Electronic Journal of Statistics
Vol. 3 (2009) 1531–1541
ISSN: 1935-7524
DOI: 10.1214/09-EJS541

Recycling physical random numbers

Art B. Owen[*]

Stanford University
e-mail: owen@stat.stanford.edu

Abstract: Physical random numbers are not as widely used in Monte Carlo integration as pseudo-random numbers are. They are inconvenient for many reasons. If we want to generate them on the fly, then they may be slow. When we want reproducible results from them, we need a lot of storage. This paper shows that we may construct $N = n(n-1)/2$ pairwise independent random vectors from n independent ones, by summing them modulo 1 in pairs. As a consequence, the storage and speed problems of physical random numbers can be greatly mitigated. The new vectors lead to Monte Carlo averages with the same mean and variance as if we had used N independent vectors. The asymptotic distribution of the sample mean has a surprising feature: it is always symmetric, but never Gaussian. This follows by writing the sample mean as a degenerate U-statistic whose kernel is a left-circulant matrix. Because of the symmetry, a small number B of replicates can be used to get confidence intervals based on the central limit theorem.

Received November 2009.

1. Introduction

When it comes to Monte Carlo simulation, physically based random numbers are the poor cousin of pseudo-random numbers. Most physical random number generators have comparatively cumbersome interfaces. Some of them are slow, although http://true-random.com/ is fast. But all of them have problems with reproducibility. To reproduce a simulation with physical random numbers, we would need to store them. Because truly random numbers cannot be compressed, the storage required could be very large. See L'Ecuyer (2009) for a survey of random and pseudo-random number generation. Because of these well-known shortcomings, the great majority of simulations take place with pseudo-random numbers. Physical random numbers do have their place however. They are still used in a small percentage of Monte Carlo applications, and there is a market for devices that produce them. For example, when we are concerned that a flaw in the pseudo-random number generator might interact with a feature of the problem, we can replace the pseudo-random numbers by physically random ones and rerun the example.

This article investigates a strategy to mitigate the storage disadvantage of physical random numbers, by summing pairs of (vectors of) uniform random numbers modulo 1. In this way, n physical random inputs can be used to get answers comparable to what we would get from $N = n(n-1)/2$ independent

[*] Supported by grant number DMS-0906056 from the U.S. National Science Foundation.

random inputs. The CPU cost is still $n(n-1)/2$ function evaluations, but storage requirements are greatly reduced, and we end up with reproducibility for physical random numbers.

We suppose that the random numbers are being used for Monte Carlo integration, as follows. There is a function f defined on $[0,1)^d$, and we seek to approximate the integral $\mu = \int_{[0,1)^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. We will assume that $f(\boldsymbol{x}) \in \mathbb{R}$. Extensions to vector valued f are straightforward. As written, $\mu = \mathbb{E}(f(\boldsymbol{X}))$ for $\boldsymbol{X} \sim \mathbf{U}[0,1)^d$. Many expectations of functions of non-uniform random variables on the unit cube and other domains, can be cast into this framework, by techniques described in Devroye (1986). We assume that $\sigma^2 = \mathrm{Var}(f(\boldsymbol{X})) < \infty$.

Forming all pairwise sums of n independent $\mathbf{U}[0,1)^d$ random variables and taking their remainder modulo 1, yields $N = \binom{n}{2}$ composite random vectors $\boldsymbol{X}_1, \dots, \boldsymbol{X}_N$. The statistic we use is $\bar{Y} = (1/N)\sum_{i=1}^N Y_i$ for $Y_i = f(\boldsymbol{X}_i)$.

Section 2 gives more details about the construction. Section 3 gives basic statistical properties of this method. The combined inputs are pairwise independent from $\mathbf{U}[0,1)^d$. It follows that $\mathbb{E}(\bar{Y}) = \mu$, $\mathrm{Var}(\bar{Y}) = \sigma^2/N$, and the usual variance estimate s^2 satisfies $\mathbb{E}(s^2) = \sigma^2$. The estimate $\bar{Y}$ is a degenerate U statistic whose asymptotic distribution is that of a weighted sum of centered independent $\chi^2_{(1)}$ random variables.

Section 4 makes a small empirical comparison of IID sampling versus pairwise and three-fold combinations. A surprising symmetry turns up in the QQ plots of the examples even for a lognormally distributed $f(X)$. Section 5 shows that this symmetry is not special to the lognormal distribution, but can instead be explained via recent results in the spectra of circulant matrices. Section 6 gives conclusions.

2. Notation

For $U, V \in [0,1)$, their sum modulo 1 is

$$U \oplus V = U + V - \lfloor U + V \rfloor,$$

where $\lfloor z \rfloor$ is the greatest integer less than or equal to $z \in \mathbb{R}$. For $\boldsymbol{U}, \boldsymbol{V} \in [0,1)^d$ define $\boldsymbol{U} \oplus \boldsymbol{V} \in [0,1)^d$ componentwise.

Given $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n \overset{\text{ind}}{\sim} \mathbf{U}[0,1)^d$ sampled from a source of random numbers, we want to obtain all of their pairwise sums modulo 1. A convenient iteration for that purpose is

```
$i \leftarrow 0$
for $r = 1, \dots, n-1$
for $s = r+1, \dots, n$
   $i \leftarrow i + 1$
   $\boldsymbol{X}_i \leftarrow \boldsymbol{U}_r \oplus \boldsymbol{U}_s$
end double for loop
```
(1)

We ordinarily use $\boldsymbol{X}_i$ right after it is generated, so we only have to store $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$. We will not need an explicit expression for i in terms of r and s, or for r and s in terms of i.

More generally, for $2 \le m \le n$ we can form $N_m = \binom{n}{m}$ points $\boldsymbol{X}_1, \dots, \boldsymbol{X}_{N_m}$ by summing all distinct m-tuples of $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$, modulo 1. It is easy to generalize (1) to triple and higher order sums. Ordinary IID sampling corresponds to $m = 1$.

3. Statistical properties

Here we give basic statistical properties for the pairwise recycled uniform vectors. Proposition 1 shows that the combined Monte Carlo inputs are pairwise independent. Then Proposition 2 shows how this suffices to get the low order moments right.

Proposition 1. *For dimension $d \ge 1$ and $n \ge m \ge 1$, let $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ be IID $\mathbf{U}[0,1)^d$ random variables. Suppose that $\boldsymbol{X}_i$ for $i = 1, \dots, N_m$ comprise all $N_m = \binom{n}{m}$ distinct sums of the form $\boldsymbol{U}_{r_1} \oplus \boldsymbol{U}_{r_2} \oplus \cdots \oplus \boldsymbol{U}_{r_m}$ for $1 \le r_1 < r_2 < \cdots < r_m \le n$. Then $\boldsymbol{X}_i$ are pairwise independent.*

Proof. Write $\boldsymbol{X}_i = \oplus_{k=1}^m \boldsymbol{U}_{r_k(i)}$ and $\boldsymbol{X}_j = \oplus_{k=1}^m \boldsymbol{U}_{r_k(j)}$. Assume that $i \ne j$. Then $\boldsymbol{X}_i$ contains at least one summand $\boldsymbol{U}_r$ that is not used in $\boldsymbol{X}_j$. Without loss of generality suppose that it is $\boldsymbol{U}_{r_m(i)}$. The distribution of $\boldsymbol{X}_i = \oplus_{k=1}^m \boldsymbol{U}_{r_k(i)}$ given $\boldsymbol{U}_{r_1(i)}, \dots, \boldsymbol{U}_{r_{m-1}(i)}$ is $\mathbf{U}[0,1)^d$. Let A and B be Borel subsets of $[0,1)^d$. Then

$$\begin{aligned}
\Pr(\boldsymbol{X}_i \in A, \boldsymbol{X}_j \in B) &= \mathbb{E}\bigl(\Pr(\boldsymbol{X}_i \in A, \boldsymbol{X}_j \in B \mid \boldsymbol{U}_s, \forall s \ne r_m(i))\bigr) \\
&= \mathbb{E}\bigl(\mathbf{1}_{\oplus_{k=1}^m \boldsymbol{U}_{r_k(j)} \in B} \Pr(\boldsymbol{X}_i \in A \mid \boldsymbol{U}_s, \forall s \ne r_m(i))\bigr) \\
&= \mathbb{E}\bigl(\mathbf{1}_{\boldsymbol{X}_j \in B}\, \mathbf{vol}(A)\bigr) \\
&= \Pr(\boldsymbol{X}_i \in A) \Pr(\boldsymbol{X}_j \in B). \square
\end{aligned}$$

The random variables $\boldsymbol{X}_1, \dots, \boldsymbol{X}_N$ are pairwise independent, but for $m > 1$ they are not generally independent. For example

$$(\boldsymbol{U}_1 \oplus \boldsymbol{U}_2) + (\boldsymbol{U}_3 \oplus \boldsymbol{U}_4) - (\boldsymbol{U}_1 \oplus \boldsymbol{U}_3) - (\boldsymbol{U}_2 \oplus \boldsymbol{U}_4)$$

is always a vector of integers.

Pairwise independent random variables satisfy many of the key properties we need in Monte Carlo integration. Suppose that $Y = f(\boldsymbol{X})$ for $\boldsymbol{X} \sim \mathbf{U}[0,1)^d$. Let $\mu = \mathbb{E}(Y)$ and suppose that $\sigma^2 = \mathrm{Var}(Y) < \infty$. From pairwise independent random vectors $\boldsymbol{X}_1, \dots, \boldsymbol{X}_N \sim \mathbf{U}[0,1)^d$ we get pairwise independent $Y_i = f(\boldsymbol{X}_i)$. These have the right low order moments as shown next.

Proposition 2. *For $N \ge 2$, let $Y_1, \dots, Y_N$ be pairwise independent random variables with common mean μ and common variance $\sigma^2 < \infty$. Define $\bar{Y} = \frac{1}{N}\sum_{i=1}^N Y_i$ and $s^2 = \frac{1}{N-1}\sum_{i=1}^N (Y_i - \bar{Y})^2$. Then*

$$\mathbb{E}(\bar{Y}) = \mu, \quad \mathrm{Var}(\bar{Y}) = \frac{\sigma^2}{N}, \quad \textit{and} \quad \mathbb{E}(s^2) = \sigma^2.$$

Proof. The first part is obvious, by linearity of expectations. The second and third parts follow easily because $\mathbb{E}(Y_1Y_2) = \mathbb{E}(Y_1)\mathbb{E}(Y_2)$ for pairwise independent random variables Y_1 and Y_2. □

Proposition 2 shows that we can use N pairwise independent random variables to get unbiased Monte Carlo estimates with the same variance as with N fully independent random variables. Furthermore we can get an unbiased estimate of that variance.

Usually in a Monte Carlo integration problem we ask for more than the moment properties in Proposition 2. To get an asymptotic confidence interval for μ, we want a central limit theorem. Sequences of pairwise independent random variables do not always satisfy a central limit theorem, even when the individual variables are identically distributed and have finite variance. For an extreme counterexample, see Romano and Siegel (1986, Chapter 5) who construct 2^n pairwise independent random bits from n independent ones.

Suppose that $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n \overset{\text{ind}}{\sim} \mathbf{U}[0,1)^d$ and $\boldsymbol{X}_1, \dots, \boldsymbol{X}_{N_m}$ are constructed as described above. Let $Y = f(\boldsymbol{X})$, where $f \in L^2[0,1)^d$. The estimate

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} f(\boldsymbol{X}_i)$$

can be written as the U-statistic

$$\bar{Y} = \frac{1}{\binom{n}{m}} \sum_{1\le r_1<r_2<\cdots<r_m\le n} \Psi(\boldsymbol{U}_{r_1}, \boldsymbol{U}_{r_2}, \dots, \boldsymbol{U}_{r_m}), \tag{2}$$

where

$$\Psi(\boldsymbol{U}_{r_1}, \boldsymbol{U}_{r_2}, \dots, \boldsymbol{U}_{r_m}) = f(\boldsymbol{U}_{r_1} \oplus \boldsymbol{U}_{r_2} \oplus \cdots \oplus \boldsymbol{U}_{r_m}).$$

Hoeffding (1948) gives a central limit theorem for U-statistics. In this setting $\sqrt{n}(\bar{Y} - \mu) \to \mathcal{N}(0, \tau^2)$ but here $\sqrt{n}\mathrm{Var}(\bar{Y}) = \sigma^2\sqrt{n/N}$ and so the limit has $\tau^2 = 0$. This U-statistic is degenerate and the central limit theorem for it does not help us set confidence intervals.

The limiting distribution for degenerate U-statistics of order $m = 2$ is a weighted sum of independent centered chisquares. It uses the eigenvalues λ_j of $\Psi(\cdot,\cdot) - \mu$ where an eigenvalue-eigenfunction pair (λ, g) satisfies

$$\int_{[0,1)^d} (f(\boldsymbol{V} \oplus \boldsymbol{U}) - \mu) g(\boldsymbol{U})\,\mathrm{d}\boldsymbol{U} = \lambda g(\boldsymbol{V}).$$

The function $g(\boldsymbol{V}) = 1$ is an eigenfunction with eigenvalue 0. By convention we will call this eigenpair (λ_0, g_0).

Theorem 1. *For integer $d \ge 1$ and $r = 1, \dots, n$, let $\boldsymbol{U}_r \overset{\text{ind}}{\sim} \mathbf{U}[0,1)^d$. For $f \in L^2[0,1)^d$, let $\mu = \mathbb{E}(f(\boldsymbol{U}_1))$, assume $\sigma^2 = \mathbb{E}((f(\boldsymbol{U}_1) - \mu)^2) > 0$, and let $\bar{Y} = \binom{n}{2}^{-1}\sum_{r=1}^{n-1}\sum_{s=r+1}^{n} f(\boldsymbol{U}_r \oplus \boldsymbol{U}_s)$. Then*

$$n(\bar{Y} - \mu) \xrightarrow{d} \sum_{j=1}^{\infty} \lambda_j (Z_j^2 - 1) \tag{3}$$

as $n \to \infty$ *where* $Z_j \overset{\text{ind}}{\sim} \mathcal{N}(0,1)$*, where* λ_j *for* $j \geq 0$ *are the eigenvalues of* $\Psi(\cdot,\cdot) = f(\cdot \oplus \cdot) - \mu$*.*

Proof. See Gregory (1977) or Serfling (1980, Chapter 5.5). □

Notice that the sum in (3) does not include the zeroth eigenvalue. Combining (3) with $N = n(n-1)/2 = (n^2/2)(1+o(1))$ find $\sqrt{N}(\bar{Y} - \mu) \xrightarrow{d} \sqrt{2}\sum_{j=1}^{\infty} \lambda_j(Z_j^2 - 1)$. Therefore

$$\sigma^2 = 4\sum_{j=1}^{\infty} \lambda_j^2$$

because $\mathrm{Var}(Z_j^2) = 2$.

The degree of nonnormality in the limiting distribution (3) is quite mild. The skewness of a weighted sum of independent $\chi^2_{(1)}$ random variables must be between $-\sqrt{8}$ and $\sqrt{8}$ because $\sqrt{8}$ is the skewness of the $\chi^2_{(1)}$ distribution. The kurtosis of that weighted sum must be between 0 and 12, the kurtosis of $\chi^2_{(1)}$.

Degenerate U-statistics for $m \geq 3$ do not satisfy a central limit theorem either. See Arcones and Giné (1993).

The lack of a central limit theorem is easily mended. We take B independent replicates of the whole process and average them. Specifically let $\boldsymbol{U}_{rb} \overset{\text{ind}}{\sim} \mathbf{U}[0,1)^d$ for $r = 1,\dots,n$ and $b = 1,\dots,B$. Let $Y_{ib} = f(\oplus_{k=1}^{m} \boldsymbol{U}_{r_k(i)b})$ and then put

$$\bar{Y} = \frac{1}{B\binom{n}{m}} \sum_{b=1}^{B} \sum_{i=1}^{\binom{n}{m}} Y_{ib}$$

and

$$s^2 = \frac{1}{B\binom{n}{m} - 1} \sum_{b=1}^{B} \sum_{i=1}^{\binom{n}{m}} (Y_{ib} - \bar{Y})^2.$$

Then

$$\mathbb{E}(\bar{Y}) = \mu, \quad \mathrm{Var}(\bar{Y}) = \frac{\sigma^2}{B\binom{n}{m}}, \quad \text{and} \quad \mathbb{E}(s^2) = \sigma^2$$

by the same pairwise independence properties used in Proposition 2. We could also form B separate averages $\bar{Y}_b$ and take $\hat{\sigma}^2 = \frac{\binom{n}{m}}{B-1}\sum_{b=1}^{B}(\bar{Y}_b - \bar{Y})^2$, but this estimate would have only $B-1$ degrees of freedom.

Using B replicates we consume only nB uniform random vectors to obtain $B\binom{n}{m}$ pairwise independent ones. For $m = 2$, taking a fixed $n > 2000$ leads to at least a 1000 fold reduction in the number of independent random vectors that we need to store. For fixed n and N, we have a central limit theorem

$$\lim_{B\to\infty} \Pr\bigl(\sqrt{NB}(\bar{Y} - \mu)/\sigma \leq z\bigr) = \Phi(z).$$

4. Numerical comparison

Here we make a small numerical inspection of random vector recycling. It is convenient to compare the methods with $m = 2$ and $m = 3$ using the same value of N. There are only three values of N in the range $10 < N < 10^6$ that can be attained as both $\binom{n_2}{2}$ and $\binom{n_3}{3}$ for some integers n_2 and n_3. They are listed below:

n_2	n_3	N
16	10	120
56	22	1540
120	36	7140

To make the comparison we use $N = 1540$. This sample size is small enough to allow many replications. The distribution of $\bar{Y}$ may be sensitive to that of Y_i. Two quite different example distributions are used for Y_i. The first is the lognormal distribution, which we get via $f(X) = \exp(\Phi^{-1}(X))$ for $X \in [0, 1)$. The second is the uniform distribution which we get via $f(X) = X$. The mean, variance, skewness, and kurtosis of these two distributions are as follows:

	Uniform	Log normal
μ	1/2	1.649
σ^2	1/12	4.671
γ	0	4.874
κ	$-6/5$	110.936

The results of 10,000 independent replicated computations of $\bar{Y}$ from these methods are displayed in Figure 1. Taking $m = 1$ corresponds to sampling N independent values of Y_i. For $m = 1, 2, 3$ the distribution of $\bar{Y}$ is nearly normal in the center, but starts to depart in the tails, where it matters most. For $m = 2$ and 3 the distributions are nearly symmetric while for $m = 1$ and the log normal distribution, the distribution of $\bar{Y}$ retains some of the skewness of Y_i. As we might expect, the non-normality is more severe with $m = 3$ than with $m = 2$. Surprisingly, the non-normality is more severe for $Y_i \sim \mathbf{U}(0, 1)$ than for log normal Y_i.

Since $m = 3$ gives greater skewness, and is not covered by the sum of chi-squareds result in Theorem 1, we focus on the case $m = 2$, which should create enough pairwise independent vectors for applications. For $m = 2$ we group the 10,000 independent replicates into groups of $B = 4$ and $B = 10$. Figure 2 shows QQ plots for the averages of these replicates. Even B as small as 10 gives a very nearly normal distribution.

A QQ plot (not shown) for $\bar{Y}$ from 1,000 independent samples for $m = 2$ and $N = 7140$ was similarly symmetric, had slightly lighter tails than that for $N = 1540$, but was clearly not normal.

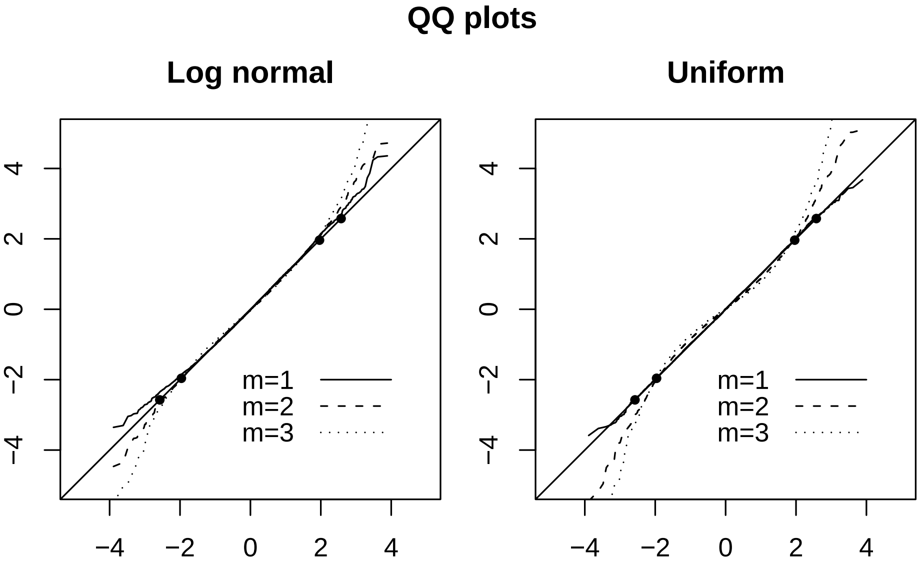

FIG 1. *QQ plots to show the effect of $m \in \{1,2,3\}$ in random vector recycling. The distribution $\sqrt{N}(\bar{Y}-\mu)/\sigma$ is plotted versus normal quantiles. All samples had $N=\binom{22}{3}=\binom{56}{2}=1540$ function evaluations. The 45 degree line is given as a reference. Points corresponding to 95% and 99% confidence interval endpoints are plotted on it.*

5. Symmetry of $\bar{Y}$

In the numerical examples, the QQ-plots of $\bar{Y}$ appeared to be symmetric, even when the test function was the log-normal inverse CDF $\exp(\Phi^{-1}(u))$. Such symmetry is consistent with the limit distribution (3) only when the eigenvalues, or at least the dominant ones, come in pairs of opposite sign. Symmetry, if it holds generally, is useful because it means that the central limit approximation will be more accurate for a small number B of replicates than it would otherwise be.

To investigate the eigenvalues we first consider a 1 dimensional problem with $\Psi(u_1,u_2)=f(u_1\oplus u_2)-\mu$ for $u_1,u_2\in[0,1)$ and $\mu=\int_0^1 f(u)\,\mathrm{d}u$. Without loss of generality, assume that $\mu=0$ for this section. The d dimensional case will be similar as remarked below.

Let G be a large odd integer and define $z_j=(j+1/2)/G$ for $j=0,\dots,G-1$. The values z_j are from a midpoint rule on $[0,1)$. We will use the approximation

$$\int_0^1 \Psi(v,u)g(u)\,\mathrm{d}u \doteq \frac{1}{G}\sum_{j=0}^{G-1}\Psi(v,z_j)g(z_j),$$

at points $v=z_k$ for $0\le k<G$. As a result, we study the eigenvalues of Ψ by looking at those of the $G\times G$ matrix $G^{-1}\Psi^G$ where

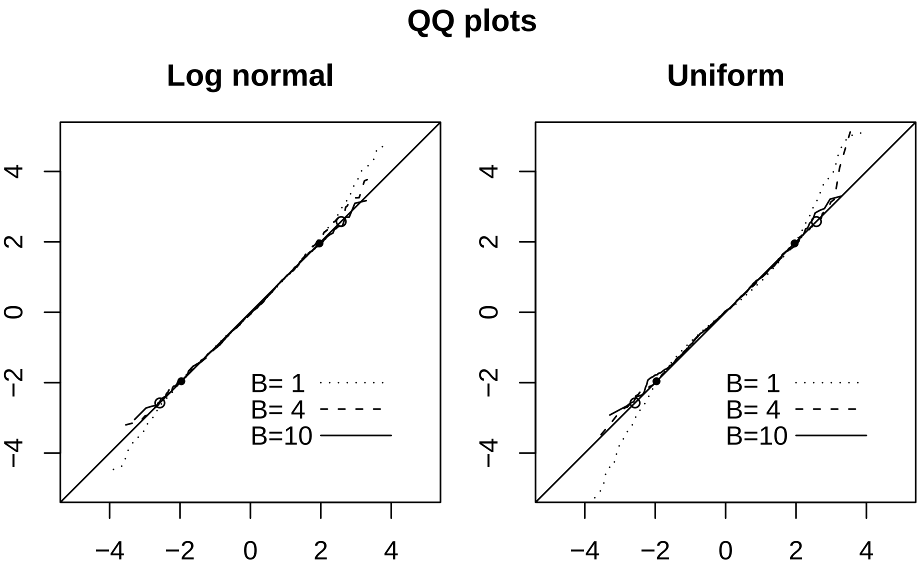

FIG 2. *QQ plots to show the effect of averaging* $B \in \{1, 4, 10\}$ *independent replicates of random vector recycling. The recycling scheme averaged* $1540 = \binom{56}{2}$ *pairwise independent vectors constructed from* $n = 56$ *independent ones. Here* $\sqrt{NB}(\bar{Y} - \mu)/\sigma$ *is plotted versus normal quantiles. The 45 degree line is given as a reference. Points corresponding to 95% and 99% confidence interval endpoints are plotted on it.*

$$\Psi^G_{jk} = f(z_j \oplus z_k) = f((j + k + 1)/G \text{ mod } 1) = f(z_0 \oplus z_{j+k \text{ mod } G}).$$

Let $a_j = f(z_0 \oplus z_j)$. Then

$$\Psi^G = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{G-1} \\ a_1 & a_2 & a_3 & \dots & a_0 \\ a_2 & a_3 & a_4 & \dots & a_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{G-1} & a_{G-2} & a_{G-3} & \dots & a_0 \end{pmatrix}$$

is a left-circulant matrix (Davis, 1979). Each row is the previous one shifted left one position with wraparound. The better known right-circulant matrices shift each row to the right, and are thus a subfamily of Toeplitz matrices.

The spectral decomposition of left-circulant matrices was recently found by Karner et al. (2003), giving a more explicit version of results in Davis (1979). We restate one of their results.

Theorem 2 (Theorem 3.6 of Karner et al. (2003)). *The eigenvalues of* $G^{-1}\Psi_G$ *are* $\lambda_0 = \frac{1}{G}\sum_{k=0}^{G-1} a_k$ *and*

$$\pm\Bigl|\frac{1}{G}\sum_{k=0}^{G-1} a_k \exp\bigl(-jk\sqrt{-1}/G\bigr)\Bigr|, \quad \mathit{for} \quad j = 1, \dots, (G-1)/2. \tag{4}$$

The sums in (4) are real because Ψ_G is a symmetric matrix. As a result, the eigenvalues of Ψ_G come as pairs of real numbers with opposite signs except for

$$a_0 = \frac{1}{G}\sum_{j=0}^{G-1} a_j = \frac{1}{G}\sum_{j=0}^{G-1} \Psi(z_0, z_j) \to \int_0^1 f(u)\,\mathrm{d}u = 0 \tag{5}$$

as $G \to \infty$. The limit in (5) holds if a midpoint rule is asymtotically correct for $\int_0^1 f(u)\,\mathrm{d}u$. It suffices for f to be Riemann integrable, but that is not necessary, as convergence also holds for some unbounded integrands.

For an even number G, the eigenvalues of $G^{-1}\Psi^G$ come in pairs apart from the 0'th and the $G/2$'th one, which is $G^{-1}\sum_{j=0}^{G-1}(-1)^j a_j \to 0$. See Karner et al. (2003).

When $d > 1$, and G is odd, the eigenvalues still come in pairs with opposite sign. The generalization of Ψ^G is then a d-fold Kronecker product of left circulant matrices. See van der Mee et al. (2006) for an example using multiindex Toeplitz matrices with a similar Kronecker product structure. The eigenvalues of the Kronecker product are products of the eigenvalues of its matrix factors and so the spectrum is still symmetric apart from the eigenvalue for the constant eigenvector, which approaches $\int_{[0,1)^d} f(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} = 0$.

6. Conclusions

Given n independent vectors $\boldsymbol{U}_i \sim \mathbf{U}[0,1)^d$ we can recycle them over and over to make $N_m = \binom{n}{m}$ pairwise independent random variables $\boldsymbol{X}_i$. The variance of Monte Carlo integration is not adversely affected when we substitute these pairwise independent for genuinely independent ones. The estimate $\bar{Y}$ is unbiased with the same variance as for independent variables and the customary variance estimate for $\bar{Y}$ is unbiased. These moment results hold for fixed m as $n \to \infty$. If m increases with n, perhaps $m = \lfloor n/2 \rfloor$, then Propositions 1 and 2 still hold, but Theorem 1 does not apply unless $m = 2$. Because results for $m = 3$ appeared worse than for $m = 2$, it is reasonable to use a small fixed m like $m = 2$.

To get confidence intervals based on the central limit theorem, it is necessary to average several replicates of the recycled vectors. Due in part to a surprising symmetry in the asymptotic distribution of $\bar{Y}$, shown for $m = 2$, only a small number of replicates are needed. Even $m = 2$ is enough to generate a very large number N of pairwise independent vectors from a modest number n of independent ones.

Pairing up random vectors works for d dimensional integration, but it is crucial to the analysis that the components X_{ij} for $j = 1, \dots, d$ of each point $\boldsymbol{X}_i$ be truly independent. We arranged this by making each $\boldsymbol{U}_i$ have d independent components. In particular, nothing in this article is meant to support turning

n independent scalar uniform random variables into N pairwise independent scalars before forming vectors of dimension d.

Another route to pairwise independent random vectors is to take $\boldsymbol{U}_r \overset{\text{ind}}{\sim} \mathbf{U}[0,1)^d$ for $1 \le r \le n$, where n is even, and form $n^2/4$ pairs $\boldsymbol{U}_r \oplus \boldsymbol{U}_s$ for $1 \le r \le n/2$ and $n/2 < s \le n$. The resulting statistic $\bar{Y}$ is a generalized U-statistic, and once again, is degenerate. Theorem 1 does not apply to it. From Lemma B of Serfling (1980, Section 5.2.2) we can find that $\mathbb{E}((\sqrt{N}(\bar{Y}-\mu))^k) = O(1)$ for $k \ge 3$ when $\mathbb{E}(|f(\boldsymbol{U})|^k) < \infty$. Unless the implied constant is zero when $k = 3$ a strategy based on generalized U-statistics will also require independent replicates in order to satisfy a central limit theorem. Even if that constant is 0, the Lemma B does not yield a central limit theorem for generalized U statistics.

Pseudorandom numbers have many advantages compared to physical ones. Indeed the simulations in Section 4 were done with pseudo-random numbers. It is also known that some sources of physical random numbers fail tests of randomness such as Marsaglia's diehard battery of tests. But when one wants to use physical random numbers, the problems of large storage needs can be greatly mitigated by pooling the random numbers together into pairwise independent vectors.

Finally, the problem considered here is similar to one that arises in randomized quasi-Monte Carlo. Scrambled digital nets with the random scrambling proposed in Owen (1995) satisfy a central limit theorem (Loh, 2003). The random linear scrambles of Matoušek (1998) are simpler to implement and require much less storage. Random linear permutations have the same first and second order marginal distributions as uniform random permutations, and scrambled nets using them have the same mean and variance as with uniform random permutations. The third and higher order margins of random linear permutations are different from the uniform ones (when more than 3 items are permuted). Thus they may fail to satisfy a central limit theorem, but no proof or counter example has yet been published. There is some empirical evidence that the distributions could be different: Hong et al. (2003) have found that the two scramblings yield quite different distributions for a related quantity (the mean squared discrepancy) of quadrature points.

Acknowledgments

I thank Robert Serfling, Jim Wilson and Fred Hickernell for helpful comments.

References

Arcones, M. A. and Giné, E. (1993). Limit Theorems for U-Processes. *Annals of Probability* **21** 1494–1542. MR1235426

Davis, P. J. (1979). *Circulant Matrices*. Wiley, New York. MR0543191

Devroye, L. (1986). *Non-uniform Random Variate Generation*. Springer. MR0836973

Gregory, G. (1977). Large sample theory for U-statistics and tests of fit. *The Annals of Statistics* **5** 110–123. MR0433669

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics* **19** 293–325. MR0026294

Hong, H. S., Hickernell, F. J. and Wei, G. (2003). The distributio nof the discrepancy of scrambled digital (t, m, s)–nets. *Mathematics and Computers in Simulation* **62** 335–345. MR1988381

Karner, H., Schneid, J. and Ueberhuber, C. W. (2003). Spectral decomposition of real circulant matrices. *Linear Algebra and its Applications* **367** 301–311. MR1976927

L'Ecuyer, P. (2009). Pseudorandom number generators. In *Encyclopedia of quantitative finance* (R. Cont, ed.) Wiley, New York.

Loh, W.-L. (2003). On the asymptotic distribution of scrambled net quadrature. *Annals of Statistics* **31** 1282–1324. MR2001651

Matoušek, J. (1998). On the L^2–discrepancy for anchored boxes. *Journal of Complexity* **14** 527–556. MR1659004

Owen, A. B. (1995). Randomly Permuted (t, m, s)-Nets and (t, s)-Sequences. In *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing* (H. Niederreiter and P. J.-S. Shiue, eds.) 299–317. Springer-Verlag, New York. MR1445791

Romano, J. P. and Siegel, A. F. (1986). *Counterexamples in probability and statistics*. Wadsworth and Brooks/Cole, Belmont CA. MR0831223

Serfling, R. J. (1980). *Approximation theorems of mathematical statistics*. Wiley, New York. MR0595165

van der Mee, C., Rodriguez, G. and Seatzu, S. (2006). Fast superoptimal preconditioning of multiindex Toeplitz matrices. *Linear Algebra and its Applications* **418** 576–590. MR2260212