{A}_j \cap \mathcal{A}_l = \emptyset$ whenever $j \neq l$.

For large n, the points X_i are well balanced in all coordinate projections X^u, where $u \subseteq \mathcal{A}_j$ for some j, but are not especially well balanced in any such projections where u has a nonempty intersection with two or more $\mathcal{A}_j$. Accordingly, from (8) we expect that for large n the variance of $\hat{I}$ should be

$$\frac{1}{n}\left(\sigma^2 - \sum_{j=1}^{k}\sum_{u\subseteq\mathcal{A}_j}\sigma_u^2\right) + o\left(\frac{1}{n}\right). \tag{15}$$

At least (8) suggests that f_u for $u \subseteq \mathcal{A}_j$ should not contribute to var$(\hat{I})$ under LSS. Under mild conditions, Owen (1998) shows that (15) holds.

The implication is as follows. If the integrand has, or can be engineered to have, almost all of its ANOVA variance contained within subsets $\mathcal{A}_j$ of input variables, then an enormous variance reduction can be obtained. In the extreme, one gets an s dimensional error rate for a d dimensional problem. In the worst case though, it may be true that almost none of the variance derives from σ_u^2 with $u \subseteq \mathcal{A}_j$. In this event, LSS fails softly, giving a variance that is essentially the same as simple MC would have given.

5 CONCLUSIONS

By employing randomness, it is possible to improve QMC in three ways. First, by replication, one can get sample based error estimates. Second, for the case of scrambled nets, one can introduce cancellation that improves accuracy. Third, by Latin supercube sampling, one can employ low dimensional rules on high dimensional problems.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation.

REFERENCES

Acworth, P., M. Broadie, and P. Glasserman. 1997. A comparison of some Monte Carlo techniques for option pricing, in *Monte Carlo and quasi-Monte Carlo methods '96*, H. Niederreiter, P. Hellekalek, G. Larcher and P. Zinterhof, ed.s. New York: Springer.

Caflisch, R., W. Morokoff, and A. Owen. 1997. Valuation of Mortgage Backed Securities using Brownian Bridges to Reduce Effective Dimension. *Journal of Computational Finance*. 1: 27–46.

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

Davis,P. and P. Rabinowitz. 1984. *Methods of Numerical Integration (2nd Ed.)*. Academic Press.

Fox, B. 1996. *Generating Poisson processes by quasi-Monte Carlo*. Boulder CO: SIM-OPT Consulting.

Hickernell, F. 1996. The Mean Square Discrepancy of Randomized Nets. *ACM Transactions on Modeling and Computer Simulation*. 6: 274–296.

McKay, M., R. Beckman and W. Conover. 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 21(2): 239–45.

Montgomery, D. 1984. *Design and Analysis of Experiments (Second Edition)*. John Wiley & Sons.

Morokoff, W. and R. Caflisch. 1995. Quasi-Monte Carlo Integration. *J. Comp. Phys.* 122: 218–230.

Niederreiter, H. 1992. *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia: SIAM.

Okten, G. 1996. *A probabilistic result on the discrepancy of a hybrid-Monte Carlo sequence and applications*. Technical report, Mathematics Department, Claremont Graduate School.

Owen, A. 1992. Orthogonal arrays for computer experiments, integration and visualization. *Statistica Sinica*. 2:439–452.

Owen, A. 1994. Lattice sampling revisited: Monte Carlo variance of means over randomized orthogonal arrays. *The Annals of Statistics*. 22:930–945.

Owen, A. 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. New York: Springer-Verlag.

Owen, A. 1997a. Monte Carlo Variance of Scrambled Equidistribution Quadrature. *SIAM Journal of Numerical Analysis*. 34(5):1884–1910.

Owen, A. 1997b. Scrambled Net Variance for Integrals of Smooth Functions. *The Annals of Statistics* 25(4):1541–1562.

Owen, A. 1998. Latin supercube sampling for very high dimensional simulations. *ACM Transactions on Modeling and Computer Simulation*. (To appear)

Paskov, S. and J. Traub. 1995. Faster Valuation of Financial Derivatives. *The Journal of Portfolio Management* Pp. 113–120.

Patterson, H. D. 1954. The errors of lattice sampling. *Journal of the Royal Statistical Society, Series B*, 16:140–149.

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

Spanier, J. 1995. Quasi-Monte Carlo Methods for Particle Transport Problems, in *Monte Carlo and Quasi-Monte*

Carlo Methods in Scientific Computing, H. Niederreiter and P. J.-S. Shiue ed.s, pp. 121–148. New York: Springer-Verlag.

Stein, M. 1987. Large Sample Properties of Simulations Using Latin Hypercube Sampling *Technometrics* 29(2):143–151.

AUTHOR BIOGRAPHY

ART B. OWEN is a professor in the Department of Statistics at Stanford University. He received a B. Math. degree from the University of Waterloo and a Ph.D. degree in Statistics from Stanford University. He is a member of the A.S.A. and a fellow of the I.M.S. His research interests include computationally intensive statistical methods applied to high dimensional numerical problems and computational finance.