ACTA ARITHMETICA
125.1 (2006)

On pseudorandom binary lattices

by

P. Hubert (Marseille), C. Mauduit (Marseille) and
A. Sárközy (Budapest)

1. Introduction. Recently in a series of papers a new constructive approach has been developed to study pseudorandomness of binary sequences

$$E_N = \{e_1, \dots, e_N\} \in \{-1, 1\}^N. \tag{1.1}$$

In particular, in [5] Mauduit and Sárközy first introduced the following measures of pseudorandomness: the *well-distribution measure* of E_N is defined by

$$W(E_N) = \max_{a,b,t} \Big| \sum_{j=0}^{t-1} e_{a+jb} \Big| \tag{1.2}$$

where the maximum is taken over all $a, b, t \in \mathbb{N}$ with $1 \le a \le a+(t-1)b \le N$, and the *correlation measure* of order k of E_N is defined as

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

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

$$Q_k(E_N) = \max_{a,b,t,\boldsymbol{D}} \Big| \sum_{j=0}^{t} e_{a+jb+d_1} \cdots e_{a+jb+d_k} \Big| \tag{1.3}$$

where the maximum is over all a, b, t and $\boldsymbol{D} = (d_1, \dots, d_k)$ such that all the subscripts $a + jb + d_l$ belong to $\{1, \dots, N\}$. Then the sequence E_N is considered to be a "good" pseudorandom sequence if both $W(E_N)$ and

2000 *Mathematics Subject Classification*: Primary 11K45.

Research of the third author partially supported by Hungarian National Foundation for Scientific Research, Grants No. T 043623 and T 049693, and by French–Hungarian EGIDE–OMKFHÁ exchange program Balaton F-2/03. This paper was written while he was visiting the Institut de Mathématiques de Luminy, Marseille.

$C_k(E_N)$ (at least for "small" k) are "small" in terms of N (in particular, both are $o(N)$ as $N \to \infty$). Indeed, later Cassaigne, Mauduit and Sárközy [2] showed that this terminology is justified since for almost all $E_N \in \{-1, 1\}^N$, both $W(E_N)$ and $C_k(E_N)$ are less than $N^{1/2}(\log N)^c$. (See also [1].) It was also shown in [5] that the Legendre symbol forms a "good" pseudorandom sequence. Later several further sequences were tested for pseudorandomness, and further constructions were given for sequences with good pseudorandom properties. In some other papers the measures of pseudorandomness were studied.

The work above was motivated by two facts: first, pseudorandom binary sequences have many applications (e.g., in stream chiper type cryptosystems) and secondly, the theory of pseudorandomness can be utilized in number theory to study certain sequences and phenomena. Several-dimensional analogs of pseudorandom binary sequences (which we will call binary lattices) also have many applications in cryptography (e.g., in encrypting "bitmaps"), steganography and watermarking (see, e.g., [4], [6]–[9] and the numerous further references therein for related work), and one may expect that a theory of pseudorandomness in several dimensions also could be utilized in number theory. Therefore in this paper we will extend the theory of pseudorandomness described above from one dimension to several dimensions. (In applications the dimension is usually 2, sometimes 3; however, since the general case of n dimensions can be handled without additional difficulties, we will consider it here.) This is not just a routine generalization since there are two difficulties. First, in one dimension in the best constructions one needs estimates for character sums with general term $\chi(f(x))$ where χ is a multiplicative character modulo p and $f(x) \in \mathbb{F}_p[x]$, and sums of this type can be estimated by using Weil's theorem [10]. On the other hand, in two dimensions the analogous constructions would lead to double character sums with general term $\chi(f(x, y))$ so that one would need Kac's theorem which, because of strong nonsingularity assumptions, is not flexible enough for this purpose. (We will get around this difficulty by considering finite fields $\mathbb{F}_{p^n}$ as vector spaces over $\mathbb{F}_p$, and using a principle due to Davenport and Lewis [3] and recently generalized and expressed in a convenient form by Winterhof [11].) Secondly, for $n > 1$ the n-dimensional lattices have no natural ordering (like the ordering of $\mathbb{N}$ in one dimension); this will lead to some difficulties in studying the truly random case.

First in Section 2 we formulate the problem and introduce the measures of pseudorandomness. In Section 3 we study these measures in the "truly random" case. Finally, in Section 4 we present a construction which is "good" in terms of these new measures.

2. Formulating the problem in several dimensions and introducing new measures. Let I_N^n denote the set of n-dimensional vectors all of whose coordinates are in $\{0, 1, \dots, N-1\}$:

$$I_N^n = \{\boldsymbol{x} = (x_1, \dots, x_n) : x_1, \dots, x_n \in \{0, 1, \dots, N-1\}\}.$$

This set forms a (truncated) n-dimensional lattice so we may call it the *n-dimensional N-lattice* or briefly (if n is fixed) the *N-lattice*. Then the binary sequences of the form (1.1) can be considered as functions of type

$$e_x = \eta(x) : I_N^1 \to \{-1, 1\}. \tag{2.1}$$

Thus clearly the natural n-dimensional extension is to study the pseudorandomness of functions of type

$$\eta = \eta(\boldsymbol{x}) : I_N^n \to \{-1, 1\}. \tag{2.2}$$

Such a function can be visualized as the lattice points of the N-lattice decorated by the two symbols $+$ or $-$, so we may call them *binary N-lattices* or briefly *binary lattices*.

In order to introduce the measures of pseudorandomness of binary lattices one might like to adopt the one-dimensional definitions. However, in one dimension different versions of the well-distribution measure and correlation measure are quite frequently used, so we focused on these two measures, although the study of the combined measure would provide more information on the sequence studied. In several dimensions the well-distribution measure and correlation measure have no standard frequently used analogs, so here we will introduce and study only a generalization of the combined measure Q_k.

If $\eta = \eta(\boldsymbol{x})$ is an n-dimensional binary N-lattice of the form (2.2), $k \in \mathbb{N}$, and $\boldsymbol{u}_i$ $(i = 1, \dots, n)$ denotes the n-dimensional unit vector whose ith coordinate is 1 and the other coordinates are 0, then write

$$Q_k(\eta) = \max_{\boldsymbol{B}, \boldsymbol{d}_1, \dots, \boldsymbol{d}_k, \boldsymbol{T}} \Big| \sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_1) \cdots \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_k) \Big| \tag{2.3}$$

where the maximum is taken over all n-dimensional vectors $\boldsymbol{B} = (b_1, \dots, b_n)$, $\boldsymbol{d}_1, \dots, \boldsymbol{d}_k$, $\boldsymbol{T} = (t_1, \dots, t_n)$ whose coordinates are non-negative integers, $b_1, \dots, b_n$ are non-zero, $\boldsymbol{d}_1, \dots, \boldsymbol{d}_k$ are distinct, and all the points $j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_i$ occurring in the multiple sum belong to the n-dimensional N-lattice I_N^n. We will call $Q_k(\eta)$ the *pseudorandom* (briefly PR) *measure* of order k of η.

Note that in the one-dimensional special case $Q_1(\eta)$ is the same as the well-distribution measure (1.2), and for every $k \in \mathbb{N}$, $Q_k(\eta)$ is the combined

measure (1.3). Then a binary N-lattice η is considered to be a "good" pseudorandom binary lattice if the PR measure of order k of η is "small" in terms of N (in particular, $Q_k(\eta) = o(N^n)$ as $N \to \infty$) for small k. This terminology will be justified by Theorem 1 in the next section.

3. The pseudorandom measures for truly random binary lattices. In this section we will estimate $Q_k(\eta)$ for a truly random binary lattice. More precisely, assume that $N \in \mathbb{N}$, $n \in \mathbb{N}$, write $Z = |I_N^n| = N^n$, denote the elements of I_N^n by $\boldsymbol{x}_1, \dots, \boldsymbol{x}_Z$, and then choose each of the binary lattices η of the form (2.2) with the same probability 2^{-Z}, i.e., define η so that $\eta(\boldsymbol{x}_1), \dots, \eta(\boldsymbol{x}_Z)$ are independent random variables with

$$P(\eta(\boldsymbol{x}_i) = 1) = P(\eta(\boldsymbol{x}_i) = -1) = 1/2. \tag{3.1}$$

We will prove:

THEOREM 1. *If $k \in \mathbb{N}$ and $\varepsilon > 0$, then there are numbers $N_0 = N_0(k, \varepsilon)$ and $\delta = \delta(k, \varepsilon) > 0$ such that for $N > N_0$ we have*

$$P(Q_k(\eta) > \delta N^{n/2}) > 1 - \varepsilon \tag{3.2}$$

and

$$P(Q_k(\eta) > (KN^n \log N^n)^{1/2}) < \varepsilon, \tag{3.3}$$

where $K = 81k$.

Proof. If $k = 1$, then (3.2) follows from

$$P(Q_k(\eta) > \delta N^{n/2}) > P\Big(\Big|\sum_{j_1=0}^{N-1} \cdots \sum_{j_n=0}^{N-1} \eta(j_1\boldsymbol{u}_1 + \cdots + j_n\boldsymbol{u}_n)\Big| > \delta N^{n/2}\Big)$$
$$= P\Big(\Big|\sum_{i=1}^{N^n} \eta(\boldsymbol{x}_i)\Big| > \delta N^{n/2}\Big),$$

(3.1) and the central limit theorem.

If $k \geq 2$, then consider the n-fold sum in (2.3) with $t_1 = [N/2k] - 1$, $t_2 = \cdots = t_n = N - 1$, $b_1 = k$, $b_2 = \cdots = b_n = 1$, $\boldsymbol{d}_i = (i-1)\boldsymbol{u}_1$ for $i = 1, \dots, k-1$ and $\boldsymbol{d}_k = [N/2]\boldsymbol{u}_1$. Then clearly, for $0 \leq j_1 \leq t_1, \dots, 0 \leq j_n \leq t_n$, $1 \leq i \leq k$ we have

$$j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_i \in I_N^n$$

and thus, indeed, the sum

$$S = \sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_1)$$
$$\cdots \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_k)$$

is as considered in (2.3). Moreover, it is easy to see that for distinct $(n+1)$-tuples $(j_1,\dots,j_n,i)$ we obtain different vectors $j_1b_1\boldsymbol{u}_1+\dots+j_nb_n\boldsymbol{u}_n+\boldsymbol{d}_i$, thus all the factors $\eta(\dots)$ in this sum are independent random variables of type (3.1). Now fixing the values of the first $k-1$ random variables η in each term of S and denoting the vectors $j_1b_1\boldsymbol{u}_1+\dots+j_nb_n\boldsymbol{u}_n+\boldsymbol{d}_k$ in S by $\boldsymbol{v}_1,\dots,\boldsymbol{v}_{(t_1+1)\cdots(t_n+1)}$, we get a sum of the form

$$S'=\sum_{1\le l\le(t_1+1)\cdots(t_n+1)} e_l\eta(\boldsymbol{v}_l)$$

where $e_l\in\{-1,1\}$ for each l. Writing $\xi_l=e_l\eta(\boldsymbol{v}_l)$, this becomes

$$S'=\sum_{1\le l\le[N/2k]N^{n-1}}\xi_l$$

where the ξ_l's are independent random variables with distribution

$$P(\xi_l=1)=P(\xi_l=-1)=1/2.$$

By the central limit theorem, there are $N_0=N_0(k,\varepsilon)$ and $\delta=\delta(k,\varepsilon)>0$ such that

$$P(|S'|>\delta N^{n/2})=P\Big(\Big|\sum_{1\le l\le[N/2k]N^{n-1}}\xi_l\Big|>\delta N^{1/2}\Big)>1-\varepsilon.$$

This is so under the condition that certain random variables $\eta(\dots)$ in S are fixed as described above, and this holds uniformly for any fixed $\eta(\dots)$ values, which implies (3.2).

In order to prove (3.3), we will need the following lemma:

LEMMA 1. *Let $r,k,M,Z\in\mathbb{N}$. Assume that $E_Z=\{e_1,\dots,e_Z\}$ is a set of independent random variables of type* (3.1), *i.e.*,

$$P(e_i=1)=P(e_i=-1)=1/2. \tag{3.4}$$

Assume also that $y_1,\dots,y_M$ are random variables of the form

$$y_l=e_{i(l,1)}e_{i(l,2)}\cdots e_{i(l,k)}\quad \text{for } l=1,\dots,M \tag{3.5}$$

where

$$i(l,j)\in\{1,\dots,Z\}\quad \text{for } 1\le l\le M,\ 1\le j\le k, \tag{3.6}$$

$$i(l,j_1)\ne i(l,j_2)\quad \text{for } 1\le l\le M,\ 1\le j_1<j_2\le k, \tag{3.7}$$

and

$$i(l,1)\ne i(l+j,m)\quad \text{for } 1\le l\le M,\ 1\le j\le M-l,\ 1\le m\le k. \tag{3.8}$$

Then

$$\mathbb{E}\Big(\Big(\sum_{n=1}^{M}y_n\Big)^{2r}\Big)\le 2^{1-M}\sum_{h=0}^{[M/2]}\binom{M}{h}(M-2h)^{2r}. \tag{3.9}$$

($\mathbb{E}(\xi)$ *denotes the expectation of the random variable* ξ.)

Proof. Both the proof of the lemma and the completion of the proof of (3.3) are similar to the proof of the upper bound in Theorem 2 of [2], so we omit some details. However, there is a significant difference: while in the one-dimensional case we may use the natural ordering of the positive integers, in several dimensions there is no natural ordering, so we have to use an artificial one, which leads to certain complications (in particular, this explains the role of condition (3.8) in the lemma).

By the multinomial theorem we have

$$\mathbb{E}\Big(\Big(\sum_{n=1}^{M} y_n\Big)^{2r}\Big) = \mathbb{E}\Big(\sum_{t=1}^{2r} \sum_{1\le i_1<\dots<i_t\le M} \sum_{\substack{j_1+\dots+j_t=2r\\ 1\le j_1,\dots,j_t}} \frac{(2r)!}{j_1!\cdots j_t!}\, y_{i_1}^{j_1}\cdots y_{i_t}^{j_t}\Big)$$
$$= \sum_{t=1}^{2r} \sum_{1\le i_1<\dots<i_t\le M} \sum_{\substack{j_1+\dots+j_t=2r\\ 1\le j_1,\dots,j_t}} \frac{(2r)!}{j_1!\cdots j_t!}\, \mathbb{E}(y_{i_1}^{j_1}\cdots y_{i_t}^{j_t}).$$

Observe that for each i we have $y_i \in \{-1,1\}$, so y_i^j depends only on the parity of j: $y_i^j = 1$ if j is even and $y_i^j = y_i$ if j is odd. Let $\sum_1$ denote the contribution of those terms for which at least one of $j_1,\dots,j_t$ is odd, and $\sum_2$ the remaining ones, so that

$$\mathbb{E}\Big(\Big(\sum_{n=1}^{M} y_n\Big)^{2r}\Big) = \sum\nolimits_1 + \sum\nolimits_2 . \tag{3.10}$$

In $\sum_1$ in each term the last factor can be replaced by a factor of the form

$$\mathbb{E}(y_{s_1}\cdots y_{s_u}) \quad \text{with } s_1 < \dots < s_u.$$

By (3.5) here we may replace each y_{s_h} by $e_{i(s_h,1)}\cdots e_{i(s_h,k)}$. Then by conditions (3.7) and (3.8), $e_{i(s_1,1)}$ occurs only once amongst the factors $e_{i(s_h,j)}$ with $1 \le h \le u$, $1 \le j \le k$. Thus $y_{s_1}\cdots y_{s_n}$ can be rewritten as

$$y_{s_1}\cdots y_{s_u} = e_{i(s_1,1)} e_{v_1}\cdots e_{v_p} \quad \text{with } i(s_1,1) \ne v_j \text{ for } 1 \le j \le p.$$

Since $e_1,\dots,e_Z$ are independent random variables with expectation 0 (by (3.4)), we have

$$\mathbb{E}(y_{s_1}\cdots y_{s_u}) = \mathbb{E}(e_{i(s_1,1)})\mathbb{E}(e_{v_1})\cdots\mathbb{E}(e_{v_p}) = 0.$$

It follows that

$$\sum\nolimits_1 = 0. \tag{3.11}$$

In $\sum_2$ we may replace each j_i by $2q_i$:

$$\begin{aligned}\sum\nolimits_2 &= \sum_{t=1}^{2r} \sum_{1\le i_1<\cdots<i_t\le M} \sum_{q_1+\cdots+q_t=r} \frac{(2r)!}{(2q_1)!\cdots(2q_t)!}\,\mathbb{E}(y_{i_1}^{2q_1}\cdots y_{i_t}^{2q_t}) \\ &= \sum_{t=1}^{2r} \sum_{1\le i_1<\cdots<i_t\le M} \sum_{q_1+\cdots+q_t=r} \frac{(2r)!}{(2q_1)!\cdots(2q_t)!}\,\mathbb{E}(1) \\ &= \sum_{t=1}^{2r} \sum_{1\le i_1<\cdots<i_t\le M} \sum_{q_1+\cdots+q_t=r} \frac{(2r)!}{(2q_1)!\cdots(2q_t)!}.\end{aligned}$$

This triple sum was computed in [2, p. 104]:

$$\sum\nolimits_2 = 2^{1-M} \sum_{h=0}^{[M/2]} \binom{M}{h} (M-2h)^{2r}. \tag{3.12}$$

Now (3.9) follows from (3.10)–(3.12), completing the proof of Lemma 1. ∎

We now complete the proof of (3.3) by using the moment method. Write $\boldsymbol{D} = (\boldsymbol{d}_1, \dots, \boldsymbol{d}_k)$, $\boldsymbol{B} = (b_1, \dots, b_n)$, $\boldsymbol{T} = (t_1, \dots, t_n)$,

$$\begin{aligned}V(\eta, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{T}) = \sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} &\eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_1) \\ &\cdots \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_k)\end{aligned} \tag{3.13}$$

and

$$\begin{aligned}S(r) &= \mathbb{E}\Big(\sum_{\boldsymbol{B}}\sum_{\boldsymbol{D}}\sum_{\boldsymbol{T}} (V(\eta, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{T}))^{2r}\Big) \\ &= \sum_{\boldsymbol{B}}\sum_{\boldsymbol{D}}\sum_{\boldsymbol{T}} \mathbb{E}((V(\eta, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{T}))^{2r})\end{aligned}$$

where $r = r(k, Z) \in \mathbb{N}$ will be fixed later and the triple sum is taken over all $\boldsymbol{B}, \boldsymbol{D}, \boldsymbol{T}$ as in (2.3).

For a fixed sum $V(\eta, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{T})$ denote the number of its terms by M, i.e., let

$$M = \prod_{i=1}^{n} (t_i + 1),$$

and split $S(r)$ in two parts: let $S_1(r)$ denote the contribution of the terms with $M \le Z^{1/4}$ and let $S_2(r)$ be the contribution of the terms with

$$Z^{1/4} < M \le Z \tag{3.14}$$

so that

$$S(r) = S_1(r) + S_2(r). \tag{3.15}$$

First we estimate $S_1(r)$. Clearly,

$$|V(\eta, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{T})| \le \sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} 1 = M,$$

whence

$$S_1(r) \le \sum_{\boldsymbol{B}} \sum_{\boldsymbol{D}} \sum_{\boldsymbol{T}} M^{2r} = M^{2r} \sum_{\boldsymbol{B}} \sum_{\boldsymbol{D}} \sum_{\boldsymbol{T}} 1.$$

Here $\boldsymbol{B} = (b_1, \dots, b_n)$ can be chosen in at most $N^n = Z$ ways, $\boldsymbol{D} = (\boldsymbol{d}_1, \dots, \boldsymbol{d}_k)$ in $|I_N^n|^k = Z^k$ ways and $\boldsymbol{T} = (t_1, \dots, t_n)$ in $N^n = Z$ ways so that, by the definition of $S_1(r)$,

$$S_1(r) \le M^{2r} \cdot Z \cdot Z^k \cdot Z \le Z^{r/2+k+2}. \tag{3.16}$$

In order to estimate $S_2(r)$ we will use Lemma 1. $S_2(r)$ is a triple sum (over $\boldsymbol{B}, \boldsymbol{D}, \boldsymbol{T}$) whose general term is

$$\begin{aligned} &\mathbb{E}((V(\eta, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{T}))^{2r}) \\ &\quad = \mathbb{E}\Big(\Big(\sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_1) \\ &\qquad \cdots \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_k) \Big)^{2r} \Big) \end{aligned} \tag{3.17}$$

with $M = \prod_{i=1}^n (t_i + 1)$ satisfying (3.14). This expression is of the type considered in (3.9) of Lemma 1, but to ensure that (3.8) in the lemma holds, we have to change the order of terms in this n-fold sum. We will use an ordering of I_N^n which is sometimes called the *graduated lexicographic ordering*, defined in the following way: if $(a_1, \dots, a_n), (b_1, \dots, b_n) \in I_N^n$ then we say that $(a_1, \dots, a_n) < (b_1, \dots, b_n)$ if and only if either $a_1 + \cdots + a_n < b_1 + \cdots + b_n$ or $a_1 + \cdots + a_n = b_1 + \cdots + b_n$ and $(a_1, \dots, a_n)$ is less than $(b_1, \dots, b_n)$ in the lexicographic order. This ordering has the following fundamental property:

$$\begin{aligned} &\text{if} \quad (a_1, \dots, a_n) < (b_1, \dots, b_n), \\ &\text{then} \quad (a_1 + c_1, \dots, a_n + c_n) < (b_1 + c_1, \dots, b_n + c_n) \\ &\qquad\qquad \text{for all } (c_1, \dots, c_n) \in I_N^n. \end{aligned} \tag{3.18}$$

Now we reorder the vectors $j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n$ (with $0 \le j_1 \le t_1, \dots, 0 \le j_n \le t_n$) to form an increasing sequence in the graduated lexicographic ordering, say

$$\boldsymbol{v}_1 < \cdots < \boldsymbol{v}_M. \tag{3.19}$$

Moreover, the vectors $\boldsymbol{d}_1, \dots, \boldsymbol{d}_k$ play a symmetric role, so we may assume without loss of generality that

$$\boldsymbol{d}_1 < \cdots < \boldsymbol{d}_k. \tag{3.20}$$

Then (3.17) can be rewritten as

$$\mathbb{E}((V(\eta,\boldsymbol{B},\boldsymbol{D},\boldsymbol{T}))^{2r}) = \mathbb{E}\Big(\Big(\sum_{i=1}^{M}\eta(\boldsymbol{v}_i+\boldsymbol{d}_1)\cdots\eta(\boldsymbol{v}_i+\boldsymbol{d}_k)\Big)^{2r}\Big).$$

Now we use Lemma 1 with $E_Z = \{e_1,\dots,e_Z\} = \{\eta(\boldsymbol{x}_1),\dots,\eta(\boldsymbol{x}_Z)\}$, $y_l = \eta(\boldsymbol{v}_l+\boldsymbol{d}_1)\cdots\eta(\boldsymbol{v}_l+\boldsymbol{d}_k)$ (for $l=1,\dots,M$), $e_{i(l,j)} = \eta(\boldsymbol{v}_l+\boldsymbol{d}_j)$ (for $1\le l\le M$, $1\le j\le k$). Then (3.4)–(3.6) in Lemma 1 hold trivially, and (3.7) and (3.8) hold by (3.18)–(3.20), so the lemma can be applied. We obtain

$$\mathbb{E}((V(\eta,\boldsymbol{B},\boldsymbol{D},\boldsymbol{T}))^{2r}) \le 2^{1-M}\sum_{h=0}^{[M/2]}\binom{M}{h}(M-2h)^{2r}. \tag{3.21}$$

Now we set

$$r = [2k\log Z].$$

Then as in [2, pp. 104–105], it follows from (3.21) that

$$\mathbb{E}((V(\eta,\boldsymbol{B},\boldsymbol{D},\boldsymbol{T}))^{2r}) < 4M(4rM)^r \quad \text{for } Z^{1/4} < M \le Z,$$

whence

$$\begin{aligned} S_2(r) &< \sum_{\boldsymbol{B}}\sum_{\boldsymbol{D}}\sum_{\boldsymbol{T}} 4M(4rM)^r \\ &\le \sum_{\boldsymbol{B}}\sum_{\boldsymbol{D}}\sum_{\boldsymbol{T}} 4Z(4rZ)^r = 4Z(4rZ)^r\sum_{\boldsymbol{B}}\sum_{\boldsymbol{D}}\sum_{\boldsymbol{T}} 1. \end{aligned} \tag{3.22}$$

Here $\boldsymbol{B} = (b_1,\dots,b_n)$ can be chosen in at most $N^n = Z$ ways, $\boldsymbol{D} = (\boldsymbol{d}_1,\dots,\boldsymbol{d}_k)$ in $(N^n)^k = Z^k$ ways, $\boldsymbol{T} = (t_1,\dots,t_n)$ in $N^n = Z$ ways, thus we see from (3.22) that

$$S_2(r) < 4Z^{k+3}(4rZ)^r. \tag{3.23}$$

It follows from (3.15), (3.16) and (3.23) that

$$S(r) < Z^{r/2+k+2} + 4Z^{k+3}(4rZ)^r < 5Z^{k+3}(4rZ)^r. \tag{3.24}$$

On the other hand, writing $X = 9(kZ\log Z)^{1/2} = (81kN^n\log N^n)^{1/2}$, clearly we have

$$\begin{aligned} S(r) &= \mathbb{E}\Big(\sum_{\boldsymbol{B}}\sum_{\boldsymbol{D}}\sum_{\boldsymbol{T}}(V(\eta,\boldsymbol{B},\boldsymbol{D},\boldsymbol{T}))^{2r}\Big) \\ &\ge \mathbb{E}((\max_{\boldsymbol{B},\boldsymbol{D},\boldsymbol{T}}|V(\eta,\boldsymbol{B},\boldsymbol{D},\boldsymbol{T})|)^{2r}) = \mathbb{E}((Q_k(\eta))^{2r}) \\ &\ge P(Q_k(\eta) > X)X^{2r}. \end{aligned} \tag{3.25}$$

It follows from (3.24) and (3.25) that

$$\begin{aligned} P(Q_k(\eta) > X) &< 5Z^{k+3}(4rZX^{-2})^r \\ &= 5Z^{k+3}(4[2k\log Z]ZX^{-2})^r \\ &\le 5Z^{k+3}\left(\frac{8}{81}\right)^r = 5\exp\left((k+3)\log Z - [2k\log Z]\log\frac{81}{8}\right). \end{aligned}$$

If N and thus also Z is large enough in terms of ε, then this upper bound is less than ε, and this completes the proof of Theorem 1. ∎

4. A construction. In this section we present a construction where good upper bounds can be given for the pseudorandom measures introduced in Section 2. We use the notation $e(\alpha) = e^{2\pi i\alpha}$. The letter p will denote an odd prime, $n \in \mathbb{N}$, $q = p^n$, and the quadratic character of $\mathbb{F}_q$ will be denoted by γ. Let $v_1, \dots, v_n$ be a basis of $\mathbb{F}_q$ as a vector space over $\mathbb{F}_p$. Then define

$$\eta : I_p^n \to \{-1, 1\}$$

by

$$\begin{aligned} \eta(\boldsymbol{x}) &= \eta((x_1, \dots, x_n)) \\ &= \begin{cases} \gamma(x_1v_1 + \cdots + x_nv_n) & \text{for } (x_1, \dots, x_n) \ne (0, \dots, 0), \\ 1 & \text{for } (x_1, \dots, x_n) = (0, \dots, 0), \end{cases} \end{aligned} \tag{4.1}$$

for any $x_1, \dots, x_n \in \mathbb{F}_p$.

THEOREM 2. *If p is a prime, $n, k \in \mathbb{N}$ and the n-dimensional binary p-lattice η is defined by* (4.1), *then*

$$Q_k(\eta) < kq^{1/2}(1 + \log p)^n. \tag{4.2}$$

(Note that by Theorem 1, for fixed k and n this upper bound is greater than the value of $Q_k(\eta)$ for a truly random η by at most a logarithm power of p.)

Proof of Theorem 2. We will need the following result of Winterhof:

LEMMA 2. *If $P, n, q, v_1, \dots, v_n$ are defined as above, χ is a multiplicative character of $\mathbb{F}_q$ of order $d > 1$, $f \in \mathbb{F}_q[x]$ is a non-constant polynomial which is not a dth power and which has m distinct zeros in its splitting field over $\mathbb{F}_q$, and $k_1, \dots, k_n$ are positive integers with $k_1 \le p, \dots, k_n \le p$, then, writing $B = \{\sum_{i=1}^n j_iv_i : 0 \le j_i < k_i\}$, we have*

$$\Big|\sum_{Z\in B} \chi(f(z))\Big| < mq^{1/2}(1 + \log p)^n.$$

Proof. This is a part of Theorem 2 in [11] (where its proof was based on A. Weil's theorem [10]). ∎

Now consider a multiple sum of the type occurring in (2.3) with $\boldsymbol{d}_i = (d_1^{(i)}, \dots, d_n^{(i)})$ (for $i = 1, \dots, k$):

$$\begin{aligned} S &= \sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_1) \\ &\qquad \cdots \eta(j_1 b_1 \boldsymbol{u}_1 + \cdots + j_n b_n \boldsymbol{u}_n + \boldsymbol{d}_k) \\ &= \sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} \eta((j_1 b_1 + d_1^{(1)}, \dots, j_n b_n + d_n^{(1)})) \\ &\qquad \cdots \eta((j_1 b_1 + d_1^{(k)}, \dots, j_n b_n + d_n^{(k)})), \end{aligned}$$

whence, by (4.1) and the multiplicativity of γ,

$$\begin{aligned} S &= \sum_{j_1=0}^{t_1} \cdots \sum_{j_n=0}^{t_n} \gamma((j_1(b_1 v_1) + \cdots + j_n(b_n v_n)) + (d_1^{(1)} v_1 + \cdots + d_n^{(1)} v_n)) \\ &\qquad \cdots \gamma((j_1(b_1 v_1) + \cdots + j_n(b_n v_n)) + (d_1^{(k)} v_1 + \cdots + d_n^{(k)} v_n)) \\ &= \sum_{Z \in B'} \gamma((z + z_1) \cdots (z + z_k)) = \sum_{Z \in B'} \gamma(f(z)) \end{aligned} \tag{4.3}$$

with

$$B' = \Big\{ \sum_{i=1}^{n} j_i (b_i v_i) : 0 \le j_i < t_{i+1} \Big\}, \tag{4.4}$$

$$z_i = d_1^{(i)} v_1 + \cdots + d_n^{(i)} v_n \quad \text{for } i = 1, \dots, k \tag{4.5}$$

and

$$f(z) = (z + z_1) \cdots (z + z_k). \tag{4.6}$$

Note that since $v_1, \dots, v_n$ are linearly independent over $\mathbb{F}_p$ and $b_1, \dots, b_n$ are non-zero, $b_1 v_1, \dots, b_n v_n$ are also linearly independent over $\mathbb{F}_p$, so that the box B' in (4.4) is of the same type as the box B in Lemma 2. Since the vectors $\boldsymbol{d}_1, \dots, \boldsymbol{d}_k$ are distinct, so also are the numbers $z_1, \dots, z_k$ in (4.5). It follows that the polynomial $f(z)$ in (4.6) has k distinct zeros so that it is certainly not a square (the order of the character γ is $d = 2$) and hence Lemma 2 can be applied to estimate the sum S in (4.3), yielding

$$|S| < k q^{1/2} (1 + \log p)^n,$$

whence (4.2) follows. ∎

References

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

[2] J. Cassaigne, C. Mauduit and A. Sárközy, *On finite pseudorandom binary sequences VII: The measures of pseudorandomness*, Acta Arith. 103 (2002), 97–118.
[3] H. Davenport and D. J. Lewis, *Character sums and primitive roots in finite fields*, Rend. Circ. Mat. Palermo (2) 12 (1963), 129–136.
[4] S. J. Li, C. Q. Li, G. R. Chen and X. Q. Mou, Cryptology ePrint Archive, Report 2004/376, available online at http://eprint.iacr.org/2004/376.
[5] C. Mauduit and A. Sárközy, *On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol*, Acta Arith. 82 (1997), 365–377.
[6] F. Petitcolas et C. Fontaine, *Nouveaux outils par l'évaluation des algorithmes de tatouage*, in: Tatouage de documents audiovisuels numériques, Traité IC2 série Traitement du signal et de l'image, F. Davaine and S. Pateux (eds.), Hermes-Lavaiser, 2004; Ch. 6, 195–214.
[7] *Security, Steganography and Watermarking of Multimedia Contents VI*, E. J. Delp III and P. W. Wong (eds.), Proc. SPIE 5306, SPIE, 2004.
[8] *Security, Steganography and Watermarking of Multimedia Contents VII*, E. J. Delp III and P. W. Wong (eds.), Proc. SPIE 5681, SPIE, 2005.
[9] *Security, Steganography and Watermarking of Multimedia Contents VIII*, E. J. Delp III and P. W. Wong (eds.), Proc. SPIE 6072, SPIE, 2006.
[10] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Act. Sci. Ind. 1041, Hermann, Paris, 1948.
[11] A. Winterhof, *Some estimates for character sums and applications*, Des. Codes Cryptogr. 22 (2001), 123–131.

Laboratoire d'Analyse, Topologie et Probabilités
Faculté des Sciences de Saint Jérôme
Avenue Escadrille Normandie-Niemen
F-13397 Marseille Cedex 20, France
E-mail: hubert@cmi.univ-mrs.fr

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

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

Received on 19.12.2005
and in revised form on 19.5.2006 (5108)