Publ. Math. Debrecen
71/3-4 (2007), 305–324

On pseudorandom $[0, 1)$ and binary sequences

By CHRISTIAN MAUDUIT (Marseille), HARALD NIEDERREITER (Singapore)
and ANDRÁS SÁRKÖZY (Budapest)

Abstract. This paper studies links between uniform pseudorandom sequences of real numbers in $[0, 1)$ and pseudorandom binary sequences. It is proved that good pseudorandom $[0, 1)$ sequences induce binary sequences that have small correlation and well-distribution measures. On the other hand, given a binary sequence with small combined well-distribution-correlation measure, it is shown how to construct a $[0, 1)$ sequence with small discrepancy. The special cases of linear congruential pseudorandom sequences and of Legendre symbol sequences are analyzed in more detail.

1. Introduction

(Uniform) *pseudorandom* (briefly PR) *sequences* $x_1, x_2, \dots$ *of real numbers* with $0 \le x_i < 1$ (briefly PR $[0, 1)$ sequences) play a crucial role in applications of the Monte Carlo method and have further applications. Thus, this field has been intensively studied in the last several decades. Surveys of this field are given in [4], [18], [19], and [21]. On the other hand, *pseudorandom binary sequences* also have many applications, in particular, they play an important role in cryptography. In this area the pseudorandomness is usually characterized in terms of complexity theory (see [20], [22]). In an asymptotic sense, pseudorandomness of infinite binary sequences has been considered also in the classical theory of normal numbers (see [8] and [10]). Recently another approach has been developed [13] which is closer to the standard approach used in the theory of PR $[0, 1)$ sequences.

Mathematics Subject Classification: 11B50, 11K16, 11K38, 11K45, 65C10.
Key words and phrases: uniform pseudorandom sequences, pseudorandom binary sequences, discrepancy, correlation measures, linear congruential method, Legendre symbol sequences.

In this paper our goal is to study the links between the two fields described above. We hope that this leads to a better understanding in both areas and that, consequently, the study of the constructions, methods, and tools developed in one field can be utilized in the other field as well.

Throughout this paper we will use the following notations: $\mathbb{N}$ and $\mathbb{Z}$ denote the set of the positive integers, respectively integers. The symbols $c_1, c_2, \dots$ denote positive absolute constants. We write $e(\alpha) = e^{2\pi i\alpha}$. The integer part of x, the fractional part of x, and the distance of x from the nearest integer are denoted by $[x]$, $\{x\}$, and $\|x\|$, respectively, so that $x = [x] + \{x\}$ and $\|x\| = \min(\{x\}, 1 - \{x\})$.

2. The measures of pseudorandomness

In the theory of PR $[0,1)$ sequences we usually study infinite sequences $x_1, x_2, \dots,$ and then we use our conclusions to qualify the finite subsequences $x_1, x_2, \dots, x_N$ obtained by truncating the infinite ones; observe that in practice we always work with *finite* sequences. On the other hand, in the case of PR binary sequences we always study finite sequences $e_1, e_2, \dots, e_N$ of a given length. To be able to compare the two fields, here we will restrict ourselves to finite sequences in both cases.

Let $N \in \mathbb{N}$, $X = (x_1, x_2, \dots, x_N)$ with $0 \le x_i < 1$ for $1 \le i \le N$, let $k \in \mathbb{N}$, $k \le N$, and consider the k-dimensional vectors

$$\mathbf{x}_n = (x_n, x_{n+1}, \dots, x_{n+k-1}), \qquad n = 1, 2, \dots, N - k + 1. \tag{1}$$

Then as the measure of pseudorandomness of the sequence X we use the *discrepancy*

$$D(\mathbf{x}_1, \dots, \mathbf{x}_{N-k+1}) \stackrel{\text{def}}{=} \sup_I \left| \frac{A(I; \mathbf{x}_1, \dots, \mathbf{x}_{N-k+1})}{N-k+1} - V(I) \right| \tag{2}$$

where $I = \prod_{i=1}^{k} [u_i, v_i)$ is a subinterval of $[0,1)^k$, $A(I; \mathbf{x}_1, \dots, \mathbf{x}_{N-k+1})$ denotes the number of n, $1 \le n \le N-k+1$, with $\mathbf{x}_n$ belonging to I, and $V(I) = \prod_{i=1}^{k} (v_i - u_i)$ is the volume of the interval I. To simplify the notation, we will also write

$$D(\mathbf{x}_1, \dots, \mathbf{x}_{N-k+1}) = D(X, N, k). \tag{3}$$

Then X is considered as a "good" PR $[0,1)$ sequence if $D(X, N, k)$ is "small" ($= o(1)$ if k is fixed and X runs over sequences with $N \to +\infty$).

In [13] the following measures of pseudorandomness for binary sequences were proposed. We consider binary sequences of type

$$E_N = (e_1, \dots, e_N) \in \{-1, +1\}^N.$$

Then the *well-distribution measure* of E_N is defined as

$$W(E_N) = \max_{a,b,t} \left| \sum_{j=0}^{t-1} e_{a+jb} \right|,$$

where the maximum is taken over all $a, b, t \in \mathbb{N}$ such that $a \leq a + (t-1)b \leq N$, the *normality measure of order* k of E_N is defined as

$$N_k(E_N) = \max_{X \in \{-1,+1\}^k} \max_{0<M\leq N+1-k} \left| |\{n : 0 \leq n < M, (e_{n+1}, \dots, e_{n+k}) = X\}| - \frac{M}{2^k} \right|,$$

and the *correlation measure of order* k of E_N is defined as

$$C_k(E_N) = \max_{M,\mathbf{D}} \left| \sum_{n=1}^{M} e_{n+d_1} e_{n+d_2} \cdots e_{n+d_k} \right|,$$

where the maximum is taken over all $\mathbf{D} = (d_1, \dots, d_k)$ and M such that $0 \leq d_1 < \cdots < d_k \leq N - M$. The *combined* (well-distribution-correlation) *PR-measure of order* k was also defined as

$$Q_k(E_N) = \max_{a,b,t,\mathbf{D}} \left| \sum_{j=0}^{t} e_{a+jb+d_1} e_{a+jb+d_2} \cdots e_{a+jb+d_k} \right|.$$

In [13] it was shown that

$$N_k(E_N) \leq \max_{1\leq t\leq k} C_t(E_N). \tag{4}$$

Thus, it suffices to estimate the well-distribution measure $W(E_N)$ and the correlation measures of order, say, $\leq k$; we obtain an upper bound for $N_\ell(E_N)$ with $\ell \leq k$ as a consequence of these estimates. However, the study of the normality measure $N_k(E_N)$ can be also useful: e.g., this is the case when the construction is of recursive type, thus we can control only the “local behavior” of the sequence, but not the “long-range” correlation of it.

It was proved in [3] (see also [1]) that for a truly random sequence $E_N \in \{-1,+1\}^N$ both PR measures W and C_k are “small”; more precisely, the order of magnitude of $W(E_N)$ and $C_k(E_N)$ (for fixed k) is $N^{1/2}$ and $N^{1/2}(\log N)^{c(k)}$, respectively. Thus, a sequence $E_N \in \{-1,+1\}^N$ can be considered as a “good” PR sequence if both $W(E_N)$ and $C_k(E_N)$ (for “small” k) are small: certainly they must be $o(N)$ as $N \to +\infty$, and ideally they are greater than $N^{1/2}$ only by at most a power of $\log N$; sequences of this type were constructed, e.g., in [5], [12], [15], and [23].

In the next sections we will show that any “good” PR $[0,1)$ sequence induces a relatively (but not necessarily ideally) “good” PR binary sequence and vice versa, and we will study two special examples in both directions.

3. From $[0,1)$ sequences to binary sequences in general

Suppose a sequence $X = (x_1, x_2, \dots, x_N)$ of real numbers in $[0,1)$ is given. There is a natural way to assign a binary sequence $E_N = E_N(X)$ to the sequence X: for $n = 1, 2, \dots, N$, define e_n by

$$e_n = \begin{cases} +1 & \text{if } 0 \le x_n < 1/2, \\ -1 & \text{if } 1/2 \le x_n < 1, \end{cases} \tag{5}$$

and let

$$E_N = E_N(X) = (e_1, e_2, \dots, e_N).$$

We may expect that if X is a "good" PR $[0,1)$ sequence, then the binary sequence $E_N(X)$ also possesses strong PR properties. This is not so in terms of the PR measures introduced in Section 2 as the following example shows.

Example 1. Let $N = 2M \in \mathbb{N}$ be an even number and assume that $x_1, x_2, \dots, x_M$ are independent random variables, each of them distributed according to the law

(i) $P(x_i < 0) = P(x_i \ge 1) = 0$

and

(ii) x_i is uniformly distributed in $[0,1)$ (for $i = 1, 2, \dots, M$).

Moreover, for $i = 1, 2, \dots, M$ set $x_{M+i} = x_i$, and let $X = (x_1, x_2, \dots, x_N)$. Then clearly, $D(X, N, k)$ is "small" ($= o(1)$) with probability near 1 if k is fixed. On the other hand, defining e_n by (5) we have $e_{M+i} = e_i$ for $i = 1, 2, \dots, M$, whence

$$C_2(E_N(X)) \ge \left| \sum_{n=1}^{M} e_n e_{n+M} \right| = M = \frac{N}{2},$$

so that the correlation measure of order 2 of $E_N(X)$ is large.

The explanation of this anomaly between the PR properties of X and $E_N(X)$ is that the discrepancy $D(X, N, k)$ introduced in Section 2 focuses on the "local" behavior of $X = (x_1, x_2, \dots, x_N)$, i.e., we are studying consecutive x_n's, while in the case of the correlation of $E_N(X)$ we also consider "long-range" correlation, i.e., pairs e_m, e_n with m, n far apart. One may eliminate this anomaly by also considering "long-range" discrepancy in the case of $[0,1)$ sequences. Indeed, let us extend definitions (1) and (2) in the following way.

If $0 \le d_1 < \cdots < d_k < N$, then write

$$\mathbf{x}_n(d_1, \dots, d_k) = (x_{n+d_1}, \dots, x_{n+d_k}) \quad \text{for } 1 \le n \le N - d_k$$

(so that the vector in (1) can be written as $\mathbf{x}_n = \mathbf{x}_n(0,1,\dots,k-1)$), and set

$$D[X,N,(d_1,\dots,d_k)] = D(\mathbf{x}_1(d_1,\dots,d_k),\dots,\mathbf{x}_{N-d_k}(d_1,\dots,d_k))$$
$$\stackrel{\text{def}}{=} \sup_I \left| \frac{A\big(I;\mathbf{x}_1(d_1,\dots,d_k),\dots,\mathbf{x}_{N-d_k}(d_1,\dots,d_k)\big)}{N-d_k} - V(I) \right|$$

with a notation analogous to that in (2). In particular, using the notation in (3) we have

$$D(X,N,k) = D[X,N,(0,1,\dots,k-1)].$$

For a binary sequence $E_N = (e_1, e_2, \dots, e_N)$ we will write

$$C\big(E_N, M, (d_1,\dots,d_k)\big) = \left| \sum_{n=1}^{M} e_{n+d_1} \cdots e_{n+d_k} \right|,$$

so that we have

$$C_k(E_N) = \max_{M, 0 \le d_1 < \dots < d_k \le N-M} C\big(E_N, M, (d_1,\dots,d_k)\big). \tag{6}$$

For $X = (x_1, x_2, \dots, x_N)$ and $M = 1,2,\dots,N$, write $X_M = (x_1, x_2, \dots, x_M)$. Now we will prove:

Theorem 1. *For any* $[0,1)$ *sequence* $X = (x_1,\dots,x_N)$, $k \in \mathbb{N}$, $M \in \mathbb{N}$, *and* $0 \le d_1 < \dots < d_k \le N-M$ *we have*

$$C\big(E_N(X), M, (d_1,\dots,d_k)\big) \le 2^k M D[X_{M+d_k}, M+d_k, (d_1,\dots,d_k)]. \tag{7}$$

Proof. Writing $E_N(X) = (e_1,\dots,e_N)$ we have

$$C(E_N(X), M, (d_1,\dots,d_k)) = \left| \sum_{n=1}^{M} e_{n+d_1} \cdots e_{n+d_k} \right|$$
$$= \left| \sum_{(\varepsilon_1,\dots,\varepsilon_k) \in \{-1,+1\}^k} \big|\{n\colon 1 \le n \le M, (e_{n+d_1},\dots,e_{n+d_k}) = (\varepsilon_1,\dots,\varepsilon_k)\}\big| \, \varepsilon_1 \cdots \varepsilon_k \right|.$$

Now for any $\varepsilon = (\varepsilon_1,\dots,\varepsilon_k) \in \{-1,+1\}^k$ and for $i = 1,2,\dots,k$, set

$$[u_i, v_i) = \begin{cases} [0,1/2) & \text{if } \varepsilon_i = +1, \\ [1/2,1) & \text{if } \varepsilon_i = -1, \end{cases} \tag{8}$$

and $I(\varepsilon)=\prod_{i=1}^{k}[u_i,v_i)$ so that

$$V(I(\varepsilon))=\frac{1}{2^k}.$$

Then by (5) and (8), for any n we have

$$(e_{n+d_1},\dots,e_{n+d_k})=\varepsilon$$

if and only if $\mathbf{x}_n(d_1,\dots,d_k)\in I(\varepsilon)$. It follows that

$$\begin{aligned}
&C(E_N(X),M,(d_1,\dots,d_k))\\
&=\Bigg|\sum_{\varepsilon\in\{-1,+1\}^k}|\{n:1\le n\le M,\ \mathbf{x}_n(d_1,\dots,d_k)\in I(\varepsilon)\}|\,\varepsilon_1\cdots\varepsilon_k\Bigg|\\
&=\Bigg|\sum_{\varepsilon\in\{-1,+1\}^k}A\big(I(\varepsilon);\mathbf{x}_1(d_1,\dots,d_k),\dots\mathbf{x}_M(d_1,\dots,d_k)\big)\,\varepsilon_1\cdots\varepsilon_k\Bigg|\\
&=\Bigg|\sum_{\varepsilon\in\{-1,+1\}^k}\frac{M}{2^k}\,\varepsilon_1\cdots\varepsilon_k\\
&\quad+\sum_{\varepsilon\in\{-1,+1\}^k}\left(A\left(I(\varepsilon);\mathbf{x}_1(d_1,\dots,d_k),\dots,\mathbf{x}_M(d_1,\dots,d_k)\right)-\frac{M}{2^k}\right)\varepsilon_1\cdots\varepsilon_k\Bigg|\\
&\le\sum_{\varepsilon\in\{-1,+1\}^k}\left|A\left(I(\varepsilon);\mathbf{x}_1(d_1,\dots,d_k),\dots,\mathbf{x}_M(d_1,\dots,d_k)\right)-\frac{M}{2^k}\right|\\
&\le\sum_{\varepsilon\in\{-1,+1\}^k}MD\left[X_{M+d_k},M+d_k,(d_1,\dots,d_k)\right]\\
&=2^kMD\left[X_{M+d_k},M+d_k,(d_1,\dots,d_k)\right]
\end{aligned}$$

which proves (7). □

It follows from (6) and Theorem 1 that

Corollary 1. *For any* $[0,1)$ *sequence* $X=(x_1,x_2,\dots,x_N)$, $k\in\mathbb{N}$, *and* $k\le N$ *we have*

$$C_k(E_N(X))\le 2^k\max_{M,0\le d_1<\dots<d_k\le N-M}MD\left[X_{M+d_k},M+d_k,(d_1,\dots,d_k)\right].$$

By the inequality

$$W(E_N)\le 3(NC_2(E_N))^{1/2}$$

proved in [13] (see also [6], [7]) and by (4) and Corollary 1, the PR measures $W(E_N(X))$ and $N_k(E_N(X))$ also can be estimated from above in terms of the discrepancies $D[X_U,U,(d_1,\dots,d_k)]$. However, we may get sharper upper bounds if we estimate these measures directly instead of using correlation estimates. In particular, the W measure can be estimated in the following way:

Theorem 2. *For any* $[0,1)$ *sequence* $X=(x_1,\dots,x_N)$ *we have*

$$W(E_N(X)) \le 1 + c_1 \max_{\substack{t\in\mathbb{N}\\ 2\le t\le N}} \left(t \max_{\substack{a,b\in\mathbb{N}\\ a+(t-1)b\le N}} D(x_a, x_{a+b}, \dots, x_{a+(t-1)b}) \right). \tag{9}$$

Proof. Write again $E_N(X)=(e_1,\dots,e_n)$ and assume that $a,b,t\in\mathbb{N}$, $a+(t-1)b\le N$. Then we have

$$\left|\sum_{j=0}^{t-1} e_{a+jb}\right| = |e_a| = 1 \quad \text{for } t=1, \tag{10}$$

and for every t,

$$\begin{aligned}\left|\sum_{j=0}^{t-1} e_{a+jb}\right| &= \big|2|\{j: 0\le j<t, e_{a+jb}=1\}| - t\big| \\ &= 2t\left|\frac{1}{t}\left|\left\{j: 0\le j<t, 0\le x_{a+jb}<\frac{1}{2}\right\}\right| - \frac{1}{2}\right| \\ &\le 2tD(x_a, x_{a+b}, \dots, x_{a+(t-1)b}). \end{aligned} \tag{11}$$

Now (9) follows from (10) and (11). □

4. From $[0,1)$ sequences to binary sequences in a special case

In Section 3 we studied the PR properties of the binary sequence $E_N(X)$ induced by the $[0,1)$ sequence X for general sequences X. Of course, for special sequences X one can usually go beyond the general estimates of Section 3. In this section we will study the, perhaps, most important special family of PR $[0,1)$ sequences, namely, the PR $[0,1)$ sequences generated by the *linear congruential method* introduced by Lehmer in 1949 and analyzed later in numerous papers; see, e.g., [11], [16], [17]. This method can be described in the following way.

Let $m\in\mathbb{N}$, $m\ge 2$, $y_0\in\mathbb{Z}$, $0\le y_0<m$, $\lambda\in\mathbb{Z}$, $\gcd(\lambda,m)=1$, and $r\in\mathbb{Z}$. Define the sequence $y_0,y_1,\dots$ by the linear recursion $y_{n+1}\equiv \lambda y_n + r \pmod m$ and $0\le y_{n+1}<m$ for $n=0,1,\dots$. Write $x_n=\frac{y_n}{m}$ for $n=0,1,\dots$ so that $x_n\in[0,1)$ for all n. Then the sequence $x_0,x_1,\dots$ is considered as a PR $[0,1)$ sequence generated by the linear congruential method. Here we will restrict ourselves to

the most important special case when $m = p$ is a prime number, $\lambda = g$ is a primitive root modulo p, $y_0 \neq 0$, and $r = 0$ (the "homogeneous case"), so that now we have

$$y_{n+1} \equiv g y_n \pmod{p}, \quad 0 < y_n < p \qquad \text{for } n = 0, 1, \dots,$$

whence

$$y_n \equiv y_0 g^n \pmod{p}, \quad 0 < y_n < p \qquad \text{for } n = 0, 1, \dots,$$

and

$$x_n = \frac{y_n}{p}, \quad 0 < x_n < 1 \qquad \text{for } n = 0, 1, \dots.$$

Then clearly, the $[0,1)$ sequence $x_0, x_1, \dots$ is periodic with least period length $p - 1$, so that we may restrict ourselves to the study of the sequence $X = (x_0, x_1, \dots, x_{p-2})$.

Niederreiter [16] proposed the *serial test* to study the pseudorandomness of this $[0,1)$ sequence X. This test consists of taking an $s \in \mathbb{N}$, $s \geq 2$, then considering the s-dimensional vectors $\mathbf{x}_n = (x_n, \dots, x_{n+s-1})$ with $n = 0, 1, \dots, p-2$, and computing the discrepancy $D(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{p-2})$; if the discrepancy is "small", then we say that X passes the s-dimensional serial test. He showed that X passes the s-dimensional serial test if g is an *optimal coefficient mod p* for this s, which means that the nontrivial solutions of the congruence

$$h_1 + h_2 g + h_3 g^2 + \dots + h_s g^{s-1} \equiv 0 \pmod{p}$$

in integers $h_1, h_2, \dots, h_s$ are such that the lattice point $(h_1, h_2, \dots, h_s)$ is far from the origin. We will need the following notations: for $m, h \in \mathbb{Z}$, $m \geq 2$, we set

$$r(h, m) = \begin{cases} 1 & \text{if } m \mid h, \\ m \sin \pi \|h/m\| & \text{if } m \nmid h, \end{cases}$$

and for a lattice point $\mathbf{h} = (h_1, h_2, \dots, h_s) \in \mathbb{Z}^s$ we write

$$r(\mathbf{h}, m) = \prod_{j=1}^{s} r(h_j, m).$$

(Note that $r(\mathbf{h}, m) > 0$ for all $\mathbf{h} \in \mathbb{Z}^s$.) Furthermore, $\sum\limits_{\mathbf{h} \ (\mathrm{mod}\ m)}$ denotes summation over all $\mathbf{h} = (h_1, h_2, \dots, h_s) \in \mathbb{Z}^s$ with $-\frac{m}{2} < h_j \leq \frac{m}{2}$ for $1 \leq j \leq s$ and $\sum\limits^{*}_{\mathbf{h} \ (\mathrm{mod}\ m)}$

denotes summation over all $\mathbf{h} = (h_1, h_2, \dots, h_s) \in \mathbb{Z}^s$ with $-\frac{m}{2} < h_j \leq \frac{m}{2}$ for $1 \leq j \leq s$ and $\mathbf{h} \neq \mathbf{0} = (0, 0, \dots, 0)$. We write

$$\mathbf{G} = (1, g, g^2, \dots, g^{s-1}) \in \mathbb{Z}^s, \tag{12}$$

and $\mathbf{h} \cdot \mathbf{G}$ denotes the scalar product of $\mathbf{h}$ and $\mathbf{G}$. Niederreiter [16, Corollary 3.3] proved:

Theorem A. *The discrepancy of the points $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{p-2} \in [0,1)^s$ defined above satisfies*

$$\begin{aligned} D(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{p-2}) < \frac{s}{p} &+ \frac{1}{p-1}\left(\left(\frac{2}{\pi}\log p + \frac{7}{5}\right)^s - 1\right) \\ &+ \frac{p-2}{p-1} \sideset{}{^*}\sum_{\substack{\mathbf{h} \pmod p \\ \mathbf{h}\cdot\mathbf{G} \equiv 0 \pmod p}} \frac{1}{r(\mathbf{h}, p)}. \end{aligned} \tag{13}$$

Moreover, denoting the last sum in (13) by

$$R_s(g, p) = \sideset{}{^*}\sum_{\substack{\mathbf{h} \pmod p \\ \mathbf{h}\cdot\mathbf{G} \equiv 0 \pmod p}} \frac{1}{r(\mathbf{h}, p)}, \tag{14}$$

he proved (Theorem 3.4 in [16]):

Theorem B. *For any prime p and any $s \geq 2$, there exists a primitiveroot $g_0 \mod p$ with*

$$R_s(g_0, p) < \frac{s-1}{\varphi(p-1)}\left(\left(\frac{2}{\pi}\log p + \frac{7}{5}\right)^s - 1\right),$$

where φ is Euler's totient function.

Combining Theorems A and B we obtain (see Corollary 3.5 in [16]):

Theorem C. *For any prime p and any $s \geq 2$, there exists a primitive root $g_0 \mod p$ such that the associated sequence $X = (x_0, x_1, \dots)$ satisfies*

$$D(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{p-2}) < \frac{1}{p-1}\left(1 + \frac{(p-2)(s-1)}{\varphi(p-1)}\right)\left(\frac{2}{\pi}\log p + \frac{7}{5}\right)^s.$$

Thus, we may conclude that for fixed p and s there is a well-characterized non-empty set of primitive roots $\mod p$ for which the $[0,1)$ sequence X possesses "good" PR properties in the sense that it passes the s-dimensional serial test. In the rest of this section we will be looking for the answer to the following questions: What can one say about *the PR properties of the associated binary sequences* $E_{p-1}(X)$? Is it true that there are primitive roots g for which $E_{p-1}(X)$ possesses strong PR properties? How well can we control the PR properties of these binary sequences? In answering these questions, we will restrict ourselves to the most important PR measures of binary sequences, namely, to the measures W and C_k.

The following four lemmas will be needed in the proofs of our main results. The first two lemmas are Lemma 2.2 and Lemma 2.3 in [16].

Lemma 1. *Let $\mathbf{y}_0, \dots, \mathbf{y}_{N-1}$ be N lattice points in $\mathbb{Z}^s$. Then, for any integer $m \geq 2$, the discrepancy D_N of the fractional parts of the points $(1/m)\mathbf{y}_0, \dots, (1/m)\mathbf{y}_{N-1}$ satisfies*

$$D_N \leq \frac{s}{m} + \sideset{}{^*}\sum_{\mathbf{h} \pmod{m}} \frac{1}{r(\mathbf{h}, m)} \left| \frac{1}{N} \sum_{n=0}^{N-1} e(\mathbf{h} \cdot \mathbf{y}_n / m) \right|.$$

Lemma 2. *For any integer $m \geq 2$, we have*

$$\sum_{\mathbf{h} \pmod{m}} \frac{1}{r(\mathbf{h}, m)} < \left(\frac{2}{\pi} \log m + \frac{7}{5} \right)^s.$$

Lemma 3. *Let $h \in \mathbb{Z}$ be of multiplicative order T modulo a positive integer m and let $a \in \mathbb{Z}$ with $\gcd(a, m) = 1$. Then*

$$\left| \sum_{n=1}^{T} e(ah^n/m) \right| \leq m^{1/2}.$$

Proof. This is the special case $b = 0$ of [9, Chapter 1, Theorem 10]. □

Lemma 4. *Let $h \in \mathbb{Z}$ be of multiplicative order T modulo a positive integer m. Then, for any integers $X < Y$ and any integer a with $\gcd(a, m) = 1$,*

$$\left| \sum_{X < n \leq Y} e(ah^n/m) \right| < c_3 \left(\frac{Y - X}{T} + 1 \right) m^{1/2} \log m.$$

Proof. This is Lemma 2.2 of Banks, Conflitti, Friedlander, and Shparlinski [2]. □

Now we will show that the well-distribution measure $W(E_{p-1}(X))$ is always small:

Theorem 3. *For any prime p and primitive root $g \mod p$ we have*

$$W(E_{p-1}(X)) < c_2 p^{1/2}(\log p)^2. \tag{15}$$

Proof. In order to be able to use Theorem 2, we need an upper bound on $D(x_0, x_{a+b}, \dots, x_{a+(t-1)b})$. This can be obtained by first using Lemma 1 with $s = 1$, $m = p$, and $\mathbf{y}_n = y_0 g^{a+nb}$ for $n = 0, 1, \dots, t-1$. We can assume $p \geq 3$, and then we get

$$\begin{aligned}
D(x_a, x_{a+b}, \dots, x_{a+(t-1)b}) &= D\left(\left\{\frac{y_0 g^a}{p}\right\}, \left\{\frac{y_0 g^{a+b}}{p}\right\}, \dots, \left\{\frac{y_0 g^{a+(t-1)b}}{p}\right\}\right) \\
&\leq \frac{1}{p} + \sum_{0<|h|<p/2} \frac{1}{p \sin \pi \|h/p\|} \left| \frac{1}{t} \sum_{n=0}^{t-1} e\left(\frac{h y_0 g^{a+nb}}{p}\right) \right| \\
&\leq \frac{1}{p} + 2 \sum_{0<h<p/2} \frac{1}{p(2h/p)} \left| \frac{1}{t} \sum_{n=0}^{t-1} e\left(\frac{h y_0 g^a (g^b)^n}{p}\right) \right|.
\end{aligned} \tag{16}$$

The multiplicative order T of g^b modulo p is

$$T = \frac{p-1}{\gcd(b, p-1)} \geq \frac{p-1}{b}. \tag{17}$$

By using Lemmas 3 and 4 and also (17), we obtain from (16) that

$$\begin{aligned}
D(x_a, x_{a+b}, \dots, x_{a+(t-1)b}) &\leq \frac{1}{p} + \sum_{0<h<p/2} \frac{1}{h} \left| \frac{1}{t} \sum_{j=0}^{\left[\frac{t-1}{T}\right]-1} \sum_{n=jT}^{(j+1)T-1} e\left(\frac{h y_0 g^a (g^b)^n}{p}\right) \right. \\
&\quad \left. + \frac{1}{t} \sum_{n=\left[\frac{t-1}{T}\right]T}^{t-1} e\left(\frac{h y_0 g^a (g^b)^n}{p}\right) \right| \\
&\leq \frac{1}{p} + \sum_{0<h<p/2} \frac{1}{h} \left(\frac{1}{t} \left[\frac{t-1}{T}\right] p^{1/2} + c_3(1+1) \frac{1}{t} p^{1/2} \log p \right) \\
&< \frac{1}{p} + \left(\sum_{0<h<p/2} \frac{1}{h} \right) \left(\frac{b}{p-1} p^{1/2} + c_4 \frac{1}{t} p^{1/2} \log p \right) \\
&< \frac{1}{p} + c_5 \left(\frac{b \log p}{p^{1/2}} + \frac{p^{1/2} (\log p)^2}{t} \right).
\end{aligned} \tag{18}$$

If $2 \leq t \leq p-1$ and $a + (t-1)b \leq p-1$, then we have

$$tb \leq 2(t-1)b < 2(p-1) < 2p.$$

Thus, it follows from Theorem 2 and (18) that

$$\begin{aligned} &W(E_{p-1}(X)) \\ &\quad \le 1 + c_1 \max_{\substack{t\in\mathbb{N}\\ 2\le t\le p-1}} \left(t \max_{\substack{a,b\in\mathbb{N}\\ a+(t-1)b\le p-1}} \left(\frac{1}{p} + c_5 \left(\frac{b\log p}{p^{1/2}} + \frac{p^{1/2}(\log p)^2}{t} \right)\right)\right) \\ &\quad < 1 + c_6 \max_{\substack{t\in\mathbb{N}\\ 2\le t\le p-1}} \left(\frac{t}{p} + \Big(\max_{\substack{a,b\in\mathbb{N}\\ a+(t-1)b\le p-1}} bt \Big) \frac{\log p}{p^{1/2}} + p^{1/2}(\log p)^2 \right) \\ &\quad < 1 + c_7 \max_{\substack{t\in\mathbb{N}\\ 2\le t\le p-1}} \big(1 + p^{1/2}\log p + p^{1/2}(\log p)^2\big) < c_8 p^{1/2}(\log p)^2, \end{aligned}$$

which completes the proof of Theorem 3. □

Next we will show that, on the other hand, the correlation C_2 is always large:

Proposition 1. *For any prime p and any primitive root $g \mod p$ we have*

$$C_2(E_{p-1}(X)) \ge \frac{p-1}{2}. \tag{19}$$

Proof. We can again assume $p \ge 3$. For $n = 0, 1, \dots$ we have

$$y_{n+(p-1)/2} \equiv y_0 g^{n+(p-1)/2} = y_0 g^n g^{(p-1)/2} \equiv y_n \cdot (-1) = -y_n \pmod p,$$

whence

$$y_{n+(p-1)/2} = p - y_n,$$

so that

$$x_{n+(p-1)/2} = \frac{y_{n+(p-1)/2}}{p} = \frac{p-y_n}{p} = 1 - x_n.$$

It follows that, writing $E_{p-1}(X) = (e_1, \dots, e_{p-1})$, we have

$$e_{n+(p-1)/2} = -e_n \quad \text{for } n = 1, \dots, p-1,$$

and thus

$$\begin{aligned} C_2(E_{p-1}(X)) &\ge C\left(E_{p-1}(X), \frac{p-1}{2}, \left(0, \frac{p-1}{2}\right)\right) \\ &= \left| \sum_{n=1}^{(p-1)/2} e_n e_{n+(p-1)/2} \right| = \left| \sum_{n=1}^{(p-1)/2} e_n(-e_n) \right| = \left| -\sum_{n=1}^{(p-1)/2} 1 \right| = \frac{p-1}{2} \end{aligned} \tag{20}$$

which proves (19). □

In the proof above, C_2 is made large by a long-range correlation. On the other hand, we will be able to give nontrivial upper bounds for the short-range correlations. First we extend the notations (12) and (14): for a fixed primitive root $g \mod p$ and for $\mathbf{D} = (d_1, \dots, d_k)$, $0 \le d_1 < \dots < d_k$, write

$$\mathbf{G}(\mathbf{D}) = (1, g^{d_2-d_1}, \dots, g^{d_k-d_1}) \in \mathbb{Z}^k$$

and

$$R_k(g,p,\mathbf{D}) = \sideset{}{^*}\sum_{\substack{\mathbf{h} \pmod p \\ \mathbf{h}\cdot\mathbf{G}(\mathbf{D})\equiv 0 \pmod p}} \frac{1}{r(\mathbf{h},p)}.$$

Note that in the last sum the vector $\mathbf{h}$ is k-dimensional. We will prove:

Theorem 4. *If p is a prime, g is a primitive root* $\mod p$, $k \in \mathbb{N}$, $M \in \mathbb{N}$, *and $\mathbf{D} = (d_1, \dots, d_k)$ with $0 \le d_1 < \dots < d_k \le p-1-M$, then we have*

$$C(E_{p-1}(X), M, \mathbf{D}) < c_9 p^{1/2} (\log p) \left(\frac{4}{\pi}\log p + \frac{14}{5}\right)^k + 2^k M R_k(g,p,\mathbf{D}).$$

(Note that if we consider “short-range” correlation, i.e., $d_k - d_1$ is small, then the number of terms in the sum in the definition of $R_k(g,p,\mathbf{D})$ is also small, which makes the upper bound in the theorem sharper.)

Proof. By Theorem 1 it suffices to estimate $D[X_{M+d_k}, M+d_k, \mathbf{D}]$. By Lemma 1 with $m = p$ and

$$\begin{aligned}(1/p)\mathbf{y}_n = \mathbf{x}_n(d_1,\dots,d_k) &= (x_{n+d_1-1}, \dots, x_{n+d_k-1}) \\ &\equiv (1/p)\big(y_0 g^{n+d_1-1}, \dots, y_0 g^{n+d_k-1}\big) \pmod 1\end{aligned}$$

for $n = 1, \dots, M$, we have

$$D\big[X_{M+d_k}, M+d_k, \mathbf{D}\big] \le \frac{k}{p} + \sideset{}{^*}\sum_{\mathbf{h} \pmod p} \frac{1}{r(\mathbf{h},p)} \left|\frac{1}{M}\sum_{n=1}^{M} e\big(\mathbf{h}\cdot\mathbf{x}_n(d_1,\dots,d_k)\big)\right|. \tag{21}$$

For fixed $\mathbf{h} = (h_1, \dots, h_k) \ne \mathbf{0}$, the absolute value of the inner sum is

$$\begin{aligned}&\left|\sum_{n=1}^{M} e\big(\mathbf{h}\cdot\mathbf{x}_n(d_1,\dots,d_k)\big)\right| = \left|\sum_{n=1}^{M} e\left(\sum_{j=1}^{k} h_j x_{n+d_j-1}\right)\right| \\ &= \left|\sum_{n=1}^{M} e\left(\left(\sum_{j=1}^{k} h_j y_0 g^{n+d_j-1}\right)\Big/p\right)\right| = \left|\sum_{n=0}^{M-1} e\left(y_0 g^{d_1}\left(\sum_{j=1}^{k} h_j g^{d_j-d_1}\right) g^n/p\right)\right|.\end{aligned}$$

If

$$\mathbf{h}\cdot\mathbf{G}(\mathbf{D})=\sum_{j=1}^{k}h_j g^{d_j-d_1}\equiv 0 \pmod p,$$

then the absolute value of the inner sum in (21) is equal to M. If $\mathbf{h}\cdot\mathbf{G}(\mathbf{D})\not\equiv 0 \pmod p$, then we may use Lemma 4 (note that $M\le M+d_k<p$) to obtain

$$\left|\sum_{n=0}^{M-1} e\left(y_0 g^{d_1}\left(\sum_{j=1}^{k}h_j g^{d_j-d_1}\right)g^n/p\right)\right| < 2c_3 p^{1/2}\log p.$$

Thus, it follows from (21) that

$$\begin{aligned} D\big[X_{M+d_k}, M+d_k, \mathbf{D}\big] &\le \frac{k}{p} \\ &+ \sideset{}{^*}\sum_{\substack{\mathbf{h} \pmod p\\ \mathbf{h}\cdot\mathbf{G}(\mathbf{D})\equiv 0 \pmod p}} \frac{1}{r(\mathbf{h},p)} + 2c_3\frac{p^{1/2}\log p}{M}\sideset{}{^*}\sum_{\mathbf{h} \pmod p}\frac{1}{r(\mathbf{h},p)}. \qquad (22)\end{aligned}$$

From Theorem 1, (22), and Lemma 2 we obtain

$$\begin{aligned} C(E_{p-1}(X), M, \mathbf{D}) &\le 2^k M D\big[X_{M+d_k}, M+d_k, \mathbf{D}\big] \\ &< \frac{k2^k}{p}M + 2^k M R_k(g,p,\mathbf{D}) + c_3 2^{k+1}p^{1/2}(\log p)\left(\frac{2}{\pi}\log p+\frac{7}{5}\right)^k \\ &< 2^k\left(k+2c_3p^{1/2}(\log p)\left(\frac{2}{\pi}\log p+\frac{7}{5}\right)^k\right) + 2^k M R_k(g,p,\mathbf{D}) \\ &< c_{10}p^{1/2}(\log p)\left(\frac{4}{\pi}\log p+\frac{14}{5}\right)^k + 2^k M R_k(g,p,\mathbf{D}) \end{aligned}$$

which completes the proof of Theorem 4. □

We remark that a theorem of type Theorem B, but with $R_k(g_0,p,\mathbf{D})$ (for *fixed* $\mathbf{D}$) in place of $R_s(g_0,p)$, could be proved similarly to the proof of Theorem 3.4 in [16], and the result obtained in this way could be combined with Theorem 4 above to get a "correlation analog" of Theorem C. However, the upper bounds would depend on $\mathbf{D}$ and, in particular, on d_k-d_1; if this difference is small, i.e., we are considering "short-range" correlation, then these bounds are relatively sharp, while if d_k-d_1 increases they get weaker, and if d_k-d_1 is large, i.e., we are considering "long-range" correlation, then they become trivial (as it is to be expected by (20) in the proof of Proposition 1).

5. From binary sequences to $[0,1)$ sequences

Suppose a binary sequence $E_N = (e_1, \dots, e_N) \in \{-1,+1\}^N$ is given. Then the most natural way to assign a $[0,1)$ sequence to it is the following. Consider a number $t \in \mathbb{N}$ which is "much smaller" than N (we will return to the size of it). Then let

$$y_i = \sum_{j=1}^{t} 2^{j-1} \frac{e_{(i-1)t+j}+1}{2} \quad \text{for } i = 1, 2, \dots, \left[\frac{N}{t}\right]$$

(so that $0 \le y_i < 2^t$ for all i),

$$x_i = \frac{y_i}{2^t} \quad \text{for } i = 1, 2, \dots, \left[\frac{N}{t}\right]$$

(so that $0 \le x_i < 1$ for all i), and

$$X = X(E_N, t) = (x_1, x_2, \dots, x_{[N/t]}).$$

One may hope that if E_N is a "good" PR binary sequence, then, at least for certain values, the $[0,1)$ sequence $X = X(E_N, t)$ also possesses strong PR properties. The question is how to choose the parameter t? If t is much smaller than $\frac{\log N}{\log 2}$, say, $t = o(\log N)$, then we may expect that a value $\frac{j}{2^t}$ occurs with a large frequency amongst the numbers x_i, so that X is certainly not of random type, its discrepancy is "not very small". On the other hand, if t is "much greater" than $\frac{\log N}{\log 2}$, then in general it is too difficult, usually hopelessly difficult to estimate the discrepancy. Thus, the optimal choice of t is about $\frac{\log N}{\log 2}$ (which is still difficult to handle).

The next question is: can one estimate the discrepancy $D(X(E_N, t))$ in terms of $W(E_N)$ and the $C_k(E_N)$, i.e., is it true that "small" $W(E_N)$ and $C_k(E_N)$ imply "small" $D(X(E_N, t))$? Consider the following example:

Example 2. Let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_t) \in \{-1,+1\}^t$ and $\eta = (\eta_1, \dots, \eta_{[N/t]}) \in \{-1,+1\}^{[N/t]}$ be two truly random binary sequences, and then define $E_N = (e_1, \dots, e_N) \in \{-1,+1\}^N$ by

$$e_n = e_{(i-1)t+j} = \eta_i \varepsilon_j \quad \text{for } 1 \le n = (i-1)t + j \le N,\ 1 \le i \le [N/t],\ 1 \le j \le t.$$

It is easy to see that for almost all of these sequences E_N, both W and, for all fixed k, C_k are "very small" ($< N^{1/2+\varepsilon}$). On the other hand, $X(E_N, t)$ contains only at most two distinct real numbers, namely, the numbers

$$\frac{1}{2} - \frac{1}{2^{t+1}} \pm \frac{1}{2^{t+1}} \sum_{j=1}^{t} 2^{j-1} \varepsilon_j;$$

thus clearly, $D(X(E_N, t))$ is "very large" (greater than a positive constant).

This example shows that it may occur that both $W(E_N)$ *and the* $C_k(E_N)$ *are small, however,* $D(X(E_N,t))$ *is large.*

On the other hand, in this example we clearly have

$$\left|\sum_{i=1}^{[N/t]} e_{(i-1)t+1}e_{(i-1)t+2}\right| = \left|\sum_{i=1}^{[N/t]} (\eta_i\varepsilon_1)(\eta_i\varepsilon_2)\right| = |[N/t]\varepsilon_1\varepsilon_2| = [N/t]$$

for $t \geq 2$, whence

$$Q_2(E_N) \geq [N/t],$$

so that *the combined PR measure of order* 2 *is large.* (This is a fact of independent interest: there exist binary sequences E_N such that *their* W *and* C_k *measures are small, but* Q_k *is large for some* k.)

This last remark inspires the following question: is it true that "small" $Q_k(E_N)$ imply "small" discrepancy $D(X(E_N,t))$? This time we will give an affirmative answer, i.e., we will give an upper bound for the discrepancy

$$D(X) = D(X(E_N,t)) = D(x_1, x_2, \ldots, x_{[N/t]})$$

in terms of the combined PR measures. (The higher-dimensional discrepancies of the type occurring in the serial test can be handled similarly, but the formulas and the computation become much longer and more complicated, thus we restrict ourselves to the study of the one-dimensional discrepancy.)

Theorem 5. *For any binary sequence* E_N *and any* $t \in \mathbb{N}$, $t < N$, *we have*

$$D(X) = D(X(E_N,t)) < \frac{1}{2^{t-1}} + \frac{2}{[N/t]}\sum_{v=1}^{t} Q_v(E_N).$$

PROOF. Note that each x_i, $i = 1, \ldots, [N/t]$, has the dyadic representation

$$x_i = \sum_{j=1}^{t} \frac{e_{it+1-j}+1}{2}\, 2^{-j}.$$

Now we apply Theorem 3.12 in [18] in the special one-dimensional case. This theorem provides an upper bound on the star discrepancy $D^*(X)$ of the sequence $X = X(E_N,t)$, and together with the well-known inequality $D(X) \leq 2D^*(X)$ (see [18, Proposition 2.4]) this yields

$$D(X) \leq \frac{1}{2^{t-1}} + \frac{2}{[N/t]} \sum_{\substack{\mathbf{h}\in\{0,1\}^t \\ \mathbf{h}\neq\mathbf{0}}} 2^{-d(\mathbf{h})}\left|\sum_{i=1}^{[N/t]} (-1)^{\sum_{j=1}^{t} h_j(e_{it+1-j}+1)/2}\right|.$$

Here, for a nonzero $\mathbf{h} = (h_1, \ldots, h_t) \in \{0,1\}^t$, we define $d(\mathbf{h})$ to be the largest value of j such that $h_j = 1$. We have

$$(-1)^{\sum_{j=1}^{t} h_j(e_{it+1-j}+1)/2} = \prod_{j=1}^{t} (-e_{it+1-j})^{h_j},$$

and so

$$\left| \sum_{i=1}^{[N/t]} (-1)^{\sum_{j=1}^{t} h_j(e_{it+1-j}+1)/2} \right| = \left| \sum_{i=1}^{[N/t]} \prod_{j=1}^{t} e_{it+1-j}^{h_j} \right|.$$

Therefore

$$D(X) \le \frac{1}{2^{t-1}} + \frac{2}{[N/t]} \sum_{d=1}^{t} 2^{-d} \sum_{\substack{\mathbf{h}\in\{0,1\}^t \\ d(\mathbf{h})=d}} \left| \sum_{i=1}^{[N/t]} \prod_{j=1}^{t} e_{it+1-j}^{h_j} \right|.$$

For $\mathbf{h} = (h_1, \ldots, h_t) \in \{0,1\}^t$ with $d(\mathbf{h}) = d$, let $1 \le j_1 < \cdots < j_v = d$ be those values of j with $h_j = 1$. Then

$$\left| \sum_{i=1}^{[N/t]} \prod_{j=1}^{t} e_{it+1-j}^{h_j} \right| = \left| \sum_{i=1}^{[N/t]} e_{it+1-j_1} \cdots e_{it+1-j_v} \right| \le Q_v(E_N).$$

It follows that

$$\begin{aligned} D(X) &\le \frac{1}{2^{t-1}} + \frac{2}{[N/t]} \sum_{d=1}^{t} 2^{-d} \sum_{v=1}^{d} \binom{d-1}{v-1} Q_v(E_N) \\ &= \frac{1}{2^{t-1}} + \frac{2}{[N/t]} \sum_{v=1}^{t} Q_v(E_N) \sum_{d=v}^{t} \binom{d-1}{v-1} 2^{-d}. \end{aligned}$$

For the last inner sum we obtain

$$\sum_{d=v}^{t} \binom{d-1}{v-1} 2^{-d} = 2^{-v} \sum_{d=0}^{t-v} \binom{d+v-1}{v-1} 2^{-d} < 2^{-v} \sum_{d=0}^{\infty} \binom{d+v-1}{v-1} 2^{-d}.$$

Note that for $|z| < 1$ and any $v \in \mathbb{N}$ we have

$$\sum_{d=0}^{\infty} \binom{d+v-1}{v-1} z^d = (1-z)^{-v},$$

and so

$$\sum_{d=0}^{\infty} \binom{d+v-1}{v-1} 2^{-d} = 2^v.$$

This yields

$$\sum_{d=v}^{t} \binom{d-1}{v-1} 2^{-d} < 1,$$

and the proof is complete. □

6. From binary sequences to $[0,1)$ sequences in a special case

The most important PR binary sequences are, perhaps, the Legendre symbol sequences $E_{p-1} = \{e_1, e_2, \dots, e_{p-1}\}$ defined by

$$e_n = \left(\frac{n}{p}\right) \quad \text{for } n = 1, 2, \dots, p-1, \tag{23}$$

where p is a prime number. These sequences were also studied by Mauduit and Sárközy in [13] who proved [13, Theorem 1] that

Theorem D. *There is a number p_0 such that if $p > p_0$ is a prime number, $k \in \mathbb{N}$, $k < p$, and E_{p-1} is the Legendre symbol sequence defined above, then we have*

$$Q_k(E_{p-1}) \leq 9kp^{1/2} \log p.$$

We will give the following upper bound for the discrepancy of the $[0,1)$ sequence induced by the Legendre symbol sequence:

Theorem 6. *If $p > p_0$ is a prime and $t < p-1$, then for the sequence E_{p-1} defined by* (23) *we have*

$$D(X) = D(X(E_{p-1}, t)) < \frac{1}{2^{t-1}} + 72t^3 \frac{\log p}{p^{1/2}}.$$

Taking here

$$t = \left[\frac{1}{2\log 2} \log p - \frac{4}{\log 2} \log\log p\right] \tag{24}$$

(approximatively this gives the best upper bound for $D(X)$), we obtain with a little computation (we leave the details to the reader) that

Corollary 2. *For any prime p and for the t defined by* (24) *we have*

$$D(X) = D(X(E_{p-1}, t)) < c_{11} \frac{(\log p)^4}{p^{1/2}}.$$

Proof of Theorem 6. It follows from Theorem 5 by using Theorem D that

$$\begin{aligned} D(X) = D(X(E_{p-1}, t)) &< \frac{1}{2^{t-1}} + \frac{2}{[(p-1)/t]} \sum_{v=1}^{t} Q_v(E_{p-1}) \\ &\leq \frac{1}{2^{t-1}} + \frac{4}{(p-1)/t} \sum_{v=1}^{t} 9vp^{1/2} \log p \leq \frac{1}{2^{t-1}} + \frac{36t}{(p-1)} p^{1/2} (\log p) \sum_{v=1}^{t} v \\ &\leq \frac{1}{2^{t-1}} + 72t^3 \frac{\log p}{p^{1/2}} \end{aligned}$$

which completes the proof of Theorem 6. □

Acknowledgments. This research is partially supported by the Hungarian National Foundation for Scientific Research, Grants No. T 043623 and T 049693, and by the French–Hungarian EGIDE–OMKFHÁ exchange program Balaton F-2/03. The research of the second author is supported by a DSTA grant with Temasek Laboratories in Singapore. This paper was written while the second and third author were visiting the Institut de Mathématiques de Luminy, Marseille.

References

[1] N. Alon, Y. Kohayakawa, C. Mauduit, C. G. Moreira and V. Rödl, Measures of pseudorandomness for finite sequences: typical values, *preprint.*

[2] W. D. Banks, A. Conflitti, J. B. Friedlander and I. E. Shparlinski, Exponential sums over Mersenne numbers, *Compositio Math.* **140** (2004), 15–30.

[3] J. Cassaigne, C. Mauduit and A. Sárközy, On finite pseudorandom binary sequences VII: The measures of pseudorandomness, *Acta Arith.* **103** (2002), 97–118.

[4] J. E. Gentle, Random Number Generation and Monte Carlo Methods, 2nd ed., *Springer*, *New York*, 2003.

[5] L. Goubin, C. Mauduit and A. Sárközy, Construction of large families of pseudorandom binary sequences, *J. Number Theory* **106** (2004), 56–69.

[6] K. Gyarmati, An inequality between the measures of pseudorandomness, *Annales Univ. Sci. Budapest. Eötvös Sect. Math.* **46** (2003), 157–166.

[7] K. Gyarmati, On the correlation of binary sequences, *Studia Sci. Math. Hungar.* **42** (2005), 79–93.

[8] D. E. Knuth, The Art of Computer Programming, vol. 2: Seminumerical algorithms, 3rd ed., *Addison-Wesley*, *Reading, MA*, 1998.

[9] N. M. Korobov, Exponential Sums and their Applications, *Kluwer Academic Publishers*, *Dordrecht, Boston, London*, 1992.

[10] L. Kuipers and H. Niederreiter, Uniform Distribution of Sequences, *Wiley*, *New York*, 1974.

[11] G. Marsaglia, The structure of Linear Congruential Sequences, Applications of Number Theory to Numerical Analysis, (S. K. Zaremba, ed.), *Academic Press*, *New York*, 1972, 249–285.

[12] C. Mauduit, J. Rivat and A. Sárközy, Construction of pseudorandom binary sequences using additive characters, *Monatshefte Math.* **141** (2004), 197–208.

[13] C. Mauduit and A. Sárközy, On finite pseudorandom binary sequences, I. Measure of pseudorandomness, the Legendre symbol, *Acta Arith.* **82** (1997), 365–377.

[14] C. Mauduit and A. Sárközy, On the measures of pseudorandomness of binary sequences, *Discrete Math.* **271** (2003), 195–207.

[15] C. Mauduit and A. Sárközy, Construction of pseudorandom binary sequences by using the multiplicative inverse, *Acta Math. Hungar.* **108** (2005), 239–252.

[16] H. Niederreiter, Pseudo-random numbers and optimal coefficients, *Advances in Math.* **26** (1977), 99–181.

[17] H. Niederreiter, The serial test for pseudo-random numbers generated by the linear congruential method, *Numer. Math.* **46** (1985), 51–68.

[18] H. Niederreiter, Random number generation and quasi-Monte Carlo methods, *Society for Industrial and Applied Mathematics*, *Philadelphia, PA*, 1992.

[19] H. Niederreiter, New Developments in Uniform Pseudorandom Number and Vector Generation, Vol. 106, Lect. Notes in Statistics, (H. Niederreiter and P. J.-S. Shiue, eds.), *Springer*, *New York*, 1995, 87–120.

[20] H. Niederreiter, Some Computable Complexity Measures for Binary Sequences, Sequences and their Applications, (C. Ding, T. Helleseth and H. Niederreiter, eds.), *Springer*, *London*, 1999, 67–78.

[21] H. Niederreiter and I. E. Shparlinski, Recent Advances in the Theory of Nonlinear Pseudorandom Number Generators, Monte Carlo and Quasi-Monte Carlo Methods 2000, (K.-T. Fang, F. J. Hickernell and H. Niederreiter, eds.), *Springer*, *Berlin*, 2002, 86–102.

[22] R. A. Rueppel, Stream Ciphers, Contemporary Cryptology: The Science of Information Integrity, (G. J. Simmons, ed.), *IEEE Press*, *New York*, 1992, 65–134.

[23] A. Sárközy, A finite pseudorandom binary sequence, *Studia Sci. Math. Hungar.* **38** (2001), 377–384.

CHRISTIAN MAUDUIT
INSTITUT DE MATHÉMATIQUES DE LUMINY
CNRS, UMR 6206
163, AVENUE DE LUMINY, CASE 907
13288 MARSEILLE CEDEX 9
FRANCE

E-mail: mauduit@iml.univ-mrs.fr

HARALD NIEDERREITER
DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
2 SCIENCE DRIVE 2, SINGAPORE 117543
REPUBLIC OF SINGAPORE

E-mail: nied@math.nus.edu.sg

ANDRÁS SÁRKÖZY
EÖTVÖS LORÁND UNIVERSITY
DEPARTMENT OF ALGEBRA AND NUMBER THEORY
H–1117 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/C
HUNGARY

E-mail: sarkozy@cs.elte.hu

(Received February 9, 2006)