Infinite-Dimensional Highly-Uniform Point Sets Defined via Linear Recurrences in $\mathbb{F}_{2^w}$

François Panneton and Pierre L'Ecuyer

Département d'informatique et de recherche opérationnelle
Université de Montréal
C.P. 6128, Succ. Centre-Ville,
Montréal (Québec), H3C 3J7, CANADA
panneton@iro.umontreal.ca and lecuyer@iro.umontreal.ca

Summary. We construct infinite-dimensional highly-uniform point sets for quasi-Monte Carlo integration. The successive coordinates of each point are determined by a linear recurrence in $\mathbb{F}_{2^w}$, the finite field with 2^w elements where w is an integer, and a mapping from this field to the interval $[0, 1)$. One interesting property of these point sets is that almost all of their two-dimensional projections are perfectly equidistributed. We performed searches for specific parameters in terms of different measures of uniformity and different numbers of points. We give a numerical illustration showing that using randomized versions of these point sets in place of independent random points can reduce the variance drastically for certain functions.

1 Introduction

Quasi-Monte Carlo (QMC) methods estimate an integral of the form

$$\mu = \int_{[0,1)^t} f(\mathbf{u}) d\mathbf{u}, \tag{1}$$

for a given function f, by the average

$$Q_n = \frac{1}{n} \sum_{i=0}^{n-1} f(\mathbf{u}_i), \tag{2}$$

for a highly-uniform (or low-discrepancy) point set $P_n = \{\mathbf{u}_0, \ldots, \mathbf{u}_{n-1}\} \subset [0, 1)^t$. *Randomized QMC* (RQMC) randomizes the point set P_n before computing Q_n, in a way that each individual point is uniformly distributed over $[0, 1)^t$ even though the point set as a whole keeps its high uniformity [6, 12, 8, 3].

In many practical simulation settings, f depends on a random and unbounded number of uniforms [8]. This can be covered by taking $t = \infty$, with

the understanding that f would typically depend only on a finite number of coordinates of $\mathbf{u}$ with probability 1, if we interpret $\mathbf{u}$ as an infinite sequence of independent uniform random variables over $[0, 1)$. On the other hand, most popular point set constructions (e.g., digital nets and lattice rules) usually assume a fixed (finite) value of t. There are exceptions, e.g., Korobov lattice rules and Korobov polynomial lattice rules [2, 4], where the dimension can be infinite.

In this paper, we introduce a method for constructing infinite-dimensional point sets P_n via a linear recurrence in the finite field $\mathbb{F}_{2^w}$ and a mapping from $\mathbb{F}_{2^w}$ to the interval $[0, 1)$. The construction is similar to the one used in [11] for random number generation. These point sets are dimension-stationary, i.e., their projections over a subset of coordinates depend only on the spacings between these coordinates. Moreover, most of their two-dimensional projections have maximal equidistribution. We provide a formula that gives the precise number in terms of the parameters of the recurrence.

We define several measures of uniformity for P_n in terms of its equidistribution properties, its q-value, and the distance between the closest points, in several dimensions. We report partial results of a search for good point sets in terms of these criteria. Then we try randomized versions of these point sets on a few test problems and compare them, in terms of variance reduction with respect to standard Monte Carlo (MC) simulation, with Sobol' nets randomized in the same way. In certain settings, the new point sets perform much better than the Sobol' nets.

2 Definition of the Point Sets

Our point sets are constructed as follows. The successive coordinates of each point are defined in essentially the same way as the successive random numbers in [11].

Let $q = 2^w$ for some integer $w \geq 1$ and $\mathbb{F}_q$ the finite field with q elements. We consider a linear recurrence of order r in $\mathbb{F}_q$,

$$m_n = \sum_{i=1}^{r} b_i m_{n-i}, \tag{3}$$

where r is a positive integer, $b_1, \ldots, b_r$ and $m_0, m_1, \ldots$ are in $\mathbb{F}_q$, $b_r \neq 0$, and all arithmetic is performed in $\mathbb{F}_q$. The polynomial $P(z) = z^r - \sum_{i=1}^{r} b_i z^{r-i}$ is a *characteristic polynomial* of this recurrence. It is well-known that (3) has period length $q^r - 1 = 2^{rw} - 1$ (full period) for any nonzero initial state $(m_{-r+1}, \ldots, m_0) \in \mathbb{F}_q^r$ if and only if $P(z)$ is primitive over $\mathbb{F}_q$. Regardless of the primitivity of $P(z)$, the recurrence (3) is *purely periodic*, in the sense that it has no transient state. See, e.g., [5, 6] for an account of linear recurrences in finite fields.

To construct a point set from such a recurrence, we must define a mapping from the state space $\mathbb{F}_q^r$ to the real interval $[0,1)$. This requires an explicit representation of the elements of $\mathbb{F}_q$. As in [11], we represent these elements in terms of an ordered polynomial basis, defined as follows. Let $M(z) = z^w + \sum_{i=1}^{w} a_i z^{w-i} \in \mathbb{F}_2[z]$ be an irreducible polynomial over $\mathbb{F}_2$. Then there exists an algebraic element ζ of $\mathbb{F}_q$ whose minimal polynomial over $\mathbb{F}_2$ is $M(z)$ and the ordered set $(1, \zeta, \ldots, \zeta^{w-1})$ is an *ordered polynomial basis* of $\mathbb{F}_q$ over $\mathbb{F}_2$ (see [5], Chapter 1.4). This means that any element $v \in \mathbb{F}_q$ can be written uniquely as a linear combination $v = v_1 + v_2\zeta + \cdots + v_w\zeta^{w-1}$ where $\mathbf{v} = (v_1, \ldots, v_w)^\mathsf{T} \in \mathbb{F}_2^w$. Here, we identify $\mathbb{F}_2$ with the set $\{0,1\}$ in which addition and multiplication are performed modulo 2. Thus, after $M(z)$ has been chosen, each element v of $\mathbb{F}_q$ can be represented by its corresponding binary column vector $\mathbf{v}$, called its *vector representation*. Then, as explained in [11], the recurrence (3) can be implemented by

$$\mathbf{m}_n = \sum_{i=1}^{r} A_{b_i} \mathbf{m}_{n-i} \tag{4}$$

where $\mathbf{m}_n$ is the vector representation of m_n and A_{b_i} performs the multiplication by b_i in the vector representation, for $1 \le i \le r$. Under this representation, the *state* at step n can be written as the rw-bit column vector

$$\mathbf{s}_n = (\mathbf{m}_{n-r+1}^\mathsf{T}, \ldots, \mathbf{m}_n^\mathsf{T})^\mathsf{T}.$$

From recurrence (4), we define an *output sequence* $u_0, u_1, \ldots$ in $[0,1)$ as follows:

$$\begin{aligned} \mathbf{y}_i &= (\mathbf{m}_{i\nu}^\mathsf{T}, \mathbf{m}_{i\nu+1}^\mathsf{T}, \ldots)^\mathsf{T} \;=\; (y_{i,0}, y_{i,1}, \ldots)^\mathsf{T}, \\ u_i &= \sum_{j=1}^{\infty} y_{i,j-1} 2^{-j} \end{aligned} \tag{5}$$

for $i \ge 0$, where ν is a fixed positive integer and $y_{i,0}, y_{i,1}, \ldots$ are the successive bits of $\mathbf{y}_i$. In practice, $\mathbf{y}_i$ and the expansion in (5) are necessarily truncated to a finite number of bits, but here we neglect the impact of this truncation. Let

$$P_n = \{(u_0, u_1, u_2, \ldots) : \mathbf{s}_0 \in \mathbb{F}_{2^{rw}}\} \tag{6}$$

be the set of all sequences of successive output values u_i, from all possible initial states $\mathbf{s}_0 = (\mathbf{m}_{-r+1}, \ldots, \mathbf{m}_0)$ in $\mathbb{F}_{2^{rw}}$. Since the number of states is 2^{rw} and the recurrence (3) is purely periodic, the cardinality of P_n is $n = 2^{rw}$. This P_n is our infinite-dimensional point set. Each point $\mathbf{u} \in P_n$ is in fact a periodic infinite sequence, whose period length is that of the cycle of the recurrence that corresponds to the initial state $\mathbf{s}_0$. In the case where $P(z)$ is primitive, for example, there are two cycles: one contains the single state $\mathbf{s}_0 = \mathbf{0}$ and has period 1 (it gives the point $\mathbf{u} = \mathbf{0}$) while the other contains all nonzero states and has period length $2^{rw} - 1$. In this case, all nonzero

points can be enumerated as follows: to get the next point, discard the first ν coordinates of the current point and shift all other coordinates by ν position to the left. If $P(z)$ is not primitive, there will be more cycles.

It is easily seen that this P_n is a *digital net* in base 2. Indeed, because of (4), each bit vector $\mathbf{y}_i$ is a linear function of the bit vector $\mathbf{s}_0$. That is, we can write $\mathbf{y}_i = \mathbf{C}^{(i)}\mathbf{s}_0$ for some $\infty \times rw$ binary matrix $\mathbf{C}^{(i)}$, for $i = 0, 1, 2, \ldots$. A quick examination of the definition of P_n immediately tells us that it satisfies the definition of a digital net (see [6, 7]) with generating matrices $\mathbf{C}^{(0)}$, $\mathbf{C}^{(1)}$, ... This net is infinite-dimensional. The sequence of generating matrices is periodic and the successive rows of any $\mathbf{C}^{(i)}$ also form a periodic sequence. If we replace $\mathbf{s}_0$ by the jth canonical vector $\mathbf{e}_j$, the corresponding $\mathbf{y}_i$ gives us the jth column of $\mathbf{C}^{(i)}$. Since the recurrence is purely periodic, there must be a one-to-one correspondance between $\mathbf{s}_0$ and the first rw bits of $\mathbf{y}_i$ for each i. This implies that the first rw rows of $\mathbf{C}^{(i)}$ must be linearly independent over $\mathbb{F}_2$. Thus, the first rw bits of any given coordinate u_j of the points of P_n take all possible 2^{rw} values exactly once. That is, if the binary expansion in (5) is truncated to its first rw bits, then each one-dimensional projection of P_n is the set $\{0, 1/n, \ldots, (n-1)/n\}$.

Some may argue that this type of infinite-dimensional point set is not very interesting because of the periodicity of the point coordinates. However, in practice, P_n would typically be *randomized* to get an unbiased estimator of μ, and the randomization would normally destroy the periodicity. For example, one simple randomization is a *random binary digital shift*: generate a single random point $\mathbf{U}$ uniformly distributed in $[0, 1)^\infty$ and add it to each point of P_n by a bitwise exclusive-or of each coordinate [3]. After this randomization, every individual point of P_n is a random point uniformly distributed over $[0, 1)^\infty$, whereas P_n preserves all its $\mathbf{p}$-equidistribution properties, as defined in the next section. The successive coordinates of the randomized points are no longer periodic.

3 Measures of uniformity

To measure the uniformity of P_n, we will examine its projections over finite subsets of the coordinates. For each such projection, we obtain a finite-dimensional point set, say a point set Q_n over the t-dimensional hypercube $[0, 1)^t$. Several figures of merit can be adopted to measure the uniformity of such a point set Q_n [6, 3]. The measures considered in this paper are based on $\mathbf{p}$-equidissections of the unit hypercube $[0, 1)^t$ and on the minimal distance between the points of Q_n. We recall definitions that can be found, e.g., in [3] and at other places.

Let $\mathbf{p} = (p_1, \ldots, p_t)$ be a vector of positive integers such that $p = p_1 + \ldots + p_t \leq k$. A $\mathbf{p}$-*equidissection* is a partition of the unit hypercube in rectangular *cells* aligned with the axes, of equal volume 2^{-p}, defined by dividing the interval $[0, 1)$ along the i-th coordinate into 2^{p_i} equal parts, for each i. A

p-equidissection such that $p_1 = \ldots = p_t = \ell$ is called an *ℓ-equidissection*. A set Q_n with $n = 2^k$ is said to be **p**-*equidistributed* if every cell defined by the **p**-equidissection contains exactly 2^{k-p} points from Q_n. It is *ℓ-equidistributed* if it is **p**-equidistributed for $p_1 = \ldots = p_t = \ell$.

A point set $Q_n \subset [0,1)^t$ with $n = 2^k$ points is a (q,k,t)-*net* (in base 2) if it is **p**-equidistributed for every **p**-equidissection such that $p_1 + \ldots + p_t \leq k - q$. The smallest q such that Q_n forms a (q,k,t)-net is called the *q-value* of Q_n. We denote it by q_t. Generally speaking, a smaller q-value means a more uniform point set.

The largest ℓ such that Q_n is ℓ-equidistributed is called its *resolution* and is denoted ℓ_t (in t dimensions). We have the upper bound $\ell_t \leq \ell_t^* \stackrel{\text{def}}{=} \lfloor k/t \rfloor$. We define the *resolution gap* in t dimensions as

$$\Lambda_t = \lfloor k/t \rfloor - \ell_t.$$

A smaller resolution gap means a more uniform point set.

Equidistribution in **p**-equidissections has its limitations in measuring the uniformity of a point set. For example, if a point **u** is a common corner for 2^t cells in t dimensions, then up to 2^t distinct points of Q_n can be arbitrarily close to **u**, one in each cell. Thus, despite good equidistribution properties, one may have a cluster of several points that are almost identical to each other. To prevent this, one may consider the *minimal distance* of Q_n under the L_p norm, defined as

$$d_p^*(Q_n) = \min\{d_p(\mathbf{x},\mathbf{y}) : \mathbf{x},\mathbf{y} \in Q_n, \mathbf{x} \neq \mathbf{y}\},$$

where $d_p(\mathbf{x},\mathbf{y})$ is the L_p-distance between **x** and **y**. A large value of $d_p^*(Q_n)$ means that all points are far away from each other, and are thus more evenly spread over the hypercube.

Here, instead of $d_p^*(Q_n)$, we use a related figure of merit defined as follows. Two cells defined by a **p**-equidissection are *adjacent* if they have at least one corner in common. A point set $Q_n \subset [0,1)^t$ is said to be *neighbor-free in resolution ℓ* if in the ℓ-equidissection, no cell contains more that one point from Q_n and every cell that contains one point is adjacent to no other such cell. The smallest value of ℓ such that Q_n is neighbor-free is called the *neighbor-free resolution* and is denoted by v_t. A lower bound on v_t is $\lceil k/t \rceil + 1$. We define the *neighbor-free gap* as

$$\Gamma_t = v_t - \lceil k/t \rceil - 1.$$

The neighbor-free resolution is linked to the minimal distance by the inequalities

$$\begin{aligned} 2^{-v_t} &< d_2^*(Q_n) &&< 2^{-v_t+2}\sqrt{t}, \\ 2^{-v_t} &< d_\infty^*(Q_n) &&< 2^{-v_t+2}, \end{aligned}$$

proved in [10]. We want v_t (or equivalently, Γ_t) to be as small as possible.

We now return to our infinite-dimensional point set P_n. For any subset of coordinates $J = \{j_1, j_2, \ldots, j_i\}$, where $0 \leq j_1 < j_2 < \cdots < j_i < t$, we define

$P_n(J)$ as the i-dimensional projection of P_n over these coordinates. Figures of merit that take into account the uniformity of projections are discussed in [2, 3], for example. Giving special attention to the most important projections often has a significant impact on the performance of RQMC. The most important projections depend on the problem in general, but they are often of small dimension, and associated with coordinate numbers that are close to each other.

For any given family $\mathcal{J}$ of projections, we define

$$\Delta(P_n, \mathcal{J}, C) = \max_{J \in \mathcal{J}} C(P_n(J))$$

and

$$\Theta(P_n, \mathcal{J}, C) = \sum_{J \in \mathcal{J}} C(P_n(J)),$$

where $C(P_n(J))$ can be either q_i, Λ_i, or Γ_i, for $i = |J|$. The criterion $\Delta(P_n, \mathcal{J}, C)$ looks at the *worst-case* projection in $\mathcal{J}$, whereas $\Theta(P_n, \mathcal{J}, C)$ considers the *average* instead.

4 Guaranteed Uniformity of Certain Projections

For the point sets defined in (6), each one-dimensional projection contains exactly one point in each of the intervals $[0, 1/n), [1/n, 2/n), \ldots, [(n-1)/n, 1)$. Moreover, because of the way P_n is defined via a recurrence, for any given set of non-negative integers $J = \{j_1, j_2, \ldots, j_i\}$, the projections $P_n(\{j_1+j, \ldots, j_i+j\})$ are identical for all $j \geq 0$. That is, the point set is *dimension-stationary* [2].

The following proposition, on the equidistribution of two-dimensional projections, is proved in [10].

Proposition 1. *Suppose that the minimal polynomial $P(z)$ of the recurrence (3) over $\mathbb{F}_{2^w}$ is a primitive polynomial. Let $h = \mathrm{lcm}((2^k - 1)/(2^w - 1), \nu)/\nu$, where* lcm *means the least common multiple. Then, the two-dimensional projection $P_n(\{j_1, j_1 + j\})$ is w-equidistributed if and only if j is not a multiple of h.*

As an illustration, consider a point set P_n of cardinality $n = 2^{16}$, obtained by taking $r = 2$, $w = 8$, and $\nu = 13$. In that case, $h = \mathrm{lcm}((2^{16} - 1)/(2^8 - 1), 13)/13 = \mathrm{lcm}(257, 13)/13 = 257$. This means that among all two-dimensional projections of the form $P_n(\{0, j\})$, exactly 65280 out of 65535 (i.e., all but 1 out of every 257) are 8-equidistributed (which is the best possible two-dimensional equidistribution for 2^{16} points).

5 A Search for Good Point Sets

We made extensive computer searches for good point sets in terms of the general figures of merit defined in Section 3, for various values of n. A small subset of the results, for $n = 2^{14}$ and 2^{16}, is given in Table 1. The elements of the finite field $\mathbb{F}_{2^w}$ are represented using the hexadecimal notation and the polynomial basis (as in [11]).

Table 1. Point sets with cardinality 2^{14} and 2^{16}.

Number	r	w	k	$M(z)$	ν	$\mathbf{b}_1$	$\mathbf{b}_2$	$\mathbf{b}_3$	$\mathbf{b}_4$	$\mathbf{b}_5$	$\mathbf{b}_6$	$\mathbf{b}_7$	$\Delta(S,C)$	$\Theta(S,C)$	S	C
1	2	7	14	77	152	73	52	–	–	–	–	–	1	12	$\mathcal{J}_1$	Λ_t
2	4	4	16	9	842	3	e	0	e	–	–	–	1	32	$\mathcal{J}_1$	Λ_t
3	7	2	14	3	548	2	0	0	2	1	0	1		12	$\mathcal{J}_1$	Λ_t
4	4	4	16	c	286	4	9	e	4	–	–	–		31	$\mathcal{J}_1$	Λ_t
5	7	2	14	3	468	2	0	1	1	0	1	3	7	934	$\mathcal{J}_1$	q_t
6	4	4	16	9	883	0	4	e	b	–	–	–	9	989	$\mathcal{J}_1$	q_t
7	7	2	14	3	236	3	2	0	0	0	3	1		889	$\mathcal{J}_1$	q_t
8	4	4	16	9	816	0	3	d	3	–	–	–		959	$\mathcal{J}_1$	q_t
9	7	2	14	3	199	1	0	3	0	1	1	1	4	303	$\mathcal{J}_2$	Γ_t
10	4	4	16	c	675	b	f	0	9	–	–	–	4	295	$\mathcal{J}_2$	Γ_t
11	2	7	14	5f	101	30	1f	–	–	–	–	–		302	$\mathcal{J}_2$	Γ_t
12	2	8	16	d8	702	88	da	–	–	–	–	–		294	$\mathcal{J}_2$	Γ_t

The sets of projections considered in the figures of merit were of the form

$$\mathcal{J} = \mathcal{J}(s, t_1, \ldots, t_s) = \left(\bigcup_{i=1}^{s} \{\{j_1, \ldots, j_i\}, 0 = j_1 \le \cdots \le j_i < t_i\} \right) \bigcup \{\{0, \ldots, j\}, 0 \le j < t_1\}.$$

They are the projections defined by j successive coordinates for j up to t_1, the two-dimensional projections with coordinates less that t_2, the three-dimensional projections with coordinates less that t_3, and so on. This type of $\mathcal{J}$ was also considered in [2]. Let us denote $\mathcal{J}(5, k, 24, 16, 8, 8)$ by $\mathcal{J}_1(k)$ and $\mathcal{J}(3, 3, 24, 16)$ by $\mathcal{J}_2$.

The parameters reported in Table 1 are for the criteria $\Delta(P_n, \mathcal{J}_1(k), \Lambda_t)$, $\Theta(P_n, \mathcal{J}_1(k), \Lambda_t)$, $\Delta(P_n, \mathcal{J}_1(k), q_t)$, $\Theta(P_n, \mathcal{J}_1(k), q_t)$, $\Delta(P_n, \mathcal{J}_2, \Gamma_t)$, and $\Theta(P_n, \mathcal{J}_2, \Gamma_t)$. More extensive tables of parameters are given in [10]. The effectiveness of these point sets will be assessed empirically for simple examples in the next section.

6 Examples

We report the results of simple numerical experiments where the point sets of Table 1 perform quite well for integrating certain multivariate functions in a RQMC scheme. We compare their performance with that of Sobol' nets when both are randomized by a random binary digital shift only (see, e.g., [3] and [9] for a definition and discussions of other randomization methods). In both cases, we estimate the variance per run, i.e., n times the variance of the average over the n points, and compare it with the empirical variance of standard MC. The *variance reduction factor* reported is the ratio of the MC variance over the RQMC variance per run.

6.1 A Markov Chain

We consider a Markov chain with state $(i, c, \mathbf{U})$ where $i \in \{0, 1, 2\}$, c is an integer, and $\mathbf{U} = (u_1, u_2, \ldots)$ is an infinite sequence with elements in $[0, 1)$. The chain starts in state $i = 1$, $c = 0$ and $\mathbf{U} = (1, 1, \ldots)$. To determine the next state, we generate $U \sim U(0, 1)$, a uniformly distributed random variable. If $U < p_{i,i+1}$ then $i \leftarrow (i + 1) \bmod 3$, otherwise $i \leftarrow (i - 1) \bmod 3$. At each step, we increase c by one and update $\mathbf{U}$ as $\mathbf{U} = (U, u_1, u_2, \ldots)$. When $c \geq 300$, $i = 2$, and $1 - p_3 < U \leq 1$, the chain terminates. In our numerical experiments, we also terminate the chain whenever $i = 360$, in order to be able to compare with the Sobol' nets, for which we have an implementation only for up to 360 dimensions. At each step, there is a cost $f_i(\mathbf{U})$, for some functions f_i that depend on only two coordinates of $\mathbf{U}$. The goal is to estimate the expected total cost, $\mu = \mathrm{E}[C]$. Figure 1 illustrates the behavior of the chain. We can view this Markov chain as a way of randomly sampling two-dimensional projections of the point set P_n, and summing up the values of $u_i u_j$ observed on these projections.

We consider two cases for the choice of the f_j's in our experiments. In both cases, $p_{i,j} = 1/2$ for $0 \leq i, j \leq 2$ and $p_3 = 1/2$. In the first case, we take $f_0(\mathbf{U}) = u_1 u_9$, $f_1(\mathbf{U}) = u_2 u_8$, and $f_2(\mathbf{U}) = u_3 u_7$. In the second case, we take $f_0(\mathbf{U}) = u_1 u_2$, $f_1(\mathbf{U}) = u_2 u_3$, and $f_2(\mathbf{U}) = u_1 u_3$.

We also give the results when we do not stop the chain when $i = 360$ ("Case 1(b)" and "Case 2(b)"). In these cases, the dimension is not bounded and our implementation of the Sobol' nets cannot be used.

The empirical variance reductions of RQMC compared with MC are given in Table 2. These improvement factors are quite large, and much larger for our new point sets than for the Sobol' nets. For most point sets, the variance reduction factors is slightly lower in the "(b)" cases but, for some, the trend is reversed (like point set number four).

6.2 Some multivariate functions

Here, we consider the following two functions f, defined over the unit hypercube $[0, 1)^t$:

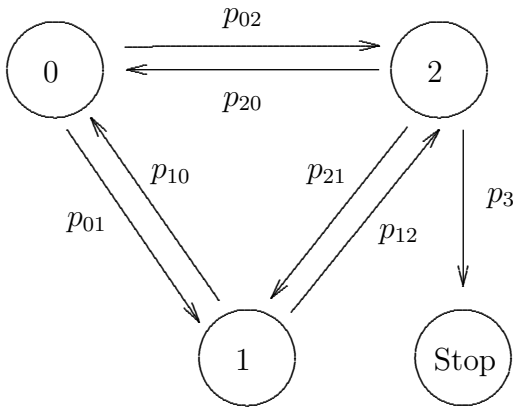

Fig. 1. Evolution of i for our Markov chain.

Table 2. Variance reduction factors of RQMC compared with MC for the Markov chain

Number	Case 1	Case 2	Case 1(b)	Case 2(b)
Sobol, $n = 2^{14}$	5	28	X	X
Sobol, $n = 2^{16}$	39	37	X	X
1	1000	1400	1200	1300
2	4900	2500	4600	2100
3	1500	1200	1400	910
4	1300	1400	1800	2100
5	1300	730	1100	910
6	1900	160	1800	180
7	550	1200	470	1000
8	1400	1200	1200	900
9	10	680	8	880
10	4200	1500	3900	1400
11	22	870	20	900
12	470	270	430	250

$$f(\mathbf{u}) = f_1(\mathbf{u}) = \sqrt{\frac{2}{t(t-1)}} \sum_{j=0}^{t-1} \sum_{i=0}^{j-1} g(u_i) g(u_j)$$

where $g(x) = 27.20917094x^3 - 36.19250850x^2 + 8.983337562x + 0.7702079855$, and

$$f(\mathbf{u}) = f_2(\mathbf{u}) = \sum_{i=0}^{n-1} \left(1 - \prod_{j=0}^{m-1} 2u_{im+j} \right)$$

for $m = 5$, $n = 20$, and $t = 100$. Function f_1, which is from [1], is a sum of functions defined on two-dimensional projections and f_2, taken from [10], is a

sum of functions that depend on projections in five dimensions. Table 3 reports the empirical variance reduction factors observed for these two functions. For certain point sets, the reduction factors are enormous and much better than for Sobol' nets.

Table 3. Variance reduction factors for functions f_1 and f_2.

Number	f_1	f_2
Sobol, $n = 2^{14}$	1.7	820
Sobol, $n = 2^{16}$	0.9	220
1	5×10^4	2×10^4
2	24	2×10^5
3	370	4×10^7
4	9500	800
5	19	2×10^8
6	80	1×10^4
7	10	1×10^9
8	1×10^5	1×10^9
9	630	1×10^9
10	7700	8×10^5
11	580	2×10^5
12	4×10^5	5×10^8

7 Conclusion

In this paper, we have introduced new point sets for quasi-Monte Carlo integration that are very flexible because of their infinite dimensionality. We have provided parameters for point sets that are uniform for many preselected projections and tested them with simple functions to integrate. Tables 2 and 3 show that the point sets selected are efficient in integrating the selected functions. A nice surprise revealed by these tables is the relatively good performance of the point sets (numbers 8 to 12) selected via the minimal distance criteria. It indicates that this uniformity criterion is worth considering for quasi-Monte Carlo applications.

8 Acknowledgments

This work has been supported by NSERC-Canada and FQRNT-Québec scholarships to the first author and by NSERC-Canada grant No. ODGP0110050, FQRNT-Québec grant No. 02ER3218, and a Canada Research Chair to the second author. We thank Alexander Keller, who suggested considering the minimal distance to measure uniformity, and the Editor Harald Niederreiter.

References

1. L. Kocis and W. J. Whiten. Computational investigations of low-discrepancy sequences. *ACM Transactions on Mathematical Software*, 23(2):266–294, June 1997.
2. P. L'Ecuyer and C. Lemieux. Variance reduction via lattice rules. *Management Science*, 46(9):1214–1235, 2000.
3. P. L'Ecuyer and C. Lemieux. Recent advances in randomized quasi-Monte Carlo methods. In M. Dror, P. L'Ecuyer, and F. Szidarovszky, editors, *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, pages 419–474. Kluwer Academic Publishers, Boston, 2002.
4. C. Lemieux and P. L'Ecuyer. Randomized polynomial lattice rules for multivariate integration and simulation. *SIAM Journal on Scientific Computing*, 24(5):1768–1789, 2003.
5. R. Lidl and H. Niederreiter. *Introduction to Finite Fields and Their Applications*. Cambridge University Press, Cambridge, revised edition, 1994.
6. H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63 of *SIAM CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, 1992.
7. H. Niederreiter. Constructions of (t, m, s)-nets and (t, s)-sequences. *Finite Fields and Their Applications*, 2005. To appear.
8. A. B. Owen. Latin supercube sampling for very high-dimensional simulations. *ACM Transactions on Modeling and Computer Simulation*, 8(1):71–102, 1998.
9. A. B. Owen. Variance with alternative scramblings of digital nets. *ACM Transactions on Modeling and Computer Simulation*, 13(4):363–378, 2003.
10. F. Panneton. *Construction d'ensembles de points basée sur des récurrences linéaires dans un corps fini de caractéristique 2 pour la simulation Monte Carlo et l'intégration quasi-Monte Carlo*. PhD thesis, Département d'informatique et de recherche opérationnelle, Université de Montréal, Canada, August 2004.
11. F. Panneton and P. L'Ecuyer. Random number generators based on linear recurrences in F_{2^w}. In H. Niederreiter, editor, *Monte Carlo and Quasi-Monte Carlo Methods 2002*, pages 367–378, Berlin, 2004. Springer-Verlag.
12. I. H. Sloan and S. Joe. *Lattice Methods for Multiple Integration*. Clarendon Press, Oxford, 1994.