

## ON THE ERROR DISTRIBUTION FOR RANDOMLY-SHIFTED LATTICE RULES

Pierre L'Ecuyer

DIRO, Université de Montréal  
C.P. 6128, Succ. Centre-Ville  
Montréal (Québec), H3C 3J7, CANADA

Bruno Tuffin

INRIA Rennes Bretagne Atlantique  
Campus de Beaulieu  
35042 Rennes Cedex, FRANCE

### ABSTRACT

Randomized quasi-Monte Carlo (RQMC) methods estimate the expectation of a random variable by the average of  $n$  dependent realizations of it. In general, due to the strong dependence, the estimation error may not obey a central limit theorem. Analysis of RQMC methods have so far focused mostly on the convergence rates of asymptotic worst-case error bounds and variance bounds, when  $n \rightarrow \infty$ , but little is known about the limiting distribution of the error. Here we examine this limiting distribution for the special case of a randomly-shifted lattice rule, when the integrand is smooth. We start with simple one-dimensional functions, where we show that the limiting distribution is uniform over a bounded interval if the integrand is non-periodic, and has a square root form over a bounded interval if the integrand is periodic. In higher dimensions, for linear functions, the distribution function of the properly standardized error converges to a spline of degree equal to the dimension.

### 1 INTRODUCTION

When an unknown mathematical expectation  $\mu$  is estimated by an average of  $n$  independent and identically distributed replicates of an unbiased estimator, as frequently done in simulation experiments, it is common practice to invoke the central-limit theorem (CLT) to assess the estimation error and compute a confidence interval on  $\mu$  (Asmussen and Glynn 2007). But when the replicates are not independent, which happens for example in generalized antithetic variates or randomized quasi-Monte Carlo (RQMC) settings, then the CLT does not always hold. In this situation, when the limiting distribution of the (properly standardized) average is not normal, it is of interest to understand what it is. The aim of this paper is to make a few steps in this direction.

Recall that any estimator computed by a simulation program in practice can be written as a function of a sequence of independent  $U(0, 1)$  random variables, i.e., of a random point in the  $s$ -dimensional unit hypercube  $(0, 1)^s$  for some integer  $s$ , where  $s$  represents the number of calls to the random number generator in the simulation (Lemieux and L'Ecuyer 2000). The Monte Carlo method averages  $n$  independent replicates of the underlying estimator, obtained from  $n$  independent random points in  $(0, 1)^s$ . In the RQMC method, the  $n$  independent points are replaced by  $n$  dependent random points, each uniformly distributed over  $(0, 1)^s$ , and that cover the unit hypercube more evenly than typical independent random points (Niederreiter 1992, L'Ecuyer 2009). Two popular ways of achieving this are randomly-shifted lattice rules and digitally-shifted nets (L'Ecuyer 2009).

A CLT is known to hold for certain special cases of RQMC methods that involve a large amount of randomization for the dependent points, such as Latin hypercube sampling (LHS) (Owen 1992) and digital nets with full nested scrambling (Loh 2003). In the LHS case, a bound on the total variation convergence to the normal distribution is also available. See Loh (2005) for a survey. But LHS is not one of the most powerful RQMC methods, because it ensures good uniformity only for the one-dimensional projections of the  $s$ -dimensional points, and nested scrambling is not widely used because it is time-consuming.

Here we examine the case where  $\mu$  is estimated by a randomly-shifted lattice rule. The problem is difficult, so we start with a simplified situation where  $\mu$  is defined by the integral of some function  $f$  over the interval  $(0, 1)$ . It turns out that for a non-periodic smooth function  $f$ , the asymptotic distribution of the properly standardized average is not normal, but uniform.

Thus, a CLT does not hold for  $s = 1$ , and therefore there is no hope in proving it for general  $s$ . When  $f$  is periodic in the sense that  $f(0) = f(1)$ , the convergence is faster and the limiting distribution of the standardized average has a square-root distribution function over a finite interval (it behaves as the square of a centered uniform). In higher dimensions, for linear functions, the distribution function of the properly standardized error converges to a spline of degree equal to the dimension.

In Sections 2 and 3, we recall the definitions of RQMC methods and randomly-shifted lattice rules, and introduce useful notation. In Section 4, we derive our results for the one-dimensional case and give several examples. In Section 5, we examine the general  $s$ -dimensional case, for which the limiting distributions turn out to be more complicated to handle. We estimate and plot the distributions for some examples. A conclusion follows.

## 2 RANDOMIZED QUASI-MONTE CARLO

Suppose we want to estimate a mathematical expectation  $\mu$  defined by

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

for some function  $f : [0, 1]^s \rightarrow \mathbb{R}$ , where  $[0, 1]^s$  is the  $s$ -dimensional unit hypercube and  $\mathbf{u} = (u_1, \dots, u_s)$  represents a point in this cube. An RQMC estimator of  $\mu$  has the form

$$\hat{\mu}_{n,\text{rqmc}} = \frac{1}{n} \sum_{i=0}^{n-1} f(\mathbf{U}_i). \tag{2}$$

where each point  $\mathbf{U}_i$  has the uniform distribution over the hypercube  $(0, 1)^s$ , but the entire point set  $P_n = \{\mathbf{U}_0, \dots, \mathbf{U}_{n-1}\} \subset (0, 1)^s$  covers the hypercube more uniformly than typical independent random points. A meaningful definition of “more uniformly” requires a specific definition of a figure of merit to measure the uniformity. For such definitions and general background on quasi-Monte Carlo methods and RQMC, the reader is referred to [Niederreiter \(1992\)](#), [Sloan and Joe \(1994\)](#), [L'Ecuyer and Lemieux \(2002\)](#), [L'Ecuyer \(2009\)](#).

The RQMC estimator (2) has expectation  $\mu$  and variance

$$\text{Var}[\hat{\mu}_{n,\text{rqmc}}] = \mathbb{E}[(\hat{\mu}_{n,\text{rqmc}} - \mu)^2]. \tag{3}$$

The usual way of estimating this variance and computing a confidence interval on  $\mu$  is to obtain  $m$  independent realizations of  $\hat{\mu}_{n,\text{rqmc}}$ , say  $X_1, \dots, X_m$ , based on  $m$  independent randomizations of  $P_n$ , and compute their sample mean  $\bar{X}_m$  and their sample variance and  $S_{x,m}^2$ . One has  $\mathbb{E}[\bar{X}_m] = \mu$  and  $\mathbb{E}[S_{x,m}^2] = m \text{Var}[\bar{X}_m]$  ([L'Ecuyer and Lemieux 2000](#)). By assuming that  $\bar{X}_m$  is approximately normally distributed, one can compute a confidence interval on  $\mu$  in a standard way. When this assumption is not true, the normal approximation might still be good for  $\bar{X}_m$  if  $m$  is large (due to the central-limit effect when  $m \rightarrow \infty$ ), but perhaps not if  $m$  is small, which is the usual case. This motivates the need for a better understanding of the distribution of  $\hat{\mu}_{n,\text{rqmc}}$ .

## 3 RANDOMLY-SHIFTED LATTICE RULES

An *integration lattice* is a discrete vector space of the form

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

where  $\mathbf{v}_1, \dots, \mathbf{v}_s \in \mathbb{R}^s$  are linearly independent over  $\mathbb{R}$  and where  $L_s$  contains  $\mathbb{Z}^s$ , the set of integer vectors. The *dual lattice* to  $L_s$  is

$$L_s^* = \{ \mathbf{h} \in \mathbb{R}^s : \mathbf{h}^t \mathbf{v} \in \mathbb{Z} \text{ for all } \mathbf{v} \in L_s \},$$

where “ $t$ ” means “transposed” and all vectors are assumed to be column vectors. It is known that all coordinates of all vectors of  $L_s^*$  are integers.

Let  $P_n = L_s \cap [0, 1]^s = \{\mathbf{u}_0, \dots, \mathbf{u}_{n-1}\}$  be the set of lattice points that belong to the unit hypercube  $[0, 1]^s$ , where  $n$  denotes the cardinality of  $P_n$ . One can show that all coordinates of all vectors of  $L_s$  must be multiples of  $1/n$ . A *lattice rule* is a numerical integration method that approximates  $\mu$  by the average of  $f(\mathbf{u}_0), \dots, f(\mathbf{u}_{n-1})$ . Further details on integration lattices, lattice rules, and ways of measuring the uniformity of  $P_n$  for an integration lattice, can be found in Sloan and Joe (1994), Tuffin (1998), L'Ecuyer and Lemieux (2000), for example.

To apply a *random shift modulo 1* to  $P_n$ , we generate a single point  $\mathbf{U}$  uniformly over  $(0, 1)^s$  and add it to each point of  $P_n$ , modulo 1, coordinate-wise (Cranley and Patterson 1976, Sloan and Joe 1994, L'Ecuyer and Lemieux 2000). For a lattice rule, this is the same as randomly shifting  $L_s$  and then taking the intersection with the unit hypercube  $[0, 1]^s$ . The shift preserves the global uniformity of the point set  $P_n$  and each point  $\mathbf{U}_i = (\mathbf{u}_i + \mathbf{U}) \bmod 1$  of the randomized version of  $P_n$  is uniformly distributed over  $[0, 1]^s$ .

For a randomly-shifted lattice rule, the integration error is the random variable

$$g_n(\mathbf{U}) \stackrel{\text{def}}{=} \hat{\mu}_{n,\text{rqmc}} - \mu, \tag{4}$$

which is a function of the random shift  $\mathbf{U}$ . We are interested in the distribution of this random variable when  $n$  is large. Suppose that the function  $f$  has the Fourier expansion

$$f(\mathbf{u}) = \sum_{\mathbf{h} \in \mathbb{Z}^s} \hat{f}(\mathbf{h}) \exp(2\pi i \mathbf{h}^t \mathbf{u}), \tag{5}$$

with Fourier coefficients

$$\hat{f}(\mathbf{h}) = \int_{(0,1)^s} f(\mathbf{u}) \exp(-2\pi i \mathbf{h}^t \mathbf{u}) d\mathbf{u},$$

where  $i = \sqrt{-1}$ . Then one can show that the Fourier coefficients of  $g_n$  are  $\hat{g}_n(\mathbf{h}) = \hat{f}(\mathbf{h})$  if  $\mathbf{0} \neq \mathbf{h} \in L_s^*$ , and  $\hat{g}_n(\mathbf{h}) = 0$  otherwise (L'Ecuyer and Lemieux 2000, Sloan and Joe 1994). Therefore, the Fourier expansion of  $g_n$  can be written (where it exists) as

$$g_n(\mathbf{u}) = \sum_{\mathbf{0} \neq \mathbf{h} \in L_s^*} \hat{f}(\mathbf{h}) \exp(2\pi i \mathbf{h}^t \mathbf{u}). \tag{6}$$

This expansion may offer a possible path to obtain the distribution of  $g_n(\mathbf{U})$ , or its asymptotic when  $n \rightarrow \infty$ , if we can handle the right side of (6). We will return to this in Section 5. In the next section, we will take a different path.

#### 4 ERROR DISTRIBUTION IN ONE DIMENSION

We start by studying the distribution of  $g(\mathbf{U})$  for the simple case of a randomly-shifted lattice rule in one dimension ( $s = 1$ ). In this case, the random shift can be generated uniformly over  $[0, 1/n)$ . That is, the randomly-shifted points can be written as  $\{U/n, (1+U)/n, \dots, (n-1+U)/n\}$  where  $U$  has the uniform distribution over  $[0, 1)$ . We will write a Taylor expansion of the error over each interval of length  $1/n$ , and sum up these terms to get a handle on the total error and its distribution.

Suppose that  $f$  has a bounded third derivative over the interval  $[0, 1]$ . Then we can write its Taylor expansion around the center  $x_i = (i + 1/2)/n$  of each interval  $[i/n, (i + 1)/n)$  as

$$f(u) = f(x_i) + (u - x_i) f'(x_i) + \frac{(u - x_i)^2}{2} f''(x_i) + e_i$$

for  $i/n \leq u < (i + 1)/n$  and  $i = 0, \dots, n - 1$ , where  $|e_i| \leq \sup_{i/n \leq u \leq (i+1)/n} |f'''(u)|/48n^3$ . Approximating  $\mu = \int_0^1 f(u) du$  by the average of  $f(x_0), \dots, f(x_{n-1})$  is called the *midpoint rule*, and it is known (Davis and Rabinowitz 1984) that the integration error for this deterministic rule is

$$\frac{-1}{24n^2} \int_0^1 f''(u) du + O(n^{-3}).$$

Because  $f'$  is continuous under our assumption, we also have that  $\int_0^1 f''(u)du = f'(1) - f'(0)$ , and then

$$\frac{1}{n} \sum_{i=0}^{n-1} f(x_i) = \mu - \frac{f'(1) - f'(0)}{24n^2} + O(n^{-3}). \tag{7}$$

Applying the same reasoning with  $f'$  and  $f''$  in place of  $f$ , we also obtain

$$\frac{1}{n} \sum_{i=0}^{n-1} f'(x_i) = f(1) - f(0) - \frac{[f''(1) - f''(0)]}{24n^2} + O(n^{-3}) \tag{8}$$

when  $f''$  is continuous over  $[0, 1]$ , and

$$\frac{1}{n} \sum_{i=0}^{n-1} f''(x_i) = f'(1) - f'(0) - \frac{[f'''(1) - f'''(0)]}{24n^2} + O(n^{-3}) \tag{9}$$

when  $f'''$  is continuous over  $[0, 1]$ . The expressions on the left of the two previous displays are approximations of the average derivative and average second derivative over the interval  $(0, 1)$ , respectively, by the midpoint rule. Putting these ingredients together, we obtain that the integration error is

$$\begin{aligned} g_n(U) &= \frac{1}{n} \sum_{i=0}^{n-1} \left[ f(x_i) - \mu + \frac{(U - 1/2)f'(x_i)}{n} + \frac{(U - 1/2)^2 f''(x_i)}{2n^2} + e_i \right] \\ &= (U - 1/2) \frac{f(1) - f(0)}{n} + [(U - 1/2)^2 - 1/12] \frac{f'(1) - f'(0)}{2n^2} + O(n^{-3}). \end{aligned} \tag{10}$$

In fact, (10) is a second order Euler-MacLaurin expansion of  $g_n(U)$ . The expansion of order  $m$ , for a function  $f$  having  $m + 1$  continuous derivatives, is (Davis and Rabinowitz 1984, Page 136) or (Dahlquist and Bjorck 2008, page 303):

$$g_n(U) = \sum_{k=1}^m B_k(U) \frac{f^{(k-1)}(1) - f^{(k-1)}(0)}{k!n^k} + O(n^{-(m+1)}), \tag{11}$$

where  $B_k$  is the  $k$ th Bernoulli polynomial, defined by

$$B_k(u) = \sum_{j=0}^k \frac{1}{j+1} \sum_{\ell=0}^j (-1)^\ell \binom{j}{\ell} (u + \ell)^k$$

(Abramowitz and Stegun 1970). This gives  $B_0(u) = 1$ ,  $B_1(u) = u - 1/2$ ,  $B_2(u) = u^2 - u + 1/6$ ,  $B_3(u) = u^3 - (3/2)u^2 + (1/2)u$ ,  $B_4(u) = u^4 - 2u^3 + u^2 - 1/30$ , and so on. A examination of the proof in (Dahlquist and Bjorck 2008) reveals that  $f^{(m+1)}$  does not have to be continuous; it suffices that it is integrable over  $[0, 1]$ , i.e., that  $\int_0^1 f^{(m+1)}(u)du < \infty$ .

**Proposition 1** (i) If  $f$  has an integrable second derivative over  $[0, 1]$  and  $f(1) \neq f(0)$ , then

$$W_n^{(1)} \stackrel{\text{def}}{=} \frac{ng_n(U)}{f(1) - f(0)} = U - 1/2 + O(1/n),$$

so  $W_n^{(1)} + 1/2$  converges in distribution to the  $U(0, 1)$  random variable  $U$  when  $n \rightarrow \infty$ .

(ii) If  $f$  has an integrable third derivative over  $[0, 1]$ ,  $f(1) = f(0)$ ,  $f'(1) \neq f'(0)$ , and if we define

$$W_n^{(2)} \stackrel{\text{def}}{=} \frac{2n^2 g_n(U)}{f'(1) - f'(0)} = (U - 1/2)^2 - 1/12 + O(1/n),$$

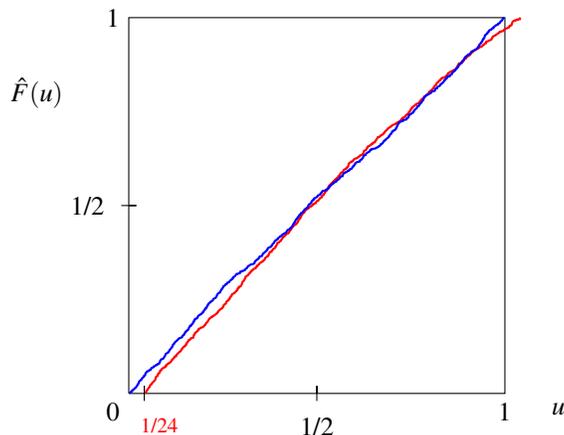


Figure 1: Empirical distribution  $\hat{F}$  of 1000 replicates of  $ng_n(U) + 1/2$  for one-dimensional lattices with  $n = 4$  points (in red) and  $n = 1024$  points (in blue), for  $f(u) = u^2$ .

then  $W_n^{(2)} + 1/12$  converges in distribution to the random variable  $(U - 1/2)^2$ , which has density  $1/\sqrt{x}$  and distribution function  $2\sqrt{x}$  for  $0 < x < 1/4$  and 0 elsewhere. This means that  $2[W_n^{(2)} + 1/12]^{1/2}$  converges in distribution to a uniform random variable over  $(0, 1)$ .

(iii) If  $f$  has  $m + 1$  derivatives over  $[0, 1]$ ,  $\int_0^1 f^{(m+1)}(u)du < \infty$ , and  $f^{(k)}(1) = f^{(k)}(0)$  for  $k = 0, \dots, m - 2$ , but  $f^{(m-1)}(1) \neq f^{(m-1)}(0)$ , then  $g_n(U) = B_m(U)[f^{(m-1)}(1) - f^{(m-1)}(0)]/(m!n^m) + O(n^{-(m+1)})$ , and therefore

$$W_n^{(m)} \stackrel{\text{def}}{=} \frac{m!n^m g_n(U)}{f^{(m-1)}(1) - f^{(m-1)}(0)} = B_m(U) + O(1/n)$$

converges in distribution to the random variable  $B_m(U)$ . For  $m > 2$ , this random variable is not monotone in  $U$  and it has a more complicated distribution than in the previous cases.

*Proof.* In case (i), we have  $g_n(U) = (U - 1/2)(f(1) - f(0))/n + O(n^{-2})$ , from which  $W_n^{(1)} = U - 1/2 + O(1/n)$ . In case (ii),  $g_n(U) = [(U - 1/2)^2 - 1/12](f'(1) - f'(0))/(2n^2) + O(n^{-3})$ , and therefore  $W_n^{(2)} = (U - 1/2)^2 - 1/12 + O(1/n)$ , and so  $[W_n^{(2)} + 1/12]^{1/2}$  converges to  $|U - 1/2|$ , which is uniform over  $(0, 1/2)$ . The result follows. Case (iii) follows similarly from (11).  $\square$

**Example 1** Let  $f(u) = u^2$ . We have  $f(1) - f(0) = 1$ , so we are in case (i) and  $W_n^{(1)} + 1/2 = ng_n(U) + 1/2$  is approximately  $U(0, 1)$ , with  $O(1/n)$  approximation error. In fact, from (10), this  $O(1/n)$  approximation error is bounded by  $|(u^2 - u + 1/6)(f'(1) - f'(0))|/(2n) + O(n^{-2}) \leq 1/(12n) + O(n^{-2})$ . Figure 1 shows the empirical distribution  $\hat{F}$  of 1000 independent replications of  $ng_n(U) + 1/2$  for one-dimensional lattices with  $n = 4$  and  $n = 2^{10} = 1024$  points. Recall that these lattices have point sets  $\{0, 1, \dots, (n - 1)/n\}$ . The observed distribution with  $n = 1024$  is already very close to uniform over  $(0, 1)$ . To understand the right shift of the curve for  $n = 4$  (compared with the uniform distribution function), observe that for this function, we have exactly  $ng_n(U) + 1/2 = U + ((U - 1/2)^2 - 1/12)/n$ , which has expectation  $1/2$  but takes its values in the interval  $(1/6n, 1 + 1/6n)$ . For  $n = 4$ , the values are in the interval  $(1/24, 1 + 1/24)$ , as observed in the figure. For  $n = 1024$ , the shift is too small to be visible to the eye.  $\square$

**Example 2** Let  $f(u) = u(1 - u)$ . We have  $f(1) = f(0) = 0$  and  $f'(1) - f'(0) = -2$ , so we are in case (ii) and  $W_n^{(2)} + 1/12 = -n^2 g_n(U) + 1/12$  converges to  $(U - 1/2)^2$  in distribution when  $n \rightarrow \infty$ . In fact, here we have  $f'''(u) = 0$ , and it can be shown that  $-n^2 g_n(U) + 1/12$  has exactly the same distribution as  $(U - 1/2)^2$ . That is,  $2[-n^2 g_n(U) + 1/12]^{1/2}$  has exactly the  $U(0, 1)$  distribution. Figure 2 shows the empirical distribution  $\hat{F}$  of 1000 independent replications of  $2[-n^2 g_n(U) + 1/12]^{1/2}$  for one-dimensional lattices with  $n = 4$  and  $n = 1024$  points.  $\square$

**Example 3** Take  $f(u) = u^3 + u^2 - 2u$ . We have  $f(0) = f(1) = 0$ ,  $f'(1) - f'(0) = 5$ , and  $f'''$  does not vanish. Here,  $2n^2 g_n(U)/5 + 1/12$  has approximately the same distribution as  $(U - 1/2)^2$ , and  $2[2n^2 g_n(U)/5 + 1/12]^{1/2}$  is approximately  $U(0, 1)$ , but this distribution is not exact. Figure 3 shows the empirical distribution  $\hat{F}$  of 1000 independent replications of  $2[2n^2 g_n(U)/5 + 1/12]^{1/2}$  for one-dimensional lattices with  $n = 64$  and  $n = 16324$  points.  $\square$

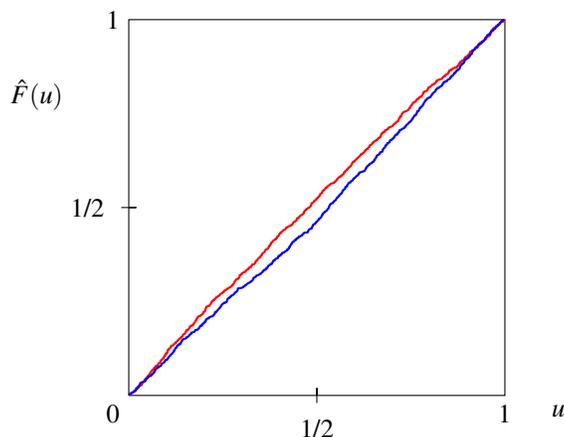


Figure 2: Empirical distribution  $\hat{F}$  of 1000 replicates of  $2[-n^2g_n(U) + 1/12]^{1/2}$  for one-dimensional lattices with  $n = 4$  points (in red) and  $n = 1024$  points (in blue), for  $f(u) = u(1 - u)$ .

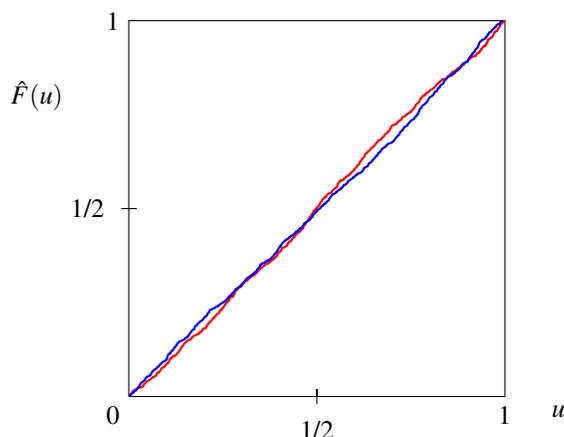


Figure 3: Empirical distribution  $\hat{F}$  of 1000 replicates of  $2[2n^2g_n(U)/5 + 1/12]^{1/2}$  for one-dimensional lattices with  $n = 64$  points (in red) and  $n = 2^{14} = 16384$  points (in blue), for  $f(u) = u^3 + u^2 - 2u$ .

**Example 4** Consider now  $f(u) = \sin(2\pi mu)$  for an arbitrary positive integer  $m$ . In this case, we have  $f^{(k)}(1) = f^{(k)}(0)$  for all  $k \geq 0$ . This implies that  $g_n(U)$  converges faster than any polynomial in  $1/n$ . In fact, if  $n$  is even and  $n \geq 2m$ ,  $g_n(U) = 0$  (exactly) for all  $U$ . In this case, this function is integrated exactly, regardless of the shift.  $\square$

It is interesting to note that an integrand  $f$  with  $f(0) \neq f(1)$  can be transformed into an integrand  $\tilde{f}$  having the same integral and with  $\tilde{f}(0) = \tilde{f}(1)$  (a periodic function), by defining  $\tilde{f}(1 - u) = \tilde{f}(u) = f(2u)$  for  $0 \leq u \leq 1/2$ . With this, we can switch from case (i) to case (ii) in Proposition 1, i.e., from  $O(n^{-1})$  to  $O(n^{-2})$  convergence for  $g_n(U)$ . The resulting function  $\tilde{f}$  also symmetric with respect to  $u = 1/2$ . This transformation is equivalently achieved by keeping  $f$  unchanged and transforming the randomized points  $U_i$  via  $\tilde{U}_i = 2U_i$  if  $U_i < 1/2$  and  $\tilde{U}_i = 2(1 - U_i)$  if  $U_i \geq 1/2$ . This is called the baker's transformation; it stretches the points  $U_i$  by a factor of two and then folds back those that exceed 1. After applying this transformation, the lattice points become locally antithetic in each interval of the form  $[i/n, (i+2)/n]$  if  $n$  and  $i$  are even, in the sense that they are at equal distance from the center of the interval, on each side. As a result, they integrate exactly any linear function over this interval. This holds for every such interval, so a piecewise-linear approximation which is linear over each interval is integrated exactly.

**Example 5** We apply the baker's transformation to the randomly-shifted points to integrate  $f(u) = u^2$ , the same function as in Example 1. The transformed function is  $\tilde{f}(u) = 4u^2$  for  $u \leq 1/2$  and  $\tilde{f}(u) = 4(1 - u)^2$  for  $u > 1/2$ . We

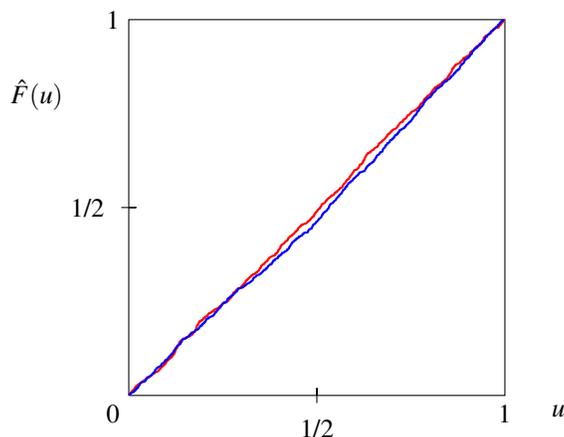


Figure 4: Empirical distribution of 1000 replicates of  $2[n^2g_n(U)/4 + 1/12]^{1/2}$  for lattices with 16 and 1024 points ( $n$  is even), for  $f(u) = u^2$  with the baker's transformation.

obtain  $\int_0^1 \tilde{f}'(u)du = 0$ , but  $\int_0^1 \tilde{f}''(u)du = 8 \neq f'(1) - f'(0)$ , which can be explained by the fact that  $\tilde{f}'$  is discontinuous at  $u = 1/2$ . If  $n$  is even, this point of discontinuity lies at the boundary between two intervals, and the result of case (ii) of Proposition 1 still stands with  $f'(1) - f'(0)$  replaced by 8. That is,  $n^2g_n(U)/4 + 1/12$  converges to  $(U - 1/2)^2$ , and then  $2[n^2g_n(U)/4 + 1/12]^{1/2}$  is approximately  $U(0, 1)$ . Figure 4 shows the empirical distribution of 1000 independent replications of  $2[n^2g_n(U)/4 + 1/12]^{1/2}$  for one-dimensional lattices with  $n = 16$  and  $n = 1024$  points. If  $n$  is odd, then it can be shown that the middle interval contributes another non-negligible quantity to  $n^2g_n(U)/4$ .  $\square$

The integrands  $f$  encountered in practice are often discontinuous, and sometimes they are also unbounded and have unbounded derivatives. The next examples give simple illustration of when can happen in these cases, in one dimension.

**Example 6** Consider the step function  $f(u) = 0$  for  $u < a$  and  $f(u) = 1$  for  $u \geq a$ , for some constant  $a \in (0, 1)$ . For a given  $n$ , let  $\delta(n) = \lceil na \rceil/n - a$ . The integration error with a randomly shifted lattice rule is then  $g_n(U) = -\delta(n)$  if  $U/n < 1/n - \delta(n)$ , and  $g_n(U) = 1/n - \delta(n)$  if  $U/n \geq 1/n - \delta(n)$ , where  $U$  is  $U(0, 1)$ . Thus, the error is distributed over only two possible values, and the variance is  $O(n^{-2})$ , unless  $a$  is a multiple of  $1/n$ , in which case there is no error.

More generally, if  $f$  is twice continuously differentiable, except at a single point  $a$  where it has a jump. Then  $f$  can be written as  $f = f_j + f_c$ , where  $f_j$  is a step function as above and  $f_c$  is a smooth function that obeys Proposition 1. The error  $g_n(U)$  is then the sum of two terms: an  $O(1/n)$  error distributed over two values as above, and another term obtained by applying Proposition 1 to  $f_c$ .  $\square$

**Example 7** We now consider the unbounded function  $f(u) = g(F^{-1}(u))$  where  $F$  is the distribution function of a random variable with infinite support  $[0, \infty)$ . Let  $n_0$  be a (fixed) large integer and let  $b = 1 - 1/n_0$ . Suppose that  $n \rightarrow \infty$  while  $n$  remains a multiple of  $n_0$ . The error for the integral over the interval  $[0, b]$  has a distribution that obeys a modified version of Proposition 1, with  $f^{(k-1)}(1)$  replaced by  $f^{(k-1)}(b)$ , and  $n$  replaced by  $n - n/n_0$ . Over the last interval  $[b, 1)$ , the distribution is more complicated, and the error is unbounded. In fact, the error over the interval  $[(n-1)/n, 1)$  is distributed as a random variable generated from the tail of the distribution  $F$ , so the total error will have a right tail that resembles the right tail of  $F$ . It will not be uniform.  $\square$

## 5 ERROR DISTRIBUTION FOR MULTI-DIMENSIONAL INTEGRALS

In  $s$  dimensions, recall that the Fourier expansion of  $g_n(\mathbf{u})$ , for an arbitrary shift  $\mathbf{u} \in [0, 1)^s$ , is

$$g_n(\mathbf{u}) = \sum_{\mathbf{0} \neq \mathbf{h} \in L_s^*} \hat{f}(\mathbf{h}) \exp(2\pi i \mathbf{h} \cdot \mathbf{u}). \tag{12}$$

Let  $\mathbf{w}_1, \dots, \mathbf{w}_s$  be a basis of the dual lattice  $L_s^*$  and let  $\mathbf{W}$  be an  $s \times s$  matrix with columns  $\mathbf{w}_1, \dots, \mathbf{w}_s$ . Note that all these vectors have integer coordinates. Thus,  $L_s^*$  is the set of all integer vectors  $\mathbf{h}$  that can be written as

$$\mathbf{h} = \mathbf{W}\mathbf{z} = \sum_{j=1}^s z_j \mathbf{w}_j$$

for  $\mathbf{z} = (z_1, \dots, z_s)^t \in \mathbb{Z}^s$ . By summing over the vectors  $\mathbf{z}$  instead of the vectors  $\mathbf{h}$  (a change of variable), we can rewrite

$$g_n(\mathbf{u}) = \sum_{\mathbf{0} \neq \mathbf{z} \in \mathbb{Z}^s} \hat{f}(\mathbf{W}\mathbf{z}) \exp(2\pi i \mathbf{u}^t \mathbf{W}\mathbf{z}) = \sum_{\mathbf{0} \neq \mathbf{z} \in \mathbb{Z}^s} \hat{v}_n(\mathbf{z}) \exp(2\pi i \mathbf{u}^t \mathbf{W}\mathbf{z}) = v_n(\mathbf{u}^t \mathbf{W}) \tag{13}$$

for some function  $v_n$  whose Fourier coefficients are  $\hat{v}_n(\mathbf{z}) = \hat{f}(\mathbf{W}\mathbf{z})$ . This function  $v_n$  depends on  $n$  and on the selected lattice (via  $\mathbf{W}$ ).

If we can recover  $v_n$  explicitly from its Fourier coefficients, this would give an explicit formula for  $g_n(\mathbf{u})$ , the error as a function of the shift. If this  $g_n$  turns out to be not too complicated (which could occur in very simple situations), we may also be able to approximate its distribution. This remains to be explored and studied. For now, we give some small examples.

**Example 8** In one dimension, the dual lattice contains all multiples of  $n$ . Thus, the matrix  $\mathbf{W}$  is one by one and contains the single entry  $n$ . The Fourier coefficients of  $v_n$  are then  $\hat{v}_n(z) = \hat{f}(nz)$ , for  $z \in \mathbb{Z}$ , and we have  $g_n(u) = v_n(nu)$ .

As a special case, if  $B_k$  is the  $k$ th Bernoulli polynomial, then its Fourier expansion is

$$B_k(u) = \sum_{0 \neq h \in \mathbb{Z}} \frac{-k!}{(2\pi i h)^k} e^{2\pi i h u}$$

for  $u \in (0, 1)$ , and we have

$$\hat{v}_n(z) = \frac{-k!}{(2\pi i n z)^k} = \hat{B}_k(z) n^{-k}.$$

We can then conclude that  $v_n(u) = n^{-k} B_k(u)$  and  $g_n(u) = v_n(nu) = n^{-k} B_k(nu \bmod 1)$  for all  $u \in (0, 1)$ . But if  $U$  is  $U(0, 1)$ , then  $nU \bmod 1$  is also  $U(0, 1)$ , and therefore  $n^k g_n(U)$  has exactly the same distribution as  $B_k(U)$ . This agrees with Proposition 1.  $\square$

**Example 9** Suppose  $f : [0, 1]^s \rightarrow \mathbb{R}$  is a linear combination of  $s$  Bernoulli polynomials, of the form

$$f(\mathbf{u}) = f(u_1, \dots, u_s) = \sum_{j=1}^s a_j B_{m_j}(u_j)$$

for some arbitrary real numbers  $a_j$ . This is just a sum of one-dimensional functions, and the total error  $g_n(\mathbf{U})$  is the sum of the integration errors made in approximating the integrals of these one-dimensional functions. From the previous example,  $g_n(\mathbf{U})$  has exactly the same distribution as  $\sum_{j=1}^s a_j n^{-m_j} B_{m_j}(U_j)$ , where  $\mathbf{U} = (U_1, \dots, U_n)$ .

The case where  $f$  is a linear function corresponds to the special case where  $m_j = 1$  for all  $j$ . Then,  $ng_n(\mathbf{U})$  has the same distribution as  $W = \sum_{j=1}^s a_j (U_j - 1/2)$ , which is a linear combination of independent uniform random variables. Theorem 1 of Barrow and Smith (1979) tells us that the (exact) cumulative distribution function of  $W$  is a non-decreasing spline of degree  $s$  with  $s - 1$  continuous derivatives, and its support is the finite interval  $[-b, b]$ , where  $b = \sum_{j=1}^s a_j / 2$ . This distribution function is exactly

$$F(w) = \mathbb{P}[W \leq w] = \frac{1}{s! \prod_{j=1}^s a_j} \sum_{J \subseteq \{1, \dots, s\}} (-1)^{|J|} \left[ \max \left( 0, w + b - \sum_{j \in J} a_j \right) \right]^s \tag{14}$$

for  $-b \leq w \leq b$ . Its density is given by a spline of degree  $s - 1$ .  $\square$

**Example 10** As a concrete illustration of the previous example, let  $s = 2$  and  $f(u_1, u_2) = 5u_1 + 3u_2 - 4 = 5B_1(u_1) + 3B_1(u_2)$ . Note that the constant  $-4$  has no effect on the error, so we can forget it. Here, we have  $b = 4$  and the distribution

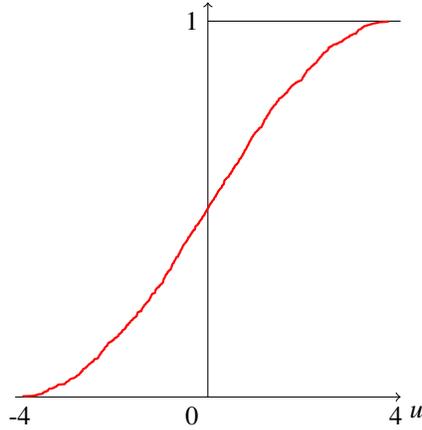


Figure 5: Empirical distribution  $\hat{F}$  of 1000 replicates of  $ng_n(U_1, U_2)$  for two-dimensional lattices with  $n = 1024$  points, for  $f(u_1, u_2) = 5u_1 + 3u_2$ .

function  $F_w$  of  $ng_n(\mathbf{U}) = ng_n(U_1, U_2)$  is a spline of degree 2 defined by

$$F_w(w) = \begin{cases} (w+4)^2/30 & \text{for } -4 \leq w \leq -1, \\ (6w+15)/30 & \text{for } -1 \leq w \leq 1, \\ -(w+1)^2/30 + (w-1)/3 + 25/30 & \text{for } 1 < w \leq 4. \end{cases}$$

The corresponding density is a “table mountain” function, defined by

$$f_w(w) = \begin{cases} (4+w)/15 & \text{for } -4 \leq w \leq -1, \\ 1/5 & \text{for } -1 \leq w \leq 1, \\ (4-w)/15 & \text{for } 1 < w \leq 4. \end{cases}$$

Figure 5 displays the empirical distribution  $\hat{F}$  of  $ng_n(U_1, U_2)$  obtained from 1000 independent replications of a randomly-shifted lattice rule with  $n = 1024$ ,  $\mathbf{v}_1 = (1/n, 271/n)$ , and  $\mathbf{v}_2 = (0, 1)$ . In Figure 6, we plot the empirical distribution of the 1000 replicates of  $F_w(ng_n(U_1, U_2))$ , which really looks uniform over  $(0, 1)$ , as expected.  $\square$

**Example 11** Suppose now that  $f: [0, 1]^s \rightarrow \mathbb{R}$  is a product of Bernoulli polynomials as follows:

$$f(\mathbf{u}) = f(u_1, \dots, u_s) = \prod_{j=1}^s B_{m_j}(u_j).$$

Its Fourier expansion is then

$$f(\mathbf{u}) = \sum_{\mathbf{0} \neq \mathbf{h} \in \mathbb{Z}^s} \prod_{j=1}^s \frac{-m_j!}{(2\pi i h_j)^{m_j}} e^{2\pi i h_j u_j} = \frac{(-1)^s m_1! \cdots m_s!}{(2\pi i)^{m_1 + \cdots + m_s}} \sum_{\mathbf{0} \neq \mathbf{h} \in \mathbb{Z}^s} \frac{1}{h_1^{m_1} \cdots h_s^{m_s}} e^{2\pi i \mathbf{h}^t \mathbf{u}}$$

for  $\mathbf{u} = (u_1, \dots, u_s)^t \in (0, 1)^s$ . In this case, the Fourier coefficient  $\hat{v}_n(\mathbf{z}) = \hat{f}(\mathbf{W}\mathbf{z})$  can be written as

$$\hat{v}_n(\mathbf{z}) = \prod_{j=1}^s \frac{-m_j!}{(2\pi i)^{m_j} (\mathbf{w}^{(j)} \mathbf{z})^{m_j}},$$

where  $\mathbf{w}^{(j)}$  is the  $j$ -th row of the matrix  $\mathbf{W}$ . Here,  $(\mathbf{w}^{(j)} \mathbf{z})^{m_j}$  is a multivariate polynomial in  $\mathbf{z}$ .  $\square$

**Example 12** Let  $s = 2$  and  $f(u_1, u_2) = B_1(u_1)B_1(u_2) = (u_1 - 1/2)(u_2 - 1/2)$ . We do not have an exact expression for the distribution of  $g_n(\mathbf{U})$  in this case, but Figure 7 shows the empirical distribution obtained with the two-dimensional

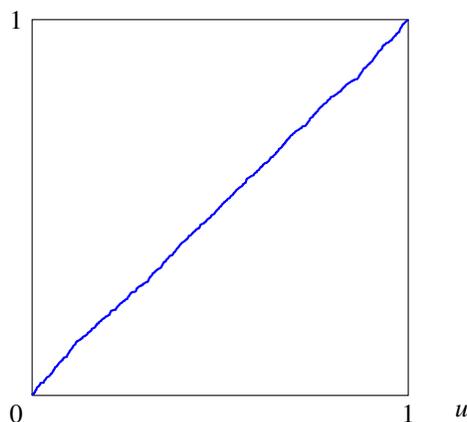


Figure 6: Empirical distribution  $\hat{F}$  of 1000 replicates of  $F_w(n g_n(U_1, U_2))$  for two-dimensional lattices with  $n = 1024$  points, for  $f(u_1, u_2) = 5u_1 + 3u_2$ .

lattice with  $n = 8192$ ,  $\mathbf{v}_1 = (1/n, 1530/n)$ , and  $\mathbf{v}_2 = (0, 1)$ . It resembles a spline of degree 2. This seems to indicate that the error is dominated by some linear combination of  $U_1$  and  $U_2$ .  $\square$

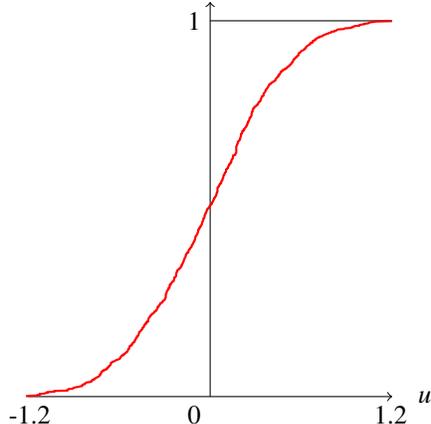


Figure 7: Empirical distribution  $\hat{F}$  of 1000 replicates of  $n g_n(U_1, U_2)$  for two-dimensional lattices with  $n = 8192$  points, for  $f(u_1, u_2) = B_1(u_1)B_1(u_2) = (u_1 - 1/2)(u_2 - 1/2)$ .

## 6 CONCLUSION

We made a few steps toward understanding the limiting distributions of RQMC estimators obtained via a randomly-shifted lattice rule. The examples analyzed in this paper are admittedly very simple. Further analysis and empirical exploration for higher-dimensional functions is certainly needed. Other RQMC methods than lattice rules also need examination. For example, what we did for randomly-shifted lattices using Fourier expansions in this paper can be transposed to digital nets with random digital shifts, if we replace the Fourier expansions by Walsh expansions, and using the correspondence explained in L'Ecuyer (2004), for example.

## ACKNOWLEDGMENTS

This work has been supported by an NSERC-Canada Discovery Grant and a Canada Research Chair to the first author, INRIA's associated team MOCQUASIN to both authors, and the EuroNF Network of Excellence to the second author. David Munger helped to correct some mistakes in an earlier version.

## REFERENCES

- Abramowitz, M., and I. A. Stegun. 1970. *Handbook of mathematical functions*. New York: Dover.
- Asmussen, S., and P. W. Glynn. 2007. *Stochastic simulation*. New York: Springer-Verlag.
- Barrow, D. L., and P. W. Smith. 1979. Spline notation applied to a volume problem. *The American Mathematical Monthly* 86:50–51.
- Cranley, R., and T. N. L. Patterson. 1976. Randomization of number theoretic methods for multiple integration. *SIAM Journal on Numerical Analysis* 13 (6): 904–914.
- Dahlquist, G., and A. Bjorck. 2008. *Numerical methods in scientific computing*. Philadelphia: Society for Industrial and Applied Mathematics.
- Davis, P., and P. Rabinowitz. 1984. *Methods of numerical integration*. second ed. New York: Academic Press.
- L'Ecuyer, P. 2004. Polynomial integration lattices. In *Monte Carlo and Quasi-Monte Carlo Methods 2002*, ed. H. Niederreiter, 73–98. Berlin: Springer-Verlag.
- L'Ecuyer, P. 2009. Quasi-Monte Carlo methods with applications in finance. *Finance and Stochastics*. To appear.
- L'Ecuyer, P., and C. Lemieux. 2000. Variance reduction via lattice rules. *Management Science* 46 (9): 1214–1235.
- L'Ecuyer, P., and C. Lemieux. 2002. Recent advances in randomized quasi-Monte Carlo methods. In *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, ed. M. Dror, P. L'Ecuyer, and F. Szidarovszky, 419–474. Boston: Kluwer Academic.
- Lemieux, C., and P. L'Ecuyer. 2000. A comparison of Monte Carlo, lattice rules and other low-discrepancy point sets. In *Monte Carlo and Quasi-Monte Carlo Methods 1998*, ed. H. Niederreiter and J. Spanier, 326–340. Berlin: Springer-Verlag.
- Loh, W.-L. 2003. On the asymptotic distribution of scramble nets quadratures. *Annals of Statistics* 31:1282–1324.
- Loh, W.-L. 2005. On the asymptotic distribution of some randomized quadrature rules. In *Stein's Method and Applications*, ed. C. Stein, A. D. Barbour, and L. H. Y. Chen, Volume 5, 209–222. World Scientific.
- Niederreiter, H. 1992. *Random number generation and quasi-Monte Carlo methods*, Volume 63 of *SIAM CBMS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia, PA: SIAM.
- Owen, A. B. 1992. A central limit theorem for Latin hypercube sampling. *Journal of the Royal Statistical Society B* 54 (2): 541–551.
- Sloan, I. H., and S. Joe. 1994. *Lattice methods for multiple integration*. Oxford: Clarendon Press.
- Tuffin, B. 1998. Variance reduction order using good lattice points in Monte Carlo methods. *Computing* 61:371–378.

## AUTHOR BIOGRAPHIES

**PIERRE L'ECUYER** is Professor in the Département d'Informatique et de Recherche Opérationnelle, at the Université de Montréal, Canada. He holds the Canada Research Chair in Stochastic Simulation and Optimization. He is a member of the CIR-RELT and GERAD research centers. His main research interests are random number generation, quasi-Monte Carlo methods, efficiency improvement via variance reduction, sensitivity analysis and optimization of discrete-event stochastic systems, and discrete-event simulation in general. He is currently Associate/Area Editor for *ACM Transactions on Modeling and Computer Simulation*, *ACM Transactions on Mathematical Software*, *Statistics and Computing*, *Management Science*, *International Transactions in Operational Research*, *The Open Applied Mathematics Journal*, and *Cryptography and Communications*. He obtained the *E. W. R. Steacie* fellowship in 1995-97, a *Killam* fellowship in 2001-03, and became an INFORMS Fellow in 2006. His recent research articles are available on-line from his web page: <http://www.iro.umontreal.ca/~lecuyer>.

**BRUNO TUFFIN** received his PhD degree in applied mathematics from the University of Rennes 1 (France) in 1997. Since then, he has been with INRIA in Rennes. He spent 8 months as a postdoc at Duke University in 1999. His research interests include developing Monte Carlo and quasi-Monte Carlo simulation techniques for the performance evaluation of telecommunication systems, and developing new Internet-pricing schemes. He is currently Associate Editor for *INFORMS Journal on Computing* and *Mathematical Methods of Operations Research*.