Pseudorandomness of finite binary sequences and lattices

Dissertation submitted to
The Hungarian Academy of Sciences
for the degree "Doctor of the HAS"

Katalin Gyarmati

Eötvös Loránd University
Budapest

2013

dc_603_12

Contents

1 Introduction

In the last hundred years some important applications such as Monte Carlo methods, wireless communications or famous encrypting algorithms (e.g. Vernam cipher) inspired intensive study of pseudorandomness of different objects. Initially, random and pseudorandom objects were generated by physical methods, but these methods have several disadvantages: they are slow, expensive, it is difficult to store the data and their pseudorandom properties cannot be proved mathematically. In order to avoid these difficulties, pseudorandom objects are generated nowdays from a small secret key by mathematical algorithms, with the intent that they appear random to a computationally bounded adversary.

Different approaches and definitions of pseudorandomness exist. Menezes, Oorschot and Vanstone [81] wrote an excellent monograph about these approaches. The most frequently used interpretation of pseudorandomness is based on complexity theory; Goldwasser [30] wrote a survey paper about this approach. In this approach usually sequences of length tending to infinity are tested while in the applications only *finite* sequences are used. Unfortunately, most of the results are based on certain unproved hypotheses (such as the difficulty of factorization of integers or the difficulty of the discrete logarithm problem). Finite pseudorandom $[0, 1)$ sequences have been studied by Niederreiter (see for example [86], [87], [88], [89]). Niederreiter [90] also studied random number generation and quasi-Monte Carlo methods and their connections.

In the second half of the 1990s, Christian Mauduit and András Sárközy [77] introduced a new constructive quantitative approach, in which the pseudorandomness of *finite binary* sequences is well characterized. Since then it is a fast developping area, several authors work in this field and several constructions, results and generalizations are presented in numerous papers. In [46] I gave a survey of the most important results.

In the present dissertation I will summarize my main results in the theory of pseudorandomness. Some of my results (see the papers [39], [41], [45], [49], [61], [62]) will be presented in details, but to keep the extent of the dissertation below a reasonable limit my other works (see the papers [37], [40], [42], [43], [44], [46], [47], [48], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59]) will be just briefly mentioned. Throughout the dissertation I will always name the authors of the theorems except for the ones proved by me without any coauthors.

In [77] Mauduit and Sárközy introduced the following pseudorandom measures:

Definition 1.1 (Mauduit, Sárközy) *For a binary sequence* $E_N = (e_1, \ldots, e_N) \in \{-1, +1\}^N$ *of length* N*, write*

$$U(E_N, t, a, b) = \sum_{j=0}^{t} e_{a+jb}.$$

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

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

where the maximum is taken over all a, b, t *such that* $a, b, t \in \mathbb{N}$ *and* $1 \le a \le a + tb \le N$*.*

The well-distribution measure studies how close are the frequencies of the $+1$'s and -1's in arithmetic progressions (for a binary sequence with strong pseudorandom properties these two quantities are expected to be very close.) But often it is also necessary to study the connections between certain elements of the sequence. For example, if the subsequence $(+1, +1)$ occurs much more frequently then the subsequence $(-1, -1)$, then it may cause problems in the applications, and we cannot say that our sequence has strong pseudorandom properties. In order to study connections of this type Mauduit and Sárközy [77] introduced the correlation and normality measures:

Definition 1.2 (Mauduit, Sárközy) *For a binary sequence* $E_N = (e_1, \ldots, e_N) \in \{-1, +1\}^N$ *of length* N*, and for* $D = (d_1, \ldots, d_\ell)$ *with non-negative integers* $0 \le d_1 < \cdots < d_\ell$*, write*

$$V(E_N, M, D) = \sum_{n=1}^{M} e_{n+d_1} \ldots e_{n+d_\ell}.$$

Then the correlation measure of order ℓ of E_N *is defined as*

$$C_\ell(E_N) = \max_{M,D} |V(E_N, M, D)| = \max_{M,D} \left| \sum_{n=1}^{M} e_{n+d_1} \ldots e_{n+d_\ell} \right|,$$

where the maximum is taken over all $D = (d_1, \ldots, d_\ell)$ *and* M *such that* $0 \le d_1 < \cdots < d_\ell < M + d_\ell \le N$*.*

Definition 1.3 (Mauduit, Sárközy) *For a binary sequence* $E_N = (e_1, \ldots, e_N) \in \{-1, +1\}^N$ *of length* N*, and for* $X = (x_1, \ldots, x_\ell) \in \{-1, +1\}^\ell$*,*

write

$$T(E_N, M, X) = \left|\{n : 0 \le n < M,\ (e_{n+1}, e_{n+2}, \ldots, e_{n+\ell}) = X\}\right|.$$

Then the normality measure of order ℓ of E_N *is defined as*

$$N_\ell(E_N) = \max_{M,X} \left|T(E_N, M, X) - M/2^\ell\right|,$$

where the maximum is taken over all $X = (x_1, \ldots, x_\ell) \in \{-1, +1\}^\ell$, *and* M *such that* $0 < M \le N - \ell + 1$.

We remark that *infinite* analogues of the functions U, V and T had been studied before (see, for example, [15], [68] and [91]), but the quantitative analysis of pseudorandom properties of *finite* sequences has started by the work of Mauduit and Sárközy [77].

The combined (well-distribution-correlation) pseudorandom measure [77] is a common generalization of the well-distribution and the correlation measures. This measure has an important role in the multidimensional extension of the theory of pseudorandomness (see Sections 5, 6, 7 and 8).

Definition 1.4 (Mauduit, Sárközy) *For a binary sequence* $E_N = (e_1, \ldots, e_N) \in \{-1, +1\}^N$ *of length* N, *and for* $a, b, t \in \mathbb{N}$, $D = (d_1, \ldots, d_\ell)$ *with non-negative integers* $0 \le d_1 < \cdots < d_\ell$, *write*

$$Z(E_N, a, b, t, D) = \sum_{j=0}^{t} e_{a+jb+d_1} \ldots e_{a+jb+d_\ell}.$$

Then the combined (well-distribution-correlation) measure of order ℓ of E_N *is defined as*

$$Q_\ell(E_N) = \max_{a,b,t,D} |Z(E_N, a, b, t, D)| = \max_{a,b,t,D} \left|\sum_{j=0}^{t} e_{a+jb+d_1} \ldots e_{a+jb+d_\ell}\right|,$$

where the maximum is taken over all a, b, t *and* $D = (d_1, \ldots, d_\ell)$ *such that all the subscripts* $a + jb + d_i$ *belong to* $\{1, 2, \ldots, N\}$.

When Mauduit and Sárközy introduced quantitative pseudorandom measures, their starting point was to balance the requirements possibly optimally. They decided to introduce functions which are real-valued and positive and the pseudorandom properties of the sequence are characterized by the sizes of the values of these functions. It was also an important requirement that

one should be able to present constructions for which these measures can be estimated well. It turned out that the measures W and C_ℓ do not only satisfy these criteria, but later Rivat and Sárközy [95] showed that if the values of W and C_ℓ are "small", then the outcome of many (previously used a posteriori) statistical tests is guaranteed to be (nearly) positive.

Although by W, C_ℓ, N_ℓ and Q_ℓ many pseudorandom properties of the sequence can be characterized, but obviously not all. For example, in [34] I introduced the symmetry measure in order to study symmetry properties of finite binary sequences (later the symmetry measure was generalized by Sziklai [102]). In [108] Winterhof gave an excellent survey on different pseudorandom measures and certain constructions. However it was also important to determine a not too large set of certain basic pseudorandom measures, which can guarantee the adequate security in the applications. The measures introduced by Mauduit and Sárközy seem to satisfy these criteria. In the case of binary sequences the most studied measures are W and C_ℓ, and many papers use only these measures, while in multidimensional extensions, the most important measure is Q_ℓ.

In [13] Cassaigne, Ferenczi, Mauduit, Rivat and Sárközy formulated the following principle: "The sequence E_N is considered a "good" pseudorandom sequence if these measures $W(E_N)$ and $C_\ell(E_N)$ (at least for "small" ℓ) are "small"."

Since 1997 many constructions with strong pseudorandom properties have been given by different authors. In 2007 Sárközy [97] presented a survey paper about the most important constructions.

One of the most intensively studied pseudorandom generator is the Blum-Blum-Shub generator, called this way after the name of its creators: Leonore Blum, Manuel Blum and Michael Shub. The unpredictability of this generator has been proved conditionally assuming the difficulty of integer factorization. In Section 2 I prove quantitative results by estimating the pseudorandom properties of the generated sequences. The power generator (an extended version of the Blum-Blum-Shub generator) will be defined in (2.1). If $p, \vartheta, t, k, T, n_0$ and the sequence u_n and E_N are defined as in Notation 2.1 and Construction 2.1, then my main result in Section 2 can be summarized as it follows.

Theorem 1.1

$$W(E_T) \ll p^{7/8} \log p,$$
$$N_\ell(E_T) \ll p^{7/8} \log p.$$

If T (the multiplicative order of k modulo t) is large in terms of p, then these bounds give a strong estimate for the well-distribution and normality measure. We remark that for the correlation measures we have slightly weaker estimates and only for shorter sequences.

Theorem 1.2 *For every $\delta > 0$, there exists a constant ε depending on ℓ and δ such that if N ($< T$) satisfies certain conditions depending on k and δ (see Theorem 2.3 in Section 2), then for the sequence E_N defined in Construction 2.1 we have*

$$C_\ell(E_N) < p^{1-\varepsilon}.$$

We remark that these results were proved only in the prime moduli case in [39], but they can be generalized to the composite moduli case at the price that the computations will be more complicated and (probably) the estimates for the pseudorandom measures will be slightly weaker. (We also note that while for public key cryptography composite moduli are used, in the case of pseudorandom generation usually we have better estimates in the prime moduli case.)

In [14] Cassaigne, Mauduit and Sárközy proved that for the majority of the sequences $E_N \in \{-1,+1\}^N$ the measures $W(E_N)$ and $C_\ell(E_N)$ are around $N^{1/2}$ (up to some logarithmic factors). Later Alon, Kohayakawa, Mauduit, Moreira and Rödl [4] improved on these bounds:

Theorem 1.A (Alon, Kohayakawa, Mauduit, Moreira, Rödl) *Suppose that we choose each $E_N \in \{-1,+1\}^N$ with probability $\frac{1}{2^N}$. For all $\varepsilon > 0$ there exist $N_0 = N_0(\varepsilon)$ and $\delta = \delta(\varepsilon) > 0$ such that for $N > N_0$ we have*

$$P\left(\delta\sqrt{N} < W(E_N) < \frac{1}{\delta}\sqrt{N}\right) > 1-\varepsilon.$$

Theorem 1.B (Alon, Kohayakawa, Mauduit, Moreira, Rödl) *Suppose that we choose each $E_N \in \{-1,+1\}^N$ with probability $\frac{1}{2^N}$. Then for all $0 < \varepsilon < 1/16$ there is a constant $N_0 = N_0(\varepsilon)$ such that for $N > N_0$ we have*

$$P\left(\frac{2}{5}\sqrt{N\log\binom{N}{\ell}} < C_\ell(E_N) < \frac{7}{4}\sqrt{N\log\binom{N}{\ell}}\right) > 1-\varepsilon.$$

We remark that while it is important that for a binary sequence with strong pseudorandom properties these measures should be "small", lower bounds are not required based on the following observations.

Write

$$m(N) = \min_{E_N \in \{-1,+1\}^N} W(E_N), \qquad M_\ell(N) = \min_{E_N \in \{-1,+1\}^N} C_\ell(E_N).$$

The estimate of $m(N)$ is a classical problem. In 1964 Roth [96] proved that $m(N) \gg N^{1/4}$. Upper bounds for $m(N)$ were given by Sárközy [21] and Beck [6]. Finally Matoušek and Spencer [74] showed that $m(N) \ll N^{1/4}$.

The value of $M_\ell(N)$ depends on the value of the order ℓ. Cassaigne, Mauduit and Sárközy [14] proved that $M_\ell(E_N) \ll (\ell N \log N)^{1/2}$. The results of [4] improved the implied constant factor (see Theorem 1.B). On the other hand, first Cassaigne, Mauduit and Sárközy [14] proved that $M_\ell(N) \gg \log(N/\ell)$ for even ℓ. This was improved considerably by Alon, Kohayakawa, Mauduit, Moreira and Rödl in [3] and [69], where the best lower bound is the following:

Theorem 1.C (Alon, Kohayakawa, Mauduit, Moreira, Rödl) *If ℓ is even then*

$$M_\ell(N) \geq \sqrt{\frac{1}{2}\left[\frac{N}{\ell+1}\right]}.$$

The proof of the theorem used deep linear algebraic tools. Later Anantharam [5] simplified the proof, but he obtained a slightly (by a constant factor) weaker result.

Cassaigne, Mauduit and Sárközy [14] noticed that the minimum values of correlation of odd order can be very small. Namely, for the sequence $E_N = (-1,+1,-1,+1,\dots) \in \{-1,+1\}^N$ we have $C_\ell(E_N) = 1$ for odd ℓ, since

$$e_{n+1+d_1} \cdots e_{n+1+d_\ell} = (-e_{n+d_1}) \cdots (-e_{n+d_\ell}) = (-1)^\ell e_{n+d_1} \cdots e_{n+d_\ell}.$$

Thus

$$\left| \sum_{n=1}^{M} e_{n+d_1} \cdots e_{n+d_\ell} \right| = |1-1+1-1+\dots| = \begin{cases} 1 & \text{if } M \text{ is odd,} \\ 0 & \text{if } M \text{ is even.} \end{cases}$$

So $C_\ell(E_N) = 1$ and thus $M_\ell(N) = 1$ for odd ℓ. Cassaigne, Mauduit and Sárközy [14] also observed that although for the sequence $E_N = (-1,+1,-1,+1,\dots)$, $C_3(E_N)$ is 1, the correlation measure of order 2 is large: $C_2(E_N) = \lceil \frac{N}{2} \rceil$. By solving problems of Cassaigne, Mauduit and Sárközy [14] and Mauduit [75], in [36] I proved that

$$C_2(E_N) C_3(E_N) \gg N^{2/3} \tag{1.1}$$

always holds. More generally, in [36] I proved an inequality involving correlation measures C_{2k+1} and $C_{2\ell}$ where $2k+1 > 2\ell$. Later Anantharam [5] sharpened (1.1). By extending the previous results, in [49] with Mauduit we were be able to compare correlation measures of C_{2k+1} and $C_{2\ell}$ (without the assumption $2k+1 > 2\ell$). Our main result was the following:

Theorem 1.3 (Gyarmati, Mauduit) *There is a constant $c_{k,\ell}$ depending only on k and ℓ such that if*

$$C_{2k+1}(E_N) < c_{k,\ell} N^{1/2},$$

then

$$C_{2k+1}(E_N)^{2\ell} C_{2\ell}(E_N)^{2k+1} \gg N^{2k+1},$$

where the implied constant factor depends only on k and ℓ.

This theorem has the following consequences:

Corollary 1.1 (Gyarmati, Mauduit) *If $C_{2k+1}(E_N) = O(1)$, then $C_{2\ell}(E_N) \gg N$, where the implied constant factor depends on k and ℓ.*

Corollary 1.2 (Gyarmati, Mauduit)

$$C_{2k+1}(E_N) C_{2\ell}(E_N) \gg N^{c(k,\ell)}$$

where the implied constant factor depends only on k and ℓ and where

$$c(k,\ell) = \begin{cases} 1 & \text{if } k \geq \ell, \\ \frac{1}{2} + \frac{2k+1}{4\ell} & \text{if } k < \ell. \end{cases}$$

In Section 3 I will prove Theorem 1.3 and its consequences.

First Goubin, Mauduit and Sárközy [31] succeeded in constructing large families of pseudorandom binary sequences. They also studied the pseudorandom properties of the generated sequences. Their construction was the following:

Construction 1.1 (Goubin, Mauduit, Sárközy) *Suppose that p is a prime number, and $f(x) \in \mathbb{F}_p[x]$ is a polynomial with degree $k > 0$ and no multiple zero in $\overline{\mathbb{F}}_p$. Define the binary sequence $E_p = (e_1, \ldots, e_p)$ by*

$$e_n = \begin{cases} \left(\frac{f(n)}{p}\right) & \text{for } (f(n), p) = 1 \\ +1 & \text{for } p \mid f(n). \end{cases} \tag{1.2}$$

Indeed, first Hoffstein and Lieman [64] proposed the use of polynomials $f(n)$ in (1.2) such that they are squarefree and neither even, nor odd, but they did not prove anything on the pseudorandom properties of the corresponding sequence $E_p = (e_1, \dots, e_p)$.

Ahlswede, Khachatrian, Mauduit and Sárközy [1] introduced the notion of family-complexity of families of binary sequences (in order to characterize the cryptographic applicability of the family). They proposed to use the following measure to study whether a family has "rich", "complex" structure or not:

Definition 1.5 (Ahlswede, Khachatrian, Mauduit, Sárközy) *The family complexity $C(\mathcal{F})$ 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 $1 \le i_1 < i_2 < \cdots < i_j \le N$, and for $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j \in \{-1,+1\}^j$, we have at least one $E_N = (e_1, \dots, e_N) \in \mathcal{F}$ for which*

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

In [1] in Section 3 it is proved that

Proposition 1.1 (Ahlswede, Khachatrian, Mauduit, Sárközy)

$$C(\mathcal{F}) \le \frac{\log |\mathcal{F}|}{\log 2}.$$

Ahlswede, Khachatrian, Mauduit and Sárközy [1] proved the following:

Theorem 1.D (Ahlswede, Khachatrian, Mauduit, Sárközy) *Let p be a prime. Consider all the polynomials $f(x)$ such that*

$$0 < \deg f(x) \le K$$

(where $\deg f(x)$ denotes the degree of $f(x)$) and $f(x)$ has no multiple zero in $\overline{\mathbb{F}}_p$. For each of these polynomials $f(x)$, consider the binary sequence $E_p = E_p(f) = (e_1, e_2, \dots, e_p) \in \{-1,+1\}^p$ defined by (1.2), and let $\mathcal{F}_1$ denote the family of all the binary sequences obtained in this way. Then

$$C(\mathcal{F}_1) \ge K. \tag{1.3}$$

By Proposition 1.1 it is clear that

$$|C(\mathcal{F}_1)| \le \frac{\log |\mathcal{F}_1|}{\log 2} \le \frac{K+1}{\log 2} \log p. \tag{1.4}$$

In [41] I improved on (1.3) and proved the following:

Theorem 1.4

$$C(\mathcal{F}_1) \gg K \log p.$$

By (1.4) this lower bound is sharp apart from the constant factor.

In this dissertation I will study the family complexity in Section 4 and I will prove Theorem 1.4 there.

For a truly random binary sequence the well-distribution measure and the correlation measures are small ($\ll N^{1/2}(\log N)^c$ for a sequence of length N). Several constructions have been given for which these measures are small ($\ll N^{1/2}(\log N)^c$), thus the sequence E_N has strong pseudorandom properties. But in certain applications, e.g. in cryptography, it is not enough to know that the sequence has strong pseudorandom properties, it is also important that the subsequences E_M (where E_M is of the form $(e_x, e_{x+1}, \dots, e_{x+M-1})$) also have strong pseudorandom properties for values M possibly small in terms of N. In Section 5 I will deal with this problem in case of values $M \gg N^{1/4+\varepsilon}$. Clearly, almost all sequences of length N consist a subsequence $(1, 1, \dots, 1)$ containing $c \log N$ of 1's, thus then for the correlation of subsequences of length $M = O(\log N)$ we cannot expect any non-trivial bound. It is an interesting open question for which sequences $E_N \in \{-1, +1\}^N$ with strong pseudorandom properties and for which values of M one can estimate $\max_{E_M=(e_x,e_{x+1},\dots,e_{x+M-1})\subset E_N} C_\ell(E_M)$ by a non-trivial upper bound. Note that this problem is related to the estimate of the least quadratic nonresidue ϑ_p modulo p. Burgess [12] proved that $\vartheta_p < p^{\frac{1}{4\sqrt{e}}+\varepsilon}$, and it is conjectured that ϑ_p is $O(\log p \log\log p)$. The difficulty of Burgess's proof and the gap between the conjecture and Burgess's result are pointing in that direction that probably one cannot prove a non-trivial bound for $\max_{E_M=(e_x,e_{x+1},\dots,e_{x+M-1})\subset E_N} C_\ell(E_M)$ when $M \ll N^c$ if c is a constant small enough.

This problem has important applications, for example, it may occur that, say, we want to encrypt a message of estimated length slightly less than N, thus we use an N bit sequence possessing strong pseudorandom properties. However, it may turn out that the text to be encrypted is of length less than, say, $\sqrt{N}$. In this case we use only a short part (of length $\sqrt{N}$) of the sequence, so we will need control over the pseudorandom properties of the short subsequences. In Section 5, I will construct a sequence for which the following holds:

Theorem 1.5 *There exists a sequence $E_N \in \{-1, +1\}^N$ for which we have*

$$C_\ell(E_M) \ll \ell^2 \left\lceil \frac{M}{N^{1/2}} \right\rceil N^{1/4} \log N$$

for every $M \le N$ *and* $E_M \subseteq E_N$ *(where* E_M *is of the form* $(e_x, e_{x+1}, \dots, e_{x+M-1})$*). Moreover*

$$C_\ell(E_N) \ll \ell^2 N^{1/2} (\log N)^2$$

and

$$W(E_N) \ll N^{3/4} \log N$$

holds.

This result was published in [45], here I will deal with these problems and prove Theorem 1.5 in Section 5. In order to prove this result we will need the *multidimensional theory of pseudorandomness*, and in Sections 6, 7 and 8 we will also need this theory.

The *multidimensional theory of pseudorandomness* was developed by Hubert, Mauduit and Sárközy [65]. They introduced the following definitions:

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

$$I_N^n = \{\mathbf{x} = (x_1, x_2, \dots, x_n) : \ x_i \in \{0, 1, \dots, N-1\}\}.$$

This set is called an *n-dimensional N-lattice* or briefly an *N-lattice*. In [61] 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 vectors, where the i-th coordinate of $\mathbf{u_i}$ is a positive integer and the other coordinates of $\mathbf{u_i}$ are 0, so that, writing $z_i = |\mathbf{u_i}|$, $\mathbf{u_i}$ is of the form $(0, \dots, 0, z_i, 0, \dots, 0)$. 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

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

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

In [65] the definition of binary sequences was extended to more dimensions by considering functions of type

$$e_{\mathbf{x}} = \eta(\mathbf{x}) : \ I_N^n \to \{-1, +1\}. \tag{1.5}$$

If $\mathbf{x} = (x_1, \dots, x_n)$ so that $\eta(\mathbf{x}) = \eta((x_1, \dots, x_n))$ then we will slightly simplify the notation 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. Binary 2 or 3 dimensional pseudorandom lattices can be used in encryption of digital images and in medical diagnostics.

Hubert, Mauduit and Sárközy [65] introduced the following measure of pseudorandomness of binary lattices (here we will present the definition in the same slightly modified but equivalent form as in [61]):

Definition 1.6 (Hubert, Mauduit, Sárközy) *Let*

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

be a binary lattice. Define the pseudorandom measure of order ℓ of η *by*

$$Q_\ell(\eta) = \max_{B, \mathbf{d_1}, \dots, \mathbf{d}_\ell} \left| \sum_{\mathbf{x} \in B} \eta(\mathbf{x} + \mathbf{d_1}) \dots \eta(\mathbf{x} + \mathbf{d}_\ell) \right|, \tag{1.6}$$

where the maximum is taken over all distinct $\mathbf{d_1}, \dots, \mathbf{d}_\ell \in I_N^n$ *and all box* N*-lattices* B *such that* $B + \mathbf{d_1}, \dots, B + \mathbf{d}_\ell \subseteq I_N^n$.

Then η is said to have strong pseudorandom properties, or briefly, it is considered a "good" pseudorandom lattice if for fixed n and ℓ and "large" N the measure $Q_\ell(\eta)$ is "small" (much smaller, then the trivial upper bound N^n). This terminology is justified by the fact that, as was proved in [65], for a truly random binary lattice defined on I_N^n and for fixed ℓ the measure $Q_\ell(\eta)$ is "small"; in particular, it is less than $N^{n/2}$ multiplied by a logarithmic factor.

In their first paper [65] on the multidimensional theory of pseudorandomness Hubert, Mauduit and Sárközy gave constructions for binary lattices with strong pseudorandom properties. They gave nearly optimal upper bounds for the pseudorandom measures of the lattices constructed. However, these early constructions also have disadvantages: they are rather artificial, and their implementation is complicated. Thus in [61] and [62] with my coauthors A. Sárközy and C. L. Stewart we defined a new construction which is based on the use of the Legendre symbol. This construction is much more natural and flexible than the earlier ones, and it can be implemented more easily. In Sections 6 and 7 I will present results from [61] and [62]. We will study the properties of the following:

Construction 1.2 (Gyarmati, Sárközy, Stewart) *Let* p *be an odd prime,* $f(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ *be a polynomial in two variables. Define* $\eta : \ I_p^2 \to \{-1, +1\}$ *by*

$$\eta(x_1, x_2) = \begin{cases} \left(\frac{f(x_1,x_2)}{p}\right) & \text{if } (f(x_1, x_2), p) = 1, \\ +1 & \text{if } p \mid f(x_1, x_2). \end{cases} \tag{1.7}$$

In Section 6.1 negative examples are presented: we will show that for certain polynomials $f(x_1, x_2)$ the associated binary lattice $\eta(x_1, x_2)$ has weak pseudorandom properties. It turns out that depending on the form of the polynomial we have to distinguish two different cases. More precisely, we say the following:

Definition 1.7 (Gyarmati, Sárközy, Stewart) *The polynomial $f(x_1, x_2)$ is called* degenerate *if it is of the form*

$$f(x_1, x_2) = \left(\prod_{j=1}^{r} f_j(\alpha_j x_1 + \beta_j x_2) \right) g(x_1, x_2)^2, \tag{1.8}$$

where $\alpha_j, \beta_j \in \mathbb{F}_p$, $f_j(x) \in \mathbb{F}_p[x]$ for $j = 1, \dots, r$, and $g(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$.

A polynomial $f \in \mathbb{F}_p[x_1, x_2]$ which can be expressed in the form (1.8) is said to be degenerate and otherwise it is said to be non-degenerate.

In Section 6 we analyze the non-degenerate case, while in Section 7 the degenerate case. These sections are based on the papers [61] and [62]. Next I present the main results from these two sections:

Theorem 1.6 (Gyarmati, Sárközy, Stewart) *Let $f(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ be a polynomial of degree k. Suppose that $f(x_1, x_2)$ cannot be expressed in the form (1.8) and one of the following 5 conditions holds:*
a) $f(x_1, x_2)$ is irreducible in $\mathbb{F}_p[x_1, x_2]$,
b) $\ell = 2$,
c) 2 is a primitive root modulo p,
d) $4^{k+\ell} < p$,
e) ℓ and the degree of the polynomial $f(x_1, x_2)$ in x_1 (or in x_2) are odd.
Then for the binary p-lattice η defined in (1.7) we have

$$Q_\ell(\eta) < 11k\ell p^{3/2} \log p. \tag{1.9}$$

In the case of degenerate polynomial we will define the rank of the polynomial as the smallest positive integer r for which $f(x_1, x_2)$ can be written in the form (1.8).

Theorem 1.7 (Gyarmati, Sárközy, Stewart) *Let $f(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ be a non-constant degenerate polynomial of rank r with degree k. Suppose that ℓ, the order of the pseudorandom measure is not greater than the rank r of $f(x_1, x_2)$, and one of the following 5 conditions holds:*
a) $f(x_1, x_2)$ is irreducible in $\mathbb{F}_p[x_1, x_2]$,
b) $\ell = 2$,

c) 2 is a primitive root modulo p,
d) $(4k)^\ell < p$ *or* $(4\ell)^k < p$*,*
e) ℓ *and the degree of the polynomial* $f(x_1, x_2)$ *in* x_1 *(or in* x_2*) are odd.*
Then for the binary lattice η *defined in* (1.7) *we have*

$$Q_\ell(\eta) < 11k\ell p^{3/2} \log p. \tag{1.10}$$

In Section 7.3 I also show that in case of degenerate polynomials, there is a pseudorandom measure of large order which is large:

Theorem 1.8 (Gyarmati, Sárközy, Stewart) *Let* $f \in \mathbb{F}_p[x_1, x_2]$ *be a degenerate polynomial with rank* r *and degree* m *and* n *in* x_1 *and* x_2*, respectively. Then there exists a positive integer* ℓ *with* $\ell \le 2^r$ *for which*

$$Q_\ell(\eta) \ge p^2 - 4rp^{3/2} - 2\ell(m+n)p.$$

Our upper bounds (1.9) and (1.10) are not optimal since they are significantly larger than the optimal $p(\log p)^c$. In Section 7.5 we will show that for a certain (rather special) family of polynomials the finite field construction presented in [79] is equivalent to a Legendre symbol construction of type (1.7). Thus in this case we obtain a family of binary lattices which combines the advantages of the two constructions: as in [79] we have optimal bounds, and as a Legendre symbol construction it can be implemented fast and easily.

Some authors gave further constructions of binary sequences and lattices with strong pseudorandom properties (see my survey paper [46]). The constructions based on elliptic curves are especially important, see e.g. the papers of Mérai [82], [83], [84], [85], Chen [16], Chen, Li and Xiao [17] and Liu, Zhan and Wang [72].

In Sections 2, 3, 4, 5, 6 I present results from my papers [39], [49], [41], [45], [61], [62] (three of them is written jointly with my coauthors).

In Section 8 I present a short summary of 18 papers which I have been written on the theory of pseudorandomness since my PhD.

2 Pseudorandom sequences constructed by the power generator

One of the most studied pseudorandom generator is Blum-Blum-Shub, called this way after the name of its creators: Leonore Blum, Manuel Blum and Michael Shub [8]. The unpredictability of this generator has been proved assuming the difficulty of integer factorization. In this section I prove quantitative results by estimating the pseudorandom properties of the generated sequences.

Leonore Blum, Manuel Blum and Michael Shub [8] defined the *power generator* by the following:

Let $k \geq 2, m \geq 1$ and ϑ be integers such that $(\vartheta, m) = 1$. Define the sequence $\{u_n\}$ by the recurrence relation

$$u_n \equiv u_{n-1}^k \pmod{m}, \quad 0 \leq u_n \leq m-1, \quad n = 1, 2, \ldots \tag{2.1}$$

with the initial value $u_0 = \vartheta$.

The power generator has many applications in cryptography, see [8], [19], [70], [101]. In the two special cases $(k, \varphi(m)) = 1$ (where $\varphi(m)$ is the Euler function) and $k = 2$ this sequence is known as the *RSA generator* and as the *Blum-Blum-Shub generator*, respectively.

Although various properties of the power generator have been studied in a number of papers, see [8], [11], [18], [19], [23], [32], [63], [70], [81], [101], few unconditional results are known: Clearly, the sequence (2.1) becomes periodic, possible values of the period are studied in [27]. Cusick [18] proved that the rightmost bit of the Blum-Blum-Shub generator assumes values 0 and 1 almost equally often, provided that the period is large enough. Friedlander, Lieman and Shparlinski [26], proved that if the period of the RSA generator is large enough, then the elements of the sequence is uniformly distributed modulo m and a positive proportion of the rightmost and leftmost bits is uniformly distributed. Lower bounds on the linear complexity of the power generator have been given in [32], [100]. The results of this section will be also unconditional.

Notation 2.1 *Let p be a prime, $\vartheta \in \mathbb{F}_p^*$ be an element. Define the sequence u_n by* (2.1) *with a prime modulus p in place of m (then the value of u_n is fixed in the interval $[0, p-1]$). Clearly the multiplicative order of $u_n \equiv \vartheta^{k^n} \pmod{p}$ is non-increasing as $n \to \infty$. Let n_0 denote the smallest positive integer such that for $n \geq n_0$ the multiplicative order of*

$$u_n \equiv \vartheta^{k^n} \pmod{p}$$

is the same number: t. Then

$$(k, t) = 1. \tag{2.2}$$

Denote by T the multiplicative order of k modulo t.

Throughout the section we will use these notations: $p, \vartheta, t, k, T, n_0$ and the sequence $\{u_n\}$ will be as it described here. Clearly the sequence

$$u_{n_0}, u_{n_0+1}, u_{n_0+2}, \dots$$

is purely periodic with the period T.

We convert the sequence $\{u_n\}$ to a binary sequence by the parity of its last bit:

Construction 2.1 (Blum, Blum, Shub) *Define the sequence $E_N = (e_1, \dots, e_N)$ by*

$$e_n = \begin{cases} +1 & \text{if } u_n \text{ is even,} \\ -1 & \text{if } u_n \text{ is odd.} \end{cases} \tag{2.3}$$

In this section we will study the pseudorandom properties of the sequence E_N. First we will give upper bounds for the well-distribution measure and the normality measure of order ℓ. In Theorems 2.1 and 2.2 the length of the sequence is T (defined in Notation 2.1), which is the period of the power generator.

Theorem 2.1

$$W(E_T) \ll p^{7/8} (\log p)^2.$$

For the normality measure we have

Theorem 2.2 *For all $\varepsilon > 1/4$ we have*

$$N_\ell(E_T) \ll k^{\varepsilon(\ell-1)} p^{7/8} (\log p)^{\ell+1},$$

where the implied constant depends only on ε.

The proof of Theorems 2.1 and 2.2 will be based on extensions of theorems of Friedlander, Hansen and Shparlinski in [25] and [28].

Until very recently only the short-range correlation ($\sum_n e_{n+d_1} e_{n+d_2} \dots e_{n+d_\ell}$ for small d_i's) could be handled. By using Bourgain [9] new result, we will be able to handle the long-range correlation as well, which was out of reach until now. Thus here all the three pseudorandom

measures of the power generator are studied, and this *unconditionally* proves that the pseudorandom generator has strong pseudorandom properties.

We will estimate the correlation measure E_N defined by (2.3) for some $N < T$, so the length of the sequence will be smaller than the period of the power generator following from certain technical conditions in Bourgain [9] theorem. The exact value of the length N is defined in Theorem 2.3.

Theorem 2.3 *Suppose that $\ell^2 < p$. Denote by $N = N(\vartheta, k, \delta)$ the largest positive integer such that for all $1 \le i < j \le 2N$ we have*

$$(k^j - k^i, t) \le tp^{-\delta}. \tag{2.4}$$

Then there exists a constant $\varepsilon(\ell, \delta) = \varepsilon > 0$ depending on ℓ and δ such that for the sequence E_N of length N defined by (2.3) we have

$$C_\ell(E_N) \le p^{1-\varepsilon}. \tag{2.5}$$

The proof will be based on a recent result of Bourgain [9]. The upper bound (2.5) for the correlation measure is non-trivial if N, the length of the sequence (defined by (2.4)) is large. The following corollary studies a simple case when N is indeed large.

Corollary 2.1 *Let $p-1 = 2q$, where p and q are odd primes, ϑ be primitive root modulo p, and k be primitive root modulo q. Then for the sequence $E_{(p-3)/4}$ of length $(p-3)/4$ defined by (2.3) we have*

$$C_\ell(E_{(p-3)/4}) \le p^{1-\varepsilon},$$

where the constant $\varepsilon > 0$ depends only on ℓ.

We remark that (2.3) is not the only way to define a binary sequence $\{e_n\}$ from the sequence $\{u_n\}$. For example, Theorems 2.1, 2.2, 2.3 also hold for the sequence $E_N = (e_1, \dots, e_N)$ defined by

$$e_n = \begin{cases} +1 & \text{if } 0 \le u_n < p/2, \\ -1 & \text{if } p/2 \le u_n < p. \end{cases}$$

In Section 2.1 we will estimate certain related exponential sums and the proofs of Theorems 2.1, 2.2 and 2.3 will be completed in Section 2.2.

In this section we study the prime modulus case, i.e., $u_n \equiv u_{n-1}^k \pmod p$, where p is a prime. These results could be extended to the composite modulus case by using exponential sum estimates from [28]. Here I do not carry out the proof, since the computations would be similar but more complicated.

However, it may happen that the power generator has stronger pseudorandom properties in the prime modulus case than in the composite modulus case. This situation indeed happens for the Jacobi symbol sequence

$$E_m = \left(\left(\frac{f(1)}{m}\right), \left(\frac{f(2)}{m}\right), \ldots, \left(\frac{f(m)}{m}\right)\right), \qquad f(x) \in \mathbb{Z}_m[x].$$

Goubin, Mauduit and Sárközy [31] proved that under certain conditions on the polynomial $f(x)$, this sequence has strong pseudorandom properties if m is a prime: $W(E_m), C_\ell(E_m) \ll m^{1/2} \log m$. If m is a product of two different odd primes, then Rivat and Sárközy [94] proved that for all polynomial $f(x) \in \mathbb{Z}_m[x]$ we have $C_4(E_m) \gg m$. The situation is very similar in case of some other constructions, see e.g. the paper of Liu, Zhan and Wang [73].

Throughout the section we write $e_p(a) = \exp(2\pi i \frac{a}{p})$.

2.1 Exponential sums

J. Friedlander, J. Hansen and I. Shparlinski gave an upper bound for the sum $\sum_{x=1}^{T} e_p(a\vartheta^{k^x})$. Later Friedlander and Shparlinski [28] extended this result to the sum $\sum_{x=1}^{T} e_p(a_1\vartheta^{k^x} + a_2\vartheta^{k^{x+1}} \cdots + a_r\vartheta^{k^{x+r-1}})$. Here we will study the extension this result to general powers and incomplete sums. First we will study the incomplete sum analog of the result in [28].

Lemma 2.1 *Let t, T be as in Notation 2.1. Let $\varepsilon_1 > 1/4$ and suppose that $t > p^{1/2+\delta}$ for a constant $\delta > 0$. Let $a_i \in \mathbb{F}_p$, $L, M \in \mathbb{N}$ with $L \le T$. Then*

$$\left| \sum_{x=M+1}^{M+L} e_p(a_1\vartheta^{k^x} + a_2\vartheta^{k^{x+1}} + \cdots + a_r\vartheta^{k^{x+r-1}}) \right| \ll k^{\varepsilon_1(r-1)} T^{1/4} t^{1/2} p^{1/8} \log p,$$

where the implied constant depends only on δ and ε_1. In the special case $r = 1$ we obtain

$$\left| \sum_{x=M+1}^{M+L} e_p(a_1\vartheta^{k^x}) \right| \ll T^{1/4} t^{1/2} p^{1/8} \log p,$$

where the implied constant depends only on δ.

Using J. Bourgain's result [9], we will prove:

Lemma 2.2 *For $1 \le i \le r$ let $h_i \in \mathbb{Z}_{p-1}$, $\vartheta_i = \vartheta^{h_i}$ and $a_i \in \mathbb{F}_p^*$ where $(h_1, \ldots, h_r, p-1) = 1$ also holds. Then the sequence*

$$\{a_1\vartheta_1^{k^x} + \cdots + a_r\vartheta_r^{k^x}\}$$

becomes periodic with period T (where T is defined in Notation 2.1). Denote by

$$N(\vartheta_1, \dots, \vartheta_r, k, \delta) = N$$

the largest positive integer N such that $N \le T$, for all $0 \le i \le N$, $1 \le j \le r$

$$(k^i h_j, t) \le t p^{-\delta}, \tag{2.6}$$

and for all pairs $\{i_1, j_1\}, \{i_2, j_2\}$ with $1 \le i_1, i_2 \le N$, $1 \le j_1 \le j_2 \le r$ we have

$$(k^{i_1} h_{j_1} - k^{i_2} h_{j_2}, t) \le t p^{-\delta} \quad \text{or} \quad (k^{i_1} h_{j_1} - k^{i_2} h_{j_2}, t) = t. \tag{2.7}$$

If there is no such N define $N(\vartheta_1, \dots, \vartheta_r, k, \delta) = N$ by 1.

Let $L, M \in \mathbb{N}$ with $L \le T$. Then there exists a constant $\varepsilon(r, \delta) = \varepsilon_2 \ge 0$ depending on only r (the number of ϑ_i's) and δ such that:

$$\left| \sum_{x=M}^{M+L} e_p(a_1 \vartheta_1^{k^x} + \cdots + a_r \vartheta_r^{k^x}) \right| \ll (tT)^{1/2} \left(p^{-\varepsilon_2} + \frac{(r+1)^{r/2}}{N^{1/2}} \right) \log p.$$

Moreover, in the special case $(h_1, t) = 1$ we may replace the term $(r+1)^{r/2}$ by $(r+1)^{1/2}$:

$$\left| \sum_{x=M}^{M+L} e_p(a_1 \vartheta_1^{k^x} + \cdots + a_r \vartheta_r^{k^x}) \right| \ll (tT)^{1/2} \left(p^{-\varepsilon_2} + \frac{(r+1)^{1/2}}{N^{1/2}} \right) \log p,$$

where the implied constant factors are absolute.

Proof of Lemma 2.1 and Lemma 2.2

We will use the following deep theorem of Bourgain [9]:

Lemma 2.3 (Bourgain) *Let p be a prime. Given $r \in \mathbb{Z}^+$ and $\delta > 0$, there is an $\varepsilon = \varepsilon(r, \delta) > 0$ satisfying the following property: If*

$$f(x) = a_1 x^{k_1} + \cdots + a_r x^{k_r} \in \mathbb{Z}[x] \quad \text{and } (a_i, p) = 1$$

where the exponents $1 \le k_i \le p-1$ satisfy

$$\begin{aligned} (k_i, p-1) &< p^{1-\delta} \text{ for all } 1 \le i \le r \\ (k_i - k_j, p-1) &< p^{1-\delta} \text{ for all } 1 \le i \ne j \le r \end{aligned} \tag{2.8}$$

then

$$\left| \sum_{x=1}^{p-1} e_p(f(x)) \right| < p^{1-\varepsilon}.$$

Proof of Lemma 2.3

See in [9].

In order to prove Lemma 2.1 and Lemma 2.2 first we need estimates for complete sums.

First we give an upper bound for n_0 defined in Notation 2.1. Let ord ϑ denote the multiplicative order of ϑ modulo p. n_0 is the smallest integer for which $(k^{n_0}, \text{ord } \vartheta)$ is maximal. From this

$$n_0 \leq \frac{\log \text{ord } \vartheta}{\log 2} < 1.45 \log p. \tag{2.9}$$

We will deduce the first two statements of Lemma 2.4 from Bourgain's theorem (Lemma 2.3), while the third part will be proved by extending an argument of Friedlander and Shparlinski [28].

Lemma 2.4 *Let $\vartheta_1, \ldots, \vartheta_r \in \mathbb{F}_p$ and $N(\vartheta_1, \ldots, \vartheta_r, k, \delta) = N$ as in Lemma 2.2, $j \in \mathbb{Z}_T$. Then there exists a constant $\varepsilon(r, \delta) = \varepsilon_2 \geq 0$ depending on only r and δ such that:*

$$\left| \sum_{x=n_0}^{n_0-1+T} e_p(a_1 \vartheta_1^{k^x} + \cdots + a_r \vartheta_r^{k^x}) e_T(jx) \right| \ll (tT)^{1/2} \left(p^{-\varepsilon_2} + \frac{(r+1)^{r/2}}{N^{1/2}} \right). \tag{2.10}$$

If $(h_1, t) = 1$ (where h_1 is defined by $\vartheta_1 \equiv \vartheta^{h_1} \pmod p$), then we may replace the term $(r+1)^{r/2}$ by $(r+1)^{1/2}$:

$$\left| \sum_{x=n_0}^{n_0-1+T} e_p(a_1 \vartheta_1^{k^x} + \cdots + a_r \vartheta_r^{k^x}) e_T(jx) \right| \ll (tT)^{1/2} \left(p^{-\varepsilon_2} + \frac{(r+1)^{1/2}}{N^{1/2}} \right), \tag{2.11}$$

where the implied constants are absolute.

If $\vartheta_i = \vartheta^{k^i}$ for $1 \leq i \leq r$ then there exists an upper bound, where the exponent of p is given: Suppose that $\varepsilon_1 > 1/4$ and $t > p^{1/2+\delta}$ for a constant $\delta > 0$, then

$$\left| \sum_{x=n_0}^{n_0-1+T} e_p(a_1 \vartheta^{k^x} + a_2 \vartheta^{k^{x+1}} + \cdots + a_r \vartheta^{k^{x+r-1}}) e_T(jx) \right| \ll k^{\varepsilon_1(r-1)} T^{1/4} t^{1/2} p^{1/8}, \tag{2.12}$$

where the implied constant depends only on ε_1 and δ.

Proof of Lemma 2.4

The proof is similar to the proof of Theorem 8 in [25] in the special case $\nu = 1$, but in order to prove (2.10) and (2.11) we use Bourgain's theorem in place of Weil's theorem.

Let $S = \left|\sum_{x=n_0}^{n_0-1+T} e_p(a_1\vartheta_1^{k^x} + \cdots + a_r\vartheta_r^{k^x})e_T(jx)\right|$ and $\mathcal{K} \subseteq \{k^1, \ldots, k^T\}$. For $y = k^v \in \mathcal{K}$ denote v by $\mathrm{ind}_k y$. Clearly,

$$S = \frac{1}{|\mathcal{K}|}\left|\sum_{y\in\mathcal{K}}\sum_{x=n_0}^{n_0-1+T} e_p(a_1\vartheta_1^{yk^x} + \cdots + a_r\vartheta_r^{yk^x})e_T(j(x+\mathrm{ind}_k y))\right|.$$

By the Cauchy-Schwartz inequality we have

$$S \le \frac{T^{1/2}}{|\mathcal{K}|}\left(\sum_{x=n_0}^{n_0-1+T}\left|\sum_{y\in\mathcal{K}} e_p(a_1\vartheta_1^{yk^x} + \cdots + a_r\vartheta_r^{yk^x})e_T(j\mathrm{ind}_k y)\right|^2\right)^{1/2}.$$

We recall that $\vartheta_i \equiv \vartheta^{h_i} \pmod p$, where $(h_1, \ldots, h_r, p-1) = 1$. Let $d = (p-1)/t$. Since the order of ϑ^{k^x} is t for $n_0 \le x$, for each of these powers ϑ^{k^x}, there exist precisely d values of $z \in \mathbb{F}_p^*$ such that $\vartheta^{k^x} \equiv z^d \pmod p$. Thus

$$\begin{aligned} S &\le \frac{T^{1/2}}{|\mathcal{K}|\,d^{1/2}}\left(\sum_{z=1}^{p-1}\left|\sum_{y\in\mathcal{K}} e_p(a_1 z^{yh_1 d} + \cdots + a_r z^{yh_r d})e_T(j\mathrm{ind}_k y)\right|^2\right)^{1/2} \\ &\le \frac{T^{1/2}}{|\mathcal{K}|\,d^{1/2}}\left(\sum_{y\in\mathcal{K}}\sum_{x\in\mathcal{K}}\Big|\sum_{z=1}^{p-1} e_p(a_1 z^{yh_1 d} + \cdots + a_r z^{yh_r d} - \right. \\ &\left. - a_1 z^{xh_1 d} - \cdots - a_r z^{xh_r d})\Big|\right)^{1/2}. \end{aligned}$$

For given $y, x \in \mathcal{K}$ define the polynomial $g_{y,x}(z) \in \mathbb{F}_p[z]$ by

$$g_{y,x}(z) \stackrel{\mathrm{def}}{=} a_1 z^{yh_1 d} + \cdots + a_r z^{yh_r d} - a_1 z^{xh_1 d} - \cdots - a_r z^{xh_r d}.$$

Denote by $g_{y,x}(z) \equiv c$ that the polynomial $g_{y,x}(z) \in \mathbb{F}_p[z]$ is identically constant. Then

$$S \le \frac{T^{1/2}}{|\mathcal{K}|\,d^{1/2}}\left(\sum_{x\in\mathcal{K}}\sum_{y\in\mathcal{K}}\left|\sum_{z=1}^{p-1} e_p(g_{y,x}(z))\right|\right)^{1/2},$$

$$S \le \frac{T^{1/2}}{|\mathcal{K}|\, d^{1/2}} \left(\sum_{\substack{x,y \in \mathcal{K} \\ g_{y,x}(z) \not\equiv c}} \left| \sum_{z=1}^{p-1} e_p(g_{y,x}(z)) \right| + \sum_{\substack{x,y \in \mathcal{K} \\ g_{y,x}(z) \equiv c}} p \right)^{1/2}. \tag{2.13}$$

Next we estimate the number of the pairs $y, x \in \mathcal{K}$ with $g_{y,x}(z) \equiv c$. Clearly, then apart from the multiplicity, the set $\{yh_1 d, \dots, yh_r d\} \setminus \{0\}$, contains the same residue classes modulo $p-1$ as the set $\{xh_1 d, \dots, xh_r d\} \setminus \{0\}$. So the set $\{yh_1, \dots, yh_r\} \setminus \{0\}$ contains the same residue classes modulo t as the set $\{xh_1, \dots, xh_r\} \setminus \{0\}$. We will use the following lemma.

Lemma 2.5 *For given $x \in \mathcal{K}$ at most $(r+1)^r$ pieces of $y \in \mathcal{K}$ exist such that the sets $\{xh_1, \dots, xh_r\} \setminus \{0\}$, $\{yh_1, \dots, yh_r\} \setminus \{0\}$ contain the same residue classes modulo t apart from the multiplicity. If $(h_1, t) = 1$ then at most $r+1$ pieces of $y \in \mathcal{K}$ exist with this property.*

Proof of Lemma 2.5 Define h_{r+1} by 0. Then for every $1 \le i \le r$ there exists a $1 \le j(i) \le r+1$ such that

$$yh_i \equiv xh_{j(i)} \pmod{t}. \tag{2.14}$$

This congruence determines y uniquely modulo $\frac{t}{(t,h_i)}$. As i runs through the numbers $1, 2, \dots, r$, by the Chinese Remainder Theorem we get that y is uniquely determined modulo $\frac{t}{(t,h_1,\dots,h_r)} = t$ (since $(h_1, \dots, h_r, p-1) = 1$). In the special case $(h_1, t) = 1$ the first congruence $yh_1 \equiv xh_{j(1)} \pmod{t}$ determines y uniquely. The elements of $\mathcal{K}$ are distinct modulo t, thus if the congruences in (2.14) are given, then at most one $y \in \mathcal{K}$ exists with the desired property. Since each $j(i)$ may take $r+1$ different values, from this follows the lemma.

We return to the proof of Lemma 2.4. Define the constant $c(r)$ by

$$c(r) = \begin{cases} r+1 & \text{if } (h_1, t) = 1, \\ (r+1)^r & \text{otherwise.} \end{cases} \tag{2.15}$$

By Lemma 2.5, for fixed $x \in \mathcal{K}$ at most $c(r)$ pieces of y exist with $g_{y,x}(z) \equiv c$. $x \in \mathcal{K}$ may take $|\mathcal{K}|$ different values, thus at most $c(r)\,|\mathcal{K}|$ pairs (y, x) exists such that $g_{y,x}(z) \equiv c$. By this and (2.13) we get

$$S \le \frac{T^{1/2}}{|\mathcal{K}|\, d^{1/2}} \left(\sum_{\substack{x,y \in \mathcal{K} \\ g_{y,x}(z) \not\equiv c}} \left| \sum_{z=1}^{p-1} e_p(g_{y,x}(z)) \right| + c(r)\,|\mathcal{K}|\, p \right)^{1/2}.$$

Let Q

$$Q \stackrel{\text{def}}{=} \max_{\substack{x,y\in\mathcal{K}\\ g_{y,x}(z)\not\equiv c}} \left|\sum_{z=1}^{p-1} e_p(g_{y,x}(z))\right|.$$

Then

$$S \le \frac{T^{1/2}}{|\mathcal{K}|\, d^{1/2}} \left(|\mathcal{K}|^2\, Q + c(r)\,|\mathcal{K}|\, p\right)^{1/2} \le \left(\frac{TQ}{d}\right)^{1/2} + \left(c(r)\frac{Tp}{|\mathcal{K}|\, d}\right)^{1/2}. \quad (2.16)$$

In order to prove (2.10) and (2.11) we choose $\mathcal{K} = \{k^1, \dots, k^N\}$ with $N = N(\vartheta_1, \dots, \vartheta_r, k, \delta)$. Then $|\mathcal{K}| = N$. For $x, y \in \mathcal{K}$ and $\frac{p-1}{t} = d$ by (2.6) we have

$$(dxh_j, p-1) = d(xh_j, t) \le dtp^{-\delta} < p^{1-\delta}. \quad (2.17)$$

Clearly (2.17) also holds with y in place of x. Similarly, by (2.7)

$$(dxh_{j_1} - dyh_{j_2}, p-1) = d(xh_{j_1} - yh_{j_2}, t) \begin{cases} \le dtp^{-\delta} < p^{1-\delta} & \text{or} \\ = dt = p-1. \end{cases}$$

Thus (2.8) holds for the polynomial $g_{y,x}(z) \in \mathbb{F}_p[z]$ and we may use Lemma 2.3 since $g_{y,x}(z) \not\equiv c$. Then

$$Q \le p^{1-\varepsilon_2}.$$

By this, (2.15), (2.16), $t = \frac{p-1}{d}$ and $|\mathcal{K}| = N$ we get:

$$S \ll \left(\frac{Tp^{1-\varepsilon_2}}{d}\right)^{1/2} + \left(c(r)\frac{Tp}{Nd}\right)^{1/2} \ll (Tt)^{1/2}(p^{-\varepsilon_2} + c(r)^{1/2}N^{-1/2})$$

which proves (2.10) and (2.11) in Lemma 2.4.

In order to get (2.12) we recall the proof of Friedlander and Shparlinski [28]. Consider the special case $h_i = k^{i-1}$ for $1 \le i \le r$. In order to estimate Q in this special case we need Weil's theorem for character sums, which we present in the following form:

Lemma 2.6 (Weil) *For any prime p, and any polynomial $f(x) \in \mathbb{F}_p[x]$ of degree $D \ge 1$ which is not identically constant, the bound*

$$\left|\sum_{x=1}^{p} e_p(f(x))\right| \le Dp^{1/2}.$$

holds.

Proof of Lemma 2.6

This lemma can be deduced from Weil's theorem. See [106], an elementary proof can be found in [99].

We will also need the following lemma of Friedlander, Hansen and Shparlinski [25]:

Lemma 2.7 (Friedlander, Hansen, Shparlinski) *For any set $\mathcal{W} \subseteq \mathbb{Z}_t^*$ of cardinality $|\mathcal{W}| = W$, any fixed $\delta > 0$ and any integer $h \ge t^\delta$, there exists an integer $a \in \mathbb{Z}_t^*$, such that the congruence*

$$ak \equiv b \pmod t, \quad k \in \mathcal{W},\ 0 \le b \le h-1 \tag{2.18}$$

has

$$L_a(h) \gg \frac{Wh}{t}$$

solutions.

Proof of Lemma 2.7

This is Lemma 2 in [25].

We return to the proof of (2.12) in Lemma 2.4. Let $\varepsilon_1 > 1/4$. If $k^{\varepsilon_1(r-1)} > \frac{T^{3/4}}{t^{1/2}p^{1/8}}$, then using the trivial estimate we obtain $S \le T \le k^{\varepsilon_1(r-1)}t^{1/2}T^{1/4}p^{1/8}$ which was to be proved. Thus we may suppose

$$k^{(r-1)/2} \le \frac{T^{3/(8\varepsilon_1)}}{t^{1/(4\varepsilon_1)}p^{1/(16\varepsilon_1)}}. \tag{2.19}$$

Set

$$h = \left[\frac{(r+1)^{1/2}t}{T^{1/2}k^{(r-1)/2}p^{1/4}}\right]. \tag{2.20}$$

Then by (2.19), $T \le t$ and $t > p^{1/2+\delta}$ we have

$$\begin{aligned} h &\gg \frac{t}{T^{1/2}\frac{T^{3/(8\varepsilon_1)}}{t^{1/(4\varepsilon_1)}p^{1/(16\varepsilon_1)}}p^{1/4}} = \frac{t^{1+1/(4\varepsilon_1)}}{T^{1/2+3/(8\varepsilon_1)}p^{1/4-1/(16\varepsilon_1)}} \gg \frac{t^{1/2-1/(8\varepsilon_1)}}{p^{1/4-1/(16\varepsilon_1)}} \\ &= \left(\frac{t}{p^{1/2}}\right)^{1/2-1/(8\varepsilon_1)} \gg t^{\frac{2\delta}{1+2\delta}(1/2-1/(8\varepsilon_1))}, \end{aligned}$$

thus we may use Lemma 2.7. Let $\mathcal{W} = \{k^1, \dots, k^T\}$. We select a as in Lemma 2.2. Let now $\mathcal{K}$ denote the subset of $\mathcal{W}$ which satisfies the corresponding congruence (2.18). Then the degree of the polynomial $g_{y,x}(z^a)$ is less than

$hk^{r-1}d$. By this and Lemma 2.6 we have

$$Q = \max_{\substack{x,y\in\mathcal{K}\\ g_{y,x}(z)\equiv 0}} \left|\sum_{z=1}^{p-1} e_p(g_{x,y}(z))\right| = \max_{\substack{x,y\in\mathcal{K}\\ g_{y,x}(z)\equiv 0}} \left|\sum_{z=1}^{p-1} e_p(g_{x,y}(z^a))\right| \le hk^{r-1}dp^{1/2}. \tag{2.21}$$

By Lemma 2.7

$$|\mathcal{K}| \gg \frac{Th}{t}. \tag{2.22}$$

By (2.2) and (2.15) we have $c(r) = r + 1$. By this, (2.16), (2.20), (2.21), (2.22) and $t = \frac{p-1}{d}$ we get

$$\begin{aligned} S &\le \left(\frac{Thk^{r-1}dp^{1/2}}{d}\right)^{1/2} + \left(\frac{c(r)Tp}{|\mathcal{K}|\,d}\right)^{1/2} \\ &\le \left(Tk^{r-1}hp^{1/2}\right)^{1/2} + \left(\frac{(r+1)Tt}{|\mathcal{K}|}\right)^{1/2} \\ &\le \left(Tk^{r-1}hp^{1/2}\right)^{1/2} + \left(\frac{(r+1)t^2}{h}\right)^{1/2} \\ &\ll \left((r+1)k^{r-1}Tt^2p^{1/2}\right)^{1/4}, \end{aligned}$$

which was to be proved.

We return to the proof of Lemma 2.1 and Lemma 2.2. Let

$$S = \left|\sum_{x=M}^{M+L} e_p(a_1\vartheta_1^{k^x} + \cdots + a_r\vartheta_r^{k^x})\right|.$$

We will suppose $M \ge n_0$, since by (2.9) the contribution of the terms of $M \le x \le n_0$ in S is small, at most $n_0 \le 1.45\log p$. Using

$$\sum_{j=1}^{T} e_T(nj) = \begin{cases} T & \text{if } T \mid n, \\ 0 & \text{otherwise,} \end{cases}$$

we get

$$S = \frac{1}{T}\left|\sum_{y=n_0}^{n_0-1+T} e_p(a_1\vartheta_1^{k^y} + \cdots + a_r\vartheta_r^{k^y}) \sum_{x=M}^{M+L}\sum_{j=1}^{T} e_T((y-x)j)\right|$$
$$= \frac{1}{T}\sum_{j=1}^{T}\left|\sum_{x=M}^{M+L} e_T(-jx)\right|\left|\sum_{y=n_0}^{n_0-1+T} e_p(a_1\vartheta_1^{k^y} + \cdots + a_r\vartheta_r^{k^y})e_T(jy)\right|. \quad (2.23)$$

Let

$$Q = \max_j \left|\sum_{y=n_0}^{n_0-1+T} e_p(a_1\vartheta_1^{k^y} + \cdots + a_r\vartheta_r^{k^y})e_T(jy)\right|.$$

By (2.23) we have

$$S \le \frac{1}{T}\sum_{j=1}^{T}\left|\sum_{x=M}^{M+L} e_T(-jx)\right| Q. \quad (2.24)$$

By Lemma 2.4 there exists a constant $\varepsilon_2 > 0$ depending only on r and δ such that

$$Q \ll (tT)^{1/2}(p^{-\varepsilon_2} + (c(r))^{1/2}N^{-1/2}), \quad (2.25)$$

where the constant $c(r)$ is defined by (2.15). Moreover in the special case $\vartheta_i = \vartheta^{k^{i-1}}$ for $1 \le i \le r$ we get that for every $\varepsilon_1 > 1/4$

$$Q \ll k^{\varepsilon_1(r-1)}t^{1/2}T^{1/4}p^{1/8} \quad (2.26)$$

also holds.

By the sum of geometric progression, the triangle-inequality and $|1 - e(x)| \ge 4 \parallel x \parallel$ we have

$$\sum_{j=1}^{T}\left|\sum_{x=0}^{L} e_T(-jx)\right| \le \sum_{j=1}^{T}\frac{2}{|1-e(j/T)|} \le \frac{1}{2}\sum_{j=1}^{T}\frac{1}{\parallel j/T \parallel} \le \sum_{j=1}^{[(T+1)/2]}\frac{1}{\parallel j/T \parallel}$$
$$= \sum_{j=1}^{[(T+1)/2]}\frac{j}{T} \ll T\log T. \quad (2.27)$$

By (2.24), (2.25), (2.26) and (2.27) we get the statements of Lemma 2.1 and Lemma 2.2.

Remark 2.1 *In fact, using the results of Friedlander, Hansen and Shparlinski [25], the following can be proved: if $t > p^{1/2+\delta}$ for all integer $\nu \ge 1$ we*

have:

$$\left|\sum_{x=M}^{M+L} e_p(a_1\vartheta^{k^x}+a_2\vartheta^{k^{x+1}}+\cdots+a_r\vartheta^{k^{x+r-1}})\right| \ll T^{1-\frac{2\nu+1}{2\nu(\nu+1)}}t^{\frac{1}{2\nu}}p^{\frac{1}{4(\nu+1)}}\log T.$$

Here, we presented the proof only in the special case $\nu=1$.

2.2 Proofs of Theorems 2.1-2.3

In order to express the terms of the sequence E_N we will use additive characters as in [76]. We will use the following representation:

Lemma 2.8 (Mauduit, Rivat, Sárközy) *For* $n\in\mathbb{N}$ $r_p(n)$ *denotes the unique* $r\in\{0,\dots,p-1\}$ *for which* $n\equiv r \pmod p$. *Then for odd integer* p, *there exists a function* $\nu_p(a,x):\mathbb{Z}\times\mathbb{Z}\to\mathbb{C}$ *such that*

$$\frac{1}{p}\sum_{|a|<p/2}\nu_p(a,x)e_p(an)=\begin{cases}+1 & \text{if } r_p(n)\equiv x \pmod 2,\\ 0 & \text{if } r_p(n)\not\equiv x \pmod 2,\end{cases}$$

and the function $\nu_p(a,x)$ *satisfies*

$$\nu_p(0,x)=\begin{cases}\frac{p+1}{2} & \text{if } x\equiv 0 \pmod 2,\\ \frac{p-1}{2} & \text{if } x\equiv 1 \pmod 2.\end{cases} \tag{2.28}$$

Furthermore, for $1\le|a|<p/2$ *we have*

$$|\nu_p(a,x)|\ll\frac{p}{\min\{a,p-2a\}}. \tag{2.29}$$

Proof of Lemma 2.8

Since for $r\in\mathbb{Z}$, we have

$$\frac{1}{p}\sum_{|a|<p/2}e_p(a(n-r))=\begin{cases}1 & \text{if } n\equiv r \pmod p,\\ 0 & \text{otherwise,}\end{cases}$$

for $0\le n\le p-1$ we have

$$\frac{1}{p}\sum_{|a|<p/2}\left(\sum_{\substack{r\equiv x \pmod 2,\\ 0\le r\le p-1}}e_p(-ar)\right)e_p(an)=\begin{cases}1 & \text{if } n\equiv x \pmod 2,\\ 0 & \text{otherwise.}\end{cases}$$

Thus we may define $\nu_p(a,x)$ by

$$\nu_p(a,x) \stackrel{\text{def}}{=} \sum_{\substack{r\equiv x \pmod 2,\\ 0\le r\le p-1}} e_p(-ar).$$

From this immediately follows (2.28). By computing the geometric sum above, using the triangle inequality and $|1-e(x)| \ge 4 \parallel x \parallel$ we get (2.29).

Writing $\nu(a) = \nu(a,0) - \nu(a,1)$ from Lemma 2.8 we get immediately:

Lemma 2.9 (Mauduit, Rivat, Sárközy) *For $0 \le n \le p-1$ and an odd integer p, we have*

$$\frac{1}{p}\sum_{|a|<p/2} \nu_p(a)e_p(an) = \begin{cases} +1 & \text{if } r_p(n) \equiv 0 \pmod 2,\\ -1 & \text{if } r_p(n) \equiv 1 \pmod 2, \end{cases}$$

where the function $\nu_p(a)$ satisfies

$$\nu_p(0) = 1, \quad |\nu_p(a)| \ll \frac{p}{\min\{a, p-2a\}} \quad (1 \le |a| < p/2).$$

Proof of Theorem 2.1

If $t \le p^{7/8}$ Theorem 2.1 and 2.2 are trivial, since all pseudorandom measures of E_T are less or equal than $T \le t \le p^{7/8}$. Thus we may suppose that

$$t > p^{7/8}. \tag{2.30}$$

We have to prove that for any $0 \le b < p$, $0 \le c < b$, $1 \le M < T$, we have the estimate

$$\left|\sum_{\substack{j\\ c+jb\le M}} e_{c+jb}\right| \ll p^{7/8}(\log p)^2.$$

By Lemma 2.9 we have

$$\left|\sum_{\substack{j\\ c+jb\le M}} e_{c+jb}\right| = \frac{1}{p}\sum_{|a|<p/2} \nu_p(a) \sum_{\substack{j\\ c+jb\le M}} e_p(au_{c+jb})$$

Since $u_{c+jb} \equiv \left(\vartheta^{(k^c)}\right)^{(k^b)^j} \pmod{p}$, the multiplicative order of k^b modulo t is larger or equal than T/b and by (2.30) we may use Lemma 2.1 and obtain

$$\left|\sum_{\substack{j \\ c+jb}} e_p(au_{c+jb})\right| \ll T^{1/4}t^{1/2}p^{1/8} \ll p^{7/8}\log p.$$

Thus

$$\left|\sum_{\substack{x \\ r+xm\le M}} e_{r+xm}\right| \ll \frac{1}{p}\left(\sum_{1\le|a|<p/2}|\nu_a(p)|\right)p^{7/8}\log p + |\nu_p(0)|\,, \tag{2.31}$$

By Lemma 2.9 $\nu_p(0) = 1$ and $\sum_{1\le|a|<p/2}|\nu_a(p)| \ll \sum_{1\le|a|<p/4+1}\frac{p}{a} \ll p\log p$, so the theorem follows from this and (2.31).

Proof of Theorem 2.2

By Lemma 2.8 for $M \le T-\ell+1$ we have

$$Z(E_T, M, X) = \frac{1}{p^\ell}\sum_{|a_1|<p/2}\cdots\sum_{|a_\ell|<p/2}\nu_p(a_1, u_{n+1})\cdots\nu_p(a_\ell, u_{n+\ell})$$
$$\sum_{n<M} e_p(a_1u_{n+1}+\cdots+a_\ell u_{n+\ell}). \tag{2.32}$$

If $(a_1,\dots,a_\ell) = (0,\dots,0)$ then trivially

$$\left|\sum_{n<M} e_p(a_1u_{n+1}+\cdots+a_\ell u_{n+\ell})\right| = M-1. \tag{2.33}$$

By Lemma 2.8 we have

$$\frac{(p-1)^\ell}{2^\ell} \le |\nu_p(0,u_{n+1})\cdots\nu_p(0,u_{n+\ell})| \le \frac{(p+1)^\ell}{2^\ell}. \tag{2.34}$$

By (2.30) we may use Lemma 2.1 and for all $\varepsilon_1 > 1/4$ we have that if $(a_1,\dots,a_\ell) \ne (0,\dots,0)$ then

$$\left|\sum_{n<M} e_p(a_1u_{n+1}+\cdots+a_\ell u_{n+\ell})\right| = \left|\sum_{n<M} e_p(a_1\vartheta^{k^{n+1}}+\cdots+a_\ell\vartheta^{k^{n+\ell}})\right|$$
$$\ll k^{\varepsilon_1(r-1)}T^{1/4}t^{1/2}p^{1/8\log p} \ll k^{\varepsilon_1(r-1)}p^{7/8}\log p, \tag{2.35}$$

where the implied constant depends only on ε_1. By (2.32), (2.33), (2.35) and the triangle inequality we have

$$\left|Z(E_T,M,X)-M/2^\ell\right| \le \frac{1}{p^\ell}\Bigg|\sum_{\substack{(a_1,\dots,a_\ell)\neq(0,\dots,0),\\ |a_i|<p/2\ (1\le i\le \ell)}} \nu_p(a_1,u_{n+1})\cdots\nu_p(a_\ell,u_{n+\ell})$$
$$\sum_{n<M} e_p(a_1u_{n+1}+\cdots+a_\ell u_{n+\ell})\Bigg| + \frac{1}{p^\ell}\left|\frac{(p+1)^\ell}{2^\ell}(M-1)-\frac{M}{2^\ell}\right|.$$

Since $\ell<p$ we have

$$\frac{1}{p^\ell}\left|\frac{(p+1)^\ell}{2^\ell}(M-1)-\frac{M}{2^\ell}\right| \le \left(\frac{(p+1)^\ell}{p^\ell}-1\right)\frac{M}{2^\ell} \le \frac{e\ell M}{p2^\ell} \le \frac{e\ell}{2^\ell} < 1.5.$$

If $(a,p)=1$ let $\mu_p(a)=\frac{p}{\min\{a,p-2a\}}$ and let $\mu_p(0)=\frac{p+1}{2}$. Then by Lemma 2.8 $\nu_p(a,u_{n+i})\le\mu(a)$. By this and (2.34) we have

$$\left|Z(E_T,M,X)-M/2^\ell\right| \ll \frac{1}{p^\ell}\left(\left|\sum_{|a|<p/2}\mu_p(a)\right|^\ell k^{\varepsilon_1(r-1)}p^{7/8}\log p\right)+1.5.$$

Using $\left|\sum_{|a|<p/2}\mu_p(a)\right| \ll \sum_{1\le|a|\le p/4+1}\frac{p}{a}\ll\log p$, we get the theorem.

Proof of Theorem 2.3

Theorem 2.3 is trivial if $N\le p^{1/2}$. Thus we may suppose that

$$N>p^{1/2}. \tag{2.36}$$

By Lemma 2.9 for $M<p$ and $0\le d_1<\cdots<d_\ell\le p-M$ we have

$$\sum_{n\le M} e_{n+d_1}\dots e_{n+d_\ell} = \frac{1}{p^\ell}\sum_{|a_1|<p/2}\cdots\sum_{|a_\ell|<p/2}\nu_p(a_1)\cdots\nu_p(a_\ell)$$
$$\sum_{n<M} e_p(a_1u_{n+d_1}+\cdots+a_\ell u_{n+d_\ell})$$

If $(a_1,\dots,a_\ell)\neq(0,\dots,0)$ we may use Lemma 2.2 with $h_1=k^{d_1},\dots,h_\ell=k^{d_\ell}$. By (2.2) $(h_i,t)=(k,t)=1$, thus we obtain

$$\begin{aligned}&\left|\sum_{n<M}e_p(a_1u_{n+d_1}+\cdots+a_\ell u_{n+d_\ell})\right|\\&=\left|\sum_{n<M}e_p(a_1\left(\vartheta^{k^{d_1}}\right)^{k^n}+\cdots+a_\ell\left(\vartheta^{k^{d_\ell}}\right)^{k^n})\right|\\&\ll(tT)^{1/2}\left(p^{-\varepsilon_2}+\frac{(r+1)^{1/2}}{N^{1/2}}\right)\log p.\end{aligned}$$

By (2.36) and $r^2<p$ we have

$$\left|\sum_{n<M}e_p(a_1u_{n+d_1}+\cdots+a_\ell u_{n+d_\ell})\right|\ll p^{1-\varepsilon_2/2},$$

where the implied constant depends only on ε_2. Thus

$$\left|\sum_{n\le M}e_{n+d_1}\dots e_{n+d_\ell}\right|=\frac{1}{p^\ell}\left(\left(\sum_{|a|<p/2}|\nu_p(a)|\right)^\ell p^{1-\varepsilon_2/2}+M\right).$$

Using $\left|\sum_{|a|<p/2}\nu_p(a)\right|\ll\left|\sum_{|a|<p/4+1}\frac{p}{a}\right|\ll p\log p$, we get

$$C_\ell(E_N)\le c_1p^{1-\varepsilon_2/4},$$

where the constant c_1 depends only on ε_2. From this for large $p>p_0$ follows the theorem, while for small $p\le p_0$ the theorem is trivial with an $\varepsilon>0$ for which $N<p^{1-\varepsilon}$ if $p<p_0$. Such $\varepsilon>0$ exists, since $N<p$.

Proof of Corollary 2.1

Since q is a prime, $t=q$ or $t=2q$. k is a primitive root modulo q, thus for $1\le i<j\le q-1$ we have

$$(k^j-k^i,t)=1\quad\text{or}\quad(k^j-k^i,t)=2$$

which is less than $tp^{-\delta}$ for $\delta<1/2$. Thus (2.4) holds with $N=(p-3)/4$ and using Theorem 3 we get the corollary.

3 On the correlation of binary sequences

Since 1997 numerous papers have been written on the theory of pseudorandomness. In the majority of these papers special sequences are constructed and/or tested for pseudorandomness (see [52] and [97] for references), while, for example in [3], [4], [14], [33], [34], [36], [69], [78] and [102] the measures of pseudorandomness are studied. In [46] I gave a survey paper on the most important results related to these measures.

In [14] Cassaigne, Mauduit and Sárközy compared correlations of different order. They proved the following

Theorem 3.A (Cassaigne, Mauduit, Sárközy) *a) For $k, \ell, N \in \mathbb{N}$, $k \mid \ell$, $E_N \in \{-1,+1\}^N$ we have*

$$C_k(E_N) \le N\left(\frac{(\ell!)^{k/\ell}}{k!}\left(\frac{C_\ell(E_N)}{N}\right)^{k/\ell} + \left(\frac{\ell^2}{N}\right)^{k/\ell}\right).$$

b) If $k, N \in \mathbb{N}$ and $k \le N$, then there is a sequence $E_N \in \{-1,+1\}^N$ such that if $\ell \le N/2$, then

$$C_\ell(E_N) > (N-\ell)/k - 54k^2N^{1/2}\log N \quad \text{if } k \mid \ell$$
$$C_\ell(E_N) < 27k^2\ell N^{1/2}\log N \quad \text{if } k \nmid \ell$$

This result shows some kind of independence between C_k and C_ℓ when $k \nmid \ell$ and $\ell \nmid k$. In this section we will show a link between C_k and C_ℓ when k and ℓ have different parity.

Cassaigne, Mauduit and Sárközy [14] asked the following related question:

Problem 3.1. (Cassaigne, Mauduit, Sárközy) For $N \to \infty$, are there sequences E_N such that $C_2(E_N) = O(\sqrt{N})$ and $C_3(E_N) = O(1)$ simultaneously?

In [75] Mauduit also asked another closely related question

Problem 3.2. (Mauduit) Let $k,\ \ell \ge 2$ be integers. Is it true that for every $E_N \in \{-1,+1\}^N$ we have

$$C_{2k+1}(E_N)C_{2\ell}(E_N) \gg N$$

where the implied constant factor depends only on k and ℓ? Or at least

$$C_{2k+1}(E_N)C_{2\ell}(E_N) \gg N^{c(k,\ell)} \tag{3.1}$$

where the implied constant factor and the constant $\frac{1}{2} < c(k,\ell) \le 1$ depend only on k and ℓ?

In [36] I solved both Problem 3.1 and Problem 3.2 in the weaker form (3.1) when $k \geq \ell$. The answer follows from the main result of [36]:

Theorem 3.B *If $k, \ell \in \mathbb{N}$, $2k+1 > 2\ell$, $N \in \mathbb{N}$ and $N > 67k^4 + 400$, then for all $E_n \in \{-1, +1\}^N$ we have*

$$\left(17\sqrt{k(2\ell+1)}\, C_{2\ell}\right)^{2k+1} + \left(17\, \frac{2k+1}{2\ell}\right)^{\ell} N^{2k-\ell} C_{2k+1}^2 \geq \frac{1}{9} N^{2k-\ell+1}$$

It follows trivially that

Corollary 3.A *If $k, \ell \in \mathbb{N}$, $\log N \geq 2k+1 > 2\ell$, $N \in \mathbb{N}$ and $N > 67k^4+400$, $E_n \in \{-1, +1\}^N$ and*

$$C_{2\ell}(E_N) < \frac{1}{20\sqrt{k(2\ell+1)}} N^{1-\ell/(2k+1)}$$

then we have

$$C_{2k+1}(E_N) > \frac{1}{8}\left(\frac{2\ell}{17(2k+1)}\right)^{\ell/2} N^{1/2}.$$

Corollary 3.B *If $k, \ell \in \mathbb{N}$, $2k+1 > 2\ell$ then*

$$C_{2k+1}(E_N) C_{2\ell}(E_N) \gg N^{1-\ell/(2k+1)}$$

where the implied constant factor depends only on k and ℓ. (This is the case $c(k,\ell) = 1 - \frac{\ell}{2k+1} > \frac{1}{2}$ in Problem 3.2.)

Later Anantharam [5] sharpened Theorem 3.A and he proved the following:

Theorem 3.C (Anantharam)

$$C_3(E_N) C_2(E_N) \geq \frac{2}{25} N.$$

Theorem 3.C solves Problem 3.2 in the stronger form in the special case $(2k+1, 2\ell) = (3, 2)$, so (3.1) holds with $c = 1$.

3.1 Results

In this section we will generalize the earlier results. Theorem 3.B studies only the case $2k+1 > 2\ell$ while Theorem 3.C involves only C_2 and C_3. Here

we study the general case, when there is no restriction of the order of the correlation measures. The proof uses methods from [5] and [36]. We will prove the following:

Theorem 3.1 (Gyarmati, Mauduit) *There is a constant $c_{k,\ell}$ depending only on k and ℓ such that if*

$$C_{2k+1}(E_N) < c_{k,\ell} N^{1/2}, \tag{3.2}$$

then

$$C_{2k+1}(E_N)^{2\ell} C_{2\ell}(E_N)^{2k+1} \gg N^{2k+1}, \tag{3.3}$$

where the implied constant factor depends only on k and ℓ.

Remark 3.1 Theorem 3.1 is optimal: For $E_N = (+1, -1, +1, -1, +1 \dots)$ we have $C_{2k+1}(E_N) = 1$ and $C_{2\ell}(E_N) = N - 2\ell + 1$.

Remark 3.2 It is an important question whether condition (3.2) is necessary in Theorem 3.1. Cassaigne, Mauduit and Sárközy [14] proved that for every ε and $N > N_0(\varepsilon)$

$$C_{2k+1}(E_N), C_{2\ell}(E_N) \ll N^{1/2} (\log N)^{1/2} \tag{3.4}$$

holds with probability $1 - \varepsilon$. Fix a sequence E_N for which (3.4) indeed holds and N is large enough. From (3.3) and (3.4)

$$N^{\ell+k+1/2} (\log N)^{\ell+k+1/2} \gg N^{2k+1} \tag{3.5}$$

follows. Since (3.5) is true for an N large enough we get from (3.5):

$$\ell + k + 1/2 \geq 2k + 1$$

and thus

$$2\ell \geq 2k + 1.$$

But in Theorem 3.1 2ℓ can be less than $2k + 1$ so we need an additional assumption on the size of $C_{2k+1}(E_N)$ and $C_{2\ell}(E_N)$.

Let us see some corollaries of Theorem 3.1.

Corollary 3.1 (Gyarmati, Mauduit) *Suppose that $C_{2\ell}(E_N) \ll N^{1/2} (\log N)^{1/2}$, then*

$$C_{2k+1}(E_N) \gg \min \left\{ N^{1/2}, \ \frac{N^{(2k+1)/(4\ell)}}{(\log N)^{(2k+1)/(4\ell)}} \right\}$$

where the implied constant factor depends on k and ℓ.

Corollary 3.2 (Gyarmati, Mauduit) *If $C_{2k+1}(E_N) = O(1)$, then*

$$C_{2\ell}(E_N) \gg N,$$

where the implied constant factor depends on k and ℓ.

Corollary 3.3 (Gyarmati, Mauduit)

$$C_{2k+1}(E_N)C_{2\ell}(E_N) \gg N^{c(k,\ell)}$$

where the implied constant factor depends only on k and ℓ and where

$$c(k,\ell) = \begin{cases} 1 & \text{if } k \geq \ell, \\ \frac{1}{2} + \frac{2k+1}{4\ell} & \text{if } k < \ell. \end{cases}$$

Remark 3.3 Corollary 3.3 solves Problem 3.2 in the stronger form when $k \geq \ell$ and in the weaker form (3.1) when $k < \ell$.

These results can be extended to the multidimensional case, for the details see the paper [49].

3.2 Proof of Theorem 3.1

Let $L = [N/2]$ and $1 \leq M \leq N/2$ be integers, where the value of M will be fixed later. Consider the following equation

$$\begin{aligned} \mathcal{A} &\stackrel{\text{def}}{=} \sum_{1 \leq n_1 < n_2 < \cdots < n_{2k+1} \leq L} \sum_{1 \leq d_1 < d_2 < \cdots < d_{2\ell} \leq M} \prod_{j=1}^{2\ell} \prod_{i=1}^{2k+1} e_{n_i + d_j} \\ &= \sum_{1 \leq d_1 < d_2 < \cdots < d_{2\ell} \leq M} \sum_{1 \leq n_1 < n_2 < \cdots < n_{2k+1} \leq L} \prod_{i=1}^{2k+1} \prod_{j=1}^{2\ell} e_{n_i + d_j} \stackrel{\text{def}}{=} \mathcal{B}. \end{aligned}$$

We will use the following lemmas proved by me in [36].

Lemma 3.1 *For all $t, A \in \mathbb{N}$, $t \leq A$ there is a polynomial $p_{t,A}(x) \in \mathbb{Q}[x]$ with the degree t such that if $x_1, x_2, \ldots, x_A \in \{-1, +1\}$ then*

$$p_{t,A}(x_1 + \cdots + x_A) = \sum_{1 \leq i_1 < i_2 < \cdots < i_t \leq A} x_{i_1} x_{i_2} \ldots x_{i_t}.$$

Denote the coefficients of $p_{t,A}$ by $a_{r,t,A}$:

$$p_{t,A}(x) = a_{t,t,A}x^t + a_{t-1,t,A}x^{t-1} + \cdots + a_{0,t,A}.$$

Then $a_{r,t,A} = 0$ if $r \not\equiv t \pmod 2$, and $(-1)^{(t-r)/2}a_{r,t,A} \geq 0$ if $r \equiv t \pmod 2$. If t is even we also have:

$$a_{0,t,A} = (-1)^{t/2}\binom{A/2}{t/2}.$$

Proof of Lemma 3.1 This is Lemma 2 in [36].

Lemma 3.2

$$|a_{r,t,A}| \leq A^{(t-r)/2}.$$

Proof of Lemma 3.2 This follows from Lemma 3 and Lemma 5 in [36]. (Indeed in [36] by Lemma 3 we get $|a_{r,t,A}| \leq d_{i,j}A^{(t-r)/2}$. In [36] ω_j is defined by $d_{0,j} + d_{1,j} + \cdots + d_{j,j}$ in Lemma 4 and in Lemma 5 $d_{i,j} \leq \omega_j \leq 1$ is proved.)

Next we return to the proof of Theorem 3.1.

First we rearrange $\mathcal{A}$. For a moment we fix the value of $n_1, n_2, \ldots, n_{2k+1}$ in the first sum. Next we use Lemma 3.1 with $t = 2\ell$, $A = M$ and $x_u = \prod_{i=1}^{2k+1} e_{n_i+u}$ for $1 \leq u \leq M$. We get

$$\begin{aligned}\mathcal{A} &= \sum_{1\leq n_1<n_2<\cdots<n_{2k+1}\leq L}\ \sum_{1\leq d_1<d_2<\cdots<d_{2\ell}\leq M}\ \prod_{j=1}^{2\ell}\prod_{i=1}^{2k+1} e_{n_i+d_j} \\ &= \sum_{1\leq n_1<n_2<\cdots<n_{2k+1}\leq L} p_{2\ell,M}\left(\sum_{u=1}^{M}\prod_{i=1}^{2k+1} e_{n_i+u}\right).\end{aligned}$$

Similarly we rearrange $\mathcal{B}$. For a moment we fix the value of $d_1, d_2, \ldots, d_{2\ell}$ in the first sum. Next we use Lemma 3.1 with $t = 2k+1$, $A = L$ and $x_u = \prod_{j=1}^{2\ell} e_{u+d_j}$ for $1 \leq u \leq M$. We get

$$\begin{aligned}\mathcal{B} &= \sum_{1\leq d_1<d_2<\cdots<d_{2\ell}\leq M}\ \sum_{1\leq n_1<n_2<\cdots<n_{2k+1}\leq L}\ \prod_{i=1}^{2k+1}\prod_{j=1}^{2\ell} e_{n_i+d_j} \\ &= \sum_{1\leq d_1<d_2<\cdots<d_{2\ell}\leq M} p_{2k+1,L}\left(\sum_{u=1}^{L}\prod_{j=1}^{2\ell} e_{u+d_j}\right).\end{aligned}$$

We denoted the coefficients of $p_{t,A}(x)$ by $a_{r,t,A}$ in Lemma 3.1. Using these

notations we get

$$\sum_{1\le n_1<n_2<\cdots<n_{2k+1}\le L}\left(a_{2\ell,2\ell,M}\left(\sum_{u=1}^{M}\prod_{i=1}^{2k+1}e_{n_i+u}\right)^{2\ell}\right.$$
$$\left.+\,a_{2\ell-1,2\ell,M}\left(\sum_{u=1}^{M}\prod_{i=1}^{2k+1}e_{n_i+u}\right)^{2\ell-1}+\cdots+a_{0,2\ell,M}\right)$$
$$=\sum_{1\le d_1<d_2<\cdots<d_{2\ell}\le M}\left(a_{2k+1,2k+1,L}\left(\sum_{u=1}^{L}\prod_{j=1}^{2\ell}e_{u+d_j}\right)^{2k+1}\right.$$
$$\left.+\,a_{2k,2k+1,L}\left(\sum_{u=1}^{L}\prod_{j=1}^{2\ell}e_{u+d_j}\right)^{2k}+\cdots+a_{0,2k+1,L}\right). \quad (3.6)$$

By Lemma 3.1 $a_{0,2k+1,L}=0$. From this and (3.6) we get

$$\sum_{1\le d_1<d_2<\cdots<d_{2\ell}\le M}\left(a_{2k+1,2k+1,L}\left(\sum_{u=1}^{L}\prod_{j=1}^{2\ell}e_{u+d_j}\right)^{2k+1}\right.$$
$$\left.+\,a_{2k,2k+1,L}\left(\sum_{u=1}^{L}\prod_{j=1}^{2\ell}e_{u+d_j}\right)^{2k}+\cdots+a_{1,2k+1,L}\left(\sum_{u=1}^{L}\prod_{j=1}^{2\ell}e_{u+d_j}\right)\right)$$
$$-\sum_{1\le n_1<n_2<\cdots<n_{2k+1}\le L}\left(a_{2\ell,2\ell,M}\left(\sum_{u=1}^{M}\prod_{i=1}^{2k+1}e_{n_i+u}\right)^{2\ell}\right.$$
$$\left.+\,a_{2\ell-1,2\ell,M}\left(\sum_{u=1}^{M}\prod_{i=1}^{2k+1}e_{n_i+u}\right)^{2\ell-1}+\cdots+a_{1,2\ell,M}\left(\sum_{u=1}^{M}\prod_{i=1}^{2k+1}e_{n_i+u}\right)\right)$$
$$=\sum_{1\le n_1<n_2<\cdots<n_{2k+1}\le L}a_{0,2\ell,M}.$$

Again by Lemma 3.1 there is a constant c_1 depending only on k and ℓ such

that

$$\Bigg| \sum_{1\le d_1<d_2<\cdots<d_{2\ell}\le M} \Bigg(a_{2k+1,2k+1,L} \Bigg(\sum_{u=1}^{L} \prod_{j=1}^{2\ell} e_{u+d_j} \Bigg)^{2k+1} + a_{2k,2k+1,L} \Bigg(\sum_{u=1}^{L} \prod_{j=1}^{2\ell} e_{u+d_j} \Bigg)^{2k} + \cdots + a_{1,2k+1,L} \Bigg(\sum_{u=1}^{L} \prod_{j=1}^{2\ell} e_{u+d_j} \Bigg) \Bigg) - \sum_{1\le n_1<n_2<\cdots<n_{2k+1}\le L} \Bigg(a_{2\ell,2\ell,M} \Bigg(\sum_{u=1}^{M} \prod_{i=1}^{2k+1} e_{n_i+u} \Bigg)^{2\ell} + a_{2\ell-1,2\ell,M} \Bigg(\sum_{u=1}^{M} \prod_{i=1}^{2k+1} e_{n_i+u} \Bigg)^{2\ell-1} + \cdots + a_{1,2\ell,M} \Bigg(\sum_{u=1}^{M} \prod_{i=1}^{2k+1} e_{n_i+u} \Bigg) \Bigg) \Bigg| \ge c_1 L^{2k+1} M^{\ell}. \tag{3.7}$$

By Lemma 3.1 $a_{r,t,A} = 0$ if $r \not\equiv t \pmod 2$. Using this and the triangle-inequality we get from (3.7)

$$\sum_{1\le d_1<d_2<\cdots<d_{2\ell}\le M} \sum_{\substack{r=1 \\ r\equiv 1 \pmod 2}}^{2k+1} |a_{r,2k+1,L}| \left| \sum_{u=1}^{L} \prod_{j=1}^{2\ell} e_{u+d_j} \right|^r + \sum_{1\le n_1<n_2<\cdots<n_{2k+1}\le L} \sum_{\substack{r=2 \\ r\equiv 0 \pmod 2}}^{2\ell} |a_{r,2\ell,M}| \left| \sum_{u=1}^{M} \prod_{i=1}^{2k+1} e_{n_i+u} \right|^r \ge c_1 L^{2k+1} M^{\ell}. \tag{3.8}$$

By the definition of the correlation measures we have

$$\left| \sum_{u=1}^{L} \prod_{j=1}^{2\ell} e_{u+d_j} \right| \le C_{2\ell}(E_N),$$

$$\left| \sum_{u=1}^{M} \prod_{i=1}^{2k+1} e_{n_i+u} \right| \le C_{2k+1}(E_N).$$

By this and (3.8) we get

$$\sum_{1\le d_1<d_2<\cdots<d_{2\ell}\le M}\ \sum_{\substack{r=1\\ r\equiv 1 \pmod 2}}^{2k+1} |a_{r,2k+1,L}|\, C_{2\ell}(E_N)^r$$

$$+\sum_{1\le n_1<n_2<\cdots<n_{2k+1}\le L}\ \sum_{\substack{r=2\\ r\equiv 0 \pmod 2}}^{2\ell} |a_{r,2\ell,M}|\, C_{2k+1}(E_N)^r \ge c_1 L^{2k+1}M^{\ell}.$$

By this and Lemma 3.2

$$M^{2\ell}\sum_{\substack{r=1\\ r\equiv 1 \pmod 2}}^{2k+1} L^{(2k+1-r)/2}C_{2\ell}(E_N)^r + L^{2k+1}\sum_{\substack{r=2\\ r\equiv 0 \pmod 2}}^{2\ell} M^{(2\ell-r)/2}C_{2k+1}(E_N)^r$$
$$\ge c_1 L^{2k+1}M^{\ell}. \tag{3.9}$$

In order to prove Theorem 3.1 we will use Theorem 1.C as a lemma.

Lemma 3.3 (Alon, Kohayakawa, Mauduit, Moreira, Rödl)

$$C_{2\ell}(E_N) \gg N^{1/2}$$

where the implied constant factor depends only on ℓ.

Proof of Lemma 3.3 See in [3] and [69].

By this for $1 \le r \le 2k+1$ we have

$$L^{(2k+1-r)/2}C_{2\ell}(E_N)^r \ll C_{2\ell}(E_N)^{2k+1}.$$

Using this and (3.9) we get there is a constant c_2 depending only on k and ℓ such that

$$c_2M^{2\ell}C_{2\ell}(E_N)^{2k+1} + L^{2k+1}\sum_{\substack{r=2\\ r\equiv 0 \pmod 2}}^{2\ell} M^{(2\ell-r)/2}C_{2k+1}(E_N)^r$$
$$\ge c_1 L^{2k+1}M^{\ell}. \tag{3.10}$$

Now we fix the value of M. Let $M = c_3C_{2k+1}(E_N)^2$, where the value of the constant c_3 will depend only on k and ℓ. We choose the value of c_3 such that

$$\left\lceil \max_{2\le r\le 2\ell}\left(\frac{\ell+1}{c_1}\right)^{2/r}\right\rceil \le c_3.$$

Then

$$M^{(2\ell-r)/2}C_{2k+1}(E_N)^r \le \frac{c_1}{\ell+1}M^\ell \tag{3.11}$$

holds. Now we fix the constant $c_{k,\ell}$ in Theorem 3.1, we put $c_{k,\ell} = \frac{1}{2c_3}$. Then $2c_3C_{2k+1}(E_N)^2 \le N$, so $M \le N/2$ indeed. By (3.10) and (3.11) we get

$$\begin{aligned} c_2M^{2\ell}C_{2\ell}(E_N)^{2k+1} + L^{2k+1}\frac{c_1\ell}{\ell+1}M^\ell &\ge c_1L^{2k+1}M^\ell \\ c_2M^{2\ell}C_{2\ell}(E_N)^{2k+1} &\ge \frac{c_1}{\ell+1}L^{2k+1}M^\ell \\ M^{2\ell}C_{2\ell}(E_N)^{2k+1} &\ge \frac{c_1}{c_2(\ell+1)}L^{2k+1}M^\ell. \end{aligned}$$

Writing $L = [N/2]$ and $M = c_3C_{2k+1}(E_N)^2$ we get

$$\begin{aligned} c_3^{2\ell}C_{2k+1}(E_N)^{4\ell}C_{2\ell}(E_N)^{2k+1} &\ge \frac{c_1}{c_2(\ell+1)}\left[\frac{N}{2}\right]^{2k+1} c_3^\ell C_{2k+1}(E_N)^{2\ell} \\ C_{2k+1}(E_N)^{2\ell}C_{2\ell}(E_N)^{2k+1} &\gg N^{2k+1} \end{aligned}$$

which was to be proved.

The proofs of Corollaries 3.1 and 3.2 are immediate from Theorem 3.1.

3.3 Proof of Corollary 3.3

If $C_{2k+1}(E_N) \gg N^{1/2}$ then Corollary 3.3 is trivial since by Lemma 3.3 $C_{2\ell}(E_N) \gg N^{1/2}$ also holds and then $C_{2k+1}(E_N)C_{2\ell}(E_N) \gg N$. Thus we may assume that $C_{2k+1}(E_N) \ll N^{1/2}$

If $k < \ell$ by Theorem 3.1 and Lemma 3.3:

$$\begin{aligned} (C_{2k+1}(E_N)C_{2\ell}(E_N))^{2\ell} &= C_{2k+1}(E_N)^{2\ell}C_{2\ell}(E_N)^{2k+1}C_{2\ell}(E_N)^{2\ell-(2k+1)} \\ &\gg N^{2k+1}C_{2\ell}(E_N)^{2\ell-(2k+1)} \\ &\gg N^{2k+1}N^{\ell-k-1/2} = N^{\ell+k+1/2}, \end{aligned}$$

so that

$$C_{2k+1}(E_N)C_{2\ell}(E_N) \gg N^{1/2+(2k+1)/(4\ell)}.$$

If $k \geq \ell$ then by Theorem 3.1

$$\begin{aligned}(C_{2k+1}(E_N)C_{2\ell}(E_N))^{2k+1} &= C_{2k+1}(E_N)^{2\ell}C_{2\ell}(E_N)^{2k+1}C_{2k+1}(E_N)^{2k-2\ell+1} \\ &\gg N^{2k+1}C_{2k+1}(E_N)^{2k-2\ell+1} \\ &\gg N^{2k+1},\end{aligned}$$

so that

$$C_{2k+1}(E_N)C_{2\ell}(E_N) \gg N.$$

4 On the complexity of a family related to the Legendre symbol

In this section we study large families of finite, binary sequences

$$E_N = (e_1, e_2, \ldots, e_N) \in \{-1, +1\}^N.$$

In many applications it is not enough to know that the family contains many binary sequences with strong pseudorandom properties; it is also important that the family has a "rich", "complex" structure, there are many "independent" sequences in it. Ahlswede, Khachatrian, Mauduit and Sárközy [1] introduced the notion of *f-complexity* ("*f*" for family):

Definition 4.1 (Ahlswede, Khachatrian, Mauduit, Sárközy)
The family complexity $C(\mathcal{F})$ 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 $1 \le i_1 < i_2 < \cdots < i_j \le N$, and for $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_j \in \{-1, +1\}^j$, we have at least one $E_N = (e_1, \ldots, e_N) \in \mathcal{F}$ for which

$$e_{i_1} = \varepsilon_1,\ e_{i_2} = \varepsilon_2, \ldots, e_{i_j} = \varepsilon_j.$$

In [1] in Section 3 the following is proved: In order to get an upper bound for $C(\mathcal{F})$, we take all specifications of the form

$$e_1 = \varepsilon_1,\ e_2 = \varepsilon_2, \ldots, e_{C(\mathcal{F})} = \varepsilon_{C(\mathcal{F})}. \tag{4.1}$$

By the definition of f-complexity, for such a specification, there is a sequence $E \in \mathcal{F}$ for which (4.1) holds. $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{C(\mathcal{F})}$ may take $2^{C(\mathcal{F})}$ different values, thus,

$$2^{C(\mathcal{F})} \le |\mathcal{F}|.$$

So:

Proposition 4.1 (Ahlswede, Khachatrian, Mauduit, Sárközy)

$$C(\mathcal{F}) \le \frac{\log |\mathcal{F}|}{\log 2}.$$

Numerous binary sequences have been tested for pseudorandomness by J. Cassaigne, Z. Chen, X. Du, L. Goubin, K. Gyarmati, S. Ferenczi, S. Li, H. Liu, C. Mauduit, L. Mérai, J. Rivat and A. Sárközy. However, the first constructions produced only "few" pseudorandom sequences, usually for a fixed

integer N, the construction provided only one pseudorandom sequence E_N of length N. L. Goubin, C. Mauduit, A. Sárközy [31] succeeded in constructing large families of pseudorandom binary sequences. Their construction was the following:

Construction 4.1 (Goubin, Mauduit, Sárközy) *Suppose that p is a prime number, and $f(x) \in \mathbb{F}_p[x]$ is a polynomial with degree $k > 0$ and no multiple zero in $\overline{\mathbb{F}}_p$. Define the binary sequence $E_p = (e_1, \dots, e_p)$ by*

$$e_n = \begin{cases} \left(\frac{f(n)}{p}\right) & \text{for } (f(n), p) = 1 \\ +1 & \text{for } p \mid f(n). \end{cases} \tag{4.2}$$

Ahlswede, Khachatrian, Mauduit and Sárközy [1] proved the following:

Theorem 4.A (Ahlswede, Khachatrian, Mauduit, Sárközy) *Let p be a prime. Consider all the polynomials $f(x)$ such that*

$$0 < \deg f(x) \le K$$

(where $\deg f(x)$ denotes the degree of $f(x)$) and $f(x)$ has no multiple zero in $\overline{\mathbb{F}}_p$. For each of these polynomials $f(x)$, consider the binary sequence $E_p = E_p(f) = (e_1, e_2, \dots, e_p) \in \{-1, +1\}^p$ defined by (4.2), and let $\mathcal{F}_1$ denote the family of all the binary sequences obtained in this way. Then

$$C(\mathcal{F}_1) \ge K. \tag{4.3}$$

By Proposition 4.1 it is clear that

$$|C(\mathcal{F}_1)| \le \frac{\log |\mathcal{F}_1|}{\log 2} \le \frac{K+1}{\log 2} \log p.$$

We will improve on (4.3) and we will prove the following:

Theorem 4.1 *For the family defined in Theorem 4.1 we have*

$$C(\mathcal{F}_1) \ge \frac{K-1}{2 \log 2} \log p - O(K \log(K \log p)). \tag{4.4}$$

4.1 Proof of Theorem 4.1

In this proof c_1, c_2 will denote absolute constants. For $K \ge p^{1/2}/\log p$ the right-hand side of (4.4) is negative, so the theorem is trivial. Thus we may suppose that

$$K < p^{1/2}/\log p. \tag{4.5}$$

Let k be the greatest odd integer with $k \le K$. Let

$$j \le \frac{k}{2\log 2}\log p - \frac{c_1 k}{\log 2}\log(k\log p), \tag{4.6}$$

where we will fix the value of the absolute constant c_1 later. Suppose that we have the specification

$$e_{n_1} = \varepsilon_1,\ e_{n_2} = \varepsilon_2, \ldots, e_{n_j} = \varepsilon_j. \tag{4.7}$$

Let $I = \{n_1, n_2, \ldots, n_j\}$. We will consider all polynomials $f(x)$ of the form

$$f_{a_1,a_2,\ldots,a_k}(x) = (x-a_1)(x-a_2)\cdots(x-a_k) \tag{4.8}$$

with $a_i \notin I$, and we will prove by a counting argument that there is at least one k-tuple $a_1, a_2, \ldots, a_k$ (where $a_i \notin I$) for which the sequence E_p defined by (4.2) with $f_{a_1,a_2,\ldots,a_k}(x)$ in place of $f(x)$ satisfies (4.7). Suppose that $\beta_1, \beta_2, \ldots, \beta_t$ are the roots of $f(x)$ which have odd multiplicity in the factorization of $f(x)$. Since the degree of $f(x)$ is odd, t the number of these roots are also odd, so $t \ge 1$. Let $f_1(x) = (x-\beta_1)(x-\beta_2)\ldots(x-\beta_t)$. Then $f_1(x)$ has no multiple zero and the sequence E_p' defined by (4.2) with $f_1(x)$ in place of $f(x)$ satisfies (4.7).

Since this will be true for every $j \le \frac{k}{2\log 2}\log p - \frac{c_1 k}{\log 2}\log(k\log p)$ from this

$$C(\mathcal{F}) \ge \left[\frac{k}{2\log 2}\log p - \frac{c_1 k}{\log 2}\log(k\log p)\right] \ge \frac{K-1}{2\log 2}\log p - c_2 K\log(K\log p)$$

follows.

Now consider a k-tuple $a_1, a_2, \ldots, a_k$ with $a_i \notin I$, and consider the corresponding polynomial

$$f_{a_1,a_2,\ldots,a_k}(x) = (x-a_1)(x-a_2)\cdots(x-a_k).$$

Define the sequence $E_p = (e_1, e_2, \ldots, e_p)$ by

$$e_n = \begin{cases} \left(\frac{f_{a_1,a_2,\ldots,a_k}(n)}{p}\right) & \text{if } (f_{a_1,\ldots,a_k}(n), p) = 1, \text{ so } n \ne a_i \text{ for } 1 \le i \le k, \\ 1 & \text{if } p \mid f_{a_1,\ldots,a_k}(n), \text{ so } n = a_i \text{ for some } 1 \le i \le k. \end{cases} \tag{4.9}$$

Clearly,

$$\frac{1}{2}(1 + \varepsilon_i e_{n_i}) = \begin{cases} 1 & \text{if } e_{n_i} = \varepsilon_i, \\ 0 & \text{if } e_{n_i} = -\varepsilon_i. \end{cases}$$

If $n_i \neq a_s$ for $1 \leq s \leq l$ then

$$\frac{1}{2}\left(1+\varepsilon_i\left(\frac{(n_i-a_1)(n_i-a_2)\cdots(n_i-a_k)}{p}\right)\right) = \begin{cases} 1 & \text{if } e_{n_i}=\varepsilon_i, \\ 0 & \text{if } e_{n_i}=-\varepsilon_i. \end{cases}$$

Let N be the number of polynomials $f_{a_1,a_2,\ldots,a_k}(x) \in \mathbb{F}_p[x]$ with $a_1, a_2, \ldots, a_k \in \mathbb{F}_p \setminus I$ such that for the sequence (4.9) specification (4.7) holds. Then

$$N = \sum_{\substack{a_1=0 \\ a_1\notin I}}^{p-1}\sum_{\substack{a_2=0 \\ a_2\notin I}}^{p-1}\cdots\sum_{\substack{a_k=0 \\ a_k\notin I}}^{p-1}\frac{1}{2^j}\prod_{i=1}^{j}\left(1+\varepsilon_i\left(\frac{(n_i-a_1)(n_i-a_2)\cdots(n_i-a_k)}{p}\right)\right). \tag{4.10}$$

Here

$$\begin{aligned} A(a_1,\ldots,a_k) &\stackrel{\text{def}}{=} \prod_{i=1}^{j}\left(1+\varepsilon_i\left(\frac{(n_i-a_1)\cdots(n_i-a_k)}{p}\right)\right) \\ &= 1+\sum_{\ell=1}^{j}\sum_{1\leq i_1<i_2<\cdots<i_\ell\leq j}\varepsilon_{i_1}\varepsilon_{i_2}\cdots\varepsilon_{i_\ell}\left(\frac{(n_{i_1}-a_1)\cdots(n_{i_1}-a_k)}{p}\right) \\ &\quad \left(\frac{(n_{i_2}-a_1)\cdots(n_{i_2}-a_k)}{p}\right)\cdots\left(\frac{(n_{i_\ell}-a_1)\cdots(n_{i_\ell}-a_k)}{p}\right). \end{aligned}$$

The Legendre symbol is multiplicative, thus

$$\begin{aligned} A(a_1,\ldots,a_k) = 1+\sum_{\ell=1}^{j}\sum_{1\leq i_1<i_2<\cdots<i_\ell\leq j}\varepsilon_{i_1}\varepsilon_{i_2}\cdots\varepsilon_{i_\ell} \\ \prod_{j=1}^{k}\left(\frac{(n_{i_1}-a_j)(n_{i_2}-a_j)\ldots(n_{i_\ell}-a_j)}{p}\right). \end{aligned}$$

Writing this in (4.10) we get

$$\begin{aligned} N = \sum_{\substack{a_1=0 \\ a_1\notin I}}^{p-1}\cdots\sum_{\substack{a_k=0 \\ a_k\notin I}}^{p-1}\frac{1}{2^j}\Bigg(1+\sum_{\ell=1}^{j}\sum_{1\leq i_1<i_2<\cdots<i_\ell\leq j}\varepsilon_{i_1}\cdots\varepsilon_{i_\ell} \\ \prod_{t=1}^{k}\left(\frac{(n_{i_1}-a_t)(n_{i_2}-a_t)\ldots(n_{i_\ell}-a_t)}{p}\right)\Bigg), \end{aligned}$$

$$N = \frac{(p-|I|)^k}{2^j} + \frac{1}{2^j} \sum_{\substack{a_1=0 \\ a_1 \notin I}}^{p-1} \cdots \sum_{\substack{a_k=0 \\ a_k \notin I}}^{p-1} \sum_{\ell=1}^{j} \sum_{1 \le i_1 < i_2 < \cdots < i_\ell \le j} \varepsilon_{i_1} \cdots \varepsilon_{i_\ell}$$
$$\prod_{t=1}^{k} \left(\frac{(n_{i_1} - a_t)(n_{i_2} - a_t) \ldots (n_{i_\ell} - a_t)}{p} \right)$$
$$= \frac{(p-j)^k}{2^j} + \frac{1}{2^j} \sum_{\ell=1}^{j} \sum_{1 \le i_1 < i_2 < \cdots < i_\ell \le j} \varepsilon_{i_1} \cdots \varepsilon_{i_\ell} \sum_{\substack{a_1=0 \\ a_1 \notin I}}^{p-1} \cdots \sum_{\substack{a_k=0 \\ a_k \notin I}}^{p-1}$$
$$\prod_{t=1}^{k} \left(\frac{(n_{i_1} - a_t)(n_{i_2} - a_t) \ldots (n_{i_\ell} - a_t)}{p} \right)$$
$$= \frac{(p-j)^k}{2^j} + \frac{1}{2^j} \sum_{\ell=1}^{j} \sum_{1 \le i_1 < i_2 < \cdots < i_\ell \le j} \varepsilon_{i_1} \cdots \varepsilon_{i_\ell}$$
$$\left(\sum_{\substack{a=0 \\ a \notin I}}^{p-1} \frac{(n_{i_1} - a)(n_{i_2} - a) \ldots (n_{i_\ell} - a)}{p} \right)^k . \tag{4.11}$$

Lemma 4.1 (Weil) *Suppose that p is a prime, χ is a non-principal character modulo p of order d, $f \in \mathbb{F}_p[x]$ has s distinct roots in $\overline{\mathbb{F}}_p$, and it is not a constant multiple of the d-th power of a polynomial over $\mathbb{F}_p$. Then:*

$$\left| \sum_{n \in \mathbb{F}_p} \chi(f(n)) \right| < s p^{1/2}.$$

Poof of Lemma 4.1

This is Weil's theorem, see [106].

By the triangle-inequality and by Lemma 4.1:

$$\left| \sum_{\substack{a=0 \\ a \notin I}}^{p-1} \left(\frac{(n_{i_1} - a)(n_{i_2} - a) \ldots (n_{i_\ell} - a)}{p} \right) \right| \le$$
$$\left| \sum_{a=0}^{p-1} \left(\frac{(n_{i_1} - a)(n_{i_2} - a) \ldots (n_{i_\ell} - a)}{p} \right) \right| + j \le \ell p^{1/2} + j \le j p^{1/2} + |I| .$$

Thus by (4.11) and the triangle-inequality

$$N \geq \frac{(p-j)^k}{2^j} - \frac{1}{2^j}\sum_{\ell=1}^{j}\sum_{1\leq i_1<i_2<\cdots<i_\ell\leq j}(jp^{1/2}+j)^k = \frac{(p-j)^k}{2^j} - (jp^{1/2}+j)^k.$$

Thus $N > 0$ follows from

$$\begin{aligned}\frac{p-j}{2^{j/k}} &> jp^{1/2}+j\\ p &> 2^{j/k}(jp^{1/2}+j)+j. \qquad (4.12)\end{aligned}$$

Thus it remains to prove (4.12). By (4.6)

$$\begin{aligned}2^{j/k}(jp^{1/2}+j)+j &\leq 2^{\left(\frac{1}{2\log 2}\log p-\frac{c_1}{\log 2}\log(k\log p)\right)}\left(\frac{k}{2\log 2}p^{1/2}\log p+\frac{k}{2\log 2}\log p\right)\\ &+\frac{k}{2\log 2}\log p \leq \frac{p^{1/2}}{(k\log p)^{c_1}}\left(\frac{k\log p}{2\log 2}p^{1/2}+\frac{kp^{1/2}\log p}{2\cdot 3^{1/2}\log 2}\right)\\ &+\frac{k}{2\log 2}\log p \leq \frac{p^{1/2}}{(k\log p)^{c_1}}1.138(k\log p)p^{1/2}+\frac{k}{2\log 2}\log p.\end{aligned}$$

By this and (4.5)

$$\begin{aligned}2^{j/k}(jp^{1/2}+j)+j &\leq 1.138\frac{p}{(k\log p)^{c_1-1}}+\frac{p^{1/2}}{2\log 2}\\ &\leq 1.138\frac{p}{(k\log p)^{c_1-1}}+\frac{p}{2\cdot 3^{1/2}\log 2}\\ &\leq 1.138\frac{p}{(k\log p)^{c_1-1}}+0.414p.\end{aligned}$$

For $c_1 = 9$ we have

$$2^{j/k}(jp^{1/2}+j)+j \leq 1.138\frac{p}{(\log 3)^8}+0.414p < p$$

which proves (4.12). Thus for $j \leq \frac{k}{2\log 2}\log p - \frac{9k}{\log 2}\log(k\log p)$ we have that (4.12) holds. Then $N > 0$. So there is a sequence E_p for which specification (4.7) holds. Thus we proved

$$C(\mathcal{F}_1) \geq \left[\frac{k}{2\log 2}\log p - \frac{9k}{\log 2}\log(k\log p)\right] \geq \frac{K-1}{2\log 2}\log p - O(K\log(K\log p)).$$

5 On the correlation of subsequences

A sequence E_N is considered a "good" pseudorandom sequence if each of these measures $W(E_N)$, $C_\ell(E_N)$ (at least for small ℓ) is "small" in terms of N (in particular all are $o(N)$ as $N \longrightarrow \infty$). Indeed, it was proved in [14] that for a truly random sequence $E_N \subseteq \{-1,+1\}^N$ each of these measures is $\ll \sqrt{N \log N}$ and $\gg \sqrt{N}$. Later these bounds were sharpened by Alon, Kohayakawa, Mauduit, Moreira and Rödl [4] (see Theorems 1.A and 1.B).

Numerous binary sequences have been tested for pseudorandomness by several authors. In the best constructions we have $W(E_N) \ll \sqrt{N}(\log N)^{c_1}$ and $C_\ell(E_N) \ll \sqrt{N}(\log N)^{c_\ell}$ with positive constants c_1 and c_ℓ. From this it follows that

$$|U(E_N, t, a, b)| \ll N^{1/2}(\log N)^{c_1} \tag{5.1}$$

and

$$|V(E_N, M, D)| \ll N^{1/2}(\log N)^{c_\ell} \tag{5.2}$$

(for all t, a, b, M, D). For every M and t, we trivially have

$$\max_{E_N \in \{-1,+1\}^N} |U(E_N, t, a, b)| = t,$$
$$\max_{E_N \in \{-1,+1\}^N} |V(E_N, M, D)| = M.$$

If $|U(E_N, t, a, b)|$ is large compared with t or $|V(E_N, M, D)|$ is large compared with M, then it may occur that our sequence E_N has a "part" with weak pseudorandom properties. Indeed, if t or M is smaller than $\sqrt{N}$ then the estimates (5.1) and (5.2) are trivial. Thus it may occur that, say, we want to encrypt a message of estimated length slightly less than N, thus we use an N bit sequence possessing strong pseudorandom properties. However, it may turn out that the text to be encrypted is of length less than, say, $\sqrt{N}$. In this case we use only a short part (of length $\sqrt{N}$) of the sequence although we do not have any control over the pseudorandom properties of the short subsequences. In this section we would like to present constructions with non-trivial estimates for $V(E_N, M, D)$ in case of small M's.

Theorem 5.1 *For every N there is a binary sequence $E_N \in \{-1,+1\}^N$ such that if $D = (d_1, d_2, \dots, d_\ell)$ and $M \le N^{1/2}$ are such that $0 \le d_1 < d_2 < \cdots < d_\ell < M + d_\ell \le N$, then we have*

$$|V(E_N, M, D)| \ll \ell^2 N^{1/4} \log N. \tag{5.3}$$

From this follows that for $1 \le M \le N$ *we have*

$$|V(E_N, M, D)| \ll \ell^2 \left\lceil \frac{M}{N^{1/2}} \right\rceil N^{1/4} \log N.$$

Corollary 5.1 *For the binary sequence* $E_N \in \{-1, +1\}^N$ *constructed in the proof of Theorem 5.1 we have*

$$C_\ell(E_M) \ll \ell^2 \left\lceil \frac{M}{N^{1/2}} \right\rceil N^{1/4} \log N \tag{5.4}$$

for every $M \le N$ *and* $E_M \subseteq E_N$ *(so* E_M *is of the form* $(e_x, e_{x+1}, \dots, e_M)$*).*

It is an interesting question whether similar results hold for $U(E_N, t, a, b)$?

Theorem 5.1 is not optimal in the sense that it follows from (5.4) for the sequence E_N which satisfies the conditions of Theorem 5.1 that

$$C_\ell(E_N) \ll \ell^2 N^{3/4} \log N,$$

while in the best constructions we have $C_\ell(E_N) \ll N^{1/2} (\log N)^{c_\ell}$. Next we will show the existence of such a sequence.

Theorem 5.2 *For every* N *there is a binary sequence* $E_N \in \{-1, +1\}^N$ *such that if* $D = (d_1, d_2, \dots, d_\ell)$ *and* $M \le N^{1/2}$ *satisfy* $0 \le d_1 < d_2 < \cdots < d_\ell < M + d_\ell \le N$*, then we have*

$$|V(E_N, M, D)| \ll \ell^2 N^{1/4} \log N. \tag{5.5}$$

Moreover

$$C_\ell(E_N) \ll \ell^2 N^{1/2} (\log N)^2 \tag{5.6}$$

and

$$W(E_N) \ll N^{3/4} \log N \tag{5.7}$$

holds.

From (5.5) *follows that for* $1 \le M \le N$ *we have*

$$|V(E_N, M, D)| \ll \ell^2 \left\lceil \frac{M}{N^{1/2}} \right\rceil N^{1/4} \log N.$$

Corollary 5.2 *For the binary sequence* $E_N \in \{-1, +1\}$ *constructed in the*

proof of Theorem 5.2 we have

$$C_\ell(E_M) \ll \ell^2 \left\lceil \frac{M}{N^{1/2}} \right\rceil N^{1/4} \log N, \tag{5.8}$$

$$W(E_M) \ll \left\lceil \frac{M}{N^{1/2}} \right\rceil^{1/2} M^{1/2} N^{1/8} (\log N)^{1/2}, \tag{5.9}$$

for every $M \le N$ and $E_M \subseteq E_N$ (where E_M is of the form $(e_x, e_{x+1}, \dots, e_{x+M-1})$). Moreover

$$C_\ell(E_N) \ll \ell^2 N^{1/2} (\log N)^2$$

and

$$W(E_N) \ll N^{3/4} \log N$$

holds.

The proofs of Theorems 5.1 and 5.2 are constructive. The construction in Theorem 5.2 uses two-dimensional binary lattices (see Section 1). In [50] we reduced the two dimensional case to the one dimensional one by the following way: To any 2-dimensional binary N-lattice

$$\eta(\underline{x}) : \ I_N^2 \to \{-1, +1\} \tag{5.10}$$

we may assign a unique binary sequence $E_{N^2} = E_{N^2}(\eta) = (e_1, e_2, \dots, e_{N^2}) \in \{-1, +1\}^N$ by taking the first (from the bottom) row of the lattice (5.10) then we continue the binary sequence by taking the second row of the lattice, then the third row follows, etc.; in general, we set

$$e_{iN+j} = \eta((j-1, i)) \text{ for } i = 0, 1, \dots, N-1, \ j = 1, 2, \dots, N. \tag{5.11}$$

In [50] we asked if it is true that if $E_{N^2}(\eta)$ is a "good" pseudorandom binary *sequence* then η is a "good" pseudorandom 2-dimensional lattice? The answer to this question is negative; we showed that it may occur that the pseudorandom measures of the sequence $E_{N^2}(\eta)$ are small, however, the corresponding pseudorandom measures of the lattice η are large. Here we study the opposite. We will prove that if the lattice η has small correlation measure, then the corresponding $E_N^2(\eta)$ sequence has small correlation measures as well.

Theorem 5.3 *Let η be an arbitrary binary lattice. Then*

$$C_\ell(E_{N^2}(\eta)) \le (\ell + 2) C_\ell(\eta).$$

By $C_\ell(\eta) \le Q_\ell(\eta)$ it follows that

Corollary 5.3 *Let η be an arbitrary binary lattice. Then*

$$C_\ell(E_{N^2}(\eta)) \le (\ell+2)Q_\ell(\eta).$$

In the proof of Theorem 5.2 we will use Theorem 5.3. But Theorem 5.3 is of independent interest: by using Theorem 5.3 we can construct pseudorandom binary sequences by using pseudorandom binary lattices.

We remark that one may obtain similar results for shorter intervals in Theorem 5.2: If t is an integer then for $M \le N^{1/t}$ we have

$$|V(E_N, M, D)| \ll N^{1/(2t)} \log N$$

in place of (5.5) while $C_\ell(E_N) \ll N^{1/2}(\log N)^{c_\ell}$ and $W(E_N) \ll N^{3/4}(\log N)^{c_1}$ also holds. However the proof of this result would be lengthy (we would need more sophisticated sums as the ones in Lemma 5.4 and the relation between the pseudorandom measures of the binary lattices and the associated binary sequences is more complicated) thus we omit here the details, but one might like to return to this problem in a subsequent paper.

Throughout the section $[a, b]$ will denote the set $\{a, a+1, \dots, b\}$.

5.1 Proofs

Proof of Theorem 5.1

For $N = 2$ the theorem is trivial. For $N \ge 3$ by Chebysev's theorem there exists an odd prime p such that

$$N^{1/2} < p < 2N^{1/2}. \tag{5.12}$$

For an irreducible polynomial $f(x) \in \mathbb{F}_p[x]$ of degree $k \ge 2$, we define a binary sequence $E_p(f) = (e_1, e_2, \dots, e_p)$ by the following way:

$$e_n = \left(\frac{f(n)}{p}\right).$$

(We remark that since f is irreducible, for an integer n, $f(n)$ is never divisible by p thus $\left(\frac{f(n)}{p}\right)$ always assumes ± 1.) Next we will construct a pseudorandom binary sequence for which (5.3) holds. Let $f_1(x), f_2(x), \dots, f_p(x)$ be different irreducible polynomials of degree $k \ge 2$ and for $1 \le i \le p$ let $f_i(x)$ be of the form

$$f_i(x) = x^k + a_{i,k-2}x^{k-2} + a_{i,k-3}x^{k-3} + \cdots + a_{i,0} \tag{5.13}$$

with $a_{i,j} \in \mathbb{F}_p$. (so the coefficient of x^{k-1} is 0 in $f_i(x)$). We remark that the

number of monic irreducible polynomials of degree $k < p$ over the finite field $\mathbb{F}_q$ is

$$L_q(k) = \frac{1}{k}\sum_{d|k}\mu\left(\frac{k}{d}\right)q^d$$

see [29, pp. 602-629]. For $k \geq 4$

$$L_q(k) \geq \frac{1}{k}q^k - \frac{1}{k}\sum_{d=1}^{[k/2]}q^d \geq \frac{1}{k}q^k - \frac{1}{k}q\frac{q^{k/2}-1}{q-1} \geq \frac{1}{k}\left(q^k - q^{(k+2)/2}\right) \geq \frac{1}{2k}q^k.$$

For every $j \in \mathbb{F}_q$ consider $f(x+j)$. Between these q different irreducible polynomials there is exactly one which is of the form

$$f(x+j) = x^k + a_{k-2}x^{k-2} + \cdots + a_0$$

(so the coefficient of x^{k-1} is 0 in $f(x+j)$). Thus for $k \geq 4$ and $p \geq 3$ the number of irreducible polynomials which are of the form $x^k + a_{k-2}x^{k-2} + \cdots + a_0$ is

$$N_q(k) \stackrel{\text{def}}{=} \frac{1}{q}L_q(k) \geq \frac{1}{2k}q^{k-1}. \tag{5.14}$$

For $k \geq 4$, $p \geq 3$ we have $N_p(k) \geq p$, thus there exist p different irreducible polynomials $f_1(x), f_2(x), \ldots, f_p(x)$ which are of the form (5.13). Let

$$E_{p^2} \stackrel{\text{def}}{=} (E_p(f_1), E_p(f_2), \ldots, E_p(f_p)) \tag{5.15}$$

where E_{p^2} is a binary sequence of length p^2 obtained by writing the elements of $E_p(f_1), E_p(f_2), \ldots, E_p(f_p)$ successively. Let $E_{p^2} = (e_1, e_2, \ldots, e_{p^2})$ and since by (5.12) we have

$$N < p^2 < 4N,$$

we may define E_N by the sequence of the first N elements of E_{p^2}:

$$E_N = (e_1, e_2, \ldots, e_N).$$

If $M < p$, $D = (d_1, \ldots, d_\ell)$

$$\begin{aligned} &V(E_N, M, D) = V(E_{p^2}, M, D) \\ &= e_{1+d_1}e_{1+d_2}\ldots e_{1+d_\ell} + e_{2+d_1}e_{2+d_2}\ldots e_{2+d_\ell} + \cdots + e_{M+d_1}e_{M+d_2}\ldots e_{M+d_\ell}. \end{aligned}$$

Next we will prove that for each $1 \leq i \leq \ell$ and $1 \leq n < M$, there exist

integers a_i, b_i and intervals $I_i = \{1, 2, \ldots, b_i\}$ and $J_i = \{b_i+1, b_i+2, \ldots, M\}$ such that

$$e_{n+d_i} = \begin{cases} \left(\frac{f_{a_i}(n+d_i)}{p}\right) & \text{if } n \in I_i, \\ \left(\frac{f_{a_i+1}(n+d_i)}{p}\right) & \text{if } n \in J_i, \end{cases} \tag{5.16}$$

(if $b_i = M$ then $J_i = \emptyset$). Indeed, let $m_p(x)$ denote the least nonnegative integer with

$$x \equiv m_p(x) \pmod{p},$$

so $0 \le m_p(x) \le p-1$. Then

$$n + d_i = \left[\frac{n+d_i-1}{p}\right] p + m_p(n+d_i-1) + 1.$$

Thus

$$e_{n+d_i} = f_{\left[\frac{n+d_i-1}{p}\right]+1}(m_p(n+d_i-1)+1) = f_{\left[\frac{n+d_i-1}{p}\right]+1}(n+d_i). \tag{5.17}$$

In (5.16) $0 \le n \le M < p$. Let $d_i = q_i p + s_i$ where $0 \le s_i \le p-1$. Then

$$\begin{aligned} \left[\frac{n+d_i-1}{p}\right] &= \left[\frac{q_i p + s_i + n - 1}{p}\right] = q_i + \left[\frac{s_i+n-1}{p}\right] \\ &= \begin{cases} q_i & \text{if } n \le p - s_i, \\ q_i + 1 & \text{if } n > p - s_i, \end{cases} \end{aligned} \tag{5.18}$$

which proves (5.16) with $a_i = q_i + 1$ and $b_i = \max\{p - s_i, M\}$, so $I_i = [1, b_i]$, $J_i = [b_i{+}1, M]$ (if $b_i = M$ then $J_i = \emptyset$). Then $\{1, b_1{+}1, b_2{+}1, \ldots, b_\ell{+}1, M{+}1\}$ is a multiset which contains integers $1 = c_1 < c_2 < \cdots < c_m = M+1$ where

$$m \le \ell + 2. \tag{5.19}$$

Then $[0, M] = \cup_{j=1}^{m-1} [c_j, c_{j+1} - 1]$. By the definition of the c_j's, $c_j < b_i + 1 < c_{j+1}$ is not possible, thus $c_{j+1} - 1 \le b_i$ or $b_i \le c_j - 1$, so $[c_j, c_{j+1} - 1] \subseteq [0, b_i]$ or $[c_j, c_{j+1} - 1] \subseteq [b_i + 1, M]$. Hence

$$V(E_N, M, D) = \sum_{n=1}^{M} e_{n+d_1} \ldots e_{n+d_\ell} = \sum_{j=1}^{m-1} \sum_{n \in [c_j, c_{j+1}-1]} e_{n+d_1} \ldots e_{n+d_\ell}. \tag{5.20}$$

Now each interval $[c_j, c_{j+1} - 1]$ is either $\subseteq I_i$ or $\subseteq J_i$ for every $1 \le i \le \ell$. Thus for every $d_1, d_2, \ldots, d_\ell$ and for every interval $[c_j, c_{j+1} - 1]$ there exists fixed numbers $h_1, h_2, \ldots, h_\ell$ (depending only on $d_1, d_2, \ldots, d_\ell$ and j) such that

for $n \in [c_j, c_{j+1} - 1]$

$$e_{n+d_1} e_{n+d_2} \dots e_{n+d_\ell} = \left(\frac{f_{h_1}(n+d_1)}{p}\right)\left(\frac{f_{h_2}(n+d_2)}{p}\right)\dots\left(\frac{f_{h_\ell}(n+d_\ell)}{p}\right)$$
$$= \left(\frac{f_{h_1}(n+d_1) f_{h_2}(n+d_2)\dots f_{h_\ell}(n+d_\ell)}{p}\right).$$

Next we estimate

$$\sum_{n\in[c_j,c_{j+1}-1]} e_{n+d_1} e_{n+d_2} \dots e_{n+d_\ell}$$
$$= \sum_{n\in[c_j,c_{j+1}-1]} \left(\frac{f_{h_1}(n+d_1) f_{h_2}(n+d_2)\dots f_{h_\ell}(n+d_\ell)}{p}\right).$$

Here $f_{h_1}(x+d_1), \dots, f_{h_\ell}(x+d_\ell)$ are different polynomials. Indeed if

$$f_{h_r}(x+d_r) = f_{h_t}(x+d_t),$$

then substituting $x + d_r$ by x we get

$$f_{h_r}(x) = f_{h_t}(x + d_t - d_r). \tag{5.21}$$

It is easy to see that there is exactly one among the polynomials $f_{h_t}(x), f_{h_t}(x+1), \dots, f_{h_t}(x+p-1)$ for which the coefficient of x^{k-1} is 0, and this one is $f_{h_t}(x)$. From this and (5.21) follows that

$$d_r \equiv d_t \pmod{p}. \tag{5.22}$$

Thus from (5.21) we get

$$f_{h_r}(x) = f_{h_t}(x).$$

Since the polynomials $f_1, f_2, \dots, f_\ell$ are different, from this

$$h_r = h_t \tag{5.23}$$

follows. Now we compute the value $h_r = h_t$. By (5.17) for $n \in [c_j, c_{j+1} - 1]$ $e_{n+d_r} = f_{h_r}(n+d_r)$, $e_{n+d_t} = f_{h_t}(n+d_t)$ where

$$h_r = \left[\frac{n+d_r-1}{p}\right] + 1,$$
$$h_t = \left[\frac{n+d_t-1}{p}\right] + 1. \tag{5.24}$$

By (5.23) and (5.24)

$$\left[\frac{n+d_r-1}{p}\right] = \left[\frac{n+d_t-1}{p}\right]. \tag{5.25}$$

Now

$$n+d_r = q_r p + s_r,\ n+d_t = q_t p + s_t \tag{5.26}$$

where $0 \le s_r, s_t \le p-1$. By (5.22)

$$s_r = s_t. \tag{5.27}$$

Now

$$\left[\frac{n+d_r-1}{p}\right] + 1 = \left[\frac{q_r p + s_r - 1}{p}\right] + 1 = q_r + 1 + \left[\frac{s_r-1}{p}\right].$$

Similarly

$$\left[\frac{n+d_t-1}{p}\right] = q_t + 1 + \left[\frac{s_t-1}{p}\right].$$

By this, (5.25) and (5.27) we have

$$q_r = q_t.$$

By this, (5.26) and (5.27)

$$d_r = d_t,$$

which is a contradiction. So indeed, the irreducible polynomials $f_{h_1}(x + d_1), \ldots, f_{h_\ell}(x + d_\ell)$ are different. Thus the product $f_{h_1}(x + d_1)f_{h_2}(x + d_2) \ldots f_{h_\ell}(x + d_\ell)$ is not of the form $cg^2(x)$. We will use the following lemma:

Lemma 5.1 (Winterhof) *Suppose that p is a prime, χ is a non-principal character modulo p of order d, $f \in \mathbb{F}_p[x]$ has s distinct roots in $\overline{\mathbb{F}}_p$, and it is not a constant multiple of the d-th power of a polynomial over $\mathbb{F}_p$. Let y be a real number with $0 < y \le p$. Then for any $x \in \mathbb{R}$:*

$$\left|\sum_{x<n\le x+y} \chi(f(n))\right| < p^{1/2}\log p. \tag{5.28}$$

Poof of Lemma 5.1

This lemma is the one-dimensional case of Lemma 5.10 due to Winterhof [107], who derived it from Weil theorem [106]. We mention that a slightly weaker version of the lemma can be found in Lemma 1 in [2] where $9sp^{1/2}\log p$

is proved in place of the right hand side of (5.28). (In the case $f(x) = x$ the best constant factor is achieved by Bourgain, Cochrane, Paulhus and C. Pinner in [10], and their method also works for higher degree polynomials.)

Since later in the proof we will also use Weil's theorem, we state it here as a lemma (see in [71] and [106]):

Lemma 5.2 (Weil) *Suppose that p is a prime, χ is a non-principal character modulo p of order d, $f \in \mathbb{F}_p[x]$ has s distinct roots in $\overline{\mathbb{F}}_p$, and it is not a constant multiple of the d-th power of a polynomial over $\mathbb{F}_p$. Then:*

$$\left| \sum_{n \in \mathbb{F}_p} \chi(f(n)) \right| < sp^{1/2}.$$

By Lemma 5.1 we get

$$\begin{aligned} &\sum_{n \in [c_j, c_{j+1}-1]} e_{n+d_1} e_{n+d_2} \dots e_{n+d_\ell} \\ &= \sum_{n \in [c_j, c_{j+1}-1]} \left(\frac{f_{h_1}(n+d_1) f_{h_2}(n+d_2) \dots f_{h_\ell}(n+d_\ell)}{p} \right). \\ &< \ell k p^{1/2} \log p. \end{aligned}$$

By (5.19) and (5.20) we get

$$|V(E_N, M, D)| \ll \ell^2 k p^{1/2} \log p \ll \ell^2 k N^{1/4} \log N. \tag{5.29}$$

Since k, the degree of the polynomials $f_1(x), f_2(x), \dots, f_p(x)$ can be chosen as $k = 4$, from (5.29) we get (5.3), which was to be proved.

Proof of Theorem 5.2 First we will need some technical preparation in order to be able to estimate character sums of the type which appear later in the proof of our theorem. First Katz [67] and Perelmuter-Shparlinski [92] studied character sums over subfields of a finite field. Their result was generalized by Wan [105] who proved the following very general theorem:

Lemma 5.3 (Wan) *Let the $f_i(T)$ with $1 \le i \le n$ be pairwise coprime polynomials. Let D be the degree of the largest squarefree divisor of $\prod_{i=1}^n f_i(T)$. Let χ_i be a multiplicative character of the field $\mathbb{F}_{q^m}$ for $1 \le i \le n$. Suppose that for some $1 \le i \le n$, there is a root ξ_i of multiplicity m_i of $f_i(T)$ such that the character χ^{m_i} is non-trivial on the set $Norm_{\mathbb{F}_{q^m}[\xi_i]/\mathbb{F}_{q^m}}(\mathbb{F}_q[\xi])$. Then we have*

$$\left| \sum_{a \in \mathbb{F}_q} \chi_1(f_1(a)) \dots \chi_n(f_n(a)) \right| \le (mD-1) q^{1/2}.$$

Part a) of the following lemma is a consequence of Lemma 5.3, while the estimate in part b) - the incomplete case - is new and I will derive it directly from Weil's theorem. (At the same time I will also give an alternative proof for part a), since in order to do so I just need to add one more sentence to the proof of part b).)

Lemma 5.4 *Let p be an odd prime, $q = p^2$ and denote the quadratic character of $\mathbb{F}_q$ by γ. Clearly $\mathbb{F}_p \subseteq \mathbb{F}_q$. Let $I = [a, a+1, a+2, \ldots, b] \subseteq \mathbb{F}_p$ and $f(x) \in \mathbb{F}_q[x]$ be a polynomial which is not of the form $cg(x)h^2(x)$ with $c \in \mathbb{F}_q$, $g(x) \in \mathbb{F}_p[x]$ and $h(x) \in \mathbb{F}_q[x]$. Suppose that $f(x)$ has m distinct zeros in its splitting field over $\mathbb{F}_p$. Then*

$$a) \quad \left| \sum_{x \in \mathbb{F}_p} \gamma(f(x)) \right| \leq 2mp^{1/2}, \tag{5.30}$$

$$b) \quad \left| \sum_{x \in I} \gamma(f(x)) \right| \leq 2mp^{1/2}(1 + \log p). \tag{5.31}$$

Proof of Lemma 5.4 Let $n \in \mathbb{F}_p$ be a quadratic non-residue modulo p, so

$$\left(\frac{n}{p}\right) = -1. \tag{5.32}$$

The polynomial $x^2 - n \in \mathbb{F}_q[x] = \mathbb{F}_{p^2}[x]$ is reducible in $\mathbb{F}_q[x]$, let $\theta \in \mathbb{F}_q$ be an element for which

$$\theta^2 = n \tag{5.33}$$

in $\mathbb{F}_q$. Since n is quadratic non-residue modulo p, $\theta \notin \mathbb{F}_p$. Then $\{1, \theta\}$ is a basis of $\mathbb{F}_q$ over $\mathbb{F}_p$, so every element of $\mathbb{F}_q$ can be written uniquely in the form $x + \theta y$ with $x, y \in \mathbb{F}_p$. Then define the conjugate of $x + \theta y$ by

$$\overline{x + \theta y} \stackrel{\text{def}}{=} x - \theta y.$$

Then for $a, b \in \mathbb{F}_q$ we have

$$\overline{ab} = \overline{a} \cdot \overline{b},$$
$$\overline{a + b} = \overline{a} + \overline{b},$$

and

$$a\overline{a} \in \mathbb{F}_p. \tag{5.34}$$

It is easy to check that

$$\overline{x + \theta y} = (x + \theta y)^p, \tag{5.35}$$

since by using the Euler lemma for $x, y \in \mathbb{F}_p$ we have

$$\begin{aligned}(x+\theta y)^p &= x^p + \theta^p y^p = x^p + \left(\theta^2\right)^{p-1/2}\theta y^p = x + \left(\theta^2\right)^{p-1/2}\theta y \\ &= x + n^{(p-1)/2}\theta y = x + \left(\frac{n}{p}\right)\theta y = x - \theta y.\end{aligned}$$

Thus the conjugation is an automorphism of $\mathbb{F}_q$ which can be extended to an automorphism of $\overline{\mathbb{F}}_q$ by

$$\begin{aligned}\overline{\mathbb{F}}_q &\to \overline{\mathbb{F}}_q, \\ \varepsilon &\to \varepsilon^p.\end{aligned}$$

This is the Froebenius automorphism.

Lemma 5.5 *For $x, y \in \mathbb{F}_p$*

$$\gamma(x+\theta y) = \left(\frac{(x+\theta y)\overline{(x+\theta y)}}{p}\right) = \left(\frac{x^2 - ny^2}{p}\right).$$

Proof of Lemma 5.5 Using (5.35) and the Euler lemma we get

$$\begin{aligned}\gamma(x+\theta y) &= (x+\theta y)^{(q-1)/2} = (x+\theta y)^{\left(p^2-1\right)/2} \\ &= (x+\theta y)^{\left(p^2-p\right)/2}(x+\theta y)^{(p-1)/2} \\ &= \left((x+\theta y)^p\right)^{(p-1)/2}(x+\theta y)^{(p-1)/2} \\ &= \left(\overline{x+\theta y}\right)^{(p-1)/2}(x+\theta y)^{(p-1)/2} \\ &= (x-\theta y)^{(p-1)/2}(x+\theta y)^{(p-1)/2} \\ &= \left(x^2 - \theta^2 y^2\right)^{(p-1)/2} \\ &= \left(x^2 - ny^2\right)^{(p-1)/2},\end{aligned}$$

which proves Lemma 5.5.

By Lemma 5.5

$$\sum_{x\in I}\gamma(f(x)) = \sum_{x\in I}\left(\frac{f(x)\overline{f(x)}}{p}\right).$$

Since $I \subseteq \mathbb{F}_p$, if $f(x) = a_k x^k + \cdots + a_o$, then

$$\sum_{x\in I}\left(\frac{f(x)\overline{f(x)}}{p}\right) = \sum_{x\in I}\left(\frac{\left(a_k x^k + \cdots + a_o\right)\overline{\left(a_k x^k + \cdots + a_o\right)}}{p}\right)$$
$$= \sum_{x\in I}\left(\frac{\left(a_k x^k + \cdots + a_o\right)\left(\overline{a_k} x^k + \cdots + \overline{a_o}\right)}{p}\right).$$

Here the coefficients of $\left(a_k x^k + \cdots + a_o\right)\left(\overline{a_k} x^k + \cdots + \overline{a_o}\right)$ are in $\mathbb{F}_p$, since $f(x)$ can be written in the form $p(x)+\theta r(x)$ with $p(x), r(x) \in \mathbb{F}_p[x]$ and then $\overline{f(x)} = \overline{a_k} x^k + \cdots + \overline{a_o} = p(x) - \theta r(x)$ so $f(x)\overline{f(x)} = (p(x) + \theta r(x))(p(x) - \theta r(x)) = p^2(x) - nq^2(x) \in \mathbb{F}_p[x]$.

Let $b(x) = \left(a_k x^k + \cdots + a_o\right)\left(\overline{a_k} x^k + \cdots + \overline{a_o}\right)$. Then

$$\sum_{x\in I}\left(\frac{f(x)\overline{f(x)}}{p}\right) = \sum_{x\in I}\left(\frac{b(x)}{p}\right).$$

Here we need Weil's theorem. If the conditions of Lemma 5.1 and Lemma 5.2 hold, then using these lemmas we get (5.30) and (5.31) which was to be proved. So indeed, we need to prove that the conditions of Lemma 5.1 and Lemma 5.2 hold for $b(x)$, so $b(x)$ is not of the form $ch^2(x)$, with $c \in \mathbb{F}_p$, $h(x) \in \mathbb{F}_p[x]$.

Let

$$f(x) = a_k(x-\varepsilon_1)(x-\varepsilon_2)\ldots(x-\varepsilon_k)$$

where $a_k \in \mathbb{F}_q$, $\varepsilon_1, \ldots, \varepsilon_k \in \overline{\mathbb{F}}_p$. Then for $x \in \mathbb{F}_p$

$$\overline{f(x)} = \overline{a_k}\,(\overline{x} - \overline{\varepsilon_1})\cdots(\overline{x} - \overline{\varepsilon_k})$$
$$= \overline{a_k}\,(x - \varepsilon_1^p)\cdots(x - \varepsilon_k^p).$$

Then $b(x) = f(x)\overline{f(x)} = a_k\overline{a_k}(x-\varepsilon_1)\cdots(x-\varepsilon_k)(x-\varepsilon_1^p)\cdots(x-\varepsilon_k^p)$. Clearly by (5.34) we have $a_k\overline{a_k} \in \mathbb{F}_p$. The next question is that when is a product $(x-\varepsilon_1)\cdots(x-\varepsilon_k)\,(x-\varepsilon_1^p)\cdots(x-\varepsilon_k^p)$ of the form $n^2(x)$ with $n(x) \in \mathbb{F}_p[x]$. Let $\alpha_1, \alpha_2, \ldots, \alpha_t$ be the different elements among $\varepsilon_1, \ldots, \varepsilon_k$ which have odd multiplicity in the factorization of $f(x) = a_k(x-\varepsilon_1)\ldots(x-\varepsilon_k)$. Writing $g(x) = (x-\alpha_1)\ldots(x-\alpha_t)$ we get that $f(x)$ is of the form $a_k g(x)h^2(x)$ where $g(x)$ has no multiple roots and $g(x), h(x) \in \overline{\mathbb{F}}_p[x]$. Then

$$b(x) = a_k\overline{a_k}(x-\alpha_1)\ldots(x-\alpha_t)(x-\alpha_1^p)\ldots(x-\alpha_t^p)s^2(x)$$

with $s(x) \in \overline{\mathbb{F}}_p[x]$. Here $(x-\alpha_1)\ldots(x-\alpha_t)(x-\alpha_1^p)\ldots(x-\alpha_t^p)$ is of the form

$u^2(x)$ with $u(x) \in \overline{\mathbb{F}}_p[x]$ if and only if $\{\alpha_1, \alpha_2, \ldots, \alpha_t\} = \{\alpha_1^p, \alpha_2^p, \ldots, \alpha_t^p\}$. If $\{\alpha_1, \alpha_2, \ldots, \alpha_t\} = \{\alpha_1^p, \alpha_2^p, \ldots, \alpha_t^p\}$ then for every symmetric polynomial $v \in \mathbb{F}_p[x_1, x_2, \ldots, x_t]$ we have

$$v(\alpha_1, \ldots, \alpha_t) = v(\alpha_1^p, \ldots, \alpha_t^p) = v^p(\alpha_1, \ldots, \alpha_t).$$

Thus $v(\alpha_1, \ldots, \alpha_t) \in \mathbb{F}_p$. So the coefficients of $g(x) = (x - \alpha_1) \ldots (x - \alpha_t)$ are the elements of $\mathbb{F}_p$. Thus the coefficients of $h^2(x) = \frac{f(x)}{a_k g(x)}$ are in $\mathbb{F}_q$.

Let $h(x) = x^f + b_{f-1}x^{f-1} + \cdots + b_0$. We will prove by induction that $b_{f-i} \in \mathbb{F}_q$. Indeed the coefficient of x^{2f-1} in $h^2(x)$ is $2b_{f-1}$, thus $b_{f-1} \in \mathbb{F}_q$. Suppose that $b_{f-1}, b_{f-2}, \ldots, b_{f-v} \in \mathbb{F}_p$. We will prove that $b_{f-v-1} \in \mathbb{F}_p$ also holds. Indeed the coefficient of x^{2f-v-1} is of the form $2b_{f-v-1} + j(b_{f-1}, b_{f-2}, \ldots, b_{f-v})$ with $j \in \mathbb{F}_p[x_1, x_2, \ldots, x_v]$. Thus $2b_{f-v-1} + j(b_{f-1}, b_{f-2}, \ldots, b_{f-v})$ is in $\mathbb{F}_q$, and by the inductive hypothesis $j(b_{f-1}, b_{f-2}, \ldots, b_{f-v})$ is in $\mathbb{F}_q$, thus b_{f-v-1} is in $\mathbb{F}_q$. So we proved that $h(x) \in \mathbb{F}_q[x]$. Thus $b(x) = a_k\overline{a_k}(x - \varepsilon_1) \ldots (x - \varepsilon_k)(x - \overline{\varepsilon_1}) \ldots (x - \overline{\varepsilon_k})$ is of the form $cn^2(x)$ with $c \in \mathbb{F}_q$, $n(x) \in \mathbb{F}_q[x]$ if and only if $f(x)$ is of the form $cg(x)h^2(x)$ with $c \in \mathbb{F}_q$, $g(x) \in \mathbb{F}_p[x]$, $h(x) \in \mathbb{F}_q[x]$, which was to be proved.

In order to prove Theorem 5.2 we need one more lemma. Namely:

Lemma 5.6 *Let $f(x) \in \mathbb{F}_{p^2}[x]$ be an irreducible polynomial in $\mathbb{F}_{p^2}[x]$ of degree k, which is of the form*

$$f(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_0,$$

where $a_{k-1} \in \mathbb{F}_p$ but $f(x) \notin \mathbb{F}_p[x]$, so there is an $1 \leq i \leq k-2$ such that $a_i \notin \mathbb{F}_p$. Then for $d_1, d_2, \ldots, d_\ell \in \mathbb{F}_{p^2}$ we have

$$f(x + d_1)f(x + d_2) \ldots f(x + d_\ell) \notin \mathbb{F}_p[x].$$

Proof of Lemma 5.6 Every $f(x) \in \mathbb{F}_{p^2}[x]$ can be uniquely written in the form

$$f(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_0$$

with $a_i \in \mathbb{F}_{p^2}$. Then define

$$\tau(f(x)) \stackrel{\text{def}}{=} \overline{a_k}x^k + \overline{a_{k-1}}x^{k-1} + \cdots + \overline{a_0}.$$

Clearly,

$$\begin{aligned} \tau(\tau(f(x))) &= f(x) \\ \tau(f(x) + g(x)) &= \tau(f(x)) + \tau(g(x)) \\ \tau(f(x)g(x)) &= \tau(f(x))\tau(g(x)). \end{aligned}$$

Lemma 5.7 *If $f(x) \in \mathbb{F}_{p^2}[x]$ is irreducible in $\mathbb{F}_{p^2}[x]$, then $\tau(f(x)) \in \mathbb{F}_{p^2}[x]$ is also irreducible in $\mathbb{F}_{p^2}[x]$.*

Proof of Lemma 5.7 Whenever

$$\tau(f(x)) = g(x)h(x) \text{ with } g(x), h(x) \in \mathbb{F}_{p^2}[x],$$

then

$$f(x) = \tau(\tau(f(x))) = \tau(g(x))\tau(h(x)).$$

Since $f(x)$ is irreducible it follows that $\tau(f(x))$ or $\tau(g(x))$ is constant. From this follows that $f(x)$ or $g(x)$ is constant. But then $\tau(f(x))$ is irreducible.

Lemma 5.8 *If $f(x) \in \mathbb{F}_{p^2}[x]$ is an irreducible polynomial in $\mathbb{F}_{p^2}[x]$ with leading coefficient 1, but $f(x) \notin \mathbb{F}_p[x]$ then $g(x) \stackrel{def}{=} f(x)\tau(f(x))$ is in $\mathbb{F}_p[x]$ and $g(x)$ is irreducible in $\mathbb{F}_p[x]$.*

Proof of Lemma 5.8 Define n and θ as in (5.32) and (5.33). Then every $f(x) \in \mathbb{F}_{p^2}[x]$ can be uniquely written in the form

$$f(x) = a(x) + \theta b(x)$$

with $a(x), b(x) \in \mathbb{F}_p[x]$. Then

$$\tau(f(x)) = a(x) - \theta b(x).$$

Thus

$$f(x)\tau(f(x)) = (a(x) + \theta b(x))(a(x) - \theta b(x)) = a^2(x) - nb^2(x) \in \mathbb{F}_p[x].$$

Suppose that $f(x)\tau(f(x))$ is not irreducible in $\mathbb{F}_p[x]$, so

$$f(x)\tau(f(x)) = g(x)h(x) \tag{5.36}$$

with $g(x), h(x) \in \mathbb{F}_p[x]$, where the leading coefficients of $g(x)$ and $h(x)$ are 1 and $\deg g(x), \deg h(x) \geq 1$. Then (5.36) also holds in $\mathbb{F}_{p^2}[x]$ since $\mathbb{F}_p \subseteq \mathbb{F}_{p^2}$. But there is a unique factorization in $\mathbb{F}_{p^2}[x]$, and $f(x)$ and $\tau(f(x))$ are irreducible polynomials in $\mathbb{F}_{p^2}[x]$ with leading coefficients 1, thus

$$f(x) = g(x),\ \tau(f(x)) = h(x)$$

or

$$f(x) = h(x),\ \tau(f(x)) = g(x).$$

In both cases we get $f(x) \in \mathbb{F}_p[x]$, which is a contradiction.

Now we are ready to prove Lemma 5.6. Suppose that

$$f(x+d_1)\dots f(x+d_\ell) \in \mathbb{F}_p[x].$$

Let $\alpha \in \overline{\mathbb{F}}_p$ be a root of $f(x+d_1)$, then $f(\alpha+d_1)=0$, thus

$$f(\alpha+d_1)\dots f(\alpha+d_\ell)=0.$$

But then the minimal polynomial of α in $\mathbb{F}_p[x]$ divides $f(x+d_1)\dots f(x+d_\ell) \in \mathbb{F}_p[x]$. Next we determine the minimal polynomial of α in $\mathbb{F}_p[x]$. α is a root of $f(x+d_1)\tau(f(x+d_1))$ and by Lemma 5.8 this polynomial is irreducible in $\mathbb{F}_p[x]$. So the minimal polynomial of α is $f(x+d_1)\tau(f(x+d_1))$ in $\mathbb{F}_p[x]$. Thus

$$f(x+d_1)\tau(f(x+d_1)) \mid f(x+d_1)\dots f(x+d_\ell) \text{ in } \mathbb{F}_p[x].$$

But $\mathbb{F}_p[x] \subseteq \mathbb{F}_{p^2}[x]$, so

$$f(x+d_1)\tau(f(x+d_1)) \mid f(x+d_1)\dots f(x+d_\ell) \text{ in } \mathbb{F}_{p^2}[x].$$

Thus

$$\tau(f(x+d_1)) \mid f(x+d_2)\dots f(x+d_\ell) \text{ in } \mathbb{F}_{p^2}[x].$$

By Lemma 5.7, $\tau(f(x+d_1))$ is irreducible in $\mathbb{F}_{p^2}[x]$ and its leading coefficient is 1, thus by the unique factorization in $\mathbb{F}_{p^2}[x]$, there is an $2 \le i \le \ell$ such that

$$\tau(f(x+d_1)) = f(x+d_i).$$

Without the loss of generality we may assume

$$\tau(f(x+d_1)) = f(x+d_2). \tag{5.37}$$

By the definition of $f(x)$ it is of the form

$$f(x) = a_k x^k + a_{k-1}x^{k-1} + \dots + a_0$$

where $a_k = 1$, $a_{k-1} \in \mathbb{F}_p[x]$. Then

$$\begin{aligned} f(x+d_1) = \sum_{i=0}^{k} \Bigg(&\binom{k}{i} a_k d_1^{k-i} + \binom{k-1}{i} a_{k-1} d_1^{k-1-i} \\ &+ \binom{k-2}{i} a_{k-2} d_1^{k-2-i} + \dots + \binom{i}{i} a_i \Bigg) x^i \end{aligned}$$

and

$$f(x+d_2)=\sum_{i=0}^{k}\left(\binom{k}{i}a_kd_2^{k-i}+\binom{k-1}{i}a_{k-1}d_2^{k-1-i}\right.$$
$$\left.+\binom{k-2}{i}a_{k-2}d_2^{k-2-i}+\cdots+\binom{i}{i}a_i\right)x^i$$

By the definition of τ

$$\tau(f(x+d_1))=\sum_{i=0}^{k}\left(\binom{k}{i}\overline{a_k}\overline{d_1}^{k-i}+\binom{k-1}{i}\overline{a_{k-1}}\overline{d_1}^{k-1-i}\right.$$
$$\left.+\binom{k-2}{i}\overline{a_{k-2}}\overline{d_1}^{k-2-i}+\cdots+\binom{i}{i}\overline{a_i}\right)x^i.$$

By (5.37) we get that for $0\le i\le k$

$$\binom{k}{i}\overline{a_k}\overline{d_1}^{k-i}+\binom{k-1}{i}\overline{a_{k-1}}\overline{d_1}^{k-1-i}+\binom{k-2}{i}\overline{a_{k-2}}\overline{d_1}^{k-2-i}+\cdots+\binom{i}{i}\overline{a_i}$$
$$=\binom{k}{i}a_kd_2^{k-i}+\binom{k-1}{i}a_{k-1}d_2^{k-1-i}+\binom{k-2}{i}a_{k-2}d_2^{k-2-i}+\cdots+\binom{i}{i}a_i. \tag{5.38}$$

For $i=k-1$ this gives

$$\binom{k}{k-1}\overline{a_k}\overline{d_1}+\binom{k-1}{k-1}\overline{a_{k-1}}=\binom{k}{k-1}a_kd_2+\binom{k-1}{k-1}a_{k-1}. \tag{5.39}$$

By the conditions of Lemma 5.6 we have $a_k=1$ and $a_{k-1}\in\mathbb{F}_p$, thus $\overline{a_k}=a_k$ and $\overline{a_{k-1}}=a_{k-1}$, so from (5.39)

$$\overline{d_1}=d_2 \tag{5.40}$$

follows.

Next we prove by induction that $a_i\in\mathbb{F}_p$. Indeed, by the conditions of Lemma 5.6, a_k and $a_{k-1}\in\mathbb{F}_p$. Next suppose that $a_k,a_{k-1},\dots,a_{i+1}\in\mathbb{F}_p$. We will prove that $a_i\in\mathbb{F}_p$. Indeed by $a_k,a_{k-1},\dots,a_{i+1}\in\mathbb{F}_p$ then

$$a_k=\overline{a_k},a_{k-1}=\overline{a_{k-1}},\dots,a_{i+1}=\overline{a_{i+1}}$$

By this, (5.38) and (5.40) we get

$$\overline{a_i} = a_i,$$

so $a_i \in \mathbb{F}_p$ which was to be proved. Thus $a_k, a_{k-1}, \ldots, a_0 \in \mathbb{F}_p$. But then $f(x) \in \mathbb{F}_p[x]$, which is contradiction. Thus we proved Lemma 5.6.

Next we return to the proof of Theorem 5.2. For $N = 2$ the theorem is trivial. For $N \geq 3$ let p be an odd prime for which

$$N^{1/2} < p < 2N^{1/2}. \tag{5.41}$$

(By Chebysev's theorem such a prime p exists.) Let $q = p^2$ and let n be a quadratic non-residue modulo p, so $\left(\frac{n}{p}\right) = -1$. Let $\theta \in \mathbb{F}_{p^2}$ be a number for which

$$\theta^2 = n$$

in $\mathbb{F}_q$. Then $\{1, \theta\}$ is a basis of $\mathbb{F}_q$ over $\mathbb{F}_p$.

Let $f(x)$ be an irreducible polynomial of degree $k \geq 2$ which is of the form

$$f(x) = x^k + a_{k-2}x^{k-2} + \cdots + a_0$$

(so the coefficient of the term x^{k-1} is 0) but

$$f(x) \notin \mathbb{F}_p[x].$$

By (5.14) the number of such polynomials is

$$R \stackrel{\text{def}}{=} N_{p^2}(k) - N_p(k) \geq \frac{1}{2k}p^{2k-1} - \frac{1}{k}p^{k-1} > 0,$$

thus such a polynomial exists, indeed.

Define the binary lattice $\eta : I_p^2 \to \{-1, +1\}$ by

$$\eta(\underline{x}) = \eta((x_1, x_2)) = \gamma(f(x_1 + \theta x_2)).$$

Lemma 5.9

$$Q_\ell(\eta) \leq k\ell\left(p(1 + \log p)^2\right) \ll k\ell N^{1/2}(\log N)^2. \tag{5.42}$$

Proof of Lemma 5.9 We remark that this construction is a shifted version of the construction in Theorem 1 in [79]. We cannot use Theorem 1 in [79] because none of the conditions a), b) and c) holds in Theorem in [79]. However, similarly to the proof of Theorem 1 in [79], it is easy to prove that

(5.42) holds:

Write $\boldsymbol{d}_i = (d_1^{(i)}, d_2^{(i)})$ (for $i = 1, \dots, \ell$), and consider the general term of the n-fold sum in (1.6):

$$\sum_{\mathbf{x}\in B} \eta(\mathbf{x}+\mathbf{d_1})\dots\eta(\mathbf{x}+\mathbf{d}_\ell)$$
$$= \sum_{j_1=0}^{[t_1/b_1]} \sum_{j_2=0}^{[t_2/b_2]} \eta\big((j_1b_1+d_1^{(1)}, j_2b_2+d_2^{(1)})\big)\dots\eta\big((j_1b_1+d_1^{(\ell)}, j_2b_2+d_2^{(\ell)})\big), \tag{5.43}$$

where B is a box-lattice of form

$$B = \{\mathbf{x} = (j_1b_1, j_2b_2) :\ 0 \le j_1b_1 \le t_1 (< p),\ 0 \le j_2b_2 \le t_2 (< p),\ j_1, j_2 \in \mathbb{N}\}.$$

Now write

$$z = j_1b_1 + j_2b_2\theta \tag{5.44}$$

so that z belongs to the box

$$B' = \{j_1b_1 + j_2b_2\theta :\ 0 \le j_1b_1 \le t_1,\ 0 \le j_2b_2 \le t_2,\ j_1, j_2 \in \mathbb{N}\}, \tag{5.45}$$

and set

$$z_i = d_1^{(i)} + d_2^{(i)}\theta. \tag{5.46}$$

If $z \in B'$ then $f(z+z_1)\dots f(z+z_k) \neq 0$, and by the definition of η and the multiplicativity of γ, the product in (5.43) is

$$\gamma\big(f(z+z_1)\big)\dots\gamma\big(f(z+z_k)\big) = \gamma\big(f(z+z_1)\dots f(z+z_k)\big).$$

Then from (5.43) we get

$$\sum_{\mathbf{x}\in B} \eta(\mathbf{x}+\mathbf{d_1})\dots\eta(\mathbf{x}+\mathbf{d}_\ell) = \sum_{z\in B'} \gamma\left(f(z+z_1)\dots f(z+z_\ell)\right) \tag{5.47}$$

Now we need the following result of Winterhof:

Lemma 5.10 (Winterhof) *Let p be a prime, $n \ge 1$ be an integer, $q = p^n$ and let $v_1, v_2, \dots, v_n$ be a basis of the vector space $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$. Let χ be a multiplicative character of $\mathbb{F}_q$ of order $d > 1$, $f \in \mathbb{F}_q[x]$ be a nonconstant polynomial which is not a d-th power of a polynomial of $\overline{\mathbb{F}}_p[x]$ and which has m distinct zeros in its splitting field over $\mathbb{F}_q$, and $k_1, \dots, k_n$ are non-negative integers with $k_1 \le p, \dots, k_n \le p$, then, writing $B = \Big\{\sum_{i=1}^{n} x_iv_i :\ 0 \le j_i < k_i\Big\}$,*

we have

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

Proof of Lemma 5.10 This is a part of Theorem 2 in [107] (where its proof was based on A. Weil's theorem [106]).

Write $h(z) = f(z+z_1)\dots f(z+z_k)$. Then in order to prove (5.42), it suffices to show:

Lemma 5.11 *$h(x)$ has at least one zero in $\overline{\mathbb{F}}_p$ whose multiplicity is odd.*

Proof of Lemma 5.11 Since $z_1, z_2, \dots, z_\ell$ are different the irreducible polynomials $f(z+z_1), \dots, f(z+z_\ell)$ are different. (Indeed, the coefficients of x^{k-1} are different.) So $h(x)$ has a zero in $\overline{\mathbb{F}}_q$ whose multiplicity is odd. Thus $h(x)$ cannot be the constant multiple of a square. Applying Lemma 5.10 we obtain from (5.47)

$$\sum_{x\in B}\eta(\mathbf{x}+\mathbf{d_1})\dots\eta(\mathbf{x}+\mathbf{d}_\ell) \ll k\ell p(1+\log p)^2 \ll k\ell N^{1/2}(\log N)^2,$$

which was to be proved.

In [50] we reduced the two dimensional case to the one dimensional one by the following way: To any 2-dimensional binary p-lattice

$$\eta(\underline{x}):\ I_p^2 \to \{-1,+1\} \tag{5.48}$$

we may assign a unique binary sequence $E_{p^2} = E_{p^2}(\eta) = (e_1, e_2, \dots, e_{p^2}) \in \{-1,+1\}^{p^2}$ by taking the first (from the bottom) row of the lattice (5.48) then we continue the binary sequence by taking the second row of the lattice, then the third row follows, etc.; in general, we set

$$e_{ip+j} = \eta((j-1,i)) = \gamma(f((j-1)+i\theta))$$
$$\text{for } i = 0,1,\dots,p-1,\ j = 1,2,\dots,p.$$

Thus we obtain a sequence of length p^2

$$E_{p^2} \stackrel{\text{def}}{=} (e_1, e_2, \dots, e_{p^2}).$$

Now $N < p^2 < 4N$. Consider the first N elements of E_{p^2}, they form a sequence of length N:

$$E_N \stackrel{\text{def}}{=} (e_1, e_2, \dots, e_N).$$

We state that E_N satisfies the conditions of the lemma.

First we estimate $|V(E_N, M, D)|$. Let $m_p(x)$ denote the unique integer x for which

$$m_p(x) \equiv x \pmod{p}, \quad 0 \le m_p(x) < p.$$

Then

$$e_{n+d_i} = e_{\left[\frac{n+d_i-1}{p}\right]p+m_p(n+d_i-1)+1}$$

and so

$$\begin{aligned} e_{n+d_i} &= \eta\left(m_p(n+d_i-1), \left[\frac{n+d_i-1}{p}\right]\right) \\ &= \gamma\left(f\left(n+d_i-1+\left[\frac{n+d_i-1}{p}\right]\theta\right)\right). \end{aligned} \tag{5.49}$$

If $1 \le n \le M < p$ then $\left[\frac{n+d_i-1}{p}\right]$ may take two different values, namely q_i and q_i+1. Indeed, define q_i and s_i by $d_i = q_i p + s_i$ where $0 \le s_i \le p-1$. Then

$$\begin{aligned} \left[\frac{n+d_i-1}{p}\right] &= \left[\frac{q_i p + s_i + n - 1}{p}\right] = q_i + \left[\frac{s_i+n-1}{p}\right] \\ &= \begin{cases} q_i & \text{if } n \le p - s_i, \\ q_i+1 & \text{if } n > p - s_i. \end{cases} \end{aligned}$$

Moreover there exists a number $b_i = \min\{M, p-s_i\}$ such that for $n \le b_i \le M$ $\left[\frac{n+d_i-1}{p}\right] = q_i$ and for $b_i < n \le M$ we have $\left[\frac{n+d_i-1}{p}\right] = q_i+1$. Let $I_i = [0, b_i]$, $J_i = [b_i+1, M]$ (if $b_i = M$ then $J_i = \emptyset$).

Then $\{1, b_1+1, b_2+1, \dots, b_\ell+1, M+1\}$ is a multiset which contains integers $1 = c_1 < c_2 < \cdots < c_m = M+1$ with $m \le \ell+2$. Then $[0, M] = \cup_{j=1}^{m-1}[c_j, c_{j+1}-1]$.

$$V(E_N, M, D) = \sum_{n=1}^{M} e_{n+d_1} \dots e_{n+d_\ell} = \sum_{j=1}^{m-1} \sum_{n \in [c_j, c_{j+1}-1]} e_{n+d_1} \dots e_{n+d_\ell} \tag{5.50}$$

By the definition of the c_j's, $c_j < b_i+1 < c_{j+1}$ is not possible, thus $c_{j+1}-1 \le b_i$ or $b_i \le c_j - 1$, so $[c_j, c_{j+1}-1] \subseteq [0, b_i]$ or $[c_j, c_{j+1}-1] \subseteq [b_i+1, M]$. Each interval $[c_j, c_{j+1}-1]$ is either $\subseteq I_i$ or $\subseteq J_i$ for every $1 \le i \le \ell$. Thus for every $d_1, d_2, \dots, d_\ell$ and for every interval $[c_j, c_{j+1}-1]$ there exist fixed numbers $h_1, h_2, \dots, h_\ell$ (depending only on $d_1, d_2, \dots, d_\ell$ and j) such that for

$n \in [c_j, c_{j+1} - 1]$

$$\begin{aligned} e_{n+d_1} e_{n+d_2} \dots e_{n+d_\ell} &= \gamma\left(f(n + d_1 - 1 + h_1\theta)\right) \gamma\left(f(n + d_2 - 1 + h_2\theta)\right) \dots \\ &\quad \gamma\left(f(n + d_\ell - 1 + h_\ell\theta)\right) \\ &= \gamma\Big(f(n + d_1 - 1 + (h_1 + 1)\theta) f(n + d_2 - 1 + (h_2 + 1)\theta) \\ &\quad \dots f(n + d_\ell - 1 + (h_\ell + 1)\theta)\Big). \end{aligned}$$

Hence

$$\begin{aligned} &\sum_{n \in [c_j, c_{j+1} - 1]} e_{n+d_1} \dots e_{n+d_\ell} \\ &= \sum_{n \in [c_j, c_{j+1} - 1]} \gamma\left(f(n + d_1 - 1 + h_1\theta) \cdots f(n + d_\ell - 1 + h_\ell\theta)\right). \end{aligned} \tag{5.51}$$

Next we prove that the irreducible polynomials $f(x + d_1 - 1 + h_1\theta), \cdots, f(x + d_\ell - 1 + h_\ell\theta)$ are different. Since if $i \neq j$ and

$$f(x + d_i - 1 + h_i\theta) = f(x + d_j - 1 + h_j\theta),$$

then

$$h_i \equiv h_j \pmod{p} \text{ and } d_i \equiv d_j \pmod{p}. \tag{5.52}$$

This can be proved by considering the coefficient x^{k-1} in the polynomials $f(x + d_i - 1 + h_i\theta)$ and $f(x + d_j - 1 + h_j\theta)$. By (5.49) we have $h_i = \left[\frac{n+d_i-1}{p}\right]$ and $h_j = \left[\frac{n+d_j-1}{p}\right]$ for $n \in [c_j, c_{j+1} - 1]$. $h_i \equiv h_j \pmod{p}$, by $0 \leq h_i = \left[\frac{n+d_i-1}{p}\right]$, $h_j = \left[\frac{n+d_j-1}{p}\right] < p$ then $h_i = h_j$. So for $n \in [c_j, c_{j+1} - 1]$

$$\left[\frac{n + d_i - 1}{p}\right] = \left[\frac{n + d_j - 1}{p}\right] \tag{5.53}$$

By (5.52),

$$n + d_i - 1 \equiv n + d_j - 1 \pmod{p}. \tag{5.54}$$

We get from (5.53) and (5.54) that

$$n + d_i - 1 = n + d_j - 1$$

So

$$d_i = d_j$$

which is a contradiction. Thus

$$q_j(x) \stackrel{\text{def}}{=} f(x+d_1-1+h_1\theta)f(x+d_2-1+h_2\theta)\cdots f(x+d_\ell-1+h_\ell\theta) \quad (5.55)$$

has no multiple root. Here by definition $f(x) \notin \mathbb{F}_p[x]$, by using Lemma 5.6 $q_j(x) \notin \mathbb{F}_p[x]$ and it has no multiple root. Thus it is not of the form $cg(x)h^2(x)$ with $c \in \mathbb{F}_p$, $g(x) \in \mathbb{F}_p[x]$, $h(x) \in \mathbb{F}_q[x]$. By the triangle inequality, Lemma 5.4, (5.50), (5.51) and (5.55) we get

$$\begin{aligned} |V(E_N, M, D)| &\le \sum_{j=1}^{m-1} \left| \sum_{n\in[c_j, c_{j+1}-1]} \gamma(q_j(n)) \right| \ll \sum_{j=1}^{m-1} (\deg\, q_j) p^{1/2} \log p \\ &\ll \ell(\deg\, q_j) p^{1/2} \log p \ll \ell^2 k p^{1/2} \log p \\ &\ll \ell^2 k N^{1/4} \log N \end{aligned}$$

which proves (5.5), since we may choose $\deg f = k$ as $k = 4$.

Next we prove (5.6). By Lemma 5.9 we have $Q_\ell(\eta) \ll k\ell N^{1/2}(\log N)^2$. By Theorem 5.3 (which we will prove later) $C_\ell(E_N) \ll C_\ell(E_{p^2}) \ll k\ell^2 N^{1/2}(\log N)^2 \ll k\ell N^{1/2}(\log N)^2$, since k can be chosen as $k = 4$ this proves (5.6).

Next we prove (5.7). We split E_N into $\left[\frac{N-1}{p}\right] + 1$ different subsequences: $E^{(1)} = (e_1, e_2, \dots, e_p)$, $E^{(2)} = (e_{p+1}, e_{p+2}, \dots, e_{2p})$,..., $E^{\left(\left[\frac{N-1}{p}\right]+1\right)} = (e_{\left(\left[\frac{N-1}{p}\right]p+1\right)}, \dots, e_N)$. By the triangle-inequality

$$W(E_N) \le \sum_{j=1}^{\left[\frac{N-1}{p}\right]+1} W(E_j). \quad (5.56)$$

Here $E_j = (e_{(j-1)p+1}, \dots, e_{jp}) = (f_1, f_2, \dots, f_p)$ for $1 \le j \le \left[\frac{N-1}{p}\right]$ and $E_j = (e_{(j-1)p+1}, \dots, e_N) = (f_1, f_2, \dots, f_{N-(j-1)p})$ for $j = \left[\frac{N-1}{p}\right] + 1$.

In [78] Mauduit and Sárközy proved that $W(E_N) \le \sqrt{NC_2(E_N)}$. By this and using (5.6) for $\ell = 2$ we get (5.7), which completes the proof of Theorem 5.2. We also remark that by using the same argument and (5.8) we get (5.9) in Corollary 5.2.

Proof of Theorem 5.3 For $x \in \mathbb{Z}$ let

$$x = r_N(x)N + m_N(x)$$

where $m_N(x) \equiv x \pmod N$, $0 \le m_N(x) \le N-1$ and $r_N(x) = \left[\frac{x}{N}\right]$.

By definition

$$e_{xN+y+1} = \eta(y, x) \text{ for } 0 \le x \le N-1,\ 0 \le y \le N-1$$

and thus

$$e_n = \eta(m_N(n-1), r_N(n-1)).$$

Then for $1 \le i \le \ell$

$$e_{n+d_i} = \eta(m_N(n+d_i-1), r_N(n+d_i-1)). \tag{5.57}$$

Here

$$n + d_i - 1 = (r_N(n-1) + r_N(d_i))N + m_N(n-1) + m_N(d_i).$$

Thus if $0 \le m_N(n-1) + m_N(d_i) \le N-1$ then

$$\begin{aligned} r_N(n+d_i-1) &= r_N(n-1) + r_N(d_i) \\ m_N(n+d_i-1) &= m_N(n-1) + m_N(d_i) \end{aligned}$$

and if $N \le m_N(n-1) + m_N(d_i)$ then

$$\begin{aligned} r_N(n+d_i-1) &= r_N(n-1) + r_N(d_i) + 1 \\ m_N(n+d_i-1) &= m_N(n-1) + m_N(d_i) - N. \end{aligned}$$

Thus we get that there exists an $a_i \stackrel{\text{def}}{=} N - 1 - m_N(d_i)$ such that for $m_N(n-1) \le a_i$

$$\begin{aligned} r_N(n+d_i-1) &= r_N(n-1) + r_N(d_i) \\ m_N(n+d_i-1) &= m_N(n-1) + m_N(d_i) \end{aligned} \tag{5.58}$$

and for $a_i + 1 \le m_N(n-1)$

$$\begin{aligned} r_N(n+d_i-1) &= r_N(n-1) + r_N(d_i) + 1 \\ m_N(n+d_i-1) &= m_N(n-1) + m_N(d_i) - N. \end{aligned} \tag{5.59}$$

Then $\{1, a_1+1, a_2+1, \ldots, a_\ell+1, m_N(M-1)+1, N\}$ is a multiset which contains integers $1 = c_1 < c_2 < \cdots < c_m \le N$ where $m \le \ell + 3$. By (5.58) and (5.59) we get that for $c_j \le n \le c_{j+1} - 1$ there exist numbers $b_{i,j}$ and $f_{i,j}$

such that

$$r_N(n+d_i-1) = r_N(n) + r_N(d_i-1) + b_{i,j}$$
$$m_N(n+d_i-1) = m_N(n) + m_N(d_i-1) - f_{i,j} \tag{5.60}$$

where $b_{i,j} \in \{0,1\}$ and $f_{i,j} \in \{0,N\}$. Moreover, if $b_{i,j}=0$ then $f_{i,j}=0$ and if $b_{i,j}=1$ then $f_{i,j}=N$. Now

$$[0,M] =$$
$$= \{n = TN+x+1:\ T=0,1,\dots,\left[\frac{M-1}{N}\right],\ x=0,1,\dots,m_N(M-1)\}$$
$$\cup\{n=TN+x+1:\ T=0,1,\dots,\left[\frac{M-1}{N}\right]-1,\ x=m_N(M-1)+1,$$
$$\dots,N-1\}$$

Thus

$$[0,M] = \cup_{j=1}^{m-1}\{n:\ n = r_N(N-1)N + m_N(n-1)+1,$$
$$c_j \le m_N(n-1) \le c_{j+1}-1,\ r_N(n-1)\in\{0,1,2,\dots,T_j\}\} \tag{5.61}$$

where $T_j = \left[\frac{M-1}{N}\right]$ if $c_{j+1} \le m_N(M-1)+1$ and $T_j = \left[\frac{M-1}{N}\right]-1$ if $m_N(M-1)+1 \le c_j$. (Since $m_N(M-1)+1 \in \{c_1,c_2,\dots,c_m\}$ and $c_1<c_2<\cdots<c_m$ thus $c_j < m_N(M-1)+1 < c_{j+1}$ is not possible.) By this, (5.57) and (5.58)

$$V(E_N,M,D) = \sum_{n=1}^{M} e_{n+d_1}\dots e_{n+d_\ell} = \sum_{j=1}^{m-1}\sum_{\substack{c_j\le m_N(n-1)\le c_{j+1}-1\\ 1\le n\le M}} e_{n+d_1}\dots e_{n+d_\ell}$$
$$= \sum_{j=1}^{m-1}\sum_{\substack{c_j\le m_N(n-1)\le c_{j+1}-1\\ 1\le n\le M}}$$
$$\prod_{i=1}^{\ell}\eta(m_N(n-1)+m_N(d_i)-f_{i,j}, r_N(n-1)+r_N(d_i)+b_{i,j})$$

By (5.61)

$$\{(m_N(n-1), r_N(n-1)):\ 1\le n\le M \text{ and } c_j \le m_N(n-1)\le c_{j+1}-1\} =$$
$$\{(x,y):\ 0\le x\le T_j \text{ and } c_j\le y\le c_{j+1}-1\}.$$

Using this and (5.60) we get

$$V(E_N, M, D) = \sum_{j=1}^{m-1} \sum_{x=0}^{T_j} \sum_{y=c_j}^{c_{j+1}-1} \prod_{i=1}^{\ell} \eta(x + m_N(d_i) - f_{i,j}, y + r_N(d_i) + b_{i,j}) \le (m-1)Q_\ell(\eta) \le (\ell+2)Q_\ell(\eta)$$

which was to be proved. Here we used the fact that the pairs $(m_N(d_i) - f_{i,j}, r_N(d_i) + b_{i,j})$ are different for fixed j as i runs over $1, 2, \ldots, \ell$. Indeed if

$$(m_N(d_{i_1}) - f_{i_1,j},\ r_N(d_{i_1}) + b_{i_1,j}) = (m_N(d_{i_2}) - f_{i_2,j},\ r_N(d_{i_2}) + b_{i_2,j}),$$

then

$$N(r_N(d_{i_1}) + b_{i_1,j}) + m_N(d_{i_1}) - f_{i_1,j} = N(r_N(d_{i_2}) + b_{i_2,j}) + m_N(d_{i_2}) - f_{i_2,j}.$$

Since if $b_{i,j} = 0$ then $f_{i,j} = 0$ and if $b_{i,j} = 1$ then $f_{i,j} = N$, from this we get

$$\begin{aligned} N r_N(d_{i_1}) + m_N(d_{i_1}) &= N r_N(d_{i_2}) + m_N(d_{i_2}) \\ d_{i_1} &= d_{i_2} \end{aligned}$$

which is a contradiction.

6 On Legendre symbol lattices (the non-degenerate case)

Pseudorandom binary sequences have many important applications. In particular, they are used as a key stream in the classical stream cipher called the Vernam cipher.

In one dimension, hence in the case of binary sequences, many good constructions have been given. Typically, the really good constructions involve $\mathbb{F}_p$, additive or multiplicative characters and polynomials, and the crucial tool in the estimation of the pseudorandom measures is Weil's theorem. Unfortunately, this approach in its original form does not readily apply in higher dimensions. The difficulty is that in n dimensions constructions involving $\mathbb{F}_p$, characters and polynomials $f(x_1, x_2, \ldots, x_n) \in \mathbb{F}_p[x_1, x_2, \ldots, x_n]$, lead naturally to the n-dimensional analogues of Weil's theorem. In particular they lead to the theorem of Deligne. While Fouvry and Katz [24] have simplified the requirements for applying Deligne's theorem the inconvenient assumption of nonsingularity is still required in order to obtain sharp bounds.

In spite of these difficulties, in [65] and [80] good n-dimensional constructions were presented. In these papers the authors got around the difficulty described above in the following way. Finite fields $\mathbb{F}_q$ with $q = p^n$ and polynomials $G(x) \in \mathbb{F}_q[x]$ are considered. Character sums involving $G(x)$ and characters of $\mathbb{F}_q$ can be estimated by Weil's theorem so that no nonsingularity assumption is needed. On the other hand, if $e_1, e_2, \ldots, e_n$ is a basis in $\mathbb{F}_q$, then every $x \in \mathbb{F}_q$ has a unique representation in the form $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$ with $x_1, x_2, \ldots, x_n \in \mathbb{F}_p$. Then $g(x_1, x_2, \ldots, x_n) = G(x_1e_1 + x_2e_2 + \cdots + x_ne_n) \in \mathbb{F}_q[x_1, x_2, \ldots, x_n]$ is a well-defined polynomial, and the estimate of n-fold character sums involving $g(x_1, x_2, \ldots, x_n)$ can be reduced to the estimate of character sums over $\mathbb{F}_q$ involving G, so that Weil's theorem can be used. (This principle goes back to Davenport and Lewis [20].)

This detour enables one to give sharp upper bounds, but it also has considerable disadvantages. In particular, in this way we get rather artificial constructions. More natural constructions cannot be tested with this approach. Secondly, the implementation of these artificial constructions is more complicated. Thus one might like to look for a trade-off between applicability of the method and sharpness of the result, in other words, for a method which is much more flexible and applicable at the expense of providing weaker but still nontrivial upper bounds. We will show that in the case when $n = 2$, there is such a method, based on the techniques introduced by Gyarmati and Sárközy [60] to estimate certain related character sums. This method allows

us to give a simple description of the exceptional polynomials, see Section 6.1. But the price paid for the flexibility of this method is that the upper bounds are not optimal usually. For a two dimensional p-lattice they are, up to logarithmic factors, $p^{3/2}$ instead of the optimal bound of p. On the other hand, they improve on the trivial bound of p^2 considerably. Here we mention that in Section 7 we will be able to show that for a certain (rather special) family of polynomials the finite field construction presented in [79] is equivalent to a Legendre symbol construction of type (6.2). Thus in this case we obtain a family of binary lattices which combines the advantages of the two constructions: as in [79] we have optimal bounds, and as a Legendre symbol construction it can be implemented fast and easily.

In Sections 6 and 7 I present results from [61] and [62], where with my coauthors Cameron L. Stewart and András Sárközy we studied a construction based on the Legendre symbol:

In one dimension the best and most intensively studied construction is based on the use of the Legendre symbol, see [31], [64], [77], [98]. Let p be a prime, $f(x) \in \mathbb{F}_p[x]$ be a polynomial, and define the sequence $E_p = (e_1, \dots, e_p)$ by

$$e_n = \begin{cases} \left(\frac{f(n)}{p}\right) & \text{if } (f(n), p) = 1, \\ +1 & \text{if } p \mid f(n). \end{cases} \tag{6.1}$$

We will identify the elements of $\mathbb{F}_p$ with the residue classes modulo p, and we will not distinguish between the residue classes and their representing elements. The natural two dimensional extension of this construction is the following.

Construction 6.1 (Gyarmati, Sárközy, Stewart) *Let p be an odd prime, $f(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ be a polynomial in two variables. Define $\eta : I_p^2 \to \{-1, +1\}$ by*

$$\eta(x_1, x_2) = \begin{cases} \left(\frac{f(x_1,x_2)}{p}\right) & \text{if } (f(x_1, x_2), p) = 1, \\ +1 & \text{if } p \mid f(x_1, x_2). \end{cases} \tag{6.2}$$

First, in Section 6.1, we will show that in two dimensions there are new difficulties arising, and there are many "bad" polynomials $f(x_1, x_2)$. Then, in Section 6.2, we will formulate Theorem 6.1, our main result. We will also present several sufficient criteria for a polynomial $f(x_1, x_2)$ for which the corresponding binary p-lattice (6.2) possesses strong pseudorandom properties. The rest of this section will be devoted to the proof of this main result.

In Section 7 we will study (6.2) in the case when $f(x_1, x_2)$ is one of the degenerate polynomials described in Section 6.1. Moreover, we will also study

implementation problems related to some constructions based on Theorem 6.1.

6.1 Negative examples

In this section we will present examples of polynomials $f(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ for which the binary p-lattice defined in (6.2) has weak pseudo-random properties.

Example 6.1 (Gyarmati, Sárközy, Stewart) If

$$f(x_1, x_2) = c\,(g(x_1, x_2))^2$$

with $c \in \mathbb{F}_p$, $g(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$, then every element of the lattice defined in (6.2) is $\left(\frac{c}{p}\right)$ except the zeros of $f(x_1, x_2)$. It follows that if the degree of $f(x_1, x_2)$ is not very large, then $Q_1(\eta)$ is large.

Example 6.2 (Gyarmati, Sárközy, Stewart) If $f(x_1, x_2) = g(x_1)$ with a polynomial $g(x) \in \mathbb{F}_p[x]$ of one variable, then we have

$$\eta(x_1, x_2)\eta(x_1, x_2 + 1) = \left(\frac{g(x_1)}{p}\right)\left(\frac{g(x_1)}{p}\right) = +1$$

(except the zeros of $g(x_1)$) from which it follows that $Q_2(\eta)$ is large.

Example 6.3 (Gyarmati, Sárközy, Stewart) If $f(x_1, x_2) = g(x_1)h(x_2)$ with polynomials $g(x), h(x) \in \mathbb{F}_p[x]$, then it can be shown by a little computation that $Q_4(\eta)$ is large.

The polynomials $f(x_1, x_2)$ occurring in Examples 6.1-6.3 are special cases of the following:

Definition 6.1 (Gyarmati, Sárközy, Stewart) *The polynomial $f(x_1, x_2)$ is called* degenerate *if it is of the form*

$$f(x_1, x_2) = \left(\prod_{j=1}^{r} f_j(\alpha_j x_1 + \beta_j x_2)\right) g(x_1, x_2)^2, \tag{6.3}$$

where $\alpha_j, \beta_j \in \mathbb{F}_p$, $f_j(x) \in \mathbb{F}_p[x]$ for $j = 1, \ldots, r$, and $g(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$.

A polynomial $f \in \mathbb{F}_p[x_1, x_2]$ which can be expressed in the form (6.3) is said to be degenerate and otherwise it is said to be non-degenerate.

As Examples 6.1, 6.2 and 6.3 show, if f is degenerate then it may be that the associated binary p-lattice (6.2) has weak pseudorandom properties. We shall analyse the situation when f is degenerate in more details in Section 7. In the balance of this section we shall restrict our attention to binary p-lattices (6.2) for which f is non-degenerate.

6.2 Sufficient conditions

In one dimension Goubin, Mauduit and Sárközy [31] gave sufficient conditions on the polynomial $f(x)$ to guarantee small pseudorandom measures. Let $\overline{\mathbb{F}}_p$ denote an algebraic closure of $\mathbb{F}_p$.

Theorem 6.A (Goubin, Mauduit, Sárközy) *Let $f(x) \in \mathbb{F}_p[x]$ be a polynomial of degree $k(> 0)$ which has no multiple zero in $\overline{\mathbb{F}}_p$. Define the sequence $E_p \in \{-1, +1\}^p$ by (6.1). Then $W(E_p)$, the "well-distribution measure" of E_p, satisfies*

$$W(E_p) < 10kp^{1/2} \log p.$$

Moreover assume that one of the following 3 conditions holds:
a) $\ell = 2$,
b) 2 is a primitive root modulo p,
c) $(4k)^\ell < p$ or $(4\ell)^k < p$,
Then $C_\ell(E_p)$, "the correlation measure of order ℓ," satisfies

$$C_\ell(E_p) \le 10k\ell p^{1/2} \log p.$$

(See [77] for the definition of well-distribution measure and correlation measure.)

We extend their result to the 2 dimensional case:

Theorem 6.1 (Gyarmati, Sárközy, Stewart) *Let $f(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ be a polynomial of degree k. Suppose that $f(x_1, x_2)$ cannot be expressed in the form (6.3) and one of the following 5 conditions holds:*
a) $f(x_1, x_2)$ is irreducible in $\mathbb{F}_p[x_1, x_2]$,
b) $\ell = 2$,
c) 2 is a primitive root modulo p,
d) $4^{k+\ell} < p$,
e) ℓ and the degree of the polynomial $f(x_1, x_2)$ in x_1 (or in x_2) are odd.
Then for the binary p-lattice η defined in (6.2) we have

$$Q_\ell(\eta) < 11k\ell p^{3/2} \log p.$$

The rest of this section is devoted to the proof of this theorem.

6.3 Proof of Theorem 6.1

For $k > p^{1/2}/10$ the theorem is trivial. Thus we may suppose that

$$k \leq p^{1/2}/10. \tag{6.4}$$

Similarly, we may suppose that

$$k^2 + \ell^2 < p, \tag{6.5}$$

otherwise the theorem is trivial since

$$4k^2\ell^2 > k^2 + \ell^2 \geq p,$$

and so

$$10k\ell p^{3/2} \log p > p^2.$$

Lemma 6.1 *If $\mathbb{F}$ is a field, then in $\mathbb{F}[x_1, x_2, \ldots, x_n]$ every polynomial has a factorization into irreducible polynomials which is unique apart from constant factors and reordering.*

Proof of Lemma 6.1 See, for example [93, Theorem 207].

If $f(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$, then we will also write $f(x_1, x_2) = f(\mathbf{x})$ with $\mathbf{x} = (x_1, x_2)$.

Lemma 6.2 (Gyarmati, Sárközy, Stewart) *Let $p \geq 5$ be a prime and χ be a multiplicative character of order d. Suppose that $h(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ is not of the form $cg(x_1, x_2)^d$ with $c \in \mathbb{F}_p$, $g(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$. Let k be the degree of $h(x_1, x_2)$. Then we have*

$$\sum_{\mathbf{x} \in B} \chi\left(h(\mathbf{x})\right) < 10kp^{3/2} \log p$$

for every 2 dimensional box p-lattice $B \subseteq I_p^2$.

We remark that the upper bound in the lemma is nearly sharp: it is easy to see that there are polynomials $h(x_1, x_2)$ of the form $h(x_1, x_2) = f(x_1)$ (so that $h(x_1, x_2)$ depends only one of the two variables) for which the left hand side of the inequality in the lemma with $\mathbb{F}_p^2$ in place of B is $> c(k)p^{3/2}$.

Proof of Lemma 6.2

It follows easily from Lemma 6.1 that $h(x_1, x_2)$ cannot be of form both $g_1(x_1)p_1(x_1, x_2)^d$ and $g_2(x_2)p_2(x_1, x_2)^d$ simultaneously with $g_1(x), g_2(x) \in \mathbb{F}_p[x]$ and $p_1(x_1, x_2)$, $p_2(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$. Thus by symmetry reasons we may suppose that $h(x_1, x_2)$ is not of the form $g_2(x_2)p_2(x_1, x_2)^d$.

Since B is a box p-lattice, write it in the form

$$B = \{\mathbf{x} = (v_1 b_1, v_2 b_2) : \ v_1, v_2 \in \mathbb{N}, \ 0 \le v_1 b_1 \le t_1, \ 0 \le v_2 b_2 \le t_2\} \qquad (6.6)$$

with $b_1, b_2 \in \mathbb{N}$ and $0 \le t_1, t_2 < p$. Then by the triangle inequality

$$\left| \sum_{\mathbf{x} \in B} \chi\left(h(\mathbf{x})\right) \right| \le \sum_{0 \le v_2 \le [t_2/b_2]} \left| \sum_{0 \le v_1 \le [t_1/b_1]} \chi\left(h(v_1 b_1, v_2 b_2)\right) \right|.$$

For fixed v_2, b_1 and b_2, the polynomial $h(v_1 b_1, v_2 b_2)$ is a polynomial of one variable in v_2. We will use the following consequence of Weil's theorem [106]:

Lemma 6.3 (Weil) *Suppose that p is a prime, χ is a non-principal character modulo p of order d, $f(x) \in \mathbb{F}_p[x]$ has s distinct roots in $\overline{\mathbb{F}}_p$, and it is not the constant multiple of the d-th power of a polynomial over $\mathbb{F}_p$. Let y be a real number with $0 < y \le p$. Then for any $x \in \mathbb{F}_p$:*

$$\left| \sum_{x < n \le x+y} \chi(f(n)) \right| < 9sp^{1/2} \log p.$$

Proof of Lemma 6.3

This is an immediate consequence of Lemma 1 in [2].

If, for fixed v_2, b_1, b_2, the polynomial $h(xb_1, v_2 b_2) \in \mathbb{F}_p[x]$ of one variable is not of the form $cg(x)^d$ with $c \in \mathbb{F}_p$, $g(x) \in \mathbb{F}_p[x]$, then by Lemma 6.3

$$\left| \sum_{0 \le v_1 \le [t_1/b_1]} \chi\left(h(v_1 b_1, v_2 b_2)\right) \right| \le 9kp^{1/2} \log p.$$

We will show that for fixed b_1 and b_2 there are only few values of v_2 for which the polynomial $h(xb_1, v_2 b_2) \in \mathbb{F}_p[x]$ is of the form $cg(x)^d$. For this we need

Lemma 6.4 (Gyarmati, Sárközy, Stewart) *Let $h(x, y) \in \mathbb{F}_p[x, y]$ be a polynomial of two variables, which is not of the form $q(y)p(x, y)^d$ with $q(y) \in \mathbb{F}_p[y]$, $p(x, y) \in \mathbb{F}_p[x, y]$. Denote by n and m the degree of the polynomial $h(x, y)$ in x and y, respectively. Then there are at most $nm + m$ values $y_0 \in \mathbb{F}_p$ such that*

$$h(x, y_0) \in \mathbb{F}_p[x]$$

is of the form $cg(x)^d$ with $c \in \mathbb{F}_p$, $g(x) \in \mathbb{F}_p[x]$.

Proof of Lemma 6.4 This is Lemma 4 in [60].

Let n and m be the degree of $h(x_1, x_2)$ in x_1 and x_2 respectively. We have assumed that $h(x_1, x_2)$ is not of the form $g_2(x_2)p_2(x_1, x_2)^d$, thus by Lemma 6.4, there are at most $nm + m$ values of v_2 such that $h(xb_1, v_2b_2)$ is of the form $cg(x)^d$ for some $c \in \mathbb{F}_p$, $g(x) \in \mathbb{F}_p[x]$. Let $\mathcal{V}$ denote the set of these v_2's. Then

$$|\mathcal{V}| \leq mn + m \leq k^2 + k. \tag{6.7}$$

By (6.6)

$$\left|\sum_{\mathbf{x}\in B} \chi\left(h(\mathbf{x})\right)\right| \leq \sum_{v_2\in\mathcal{V}} \left|\sum_{0\leq v_1\leq [t_1/b_1]} \chi\left(h(v_1b_1, v_2b_2)\right)\right| + \sum_{v_2\in\mathbb{F}_p\setminus\mathcal{V}} \left|\sum_{0\leq v_1\leq [t_1/b_1]} \chi\left(h(v_1b_1, v_2b_2)\right)\right|.$$

For $v_2 \in \mathcal{V}$ we use the trivial estimate p for the inner sum. By Lemma 6.4 and (6.7)

$$\sum_{v_2\in\mathcal{V}} \left|\sum_{0\leq v_1\leq [t_1/b_1]} \chi\left(h(v_1b_1, v_2b_2)\right)\right| \leq (k^2 + k)p.$$

For $v_2 \in \mathbb{F}_p \setminus \mathcal{V}$ we use Lemma 6.3 to deduce that

$$\sum_{v_2\in\mathbb{F}_p\setminus\mathcal{V}} \left|\sum_{0\leq v_1\leq [t_1/b_1]} \chi\left(h(v_1b_1, v_2b_2)\right)\right| < 9kp^{3/2}\log p.$$

Thus by (6.4)

$$\left|\sum_{\mathbf{x}\in B} \chi\left(h(\mathbf{x})\right)\right| < (k^2 + k)p + 9kp^{3/2}\log p < 10kp^{3/2}\log p$$

which completes the proof of Lemma 6.2.

Lemma 6.5 (Gyarmati, Sárközy, Stewart) *Suppose that $f \in \mathbb{F}_p[x_1, x_2]$ is a polynomial such that there are no distinct $\mathbf{d_1}, \ldots, \mathbf{d}_\ell \in \mathbb{F}_p^2$ with the property that $f(\mathbf{x} + \mathbf{d_1}) \ldots f(\mathbf{x} + \mathbf{d}_\ell)$ is of the form $cg(\mathbf{x})^2$ with $c \in \mathbb{F}_p$, $g \in \mathbb{F}_p[x_1, x_2]$. Let k be the degree of the polynomial $f(x_1, x_2)$. Then for the binary p-lattice η defined in* (6.3) *we have*

$$|Q_\ell(\eta)| < 11k\ell p^{3/2}\log p.$$

Proof of Lemma 6.5 We have

$$Q_\ell(\eta) = \max_{B,\mathbf{d_1},\ldots,\mathbf{d_k}} \left| \sum_{\mathbf{x}\in B} \eta(\mathbf{x}+\mathbf{d_1}) \cdots \eta(\mathbf{x}+\mathbf{d}_\ell) \right|,$$

where the maximum is taken over all distinct $\mathbf{d_1}, \ldots, \mathbf{d}_\ell \in I_p^2$ and box p-lattices B such that $B+\mathbf{d_1}, \ldots, B+\mathbf{d}_\ell \subseteq I_p^2$. Let B be the box p-lattice, $\mathbf{d_1}, \ldots, \mathbf{d}_\ell \in I_p^2$ be the vectors for which this maximum is attained so that

$$Q_\ell(\eta) = \left| \sum_{\mathbf{x}\in B} \eta(\mathbf{x}+\mathbf{d_1}) \cdots \eta(\mathbf{x}+\mathbf{d}_\ell) \right|.$$

Write $h(\mathbf{x}) = f(\mathbf{x}+\mathbf{d_1}) \cdots f(\mathbf{x}+\mathbf{d}_\ell)$, then

$$Q_\ell(\eta) \le \left| \sum_{\mathbf{x}\in B} \left(\frac{h(\mathbf{x})}{p} \right) \right| + \sum_{\substack{\mathbf{x}\in B \\ h(\mathbf{x})=0}} 1.$$

$h(\mathbf{x})$ is a polynomial of degree $k\ell$. Estimating the number of zeros of $h(\mathbf{x})$ we find that

$$\sum_{\substack{\mathbf{x}\in B \\ h(\mathbf{x})=0}} 1 \le k\ell p. \tag{6.8}$$

By assumption $h(\mathbf{x})$ is not of the form $cg(\mathbf{x})^2$ and its degree is ℓk. Thus by Lemma 6.2 and (6.8) we have

$$Q_\ell(\eta) \le 10\ell k p^{3/2} \log p + \ell k p,$$

which was to be proved.

Suppose that one of the 5 conditions in Theorem 6.1 holds. We will prove that the product

$$h(\mathbf{x}) = f(\mathbf{x}+\mathbf{d_1}) \ldots f(\mathbf{x}+\mathbf{d}_\ell)$$

cannot be the constant multiple of a perfect square. Then by Lemma 6.5 we get Theorem 6.1.

Next we will introduce three definitions (they are very similar to the ones introduced by Goubin, Mauduit and Sárközy in [31]).

Definition 6.2 (Gyarmati, Sárközy, Stewart) *Let G be a group with respect to addition. Let A and B be subsets of G and suppose that for all c in*

G the number of solutions of

$$a + b = c,$$

with a in A and b in B is even. Then (A, B) is said to have property P.

Definition 6.3 (Gyarmati, Sárközy, Stewart) *Let r, ℓ, and m be positive integers with $r, \ell \leq m$. The triple (r, ℓ, m) is said to be admissible if there are no $\mathcal{A}, \mathcal{B} \subseteq \mathbb{Z}_m$ such that $|\mathcal{A}| = r$, $|\mathcal{B}| = \ell$, and $(\mathcal{A}, \mathcal{B})$ possesses property P.*

We shall also introduce an equivalence relation on $\mathbb{F}_p[x_1, x_2]$ as in the proof of Theorem 6.A in [31].

Definition 6.4 (Gyarmati, Sárközy, Stewart) *Two polynomials $\varphi(x_1, x_2), \psi(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ are equivalent if there are $a_1, a_2 \in \mathbb{F}_p$ such that*

$$\psi(x_1, x_2) = \varphi(x_1 + a_1, x_2 + a_2).$$

Write the polynomial $f(x_1, x_2)$ in the theorem as a product of irreducible polynomials in $\mathbb{F}_p[x_1, x_2]$. (Recall that the lattice η is determined by this polynomial $f(x_1, x_2)$, the definition of η is presented in (6.2).) Let us group these factors so that in each group the equivalent irreducible factors are collected. Consider a typical group $\varphi(x_1 + a_{1,1}, x_2 + a_{2,1}), \varphi(x_1 + a_{1,2}, x_2 + a_{2,2}), \ldots, \varphi(x_1 + a_{1,s}, x_2 + a_{2,s})$. Then $f(x_1, x_2)$ is of the form

$$f(x_1, x_2) = \varphi(x_1 + a_{1,1}, x_2 + a_{2,1}) \cdots \varphi(x_1 + a_{1,s}, x_2 + a_{2,s}) g(x_1, x_2),$$

where $g(x_1, x_2)$ has no irreducible factor equivalent with any $\varphi(x_1 + a_{1,i}, x_2 + a_{2,i})$ $(1 \leq i \leq s)$.

We will use the following lemma:

Lemma 6.6 (Gyarmati, Sárközy, Stewart) *Let $\varphi(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ be nonzero and let $c, a_1, a_2 \in \mathbb{F}_p$ with $(a_1, a_2) \neq (0, 0)$ be such that*

$$\varphi(x_1, x_2) = c\varphi(x_1 + a_1, x_2 + a_2), \tag{6.9}$$

for all (x_1, x_2) in $\mathbb{F}_p^2$. Suppose that the degree of $\varphi(x_1, x_2)$ is less than p. Then there is a polynomial $g \in F_p[x]$ such that

$$\varphi(x_1, x_2) = g(a_2 x_1 - a_1 x_2). \tag{6.10}$$

Proof of Lemma 6.6 We will use repeatedly the fact that if two polynomials of degree less than p in each variable define the same polynomial function, then they must also be identical polynomials.

By considering the highest degree terms in (6.9), we get $c = 1$ so that

$$\varphi(x_1, x_2) = \varphi(x_1 + a_1, x_2 + a_2).$$

It follows from this that for every $t \in \mathbb{F}_p$

$$\varphi(x_1, x_2) = \varphi(x_1 + ta_1, x_2 + ta_2). \tag{6.11}$$

One of a_1 and a_2 is nonzero and, without loss of generality, we may suppose that $a_2 \neq 0$. Then write $\varphi(x_1, x_2)$ in the form

$$\begin{aligned}\varphi(x_1, x_2) &= \varphi(a_2^{-1}((a_2x_1 - a_1x_2) + a_1x_2), x_2) \\ &= q_n(a_2x_1 - a_1x_2)x_2^n + q_{n-1}(a_2x_1 - a_1x_2)x_2^{n-1} + \dots \\ &+ q_0(a_2x_1 - a_1x_2),\end{aligned} \tag{6.12}$$

where $q_i(x) \in \mathbb{F}_p[x]$ are polynomials of one variable. For fixed x_1, x_2 write $A = \varphi(x_1, x_2)$ and $Q_i = q_i(a_2x_1 - a_1x_2) = q_i(a_2(x_1 + ta_1) - a_1(x_2 + ta_2))$. Then by (6.11) and (6.12) for every $t \in \mathbb{F}_p$:

$$A = \varphi(x_1, x_2) = \varphi(x_1 + ta_1, x_2 + ta_2) = Q_n(x_2 + ta_2)^n + \dots + Q_0.$$

Both A and the expression on the right above are polynomials in t of degree at most p. These polynomials define the same function and so they are the same polynomials, which is possible only if $n = 0$. It follows that

$$q_0(a_2x_1 - a_1x_2) - \varphi(x_1, x_2) = Q_0 - A = 0,$$

for every $x_1, x_2 \in \mathbb{F}_p$. Since both q_0 and φ have degree less than p in x_1 and x_2, thus

$$q_0(a_2x_1 - a_1x_2) = \varphi(x_1, x_2)$$

as formal polynomials, which proves (6.10).

First we study the case when condition a) holds in Theorem 6.1, so when $f(x_1, x_2)$ is irreducible in $\mathbb{F}_p[x_1, x_2]$. As before let $\mathbf{d_1}, \dots, \mathbf{d}_\ell$ be distinct elements of I_p^2 and put $h(\mathbf{x}) = f(\mathbf{x} + \mathbf{d_1}) \cdots f(\mathbf{x} + \mathbf{d}_\ell)$. Then by Lemma 6.6 the irreducible polynomials $f(\mathbf{x} + \mathbf{d_j})$ are different since $f(x_1, x_2)$ is not of the form (6.3). By Lemma 6.1, there is unique factorization in $\mathbb{F}_p[x_1, x_2]$, thus $h(\mathbf{x})$ cannot be the constant multiple of a perfect square. By using Lemma 6.5 we get the statement.

Next we prove parts b), c) and d) in Theorem 6.1. Write $f(x_1, x_2)$ in the form $u(x_1, x_2)(v(x_1, x_2))^2$ where $u(x_1, x_2)$ is squarefree, so, in other words, there is no non-constant irreducible polynomial $h(x_1, x_2)$ with $(h(x_1, x_2))^2$ a divisor of $u(x_1, x_2)$. Since $f(x_1, x_2)$ is not of the form (6.3), in the factorization of $u(x_1, x_2)$ there is an irreducible factor $\overline{u}(x_1, x_2)$ which cannot be written in the form

$$\overline{u}(x_1, x_2) = u(\alpha x_1 + \beta x_2). \tag{6.13}$$

Consider the polynomials $\overline{u}(\mathbf{x} + \mathbf{a_i})$ for $i = 1, 2, \ldots, r$ which are equivalent with $\overline{u}(\mathbf{x})$ and appear in the factorization of $u(\mathbf{x})$.

We shall prove that $h(\mathbf{x}) = f(\mathbf{x} + \mathbf{d_1}) \cdots f(\mathbf{x} + \mathbf{d}_\ell)$ is not a constant multiple of a perfect square. We shall suppose that $h(\mathbf{x})$ is the constant multiple of a perfect square. Then $h_1(\mathbf{x}) = u(\mathbf{x} + \mathbf{d_1}) \cdots u(\mathbf{x} + \mathbf{d}_\ell)$ is also a constant multiple of a perfect square.

Write $h_1(\mathbf{x})$ as a product of irreducible polynomials in $\mathbb{F}_p[x_1, x_2]$. Then all polynomials $\overline{u}(\mathbf{x} + \mathbf{a_i} + \mathbf{d_j})$ $(1 \le i \le s,\ 1 \le j \le \ell)$ occur amongst the factors. These polynomials $\overline{u}(\mathbf{x}+\mathbf{a_i}+\mathbf{d_j})$ are equivalent, and no other factors belonging to this equivalence class will occur amongst the irreducible factors of $h_1(\mathbf{x})$. By Lemma 6.6 all polynomials $\overline{u}(\mathbf{x}+\mathbf{c})$ for $\mathbf{c} \in \mathbb{F}_p^2$ are distinct since $\overline{u}$ is not of the form (6.13). Thus in the collection, formed by the equivalent factors $\overline{u}(\mathbf{x}+\mathbf{a_i}+\mathbf{d_j})$, every polynomial must occur an even number of times. As a consequence every $c \in \mathbb{F}_p^2$ occurs an even number of times in the form $a_i + d_j$ with $1 \le i \le r$ and $1 \le j \le \ell$.

Lemma 6.7 (Gyarmati, Sárközy, Stewart) *Let $s(s-1)/2 < p$ and*

$$\mathbf{d_i} = (d_i', d_i'') \in \mathbb{F}_p^2 \quad (1 \le i \le s)$$

be different vectors. Then there exists a $\lambda \in \mathbb{F}_p^*$ such that

$$d_i' + \lambda d_i'' \in \mathbb{F}_p \quad (1 \le i \le s)$$

are different.

Proof of Lemma 6.7 Suppose that for some pair (i, j) with $1 \le i < j \le \ell$ we have

$$d_i' + \lambda d_i'' = d_j' + \lambda d_j''.$$

Then $d_i'' \ne d_j''$, otherwise we obtain $(d_i', d_i'') = (d_j', d_j'')$. Thus for every $i \ne j$ at most one λ exists such that

$$d_i' + \lambda d_i'' = d_j' + \lambda d_j''.$$

The number of pairs (i,j) with $1 \le i < j \le \ell$ is $\ell(\ell-1)/2$. Thus at most $\ell(\ell-1)/2$ values of λ exist such that

$$d_i' + \lambda d_i'' = d_j' + \lambda d_j''$$

for some $i \neq j$. Since $\ell(\ell-1)/2 < p$ the lemma follows.

We have $\mathcal{A} = \{\mathbf{a_1}, \dots, \mathbf{a_r}\}$ and $\mathcal{D} = \{\mathbf{d_1}, \dots, \mathbf{d}_\ell\} \subseteq \mathbb{F}_p^2$, where $r \le k$. By Lemma 6.7 we may choose $\lambda \in \mathbb{F}_p$ such that both sets

$$\mathcal{A}' = \{a' + \lambda a'' : \ (a', a'') \in \mathcal{A}\}$$

and

$$\mathcal{D}' = \{d' + \lambda d'' : \ (d', d'') \in \mathcal{D}\}$$

contain different elements.

Lemma 6.8 (Gyarmati, Sárközy, Stewart) $(\mathcal{A}', \mathcal{D}')$ *possesses property P.*

Proof of Lemma 6.8 In order to verify the lemma we need to prove that for any $c \in \mathbb{F}_p$ the number of solutions

$$a + d = c, \quad a \in \mathcal{A}', \ d \in \mathcal{D}' \tag{6.14}$$

is even. Indeed, it is clear that the number of solutions of (6.14) is the same as the number of solutions of

$$\begin{aligned} (a', a'') + (d', d'') = (c', c''), \quad (a', a'') \in \mathcal{A}, \ (d', d'') \in \mathcal{D} \\ c' + \lambda c'' = c. \end{aligned} \tag{6.15}$$

Since $(\mathcal{A}, \mathcal{D})$ possesses property P, for each $(c', c'') \in \mathbb{F}_p^2$ the number of solutions of the equation

$$(a', a'') + (d', d'') = (c', c''), \quad (a', a'') \in \mathcal{A}, \ (d', d'') \in \mathcal{D}$$

is even. Thus the number of solutions of the system (6.15) is also even, and equivalently, the number of solutions of (6.14) is also even. This proves Lemma 6.8.

By Lemma 6.8 $(\mathcal{A}', \mathcal{D}')$ possesses property P. Thus (r, ℓ, p) is not an admissible triple. By contrast we have the following lemma.

Lemma 6.9 (Goubin, Mauduit, Sárközy) *(i) For every prime p and $r \in \mathbb{N}$ the triple $(r, 2, p)$ is admissible.*

(ii) If p is prime, $r, \ell \in \mathbb{N}$ and

$$4^{\ell+r} < p,$$

then (r, ℓ, p) is admissible.

(iii) If p is a prime such that 2 is a primitive root modulo p, then for every pair $(r, \ell) \in \mathbb{N}$ with $r < p$, $\ell < p$ the triple (r, ℓ, p) is admissible.

Proof of Lemma 6.9 Parts (i) and (iii) are Theorem 2 in [31] while part (ii) is Theorem 2 in [79].

Since (r, ℓ, p) is not admissible parts b), c) and d) of Theorem 6.1 follow from Lemma 6.9. In the proofs of b) and d) we could have replaced Lemma 6.8 by Lemma 4 in [79], however the lemma there does not suffice to prove part c) in Theorem 6.1, thus we have preferred to prove Lemma 6.8 here.

In order to prove part e) in Theorem 6.1 we note that the degree of the polynomial $h(x_1, x_2)$ in x_1 is odd, thus it cannot be the constant multiple of a perfect square. Using Lemma 6.5 again part e) follows.

7 On Legendre symbol lattices (the degenerate case and a related construction)

In this section our goal is to continue the study of Construction 6.1. First we will analyze the degenerate case. In Section 7.1 we will analyze the structure of the degenerate polynomials $f(x_1, x_2)$, and we will introduce the notion of the normal form and rank $r = r(f)$ of such a polynomial. In Section 7.2 we will prove that if f is degenerate, $\ell \leq r = r(f)$, η is defined by (6.2) and one of four specified conditions holds, then $Q_\ell(\eta)$ is small. We will also present an algorithm for deciding whether a given polynomial $f(x_1, x_2)$ is degenerate and, if it is, for determining its normal form. In Section 7.3 we will show that here the upper bound r cannot be replaced by 2^r. In Section 7.4 we will study the implementation of Construction 6.1 and, in particular, we will construct a large family of polynomials $f(x_1, x_2)$ which are non-degenerate and satisfy the first sufficient condition in Theorem 6.1 so that the binary lattice η in (6.2) possesses strong pseudorandom properties. In particular its pseudorandom measures $Q_\ell(\eta)$ are small for ℓ not very large. Finally, in Section 7.5, we construct families of polynomials for which the bounds for the pseudorandom measures are essentially optimal.

7.1 Structure of degenerate polynomials

In this section our goal is to transform the representation (6.3) of a degenerate polynomial into another more useful one. We will need several lemmas.

Lemma 7.1 *If $\mathbb{F}$ is a field, then in $\mathbb{F}[x_1, x_2, \ldots, x_n]$ every polynomial has a factorization into irreducible polynomials which is unique apart from constant factors and reordering.*

Proof of Lemma 7.1 See, for example [93, Theorem 207]. □

Lemma 7.2 (Gyarmati, Sárközy, Stewart) *Let $g_1, g_2 \in \mathbb{F}_p[x, y]$ and $f \in \mathbb{F}_p[x]$ be non-zero polynomials. Suppose that for some $(\alpha, \beta) \in \mathbb{F}_p \times \mathbb{F}_p$*

$$g_1(x, y)g_2(x, y) = f(\alpha x + \beta y). \tag{7.1}$$

Then there exist $f_1, f_2 \in \mathbb{F}_p[x]$ such that

$$g_i(x, y) = f_i(\alpha x + \beta y)$$

for $i = 1, 2$.

Proof of Lemma 7.2 If $(\alpha, \beta) = (0, 0)$ the result is immediate. Thus we may suppose that $(\alpha, \beta) \neq (0, 0)$ and, without loss of generality, we may assume that $\alpha \neq 0$. Put

$$z = \alpha x + \beta y$$

so that $x = \alpha^{-1}z - \alpha^{-1}\beta y$. We may now define h_1, h_2 in $\mathbb{F}_p[y, z]$ by putting

$$h_i(y, z) = g_i(\alpha^{-1}z - \alpha^{-1}\beta y, y) \quad \text{for } i = 1, 2.$$

From (7.1) we find that

$$h_1(y, z)h_2(y, z) = f(z). \tag{7.2}$$

Write

$$\begin{aligned} h_1(y, z) &= u_a(z)y^a + u_{a-1}(z)y^{a-1} + \cdots + u_0(z), \\ h_2(y, z) &= v_b(z)y^b + v_{b-1}(z)y^{b-1} + \cdots + v_0(z) \end{aligned}$$

and

$$h_1(y, z)h_2(y, z) = w_{a+b}(z)y^{a+b} + w_{a+b-1}(z)y^{a+b-1} + \cdots + w_0(z)$$

where $u_a(z), v_b(z)$ are not the zero polynomial. Clearly we have

$$w_{a+b}(z) = u_a(z)v_b(z). \tag{7.3}$$

But by (7.2), $h_1(y, z)h_2(y, z)$ is a one variable polynomial in z, thus we have

$$w_{a+b}(z) = w_{a+b-1}(z) = \cdots = w_1(z) = 0 \text{ if } a + b > 0. \tag{7.4}$$

It follows from (7.3) and $u_a(z) \neq 0$, $v_b(z) \neq 0$ that $w_{a+b}(z) \neq 0$. Thus by (7.4) we have $a + b = 0$ whence $a = b = 0$. Then $h_1(y, z) = u_0(z)$, $h_2(y, z) = v_0(z)$ which completes the proof of the lemma. □

We shall identify the elements of $\mathbb{F}_p$ with the p congruence classes modulo p and shall denote the elements of $\mathbb{F}_p \times \mathbb{F}_p$ by (a, b) with a and b integers representing the congruence class of a and of b modulo p. Define the subset T of $\mathbb{F}_p \times \mathbb{F}_p$ by

$$T = \{(0, 1), (1, 0), (1, 1), (2, 1), \ldots, (p - 1, 1)\}.$$

Lemma 7.3 (Gyarmati, Sárközy, Stewart) *Let f be a non-constant degenerate polynomial in $\mathbb{F}_p[x_1, x_2]$ of degree less than p in x_1 and in x_2. Then there exist a non-zero λ in $\mathbb{F}_p$, a non-negative integer r, distinct elements*

$(\gamma_1, \delta_1), \ldots, (\gamma_r, \delta_r)$ from T, ψ in $\mathbb{F}_p[x_1, x_2]$ and squarefree non-constant polynomials $\varphi_1, \ldots, \varphi_r$ in $\mathbb{F}_p[x]$ for which

$$f(x_1, x_2) = \lambda \left(\prod_{j=1}^{r} \varphi_j(\gamma_j x_1 + \delta_j x_2) \right) \psi^2(x_1, x_2). \tag{7.5}$$

Further r is uniquely determined and the polynomials $\varphi_j(\gamma_j x_1 + \delta_j x_2)$ and $\psi(x_1, x_2)$ are unique up to constant factors and reordering of $\varphi_1(\gamma_1 x_1 + \delta_1 x_2), \ldots, \varphi_r(\gamma_r x_1 + \delta_r x_2)$.

We shall refer to a decomposition of f as in (7.5) as a normal form of f and to r as the rank of f. Notice that since $(\gamma_1, \delta_1), \ldots, (\gamma_r, \delta_r)$ are distinct elements of T we have

$$\gamma_j \delta_i - \delta_j \gamma_i \neq 0 \quad \text{for } i \neq j. \tag{7.6}$$

Proof of Lemma 7.3 Let ψ be a polynomial of largest degree for which ψ^2 divides f in $\mathbb{F}_p[x_1, x_2]$. Then since f is degenerate we may write f in the form (6.3) with ψ as above and with $(\gamma_i, \delta_i) \neq (0, 0)$ for $i = 1, \ldots, s$. Further we may suppose that $\varphi_1, \ldots, \varphi_s$ are squarefree polynomials in $\mathbb{F}_p[x]$ and that $\varphi_1 \cdots \varphi_s$ is also squarefree.

Suppose that φ is in $\mathbb{F}_p[x]$ and (γ, δ) are in $\mathbb{F}_p \times \mathbb{F}_p \backslash \{(0, 0)\}$ and define φ^* in $\mathbb{F}_p[x]$ by

$$\varphi^*(x) = \begin{cases} \varphi(\gamma x) & \text{when } \gamma \neq 0, \\ \varphi(\delta x) & \text{when } \gamma = 0. \end{cases}$$

Then

$$\varphi(\gamma x_1 + \delta x_2) = \begin{cases} \varphi^*(x_1 + \delta \gamma^{-1} x_2) & \text{if } \gamma \neq 0, \\ \varphi^*(x_2) & \text{if } \gamma = 0. \end{cases}$$

Therefore we may write

$$\varphi_1(\gamma_1 x_1 + \delta_1 x_2) \cdots \varphi_s(\gamma_s x_1 + \delta_s x_2)$$

as

$$\varphi_1^*(\gamma_1 x_1 + \delta_1 x_2) \cdots \varphi_s^*(\gamma_s x_1 + \delta_s x_2)$$

where now (γ_i, δ_i) is in T for $i = 1, \ldots, s$. We now collect and multiply together the polynomials φ_i^* for which (γ_i, δ_i) are the same to get a representation for f of the form (7.5).

Suppose that, in addition to (7.5),

$$f(x_1, x_2) = \lambda_1 \left(\prod_{j=1}^{s} \rho_j(\theta_j x_1 + \beta_j x_2) \right) \psi_1^2(x_1, x_2)$$

with $(\theta_1, \beta_1), \ldots, (\theta_s, \beta_s)$ distinct elements of T, λ_1 a non-zero element of $\mathbb{F}_p$, ψ_1 in $\mathbb{F}_p[x_1, x_2]$ and squarefree non-constant polynomials $\rho_1, \ldots, \rho_s$ in $\mathbb{F}_p[x]$. By Lemma 7.1 $\psi(x)$ is a constant times $\psi_1(x)$ since $\psi^2(x)$ and $\psi_1^2(x)$ correspond to the greatest square factor of f in $\mathbb{F}_p[x_1, x_2]$. Next note that for each j from 1 to s we may decompose $\rho_j(\theta_j x_1 + \beta_j x_2)$ into irreducibles and by Lemma 7.2

$$\rho_j(\theta_j x_1 + \beta_j x_2) = \rho_{j,1}(\theta_j x_1 + \beta_j x_2) \cdots \rho_{j,t}(\theta_j x_1 + \beta_j x_2)$$

where $\rho_{j,1}, \ldots, \rho_{j,t}$ are irreducible polynomials in $\mathbb{F}_p[x]$. Thus each irreducible $\rho_{j,k}(\theta_j x_1 + \beta_j x_2)$ occurs in the essentially unique decomposition of $\varphi_m(\gamma_m x_1 + \delta_m x_2)$ into irreducibles for some m. Notice that if a polynomial $g(x, y) = f_1(\gamma_1 x + \beta_1 y) = f_2(\gamma_2 x + \beta_2 y)$ with $f_1, f_2 \in \mathbb{F}[x]$ and $\gamma_1\beta_2 - \gamma_2\beta_1 \neq 0$ then $g(x, y)$ is a constant. (Indeed, fix $a, b, c, d \in \mathbb{F}_p$ and we will prove that $g(a, b) = g(c, d)$. Since $\gamma_1\beta_2 - \gamma_2\beta_1 \neq 0$ the system of linear equations

$$\gamma_1 x + \beta_1 y = \gamma_1 a + \beta_1 b$$
$$\gamma_2 x + \beta_2 y = \gamma_2 c + \beta_2 d$$

has a unique solution in $x, y \in \mathbb{F}_p$. Then

$$\begin{aligned} g(a, b) &= f_1(\gamma_1 a + \beta_1 b) = f_1(\gamma_1 x + \beta_1 y) = g(x, y) = f_2(\gamma_2 x + \beta_2 y) \\ &= f_2(\gamma_2 c + \beta_2 d) = g(c, d).) \end{aligned}$$

Thus, by (7.6), $(\theta_j, \beta_j) = (\gamma_m, \delta_m)$. Repeating this argument with all the irreducible factors of ρ_j and all the irreducible factors of $\varphi_m(\gamma_m x_1 + \delta_m x_z)$ we find that $\varphi_m(\gamma_m x_1 + \delta_m x_2)/\rho_j(\theta_j x_1 + \beta_j x_2)$ is a constant. From this it readily follows that $r = s$ and the result follows. □

We remark that we may determine if a polynomial f is degenerate by first replacing it with a polynomial f^* of degree at most $p - 1$ in each variable by using the fact that $x^p = x$ for all x in $\mathbb{F}_p$. We then factor f^* and write f^* as a product of irreducibles multiplied by its largest square divisor. Each irreducible must be tested to see if it is of the form $g(\gamma x + \beta y)$ with $g \in \mathbb{F}_p[x]$ and $(\gamma, \beta) \in T$. Given (γ, β) in T if suffices to check that the irreducible is constant on the lines in $\mathbb{F}_p \times \mathbb{F}_p$ given by $\gamma x + \beta y = c$ for c in $\mathbb{F}_p$ and this is a finite process. Furthermore T is a finite set. Either there is an irreducible

not of the form $g(\gamma x + \beta y)$ for any $g \in \mathbb{F}[x]$ and (γ, β) in T in which case f^* is non-degenerate or f^* is degenerate and we may produce the normal form as in the proof of Lemma 7.3.

7.2 The pseudorandom measures of small order in the degenerate case.

We will show that if $f(x_1, x_2)$ is a degenerate polynomial and the order ℓ of the pseudorandom measure Q_ℓ is not greater than the rank of f then, for the binary lattice η defined in (6.2), $Q_\ell(\eta)$ is small. In fact our estimates are the same as in the non-degenerate case studied in Theorem 6.1.

Theorem 7.1 (Gyarmati, Sárközy, Stewart) *Let $f(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ be a non-constant degenerate polynomial of reduced normal form (7.5) with degree k. Suppose that ℓ, the order of the pseudorandom measure is not greater than the rank r of $f(x_1, x_2)$, and one of the following 5 conditions holds:*
a) $f(x_1, x_2)$ is irreducible in $\mathbb{F}_p[x_1, x_2]$,
b) $\ell = 2$,
c) 2 is a primitive root modulo p,
d) $(4k)^\ell < p$ or $(4\ell)^k < p$,
e) ℓ and the degree of the polynomial $f(x_1, x_2)$ in x_1 (or in x_2) are odd.
Then for the binary lattice η defined in (6.2) we have

$$Q_\ell(\eta) < 11k\ell p^{3/2} \log p.$$

Proof of Theorem 7.1 The proof will be based on the following result.

Lemma 7.4 (Gyarmati, Sárközy, Stewart) *Suppose that $f \in \mathbb{F}_p[x_1, x_2]$ is a polynomial such that there are no distinct $\mathbf{d_1}, \dots, \mathbf{d}_\ell \in \mathbb{F}_p^2$ with the property that $f(\mathbf{x} + \mathbf{d_1}) \dots f(\mathbf{x} + \mathbf{d}_\ell)$ is of the form $cq(\mathbf{x})^2$ with $c \in \mathbb{F}_p$, $q \in \mathbb{F}_p[x_1, x_2]$. Let k be the degree of the polynomial $f(x_1, x_2)$. Then for the binary p-lattice η defined in (6.2) we have*

$$|Q_\ell(\eta)| < 11k\ell p^{3/2} \log p.$$

Proof of Lemma 7.4 This is Lemma 6.5 in Section 6.

In order to ensure the applicability of this lemma, we have to show that it follows from one of the 5 assumptions in Theorem 7.1 that there are not distinct $\mathbf{d_1}, \dots, \mathbf{d}_\ell \in \mathbb{F}_p^2$ such that the polynomial

$$h(\mathbf{x}) = f(\mathbf{x} + \mathbf{d_1}) \dots f(\mathbf{x} + \mathbf{d}_\ell)$$

is of the form $cq(\mathbf{x})^2$ with $c \in \mathbb{F}_p$, $q \in \mathbb{F}_p[x_1, x_2]$. Indeed, if this is proved, then the assumption in Lemma 7.4 holds in each of these 5 cases thus the statement of Theorem 7.1 follows from Lemma 7.4 immediately.

We will prove this by contradiction. Assume that

$$h(\mathbf{x}) = f(\mathbf{x} + \mathbf{d_1}) \cdots f(\mathbf{x} + \mathbf{d}_\ell)$$

is the constant multiple of a perfect square. Then we will prove

$$r + 1 \leq \ell,$$

where r denotes the rank of f, which contradicts our assumption.

Write

$$\mathbf{d}_i = (d_i', d_i'')$$

for $i = 1, \ldots, l$.

Suppose that f has the normal form

$$f(x_1, x_2) = \lambda \prod_{j=1}^{r} f_j(\alpha_j x_1 + \beta_j x_2)\psi^2(x_1, x_2)$$

with $\lambda \in \mathbb{F}_p \backslash \{0\}$, $(\alpha_1, \beta_1), \ldots, (\alpha_r, \beta_r)$ distinct elements of T, $f_1, \ldots, f_r$ squarefree non-constant polynomials in $\mathbb{F}_p[x]$ and $\psi \in \mathbb{F}_p[x_1, x_2]$. Then it follows that

$$\prod_{j=1}^{r} f_j(\alpha_j x_1 + \beta_j x_2 + \alpha_j d_1' + \beta_j d_1'') f_j(\alpha_j x_1 + \beta_j x_2 + \alpha_j d_2' + \beta_j d_2'') \cdots$$
$$f_j(\alpha_j x_1 + \beta_j x_2 + \alpha_j d_\ell' + \beta_j d_\ell''). \tag{7.7}$$

is a non-zero multiple of the square of a polynomial in $\mathbb{F}_p[x_1, x_2]$.

Now we will introduce an equivalence relation which is similar to the one used in the proof of Theorem 1 in [31].

Definition 7.1 *Two polynomials $\varphi(x_1, x_2)$, $\psi(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ are t-equivalent (t for translation) if there are $a_1, a_2 \in \mathbb{F}_p$ such that*

$$\psi(x_1, x_2) = \varphi(x_1 + a_1, x_2 + a_2).$$

Consider any two factors $f_{j_1}(\alpha_{j_1} x_1 + \beta_{j_1} x_2 + \alpha_{j_1} d_{v_1}' + \beta_{j_1} d_{v_1}'') = f_{j_1}^*(\alpha_{j_1} x_1 + \beta_{j_1} x_2)$ and $f_{j_2}(\alpha_{j_2} x_1 + \beta_{j_2} x_2 + \alpha_{j_2} d_{v_2}' + \beta_{j_2} d_{v_2}'') = f_{j_2}^*(\alpha_{j_2} x_1 + \beta_{j_2} x_2)$ with $j_1 \neq j_2$ on the right hand side of (7.7), factor them into irreducible polynomials, and consider an irreducible factor φ_1 of the former polynomial and φ_2 of the latter

polynomial. Then by Lemma 7.2, these irreducible factors are of the form $\varphi_1(\alpha_{j_1}x_1+\beta_{j_1}x_2)$, and $\varphi_2(\alpha_{j_2}x_1+\beta_{j_2}x_2)$. Assume that these two polynomials are t-equivalent, so that there exist $a, b \in \mathbb{F}_p$ such that

$$\begin{aligned}\varphi_1(\alpha_{j_1}x_1+\beta_{j_1}x_2) &= \varphi_2(\alpha_{j_2}(x_1+a)+\beta_{j_2}(x_2+b)) \\ &= \varphi_2((\alpha_{j_2}x_1+\beta_{j_2}x_2)+(\alpha_{j_2}a+\beta_{j_2}b)) = \varphi_3(\alpha_{j_2}x_1+\beta_{j_2}x_2)\end{aligned} \tag{7.8}$$

(where $\varphi_3(z) = \varphi_2(z+(\alpha_{j_2}a+\beta_{j_2}b))$). Both the first and last polynomial in (7.8) are in normal form, and since the normal form is unique, we must have $(\alpha_{j_1}, \beta_{j_1}) = (\alpha_{j_2}, \beta_{j_2})$ whence $j_1 = j_2$.

Thus if two factors $f_{j_1}(\alpha_{j_1}x_1+\beta_{j_1}x_2+\alpha_{j_1}d'_{v_1}+\beta_{j_1}d''_{v_1})$ and $f_{j_2}(\alpha_{j_2}x_1+\beta_{j_2}x_2+\alpha_{j_2}d'_{v_2}+\beta_{j_2}d''_{v_2})$ on the right hand side of (7.7) have t-equivalent irreducible factors then $j_1 = j_2$. But then the expression (7.7) is of the form $cq(x_1,x_2)^2$ if and only if

$$f_j(\alpha_j x_1+\beta_j x_2+\alpha_j d'_1+\beta_j d''_1)\cdots f_j(\alpha_j x_1+\beta_j x_2+\alpha_j d'_\ell+\beta_j d''_\ell)$$

is the constant multiple of a square for every $1 \le j \le r$. Writing $z = \alpha_j x_1+\beta_j x_2$ and $d^*_j(i) = \alpha_j d'_i+\beta_j d''_i \in \mathbb{F}_p^*$ we obtain for $1 \le j \le r$:

$$f_j(z+d^*_j(1))f_j(z+d^*_j(2))\cdots f_j(z+d^*_j(\ell))$$

is of the form $cq(z)^2$. Let D_j be the set of terms of the sequence $(d^*_j(1),\ldots,d^*_j(\ell))$ which occur with odd multiplicity. If D_j is not empty, then the one variable polynomial

$$\prod_{d\in D_j} f_j(z+d)$$

is also the constant multiple of a perfect square. By the proof of Lemma 7.2 in [31] this is not possible (note that by Lemma 7.2, in case a) the one-variable polynomial $f(z)$ is also irreducible) since the polynomial $f_j(z)$ is squarefree. It remains to consider the case when D_j is empty for $j = 1,\ldots,r$. Then, for $1 \le j \le r$, in the sequence

$$(\alpha_j d'_1+\beta_j d''_1,\ \alpha_j d'_2+\beta_j d''_2,\ldots,\ \alpha_j d'_\ell+\beta_j d''_\ell)$$

every term occurs with even multiplicity, hence every term occurs with multiplicity at least 2. Then for every j, there is a number $2 \le i(j) \le \ell$ such that

$$\alpha_j d'_1+\beta_j d''_1 = \alpha_j d'_{i(j)}+\beta_j d''_{i(j)}.$$

We will prove that $1, i(1), i(2), \ldots, i(r)$ are different numbers. It is clear that none of $i(1), i(2), \ldots, i(r)$ is equal to 1. It remains to prove that

$$x = i(j_1) = i(j_2) \tag{7.9}$$

is not possible. Suppose that (7.9) holds. Then

$$\begin{aligned} \alpha_{j_1} d_1' + \beta_{j_1} d_1'' &= \alpha_{j_1} d_x' + \beta_{j_1} d_x'', \\ \alpha_{j_2} d_1' + \beta_{j_2} d_1'' &= \alpha_{j_2} d_x' + \beta_{j_2} d_x''. \end{aligned}$$

Thus

$$\begin{aligned} \alpha_{j_1}(d_1' - d_x') - \beta_{j_1}(d_1'' - d_x'') &= 0, \\ \alpha_{j_2}(d_1' - d_x') - \beta_{j_2}(d_1'' - d_x'') &= 0. \end{aligned} \tag{7.10}$$

Since $(d_1', d_1'') \neq (d_x', d_x'')$ from (7.10) we obtain

$$\alpha_{j_1}\beta_{j_2} - \alpha_{j_2}\beta_{j_1} = 0,$$

from which $j_1 = j_2$ follows. Thus $1 < i(1), i(2), \ldots, i(r) \leq \ell$ and $i(1), i(2), \ldots, i(r)$ are different numbers, so that

$$r + 1 \leq \ell$$

which contradicts the conditions of Theorem 7.1 and this completes the proof of the theorem. □

7.3 The pseudorandom measures of large order in the degenerate case

In Section 7.2 we showed that in the degenerate case if $\ell \leq r$ then $Q_\ell(\eta)$ is small. Now we will prove that $Q_\ell(\eta)$ is large for some ℓ with $\ell \leq 2^r$.

Theorem 7.2 (Gyarmati, Sárközy, Stewart) *Let $f \in \mathbb{F}_p[x_1, x_2]$ be a degenerate polynomial with rank r and degree m and n in x_1 and x_2, respectively. Then there exists a positive integer ℓ with $\ell \leq 2^r$ for which*

$$Q_\ell(\eta) \geq p^2 - 4rp^{3/2} - 2\ell(m+n)p.$$

Proof of Theorem 7.2 We may assume that $r \leq p^{1/2}/4$ since otherwise the theorem is immediate. Suppose that $f(x_1, x_2)$ has the normal form

$$f(x_1, x_2) = \lambda \prod_{j=1}^{r} f_j(\alpha_j x_1 + \beta_j x_2)\psi(x_1, x_2)^2$$

with $(\alpha_1, \beta_1), \ldots, (\alpha_r, \beta_r)$ distinct elements from T. We distinguish two cases. In the first case all of the α_i's are non-zero. In the second case one of the α_i's is zero and in that case we may suppose, without loss of generality, that $(\alpha_1, \beta_1) = (0, 1)$ (since x_1 and x_2 play symmetric role and if $\alpha_1 = 0$, $\beta_1 \neq 0, 1$ then writing $y = \beta_1 x_2$, x_1, x_2 can be replaced by the variables $x_1, \frac{y}{\beta_1}$). There exists an integer γ_i with $1 \leq |\gamma_i| \leq p^{1/2} + 1$ such that $\gamma_i \alpha_i$ is congruent modulo p to a positive integer of size at most $p^{1/2}$ for $i = 1, \ldots, r$ in the first case and $i = 2, \ldots, r$ in the second case. To see this consider the first $[p^{1/2}] + 2$ multiples of α_i in $\mathbb{F}_p$. Two of them have representations which differ by at most $(p-1)/([p^{1/2}] + 1)$, so by at most $p^{1/2}$, and the difference gives the result. In the second case we may take $\gamma_1 = 1$ so $\gamma_1 \beta_1 = 1$.

Put

$$E = \{\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_r) \text{ with } \varepsilon_i \in \{0, 1\} \text{ for } i = 1, \ldots, r\}$$

and for each $\boldsymbol{\varepsilon}$ in E put

$$\mathbf{d}(\boldsymbol{\varepsilon}) = \varepsilon_1(\beta_1, -\alpha_1)\gamma_1 + \cdots + \varepsilon_r(\beta_r, -\alpha_r)\gamma_r.$$

Notice that for each $\boldsymbol{\varepsilon}$ in E, $\mathbf{d}(\boldsymbol{\varepsilon})$ has coordinates represented by integers between $-r(p^{1/2} + 1)$ and $r(p^{1/2} + 1)$.

Lemma 7.5 (Gyarmati, Sárközy, Stewart)

$$\prod_{\boldsymbol{\varepsilon} \in E} f(\mathbf{x} + \mathbf{d}(\boldsymbol{\varepsilon}))$$

is the square of a polynomial in $\mathbb{F}_p[x_1, x_2]$.

Proof of Lemma 7.5 Write

$$\overline{f_j}(x_1, x_2) = f_j(\alpha_j x_1 + \beta_j x_2),$$

for $j = 1, \ldots, r$, so that

$$f(\mathbf{x}) = \lambda \prod_{j=1}^{r} \overline{f_j}(x_1, x_2)\psi^2(x_1, x_2). \tag{7.11}$$

For each integer j with $1 \le j \le r$ we may split E into two disjoint sets E_j^0 and E_j^1 where $\boldsymbol{\varepsilon}$ in E is in E_j^0 if $\varepsilon_j = 0$ and is in E_j^1 if $\varepsilon_j = 1$. For $\boldsymbol{\varepsilon}$ in E_j^0 let $\boldsymbol{\varepsilon}^1$ denote the element of E_j^1 with the same coordinates as $\boldsymbol{\varepsilon}$ except for the j-th coordinate which is 1. Then, for $\boldsymbol{\varepsilon}$ in E_j^0,

$$\overline{f_j}(\mathbf{x} + \mathbf{d}(\boldsymbol{\varepsilon})) = \overline{f_j}(\mathbf{x} + \mathbf{d}(\boldsymbol{\varepsilon}^1))$$

and so

$$\begin{aligned}\prod_{\boldsymbol{\varepsilon} \in E} \overline{f_j}(\mathbf{x} + \mathbf{d}(\boldsymbol{\varepsilon})) &= \prod_{\boldsymbol{\varepsilon} \in E_j^0} (\overline{f_j}(\mathbf{x} + \mathbf{d}(\boldsymbol{\varepsilon}))\overline{f_j}(\mathbf{x} + \mathbf{d}(\boldsymbol{\varepsilon}^1))) \\ &= \left(\prod_{\boldsymbol{\varepsilon} \in E_j^0} \overline{f_j}(\mathbf{x} + \mathbf{d}(\boldsymbol{\varepsilon})) \right)^2.\end{aligned}$$

The result now follows from (7.11) since $|E|$ is even. □

Let D be the set of $\mathbf{d} = \mathbf{d}(\boldsymbol{\varepsilon})$ which occur with odd multiplicity among the terms $\mathbf{d}(\boldsymbol{\varepsilon})$ with $\boldsymbol{\varepsilon}$ in E. It follows from Lemma 7.5 that if D is non-empty then

$$\prod_{\mathbf{d} \in D} f(\mathbf{x} + \mathbf{d}) \tag{7.12}$$

is the square of a polynomial in $\mathbb{F}_p[x_1, x_2]$.

We claim that $(0,0)$ is in D. Certainly $\mathbf{d}(0, \ldots, 0) = (0,0)$. Further if $\boldsymbol{\varepsilon}$ is in E and $\mathbf{d}(\boldsymbol{\varepsilon}) = (0,0)$ then $\varepsilon_1\alpha_1\gamma_1 + \cdots + \varepsilon_r\alpha_r\gamma_r = 0$. Since $\alpha_i\gamma_i$ is congruent to a positive integer of size at most $p^{1/2}$ and r is at most $p^{1/2}/4$ we see that $\varepsilon_1 = \cdots = \varepsilon_r = 0$ in the first case and that $\varepsilon_2 = \cdots = \varepsilon_r = 0$ in the second case. But in the second case we find that $\mathbf{d}(\boldsymbol{\varepsilon}) = (\varepsilon_1\beta_1\gamma_1, 0) = (\varepsilon_1, 0)$ so $\varepsilon_1 = 0$. Therefore if $\boldsymbol{\varepsilon}$ is in E and $\mathbf{d}(\boldsymbol{\varepsilon}) = (0,0)$ we see that $\boldsymbol{\varepsilon} = (0, \ldots, 0)$ and this shows that $(0,0)$ is in D. Clearly, $|D| \equiv |E| \pmod 2$ and since $|E| = 2^r$ we conclude that

$$2 \le |D| \le |E| = 2^r.$$

Let $\mathbf{d} = (d_1, d_2)$ in D. Then d_1 and d_2 are integers between $-r(p^{1/2}+1)$ and $r(p^{1/2}+1)$. Put

$$d_1^1 = \min_{\mathbf{d} \in D} d_1, \quad d_2^1 = \min_{\mathbf{d} \in D} d_2$$

and

$$\mathbf{d}_0 = (d_1^1, d_2^1).$$

Then $\mathbf{d} - \mathbf{d}_0 \in I_p^2$ for $\mathbf{d} \in D$ since $r \le p^{1/2}/4$. Next put

$$B = \{(x_1, x_2) \in I_p^2 \mid 0 \le x_i < p - 2r(p^{1/2}+1) \text{ for } i = 1, 2\}.$$

Notice that

$$|B| \ge (p - 2r(p^{1/2}+1))^2 \ge p^2 - 4rp^{3/2}. \tag{7.13}$$

Put

$$F(\mathbf{x}) = \prod_{\mathbf{d}\in D} f(\mathbf{x} + \mathbf{d} - \mathbf{d}_0).$$

$F(\mathbf{x})$ is the square of a polynomial in $\mathbb{F}_p[x_1, x_2]$ by (7.12). Let $\ell = |D|$. With η defined by (6.2) we find that

$$\begin{aligned} Q_\ell(\eta) &\ge \left| \sum_{\mathbf{x}\in B} \prod_{\mathbf{d}\in D} \eta(\mathbf{x} + \mathbf{d} - \mathbf{d}_0) \right| \\ &= \left| \sum_{\substack{\mathbf{x}\in B \\ F(\mathbf{x})\neq 0}} \left(\frac{F(\mathbf{x})}{p}\right) + \sum_{\substack{\mathbf{x}\in B \\ F(\mathbf{x})=0}} \prod_{\mathbf{d}\in D} \eta(\mathbf{x} + \mathbf{d} - \mathbf{d}_0) \right| \\ &\ge \sum_{\substack{\mathbf{x}\in B \\ F(\mathbf{x})\neq 0}} 1 - \sum_{\substack{\mathbf{x}\in B \\ F(\mathbf{x})=0}} 1 \ge |\mathcal{B}| - 2 \sum_{\substack{\mathbf{x}\in \mathbb{F}_p^2 \\ F(\mathbf{x})=0}} 1. \end{aligned} \tag{7.14}$$

It is easy to see that if a polynomial $F \in \mathbb{F}_p[x_1, x_2]$ is of degree u and v in x_1 and x_2, respectively, then the number of its zeros $\mathbf{x} \in \mathbb{F}_p^2$ is at most $(u+v)p$. Thus it follows from (7.13) and (7.14) that

$$Q_\ell(\eta) \ge p^2 - 4rp^{3/2} - 2\ell(m+n)p$$

which proves Theorem 7.2. □

7.4 Generating a large family of suitable polynomials

In this section we construct a large family of polynomials which are non-degenerate.

Theorem 7.3 (Gyarmati, Sárközy, Stewart) *Let $f \in \mathbb{F}_p[x_1, x_2]$ be a polynomial of the form*

$$f(x_1, x_2) = x_1^k + x_1 x_2 g(x_1, x_2) + x_2 h(x_2) \tag{7.15}$$

with $g \in \mathbb{F}_p[x_1, x_2]$, $\deg g \leq k-3$, $h \in \mathbb{F}_p[x_2]$, $\deg h(x_2) \leq k-2$ *and* $x_2 \nmid h(x_2)$. *Then for the binary lattice* η *defined in* (6.2) *we have*

$$Q_\ell(\eta) < 11k\ell p^{3/2} \log p. \tag{7.16}$$

Proof of Theorem 7.3 We will need the following generalization of the Schönemann-Eisenstein theorem.

Lemma 7.6 *If* $f(x) = a_0x^n + \cdots + a_n$ *is a polynomial over an integral domain* R *and* $\mathfrak{a}$ *is a maximal ideal of* R *with*

$$\begin{aligned} a_0 &\not\equiv 0 \pmod{\mathfrak{a}}, \\ a_1 \equiv \cdots \equiv a_n &\equiv 0 \pmod{\mathfrak{a}}, \\ a_n &\not\equiv 0 \pmod{\mathfrak{a}^2} \end{aligned}$$

then $f(x)$ *cannot be decomposed in* $R[x]$ *into a product of non-constant factors.*

Proof of Lemma 7.6 See, for example [93, Theorem 282]. □

$R = \mathbb{F}_p[x_2]$ is an integral domain and $\mathfrak{a} =< x_2 >$ is a maximal ideal in it. Then the conditions of Lemma 7.6 hold for the polynomial $f(x_1, x_2) \in R[x_1]$ in (7.15), thus $f(x_1, x_2)$ is irreducible.

In order to use Theorem 7.1 we prove that $f(x_1, x_2)$ is not of the form (7.5). Since $f(x_1, x_2)$ is irreducible we have to prove that $f(x_1, x_2)$ is not of the form

$$f(x_1, x_2) = f_1(\alpha_1 x_1 + \beta_1 x_2). \tag{7.17}$$

Let h be the degree of f_1 and consider the terms of degree h in f_1, so

$$f_1(\alpha_1 x_1 + \beta_1 x_2) = c(\alpha_1 x_1 + \beta_1 x_2)^h + f_2(\alpha_1 x_1 + \beta_1 x_2),$$

where the degree of $f_2(\alpha_1 x_1 + \beta_1 x_2)$ is $\leq h-1$ and $c \neq 0 \in \mathbb{F}_p$. Clearly, $c(\alpha_1 x_1 + \beta_1 x_2)^h$ equals the sum of the terms of degree k of $f(x_1, x_2)$, thus by the conditions of Theorem 7.2 we have

$$c(\alpha_1 x_1 + \beta_1 x_2)^h = x_1^k.$$

We may suppose that k is less than p since the result is immediate otherwise. It then follows that $h = k$, $c = \alpha_1 = 1$ and $\beta_1 = 0$, thus from (7.17)

$$f(x_1, x_2) = f_1(x_1). \tag{7.18}$$

On the other hand $f(x_1, x_2)$ contains a power of x_2, and this contradicts (7.18). Thus $f(x_1, x_2)$ is not of the form (7.5). We have also proved that $f(x_1, x_2)$ is irreducible, and by using Theorem 7.1 a) we obtain the result.□

7.5 A Legendre symbol construction with optimal bounds

As we remarked already in [61], our upper bounds are not optimal; in particular, in (7.16) the optimal upper bound would be, up to logarithmic factors, p (with a factor depending on k and ℓ). On the other hand this construction is more natural than the ones using finite fields in [65], [79] or [80] (where the bounds are sharper), and it can be implemented faster. However, we will show that for a certain (rather special) family of polynomials the finite field construction presented in [79] is equivalent to a Legendre symbol construction of type (6.2). Thus in this case we obtain a family of binary lattices which combines the advantages of the two constructions: as in [79] we have optimal bounds, and as a Legendre symbol construction it can be implemented fast and easily.

Indeed, combining Theorems 7.1 and 7.2 of [79], we get the following result:

Theorem 7.A (Mauduit, Sárközy) *Let p be an odd prime, $n \in \mathbb{N}$, $q = p^n$, and 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. Let $f(\mathbf{x}) \in \mathbb{F}_q[\mathbf{x}]$ be a polynomial of degree k with*

$$0 < k < p \tag{7.19}$$

which has no multiple zero. Define the n-dimensional binary p-lattice $\eta(\mathbf{x}) : I_p^n \to \{-1, +1\}$ by

$$\begin{aligned} \eta(\mathbf{x}) &= \eta((x_1, \dots, x_n)) \\ &= \begin{cases} \gamma(f(x_1v_1 + \cdots + x_nv_n)) & \text{for } f(x_1v_1 + \cdots + x_nv_n) \neq 0 \\ 1 & \text{for } f(x_1v_1 + \cdots + x_nv_n) = 0. \end{cases} \end{aligned} \tag{7.20}$$

Assume also that $\ell \in \mathbb{N}$ with

$$4^{n(k+\ell)} < p. \tag{7.21}$$

Then we have

$$Q_\ell(\eta) < k\ell \left(q^{1/2}(1 + \log p)^n + 2\right). \tag{7.22}$$

Our next result follows from Theorem 7.A in the case that $n = 2$ and for

a special choice of v_1, v_2 and the polynomial f.

Theorem 7.4 (Gyarmati, Sárközy, Stewart) *Let p be an odd prime and let r be a quadratic non-residue modulo p. Then the polynomial $x^2 - r$ is irreducible over $\mathbb{F}_p$; denote one of its zeros by θ, and consider the extension of $\mathbb{F}_p$ by θ: $\mathbb{F}_p[\theta](\cong \mathbb{F}_{p^2})$. Let k and ℓ be integers which satisfy (7.19) and (7.21), and assume that $a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_k \in \mathbb{F}_p$ satisfy*

$$a_i + b_i\theta \neq a_j + b_j\theta \text{ and } a_i + b_i\theta \neq a_j - b_j\theta \text{ for } 1 \leq i < j \leq k. \tag{7.23}$$

Put

$$\tilde{f}(x_1, x_2) = \prod_{i=1}^{k} \left((x_1 - a_i)^2 - r(x_2 - b_i)^2\right) \tag{7.24}$$

and

$$\tilde{\eta}(\mathbf{x}) = \tilde{\eta}(\mathbf{x}) = \tilde{\eta}((x_1, x_2)) = \begin{cases} \left(\frac{\tilde{f}(x_1,x_2)}{p}\right) & \text{if } (\tilde{f}(x_1, x_2), p) = 1 \\ 1 & \text{if } p \mid \tilde{f}(x_1, x_2). \end{cases} \tag{7.25}$$

For each positive integer ℓ with

$$4^{2(\ell+k)} < p \tag{7.26}$$

we have

$$Q_\ell(\tilde{\eta}) < \ell k \left(p(1 + \log p)^2 + 2\right).$$

Proof of Theorem 7.4 By the definition of θ and Euler's lemma, we have

$$\theta^p = (\theta^2)^{\frac{p-1}{2}}\theta = r^{\frac{p-1}{2}}\theta = -\theta. \tag{7.27}$$

We will use Theorem 7.A with $n = 2$, $q = p^2$, $v_1 = 1$, $v_2 = \theta$, so that now the elements of $\mathbb{F}_q = \mathbb{F}_{p^2}$ are represented in the form $x_1 + x_2\theta$. Then by the generalization of Euler's lemma to $\mathbb{F}_q$ and (7.27), for $x_1 + x_2\theta \in \mathbb{F}_{p^2}^*$, so with $(x_1, x_2) \neq (0, 0)$, we have

$$\begin{aligned} \gamma(x_1 + x_2\theta) &= (x_1 + x_2\theta)^{\frac{p^2-1}{2}} = (x_1 + x_2\theta)^{\frac{p^2-p}{2}}(x_1 + x_2\theta)^{\frac{p-1}{2}} \\ &= ((x_1 + x_2\theta)^p)^{\frac{p-1}{2}}(x_1 + x_2\theta)^{\frac{p-1}{2}} = (x_1^p + x_2^p\theta^p)^{\frac{p-1}{2}}(x_1 + x_2\theta)^{\frac{p-1}{2}} \\ &= (x_1 - x_2\theta)^{\frac{p-1}{2}}(x_1 + x_2\theta)^{\frac{p-1}{2}} = (x_1^2 - x_2^2\theta^2)^{\frac{p-1}{2}} = (x_1^2 - rx_2^2)^{\frac{p-1}{2}} \\ &= \left(\frac{x_1^2 - rx_2^2}{p}\right). \end{aligned}$$

By the multiplicativity of γ and the Legendre symbol, it follows that writing

$$f(x_1 + x_2\theta) = \prod_{i=1}^{k} ((x_1 + x_2\theta) - (a_i + b_i\theta)) \tag{7.28}$$

and defining $\eta(\mathbf{x}) = \eta((x_1, x_2))$ as in (7.20) we have

$$\begin{aligned}\eta(\mathbf{x}) &= \gamma(f(x_1 + x_2\theta)) = \gamma\left(\prod_{i=1}^{k}((x_1 + x_2\theta) - (a_i + b_i\theta))\right) \\ &= \prod_{i=1}^{k} \gamma\left((x_1 + x_2\theta) - (a_i + b_i\theta)\right) = \prod_{i=1}^{k} \gamma\left((x_1 - a_i) + (x_2 - b_i)\theta\right) \\ &= \prod_{i=1}^{k} \left(\frac{(x_1 - a_i)^2 - r(x_2 - b_i)^2}{p}\right) = \left(\frac{\prod_{i=1}^{k}((x_1 - a_i)^2 - r(x_2 - b_i)^2)}{p}\right) \\ &= \left(\frac{\tilde{f}(x_1, x_2)}{p}\right) = \tilde{\eta}(\mathbf{x}) \quad (\text{for } f(x_1 + x_2\theta) \neq 0)\end{aligned} \tag{7.29}$$

with the polynomial $\tilde{f}$ and the lattice $\tilde{\eta}$ defined by (7.24) and (7.25), respectively, and trivially we have

$$\eta(\mathbf{x}) = \tilde{\eta}(\mathbf{x}) \text{ for } f(x_1 + x_2\theta) = 0. \tag{7.30}$$

By (7.23) and the definition of r, the polynomial $\tilde{f}$ has no multiple zero, and now (7.21) holds by (7.26). Thus Theorem 7.A can be applied, and then we obtain from (7.22), (7.29) and (7.30) that

$$Q_\ell(\eta) = Q_\ell(\tilde{\eta}) < \ell k \left(p(1 + \log p)^2 + 2\right)$$

which completes the proof of Theorem 7.4.

We remark that the construction in Theorem 7.4 could be extended by also considering higher degree factors in (7.28). Even more generally, we may consider polynomials f which are not given in a product form. In either case, we may use the fact that if $f(x_1 + x_2\theta) = p(x_1, x_2) + \theta q(x_1, x_2)$ (with $f(z) \in \mathbb{F}_p[z]$, $p(x_1, x_2),\ q(x_1, x_2) \in \mathbb{F}_p[x_1, x_2]$ and θ, r defined as above), then we have

$$\gamma(f(x_1 + \theta x_2)) = \gamma(p(x_1, x_2) + \theta q(x_1, x_2)) = \left(\frac{p^2(x_1, x_2) - rq^2(x_1, x_2)}{p}\right).$$

However this would make the polynomial $\tilde{f}$ in (7.24) in Theorem 7.4 much

more complicated.

Finally, we would like to discuss the implementation of the construction in Theorem 7.4. The critical point of the implementation is to find a quadratic non-residue r. If p is fixed, then it is known that the GRH implies that the least quadratic non-residue modulo p is less than $(\log p)^c$ (with some positive constant c), and since the quadratic character of a given residue can be decided in polynomial time (by using Jacobi symbols), r can be chosen as the least quadratic non-residue modulo p which can be determined in polynomial time. On the other hand, no algorithm is known for finding the least quadratic non-residue in polynomial time without any unproved hypothesis. However, in most cases one need not fix p, and this difficulty can be avoided. Namely, we may start out from the fact that if p is a prime of the form $4k-1$, then -1 is a quadratic non-residue modulo p. Thus it is worthwhile to make first a long sequence of primes $p_1 = 3 < p_2 < \cdots < p_t$ of the form $4k-1$ with say, $p_i < p_{i+1} < 2p_i$, and if we need a prime p of size about N with $p \equiv -1 \pmod 4$, then we take the first prime from this sequence greater than N, and we take $r = -1$. (If we want a large prime p of the form $4k-1$, then we may use the fact that the Mersenne primes are of the form $4k-1$.)

8 Further results

In this section, I will write a few sentences about my further papers (written partly with my coauthors) about pseudorandomness, which I have written since my PhD:

With Attila Pethő and András Sárközy in [59] we studied a pseudorandom generator based on linear recursions. This construction has several advantages: easy implementation and we were able to prove optimal bounds for the pseudorandom measures.

In [37] I sharpened some earlier estimates on the pseudorandom measures of a construction based on the discrete logarithm. (This earlier construction was defined in [38] and my PhD dissertation [35].) In [44] I extend this construction. This generalization has the interesting property that in special cases we get pseudorandom sequences based on elliptic curves.

The connection between the pseudorandomness of binary sequences and binary lattices is studied in [50]. From a two-dimensional binary N-lattice one can make a unique binary sequence of length N^2 by taking first the first row of the lattice then continuing the sequence by the second row of the matrix, etc. In [50] we showed that the lattice may have weak pseudorandom properties, however, the associated sequence has strong pseudorandom properties. In Theorem 5.3 I proved a result pointing the opposite direction, moreover if the lattice has strong pseudorandom properties, then the associated sequence also has (see Section 5).

In the applications it may occur that the initial pseudorandom sequence turns out to be not long enough, thus we have to take the concatenation or merging of it with another pseudorandom sequences. I studied this problem in [40] and [42].

Three different constructions of binary lattices with strong pseudorandom properties are given in [51]. These constructions are the two dimensional extensions and modifications of three of the most important one dimensional constructions.

In [52], [53], [54] we studied the pseudorandom measures of two-dimensional binary lattices with my coauthors. Thus in [52] we compared the different pseudorandom measures and we estimated the normality measure by the maximum of Q_ℓ measures. In [54] we studied the symmetry properties of binary lattices. Finally, in [53] we introduced the multidimensional version of the correlation measure C_ℓ and we estimated the minimum of the measures C_ℓ and Q_ℓ.

In [43] I realized that the shape of the box-lattices B in Definition 1.6 is of very special type. Sometimes we have to cover more general situation, so in [43] I introduced further new measures. I introduced the convex and line

measure and studied the connections between the new and old measures. I show that there exists a special case of the Legendre symbol construction (see Construction 1.2) for which a strong upper bound can be given for these much more general measures.

In [55], [56] with Christian Mauduit and András Sárközy we studied the following problems: In cryptographic applications sometimes is not enough that the binary lattices have strong pseudorandom properties, but it is also important that their large family contains "significantly" different lattices. The collision and avalanche effect study this property. In the one-dimensional case these notions are studied for example in the papers [7], [22], [66], [81], [103], [104]. In [55] we generalized the collision and avalanche effect for the multidimensional case, and we define new measures. In [56] we presented a further construction for which these new measures are optimal.

The linear complexity is an important and frequently used measure of unpredictability and pseudorandomness of binary sequences. In [57] and [58] we extend this notion to two dimensions. We estimated the linear complexity of a truly random binary lattice. We analyzed the connection between the linear complexity and the correlation measures, and we utilized the inequalities obtained in this way for estimating the linear complexity of an important special binary lattice. We studied connection between the linear complexity of binary lattices and of the associated binary sequences. We extend the notion of k-error linear complexity to bit lattices. Finally, we present another alternative definition of linear complexity of bit lattices.

Pseudorandomness can be defined on various different objects. In [47] and [48] with Pascal Hubert and András Sárközy we studied pseudorandom binary functions on trees.

In [46] I presented a survey of the most important results involving the new quantitative pseudorandom measures of finite binary sequences and lattices.

References

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

[2] R. Ahlswede, C. Mauduit and A. Sárközy, *Large families of pseudorandom sequences of k symbols and their complexity, Part I, Part II.*, Lecture Notes in Computer Science, Springer Berlin / Heidelberg 2006, Volume 4123/2006, 293-325.

[3] N. Alon, Y. Kohayakawa, C. Mauduit, C. G. Moreira and V. Rödl, *Measures of pseudorandomness for finite sequences: minimal values*, Combin., Probab. Comput. 15 (2005), 1-29.

[4] 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.

[5] V. Anantharam, *A technique to study the correlation measures of binary sequences*, Discrete Math. 308, 24 (2008), 6203 -6209.

[6] J. Beck, *Roth's estimate on the discrepancy of integer sequences is nearly sharp*, Combinatorica 1 (1981), 319-325.

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

[8] L. Blum, M. Blum and M. Shub, *A simple unpredictable pseudorandom number generator*, SIAM J. Comp. 15 (1986), 364-383.

[9] J. Bourgain, *Mordell's exponential sum estimate revisited*, J. Amer. Math. Soc. 18 (2005), 477-499.

[10] J. Bourgain, T. Cochrane, J. Paulhus and C. Pinner, *On the parity of k-th powers mod p, a generalization of a problem of Lehmer*, Acta Arith. 147 (2011), 173-203.

[11] J. J. Brennan and B. Geist, *Analysis of iterated modular exponentiation: The orbit of x^α mod N*, Designs, Codes and Cryptography 13 (1998), 229-245.

[12] D. A. Burgess, *The distribution of quadratic residues and non-residues*, Mathematika 4 (1957) 106-112.

[13] J. Cassaigne, S. Ferenczi, C. Mauduit, J. Rivat and A. Sárközy, *On finite pseudorandom binary sequences III: The Liouville function, I*, Acta Arith. 87 (1999), 367-384.

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

[15] J. W. S. Cassels, *On a paper of Niven and Zuckerman*, Pacific J. Math. 2 (1952), 555-557.

[16] Z. Chen, *Elliptic curve analogue of Legendre sequences*, Monatshefte für Mathematik, 154 (2008), 1-10.

[17] Z. Chen, S. Li and G. Xiao, *Construction of pseudo-random binary sequences from elliptic curves by using discrete logarithm*, in: Sequences and their applications - SETA 2006, Lecture Notes in Computer Science 4086, Berlin ; Heidelberg: Springer Verlag, 2006, 285-294.

[18] T. W. Cusick, *Properties of the x^2 mod N pseudorandom number generator*, IEEE Trans. Inform. Theory 41 (1995), 1155-1159.

[19] T. W. Cusick, C. Ding and A. Renvall, *Stream Ciphers and Number Theory*, Elsevier, North-Holland Publishing Co., Amsterdam 1998.

[20] H. Davenport and D. J. Lewis, *Character sums and primitive roots in finite fields*, Rend. Circ. Mat. Palermo (2) 12 (1963), 129-136.

[21] P. Erdős and A. Sárközy, *Some solved and unsolved problems in combinatorial number theory*, Math. Slovaca 28 (1978), 407-421 (page 415).

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

[23] R. Fischlin and C. P. Schnorr, *Stronger Security proofs for RSA and Rabin bits*, Lecture Notes in Comp. Sci., Springer-Verlag, Berlin, 1233 (1997), 267-279.

[24] E. Fouvry and N. Katz, *A general stratification theorem for exponential sums, and applications*, J. Reine Angew. Math. 540 (2001), 115-166.

[25] J. B. Friedlander, J. Hansen and I. Shparlinski, *Character sums with exponential functions*, Mathematika, 47 (2000), 75-85.

[26] J. B. Friedlander, D. Lieman and I. E. Shparlinski, *On the distribution of the RSA generator*, Proc. Itern. Conf. on Sequences and Their Applications (SETA'98), Singapore, Springer-Verlag, London 1999, 205-212.

[27] J. B. Friedlander, C. Pomerance and I. E. Shparlinski, *Period of the power generator and small values of the Carhmichael's function*, Math Comp. 70 (2001), no. 236, 1591-1605.

[28] J. B. Friedlander and I. E. Shparlinski, *On the distribution of the power generator*, Math Comp. 70 (2001), no. 236, 1575-1589.

[29] C. F. Gauss, *Untersuchungen Über höhere Arithmetik*, Chelsea publishing company, second edition, reprinted, New York 1981.

[30] S. Goldwasser, *Mathematical Foundations of Modern Cryptography: Computational Complexity Perspective*, ICM 2002, vol., I, 245-272.

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

[32] F. Griffin and I. E. Shparlinski, *On the linear complexity profile of the power generator*, IEEE Trans. Inform. Theory 46 (2000), no. 6, 2159-2162.

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

[34] K. Gyarmati, *On a pseudorandom property of binary sequences*, Ramanujan J. 8 (2004), 289-302,

[35] K. Gyarmati, *On pseudorandom binary sequences*, PhD dissertation, Budapest 2004.

[36] K. Gyarmati, *On the correlation of binary sequences*, Studia Sci. Math. Hungar. 42 (2005), 59-75.

[37] K. Gyarmati, *A note to the paper "On a fast version of a pseudorandom generator"*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. 49 (2006), 87-93.

[38] K. Gyarmati, *On a fast version of a pseudorandom generator*, General Theory of Information Transfer and Combinatorics, Lecture Notes in Computer Science, Vol. 4123, Springer, Berlin/Heidelberg 2006, 326-342.

[39] K. Gyarmati, *Pseudorandom sequences constructed by the power generator*, Period. Math. Hungar. 52 (2006) 9-26.

[40] K. Gyarmati, *Concatenation of pseudorandom binary sequences*, Period. Math. Hung. 58 (2009), 99-120.

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

[42] K. Gyarmati, *Concatenation of Legendre symbol sequences*, Studia Sci. Math. Hungarica 48 (2011), 193-204.

[43] K. Gyarmati, *On new measures of pseudorandomness of binary lattices*, Acta Math. Hung. 131 (2011), 346-359.

[44] K. Gyarmati, *Elliptic curve analogues of a pseudorandom generator*, Period. Math. Hungar. 64 (2012), 119-130.

[45] K. Gyarmati, *On the correlation of subsequences*, Unif. Distrib. Theory 7 (2012), no. 2, 169-195.

[46] K. Gyarmati, *Measures of pseudorandomness*, Finite fields and applications: character sums and polynomials, P. Charpin, A. Pott, A. Winterhof (eds.), Radon Series in Computational and Applied Mathematics, de Gruyter, to appear.

[47] K. Gyarmati, P. Hubert and A. Sárközy, *Pseudorandom binary functions on almost uniform trees*, J. Combin. Number Theory 2 (1), (2010), 1-24.

[48] K. Gyarmati, P. Hubert and A. Sárközy, *Pseudorandom binary functions on rooted plane trees*, J. Combin. Number Theory 4 (1), (2012), 1-19.

[49] K. Gyarmati and C. Mauduit, *On the correlation of binary sequences, II*, Discrete Math. 312 (2012), 811-818.

[50] K. Gyarmati, C. Mauduit and A. Sárközy, *Pseudorandom binary sequences and lattices*, Acta Arith. 135 (2008), 181-197.

[51] K. Gyarmati, C. Mauduit and A. Sárközy, *Constructions of pseudorandom binary lattices*, Uniform Distribution Theory 4 (2), (2009), 59-80.

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

[53] K. Gyarmati, C. Mauduit and A. Sárközy, *Measures of pseudorandomness of binary lattices, III. (Q_k, correlation, normality, minimal values.)*, Unif. Distrib. Theory 5 (2010), 183-207.

[54] K. Gyarmati, C. Mauduit and A. Sárközy, *Measures of pseudorandomness of finite binary lattices, II (The symmetry measures.)*, Ramanujan J. 25 (2), (2011), 155-178.

[55] K. Gyarmati, C. Mauduit and A. Sárközy, *Measures of pseudorandomness of families of binary lattices, I (Definitions, a construction using quadratic characters.)*, Publi. Math. Debrecen 79 (3), (2011), 445-460.

[56] K. Gyarmati, C. Mauduit and A. Sárközy, *Measures of pseudorandomness of families of binary lattices, II (A further construction.)*, Publi. Math. Debrecen 80 (3), (2012), 479-502.

[57] K. Gyarmati, C. Maduit and A. Sárközy, *On linear complexity of binary lattices*, Ramanujan J., to appear.

[58] K. Gyarmati, C. Maduit and A. Sárközy, *On linear complexity of binary lattices, II*, submitted.

[59] K. Gyarmati, A. Pethő and A. Sárközy, *On linear recursion and pseudorandomness*, Acta Arith. 118 (4), (2005), 359-374.

[60] K. Gyarmati and A. Sárközy, *Equations in finite fields with restricted solution sets, I. (Character sums.)*, Acta Math. Hungar. 118 (2008), 129-148.

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

[62] K. Gyarmati, A. Sárközy and C. L. Stewart, *On Legendre symbol lattices, II*, Unif. Distrib. Theory, to appear.

[63] J. Håstad and M. Näslund, *The security of individual RSA bits*, Proc 39th IEEE Symp. on Foundations of Comp. Sci., 1998, 510-519.

[64] J. Hoffstein and D. Lieman, *The distribution of the quadratic symbol in function fields and a faster mathematical stream cipher*, Progress in Computer Science and Applied Logic, Vol. 20, Birkhäuser, Verlag, Basel, 2001; pp. 59-68.

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

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

[67] N. M. Katz, *An estimate for character sums*, J. Amer. Math. Soc. 2, No. 2 (1989), 197-200.

[68] D. E. Knuth, *The Art of Computer Programming*, Vol. 2, 2nd ed., Addison-Wesley, Reading, Mass., 1981.

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

[70] J. C. Lagarias, *Pseudorandom number generators in cryptography and number theory*, Proc. Symp. in Appl. Math., Amer. Math. Soc., Providence, RI, 42 (1990), 115-143.

[71] R. Lidl and H. Niderreiter, *Finite Fields*, Second edition, Cambridge University Press. 1997.

[72] H. Liu, T. Zhan and X. Wang, *Large families of elliptic curve pseudorandom binary sequences*, Acta Arith. 140 (2009), 135-144.

[73] H. Liu, T. Zhan and X. Wang, *On the correlation of pseudorandom binary sequences with composite moduli*, Publ. Math. Debrecen 74 (2009), 195-214.

[74] J. Matoušek and J. Spencer, *Discrepancy in arithmetic progressions*, J. Amer. Math. Soc. 9 (1996), 195-204.

[75] C. Mauduit, *Construction of pseudorandom finite sequences*, unpublished lecture notes to the conference, Information Theory and Some Friendly Neighbours- ein Wunschkonzert, Bielefeld, 2003.

[76] C. Mauduit, J. Rivat and A. Sárközy, *Construction of Pseudorandom Binary Sequences Using Additive Characters*, Monatsh. Math. 141, (2004), 197-208.

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

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

[79] C. Mauduit and A. Sárközy, *On large families of pseudorandom binary lattices*, Unif. Distrib. Theory 2 (2007), no. 1, 23-37.

[80] C. Mauduit and A. Sárközy, *Construction of pseudorandom binary lattices by using the multiplicative inverse*, Monatsh. Math. 153 (2008), 217-231.

[81] A. J. Menezes, P. C. van Oorschot and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, FL, 1997.

[82] L. Mérai, *Pszeudovéletlen sorozatok és rácsok*, PhD dissertation, 2010.

[83] L. Mérai, *Construction of pseudorandom binary lattices using elliptic curves*, Proc. Amer. Math. Soc. 139 (2011), no. 2, 407-420.

[84] L. Mérai, *Construction of pseudorandom binary sequences over elliptic curves using multiplicative characters*, Publ. Math. Debrecen, 80 (2012), 199-213.

[85] L. Mérai, *Remarks on pseudorandom binary sequences over elliptic curves*, Fund. Inform. 114 (2012), 301-308.

[86] H. Niederreiter, *On the distribution of pseudo-random numbers generated by the linear congruential method.*, Math. Comp. 26 (1972), 793-795.

[87] H. Niederreiter, *On the distribution of pseudo-random numbers generated by the linear congruential method. II*, Math. Comp. 28 (1974), 1117-1132.

[88] H. Niederreiter, *On the distribution of pseudo-random numbers generated by the linear congruential method. III*, Math. Comp. 30 (1976), 571-597.

[89] H. Niederreiter, *Quasi-Monte Carlo methods and pseudorandom numbers*, Bull. Amer. Math. Soc. 84 (1978), 957-1041.

[90] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Regional Conference Series in Applied Math., Vol. 63, Soc. Industr. Applied Math., Philadelphia, 1992.

[91] I. Niven and H. S. Zuckerman, *On the definition of normal numbers*, Pacific J. Math. 1 (1951), 103-109.

[92] G. I. Perel'muter and I. Shparlinski, *Distribution of primitive roots in finite fields* Uspechi Matem. Nauk **45** (1990), no. 1, 185-186 (in Russian).

[93] L. Rédei, *Algebra*, Pergamon Press, Oxford-New York-Toronto, Ont. 1967.

[94] J. Rivat and A. Sárközy, *Modular Constructions of pseudorandom binary sequences with composite moduli*, Period. Math. Hungar. 51 (2005), 75-107.

[95] J. Rivat and A. Sárközy, *On pseudorandom sequences and their application*, Lecture Notes in Comput. Sci. 4123, General theory of information transfer and combinatorics, Springer, Berlin / Heidelberg, 2006, 343-361.

[96] K. F. Roth, *Remark concerning integer sequences*, Acta Arith. 9 (1964), 257-260.

[97] A. Sárközy, *On finite pseudorandom binary sequences and their applications in cryptography*, Tatra Mt. Math. Publ. 37 (2007), 123-136.

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

[99] W. M. Schmidt, *Equations over Finite Fields*, Springer-Verlag, Berlin · Heidelberg · New York 1976.

[100] I. E. Shparlinski, *On the linear complexity of the power generator*, Designs, Codes and Cryptography 23 (2001) no. 1, 5-10.

[101] D. R. Stinson, *Cryptography: Theory and Practice*, CRC Press, Boca Raton, FL, 1995.

[102] B. Sziklai, *On the symmetry of finite pseudorandom binary sequences*, Uniform Distribution Theory 6 (2011), 143-156.

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

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

[105] D. Wan, *Generators and irreducible polynomials over finite fields*, Math. Comput. 66 (219) (1997), 1195-1212.

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

[107] A. Winterhof, *Some estimates for character sums and applications*, Des. Codes Cryptogr. 22 (2001), 123–131.

[108] A. Winterhof, *Measures of pseudorandomness*, in: S. Boztas, (ed.), CRC Handbook of Sequences, Codes and Applications: Chapman and Hall/CRC Press, to appear.