Publl. Math. Debrecen
79/3-4 (2011), 445–460
DOI: 10.5486/PMD.2011.5066

Measures of pseudorandomness of families of binary lattices, I (Definitions, a construction using quadratic characters)

By KATALIN GYARMATI (Budapest), CHRISTIAN MAUDUIT (Marseille)
and ANDRÁS SÁRKÖZY (Budapest)

Abstract. In the last 15 years a new constructive theory of pseudorandomness of binary sequences has been developed. Later this theory was extended to n dimensions, i.e., to the study of pseudorandomness of binary lattices. In the applications it is not enough to consider single binary sequences, one also needs information on the structure of large families of binary sequences with strong pseudorandom properties. Thus the related notions of family complexity, collision and avalanche effect have been introduced. In this paper our goal is to extend these definitions to binary lattices, and we will present constructions of large families of binary lattices with strong pseudorandom properties such that these families also possess a nice structure.

1. Introduction

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

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

Mathematics Subject Classification: Primary 11K45.
Key words and phrases: pseudorandom, binary lattice, family complexity, collision and avalanche effect.
Research partially supported by Hungarian National Foundation for Scientific Research, Grants No. K67676, K72731 and PD72264, French-Hungarian exchange program "Balaton" FR-33/2009, and the Agence Nationale de la Recherche, grant N^o ANR-10-BLAN 0103.

In particular in [16] Mauduit and Sárközy first introduced the following measures of pseudorandomness: the *well-distribution measure* of E_N is defined by

$$W(E_N) = \max_{a,b,t} \left| \sum_{j=0}^{t-1} e_{a+jb} \right| \tag{1.1}$$

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

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

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

$$Q_k(E_N) = \max_{a,b,t,\mathbf{D}} \left(\sum_{j=0}^{t} e_{a+jb+d_1} \dots e_{a+jb+d_k} \right) \tag{1.2}$$

where the maximum is taken over all a, b, t and $\mathbf{D} = (d_1, \dots, d_k)$ such that all the subscripts $a + jb + d_\ell$ belong to $\{1, 2, \dots, N\}$. (Note that $Q_1(E_N) = W(E_N)$ and clearly $C_k(E_N) \leq Q_k(E_N)$.) Then the sequence E_N is considered to be a "good" pseudorandom sequence if both $W(E_N)$ and $C_k(E_N)$ (at least for "small" k) are "small" in terms of N (in particular, both are $o(N)$ as $N \longrightarrow \infty$). Indeed, later Cassaigne, Mauduit and Sárközy [4] 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 [2] and [15].) Since that many papers have been written on the pseudorandomness of special binary sequences and on the measures of pseudorandomness; a list of these papers is presented in [9].

In [13] Hubert, Mauduit and Sárközy extended this theory of pseudorandomness to n dimensions. They introduced the following definitions:

Denote by I_N^n the set of n-dimensional vectors whose coordinates are integers between 0 and $N - 1$:

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

This set is called an *n-dimensional N-lattice* or briefly an *N-lattice*. In [12] this definition was extended to more general lattices in the following way: Let $\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_n}$ be n linearly independent n-dimensional vectors over the field of the real numbers such that the i-th coordinate of $\mathbf{u_i}$ is a positive integer and the other coordinates of $\mathbf{u_i}$ are 0, so that $\mathbf{u_i}$ is of the form $(0, \dots, 0, z_i, 0, \dots, 0)$ (with $z_i \in \mathbb{N}$). Let $t_1, t_2, \dots, t_n$ be integers with $0 \le t_1, t_2, \dots, t_n < N$. Then we call the set

$$\begin{aligned} B_N^n = \{\mathbf{x} = x_1\mathbf{u_1} + \dots + x_n\mathbf{u_n} :\ & x_i \in \mathbb{N} \cup \{0\},\ 0 \le x_i\,|\mathbf{u_i}| \le t_i (< N) \\ & \text{for } i = 1, \dots, n\} \end{aligned}$$

an *n-dimensional box N-lattice* or briefly a *box N-lattice.*

In [13] the definition of binary sequences was extended to more dimensions by considering functions of type

$$\eta(\mathbf{x}) :\ I_N^n \to \{-1, +1\}.$$

If $\mathbf{x} = (x_1, \dots, x_n)$ so that $\eta(\mathbf{x}) = \eta((x_1, \dots, x_n))$ then we will simplify the notation slightly by writing $\eta(\mathbf{x}) = \eta(x_1, \dots, x_n)$. Such a function can be visualized as the lattice points of the N-lattice replaced by the two symbols $+$ and $-$, thus they are called *binary N-lattices.*

In [13] Hubert, Mauduit and Sárközy introduced the following measures of pseudorandomness of binary lattices (here we will present the definition in the same slightly modified but equivalent form as in [12]):

$$\eta : I_N^n \to \{-1, +1\}.$$

Define the pseudorandom measure of order k of η by

$$Q_k(\eta) = \max_{B, \mathbf{d_1}, \dots, \mathbf{d_k}} \left| \sum_{\mathbf{x} \in B} \eta(\mathbf{x} + \mathbf{d_1}) \cdots \eta(\mathbf{x} + \mathbf{d_k}) \right|,$$

where the maximum is taken over all distinct $\mathbf{d_1}, \dots, \mathbf{d_k} \in I_N^n$ and all box N-lattices B such that $B + \mathbf{d_1}, \dots, B + \mathbf{d_k} \subseteq I_N^n$. Note that in the one dimensional special case $Q_k(\eta)$ is the same as the combined pseudorandom measure (1.2) for every k and, in particular $Q_1(\eta)$ is the well-distribution measure W in (1.1).

Then η is said to have strong pseudorandom properties, or briefly, it is considered as a "good" pseudorandom binary lattice if for fixed n and k and "large" N the measure $Q_k(\eta)$ is "small" (much smaller, than the trivial upper bound N^n). This terminology is justified by the fact that, as it was proved in [13], for a truly random binary lattice defined on I_N^n and for fixed k the measure $Q_k(\eta)$ is "small", more precisely, it is less than $N^{n/2}$ multiplied by a logarithmic factor. As in the one-dimensional case, a list of papers written on pseudorandomness of binary lattices and on the measures of pseudorandomness is presented in [9]; see also the more recent papers [10] and [11].

In the applications one may need not just a single binary sequence resp. lattice with strong pseudorandom properties but a large family of them. Moreover, in many applications it is not enough if our family $\mathcal{F}$ is large; it can be much more important to know that $\mathcal{F}$ has a "rich", "complex" structure, there are many "independent" sequences, resp. lattices in it which are "far apart". Thus one needs quantitative measures for these properties of families of binary sequences resp. lattices. In case of binary sequences such a measure was introduced by Ahlswede, Khachatrian, Mauduit and Sárközy in [1]:

Let $\mathcal{F}$ be a family of binary sequences $E_N = (e_1, e_2, \dots, e_N) \in \{-1, +1\}^N$, and let $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j) \in \{-1, +1\}^j$ be a fixed binary sequence of length j (for some $j \leq N$), and let $1 \leq i_1 < i_2 < \dots < i_j \leq N$. If we consider binary sequences $E_N = (e_1, e_2, \dots, e_N) \in \{-1, +1\}^N$ with

$$e_{i_1} = \varepsilon_1,\ e_{i_2} = \varepsilon_2, \dots, e_{i_j} = \varepsilon_j, \tag{1.3}$$

Definition 1. (1.3) is said to be a *specification* of length j (of the binary sequence E_N).

Definition 2 ([1]). The *family complexity* or briefly *f-complexity* of a family $\mathcal{F}$ of binary sequences $E_N \in \{-1, +1\}^N$ is defined as the greatest integer j so that for any specification (1.3) (of length j) there is at least one $E_N \in \mathcal{F}$ which satisfies it. The f-complexity of $\mathcal{F}$ is denoted by $\Gamma(\mathcal{F})$. (If there is no $j \in \mathbb{N}$ with the property above, we set $\Gamma(\mathcal{F}) = 0$.)

Note that an easy consequence of the definition is that

$$2^{\Gamma(\mathcal{F})} \leq |\mathcal{F}|$$

whence

Proposition 1.

$$\Gamma(\mathcal{F}) \leq \frac{\log |\mathcal{F}|}{\log 2}. \tag{1.4}$$

Goubin, Mauduit and Sárközy [7] constructed the first large family of binary sequences with strong pseudorandom properties by using the Legendre symbol. They showed that if p is a prime number, K is "not very large" in terms of p, we consider all polynomials $f(x) \in \mathbb{F}_p[x]$ such that $0 < \deg f(x) \leq K$ and $f(x)$ has no multiple zeros, and each of these polynomials $f(x)$ we assign a binary sequence $E_p = (e_1, e_2, \dots, e_p)$ defined by

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

(where $\left(\frac{\cdots}{p}\right)$ is the Legendre symbol), then all these binary sequences possess strong pseudorandom properties (both $W(E_p)$ and $C_k(E_p)$ for k "not very large" are small). Let $\mathcal{F}$ denote the family of these binary sequences E_p. Ahlswede, Khachatrian, Mauduit and Sárközy [1] showed that the f-complexity $\Gamma(\mathcal{F})$ of this family is large. Later Gyarmati [8] improved on their lower bound by showing that $\Gamma(\mathcal{F}) > c \log |\mathcal{F}|$ with some explicit constant c; note that by (1.4), this estimate is best possible apart from the value of this constant c, and the complexity of this family is optimally large apart from the constant factor. (See also [6].)

Another important tool of studying the pseudorandomness of families of binary sequences is the notion of *collision* [3], [19], [20], [21]:

Assume that $N \in \mathbb{N}$, $\mathcal{S}$ is a given set (e.g., a set of certain polynomials or the set of all the binary sequences of a given length much less than N), to each $s \in \mathcal{S}$ we assign a unique binary sequence

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

and let $\mathcal{F} = \mathcal{F}(\mathcal{S})$ denote the family of the binary sequences obtained in this way:

$$\mathcal{F} = \mathcal{F}(\mathcal{S}) = \{E_N(s) : s \in \mathcal{S}\}. \tag{1.5}$$

Definition 3. If $s \in \mathcal{S}$, $s' \in \mathcal{S}$, $s \neq s'$ and

$$E_N(s) = E_N(s'), \tag{1.6}$$

then (1.6) is said to be a *collision* in $\mathcal{F} = \mathcal{F}(\mathcal{S})$. If there is no collision in $\mathcal{F} = \mathcal{F}(\mathcal{S})$, then $\mathcal{F}$ is said to be *collision free.*

In other words, $\mathcal{F} = \mathcal{F}(\mathcal{S})$ is collision free if we have $|\mathcal{F}| = |\mathcal{S}|$. An ideally good family of pseudorandom binary sequences is collision free. If $\mathcal{F}$ is not collision free but the number of collisions is "small", then they may cause only minor problems. A good measure of the number of collisions is the following:

Definition 4. The *collision maximum* $M = M(\mathcal{F}, \mathcal{S})$ is defined by

$$M = M(\mathcal{F}, \mathcal{S}) = \max_{E_N \in \mathcal{F}} |\{s : s \in \mathcal{S},\ E_N(s) = E_N\}|$$

(i.e., M is the maximal number of elements of $\mathcal{S}$ representing the same binary sequence E_N, and $\mathcal{F} = \mathcal{F}(\mathcal{S})$ is collision free if and only if $M(\mathcal{F}, \mathcal{S}) = 1$).

There is another related notion appearing in the literature, namely, the notion of avalanche effect (see, e.g., [3], [5], [6], [14], [20], [21]):

Definition 5. If in (1.5) we have $\mathcal{S} = \{-1,+1\}^{\ell}$, and for any $s \in \mathcal{S}$, changing any element of s changes "many" elements of $E_N(s)$ (i.e., for $s \neq s'$ many elements of the sequences $E_N(s)$ and $E_N(s')$ are different), then we speak about *avalanche effect*, and we say that $\mathcal{F} = \mathcal{F}(\mathcal{S})$ possesses the avalanche property. If $N \to \infty$ and for any $s \in \mathcal{S}$, $s' \in \mathcal{S}$, $s \neq s'$ at least $\big(\frac{1}{2} - o(1)\big) N$ elements of $E_N(s)$ and $E_N(s')$ are different, then $\mathcal{F}$ is said to possess *strict avalanche property.*

To study the avalanche property, one may introduce the following quantitative measure:

Definition 6. If $N \in \mathbb{N}$, $E_n = (e_1, \dots, e_N) \in \{-1,+1\}^N$ and $E'_n = (e'_1, \dots, e'_N) \in \{-1,+1\}^N \in \{-1,+1\}^N$, then the *distance* $d(E_N, E'_N)$ between E_N and E'_N is defined by

$$d(E_N, E'_N) = |\{n : 1 \leq n \leq N,\ e_n \neq e'_n\}|$$

(a similar notion is introduced in [3]; this is a variant of the Hamming distance). Moreover, if $\mathcal{F}$ is a family from (1.5), then the *distance minimum* $m(\mathcal{F})$ of $\mathcal{F}$ is defined by

$$m(\mathcal{F}) = \min_{\substack{s,s' \in \mathcal{S} \\ s \neq s'}} d(E_N(s), E_N(s')).$$

Applying this notion we may say that the family $\mathcal{F}$ in (1.5) is collision free if and only if $m(\mathcal{F}) > 0$, and $\mathcal{F}$ possesses the strict avalanche property if

$$m(\mathcal{F}) \geq \left(\frac{1}{2} - o(1)\right) N.$$

In [20] Tóth studied the Legendre symbol construction described after Proposition 1, and she showed that a variant of the family defined there (she replaced the condition $\deg f(x) \leq K$ by $\deg f(x) = K$) is collision free if $K < p^{1/2}/2$, and it possesses the strong avalanche effect for $p \to \infty$, $K = o(p^{1/2})$. (In [20] and [21] she also studied a further construction using additive characters, she showed that there are many collisions in it, but a large subfamily of it possesses the strong avalanche property.)

Here first in Section 2 we will generalize the above definitions to n dimensions, i.e., to binary lattices. Then in Section 3 and 4 we will study a family of binary lattices constructed by using quadratic characters of finite fields and polynomials (and we will prove the n-dimensional analogues of some results of Tóth [20], [21]). In Part II of this paper we will study two further families of binary lattices constructed by using finite fields, polynomials and the notion of the multiplicative inverse.

2. Family complexity, collision, avalanche property for families of binary lattices

Each of Definitions 1–6 can be extended easily from one dimension to n dimensions, i.e., from binary sequences to binary lattices. For the sake of completeness we will present the generalizations of these definitions without adding any comments.

Let $\mathcal{F}$ be a family of binary lattices $\eta : \ I_N^n \to \{-1,+1\}$, let $j \leq N^n$, let $\mathbf{x_1}, \mathbf{x_2}, \dots, \mathbf{x_j}$ be j distinct vectors from I_N^n, and let $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j) \in \{-1,+1\}^j$. If we consider binary lattices $\eta : \ I_N^n \to \{-1,+1\}$ with

$$\eta(\mathbf{x_1}) = \varepsilon_1,\ \eta(\mathbf{x_2}) = \varepsilon_2, \dots, \eta(\mathbf{x_j}) = \varepsilon_j, \tag{2.1}$$

then

Definition 7. (2.1) is said to be a specification of length j of η.

Definition 8. The *family complexity* or *f-complexity* of a family $\mathcal{F}$ of binary lattices $\eta : \ I_N^n \to \{-1,+1\}$, denoted by $\Gamma(\mathcal{F})$, is defined as the greatest integer j so that for any specification (2.1) of length j there is at least one $\eta \in \mathcal{F}$ which satisfies it.

Then again (1.4) holds.

Assume that $N \in \mathbb{N}$, $n \in \mathbb{N}$, $\mathcal{S}$ is a given finite set, to each $s \in \mathcal{S}$ we assign a unique binary lattice $\eta = \eta_s : \ I_N^n \to \{-1,+1\}$, and let $\mathcal{F} = \mathcal{F}(\mathcal{S})$ denote the family of the binary sequences obtained in this way:

$$\mathcal{F} = \mathcal{F}(\mathcal{S}) = \{\eta_s : s \in \mathcal{S}\}. \tag{2.2}$$

Definition 9. If $s \in \mathcal{S}$, $s' \in \mathcal{S}$, $s \neq s'$ and $\eta_s = \eta_{s'}$, then this is said to be a *collision* in $\mathcal{F} = \mathcal{F}(\mathcal{S})$. If there is no collision in $\mathcal{F} = \mathcal{F}(\mathcal{S})$, then $\mathcal{F}$ is said to be *collision free.*

(We leave the generalization of Definition 4 to the reader.)

Definition 10. If $\mathcal{F}$ is of form (2.2), and for any $s \in \mathcal{S}$ changing any element of s changes "many" elements of $\eta_s : I_N^n \to \{-1,+1\}$, then we speak about *avalanche effect*, and we say that $\mathcal{F} = \mathcal{F}(\mathcal{S})$ possesses the *avalanche property.* If for any $s \in \mathcal{S}$, $s' \in \mathcal{S}$, $s \neq s'$ at least $\left(\frac{1}{2} - o(1)\right) N^n$ elements of η_s and $\eta_{s'}$ are different, then $\mathcal{F}$ is said to possess the *strict avalanche property.*

Definition 11. If $N \in \mathbb{N}$, $n \in \mathbb{N}$, $\eta : \ I_N^n \to \{-1, +1\}$ and $\eta' : \ I_N^n \to \{-1, +1\}$, then the distance $d(\eta, \eta')$ between η and η' is defined by

$$d(\eta, \eta') = |\{(x_1, x_2, \dots, x_n) : (x_1, \dots, x_n) \in \mathbb{I}_N^n, \ \eta(x_1, \dots, x_n) \neq \eta'(x_1, \dots, x_n)\}|.$$

If $\mathcal{F}$ is a family of form (2.2), then the *distance minimum* $m(\mathcal{F})$ is defined by

$$m(\mathcal{F}) = \min_{\substack{s, s' \in \mathcal{S} \\ s \neq s'}} d(\eta_s, \eta_{s'}).$$

(So that $\mathcal{F}$ is collision free if $m(\mathcal{F}) > 0$, and it possesses the strict avalanche property if

$$m(\mathcal{F}) \geq \left(\frac{1}{2} - o(1)\right) N^n.)$$

3. A family of binary lattices constructed using quadratic characters: family complexity

Mauduit and Sárközy [17] constructed a large family of binary lattices with strong pseudorandom properties by using quadratic characters of finite fields (this construction generalizes the one dimensional constructions in [7] and [16]). They proved the following theorem:

Theorem A. *Assume that $q = p^n$ is the power of an odd prime, $f(x) \in \mathbb{F}_q[x]$ has degree ℓ with*

$$0 < \ell < p,$$

and $f(x)$ has no multiple zero in $\overline{\mathbb{F}}_q$. Denote the quadratic character of $\mathbb{F}_q$ by γ (setting also $\gamma(0) = 0$). Consider the linear vector space formed by the elements of $\mathbb{F}_q$ over $\mathbb{F}_p$, and let $v_1, \dots, v_n$ be a basis of this vector space (i.e., assume that $v_1, v_2, \dots, v_n$ are linearly independent over $\mathbb{F}_p$). Define the n dimensional binary p-lattice $\eta : \ I_p^n \to \{-1, +1\}$ by

$$\eta(\mathbf{x}) = \eta((x_1, \dots, x_n)) = \begin{cases} \gamma(f(x_1 v_1 + \dots + x_n v_n)) \\ \qquad \textit{for } f(x_1 v_1 + \dots + x_n v_n) \neq 0 \\ +1 \quad \textit{for } f(x_1 v_1 + \dots + x_n v_n) = 0. \end{cases} \tag{3.1}$$

Assume also $k \in \mathbb{N}$ and

$$4^{n(k+\ell)} < p. \tag{3.2}$$

Then we have

$$Q_k(\eta) < k\ell(q^{1/2}(1+\log p)^n + 2). \tag{3.3}$$

Indeed this is a combination of Theorems 1 and 2 in [18].

Now define p, q, n as above, and set

$$L = \frac{1}{2\log 4}\frac{\log p}{n}. \tag{3.4}$$

Let $\mathcal{F}_L$ denote the family of the binary lattices η assigned to the *monic* polynomials f satisfying the conditions in Theorem A with

$$0 < \deg f = \ell < L.$$

Then for every k with

$$k < L \tag{3.5}$$

(3.2) holds, thus by Theorem A all these lattices η satisfy (3.3) for every k satisfying (3.5), so that *all these lattices η possess strong pseudorandom properties* in this sense.

Now we will show that this family $\mathcal{F}_L$ is also of large complexity and, indeed, this is so for any number K with $0 < K < p$ in place of the number L defined by (3.4):

Theorem 1. *Assume that $q = p^n$ is the power of an odd prime, let*

$$0 < K < p,$$

and consider all the polynomials $f(x) \in \mathbb{F}_q[x]$ such that

$$0 < \deg f < K$$

and $f(x)$ has no multiple zero in $\overline{\mathbb{F}}_q$. To each of these polynomials f assign the binary lattice η defined by (3.1) as described in Theorem A, and let $\mathcal{F}_K$ denote the family of these binary lattices. Then we have

$$\Gamma(\mathcal{F}_K) > \frac{K-1}{2\log 2}\log q - cK\log(K\log q) \tag{3.6}$$

with some absolute constant c.

Note that the number of polynomials $f \in \mathbb{F}_q[x]$ with $\deg f < K$ is clearly at most q^{K+1}, thus we have

$$|\mathcal{F}_K| \le |\{f :\ f \in \mathbb{F}_q[x],\ \deg f < K\}| \le q^{K+1}. \tag{3.7}$$

It follows from (1.4) and (3.7) that

$$\Gamma(\mathcal{F}_K) \le \frac{\log|\mathcal{F}_K|}{\log 2} \le \frac{(K+1)\log q}{\log 2} \tag{3.8}$$

so that the lower bound (3.6) is best possible apart from a constant factor at most.

Proof of Theorem 1. Gyarmati's method used in the one-dimensional case in [8] can be adapted. Since a considerable part of the proof will be similar to the one in [8] thus we will leave some details to the reader. □

If c is large enough and $K \ge q^{1/2}/\log q$ then the right hand side of (3.6) is negative thus (3.6) holds trivially. Thus we may assume that

$$K < q^{1/2}/\log q. \tag{3.9}$$

Let h be the greatest odd integer with $h < K$. Let $j \in \mathbb{N}$,

$$j \le \frac{h}{2\log 2}\log q - \frac{c'h}{\log 2}\log(h\log q) \tag{3.10}$$

where we will fix the value of the absolute constant c' later. Assume that we are looking for a binary lattice $\eta \in \mathcal{F}_K$ satisfying the specification

$$\eta(\mathbf{x_1}) = \varepsilon_1,\ \eta(\mathbf{x_2}) = \varepsilon_2, \dots, \eta(\mathbf{x_j}) = \varepsilon_j. \tag{3.11}$$

Let $\varphi :\ I_p^n \to \mathbb{F}_q$ be the mapping defined so that for $\mathbf{x} = (x_1, \dots, x_n) \in I_p^n$ we have

$$\varphi(\mathbf{x}) = \varphi((x_1, \dots, x_n)) = x_1v_1 + \dots + x_nv_n \in \mathbb{F}_q.$$

Clearly, this is a bijection, and the definition of η in (3.1) can be rewritten as

$$\eta(\mathbf{x}) = \begin{cases} \gamma(f(\varphi(\mathbf{x}))) & \text{for } f(\varphi(\mathbf{x})) \ne 0 \\ +1 & \text{for } f(\varphi(\mathbf{x})) = 0. \end{cases} \tag{3.12}$$

For each of the vectors $\mathbf{x_i} \in I_p^n$ considered in (3.11) write $\varphi(\mathbf{x_i}) = y_i (\in \mathbb{F}_q)$. Then by (3.12), the specification in (3.11) can be rewritten as

$$\gamma(f(y_1)) = \varepsilon_1, \quad \gamma(f(y_2)) = \varepsilon_2, \dots, \gamma(f(y_j)) = \varepsilon_j$$

$$\text{for}\quad f(y_1), f(y_2), \dots, f(y_j) \neq 0. \tag{3.13}$$

Write $\mathcal{Y} = (y_1, y_2, \dots, y_j)$. Now let $\mathcal{A}$ denote the set of the h-tuples $(a_1, a_2, \dots, a_h)$ with $a_i \in \mathbb{F}_q \setminus \mathcal{Y}$ for $i = 1, 2, \dots, h$, and consider all the polynomials $f(z) \in \mathbb{F}_q[z]$ of the form

$$f_{a_1,a_2,\dots,a_h}(z) = (z - a_1)(z - a_2)\dots(z - a_h) \quad \text{with } (a_1, a_2, \dots, a_h) \in \mathcal{A}.$$

We will prove by a counting argument that there is at least one h-tuple $(a_1, a_2, \dots, a_h) \in \mathcal{A}$ for which the binary lattice η defined by (3.12) with $f_{a_1,a_2\dots,a_h}(z)$ in place of $f(z)$ satisfies (3.13). Let $\beta_1, \beta_2 \dots, \beta_t$ denote those zeros of $f_{a_1,a_2,\dots,a_h}(z)$ which have odd multiplicity in the factorization of it. Since the degree of $f_{a_1,a_2,\dots,a_h}(z)$ is odd, the number t of these zeros is also odd thus we have $t \geq 1$. Write $g_{a_1,a_2,\dots,a_h}(z) = (z-\beta_1)(z-\beta_2)\dots(z-\beta_t)$. Then $g_{a_1,a_2,\dots,a_h}(z)$ has no multiple zero and its degree is $t \leq h < K$ so that the binary lattice defined by (3.1) with $g_{a_1,a_2,\dots,a_h}(z)$ in place of $f(z)$ belongs to $\mathcal{F}_K$, and it satisfies the specification (3.11). Since this holds for every j satisfying (3.10), it follows that

$$\Gamma(\mathcal{F}_K) \geq \left[\frac{h}{2\log 2}\log q - \frac{c'h}{\log 2}\log(h\log q)\right]$$

which proves (3.6).

Thus, indeed, it remains to prove that there is an h-tuple $(a_1, a_2, \dots, a_h)$ for which the lattice η in (3.12) with $f_{a_1,a_2,\dots,a_h}(z)$ in place of $f(z)$ satisfies (3.13). To show this, consider a h-tuple $(a_1, a_2, \dots, a_h) \in \mathcal{A}$ and the polynomial

$$f_{a_1,a_2,\dots,a_h}(x) = (x - a_1)(x - a_2)\dots(x - a_h)$$

assigned to this h-tuple. Define the binary lattice $\eta : I_p^n \to \{-1, +1\}$ as in (3.12) with $f_{a_1,a_2,\dots,a_h}(z)$ in place of $f(z)$:

$$\eta(\mathbf{x}) = \begin{cases} \gamma(f_{a_1,a_2,\dots,a_h}(\varphi(\mathbf{x}))) & \text{if } f_{a_1,a_2,\dots,a_h}(\varphi(\mathbf{x})) \neq 0, \text{ i.e.,} \\ & \varphi(\mathbf{x}) \neq a_i \text{ for } 1 \leq i \leq h, \\ +1 & \text{if } f_{a_1,a_2,\dots,a_h}(\varphi(\mathbf{x})) = 0, \text{ i.e.,} \\ & \varphi(\mathbf{x}) = a_i \text{ for some } 1 \leq i \leq h. \end{cases} \tag{3.14}$$

Clearly,

$$\frac{1}{2}\left(1 + \varepsilon_i \eta(\mathbf{x}_i)\right) = \begin{cases} 1 & \text{if } \eta(\mathbf{x}_i) = \varepsilon_i \\ 0 & \text{if } \eta(\mathbf{x}_i) = -\varepsilon_i \end{cases} \tag{3.15}$$

for $i = 1, 2, \dots, j$. If $i = 1, 2, \dots, j$ then $\varphi(\mathbf{x}_i) = y_i$, and for $t = 1, 2, \dots, h$ we have $a_t \in \mathbb{F}_q \setminus \mathcal{Y}$ whence $a_t \neq y_i$. It follows that $f_{a_1,a_2,\dots,a_h}(\varphi(\mathbf{x_i})) = f_{a_1,a_2,\dots,a_h}(y_i) = (y_i - a_1)(y_i - a_2)\dots(y_i - a_h) \neq 0$, thus by (3.14) we have

$$\begin{aligned}\eta(\mathbf{x}_i) &= \gamma(f_{a_1,a_2,\dots,a_h}(\varphi(\mathbf{x}_i))) = \gamma(f_{a_1,a_2,\dots,a_h}(y_i)) \\ &= \gamma((y_i - a_1)(y_i - a_2)\dots(y_i - a_h)) \quad (\text{for } i = 1, 2, \dots, j).\end{aligned}$$

Thus (3.15) can be rewritten as

$$\frac{1}{2}\left(1 + \varepsilon_i \gamma\left((z_i - a_1)(z_i - a_2)\dots(z_i - a_n)\right)\right) = \begin{cases} 1 & \text{if } \eta(\mathbf{x}_i) = \varepsilon_i, \\ 0 & \text{if } \eta(\mathbf{x}_i) = -\varepsilon_i. \end{cases} \tag{3.16}$$

Let N denote the number of polynomials $f_{a_1,a_2,\dots,a_h}(x) \in \mathbb{F}_q[x]$ with $(a_1, a_2, \dots, a_h) \in \mathcal{A}$ such that for binary lattice (3.14) specification (3.11) holds. Then by (3.16) we have

$$\begin{aligned}N = &\sum_{a_1 \in \mathbb{F}_q \setminus \mathcal{Y}} \sum_{a_2 \in \mathbb{F}_q \setminus \mathcal{Y}} \dots \\ &\dots \sum_{a_h \in \mathbb{F}_q \setminus \mathcal{Y}} \frac{1}{2^j} \prod_{i=1}^{j} \left(1 + \varepsilon_i \gamma\left((y_i - a_1)(y_i - a_2)\dots(y_i - a_h)\right)\right).\end{aligned} \tag{3.17}$$

In the same way as (3.3) was deduced from (3.2) in [8], by using the multiplicativity of λ one may deduce from (3.17) that

$$\begin{aligned}N = \frac{(q-j)^h}{2^j} + \frac{1}{2^j} \sum_{\ell=1}^{j} &\sum_{1 \le i_1 < i_2 < \dots < i_\ell \le j} \varepsilon_{i_1} \varepsilon_{i_2} \dots \varepsilon_{i_\ell} \\ &\left(\sum_{a \in \mathbb{F}_q \setminus \mathcal{Y}} \gamma\left((y_{i_1} - a)(y_{i_2} - a)\dots(y_{i_\ell} - a)\right) \right)^h.\end{aligned} \tag{3.18}$$

Now we need

Lemma 1. *If $q = p^n$ is a prime power, χ is a non-principal character modulo q of order d, $f(x) \in \mathbb{F}_q[x]$ has s distinct zeros in $\overline{\mathbb{F}}_q$ and it is not the constant multiple of the d-th power of a polynomial over $\mathbb{F}_q$, then*

$$\left| \sum_{z \in \mathbb{F}_q} \chi(f(z)) \right| \le (s-1) q^{1/2}.$$

Proof of Lemma 1. This is a special case of Weil's theorem [22].

By Lemma 1 we have

$$\left|\sum_{a\in\mathbb{F}_q\setminus\mathcal{Y}}\gamma\left((y_{i_1}-a)(y_{i_2}-a)\dots(y_{i_\ell}-a)\right)\right|$$
$$\le\left|\sum_{a\in\mathbb{F}_q}\gamma\left((y_{i_1}-a)(y_{i_2}-a)\dots(y_{i_\ell}-a)\right)\right|+\sum_{a\in\mathcal{Y}}1\le \ell q^{1/2}+j\le j(q^{1/2}+1).$$

Thus it follows from (3.18) that

$$N\ge\frac{(q-j)^h}{2^j}-\frac{1}{2^j}\sum_{\ell=1}^{j}\sum_{1\le i_1<i_2<\dots<i_\ell\le j}\Big(j(q^{1/2}+1)\Big)^h>\frac{(q-j)^h}{2^j}-\Big(j(q^{1/2}+1)\Big)^h.$$

Thus in order to prove $N>0$ we have to show that

$$\frac{q-j}{2^{j/h}}>j(q^{1/2}+1)$$

or, in equivalent form,

$$q>2^{j/h}\left(jq^{1/2}+j\right)+j. \tag{3.19}$$

With p in place of q this is inequality (12) in [8] and it was shown in [8] that it follows from (5) and (6) if $c_1=9$ is chosen. Replacing p by q and c_1 by c' in these two formulas, we obtain (3.9) and (3.10) above, so that if j satisfies (3.10) then (3.19) holds whence $N>0$ follows. Thus there is a binary lattice $\eta\in\mathcal{F}_K$ which satisfies specification (3.11) and this completes the proof of Theorem 1. □

4. A family of binary lattices constructed using quadratic characters: collisions, avalanche effect

Now we will show that if K is "not very large", then the family $\mathcal{F}_K$ of binary lattices defined in Theorem 1 is collision free, and it also possesses the strict avalanche property. Again, let $q=p^n$ be a fixed odd prime power and $0<K<p$. Let $\mathcal{S}_K$ denote the set of *monic* polynomials $f(x)\in\mathbb{F}_q[x]$ such that $0<\deg f<K$. For every polynomial $f\in\mathcal{S}_K$ we consider the binary lattice η defined by (3.1) as described in Theorem A, and we denote it by η_f. Then the family $\mathcal{F}_K$ of binary lattices defined in Theorem 1 is the set of these lattices η_f:

$$\mathcal{F}_K=\mathcal{F}_K(\mathcal{S}_K)=\{\eta_f:f\in\mathcal{S}_K\}.$$

Using these notations we have

Theorem 2.
$$m(\mathcal{F}_K) > \frac{1}{2}\left(q-(2K-1)q^{1/2}-2K\right).$$

Note that if $K < \frac{1}{2}q^{1/2}$, then it follows from Theorem 2 that

$$m(\mathcal{F}_K) > \frac{1}{2}\left(q-(2K-1)q^{1/2}-q^{1/2}\right) = \frac{1}{2}\left(q-2Kq^{1/2}\right) > 0$$

and thus $\mathcal{F}_K$ is collision free. This proves

Corollary 1. *If $\mathcal{S}_K$, $\mathcal{F}_K$ are defined as above and we also have $K < \frac{1}{2}q^{1/2}$, then $\mathcal{F}_K$ is collision free.*

Moreover, if $q \to \infty$ and $K = o(q^{1/2})$ then Theorem 2 gives

$$m(\mathcal{F}_K) \geq \left(\frac{1}{2} - o(1)\right) q$$

which proves

Corollary 2. *If $\mathcal{S}_K$, $\mathcal{F}_K$ are defined as above and we have $q \to \infty$, $K = o(q^{1/2})$, then $\mathcal{F}_K$ possesses the strict avalanche property.*

Proof of Theorem 2. We will adapt Tóth's method [20]. Assume that $f, g \in \mathcal{S}_K$and $f \neq g$. Then for $\mathbf{x} \in I_p^n$ we have

$$\eta_f(\mathbf{x})\eta_g(\mathbf{x}) = \begin{cases} +1 & \text{if } \eta_f(\mathbf{x}) = \eta_g(\mathbf{x}) \\ -1 & \text{if } \eta_f(\mathbf{x}) \neq \eta_g(\mathbf{x}) \end{cases}$$

whence

$$\frac{1}{2}\left(1-\eta_f(\mathbf{x})\eta_g(\mathbf{x})\right) = \begin{cases} 0 & \text{if } \eta_f(\mathbf{x}) = \eta_g(\mathbf{x}) \\ 1 & \text{if } \eta_f(\mathbf{x}) \neq \eta_g(\mathbf{x}). \end{cases}$$

It follows that

$$\begin{aligned} d(\eta_f, \eta_g) &= \sum_{\mathbf{x}\in I_p^n} \frac{1}{2}\left(1-\eta_f(\mathbf{x})\eta_g(\mathbf{x})\right) = \frac{1}{2}\Bigg(p^n - \sum_{\mathbf{x}\in I_p^n} \eta_f(\mathbf{x})\eta_g(\mathbf{x})\Bigg) \\ &= \frac{1}{2}\Bigg(q - \sum_{\mathbf{x}\in I_p^n} \eta_f(\mathbf{x})\eta_g(\mathbf{x})\Bigg). \end{aligned}$$

Then using again the bijection $\varphi: I_p^n \to \mathbb{F}_q$ introduced at the beginning of the proof of Theorem 1, by (3.12) this can be rewritten as

$$d(\eta_f, \eta_g) = \frac{1}{2}\Bigg(q - \sum_{\substack{\mathbf{x}\in I_p^n \\ f(\varphi(\mathbf{x}))g(\varphi(\mathbf{x}))\neq 0}} \gamma(f(\varphi(\mathbf{x})))\gamma(g(\varphi(\mathbf{x}))) - \sum_{\substack{\mathbf{x}\in I_p^n \\ f(\varphi(\mathbf{x}))g(\varphi(\mathbf{x}))=0}} \eta_f(\mathbf{x})\eta_g(\mathbf{x})\Bigg)$$

$$= \frac{1}{2}\Biggl(q - \sum_{\substack{\mathbf{x}\in I_p^n \\ (fg)(\varphi(\mathbf{x}))\neq 0}} \gamma((fg)(\varphi(\mathbf{x}))) - \sum_{\substack{\mathbf{x}\in I_p^n \\ (fg)(\varphi(\mathbf{x}))=0}} \eta_f(\mathbf{x})\eta_g(\mathbf{x}) \Biggr)$$

$$\geq \frac{1}{2}\Biggl(q - \sum_{z\in\mathbb{F}_q} \gamma((fg)(z)) - \sum_{\substack{z\in\mathbb{F}_q \\ fg(z)=0}} 1 \Biggr)$$

$$> \frac{1}{2}\Biggl(q - \sum_{z\in\mathbb{F}_q} \gamma((fg)(z)) - 2K \Biggr). \tag{4.1}$$

The order of the character γ is 2, and since $f \neq g$, both polynomials are monic and f, g have no multiple zeros, thus $f(x)g(x)$ is not the constant multiple of the square of a polynomial over $\mathbf{F}_q$. Thus we may apply Lemma 1 with γ and fg in place of χ and f, respectively. Since the polynomial fg has less than $2K$ zeros in $\overline{\mathbb{F}}_p$, thus applying Lemma 1 we obtain from (4.1) that

$$d(\eta_f, \eta_g) > \frac{1}{2}\left(q - (2K-1)q^{1/2} - 2K \right)$$

which completes the proof of Theorem 2. □

References

[1] R. Ahlswede, L. H. Khachatrian and C. Mauduit, A complexity measure for families of binary sequences, *Period. Math. Hungar.* **46** (2003), 107–118.

[2] N. Alon, Y. Kohayakawa, C. Mauduit, C. G. Moreira and V. Rödl, Measures of pseudorandomness for finite sequences: typical values, *Proc. London Math. Soc.* **95** (2007), 778–812.

[3] A. Bérczes, J. Ködmön and A. Pethő, A one-way function based on norm form equations, *Period. Math. Hungar.* **49** (2004), 1–13.

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

[5] H. Feistel, W. A. Notz and J. L. Smith, Some cryptographic techniques for machine -to-machine data communications, *Proc. IEEE* **63** (1975), 1545–1554.

[6] J. Folláth, Construction of pseudorandom binary sequences using additive characters over $GF(2^k)$, II, *Period. Math. Hungar.* **60** (2010), 127–135.

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

[8] K. Gyarmati, On the complexity of a family related to the Legendre symbol, *Period. Math. Hungar.* **58** (2009), 209–215.

[9] K. Gyarmati, C. Mauduit and A. Sárközy, Measures of pseudorandomness of finite binary lattices, I. (The measures Q_k, normality), *Acta Arith.* **144** (2010), 295–313.

[10] K. Gyarmati, C. Mauduit and A. Sárközy, Measures of pseudorandomness of finite binary lattices, II. (The symmetry measures), *Ramanujan J.* **25** (2011), 155-178.

[11] K. Gyarmati, C. Mauduit and A. Sárközy, Measures of pseudorandomness of finite binary lattices, III. (Q_k, correlation, normality, minimal values), *Unif. Distrib. Theory* (*to appear*).

[12] K. Gyarmati, A. Sárközy and C. L. Stewart, On Legendre symbol lattices, *Unif. Distrib. Theory* **4** (2009), 81–95.

[13] P. Hubert, C. Mauduit and A. Sárközy, On pseudorandom binary lattices, *Acta Arith.* **125** (2006), 51–62.

[14] J. Kam and G. Davida, Structured design of substitution-permutation encryption networks, *IEEE Transactions on Computers* **28** (1979), 747–753.

[15] Y. Kohayakawa, C. Mauduit, C. G. Moreira and V. Rödl, Measures of pseudorandomness for finite sequences: minimum and typical values, Proceedings of WORDS'03, 159–169, TUCS Gen. Publ., 27, *Turku Cent. Comput. Sci., Turku,*, 2003.

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

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

[18] C. Mauduit and A. Sárközy, On large families of pseudorandom binary lattices, *Unif. Distr. Theory* **2** (2007), 23–37.

[19] A. Menezes, P. C. van Oorshot and S. Vanstone, Handbook of applied cryptography, *CRS Press, Boca Raton*, 1997.

[20] V. Tóth, Collision and avalanche effect in families of pseudorandom binary sequences, *Period. Math. Hungar.* **55** (2007), 185–196.

[21] V. Tóth, The study of collision and avalanche effect in a family of pseudorandom binary sequences, *Period. Math. Hungar.* **59** (2009), 1–8.

[22] A. Weil, Sur les courbes algébriques et les variétés qui s'en déduissent, *Actualités Sci. Ind.*, no. 1041, Hermann et Cie., Paris, 1948 (in *French*).

KATALIN GYARMATI
EÖTVÖS LORÁND UNIVERSITY
DEPARTMENT OF ALGEBRA AND NUMBER THEORY
PÁZMÁNY PÉTER SÉTÁNY 1/C
H-1117 BUDAPEST
HUNGARY

E-mail: gykati@cs.elte.hu

CHRISTIAN MAUDUIT
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

ANDRÁS SÁRKÖZY
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 December 8, 2010; revised July 30, 2011)