Randomization of Quasi-Monte Carlo Methods for Error Estimation: Survey and Normal Approximation*

Bruno Tuffin
IRISA-INRIA
Campus universitaire de Beaulieu
35042 Rennes Cedex
France
E-mail: btuffin@irisa.fr

Abstract

Monte Carlo and quasi-Monte Carlo methods are simulation techniques that have been designed to efficiently estimate integrals for instance. Quasi-Monte Carlo asymptotically outperforms Monte Carlo, but the error can hardly be estimated. We propose here to recall how hybrid Monte Carlo/Quasi-Monte Carlo have been developed to easily get error estimations, with a special emphasis on the so-called randomly shifted low discrepancy sequences. Two additional points are investigated: we illustrate that the convergence rate is not always improved with respect to Monte Carlo and we discuss the confidence interval coverage problem.

Key words: Monte Carlo, quasi-Monte Carlo, Simulation

1 Introduction

Quasi-Monte Carlo (QMC) methods [11] are deterministic numerical integration tools that have been designed by analogy with Monte Carlo (MC) simulation technique. In Quasi-Monte Carlo, the random sample of MC is replaced by a sequence of "well distributed" points (called a *low discrepancy sequence*), so that the estimation is likely to converge faster, at least on regular integrands. Whereas very promising theoretically, QMC methods suffer from two main drawbacks. First, their domain of applicability is restricted with respect to MC. This is due to the correlation structure of the low discrepancy sequence (mandatory to speed up the convergence). Second, the error estimation, possible in theory, is intractable in practice whereas MC easily provides a (statistical) confidence interval.

The goal of this paper is first to recall how hybrid QMC methods have been designed in order to circumvent the error estimation problem. Then, as contributions, an illustration that the convergence rate of hybrid QMC is not always better than MC is presented, as well as a study of the normal approximation that is used to build the confidence interval. To our knowledge, this last point has not been studied in the literature but deserves consideration since we should know how to simultaneously increase the two involved parameters (instead of one in MC) to get a correct confidence interval coverage.

*This work has been partially supported by the Sure-Paths ACI project

The layout of the paper is as follows. In Section 2 we recall the basic definitions of MC and QMC and highlight the main drawbacks of QMC. Section 3 briefly describes the randomization methods from the literature. We especially focus on the convergence results for the shifted Quasi-Monte Carlo technique (but similar results exist for the other techniques). We illustrate in Section 4 on a counter-example that, nevertheless, the convergence speed is not always faster with randomized QMC techniques (even if it does not deter from using it on such an example) and we investigate Section 5 how the two parameters involved in the technique should increase in order to obtain a satisfactory confidence interval coverage. Section 6 is devoted to the conclusions and perspectives of research.

2 Monte Carlo and quasi-Monte Carlo methods

2.1 Basic notions on MC and QMC

Assume that we wish to compute the integral

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

MC considers a sequence $(X^{(i)})_{1\leq i\leq N}$ of N random and independent vectors uniformly distributed over $[0,1)^s$. From the central limit theorem, we have a confidence interval for $\mathcal{I}$:

$$\mathcal{I} \in \left(\frac{1}{N}\sum_{i=1}^{N} f(X^{(i)}) - \frac{c_\alpha \sigma}{\sqrt{N}}, \frac{1}{N}\sum_{i=1}^{N} f(X^{(i)}) + \frac{c_\alpha \sigma}{\sqrt{N}} \right)$$

at confidence level $1-\alpha$, where $c_\alpha = \mathcal{N}^{-1}(1-\frac{\alpha}{2})$, σ^2 is the variance of the random variable $f(X)$ with X uniformly distributed over $[0,1)^s$, and $\mathcal{N}$ is the cumulative distribution function of the Gaussian law with mean 0 and standard deviation 1. The convergence speed of this method is then, on average, $O(N^{-1/2})$, independent of the dimension s of the problem.

QMC methods [11] use a deterministic sequence $\mathcal{P} = (\xi^{(n)})_n$ instead of a random one, leading to an estimation

$$\frac{1}{N}\sum_{n=1}^{N} f(\xi^{(n)}). \tag{1}$$

The *discrepancy* aims at measuring the repartition of the sequence over $[0,1)^s$ by

$$D_N^*(\mathcal{P}) = \sup_{x\in[0,1]^s} \left| \frac{\sum_{n=1}^{N} 1_B(\xi^{(n)})}{N} - \prod_{i=1}^{s} x_i \right|$$

with $x = (x_1, \cdots, x_s)$. The sequence $\mathcal{P} = (\xi^{(n)})_{n\in\mathbb{N}}$ is then equi-distributed if and only if $\lim_{N\to+\infty} D_N^*(\mathcal{P}) = 0$. Error bounds for the approximation (1) of $\mathcal{I}$ are obtained in terms of the discrepancy. For P partition of $[0,1]^s$ in subintervals, let $\Delta(f,J)$ be the alterned sum of f values at the edges of sub-interval J. The variation in sense of Vitali is defined as

$$V_{Vit}(f) = \sup_P \sum_{J\in P} |\Delta(f,J)|.$$

Number of points	10^3	10^4	10^6	10^9	10^{12}	10^{16}
Bound	1.50	$3.34\ 10^{-1}$	$2.14\ 10^{-2}$	$1.21\ 10^{-4}$	$5.83\ 10^{-7}$	$1.92\ 10^{-10}$

Table 1: Bounds of discrepancy in dimension 5

From $V_{Vit}(f)$, we define $V(f)$, the variation of f in sense of Hardy and Krause by

$$V(f) = \sum_{k=1}^{s} \sum_{1 \leq i_1 < \cdots < i_k \leq s} V_{Vit}^{(k)}(f\ ; i_1, \cdots, i_k)$$

where $V_{Vit}^{(k)}(f\ ; i_1, \cdots, i_k)$ is the variation in sense of Vitali applied to the restriction of f to the space of dimension k $\{(u_1, \cdots, u_s) \in [0,1]^s\ : u_j = 1 \text{ for } j \neq i_1, \cdots, i_k\}$. We then have the following Koksma-Hlawka error bound [11]:

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(\xi^{(n)}) - \int_{[0,1)^s} f(u)du \right| \leq V(f) D_N^*(\mathcal{P}). \tag{2}$$

A sequence $\mathcal{P} = (\xi^{(n)})_{n \in \mathbb{N}}$ is said to be a *low discrepancy sequence* if $D_N^*(\mathcal{P}) = O(N^{-1}(\ln N)^s)$ (the best-known result for infinite sequences) so that the estimation converges faster than for MC. There exist many low discrepancy sequences; we can quote for instance Halton [11], Sobol' [1, 16], SQRT, Niederreiter [11], or Faure [5] sequences.

2.2 A drawback of quasi-Monte Carlo

MC methods easily give a confidence interval for the estimation by using the central limit theorem. One could argue that QMC methods, by means of the Koksma-Hlawka bound (or other related bounds), provide a stronger (since strict) bound that also allows to estimate the error. Unfortunately, these bounds, composed of the variation of the function and the discrepancy of the sequence, are useless in practice. To illustrate this, it can first be observed that QMC does converge for functions with infinite variation such as function

$$f(x_1, x_2, x_3) = \min(x_1 + x_2 + x_3, 1)$$

(exhibiting for instance a convergence rate $O(N^{-0.67})$ for the Niederreiter sequence in base $b = 3$) in dimension $s = 3$ [3], meaning that the Koksma-Hlawka bound is not tight. Also, even if this variation in the sense of Hardy and Krause is finite, it is the sum of $2^s - 1$ variation in the sense of Vitali, each of them being at least as difficult to estimate as $\mathcal{I}$ itself, their sum being potentially large, i.e. useless, even for moderate values of s. Second, the known bounds of discrepancy, even if useful asymptotically (since they provide the convergence rate $O(N^{-1}(\ln N)^s)$), usually require a very large N in order to be $O(N^{-1}(\ln N)^s)$, though it is observed quickly for the actual error. Tables 1 and 2 respectively display the bounds for the $(0,5)$-Niederreiter sequence in base $b = 5$ and dimension $s = 5$ and the $(0,25)$-Niederreiter sequence in base $b = 27$ and dimension $s = 25$ [11]. The convergence rate is longer to be observed when the dimension increases. In dimension $s = 25$, 10^{21} numbers of points could never be sampled, and the result is still larger than 1. The bad behavior of this bound for a fixed number of points can also be emphasized by the fact that, straightforwardly from the definition, $D_N^*(\mathcal{P}) \leq 1$.

Number of points	10^3	10^9	10^{15}	10^{21}	10^{35}
Bound	$2.88\ 10^5$	$1.11\ 10^6$	$2.46\ 10^5$	$5.22\ 10^3$	$1.62\ 10^{-3}$

Table 2: Bounds of discrepancy in dimension 25

In the next sections, we see how using hybrid QMC methods would benefit to QMC in order to circumvent these drawbacks.

3 Randomization of QMC for error estimation

We review in this section the randomization techniques that allow to use the central limit theorem in order to determine/estimate the error in QMC methods. The technique can also be seen as a use of low discrepancy sequences as a variance reduction technique in MC. We will more specifically focus on the so-called shifted low discrepancy method and use it as a benchmark. For numerical comparisons between the different randomization techniques, the reader can refer to [12, 18].

3.1 Shifted low discrepancy sequences

Let X be a random variable uniformly distributed over $[0,1)^s$, and $(\xi^{(n)})_{n\in\mathbb{N}}$ a low discrepancy sequence. As randomization technique [4], consider the random variable

$$Z = \frac{1}{N}\sum_{n=1}^{N} f(\{X+\xi^{(n)}\}) \tag{3}$$

instead of $\frac{1}{N}\sum_{n=1}^{N} f(\xi^{(n)})$ in QMC, where $\{x\}$ is the vector of fractional parts of the coordinates of x. The idea is to compute the average value of I independent copies Z_i of Z,

$$\frac{1}{I}\sum_{i=1}^{I} Z_i \tag{4}$$

and to obtain a confidence interval from the usual central limit theorem. Note that this technique can be seen as a special case of the antithetic variates method described in [6] where, for $X \sim U[0,1)^s$, the estimator of $\mathcal{I}$ is

$$\frac{1}{N}\sum_{n=1}^{N} f(\omega_n(X))$$

with ω_n a one-to-one mapping of $[0,1)^s$. Shifted-QMC is the case where $\omega_n(x) = \{x+\xi^{(n)}\}$.

To compute f the same number of times, we compare the variance of (4) with the variance of the standard MC simulation with NI random variables. We will obtain a variance reduction if and only if

$$\sigma^2\left(\frac{1}{N}\sum_{n=1}^{N} f(\{X+\xi^{(n)}\})\right) < \frac{1}{N}\sigma^2(f(X)). \tag{5}$$

The convergence speed of the method has been shown in [19]. If f is a function with bounded variation, we have

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

The estimator gives also good results on functions with infinite variation. It is proved in [19] that, for any low discrepancy sequence $\mathcal{P} = (\xi^{(n)})_{n\in\mathbb{N}}$, the mean variance $\sigma^2_{N,avg}$ of Z, mean taken over the set $\mathcal{F}$ of continuous functions f on $[0,1]^s$ equipped with the Wiener measure μ_W, is in $O(N^{-2}(\log N)^{2s})$. Recall that the Wiener measure μ_W is concentrated on functions with infinite variation [10]).

The convergence speed can even be faster for special classes of functions [20]. Let $\alpha > 1$, $C > 0$ et $\forall h \in \mathbb{Z}^s$, $r(h) = \prod_{i=1}^{s} \max(1, |h_i|)$. Let $E^s_\alpha(C)$ be the set of periodic functions $f : \mathbb{R}^s \to \mathbb{R}$, with period 1 over each coordinate, such that the Fourier coefficient of rank h of function f, $\hat{f}(h) = \int_{[0,1]^s} f(x)e(-h.x)dx$, where $x.y$ is the standard innner product of $x, y \in \mathbb{R}^s$ verifies

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

Let X be a random vector with uniform distribution over $[0,1]^s$. Then, for all $\alpha > 1$, $C > 0$ and $N \geq 1$, there exists $g \in \mathbb{Z}^s$ such that

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

This work has been extended in [8, 9] where the choice of a good g is investigated.

3.2 Other randomizations

Scrambled (t,s)-sequences. Scrambled (t,s)-sequences have been introduced by A. Owen in [13, 14]. The idea is to scramble the digits of special low-discrepancy sequences, the (t,s)-sequences in base b by using random permutations for the digits, while preserving the low-discrepancy property. It basically works as follows. Each permutation is uniformly distributed on the set of $b!$ permutations of $\{0,\ldots,b-1\}$. For a point $\xi = \sum_{k=1}^{\infty} a_k b^{-k}$ in base b, the permutation used for a_1 is $\pi_.$, the permutation used for a_2 is $\pi_{.,a_1}$ (dependent on a_1 but independent of $\pi_.$), and generally the permutation used for a_k is $\pi_{.,a_{k-1},\ldots,a_1}$. Other easier permutation choices have been recently proposed. For special classes of functions the variance of the quadrature rule based on scrambled nets can be as small as $O\left(N^{-3}(\log N)^{K-1}\right)$ [15].

Random-start Halton sequences. This method [17, 21] views Halton sequence as an application of multidimensional von Neuman-Kakutani transformation [7] with orbit vector $(0,\cdots,0)$. By randomly choosing an orbit vector, we obtain randomized Halton sequences.

A generalization of this last work (at the cost of introducing a bias) could be imagined as follows. Let T be a random variable defined over $\mathbb{N}$ and note $\pi_i = P[T = i]$ $\forall i \in \mathbb{N}$. Considering an s-dimensional sequence $(\xi^{(n)})_{n\in\mathbb{N}}$ with $\xi^{(n)} = (\xi_1^{(n)},\cdots,\xi_s^{(n)})$, define $X^{(n)}(T)$ $\forall n \geq 1$ by

$$X^{(n)}(T) = \xi^{(n+T)}.$$

Then, $\forall T$, the sequence $(X^{(n)}(T))_{n\geq 1}$ is random and

$$\frac{1}{N}\sum_{n=1}^{N} f(X^{(n)}(T))$$

tends to $\mathcal{I} = \int_{[0,1]^s} f(x)dx$ as $N \to \infty$ from the properties of sequence $(\xi^{(n)})_{n\in\mathbb{N}}$. By using I independent replications T_i, $1 \geq i \leq I$ of random variable T, we can obtain a confidence interval, from the estimator

$$\bar{f}_{N,I} = \frac{1}{I}\sum_{i=1}^{I}\frac{1}{N}\sum_{n=1}^{N} f(X^{(n)}(T_i)).$$

With respect to the other randomization approaches, this generalization of the random start Halton sequence to any other sequence introduces a bias. A research activity would be to determine which choices of seed distribution π will reduce this bias.

4 Randomized-QMC might be as slow as MC

Let us see now if the convergence is always sped up using shifted low discrepancy sequences. Consider the 2-dimensional function $f(x_1, x_2) = 1$ if $x_1 < x_2$, 0 otherwise, and the low discrepancy sequence $\{(n2^{-m}, \phi_2(n)) \ : 0 \leq n < 2^m\}$, where ϕ_b is the radical inverse function in base b: $\phi_b(n) = \sum_{k=0}^{m-1} \frac{n_k}{b^{k+1}}$ if $n = \sum_{k=0}^{m-1} n_k b^k$ with $n_k \in \{0, \cdots, b-1\}$.

The convergence results are displayed Figure 1 for MC and randomized QMC. MC and randomized QMC provide similar convergence speed $O(N^{-0.5})$. The fact that randomized QMC is not faster than MC is not due to the integrand alone since the SQRT sequence

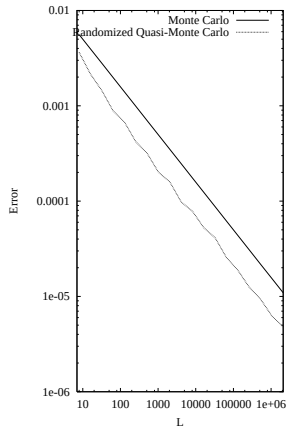


Figure 1: Convergence for $I = 10^3$ independent random variables and an increasing number N of QMC points for for RQMC with respect to the convergence with $L = NI$ replications for MC

performs better on this integration problem, with a regression analysis providing a rate $O(N^{-0.89})$. On the other hand, the sequence $\{(n2^{-m}, \phi_2(n)) \ : 0 \leq n < 2^m\}$ can also provide very could convergence results. For instance, if use it to integrate $f(x_1, x_2) =$

$\frac{\pi^2}{4}\sin(\pi x_1)\sin(\pi x_2)$, we obtain a very high convergence rate $O(N^{-1.86})$. It means that the shifted low-discrepancy sequence might be as slow as the traditional MC, at least on some functions (even if in "average", it is proved to be faster on continuous functions).

Nevertheless it has to be mentionned that, even if this example shows that the convergence rate is not always better,

- the variance is by a constant smaller using randomization;
- the simulation time decreases when using randomization (0.7 times the simulation time of MC for our example). This is due to the fact that there are less calls to the pseudo-random number generator.

For these two reasons, the randomization technique is still preferred.

5 Normal approximation and confidence interval coverage

In this section, we again concentrate on the randomly shifted low discrepancy sequences, but other randomizations can be considered as well.

Whereas Monte Carlo simulation requires only one parameter, the number I of independent random variables, randomized quasi-Monte Carlo simulation involves two parameters, still the number I of independent random variables, but also the number of points of the low discrepancy sequence N.

Thus, in MC simulation, we only increase I to obtain convergence, whereas in randomized QMC, we increase the *product* NI. Typically, users of randomized QMC usually consider a given number I of random variables (say $I = 100$), and only increase N to take benefit of the faster convergence of QMC. The number I of random variables would be only used to obtain a confidence interval based on the normal law from the central limit theorem. Nevertheless, it is not clear whether keeping a constant I and increasing N will force the estimator to converge in law to a normal distribution, or even to keep the normal approximation bounded.

To see how this convergence to the normal distribution works, let us first describe the Berry-Essen bound in MC simulation [2]. Consider a random variable X. Let $\mu = E(f(X))$, $\sigma^2 = E((f(X) - E(f(X))^2)$, $\beta = E(|f(X) - E(f(X)|^3)$, $\bar{X}_I = \frac{1}{I}\sum_{i=1}^{I} f(X^{(i)})$, for $(X^{(i)})_i$ independent copies of X, and let $\bar{F}_I(x)$ be the cumulative distribution of random variable

$$\sqrt{I}\frac{\bar{X}_I - \mu}{\sigma}$$

and $\mathcal{N}(x)$ be the Gaussian cumulative distribution with mean 0 and variance 1. We have $\forall x$,

$$\left|\bar{F}_I(x) - \mathcal{N}(x)\right| \leq C\frac{\beta}{\sigma^3\sqrt{I}}. \tag{6}$$

It means that in MC methods, when I increases, the confidence interval coverage error decreases.

To simplify the analysis, consider centered intervals. Let c_α be the $1 - \alpha/2$ quantile of the cumulative distribution of a Gaussian random variable Y with mean 0 and variance 1. Using (6), we have

$$\mathbb{P}\left(-c_\alpha - C\frac{\beta}{\sigma^3\sqrt{I}} \leq \frac{\bar{X}_I - \mu}{\sigma/\sqrt{I}} \leq c_\alpha + C\frac{\beta}{\sigma^3\sqrt{I}}\right) \geq \mathbb{P}\left(-c_\alpha \leq Y \leq c_\alpha\right) = 1 - \alpha.$$

Thus a confidence interval for μ at confidence level at least $1-\alpha$ is

$$\left(\bar{X}_I - c_\alpha \frac{\sigma}{\sqrt{I}} - \frac{c\beta}{\sigma^2 I}, \bar{X}_I + c_\alpha \frac{\sigma}{\sqrt{I}} + \frac{c\beta}{\sigma^2 I}\right).$$

If we compare it with the traditional confidence interval $\left(\bar{X}_I - c_\alpha \frac{\sigma}{\sqrt{I}}, \bar{X}_I + c_\alpha \frac{\sigma}{\sqrt{I}}\right)$, it can be observed that the half-width of the interval including the normal approximation error is about the half-width of the interval without this error for a large I since $C_\alpha \frac{\sigma}{\sqrt{I}} + \frac{c\beta}{\sigma^2 I} = C_\alpha \frac{\sigma}{\sqrt{I}} + O\left(\frac{1}{I}\right)$. Indeed, the error of considering the mean estimation Gaussian (relatively) vanishes in $O(1/\sqrt{I})$.

Consider now randomized QMC methods. Let β_N be the centred absolute moment of order three of r.v. $\frac{1}{N}\sum_{n=1}^{N} f(\{X+\xi^{(n)}\})$, σ_N be its standard deviation, $\bar{X}_{I,N} = \frac{1}{IN}\sum_{i=1}^{I}\sum_{n=1}^{N} f(\{X^{(i)}+\xi^{(n)}\})$, and $F_{N,I}(x)$ be the cumulative distribution of random variable $\sqrt{I}\frac{\bar{X}_{I,N}-\mu}{\sigma_N}$. We still have from the Berry-Essen bound that

$$\left|\bar{F}_{I,N}(x) - \mathcal{N}(x)\right| \leq C\frac{\beta_N}{\sigma_N^3\sqrt{I}}. \tag{7}$$

Following the MC analysis, a confidence interval incorporating the normal approximation error is

$$\left(\bar{X}_{N,I} - c_\alpha \frac{\sigma_N}{\sqrt{I}} - \frac{c\beta_N}{\sigma_N^2 I}, \bar{X}_{N,I} + c_\alpha \frac{\sigma_N}{\sqrt{I}} + \frac{c\beta_N}{\sigma_N^2 I}\right).$$

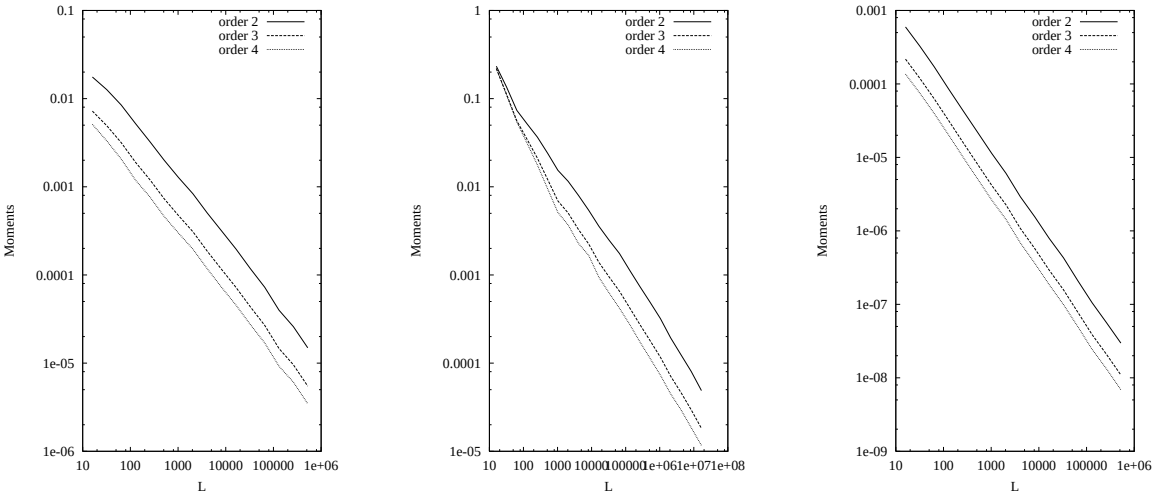
Thus the traditional interval without inclusion of the normal approximation, of half-width $c_\alpha \frac{\sigma_N}{\sqrt{I}}$, will be asymptotically correct if and only if $\frac{c\beta_N}{\sigma_N^2 I}$ decreases faster than $c_\alpha \frac{\sigma_N}{\sqrt{I}}$, meaning that we have to increase N and I such that $\frac{c\beta_N}{\sigma_N^3\sqrt{I}} \to 0$. Thus the idea to fix I and increase N is valid if and only if $\frac{c\beta_N}{\sigma_N^3} \to 0$ when $N \to \infty$.

But it must be noted that, by applying Hölder inequality, there exists a constant C_{st} such that

$$\beta_N \geq C_{st}\sigma_N^3.$$

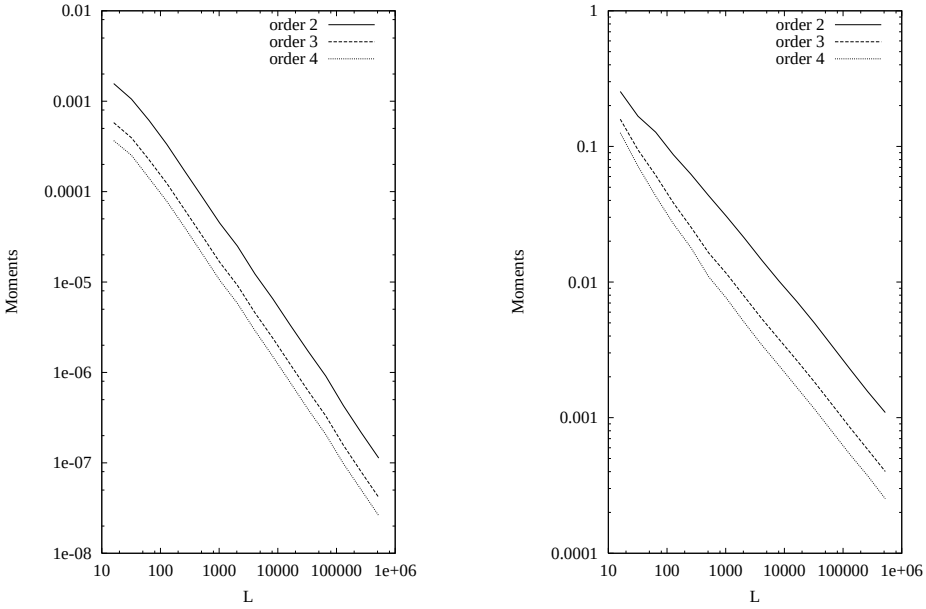
It means that the strategy of fixing I and increasing N does not improve the normal approximation, and can even introduce an additional error if σ_N^3 decreases faster than β_N. In order to investigate this last possibility, Figures 2 and 3 illustrate the convergence rate of the (normalized) second, third and even fourth moments for MC and randomized QMC on four different kinds of functions in terms of $L = NI$ for $I = 100$ and an increasing N. It can be remarked that the (normalized) moments converge at the same rates, so that, at least for these functions, β_N/σ_N^3 is bounded. At first sight, it might seem that σ_N^3 decreases slower than β_N for function $\prod_{i=1}^{10} 12(x_i - 0.5)^2$, but the curves are parallel after a while (and it cannot happen from Hölder inequality). Other examples exhibit the same results. In conclusion, we have not been able to find a counter-example showing that β_N/σ_N^3 is unbounded when $N \to \infty$, but it can be shown that it is lower-bounded.

Based on this idea that the normal approximation does not decrease to zero when $NI \to \infty$, let us now investigate how to incorporate the additional Berry-Essen term in practice. Let $\hat{\beta}_N$ and $\hat{\sigma}_N$ be estimations of β_N and σ_N.



(a) $\prod_{i=1}^{10} \frac{\pi}{2} \sin(\pi x_i)$ (b) $\prod_{i=1}^{10} 12(x_i - 0.5)^2$ (c) $\frac{1}{5} \sum_{i=1}^{10} x_i$

Figure 2:



(a) $\prod_{i=1}^{10} \frac{e^{-|x_i - 0.5|}}{2 - 2e^{0.5}}$ (b) $\prod_{i=1}^{10} 2\ 1_{[x_i > 0.5]}$

Figure 3:

- A first idea would be to integrate the $\frac{c\beta_N}{\sigma_N^2 I}$ term to improve the interval coverage:

$$\left(\bar{X}_{N,I} - C_\alpha \frac{\hat{\sigma}_N}{\sqrt{I}} - \frac{c\hat{\beta}_N}{\hat{\sigma}_N^2 I}, \bar{X}_{N,I} + C_\alpha \frac{\hat{\sigma}_N}{\sqrt{I}} + \frac{c\hat{\beta}_N}{\hat{\sigma}_N^2 I}\right)$$

- A second idea is to dynamically increase N, I so that $\frac{\hat{\beta}_N}{\hat{\sigma}_N^3 \sqrt{I}} \to 0$ and use the traditional interval

$$\left(\bar{X}_{N,I} - C_\alpha \frac{\hat{\sigma}_N}{\sqrt{I}}, \bar{X}_{N,I} + C_\alpha \frac{\hat{\sigma}_N}{\sqrt{I}}\right).$$

6 Conclusions and Perspectives

Randomized quasi-Monte Carlo methods have been designed to provide an error estimation in QMC, as usually done in MC. In this paper, we have recalled the existing convergence results of these methods, with a special focus on shifted-low discrepancy sequences. As main contributions, we have illustrated that randomized QMC does not always improve the convergence rate with respect to MC (even, it is still preferred thanks to the smaller computational time), and we have investigated the normal approximation problem, showing that, unlike for MC, it is advised to add in the confidence interval a term bounding the normal approximation error.

As directions for future research, we plan to look more closely at how simultaneously increase N and I in order to obtain a minimal confidence interval width for a given total computational time. We plan also to more deeply investigate the generalization of the random start Halton sequence that we have introduced: we wish to study how the bias can be reduced by a given choice of the random start, and how this method compares with the other randomization techniques.

References

[1] I.A.. Antonov and Saleev V.M. An economic method of computing lp_τ-sequences. *USSR Computational Math. and Math. Phys.*, 19(1):252–256, 1979.

[2] V. Bentkus and F. Götze. The Berry-Esseen bound for Student's statistic. *The Annals of Probability*, 24(1):491–503, 1996.

[3] N. Bouleau and D. Lépingle. *Numerical methods for stochastic processes.* John Wiley & Sons, 1994.

[4] R. Cranley and T.N.L. Patterson. Randomization of number theoretic methods for multiple integration. *SIAM J. Numer. Anal.*, 13(6):904–914, December 1976.

[5] H. Faure. Discrépances de suites associées à un système de numération (en dimension s). *Acta Arith.*, 41:337–351, 1982.

[6] G. S. Fishman and B.D. Huang. Antithetic Variates Revisited. *Communications of the ACM*, 26(11):964–971, 1983.

[7] J.P. Lambert. Quasi-Monte Carlo, low discrepancy sequences, and ergodic transformations. *Journal of Computational and Applied Mathemathics*, 12&13:419–423, 1985.

[8] P. L'Ecuyer and C. Lemieux. Variance reduction via lattice rules. *Management Science*, 49(6):1214–1235, 1998.

[9] C. Lemieux and P. L'Ecuyer. Efficiency improvement by lattice rules for pricing asian options. In *Proceedings of the 1998 Winter Simulation Conference*, pages 579–585, 1998.

[10] W. J. Morokoff and R. E. Caflisch. Quasi-Random Sequences and their Discrepancies. *SIAM Journal on Scientific Computing*, pages 1571–1599, December 1994.

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

[12] G. Ökten and W. Eastman. Randomized Quasi-Monte Carlo Methods in Pricing Securities. *Submitted for publication.*

[13] A. B. Owen. Randomly permuted (t, m, s)-nets and (t, s)-sequences. In Harald Niederreiter and Peter Jau-Shyong Shiue, editors, *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, volume 106 of *Lecture Notes in Statistics*, pages 299–315. Springer, 1995.

[14] A. B. Owen. Monte Carlo Variance of Scrambled Net Quadrature. *SIAM Journal of Numerical Analysis*, 34:1884–1910, 1997.

[15] A. B. Owen. Scrambled Net Variance for Integrals of Smooth Functions. *Annals of Statistic*, 25:1541–1562, 1997.

[16] I.M. Sobol'. Uniformly distributed sequences with an additional uniform property. *USSR Computational Math. and Math. Phys.*, 16(5):236–242, 1976.

[17] J. Struckmeier. Fast generation of low-discrepancy sequences. *Journal of COmputational and Applied Mathematics*, 61:29–41, 1995.

[18] B. Tuffin. On the Use of Low Discrepancy Sequences in Monte Carlo Methods. *Monte Carlo Methods and Applications*, 2(4):295–320, 1996.

[19] B. Tuffin. Variance Reductions applied to Product-Form Multi-Class Queuing Network. *ACM Transactions on Modeling and Computer Simulation*, 7(4):478–500, 1997.

[20] B. Tuffin. Variance reduction order using good lattice points in Monte Carlo methods. *Computing*, 61(4):371–378, 1998.

[21] X. Wang and F.J. Hickernell. Randomized Halton sequences. *Mathematical and Computer Modelling*, 32:887–899, 2000.