ACTA ARITHMETICA
133.2 (2008)

Pseudorandom sequences of binary vectors

by

Harald Niederreiter (Singapore), Joël Rivat (Marseille)
and András Sárközy (Budapest)

Dedicated to W. M. Schmidt on the occasion of his 75th birthday

1. Introduction. In this paper our goal is to study pseudorandomness of sequences

$$\mathbf{U}_M^{(k)} = (\mathbf{u}_1, \dots, \mathbf{u}_M) \tag{1}$$

of k-dimensional binary vectors $\mathbf{u}_n \in \{-1, +1\}^k$. In one dimension a constructive and quantitative theory of pseudorandomness has been developed in two directions: pseudorandomness of sequences of real numbers from the interval $[0, 1)$ (see e.g. [2, 15, 16, 17, 20]) and pseudorandomness of binary sequences of elements of $\{-1, +1\}$ (see e.g. [1, 3, 9, 10, 12, 13, 23]) have been studied; see also [8] for the connection between the two fields. We will show that these two theories can be extended and combined to study the pseudorandomness of sequences of binary vectors. In particular, we will show that the measures of pseudorandomness used in the two fields can be extended in this direction, we will present principles to extend the one-dimensional constructions and utilize the one-dimensional estimates in this multidimensional situation, and we will also study special constructions obtained by using these principles.

2. Discrepancy, a construction principle, examples. In the theory of pseudorandomness of *real numbers* the most frequently used measure

2000 *Mathematics Subject Classification*: Primary 11K45; Secondary 11L07, 11L40.
Key words and phrases: pseudorandomness, binary vectors, discrepancy, correlation.
Research supported by the "Balaton" French-Hungarian exchange program F-48/2006. Research of the first author partially supported by the project NUGET of the Agence Nationale de la Recherche (France). Research of the third author partially supported by the Hungarian National Foundation for Scientific Research, Grants No. K 67676 and T 049 693.

of pseudorandomness is the discrepancy. This can be easily extended to sequences of *binary vectors*.

Consider the sequence (1) of k-dimensional binary vectors (with $\mathbf{u}_n \in \{-1,+1\}^k$), and for $\mathbf{c} \in \{-1,+1\}^k$ write

$$A(\mathbf{u}_M^{(k)}, \mathbf{c}) = \#\{n \in \mathbb{Z} : 1 \le n \le M,\ \mathbf{u}_n = \mathbf{c}\}.$$

DEFINITION 1. The *discrepancy* of the sequence $\mathbf{U}_M^{(k)}$ is defined as

$$D(\mathbf{U}_M^{(k)}) = \max_{\mathbf{c}\in\{-1,+1\}^k} \left| A(\mathbf{U}_M^{(k)}, \mathbf{c}) - \frac{M}{2^k} \right|.$$

Now we will present a general principle for constructing pseudorandom sequences of binary vectors. For a prime power q, we let $\mathbb{F}_q$ denote the finite field of order q. We refer to [7, Chapter 5] and [25, Chapter II] for the theory of characters of finite fields.

CONSTRUCTION 1. Let $h, k, M \in \mathbb{N}$, $k \le h$, $\gamma_1, \dots, \gamma_M \in \mathbb{F}_{2^h}$, let $\psi : \mathbb{F}_{2^h} \to \{-1,+1\}$ be a nontrivial additive character of $\mathbb{F}_{2^h}$, and let $\beta^{(1)}, \dots, \beta^{(k)} \in \mathbb{F}_{2^h}$ be linearly independent over $\mathbb{F}_2$. Then define the sequence $\mathbf{U}_M^{(k)} = (\mathbf{u}_1, \dots, \mathbf{u}_M)$ by

$$\mathbf{u}_n = (\psi(\beta^{(1)}\gamma_n), \dots, \psi(\beta^{(k)}\gamma_n)) \in \{-1,+1\}^k \quad \text{for } n = 1, \dots, M. \tag{2}$$

THEOREM 1. *For the sequence* $\mathbf{U}_M^{(k)} = (\mathbf{u}_1, \dots, \mathbf{u}_M)$ *defined by* (2), *we have*

$$D(\mathbf{U}_M^{(k)}) \le \left(1 - \frac{1}{2^k}\right) \max_{\varphi \ne \varphi_0} \left| \sum_{n=1}^{M} \varphi(\gamma_n) \right|, \tag{3}$$

where the maximum is taken over all additive characters φ *of* $\mathbb{F}_{2^h}$ *different from the trivial character* φ_0.

Proof. Let $\mathbf{c} = (c^{(1)}, \dots, c^{(k)}) \in \{-1,+1\}^k$. Then by (2), $\mathbf{u}_n = \mathbf{c}$ if and only if

$$\psi(\beta^{(i)}\gamma_n) = c^{(i)} \quad \text{for } i = 1, \dots, k.$$

Since $c^{(i)} \in \{-1,+1\}$ and $\psi : \mathbb{F}_{2^h} \to \{-1,+1\}$ is nontrivial, there is thus an $\alpha^{(i)} \in \mathbb{F}_{2^h}$ with $c^{(i)} = \psi(\alpha^{(i)})$. Then $\mathbf{u}_n = \mathbf{c}$ if and only if

$$\psi(\beta^{(i)}\gamma_n + \alpha^{(i)}) = 1 \quad \text{for } i = 1, \dots, k.$$

It follows that

$$A(\mathbf{U}_M^{(k)}, \mathbf{c}) = \frac{1}{2^k} \sum_{n=1}^{M} \prod_{i=1}^{k} (1 + \psi(\beta^{(i)}\gamma_n + \alpha^{(i)})),$$

whence

$$(4)\qquad \left|A(\mathbf{U}_M^{(k)},\mathbf{c})-\frac{M}{2^k}\right|$$

$$=\frac{1}{2^k}\Big|\sum_{r=1}^{k}\sum_{1\le i_1<\cdots<i_r\le k}\sum_{n=1}^{M}\psi((\beta^{(i_1)}+\cdots+\beta^{(i_r)})\gamma_n+(\alpha^{(i_1)}+\cdots+\alpha^{(i_r)}))\Big|$$

$$\le\frac{1}{2^k}\sum_{r=1}^{k}\sum_{1\le i_1<\cdots<i_r\le k}\Big|\sum_{n=1}^{M}\psi((\beta^{(i_1)}+\cdots+\beta^{(i_r)})\gamma_n+(\alpha^{(i_1)}+\cdots+\alpha^{(i_r)}))\Big|$$

$$=\frac{1}{2^k}\sum_{r=1}^{k}\sum_{1\le i_1<\cdots<i_r\le k}\Big|\sum_{n=1}^{M}\psi((\beta^{(i_1)}+\cdots+\beta^{(i_r)})\gamma_n)\Big|.$$

Now consider the innermost sum. Since $\beta^{(1)},\dots,\beta^{(k)}$ are linearly independent over $\mathbb{F}_2$, we have

$$\beta^{(i_1)}+\cdots+\beta^{(i_r)}\neq 0.$$

Thus if we write

$$\varphi(\gamma)=\psi((\beta^{(i_1)}+\cdots+\beta^{(i_r)})\gamma)\quad\text{ for all }\gamma\in\mathbb{F}_{2^h},$$

then φ is a nontrivial additive character of $\mathbb{F}_{2^h}$, and the innermost sum in (4) can be rewritten as

$$\sum_{n=1}^{M}\varphi(\gamma_n).$$

Thus, it follows from (4) that

$$\begin{aligned}D(\mathbf{U}_M^{(k)})&=\max_{\mathbf{c}\in\{-1,+1\}^k}\left|A(\mathbf{U}_M^{(k)},\mathbf{c})-\frac{M}{2^k}\right|\\&\le\frac{1}{2^k}\sum_{r=1}^{k}\sum_{1\le i_1<\cdots<i_r\le k}\max_{\varphi\neq\varphi_0}\Big|\sum_{n=1}^{M}\varphi(\gamma_n)\Big|\\&=\frac{2^k-1}{2^k}\max_{\varphi\neq\varphi_0}\Big|\sum_{n=1}^{M}\varphi(\gamma_n)\Big|,\end{aligned}$$

which completes the proof of Theorem 1. ∎

There are certain important special sequences $\gamma_1,\dots,\gamma_M$ which can be generated in a fast and simple way, and for which good estimates are known for the maximum

$$m_M:=\max_{\varphi\neq\varphi_0}\Big|\sum_{n=1}^{M}\varphi(\gamma_n)\Big|$$

in (3).

EXAMPLE 1. Let $\gamma_1, \gamma_2, \ldots \in \mathbb{F}_{2^h}$ be an inversive sequence, that is, with some $\alpha, \beta \in \mathbb{F}_{2^h}$, $\alpha \neq 0$, we have

$$\gamma_n = \begin{cases} \alpha\gamma_{n-1}^{-1} + \beta & \text{if } \gamma_{n-1} \neq 0, \\ \beta & \text{if } \gamma_{n-1} = 0, \end{cases} \tag{5}$$

for $n = 2, 3, \ldots$, with an initial value γ_1. Let t be the least period of this sequence. Then by Theorem 5 (with $s = 1$) in [19] we have

$$m_M \leq 2.1\, M^{1/2} 2^{h/4} + 2^{h/2} \quad \text{for } 1 \leq M \leq t.$$

Note that t can be as large as 2^h.

EXAMPLE 2. For arbitrary $r \in \mathbb{N}$, let $\gamma_1, \gamma_2, \ldots \in \mathbb{F}_{2^h}$ be an rth-order linear recurring sequence which is purely periodic with least period t. Then by Theorem 3 and Lemma 2 in [14] we have

$$m_M \leq 2^{hr/2}(1 + \log t) \quad \text{for } 1 \leq M \leq t.$$

Note that t can be as large as $2^{hr} - 1$.

3. The well-distribution measure and the correlation measure. Now we will consider binary sequences of the form

$$E_N = (e_1, \ldots, e_N) \in \{-1, +1\}^N. \tag{6}$$

Mauduit and Sárközy [10] introduced the following measures of pseudorandomness for sequences of this type.

DEFINITION 2. The *well-distribution measure* of the sequence (6) is defined by

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

where the maximum is taken over all $a, b, t \in \mathbb{N}$ such that $1 \leq a \leq a + (t-1)b \leq N$.

DEFINITION 3. The *correlation measure of order s* of the sequence (6) is defined as

$$C_s(E_N) = \max_{M,D} \Big| \sum_{n=1}^{M} e_{n+d_1} \cdots e_{n+d_s} \Big|,$$

where the maximum is taken over all $D = (d_1, \ldots, d_s) \in \mathbb{Z}^s$ and $M \in \mathbb{N}$ such that $0 \leq d_1 < \cdots < d_s \leq N - M$.

DEFINITION 4. The *combined* (well-distribution-correlation) *measure of order s* of the sequence (6) is defined as

$$Q_s(E_N) = \max_{a,t,D} \Big| \sum_{j=0}^{t-1} e_{ja+d_1} \cdots e_{ja+d_s} \Big|,$$

where the maximum is taken over all $D = (d_1, \dots, d_s)$, a, and t such that $d_1, \dots, d_s, a, t \in \mathbb{N}$ and $d_1 < \cdots < d_s$, $(t-1)a + d_s \le N$.

In several papers, constructions have been given for binary sequences with strong pseudorandomness properties in terms of these measures of pseudorandomness. The best sequences $E_N = (e_1, \dots, e_N)$ constructed are, perhaps, the following ones.

EXAMPLE 3. Mauduit and Sárközy [10] studied the sequence defined by

$$e_n = \left(\frac{n}{p}\right) \quad \text{for } n = 1, \dots, p-1, \tag{7}$$

where p is an odd prime and $\left(\frac{n}{p}\right)$ denotes the Legendre symbol. Goubin, Mauduit, and Sárközy [3] extended (7) to

$$e_n = \begin{cases} \left(\dfrac{f(n)}{p}\right) & \text{for } 1 \le n \le p,\ \gcd(f(n), p) = 1, \\ +1 & \text{for } 1 \le n \le p,\ p \mid f(n), \end{cases} \tag{8}$$

for certain polynomials $f \in \mathbb{F}_p[x]$. (See also [22, 24].)

EXAMPLE 4. Sárközy [23] studied the sequence defined by

$$e_n = \begin{cases} +1 & \text{if } 1 \le \operatorname{ind} n \le (p-1)/2, \\ -1 & \text{if } (p+1)/2 \le \operatorname{ind} n \le p-1, \end{cases} \tag{9}$$

for $n = 1, \dots, p-1$, where p is an odd prime and $\operatorname{ind} n$ denotes the mod p index (or discrete logarithm) of n with respect to a fixed primitive root mod p. Gyarmati [4, 5, 6] extended this to

$$e_n = \begin{cases} +1 & \text{if } \gcd(f(n), p) = 1 \text{ and } 1 \le \operatorname{ind} f(n) \le (p-1)/2, \\ -1 & \text{if } \gcd(f(n), p) = 1 \text{ and } (p+1)/2 \le \operatorname{ind} f(n) \le p-1 \\ & \quad \text{or } p \mid f(n), \end{cases} \tag{10}$$

for $n = 1, \dots, p$, with certain polynomials $f \in \mathbb{F}_p[x]$.

EXAMPLE 5. Mauduit and Sárközy [13] studied the sequence with

$$e_n = \begin{cases} +1 & \text{if } \gcd(f(n), p) = 1 \text{ and } 1 \le (f(n))^{-1} < p/2, \\ -1 & \text{if } \gcd(f(n), p) = 1 \text{ and } p/2 < (f(n))^{-1} < p \\ & \quad \text{or } p \mid f(n), \end{cases} \tag{11}$$

for $n = 1, \dots, p$, with certain polynomials $f \in \mathbb{F}_p[x]$, where p is an odd prime and b^{-1} denotes the multiplicative inverse of $b \in \mathbb{F}_p^*$.

In each of the three examples above, it has been shown that both $W(E_N)$ and (for small s) $C_s(E_N)$ are small. (See also [9] and [21].)

Later Mauduit and Sárközy [11] extended the study of pseudorandomness of binary sequences to sequences over a set of K symbols. Let $K \in \mathbb{N}$,

$K \geq 2$, let $\mathcal{A} = \{a_1, \ldots, a_K\}$ be a finite set ("alphabet") of K symbols ("letters"), and consider a sequence

$$E_N = (e_1, \ldots, e_N) \in \mathcal{A}^N \tag{12}$$

of these symbols. Then a natural requirement towards pseudorandomness is that any fixed l-tuple ("word") $(a_{i_1}, \ldots, a_{i_l}) \in \mathcal{A}^l$ occurs with the expected *frequency* in certain positions in E_N. This approach leads to Definitions 5 and 6 below. Write

$$x(E_N, a, M, u, v) = \#\{j \in \mathbb{Z} : 0 \leq j \leq M-1,\ e_{u+jv} = a\}$$

for $a \in \mathcal{A}$ and $u, v, M \in \mathbb{N}$ with $u+(M-1)v \leq N$, and for $w = (a_{i_1}, \ldots, a_{i_l}) \in \mathcal{A}^l$ and $D = (d_1, \ldots, d_l)$ with nonnegative integers $d_1 < \cdots < d_l \leq N - M$,

$$g(E_N, w, M, D) = \#\{n \in \mathbb{Z} : 1 \leq n \leq M,\ (e_{n+d_1}, \ldots, e_{n+d_l}) = w\}.$$

Definition 5. The *f-well distribution* ("f" for "frequency") *measure* of the sequence (12) is defined as

$$\delta(E_N) = \max_{a,M,u,v} \left| x(E_N, a, M, u, v) - \frac{M}{K} \right|,$$

where the maximum is taken over all $a \in \mathcal{A}$ and $u, v, M \in \mathbb{N}$ with $u + (M-1)v \leq N$.

Definition 6. The *f-correlation measure of order* l of the sequence (12) is defined as

$$\gamma_l(E_N) = \max_{w,M,D} \left| g(E_N, w, M, D) - \frac{M}{K^l} \right|,$$

where the maximum is taken over all $w \in \mathcal{A}^l$, $D = (d_1, \ldots, d_l)$, and M such that $d_1, \ldots, d_l \in \mathbb{Z}$, $M \in \mathbb{N}$, and $0 \leq d_1 < \cdots < d_l \leq N - M$.

Note that in [11] another type of well-distribution measure, resp. correlation measure, was also introduced, but then it was shown that the corresponding types of measures are nearly equivalent. Besides, the definitions above are more suitable for our purpose, thus we will use only these definitions.

Observe that our problem described in Section 1 is a special case of this K-symbol situation: the sequence in (1) is composed of k-dimensional vectors in $\{-1, +1\}^k$, so now these 2^k vectors are the "symbols". Correspondingly, we may adapt Definitions 5 and 6 (with $\{-1, +1\}^k$ in place of $\mathcal{A}$) to study sequences of binary vectors of the type (1).

The f-well distribution measure δ is closely related to the discrepancy measure introduced in Section 2, but it is slightly more general than that. However, its estimation is usually similar to the estimation of the discrepancy. On the other hand, the estimation of the f-correlation measure is usually much more difficult. Thus, in many cases the best we can do is

to estimate the discrepancy (or δ). Typically, this is the case in recursive constructions where it is usually very difficult to control the "long-range" correlation. However, in the next section we will present a principle to extend one-dimensional constructions to k-dimensional constructions, so that if there are good estimates for the measures of pseudorandomness in one dimension (as in Examples 3–5), then one may expect that the k-dimensional measures can be estimated equally well. We will show that this is the case when extending Examples 3–5. We will also present a one-dimensional construction which is of recursive type, but the measures of pseudorandomness (including the correlation) are well controlled, and this property is also preserved in its k-dimensional extension.

4. Constructing sequences of binary vectors from binary sequences. Now we will show that if

$$E_N = (e_1, \dots, e_N) \in \{-1, +1\}^N \tag{13}$$

is a binary sequence with strong pseudorandomness properties and $k \in \mathbb{N}$ (we assume that k is fixed and N is "large", or at least k grows slowly in terms of N), then starting out from E_N one can construct a sequence

$$\mathbf{U} = \mathbf{U}(E_N) = (\mathbf{u}_1, \dots, \mathbf{u}_M) \tag{14}$$

of k-dimensional binary *vectors* which has pseudorandomness properties nearly equal to those of E_N.

Construction 2. Let $M = \lfloor N/k \rfloor$ and define the vectors $\mathbf{u}_n$ in (14) for $n = 1, \dots, M$ by

$$\mathbf{u}_n = (u_n^{(1)}, \dots, u_n^{(k)}), \quad u_n^{(j)} = e_{(n-1)k+j} \quad \text{for } j = 1, \dots, k. \tag{15}$$

A related idea was used also in [8]. We will show that $\delta(\mathbf{U}(E_N))$ and $\gamma_l(\mathbf{U}(E_N))$ can be estimated in terms of the measures of pseudorandomness of E_N if $\mathbf{U}(E_N)$ is given by (14) and (15).

Theorem 2. *For every binary sequence E_N of the form* (13), *we have*

$$\delta(\mathbf{U}(E_N)) \le \frac{1}{2^k} \sum_{r=1}^{k} \binom{k}{r} Q_r(E_N), \tag{16}$$

where $Q_r(E_N)$ is the combined measure of order r of the sequence E_N.

Proof. If $\mathbf{u}_n$ is of the form (15) and

$$\mathbf{a} = (\varepsilon_1, \dots, \varepsilon_k) \in \{-1, +1\}^k, \tag{17}$$

then clearly

$$\frac{1}{2^k} \prod_{j=1}^{k} (1 + u_n^{(j)} \varepsilon_j) = \frac{1}{2^k} \prod_{j=1}^{k} (1 + e_{(n-1)k+j} \varepsilon_j) = \begin{cases} 1 & \text{if } \mathbf{u}_n = \mathbf{a}, \\ 0 & \text{if } \mathbf{u}_n \ne \mathbf{a}. \end{cases} \tag{18}$$

It follows that

$$x(\mathbf{U}(E_N), \mathbf{a}, H, s, t)$$
$$= \sum_{n=0}^{H-1} \frac{1}{2^k} \prod_{j=1}^{k} (1 + e_{(s+nt-1)k+j}\varepsilon_j)$$
$$= \frac{H}{2^k} + \frac{1}{2^k} \sum_{r=1}^{k} \sum_{1 \le j_1 < \cdots < j_r \le k} \varepsilon_{j_1} \cdots \varepsilon_{j_r} \sum_{n=0}^{H-1} e_{(s+nt-1)k+j_1} \cdots e_{(s+nt-1)k+j_r},$$

whence

$$\left| x(\mathbf{U}(E_N), \mathbf{a}, H, s, t) - \frac{H}{2^k} \right|$$
$$\le \frac{1}{2^k} \sum_{r=1}^{k} \sum_{1 \le j_1 < \cdots < j_r \le k} \left| \sum_{n=0}^{H-1} e_{(s+nt-1)k+j_1} \cdots e_{(s+nt-1)k+j_r} \right|$$
$$\le \frac{1}{2^k} \sum_{r=1}^{k} \sum_{1 \le j_1 < \cdots < j_r \le k} Q_r(E_N) = \frac{1}{2^k} \sum_{r=1}^{k} \binom{k}{r} Q_r(E_N),$$

which proves (16). ∎

THEOREM 3. *For $l \in \mathbb{N}$, $l \ge 2$, and every binary sequence E_N of the form* (13), *we have*

$$\gamma_l(\mathbf{U}(E_N)) \le \frac{1}{2^{kl}} \sum_{r=1}^{kl} \binom{kl}{r} Q_r(E_N), \tag{19}$$

where $Q_r(E_N)$ is the combined measure of order r of the sequence E_N.

Proof. Write

$$\mathbf{w} = (\mathbf{v}_1, \ldots, \mathbf{v}_l)$$

with

$$\mathbf{v}_s = (\varepsilon_1^{(s)}, \ldots, \varepsilon_k^{(s)}) \in \{-1, +1\}^k \quad \text{for } 1 \le s \le l.$$

If $\mathbf{u}_i$ is of the form (15), then by (18) we have

$$g(\mathbf{U}(E_N), \mathbf{w}, H, D)$$
$$= \#\{n \in \mathbb{Z} : 1 \le n \le H, (\mathbf{u}_{n+d_1}, \ldots, \mathbf{u}_{n+d_l}) = (\mathbf{v}_1, \ldots, \mathbf{v}_l)\}$$
$$= \sum_{n=1}^{H} \frac{1}{2^{kl}} \prod_{s=1}^{l} \prod_{j=1}^{k} (1 + e_{(n+d_s-1)k+j}\varepsilon_j^{(s)}),$$

whence writing

$$\mathcal{G} = \{(s, j) : s \in \{1, \ldots, l\}, j \in \{1, \ldots, k\}\}$$

we obtain

$$\begin{aligned}
&\Big|g(\mathbf{U}(E_N),\mathbf{w},H,D)-\frac{H}{2^{kl}}\Big| \\
&\qquad=\frac{1}{2^{kl}}\Big|\sum_{n=1}^{H}\sum_{\substack{\mathcal{F}\subseteq\mathcal{G}\\ \mathcal{F}\neq\emptyset}}\prod_{(s,j)\in\mathcal{F}}\varepsilon_j^{(s)}e_{(n+d_s-1)k+j}\Big| \\
&\qquad=\frac{1}{2^{kl}}\Big|\sum_{\substack{\mathcal{F}\subseteq\mathcal{G}\\ \mathcal{F}\neq\emptyset}}\Big(\prod_{(s,j)\in\mathcal{F}}\varepsilon_j^{(s)}\Big)\Big(\sum_{n=1}^{H}\prod_{(s,j)\in\mathcal{F}}e_{(n+d_s-1)k+j}\Big)\Big| \\
&\qquad\le\frac{1}{2^{kl}}\sum_{\substack{\mathcal{F}\subseteq\mathcal{G}\\ \mathcal{F}\neq\emptyset}}\Big|\sum_{n=1}^{H}\prod_{(s,j)\in\mathcal{F}}e_{(n+d_s-1)k+j}\Big|\le\frac{1}{2^{kl}}\sum_{\substack{\mathcal{F}\subseteq\mathcal{G}\\ \mathcal{F}\neq\emptyset}}Q_{|\mathcal{F}|}(E_N) \\
&\qquad=\frac{1}{2^{kl}}\sum_{r=1}^{|\mathcal{G}|}\binom{|\mathcal{G}|}{r}Q_r(E_N)=\frac{1}{2^{kl}}\sum_{r=1}^{kl}\binom{kl}{r}Q_r(E_N),
\end{aligned}$$

which proves (19). ∎

By Theorems 2 and 3, the estimation of the measures of pseudorandomness of binary vector sequences $\mathbf{U}=\mathbf{U}(E_N)$ of the type described in (14) and (15) can be reduced to the estimation of the combined measures of E_N. In many cases, the upper bounds on these measures are similar to the upper bounds on the correlation measures and, in particular, this is the situation in Examples 3–5. First we will study Example 3. Since the estimation will be very much similar to the estimation of the correlation measure, we will give a sketch of the proof only.

THEOREM 4. *Assume that p is an odd prime, $f\in\mathbb{F}_p[x]$ is a polynomial of degree $D\ge 1$ which has no multiple root in $\overline{\mathbb{F}}_p$, and the binary sequence E_p is defined by* (8). *Let $r\in\mathbb{N}$ and suppose that one of the following conditions holds*:

(i) $r=2$;
(ii) $(4r)^D<p$;
(iii) 2 *is a primitive root mod p and $r<p$, $D<p$.*

Then for the combined measure $Q_r(E_p)$ of order r of the sequence E_p we have

$$Q_r(E_p)<10Drp^{1/2}\log p. \tag{20}$$

Combining Theorems 2 and 4, resp. Theorems 3 and 4, we obtain the following two corollaries.

COROLLARY 1. *Assume that $k \in \mathbb{N}$ with $k \geq 2$, that f, D, and E_p are defined as in Theorem 4, and that one of the following conditions holds*:

(i) $k = 2$;
(ii) $(4k)^D < p$;
(iii) 2 *is a primitive root mod* p *and* $k < p$, $D < p$.

Then for the sequence $\mathbf{U}(E_p)$ *defined by* (14) *and* (15) *we have*

$$\delta(\mathbf{U}(E_p)) < \frac{10D}{2^k} \sum_{r=1}^{k} r \binom{k}{r} p^{1/2} \log p = 5Dkp^{1/2} \log p.$$

COROLLARY 2. *Assume that $k, l \in \mathbb{N}$ with $k \geq 2$ and $l \geq 2$, that f, D, and E_p are defined as in Theorem 4, and that one of the following conditions holds*:

(i) $(4kl)^D < p$;
(ii) 2 *is a primitive root mod* p *and* $kl < p$, $D < p$.

Then for the sequence $\mathbf{U}(E_p)$ *defined by* (14) *and* (15) *we have*

$$\gamma_l(\mathbf{U}(E_p)) < \frac{10D}{2^{kl}} \sum_{r=1}^{kl} r \binom{kl}{r} p^{1/2} \log p = 5Dklp^{1/2} \log p.$$

Proof of Theorem 4. We shall first show that combining Theorems 1 and 2 in [3], under the assumptions of Theorem 4 here, one gets

$$C_r(E_p) < 10Drp^{1/2} \log p, \tag{21}$$

so that we obtain the same upper bound for C_r as the one for Q_r in (20) to be proved. The proof of (20) is also very similar to the proofs in [3]. Indeed write $f = bf_1$ with $b \in \mathbb{F}_p^*$, where f_1 is a monic polynomial. Then, on putting $\left(\frac{0}{p}\right) = 0$, the proof in [3] starts with

$$\begin{aligned} \left| \sum_{n=1}^{H} e_{n+d_1} \cdots e_{n+d_r} \right| &\leq \left| \sum_{n=1}^{H} \left(\frac{f(n+d_1)}{p} \right) \cdots \left(\frac{f(n+d_r)}{p} \right) \right| + Dr \\ &= \left| \left(\frac{b^r}{p} \right) \sum_{n=1}^{H} \left(\frac{f_1(n+d_1) \cdots f_1(n+d_r)}{p} \right) \right| + Dr \\ &= \left| \sum_{n=1}^{H} \left(\frac{h(n)}{p} \right) \right| + Dr, \end{aligned} \tag{22}$$

where $h(x) = f_1(x + d_1) \cdots f_1(x + d_r)$. Then the following is shown in [3, Lemma 2].

LEMMA 1. *If f, D, r are defined as in Theorem* 4, *then the polynomial h has at least one root in $\overline{\mathbb{F}}_p$ whose multiplicity is odd.*

This lemma ensures the applicability of the following consequence of Weil's theorem [26], [25], [3, Lemma 1].

LEMMA 2. *Suppose that p is a prime, χ is a multiplicative character of $\mathbb{F}_p$ of order $d \geq 2$, and $g \in \mathbb{F}_p[x]$ has positive degree and a factorization $g(x) = b(x - x_1)^{d_1} \cdots (x - x_s)^{d_s}$ (where $x_i \neq x_j$ for $i \neq j$) in $\overline{\mathbb{F}}_p[x]$ with $b \in \mathbb{F}_p^*$ and*

$$\gcd(d, d_1, \ldots, d_s) = 1.$$

Let X, Y be real numbers with $0 < Y \leq p$. Then

$$\Big| \sum_{X < n \leq X+Y} \chi(g(n)) \Big| < 9 \deg(g) p^{1/2} \log p.$$

By using this lemma, it follows from (22) that

$$\begin{aligned} C_r(E_p) &= \max_{H,d_1,\ldots,d_r} \Big| \sum_{n=1}^{H} e_{n+d_1} \cdots e_{n+d_r} \Big| \leq \max_{H,d_1,\ldots,d_r} \Big| \sum_{n=1}^{H} \left(\frac{h(n)}{p} \right) \Big| + Dr \\ &< 9Drp^{1/2} \log p + Dr < 10Drp^{1/2} \log p, \end{aligned}$$

which proves (21).

The "Q-analog" of this argument starts out from

$$\begin{aligned} (23) \qquad & \Big| \sum_{n=1}^{H} e_{na+d_1} \cdots e_{na+d_r} \Big| \\ & \leq \left| \sum_{n=1}^{H} \left(\frac{h(na)}{p} \right) \right| + Dr = \left| \sum_{n=1}^{H} \left(\frac{h_a(n)}{p} \right) \right| + Dr, \end{aligned}$$

where $h(n)$ is defined as in (22) and $h_a(n) = h(na)$. By Lemma 1, h has at least one root $x_0 \in \overline{\mathbb{F}}_p$ of odd multiplicity. Then $a^{-1}x_0$ is a root of the same multiplicity of $h_a(x) = h(xa)$, thus $h_a \in \mathbb{F}_p[x]$ also has a root of odd multiplicity. Then again we may apply Lemma 2 (with $\left(\frac{n}{p}\right)$ and $h_a(n)$ in place of $\chi(n)$ and $g(n)$, respectively), and we deduce from (23) that

$$\begin{aligned} Q_r(E_p) &= \max_{a,H,d_1,\ldots,d_r} \Big| \sum_{n=1}^{H} e_{na+d_1} \cdots e_{na+d_r} \Big| \\ &\leq \max_{a,H,d_1,\ldots,d_r} \left| \sum_{n=1}^{H} \left(\frac{h_a(n)}{p} \right) \right| + Dr \\ &< 9Drp^{1/2} \log p + Dr < 10Drp^{1/2} \log p, \end{aligned}$$

which completes the proof of Theorem 4. ∎

In the cases of the k-dimensional constructions obtained from (10) and (11) in Examples 4 and 5, respectively, the situation is similar: one may estimate $Q_r(E_p)$ by a small modification of the estimation of $C_r(E_p)$ in [4] and [13], respectively, thus we leave the details to the reader.

At the end of Section 3 we mentioned a one-dimensional construction of recursive type with good control over the measures of pseudorandomness. This is the following construction of Niederreiter and Rivat [18].

EXAMPLE 6. Consider the sequence $R_1, R_2, \ldots$ of rational functions over $\mathbb{F}_q$ defined by

$$R_1(X) = X, \quad R_i(X) = R_{i-1}(\alpha X^{-1} + \beta) \quad \text{for } i = 2, 3, \ldots,$$

where $\alpha \in \mathbb{F}_q^*$ and $\beta \in \mathbb{F}_q$. In [18, Lemma 1] it is proved (improving results of [19]) that the sequence of rational functions $R_1, R_2, \ldots$ (actually we have shifted the indices by 1 in the present paper for the sake of consistency) is purely periodic with least period $T \le q+1$, and that there exist distinct elements $\varepsilon_2, \ldots, \varepsilon_T$ of $\mathbb{F}_q$ such that

$$R_i(X) = \frac{(\beta - \varepsilon_i)X + \alpha}{X - \varepsilon_i} \quad \text{for } 2 \le i \le T.$$

For $2 \le i \le T$, we will consider the permutations of $\mathbb{F}_q$ defined by

$$\psi_i(\gamma) = \begin{cases} R_i(\gamma) & \text{if } \gamma \ne \varepsilon_i, \\ \beta - \varepsilon_i & \text{if } \gamma = \varepsilon_i. \end{cases}$$

Now if we also assume that $\beta \in \mathbb{F}_q^*$ and that $\gamma_1 \in \mathbb{F}_q$ satisfies

$$\gamma_1^2 \ne \beta\gamma_1 + \alpha,$$

then by Lemma 2 of [18] we know that the sequence $\gamma_1, \gamma_2, \ldots$ of elements of $\mathbb{F}_q$, defined by

$$\gamma_i = \psi_i(\gamma_1) \quad \text{for } 2 \le i \le T$$

and extended with period T, has least period T and contains at least $T-1$ distinct elements of $\mathbb{F}_q$.

Now we consider the special case where $q = p$ is an odd prime number, and we define a binary sequence

$$E_N = (e_1, \ldots, e_N), \quad 1 \le N \le T, \tag{24}$$

by

$$e_n := \begin{cases} +1 & \text{if } 0 \le \gamma_n \le (p-1)/2, \\ -1 & \text{if } (p+1)/2 \le \gamma_n \le p-1. \end{cases} \tag{25}$$

Niederreiter and Rivat [18] gave upper bounds on $W(E_N)$ and $C_s(E_N)$ under the conditions in Example 6. In particular, they showed that for $1 \le N \le T$

we have

$$C_s(E_N) < 2^s((14s)^{1/2}N^{1/2}p^{1/4} + sp^{1/2} + 8s)\left(\frac{4}{\pi^2}\log p + 1.72\right)^s + \frac{s2^sN}{p}. \tag{26}$$

We will generalize this result to the combined measure $Q_s(E_N)$.

THEOREM 5. *Under the assumptions above, we have, for $s \in \mathbb{N}$ and $1 \le N \le T$,*

$$Q_s(E_N) < 2^s((14s)^{1/2}N^{1/2}p^{1/4} + sp^{1/2} + 8s)\left(\frac{4}{\pi^2}\log p + 1.72\right)^s + \frac{s2^sN}{p}.$$

Again, combining this theorem with Theorem 2 and Theorem 3, respectively, one could estimate the measures of pseudorandomness of the vector sequence $\mathbf{U}(E_N)$ composed from the binary sequence E_N defined by (24) and (25); we leave the details of this to the reader.

Proof of Theorem 5. We extend the sequence $(\gamma_n)_{n=1}^T$ to a doubly-infinite sequence $(\gamma_n)_{n=-\infty}^\infty$ with period T. As in the proof of Theorem 5 of [18], we only need to show Lemma 3 below which holds for arbitrary finite fields $\mathbb{F}_q$. ∎

LEMMA 3. *For a nontrivial additive character φ of $\mathbb{F}_q$, for $s \in \mathbb{N}$, for integers $1 \le d_1 < \cdots < d_s \le T$, and for integers $a \ge 1$ and $t \ge 1$ with $a(t-1) + d_s \le T$, if $\mu_1, \dots, \mu_s \in \mathbb{F}_q$ are not all 0, then for the character sum*

$$S_t := \sum_{n=0}^{t-1} \varphi\Big(\sum_{i=1}^{s} \mu_i \gamma_{an+d_i}\Big)$$

we have

$$|S_t| < (14s)^{1/2}t^{1/2}q^{1/4} + sq^{1/2} + 8s.$$

Proof. By the argument at the beginning of the proof of Theorem 2 of [18], we have, for any integer L with $1 \le L \le T$ (see formula (12) of [18]),

$$L|S_t| \le W(t,L) + L^2/2 \tag{27}$$

where

$$W(t,L) := \Big|\sum_{n=0}^{t-1} \sum_{l \in \mathcal{R}(L)} \varphi\Big(\sum_{i=1}^{s} \mu_i \gamma_{a(n+l)+d_i}\Big)\Big|$$

with $\mathcal{R}(L)$ being the interval of integers l such that $-L/2 < l \le L/2$. Let ψ_1 be the identity map on $\mathbb{F}_q$ and for any $k \in \mathbb{Z}$ let $\psi_k = \psi_r$, where $r \in \{1, \dots, T\}$ is such that $k \equiv r \bmod T$. Then

$$W(t,L) \le U(t,L) + 2E(t,L),$$

where

$$U(t,L) := \sum_{n=0}^{t-1} \Big| \sum_{l \in \mathcal{R}(L)} \varphi\Big(\sum_{i=1}^{s} \mu_i \psi_{d_i+al}(\gamma_{an}) \Big) \Big|$$

and $E(t,L)$ is the number of pairs (n,l) with $0 \le n \le t-1$ and $l \in \mathcal{R}(L)$ such that there exists $i \in \{1,\dots,s\}$ with $\gamma_{an+d_i+al} \ne \psi_{d_i+al}(\gamma_{an})$.

In order to apply the same arguments as in [18], we impose the condition $1 \le L \le t$, which implies in particular that $1 \le a(L-1)+d_s \le T$. Obviously

$$E(t,L) \le \sum_{i=1}^{s} E_i(t,L),$$

where for $1 \le i \le s$ we have

$$\begin{aligned} E_i(t,L) &= \#\{(n,l) \in \{0,\dots,t-1\} \times \mathcal{R}(L) : \gamma_{an+d_i+al} \ne \psi_{d_i+al}(\gamma_{an})\} \\ &\le \#\{(n,l) \in \{0,\dots,t-1\} \times \mathcal{R}(L) : \\ &\qquad\qquad \varepsilon_{an} = \gamma_1 \text{ or } \varepsilon_{an+d_i+al} = \gamma_1 \text{ or } \gamma_{an} = \varepsilon_{d_i+al}\} \\ &\le \#\{(n,l) : \varepsilon_{an} = \gamma_1\} + \#\{(n,l) : \varepsilon_{an+d_i+al} = \gamma_1\} \\ &\quad + \#\{(n,l) : \gamma_{an} = \varepsilon_{d_i+al}\}. \end{aligned}$$

Here $\varepsilon_k = \varepsilon_r$ if $k \equiv r \bmod T$ with $r \in \{2,\dots,T\}$ and ε_1 arbitrary but not in $\mathbb{F}_q$. We observe that the condition $a(t-1)+d_s \le T$ implies that the values of an with n running through $\{0,\dots,t-1\}$ are all distinct modulo T. Since $\varepsilon_1,\dots,\varepsilon_T$ are all distinct, each of the first two sets in the last expression has at most L elements. Concerning the third set in the last expression, since all but at most two of the γ_n with $0 \le n \le t-1$ are distinct and the condition $a(L-1)+d_s \le T$ implies that all the values al with l running through $\mathcal{R}(L)$ are distinct modulo T, for each l except possibly one there is at most one corresponding n, and in the exceptional case there are at most two corresponding values of n. Hence the cardinality of the third set is at most $L+1$. Therefore

$$E(t,L) \le (3L+1)s,$$

so that, as in formula (14) of [18], we get

$$W(t,L) \le U(t,L) + (6L+2)s \le U(t,L) + 8Ls. \tag{28}$$

By the Cauchy–Schwarz inequality we obtain

$$U(t,L)^2 \le t \sum_{n=0}^{t-1} \Big| \sum_{l \in \mathcal{R}(L)} \varphi\Big(\sum_{i=1}^{s} \mu_i \psi_{d_i+al}(\gamma_{an}) \Big) \Big|^2.$$

Since the γ_{an} with $0 \le n \le t-1$ are distinct except possibly two of them

(see again Lemma 2 of [18]), we get

$$U(t,L)^2 \le 2t \sum_{\phi\in\mathbb{F}_q} \Big| \sum_{l\in\mathcal{R}(L)} \varphi\Big(\sum_{i=1}^{s} \mu_i \psi_{d_i+al}(\phi)\Big)\Big|^2$$

$$\le 2t \sum_{l,m\in\mathcal{R}(L)} \Big| \sum_{\phi\in\mathbb{F}_q} \varphi\Big(\sum_{i=1}^{s} \mu_i(\psi_{d_i+al}(\phi) - \psi_{d_i+am}(\phi))\Big)\Big|.$$

The contribution of the terms with $l = m$ is $2Ltq$. Next we consider the pairs $(l,m) \in \mathcal{R}(L)^2$ with $l \ne m$ for which there exist $i,j \in \{1,\dots,s\}$ such that

$$al - am \equiv d_i - d_j \bmod T.$$

As we have noted above, the values al with l running through $\mathcal{R}(L)$ are distinct modulo T, and so there are at most $s(s-1)L$ such pairs (l,m). For each such pair, we use the trivial bound q for the last character sum. The total contribution of these terms to the upper bound on $U(t,L)^2$ is $2s(s-1)Ltq$.

It remains to consider the pairs $(l,m) \in \mathcal{R}(L)^2$ for which

$$al - am \not\equiv d_i - d_j \bmod T \quad \text{for all } 1 \le i,j \le s. \tag{29}$$

In [18, formula (15)] the authors fixed such a pair (l,m) and their proof can be adapted to get

$$\Big| \sum_{\phi\in\mathbb{F}_q} \varphi\Big(\sum_{i=1}^{s} \mu_i(\psi_{d_i+al}(\phi) - \psi_{d_i+am}(\phi))\Big)\Big| \le 4sq^{1/2} + 2s \le 6sq^{1/2}$$

for all pairs $(l,m) \in \mathcal{R}(L)^2$ satisfying (29).

By combining all cases for the pairs $(l,m) \in \mathcal{R}(L)^2$, we obtain

$$U(t,L)^2 \le 2s^2Ltq + 12sL^2tq^{1/2}.$$

If $t \le 14sq^{1/2}$, then the result of the lemma is trivial, and so we can assume that $t > 14sq^{1/2}$. Then $L := \lfloor 2sq^{1/2} \rfloor$ satisfies the condition $1 \le L \le t$. With this choice of L we have $L > sq^{1/2}$, thus $s^2Lq < sL^2q^{1/2}$, and so

$$U(t,L)^2 < 2sL^2tq^{1/2} + 12sL^2tq^{1/2} = 14sL^2tq^{1/2}.$$

Then (28) yields

$$W(t,L) < (14s)^{1/2}Lt^{1/2}q^{1/4} + 8Ls,$$

and an application of (27) completes the proof of Lemma 3. ∎

References

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

[2] J. E. Gentle, *Random Number Generation and Monte Carlo Methods*, 2nd ed., Springer, New York, 2003.
[3] L. Goubin, Ch. Mauduit, and A. Sárközy, *Construction of large families of pseudorandom binary sequences*, J. Number Theory 106 (2004), 56–69.
[4] K. Gyarmati, *On a family of pseudorandom binary sequences*, Period. Math. Hungar. 49 (2004), 45–63.
[5] —, *On a fast version of a pseudorandom generator*, in: General Theory of Information Transfer and Combinatorics, Lecture Notes in Comput. Sci. 4123, Springer, Berlin, 2006, 326–342.
[6] —, *A note to the paper "On a fast version of a pseudorandom generator"*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math., to appear.
[7] R. Lidl and H. Niederreiter, *Finite Fields*, Cambridge Univ. Press, Cambridge, 1997.
[8] Ch. Mauduit, H. Niederreiter, and A. Sárközy, *On pseudorandom* $[0, 1)$ *and binary sequences*, Publ. Math. Debrecen 71 (2007), 305–324.
[9] Ch. Mauduit, J. Rivat, and A. Sárközy, *Construction of pseudorandom binary sequences using additive characters*, Monatsh. Math. 141 (2004), 197–208.
[10] Ch. Mauduit and A. Sárközy, *On finite pseudorandom binary sequences. I. Measure of pseudorandomness, the Legendre symbol*, Acta Arith. 82 (1997), 365–377.
[11] —, —, *On finite pseudorandom sequences of k symbols*, Indag. Math. (N.S.) 13 (2002), 89–101.
[12] —, —, *On the measures of pseudorandomness of binary sequences*, Discrete Math. 271 (2003), 195–207.
[13] —, —, *Construction of pseudorandom binary sequences by using the multiplicative inverse*, Acta Math. Hungar. 108 (2005), 239–252.
[14] H. Niederreiter, *Some new exponential sums with applications to pseudo-random numbers*, in: Topics in Number Theory, P. Turán (ed.), Colloq. Math. Soc. János Bolyai 13, North-Holland, Amsterdam, 1976, 209–232.
[15] —, *The serial test for pseudo-random numbers generated by the linear congruential method*, Numer. Math. 46 (1985), 51–68.
[16] —, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.
[17] —, *New developments in uniform pseudorandom number and vector generation*, in: Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, H. Niederreiter and P. J.-S. Shiue (eds.), Lecture Notes in Statist. 106, Springer, New York, 1995, 87–120.
[18] H. Niederreiter and J. Rivat, *On the correlation of pseudorandom numbers generated by inversive methods*, Monatsh. Math. 153 (2008), 251–264.
[19] H. Niederreiter and I. E. Shparlinski, *On the distribution of pseudorandom numbers and vectors generated by inversive methods*, Appl. Algebra Engrg. Comm. Comput. 10 (2000), 189–202.
[20] —, —, *Recent advances in the theory of nonlinear pseudorandom number generators*, in: Monte Carlo and Quasi-Monte Carlo Methods 2000, K.-T. Fang, F. J. Hickernell, and H. Niederreiter (eds.), Springer, Berlin, 2002, 86–102.
[21] S.-M. Oon, *On some pseudorandom properties of some Dirichlet characters*, Ramanujan J. 15 (2008), 19–30.
[22] J. Rivat and A. Sárközy, *On pseudorandom binary sequences and their applications*, in: General Theory of Information Transfer and Combinatorics, Lecture Notes in Computer Science 4123, Springer, Berlin, 2006, 343–361.
[23] A. Sárközy, *A finite pseudorandom binary sequence*, Studia Sci. Math. Hungar. 38 (2001), 377–384.

[24] A. Sárközy and C. L. Stewart, *On pseudorandomness in families of sequences derived from the Legendre symbol*, Period. Math. Hungar. 54 (2007), 163–173.

[25] W. M. Schmidt, *Equations over Finite Fields*: *An Elementary Approach*, Lecture Notes in Math. 536, Springer, Berlin, 1976; 2nd ed., Kendrick Press, Heber City, UT, 2004.

[26] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Actualités Sci. Ind. 1041, Hermann, Paris, 1948.

Department of Mathematics
and Risk Management Institute
National University of Singapore
2 Science Drive 2
Singapore 117543
Republic of Singapore
E-mail: nied@math.nus.edu.sg

Institut de Mathématiques de Luminy
CNRS-UMR 6206
Université de la Méditerranée
163 avenue de Luminy, Case 907
13288 Marseille Cedex 9, France
E-mail: rivat@iml.univ-mrs.fr

Department of Algebra and Number Theory
Eötvös Loránd University
Pázmány Péter sétány 1/c
H-1117 Budapest, Hungary
E-mail: sarkozy@cs.elte.hu

Received on 22.8.2007
and in revised form on 16.1.2008 (5501)