Variance Reduction Order Using Good Lattice Points in Monte Carlo Methods

Bruno Tuffin
Institut Mathmatique de Rennes
Universit de Rennes 1
Campus universitaire de Beaulieu
35042 Rennes Cedex
France
e-mail: tuffin@maths.univ-rennes1.fr

Abstract

Quasi-Monte Carlo methods and lattice rules with good lattice points give rapidly "good" approximations for numerical integration, but the error estimation is intractable in practice. In the literature, a randomization of these methods, using a combination of Monte Carlo and quasi-Monte Carlo methods, has been done to obtain a confidence interval using the Central Limit Theorem. In this paper we show that for a special class of functions with small Fourier coefficients and using good lattice points, the decreasing of the variance of the combined estimator is faster than the usual one.

Key words: Monte Carlo, Quasi-Monte Carlo, Numerical Integration, Lattice points.
AMS classification: 65C05; 11K36.

1 Introduction

Among approximation methods of multiple integrals $I(f) = \int_{[0,1]^s} f(x)dx$, Monte Carlo methods [3] and quasi-Monte Carlo methods have been extensively studied [7, 2]. Quasi-Monte Carlo methods are the ones for which the convergence speed is the fastest. They use low discrepancy sequences $(\xi^{(n)})_{n\in\mathbb{N}}$ [7] or lattice rules including good lattice points $(\{ng/N\})_{0\leq n\leq N-1}$ [7, 10]. These techniques present the advantage, with respect to Monte Carlo ones, of using a sequence of "well-distributed" points over $[0,1]^s$. Nevertheless, the error estimation is in general difficult to obtain in practice. A solution to this problem is to combine Monte Carlo and quasi-Monte Carlo methods, considering the random variable

$$Z_N = \frac{1}{N}\sum_{n=0}^{N-1} f(\{X + \xi^{(n)}\}) \tag{1}$$

where X is a random vector uniformly distributed over $[0,1]^s$ and $\{x\}$ denotes the fractional part of each coordinate of $x \in \mathbb{R}^s$. If we consider a sample of size I issued from this random variable and if we estimate its variance, we can use the central limit Theorem to obtain a confidence interval. This has been done first in [9] using low discrepancy sequences, in [1] using good lattice points and in [5] using lattice rules, but without giving any information about the variance reduction with respect to the variance of the mean of NI i.i.d random variable with the same law than $f(X)$. In [13] and more precisely in [12], it is shown that, under some restrictions on f, $\sigma^2(Z_N) = O(N^{-2}(\ln N)^{2s})$, so that the global convergence speed is $O(I^{-1/2}N^{-1}(\ln N)^s)$. Moreover efficient applications of this technique can be found in [13, 12, 11].

In this paper we show that, under more restrictive properties of the function f, i.e. $f \in E_\alpha^s(C)$ where $E_\alpha^s(C)$ with $\alpha > 1$ is defined is the next Section, and using the good lattice points, we have

$$\sigma^2\left(\frac{1}{N}\sum_{n=0}^{N-1} f\left(\left\{\frac{n}{N}g + X\right\}\right)\right) = O(N^{-2\alpha}(\ln N)^{2\alpha(s-1)+1}(\ln\ln N)^{(2\alpha-1)(s-1)}).$$

The paper is organized as follows. Section 2 reviews the major results on good lattice points. Section 3 is devoted to the results on the variance reduction order of the combined method and numerical illustrations can be found in Section 4. Finally we conclude and present directions for future research in Section 5.

2 Good lattice rules

Let $f : \mathbb{R}^s \to \mathbb{R}$ be a periodic function with period one in each coordinate and, for $y \in \mathbb{R}$, define $e(y) = e^{2\pi i y}$.

For $h \in \mathbb{Z}^s$, let

$$\hat{f}(h) = \int_{[0,1]^s} f(x)e(-h.x)dx$$

be the Fourier coefficient of rank h of function f, where $x.y$ is the standard innner product of $x, y \in \mathbb{R}^s$. Suppose that we have

$$f(x) = \sum_{h\in\mathbb{Z}^s} \hat{f}(h)e(h.x).$$

Moreover, for $h = (h_1, \cdots, h_s) \in \mathbb{Z}^s$, let

$$r(h) = \prod_{i=1}^{s} \max(1, |h_i|).$$

Define now a specific class of functions $E_\alpha^s(C)$ with respect to some properties of the Fourier coefficients.

Definition 1 *Let $\alpha > 1$ and $C > 0$ be real numbers. Then $E_\alpha^s(C)$ is the class of all periodic functions $f\ : \mathbb{R}^s \to \mathbb{R}$ with period one in each coordinate and such that*

$$|\hat{f}(h)| \leq C r(h)^{-\alpha} \text{ for all nonzero } h \in \mathbb{Z}^s.$$

It is easily seen that if $f \in E_\alpha^s(C)$, then its Fourier series is absolutely convergent and represents f [7].

It is shown in [14] that a sufficient condition for f to be in $E_\alpha^s(C)$, with $\alpha \in \mathbb{N}^*$ and with a value of C which can be given explicitly, is the following: all partial derivatives of form $\partial^{m_1+\cdots+m_s} f/(\partial x_1^{m_1} \cdots \partial x_s^{m_s})$ with $0 \leq m_i \leq \alpha - 1$ for $1 \leq i \leq s$ exist and are of bounded variation on $[0,1]^s$ in the sense of Hardy and Krause (see [2] for a definition).

Let $S = (\xi^{(n)})_{0 \leq n \leq N-1}$ be an integration lattice rule [7, chapter 5]. Define also $S^\perp$ and $P_\alpha(S)$ as

$$S^\perp = \{h \in \mathbb{Z}^s | h.\xi \in \mathbb{Z} \text{ for all } \xi \in S\}$$

and

$$P_\alpha(S) = \sum_{h \in S^\perp}{}' \, r(h)^{-\alpha}$$

where the prime indicates that the $h = 0$ term is to be omitted from the sum.

We have then [7], for any real numbers $\alpha > 1$ and $C > 0$, for any integer $N \geq 1$ and any lattice rule $S = (\xi^{(n)})_{0 \leq n \leq N-1}$,

$$\max_{f \in E_\alpha^s(C)} \left| \frac{1}{N} \sum_{n=0}^{N-1} f\left(\xi^{(n)}\right) - \int_{[0,1]^s} f(x)dx \right| = C P_\alpha(S).$$

The problem is then to find good lattice rules. The special case of good lattice points use a vector $g = (g_1, \cdots, g_s) \in \mathbb{Z}^s$, so that $S = (\{ng/N\})_{0 \leq n \leq N-1}$ and $S^\perp = \{h \in \mathbb{Z}^s | h.g \equiv 0 \bmod N\}$. In this case, we note $P_\alpha(g, N)$ instead of $P_\alpha(S)$. As we have assumed that f is periodic with period one in each coordinate, then $f(\{ng/N\}) = f(ng/N)$, so we can remove the braces.

Using Theorem 1 of [8] and a remark on p. 64, there is a constant $c' = c'(C, \alpha, s)$ such that for all $s \geq 2$ and $N \geq 2$ there exists a good lattice point g with $-N/2 < g_i \leq N/2$ for $1 \leq i \leq s$ so that $\forall f \in E_\alpha^s(C)$

$$\left| \frac{1}{N} \sum_{n=0}^{N-1} f\left(\frac{n}{N} g\right) - \int_{[0,1]^s} f(x)dx \right| \leq C P_\alpha(g, N) \leq c' \frac{(\ln N)^{\alpha(s-1)+1} (\ln \ln N)^{(\alpha-1)(s-1)}}{N^\alpha}.$$

Methods to compute good lattice points can be found in [10, 6] but are expensive in computational time. Tables for dimensions $s \leq 12$ of precalculated points can be found in [4]. Otherwise, a faster method to generate points, which are then less efficient but with sufficiently good properties nevertheless, can be found in [15].

3 Randomization and order of the variance

As we have remarked in the introduction, the error estimation is an important problem in the application of quasi-Monte Carlo methods. The use of the combined method seems then to be a good alternative. We show here that for $f \in E_\alpha^s(C)$, and using good lattice points, the variance of the random variable Z_N defined by relation (1) is in $O(N^{-2\alpha}(\ln N)^{2\alpha(s-1)+1}(\ln\ln N)^{(2\alpha-1)(s-1)})$, then better than the $O(N^{-2}(\ln N)^{2s})$ of [13, 12] and much better than the $O(N^{-1})$ of a standard Monte Carlo method.

We begin by a preliminary and key Theorem.

Theorem 1 *Let $f : \mathbb{R}^s \to \mathbb{R}$ be such that its Fourier series $\sum_{h\in\mathbb{Z}^s} \hat{f}(h)e(h.u)$ is absolutely convergent to f. If X is a random vector uniformly distributed on $[0,1]^s$, we have then*

$$\sigma^2\left(\frac{1}{N}\sum_{n=0}^{N-1} f(\{X+\xi^{(n)}\})\right) = \sum_{h\in\mathbb{Z}^s}{}' |\hat{f}(h)|^2 \left|\frac{1}{N}\sum_{n=0}^{N-1} e(h.\xi^{(n)})\right|^2. \qquad (2)$$

Proof: As a first part of the proof,

$$E\left(\frac{1}{N}\sum_{n=0}^{N-1} f(\{X+\xi^{(n)}\})\right) = \sum_{h\in\mathbb{Z}^s} \hat{f}(h)E\left(\frac{1}{N}\sum_{n=0}^{N-1} e(h.(X+\xi^{(n)}))\right) = \hat{f}(0) = I(f)$$

where $I(f)$ is the integral of f over $[0,1]^s$. Thus, if $Z_N = (1/N)\sum_{n=0}^{N-1} f(\{X+\xi^{(n)}\})$,

$$\sigma^2(Z_N) = \sum_{h,k\in\mathbb{Z}^s} \hat{f}(h)\hat{f}(k)\frac{1}{N^2}\sum_{m,n=0}^{N-1} E(e(h.(X+\xi^{(n)}))e(k.(X+\xi^{(m)}))) - (\hat{f}(0))^2.$$

But

$$E(e(h.(X+\xi^{(n)}))e(k.(X+\xi^{(m)}))) = e(h.\xi^{(n)})e(k.\xi^{(m)})E(e((h+k).X))$$

and

$$E(e((h+k).X)) = \begin{cases} 1 & \text{if } h+k=0 \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, f is a real function, so $\hat{f}(-h) = \overline{\hat{f}(h)}$, where $\overline{z}$ is the complex conjugate of z. We have then

$$\begin{aligned} \sigma^2\left(\frac{1}{N}\sum_{n=0}^{N-1} f(\{X+\xi^{(n)}\})\right) &= \sum_{h\neq 0} |\hat{f}(h)|^2 \left(\frac{1}{N^2}\sum_{m,n=0}^{N-1} e(h.(\xi^{(n)}-\xi^{(m)}))\right) \\ &= \sum_{h\neq 0} |\hat{f}(h)|^2 \left|\frac{1}{N}\sum_{n=0}^{N-1} e(h.\xi^{(n)})\right|^2. \end{aligned}$$

■

Note that every function f in $E_\alpha^s(C)$ verifies the absolute convergence of its Fourier series. Theorem 1 allows to show the following result:

Theorem 2 *Let X be a random vector uniformly distributed over $[0,1]^s$. For any real numbers $\alpha > 1$ and $C > 0$, for any integer $N \geq 1$ and s-dimensional lattice rule $S = (\xi^{(n)})_{0 \leq n \leq N-1}$, we have*

$$\max_{f \in E_\alpha^s(C)} \sigma^2 \left(\frac{1}{N} \sum_{n=0}^{N-1} f\left(\left\{ \xi^{(n)} + X \right\} \right) \right) = C^2 P_{2\alpha}(S).$$

Proof: First, let us remember that $S^\perp = \{h \in \mathbb{Z}^s | h.\xi \in \mathbb{Z} \text{ for all } \xi \in S\}$. From (2), remarking that [7, Lemma 5.21], $\forall h \in \mathbb{Z}^s$,

$$\frac{1}{N} \sum_{n=0}^{N-1} e\left(h.\xi^{(n)}\right) = \begin{cases} 1 & \text{if } h \in S^\perp \\ 0 & \text{otherwise,} \end{cases}$$

we have

$$\sigma^2 \left(\frac{1}{N} \sum_{n=0}^{N-1} f\left(\left\{ \xi^{(n)} + X \right\} \right) \right) = \sum_{h \in S^\perp}{}' |\hat{f}(h)|^2.$$

Then for every function $f \in E_\alpha^s(C)$,

$$\sigma^2 \left(\frac{1}{N} \sum_{n=0}^{N-1} f\left(\left\{ \xi^{(n)} + X \right\} \right) \right) \leq \sum_{h \in S^\perp}{}' C^2 r(h)^{-2\alpha} = C^2 P_{2\alpha}(S).$$

Now define the function f_0 by $f_0(x) = C \sum_{h \in \mathbb{Z}^s} r(h)^{-\alpha} e(h.x)$ $\forall x \in \mathbb{R}^s$. Then $f_0 \in E_\alpha^s(C)$ and

$$\sigma^2 \left(\frac{1}{N} \sum_{n=0}^{N-1} f_0\left(\left\{ \xi^{(n)} + X \right\} \right) \right) = C^2 P_{2\alpha}(S),$$

thus we obtain the result. ■

Theorem 3 *Let X be a random vector uniformly distributed over $[0,1]^s$ with $s \geq 2$. For any real numbers $\alpha > 1$, $C > 0$ and any integer $N \geq 2$, there exists $g = (g_1, \cdots, g_s)$ with $-N/2 < g_i \leq N/2$ for $1 \leq i \leq s$ and such that*

$$\max_{f \in E_\alpha^s(C)} \sigma^2 \left(\frac{1}{N} \sum_{n=0}^{N-1} f\left(\left\{ \frac{n}{N} g + X \right\} \right) \right) = O(N^{-2\alpha} (\ln N)^{2\alpha(s-1)+1} (\ln \ln N)^{(2\alpha-1)(s-1)}).$$

Proof: From Theorem 2, and using Theorem 1 of [8] and the remark on p. 64 which says that $\forall \alpha > 1$, $\forall s \geq 2$ and $\forall N \geq 2$, there exists g with $-N/2 < g_i \leq N/2$ for $1 \leq i \leq s$ such that

$$P_\alpha(g, N) = O(N^{-\alpha} (\ln N)^{\alpha(s-1)+1} (\ln \ln N)^{(\alpha-1)(s-1)}),$$

we obtain immediately the result. ■

4 Numerical illustrations

As in [15], consider the test functions

$$H_2(x_1,\cdots,x_s)=\prod_{i=1}^{s}\left(1-\frac{\pi^2}{6}+\frac{\pi^2}{2}(1-2x_i)^2\right)$$

and

$$H_1(x_1,\cdots,x_s)=\prod_{i=1}^{s}(1-2\ln(2\sin\pi x_i)).$$

The Fourier coefficients of these functions are respectively $\hat{H}_2(h) = r(h)^{-2}$ and $\hat{H}_1(h) = r(h)^{-1}$.

Tables 1 and 2 compare the confidence interval width and the efficiency $1/(\sigma^2 \times t)$ (where t is the computational time [3]) of the estimator given by the combined method with the ones obtained by the crude Monte Carlo estimator, for the evaluation of the integrals of functions H_1 and H_2 over $[0,1]^5$. The comparisons are done for I independent and identically distributed random variables and N points of a lattice for the combined method, so that $NI \approx 10^6$, with respect to $I = 10^6$ iterations for the crude Monte Carlo method, so approximately with the same number of calls to the function.

Method	Estimation	Confidence interval width	Efficiency
Monte Carlo $I = 10^6$	9.976780e-01	1.377998e-01	9.099350e+00
GLP ($g^{(1)}, N = 1069$), $I = 935$	1.002022e+00	1.102490e-01	1.115490e+01
GLP ($g^{(2)}, N = 10007$), $I = 100$	9.837324e-01	6.043583e-02	3.637265e+01

Table 1: Estimations and confidence interval widths for the estimation of $\int_{[0,1]^5} H_1(x)dx$ in dimension $s = 5$. We use $g^{(1)} = (1, 63, 762, 970, 177)$ and $g^{(2)} = (1, 198, 9183, 6967, 8507)$ of [4].

Method	Estimation	Confidence interval width	Efficiency
Monte Carlo $I = 10^6$	1.002073e+00	6.953079e-02	6.929800e+01
GLP ($g^{(1)}, N = 1069$), $I = 935$	9.997191e-01	9.065306e-03	3.619068e+03
GLP ($g^{(2)}, N = 10007$), $I = 100$	9.997344e-01	9.625719e-04	2.947505e+05

Table 2: Estimations and confidence interval widths for the estimation of $\int_{[0,1]^5} H_2(x)dx$ in dimension $s = 5$. We use $g^{(1)} = (1, 63, 762, 970, 177)$ and $g^{(2)} = (1, 198, 9183, 6967, 8507)$ of [4].

As we can see, good lattice points improve the simulation with respect to crude one and the improvement is more important for $\alpha = 2$ (function H_2) than for $\alpha = 1$ (function H_1).

5 Conclusion

We have proved that for a specific class of functions the variance reduction in the combined Monte Carlo-quasi Monte Carlo method is greater than the one without specific regularity. We have also given numerical illustrations of this reduction order. A possibility for future research is to obtain the same kind of results, but by using the Walsh or the Haar series of a function and low discrepancy sequences.

Acknowledgement

The author would like to thank the anonymous referee for its remarks which improve the quality of the paper.

References

[1] Cranley R. and Patterson T. N. L.: Randomization of number theoretic methods for multiple integration. *SIAM J. Numer. Anal.*, 13, 904-914 (1976).

[2] Drmota M. and Tichy R. F.: *Sequences, Discrepancies and Applications*, volume 1651 of *Lecture Notes in Mathematics*. Springer Verlag, Heidelberg, 1997.

[3] Fishman G. S.: *Monte Carlo: Concepts, Algorithms and Applications*. Springer-Verlag, 1997.

[4] Hua L. K. and Wang Y.: *Applications of Number Theory to Numerical Analysis*. Springer Verlag, 1981.

[5] Joe S.: Randomization of lattice rules for numerical multiple integration. *J. Comp. Appl. Math.*, 3, 299-304 (1990).

[6] Korobov N. M.: On the computation of optimal coefficients. *Soviet Math. Dokl.*, 26, 590-593 (1982).

[7] Niederreiter H.: *Random Number Generation and Quasi-Monte Carlo Methods*. CBMS-SIAM 63, Philadelphia, 1992.

[8] Niederreiter H.: Improved Error Bounds for Lattice Rules. *J. Complexity*, 9, 60-75 (1993).

[9] Shaw J. E. H.: A quasirandom approach to integration in Bayesian statistics. *Ann. Statist.*, 16, 895-914 (1988).

[10] Sloan I. H. and Kachoyan P. J.: Lattice Methods for Multiple Integration: Theory, Error Analysis and Examples. *SIAM J. Numer. Anal.*, 24, 116-128 (1987).

[11] Tuffin B.: Variance reduction technique for a cellular system with dynamic resource sharing. In *Proceedings of the 10th European Simulation Multiconference*, 467-471, Budapest (1996).

[12] Tuffin B.: *Simulation accélérée par les méthodes de Monte Carlo et quasi-Monte Carlo : théorie et applications.* PhD thesis, Université de Rennes 1, 1997.

[13] Tuffin B.: Variance Reductions applied to Product-Form Multi-Class Queuing Network. *ACM Trans. Model.Comput. Simul.*, 7, 478-500 (1997).

[14] Zaremba S. K.: Some applications of multidimensional integration by parts. *Ann. Pol. Math*, XXI, 85-96 (1968).

[15] Zinterhof P.: Gratis Lattice Points for Multidimensional Integration. *Computing*, 38, 347-353 (1987).