sciendo

uniform distribution theory

DOI: 10.2478/udt-2019–0017
Unif. Distrib. Theory **14** (2019), no.1, 103–126

QUASI-RANDOM GRAPHS, PSEUDO-RANDOM GRAPHS AND PSEUDORANDOM BINARY SEQUENCES, I. (QUASI-RANDOM GRAPHS)

József Borbély — András Sárközy

University of Óbuda, AMK, Székesfehérvár, HUNGARY

Department of Algebra and Number Theory, Eötvös Loránd University, Budapest, HUNGARY

ABSTRACT. In the last decades many results have been proved on pseudo-randomness of binary sequences. In this series our goal is to show that using many of these results one can also construct large families of quasi-random, pseudo-random and strongly pseudo-random graphs. Indeed, it will be proved that if the first row of the adjacency matrix of a circulant graph forms a binary sequence which possesses certain pseudorandom properties (and there are many large families of binary sequences known with these properties), then the graph is quasi-random, pseudo-random or strongly pseudo-random, respectively. In particular, here in Part I we will construct large families of quasi-random graphs along these lines. (In Parts II and III we will present and study constructions for pseudo-random and *strongly* pseudo-random graphs, respectively.)

Communicated by János Pintz

2010 Mathematics Subject Classification: 05C80.
Keywords: quasi-random graph, pseudo-random graph, pseudorandom binary sequence.
Research partially supported by Hungarian National Research Development and Innovation Fund K119528.

1. Introduction

The notion of pseudorandomness has many applications and it appears in many different fields of modern mathematics. Clearly, the study of the interactions between the applications in these fields may lead to interesting new results. In this series our goal is to study such an interaction: we will show that many results and constructions from the theory of *pseudorandomness of binary sequences* (motivated originally by applications in cryptography) can be adjusted to utilize them in studying *quasi-randomness and pseudo-randomness of graphs*. (We remark that in graph theory the words quasi-random and pseudo-random are spelled with hyphen while in case of binary sequences it is more customary to write the word "pseudorandom" without hyphen. We will keep this spelling here, since in this way we will be able to specify the meaning of the word "pseudorandom" just by adding or omitting the hyphen and no further explanation will be needed.)

The notion of pseudo-random graphs (in particular, the definition and properties of "(p, α)-jumbled" graphs) was introduced and studied by A. Thomason [21], [22] while the notion of quasi-random graphs was defined and investigated by F. R. K. Chung, R. L. Graham and R. M. Wilson [5], [6]. As they write in [6]: "We follow much in the spirit of the recent seminal paper of Thomason [21]." Indeed, the notions of pseudo-randomness and quasi-randomness of graphs are related, the philosophy behind them is similar, but the aims and tools of the two approaches are different and they focus on different graph properties.

First here in Part I we will study the applicability of the theory of pseudorandom binary sequences for constructing *quasi-random* graphs (while in Part II *pseudo-random*, in Part III *strongly pseudo-random* graphs will be constructed). Namely, the construction of quasi-random graphs will be simpler and more transparent, while for constructing pseudo-random graphs some adjustments and further work will be needed. In Sections 2 and 3 we will present the basic definitions, facts and results on quasi-random graphs and pseudorandom binary sequences, respectively, which will be needed later. In Section 4 we will describe the general construction principle, while in Section 5 we will present several special constructions. Throughout this series we will follow the notation and terminology used in graph theory (in particular, in [5], [6], [21], [22]) and the one used in the theory of pseudorandomness of binary sequences (in particular, in [4], [11], [16]) possibly closely. However, in some cases we will simplify the notation slightly. Moreover, in certain cases there is conflict between the notation used in the two fields. In particular, the letter p will be reserved for the edge density (in other words, for the probability of joining two vertices), thus the primes used in constructions will be denoted by q instead of the customary p.

The number of vertices of graphs is usually denoted by n, while the standard notation for the length of binary sequences is N. However, in some constructions these quantities coincide (or nearly coincide); then we will be forced to stick to one of the two letters. We will represent a finite field $\mathbb{F}_q$ (with q prime) by the field of the modulo q residue classes, and we will use the same notation for a residue class and an integer representing it.

2. Quasi-random graphs

We will write $G = (V, E)$ for a (finite) graph G with vertex set V and edge set E. If G has n vertices then we will also write $G = G(n)$. The number of edges of G is denoted by $e(G)$: $e(G) = |E|$. For $G = G(n)$, $V = \{v_1, v_2, \dots, v_n\}$, $i \in \{1, 2, \dots, n\}$ the set of the vertices joined to v_i is denoted by V_i : $V_i = \big\{v_j \in V : \{v_i, v_j\} \in E\big\}$ so that the degree of v_i, denoted by $\deg(v_i)$, is $\deg(v_i) = |V_i|$. For some other graph $G' = (V', E')$, $N^*_G(G')$ denotes the number of labeled occurrences of G' as an induced subgraph of G, and define $N_G(G')$ as the number of occurrences of G' as a (not necessarily induced) subgraph of G.

For some graph $G = G(n)$ with vertex set $V = \{v_1, v_2, \dots, v_n\}$, let $A(G) = [a(i,j)]_{i,j\in\{1,2,\dots,n\}}$ denote the *adjacency matrix* of G defined by setting $a(i,j) = 1$ if v_i and v_j are joined, and 0, otherwise. Denote the eigenvalues of $A(G)$ (which are real since $A(G)$ is symmetric) by $\lambda_1, \lambda_2, \dots, \lambda_n$ ordered so that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$.

In the abstract of their paper [5] introducing the notion of quasi-random graphs, Chung, Graham and Wilson write: "We introduce a large equivalence class of graph properties, all of which are shared by so-called random graphs. Unlike random graphs, however, it is often relatively easy to verify that a particular family of graphs possesses some property in this class." In their paper, they use the phrase "random graphs" in the sense that they consider a random graph $G = G(n)$ on n vertices so that its edges are taken independently each with probability $1/2$. The graph properties studied by them are the following:

Take a family $\mathcal{F}$ of graphs $G(n)$ with $n \to \infty$. Then the seven properties considered are:

$\mathbf{P}_1(s)$: For all graphs $H(s)$ on s vertices we have

$$N^*_G\big(H(s)\big) = \big(1 + o(1)\big)n^s 2^{-\binom{s}{2}}.$$

(So that all the $2^{\binom{s}{2}}$ labeled graphs $H(s)$ on s vertices occur asymptotically the same number of times in G.)

$\mathbf{P}_2(t)$: $e(G) \geq \big(1+o(1)\big)\frac{n^2}{4}$, $N_G(C_t) \leq \big(1+o(1)\big)\left(\frac{n}{2}\right)^t$, where C_t denotes the cycle with t vertices.

$\mathbf{P}_3$: $e(G) \geq \big(1+o(1)\big)\frac{n^2}{4}$, $\lambda_1 = \big(1+o(1)\big)\frac{n}{2}$, $\lambda_2 = o(n)$.

$\mathbf{P}_4(\varepsilon)$: For each subset $S \subseteq V$ with $|S| \geq \varepsilon n$ we have

$$e(S) = \big(1+o(1)\big)\frac{|S|^2}{4}.$$

$\mathbf{P}_5$: For each subset $S \subseteq V$ with $|S| = [n/2]$,

$$e(s) = \big(1+o(1)\big)\frac{n^2}{16}.$$

$\mathbf{P}_6$:

$$\sum_{i,j\in\{1,2,\ldots,n\}} \left|s(i,j) - \frac{n}{2}\right| = o(n^3),$$

where

$$s(i,j) = \big|\big\{x \in \{1,2,\ldots,n\} : a(i,x) = a(j,x)\big\}\big| \quad \text{for } i,j \in \{1,2,\ldots,n\}. \tag{2.1}$$

$\mathbf{P}_7$: $$\sum_{i,j\in\{1,2,\ldots,n\}} \left|\big|\{x : a(i,x) = a(j,x) = 1\}\big| - \frac{n}{4}\right| = o(n^3).$$

They proved the following theorem:

Theorem 2.1. *The following properties are equivalent for a graph $G = G(n)$:*

a) $\mathrm{P}_1(s)$ *with fixed $s \geq 4$;*
b) $\mathrm{P}_2(4)$*;*
c) $\mathrm{P}_2(t)$ *with fixed even $t \geq 4$;*
d) P_3*;*
e) $\mathrm{P}_4(\varepsilon)$ *with fixed $\varepsilon > 0$;*
f) P_5*;*
g) P_6*;*
h) P_7.

Based on this theorem, they define the notion of *quasi-random graph* in the following way:

Definition 2.1. Graphs having any (and therefore, all) of the above properties will be called quasi-random.

They write in [5]: "The same techniques can be used to establish the corresponding results for quasi-random graphs that imitate random graphs generated with a more general edge probability $p = p(n)$ (see also ref. [21])." This is certainly so, however, in their papers they stick to the case $p = 1/2$ so that Theorem 2.1 and Definition 2.1 are restricted to this case. Thus in this paper, we will also stick to this case, and we will study the case of general p only in the sequels.

3. Pseudorandom binary sequences

Pseudorandomness of binary sequences plays a crucial role in cryptography, and it has several different definitions. Here we will use the constructive and quantitative approach developed in [16] and its sequels which is most suitable for our goals.

In [16] Mauduit and the second author introduced the following measures of pseudorandomness:

Definition 3.1. Let

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

be a finite binary sequence. Then the *well-distribution measure* of E_N is defined as

$$W(E_N) = \max_{a,b,t} \left| \sum_{j=0}^{t-1} e_{a+jb} \right|, \tag{3.1}$$

where the maximum is taken over all $a, b, t \in \mathbb{N}$ such that $1 \le a \le a + (t-1)b \le N$, and for $\ell \in \mathbb{N}$, $\ell \ge 2$ the *correlation measure of order* ℓ of E_N is defined as

$$C_\ell(E_N) = \max_{M,D} \left| \sum_{n=1}^{M} e_{n+d_1} e_{n+d_2} \dots e_{n+d_\ell} \right|, \tag{3.2}$$

where the maximum is taken over all $D = (d_1, d_2, \dots, d_\ell)$ and M such that $0 \le d_1 < \dots < d_\ell \le N - M$.

Then the sequence E_N is said to have strong pseudorandom properties (or briefly, to be a "good" pseudorandom sequence) if both of these measures $W(E_N)$ and $C_\ell(E_N)$ (at least for "small" ℓ) are small in terms of N (in particular, both are $o(N)$ as $N \to \infty$). Indeed, later Cassaigne, Mauduit and the second author [4] showed that this terminology is justified since for almost all $E_N \in \{-1, +1\}^N$ both $W(E_N)$ and $C_\ell(E_N)$ are less than $N^{1/2}(\log N)^c$ (and later their results were

sharpened further in [1] and [2]). It was also shown in [16] that the Legendre symbol sequence $\left(\left(\frac{1}{p}\right),\left(\frac{2}{p}\right),\dots,\left(\frac{p-1}{p}\right)\right)$ forms a "good" pseudorandom sequence:

Theorem 3.1. *There is a number q_0 such that if $q > q_0$ is a prime number, $k \in \mathbb{N}$, $k < q$ and we write*

$$E_{q-1} = \left(\left(\frac{1}{q}\right),\left(\frac{2}{q}\right),\dots,\left(\frac{q-1}{q}\right)\right), \tag{3.3}$$

then we have

$$W(E_{q-1}) \le 9q^{1/2}\log q$$

and

$$C_k(E_{q-1}) < 9kq^{1/2}\log q. \tag{3.4}$$

Indeed, this is a special case of Theorem 1 in [16]. (The proof is based on A. Weil's theorem [24].)

Since that numerous results have been proved and constructions presented along these lines; in [11] Gyarmati gave an excellent survey of them (there are 135 references listed in her paper), see also [20]. Here we will present only a few further definitions and constructions that we will need later.

First we present 6 constructions (some of them in a simplified form) for binary sequences with strong pseudorandom properties. The first two of them are, perhaps, the best ones.

Theorem 3.2. *Assume that q is a prime number, $f(x) \in \mathbb{F}_q[x]$ ($\mathbb{F}_q$ being the field of the modulo residue classes) has degree k (> 0), $f(x)$ has no multiple zero in $\overline{\mathbb{F}}_q$ (= the algebraic closure of $\mathbb{F}_q$), and the binary sequence $E_q = (e_1, e_2, \dots, e_q)$ is defined by*

$$e_n = \begin{cases} \left(\frac{f(n)}{q}\right) & \text{for } \big(f(n), q\big) = 1, \\ +1 & \text{for } q \mid f(n), \end{cases} \tag{3.5}$$

where $\left(\frac{\dots}{q}\right)$ is the Legendre symbol. Then we have

$$W(E_q) < 10kq^{1/2}\log q\,.$$

Moreover, assume that also $\ell \in \mathbb{N}$, and one of the following assumptions holds:

(i) $\ell = 2$;

(ii) $\ell < q$, *and 2 is a primitive root modulo q;*

(iii) $(4k)^\ell < q$.

Then we also have

$$C_\ell(E_q) < 10k\ell q^{1/2}\log q\,.$$

This was proved by Goubin, Mauduit and the second author [8]; it is a combination of their Theorem 1 and 2 there. They also presented examples showing that if none of conditions (i), (ii) and (iii) holds, then $C_\ell(E_q)$ can be large.

Theorem 3.3. *Assume that q is an odd prime number, $k,\ell\in\mathbb{N}$, $k<q$, $2\le\ell<q$,*

$$k\ell < \frac{q}{2}, \tag{3.6}$$

and $f(x)\in\mathbb{F}_q[x]$ is of form

$$f(x) = (x+a_1)(x+a_2)\cdots(x+a_k), \tag{3.7}$$

where $a_1,a_2,\dots,a_k$ are pairwise distinct elements of $\mathbb{F}_q$. For $a\in\mathbb{F}_q$, $a\ne 0$ denote the multiplicative inverse of a modulo q by a^{-1} : $aa^{-1}\equiv 1 \pmod q$. Define the binary sequence $E_q=(e_1,e_2,\dots,e_q)$ by

$$e_i = \begin{cases} +1 & \text{if } (f(n),q)=1, \quad r_q\big(f(n)^{-1}\big) < \frac{q}{2}, \\ -1 & \text{if } (f(n),q)=1, \quad r_p\big(f(n)^{-1}\big) > \frac{q}{2} \quad \text{or} \quad q \mid f(n), \end{cases}$$

where for all $a\in\mathbb{Z}$, $r_q(a)$ denotes the unique integer r such that

$$r\in\{0,1,\dots,q-1\} \quad \text{and} \quad r\equiv a \pmod q. \tag{3.8}$$

Then we have

$$W(E_q) \ll kq^{1/2}(\log q)^2$$

and

$$C_\ell(E_q) \ll k\ell q^{1/2}(\log q)^{\ell+1}. \tag{3.9}$$

(Here and later $\ll$ is Vinogradov's notation: $f(n)\ll g(n)$ means that $f(n)=O\big(g(n)\big)$.) This is a special case of the combinations of Theorem 1 and Theorem 3 in the paper [17] of Mauduit and the second author. Note that Liu [13] proved another similar theorem in which (3.7) is replaced by an other assumption on $f(x)$.

The next two constructions also produce large families of binary sequences with strong pseudorandom properties, they can be handled well, and they can be adjusted easily to different situations; we will be able to profit from these properties in Parts II and III of this series. However, these constructions also have certain weak points. In [9] Gyarmati proved the following theorem (generalizing a result of the second author [19]):

Theorem 3.4. *Let q be an odd prime, g a primitive root modulo q. For $(n,q)=1$ define* ind n *(the index or discrete logarithm of n modulo q) by $n\equiv g^{\operatorname{ind} n} \pmod q$ and (to make it uniquely determined) $1\le \operatorname{ind} n\le q-1$. Let $f(x)\in\mathbb{F}[x]$ be*

a polynomial of degree k and define the binary sequence $E_q = (e_1, e_2, \dots, \ell_q)$ by

$$e_n = \begin{cases} +1 & \text{if} \quad 1 \le \text{ind } f(n) \le (q-1)/2, \\ -1 & \text{if} \quad (q+1)/2 \le \text{ind } f(n) \le q-1 \quad \text{or} \quad q \mid f(n) \end{cases} \tag{3.10}$$

for $n = 1, 2, \dots, q$. Then we have

$$W(E_q) < 38kq^{1/2}(\log q)^{1/2}.$$

Moreover, assume that $\ell \in \mathbb{N}$, and one of the following 4 conditions holds:

(i) *$f(x)$ is irreducible over $\mathbb{F}_q$;*

(ii) *if $f(x)$ has the factorization $f(x) = (\varphi_1(x))^{\alpha_1}(\varphi_2(x))^{\alpha_2} \dots \big(\varphi_u(x)\big)^{\alpha_u}$, where $\alpha_i \in \mathbb{N}$ and $\varphi_i(x)$ is irreducible over $\mathbb{F}_q$ for every $i \in \{1, 2, \dots, u\}$, then there exists a $\beta \in \mathbb{N}$ such that exactly one or two $\varphi_i(x)$ have degree β;*

(iii) *$\ell = 2$;*

(iv) *$(4\ell)^k < q$ or $(4k)^\ell < q$.*

Then we also have

$$C_\ell(E_q) < 10k\ell 4^\ell q^{1/2}(\log q)^{\ell+1}. \tag{3.11}$$

We remark that the weak point of this construction is that there is no fast algorithm for computing the discrete logarithm, thus the explicit computation of the elements of E_q is rather slow (see [10] for a faster, although more complicated version of this construction).

In [15] Mauduit, Rivat and the second author of this paper proved the following theorem:

Theorem 3.5. *Let q be a prime, $k \in \mathbb{N}$, $k \ge 2$ and $f(x) \in \mathbb{F}_q[x]$ a polynomial of degree k. Define the binary sequence $E_q = (e_1, e_2, \dots, e_q)$ by*

$$e_n = \begin{cases} +1 & \text{if} \quad 0 \le r_q\big(f(n)\big) < q/2, \\ -1 & \text{if} \quad q/2 < r_q\big(f(n)\big) < q \end{cases} \tag{3.12}$$

for all $n \in \{1, 2, \dots, q\}$ where again $r_q(a)$ is defined by (3.8)*. Then we have*

$$W(E_q) \ll kq^{1/2}(\log q)^2.$$

Moreover, if $\ell \in \mathbb{N}$ and

$$2 \le \ell \le k-1, \tag{3.13}$$

then we also have

$$C_\ell(E_q) \ll kq^{1/2}(\log q)^{\ell+1}.$$

This construction is one of the simplest ones, however, it has the weak point that a condition of type (3.13) is necessary, since the correlation of order k can be "large". More precisely, it is also proved in [15] that

Theorem 3.6. *If $t \in \mathbb{N}$ and $k = 2^t$, then there exists a constant $c = c(k) > 0$ such that if q is a prime number large enough, $f(x) \in \mathbb{F}_q[x]$ is of degree k and $E_q = (e_1, e_2, \dots, e_q) \in \{-1, +1\}^q$ is defined by (3.12), then*

$$C_k(E_q) \gg \max_{1 \le T < T+M \le q} \left| \sum_{n=T}^{T+M} e_n e_{n+1} \dots e_{n+k-1} \right| \gg cq.$$

Observe that all the constructions above are "modular" constructions with prime moduli q, i.e., we work over $\mathbb{F}_q$. There are very few constructions of other type and, indeed, just extending some of these constructions from prime moduli to "RSA type" composite moduli

$$m = pq, \quad \text{where} \quad p, q \quad \text{are primes with} \quad (2 <) < p < q < 2p, \tag{3.14}$$

the situation gets more complicated. In [18] Rivat and the second author studied, among others, the extension of construction (3.5) to the Jacobi symbol and moduli of type (3.14). They proved:

Theorem 3.7. *Assume that $m \in \mathbb{N}$ is of the form (3.14), $f(x) = a_d x^d + \cdots$ $\cdots + a_1 x + a_0 \in \mathbb{Z}[x]$,*

$$(a_d, pq) = 1, \tag{3.15}$$

$$0 < d < p (< q), \tag{3.16}$$

$f(x)$ as a polynomial over $\mathbb{F}_p$ (more exactly, the polynomial of degree d whose coefficients are the residue classes modulo p represented by $a_d, \dots, a_1$, resp. a_0) is not the constant multiple of the square of a polynomial over $\mathbb{F}_p$, and $f(x)$ as a polynomial over $\mathbb{F}_q$ is not the constant multiple of the square of a polynomial over $\mathbb{F}_q$. Define the binary sequence $E_m = (e_1, \dots, e_m)$ by

$$e_n = \begin{cases} \left(\frac{f(n)}{m}\right) & \text{for } \big(f(n), m\big) = 1, \\ +1 & \text{for } \big(f(n), m\big) > 1, \end{cases}$$

where $\left(\frac{\dots}{m}\right)$ denotes the Jacobi symbol.

Then we have

$$\begin{aligned} W(E_m) &\ll d^2 m^{1/2} \log m, \\ C_2(E_m) &\ll d m^{3/4} \end{aligned} \tag{3.17}$$

and

$$C_4(E_m) \ge m - 35 d m^{1/2} \Big(= \big(1 + o(1)\big) m \Big).$$

(See also [14].) In [18] similar extensions of other composite moduli constructions for binary sequences with strong pseudorandom properties are also studied, and similar results have been proved.

So far we have summarized the most important definitions, notations, results and facts on pseudorandomness of binary sequences that we will need in this series. However, for adapting them to study quasi-randomness and pseudo-randomness of graphs, we have to complete this survey by adding a few more definitions and facts.

We will also need the *cyclic* versions of the measures of pseudorandomness defined and applied in [12]:

Definition 3.2. If E_N is the binary sequence

$$E_N = (e_1, e_2, \dots, e_N) \in \{-1, +1\}^N, \tag{3.18}$$

then the *infinite* binary sequence

$$\overset{\circ}{E}_N = (\dots, e_{-2}, e_{-1}, e_0, e_1, e_2, \dots), \tag{3.19}$$

(infinite in both directions) is defined so that for $i \in \mathbb{Z}$ let $r(i)$ be the integer with $r(i) \equiv i \pmod N$, $1 \le r(i) \le N$, and then $e_i = e_{r(i)}$.

(In other words, $\overset{\circ}{E}_N$ is the periodic extension of E_N with period length N.)

Definition 3.3. The cyclic well-distribution measure of the sequence E_N in (3.18) is defined by

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

where the maximum is taken over all $a \in \mathbb{Z}$ and $b, t \in \mathbb{N}$ such that $(0 \le)$ $(t-1)b < N$ (and the terms e_{a+jb} are defined as in (3.19)).

Definition 3.4. The cyclic correlation measure of order ℓ of the sequence E_N in (3.18) is defined by

$$\overset{\circ}{C}_\ell(E_N) = \max_{M,D} \left| \sum_{n=1}^{M} e_{n+d_1} e_{n+d_2} \dots e_{n+d_\ell} \right|,$$

where the maximum is taken over all $D = (d_1, d_2, \dots, d_\ell)$ and M such that the d_i's are integers with $0 \le d_1 < d_2 < \dots < d_\ell < N$ and $M \in \mathbb{N}$, $M \le N$ (and the terms e_{n+d_i} are defined as in (3.19)).

Definition 3.5. If E_N is the sequence in (3.18) then we will write

$$\overset{\circ}{S}_1(E_N) = \left|\sum_{n=1}^{N} e_n\right|$$

and, for $\ell \in \mathbb{N}$, $\ell \geq 2$,

$$\overset{\circ}{S}_\ell(E_N) = \max_D \left|\sum_{n=1}^{N} e_{n+d_1} e_{n+d_2} \dots e_{n+d_\ell}\right|,$$

where the maximum is taken over all $D = (d_1, d_2, \dots, d_\ell)$ such that the d_i's are integers with $0 \leq d_1 < d_2 < \dots < d_\ell < N$ (and the terms e_{n+d_i} are defined as in (3.19)). Then, clearly, we have

$$\overset{\circ}{S}_1(E_N) \leq W(E_N)$$

and

$$\overset{\circ}{S}_\ell(E_N) \leq \overset{\circ}{C}_\ell(E_N) \qquad \text{(for every } \ell \geq 2\text{)}. \tag{3.20}$$

Moreover, by Proposition 1 in [12] we also have

$$W(E_N) \leq \overset{\circ}{W}(E_N) \leq 2W(E_N) \tag{3.21}$$

and

$$C_\ell(E_N) \leq \overset{\circ}{C}_\ell(E_N) \leq (\ell+1)C_\ell(E_N) \qquad \text{(for every } \ell \geq 2\text{)}. \tag{3.22}$$

So far we have studied binary sequences of the type (3.18) all whose elements are -1 or $+1$, although later we will work with bit sequences, i.e., with sequences belonging to $\{0,1\}^N$ (in cryptography the situation is similar). The reason of this is that for sequences of type (3.18) the formulas tend to be simpler, since the main term is frequently 0 and, on the other hand, it is easy to switch from sequences of type (3.18) to bit sequences and vice versa by using the simple bijection

$$\varphi : \{-1,+1\} \longleftrightarrow \{0,1\} \quad \text{defined so that for } e \in \{-1,+1\}$$

we have

$$\varphi(e) = \frac{1+e}{2} \qquad \text{(for } e \in \{-1,+1\}\text{)}, \tag{3.23}$$

and then we transform the binary sequence E_N in (3.18) into a bit sequence by the bijection $\Phi : \{-1,+1\}^N \longleftrightarrow \{0,1\}^N$ defined as

$$\Phi(E_N) = \Phi\big((e_1, e_2, \dots, e_n)\big) = \big(\varphi(e_1), \varphi(e_2), \dots, \varphi(e_N)\big)$$

$$\text{(for the sequence } E_N \text{ in (3.18))}, \tag{3.24}$$

we will keep this notation throughout this paper.

4. The general construction principle

We will need some more definitions.

Definition 4.1. A *circulant matrix* is a square matrix whose each row vector can be obtained from the preceding vector by rotating it one element to the right, i.e., a matrix Z of the form

$$Z = \begin{bmatrix} z_0 & z_1 & z_2 & \dots & z_{n-2} & z_{n-1} \\ z_{n-1} & z_0 & z_1 & \dots & z_{n-3} & z_{n-2} \\ z_{n-2} & z_{n-1} & z_0 & \dots & z_{n-4} & z_{n-3} \\ & & & \dots & & \\ z_1 & z_2 & z_3 & \dots & z_{n-1} & z_0 \end{bmatrix}. \tag{4.1}$$

(See [7].) Observe that such a circulant matrix is uniquely determined by its first row, i.e., by the sequence

$$Z_n = (z_0, z_1, \dots, z_{n-1}). \tag{4.2}$$

Thus we may introduce the following terminology:

Definition 4.2. The *circulant matrix* Z in (4.1) is *generated by the sequence* Z_n in (4.2), and the matrix Z generated by the sequence Z_n is denoted by $Z = Z(Z_n)$.

Definition 4.3. A *circulant* (or cyclic) *graph* is a graph whose adjacency matrix is a circulant matrix.

(See [23].)

Our main construction principle is to start out from constructions for large families of binary sequences with certain pseudorandom properties (like the constructions in Theorems 3.1–3.5 and 3.7 above), and to show that these families contain large *subfamilies* consisting of *binary sequences* which (after making, perhaps, minor adjustments) *generate circulant graphs which are quasi-random*. This principle will be formulated in a more precise form in Theorem 4.1 and Corollaries 4.1 and 4.2.

Definition 4.4. A binary sequence $F_n = \{f_0, f_1, \dots, f_{n-1}\} \in \{-1, +1\}^n$ is said to be *symmetric* if for the extended binary sequence

$$\overset{\circ}{F}_n = (\dots, f_{-2}, f_{-1}, f_0, f_1, f_2, \dots) \tag{4.3}$$

(defined as in Definition 3.2) we have

$$f_i = f_{-i} \quad \text{for all} \quad i \in \mathbb{N},$$

i.e.,

$$f_i = f_{n-i} \quad \text{for all} \quad i \in \{1, 2, \dots, n-1\}.$$

Theorem 4.1. *Assume that n is integer with $n \to \infty$,*

$$F_n = (f_0, f_1, \dots, f_{n-1}) \in \{-1, +1\}^n \tag{4.4}$$

is such that:

(i) $f_0 = -1$,
(ii) F_n *is symmetric,*

and define the sequence $F'_n = (f'_0, f'_1, \dots, f'_{n-1}) \in \{0, 1\}^n$ *by*

$$F'_n = \Phi(F_n), \tag{4.5}$$

where $\Phi : \{-1, +1\}^n \to \{0, 1\}^n$ *is the transformation defined by* (3.23) *and* (3.24). *Then by* (4.4), (i) *and* (ii) *the circulant matrix* $Z(F'_n)$ *generated by the sequence* F'_n *defined in* (4.5) *is such that every element of it is* 0 *or* 1, *the elements in its main diagonal are* 0, *and it is symmetric. Thus it is the adjacency matrix of a (uniquely determined) circulant graph* $G_n(F'_n)$. *Denote the* ℓ*th element in the* k*th row of the matrix* $Z(F'_n)$ *by* $a(k, \ell)$, *and define the function* $s(i, j)$ *by* (2.1) *and* $\overset{\circ}{S}_2(F_n)$ *by Definition 3.5. Then we have*

$$\sum_{\substack{i,j \in \{1,2,\dots,n\} \\ i \neq j}} \left| s(i,j) - \frac{n}{2} \right| \leq \frac{n(n-1)}{4} \overset{\circ}{S}_2(F_n). \tag{4.6}$$

Corollary 4.1. *If we have the same notation and assumptions as in Theorem 4.1, then we also have*

$$\sum_{\substack{i,j \in \{1,2,\dots,n\} \\ i \neq j}} \left| S(i,j) - \frac{n}{2} \right| \leq \frac{3}{4} n(n-1) C_2(F_n). \tag{4.7}$$

Corollary 4.2. *Consider a family* $\mathcal{F}$ *of binary sequences* F_n *with a sequence of positive integers* n *tending to infinity and such that they are of the form* (4.1), *satisfy* (i) *and* (ii) *in Theorem 4.1, and assume that either*

$$\overset{\circ}{S}_2(F_n) = o(n) \tag{4.8}$$

or

$$C_2(F_n) = o(n) \tag{4.9}$$

also holds. Then the circulant graphs $G_n(F'_n)$ *are quasi-random.*

P r o o f o f T h e o r e m 4.1. It follows from the definition of the matrix $Z(F'_n)$ that its k-th row is

$$\begin{aligned}\big(a(k,1),a(k,2),\dots,a(k,\ell),\dots,a(k,n)\big) &= \\ \big(f'_{n-k+1},f'_{n-k+2},\dots,f'_{n-k+\ell},\dots,f'_{n-k+n}\big) &= \\ \left(\frac{1+f_{n-k+1}}{2},\frac{1+f_{n-k+2}}{2},\dots,\frac{1+f_{n-k+\ell}}{2},\dots,\frac{1+f_{n-k+n}}{2}\right),&\end{aligned}$$

where the numbers f_x are defined in the periodic extension sense (with period length n) of F_n (as described in Definition 3.2). It follows that

$$a(k,\ell)=\frac{1+f_{n-k+\ell}}{2}. \tag{4.10}$$

Clearly, for all

$$i,j\in\{1,2,\dots,n\},\quad i\ne j, \tag{4.11}$$

we have

$$\big(a(i,x)-a(j,x)\big)^2=\begin{cases}0 & \text{if } a(i,x)=a(j,x),\\ 1 & \text{if } a(i,x)\ne a(j,x),\end{cases}$$

hence

$$1-\big(a(i,x)-a(j,x)\big)^2=\begin{cases}1 & \text{if } a(i,x)=a(j,x),\\ 0 & \text{if } a(i,x)\ne a(j,x).\end{cases} \tag{4.12}$$

It follows from (4.10) and (4.12) that

$$\begin{aligned}s(i,j)&=\big|\big\{x\in\{1,2,\dots,n\}:\ a(i,x)=a(j,x)\big\}\big|\\ &=\sum_{x=1}^{n}\Big(1-\big(a(i,x)-a(j,x)\big)^2\Big)=n-\sum_{x=1}^{n}\big(a(i,x)-a(j,x)\big)^2\\ &=n-\sum_{x=1}^{n}\left(\frac{(1+f_{n-i+x})-(1+f_{n-j+x})}{2}\right)^2\\ &=n-\frac14\sum_{x=1}^{n}(f_{n-i+x}-f_{n-j+x})^2\\ &=n-\frac14\sum_{x=1}^{n}(1-2f_{n-i+x}f_{n-j+x}+1)=\frac{n}{2}+\frac12\sum_{x=1}^{n}f_{n-i+x}f_{n-j+x},\end{aligned}$$

whence, by Definition 3.5,

$$\left|s(i,j)-\frac{n}{2}\right|=\frac12\left|\sum_{x=1}^{n}f_{n-i+x}f_{n-j+x}\right|\le\frac12\overset{\circ}{S}_2(F_n).$$

This holds for every i, j satisfying (4.11), thus we have

$$\sum_{\substack{i,j\in\{1,2,\dots,n\}\\ i\neq j}} \left|s(i,j)-\frac{n}{2}\right| \le \sum_{\substack{i,j\in\{1,2,\dots,n\}\\ i\neq j}} \frac{1}{2}\overset{\circ}{S}_2(F_n) = \frac{n(n-1)}{4}\overset{\circ}{S}_2(F_n). \qquad \square$$

Proof of Corollary 4.1. By the $\ell = 2$ special cases of (3.20) and (3.22) we have

$$\overset{\circ}{S}_2(F_n) \le \overset{\circ}{C}_2(F_n)$$

and

$$\overset{\circ}{C}_2(F_n) \le 3C_2(F_n).$$

Combining these inequalities with (4.6) in Theorem 4.1 we get the result. $\square$

(We remark that we have used $\overset{\circ}{S}_\ell$ and $\overset{\circ}{C}_\ell$ only for $\ell = 2$ so far; $\overset{\circ}{S}_\ell$ with $\ell=1$ and $\ell > 2$, and also $\overset{\circ}{W}$ will be used only in the sequels of this paper.)

Proof of Corollary 4.2. By Theorem A of Chung, Graham and Wilson it suffices to show that the graph $G(F'_n)$ possesses property P_6, i.e.,

$$\sum_{i,j\in\{1,2,\dots,n\}} \left|s(i,j)-\frac{n}{2}\right| = o(n^3). \tag{4.13}$$

If either (4.8) or (4.9) holds, then this follows trivially from Theorem 4.1 and Corollary 4.1 using also that, clearly,

$$\left|s(i,i)-\frac{n}{2}\right| = \frac{n}{2} \quad \text{for all} \quad i \in \{1,2,\dots,n\}. \qquad \square$$

To extend the applicability of Theorem 4.1 we will need two more simple lemmas.

Lemma 4.1. *If $N \in \mathbb{N}$, $N \ge 3$, $E_{N-1} = (e_1, e_2, \dots, e_{N-1}) \in \{-1,+1\}^{N-1}$ and $F_N \in \{-1,+1\}^N$ is of the form $F_N = (f_1, f_2, \dots, f_N) = (f_1, e_1, e_2, \dots e_{N-1})$, then we have*

$$C_2(F_N) \le 3C_2(E_{N-1}). \tag{4.14}$$

Proof of Lemma 4.1. By Definition 3.1 we have

$$C_2(E_{N-1}) = \max_{M,D}\left|\sum_{n=1}^{M} e_{n+d_1}e_{n+d_2}\right|, \tag{4.15}$$

where M, D are such that $D = (d_1, d_2)$ with $0 \le d_1 < d_2 \le N-1-M$ and

$$C_2(F_N) = \max\left|\sum_{n=1}^{M} f_{n+d_1} f_{n+d_2}\right|, \tag{4.16}$$

where M', D' are such that $D = (d_1, d_2)$ with $0 \le d_1 < d_2 \le N' - 1 - M'$. Observe that the sums in (4.16) are the same as the sums in (4.15) except at most a single term of absolute value 1 appearing in (4.16) which includes the factor f_1, thus the absolute values of the sums in (4.16) are greater than the absolute values of the sums in (4.15) by at most 2. It follows that

$$C_2(F_N) \le C_2(E_{N-1}) + 2 \le 3C_2(E_{N-1}). \qquad \square$$

Next we will show how to get around condition (i) in Theorem 4.1. Assume that F_n is a binary sequence of form (4.4) which satisfies (ii) in Theorem 4.1 (but (i) is not assumed now). Define

$$\overline{F}_n = \left(\overline{f}_0, \overline{f}_1, \dots, \overline{f}_{n-1}\right) \in \{-1, +1\}^n \tag{4.17}$$

by

$$\overline{F}_n = (-f_0)(f_0, f_1, \dots, f_{n-1}) = (-1, -f_0 f_1, \dots, -f_0 f_{n-1}), \tag{4.18}$$

and let

$$\overline{F}'_n = \Phi(\overline{F}_n). \tag{4.19}$$

Then all the assumptions in Theorem 4.1 (including (i)) hold with $\overline{F}_n$ in place of F_n. It follows easily from (3.2) in Definition 3.1 that we also have

$$\overset{\circ}{S}_2(\overline{F}_n) = \overset{\circ}{S}_2(F_n) \quad \text{and} \quad C_2(\overline{F}_n) = C_2(F_n).$$

Thus we obtain

Lemma 4.2. *If F_n satisfies* (4.4) *and* (ii) *in Theorem 4.1, then the uniquely determined circulant graph $G_n(\overline{F}'_n)$ satisfies each of Theorem 4.1, Corollary 4.1 and Corollary 4.2 with $\overline{F}_n$ in place of F_n.*

5. Examples for quasi-random graphs constructed using Theorems 3.1, 3.2, 3.3, 3.4, 3.5 and 3.7

In this section we will specify Theorem 4.1 to circulant graphs constructed by using (slightly modified) subfamilies of the families of the binary sequences studied in Theorems 3.1–3.5 and 3.7. Theorems 5.2 and 5.3 will be the best constructions, while the other costructions are presented here since we can use them to illustrate certain facts here and in the sequels.

Theorem 5.1. *Let q be a prime number of the form*

$$q = 4k + 1, \tag{5.1}$$

and define $F_q \in \{-1, +1\}^q$ by

$$F_q = (f_0, f_1, f_2, \ldots, f_{q-1}) = \left(-1, \left(\frac{1}{q}\right), \left(\frac{2}{q}\right), \ldots, \left(\frac{q-1}{q}\right)\right) \tag{5.2}$$

(where again $\left(\frac{\cdots}{q}\right)$ is the Legendre symbol). Then defining F'_q by

$$F'_q = \Phi(F_q), \tag{5.3}$$

the circulant graphs $G_q = G_q(F'_q)$ are quasi-random (for $q \to \infty$).

P r o o f. The sequence F_q in (5.2) satisfies (4.4) and (i) in Theorem 4.1 trivially, and (ii) (symmetry) in the theorem also holds since by (5.1) we have

$$f_i = \left(\frac{i}{q}\right) = \left(\frac{-i}{q}\right) = \left(\frac{q-i}{q}\right) = f_{q-i} \quad \text{for all} \quad i \in \{1, 2, \ldots, q-1\}.$$

Moreover, using Lemma 4.1 with $q, -1$ and $\left(\left(\frac{1}{q}\right), \left(\frac{2}{q}\right), \ldots, \left(\frac{q-1}{q}\right)\right)$ in place of N, f_1 and E_{N-1}, respectively, we obtain that

$$C_2(F_q) \le 3C_2(E_{q-1})$$

whence, by (3.4) in Theorem 3.1,

$$C_2(F_q) = O(q^{1/2} \log q) = o(q).$$

Thus by Corollary 4.2 the graphs $G_q(F'_q)$ are quasi-random. □

Remarks. It is easy to see that this graph $G_q(F'_q)$ is the Payley graph which can be defined by the adjacency matrix $[a(i,j)]_{i,j \in \{0,1,\ldots,q-1\}}$ with $a(i,j) = \left(\frac{j-i}{q}\right)$ for $i \ne j$ and $a(i,i) = 0$, and which is known to be quasi-random [6, p. 359], se also [3, Chapter XIII]. Moreover, we remark that our proof above gives the upper bound $q^{5/2} \log q$ for the sum

$$\sum_{i,j \in \{1,2,\ldots,q\}} \left| s(i,j) - \frac{q}{2} \right|.$$

The factor $\log q$ in this upper bound could be eliminated but here this slightly weaker upper bound is enough for our purpose, thus we will discuss this matter only in Part II of this series where the estimate of the deviation between certain graph related quantities and their expected value will play a more important role. (In most of the following examples the situation will be similar.)

Theorem 5.2. *Assume that q is a prime with $q \to \infty$, $t \in \mathbb{N}$ with*

$$t = o\left(\frac{q^{1/2}}{\log q}\right), \tag{5.4}$$

and let $a_1, a_2, \dots, a_t \in \mathbb{F}_q$ be such that

$$a_i \neq 0 \quad \text{for} \qquad i = 1, 2, \dots, t \tag{5.5}$$

and

$$a_i^2 \neq a_j^2 \quad \text{for} \quad i, j = 1, 2, \dots, t,\ i \neq j. \tag{5.6}$$

Define $f(x) \in \mathbb{F}_q[x]$ by

$$f(x) = \prod_{i=1}^{t} \left(x^2 - a_i^2\right) = \prod_{i=1}^{t} (x - a_i)(x + a_i) \tag{5.7}$$

and $F_q = (f_0, f_1, \dots, f_{q-1})$ by

$$f_i = \begin{cases} \left(\frac{f(i)}{q}\right) & \text{for } \big(f(i), q\big) = 1, \\ +1 & \text{for } q \mid f(i) \end{cases} \tag{5.8}$$

(for $i = 0, 1, \dots, q-1$). Define $\overline{F}_q'$ as in (4.17), (4.18) *and* (4.19) *(with q in place of n) :*

$$\overline{F}_q' = \Phi(\overline{F}_q).$$

Then the circulant graphs $G_q = G_q(\overline{F}_q')$ are quasi-random.

P r o o f. It is trivial that F_q is of form (4.4) (with q in place of n). Moreover, the function $f(x)$ is even, thus we have

$$f(a) = f(-a) = f(q - a)$$

for every $a \in \mathbb{F}_q$, thus it follows from (5.8) that

$$f_i = f_{q-i} \quad \text{for} \quad i = 1, 2, \dots, q - 1$$

so that F_q also satisfies (ii) in Theorem 4.1 (with q in place of n). Thus by Lemma 4.2, we may apply Corollary 4.2 (with q in place of n) if we have

$$C_2(F_q) = o(q). \tag{5.9}$$

To see that this holds, observe that by (5.5), (5.6) and (5.7) the zeros of the polynomial $f(x)$ are pairwise distinct so that it has no multiple zero. Thus we may apply (i) in Theorem 3.2 (with $k = 2t$, $\ell = 2$) to estimate $C_2(F_q)$, and then we obtain that

$$C_2(F_q) < 10 \cdot 2t \cdot 2q^{1/2} \log q = 40tq^{1/2} \log q. \tag{5.10}$$

The equation (5.9) follows from (5.4) and (5.10) thus, indeed, by Corollary 4.2 the circulant graphs $G_q = G_q(\overline{F}_q')$ are quasi-random. □

Theorem 5.3. *Assume that q is prime with $q \to \infty$, let $t \in \mathbb{N}$ with*

$$t = o\left(\frac{q^{1/2}}{(\log q)^3}\right), \tag{5.11}$$

and define $a_1, a_2, \dots, a_t$ and $f(x)$ as in Theorem 5.2. Define the binary sequence $F_q = (f_0, f_1, \dots, f_{q-1})$ by

$$f_i = \begin{cases} +1 & \text{if } \big(f(i), q\big) = 1,\ r_q\big(f(i)^{-1}\big) < \frac{q}{2}, \\ -1 & \text{if } \big(f(i), q\big) = 1,\ r_q\big(f(i)^{-1}\big) > \frac{q}{2} \text{ or } q \mid f(i) \end{cases}$$

(for $i = 0, 1, \dots, q-1$) where a^{-1} and $r_q(a)$ are defined as in Theorem 3.3. Define $\overline{F}_q'$ as in (4.17), (4.18) *and* (4.19) *(with q in place of n): $\overline{F}_q' = \Phi(\overline{F}_q)$. Then the circulant graphs $G_q = G_q(\overline{F}_q')$ are quasi-random.*

P r o o f. As in the proof of Theorem 5.2, it suffices to show that (5.9) holds. By our assumptions on $f(x)$, it is of form (5.5) with

$$k = \deg\, f(x) = 2t,$$

thus it follows from (5.11) that t and $\ell = 2$ satisfy (3.6) for large q, so that we may apply Theorem 3.3 with $\ell = 2$ for estimating $C_2(F_q)$. Then by (5.11) we obtain from (3.9) that

$$C_2(F_q) \ll k\ell q^{1/2}(\log q)^{\ell+1} = 4tq^{1/2}(\log q)^3 = o(q).$$

Thus all the assumptions in Corollary 4.2 hold (with (4.9) and q in place of n) so that the statement of the theorem follows from this corollary. □

Theorem 5.4. *Assume that q is prime with $q \to \infty$, $t \in \mathbb{N}$,*

$$t = o\left(\frac{q^{1/2}}{(\log q)^3}\right), \tag{5.12}$$

$g(x) \in \mathbb{F}_q[x]$ is an irreducible polynomial of degree t and

$$f(x) = g(x)g(-x). \tag{5.13}$$

Define the binary sequence $F_q = (f_0, f_1, \dots, f_{q-1})$ by

$$f_i = \begin{cases} +1 & \text{if} \quad 1 \le \text{ind}\, f(i) \le (q-1)/2, \\ -1 & \text{if} \quad (q+1)/2 \le \text{ind}\, f(i) \le q-1 \text{ or } q \mid f(i) \end{cases}$$

(for $i = 1, 2, \dots, q$; here ind *a is defined as in Theorem 3.4) and define $\overline{F}_q$ and $\overline{F}_q'$ by* (4.17), (4.18) *and* (4.19) *(with q in place of n) so that $\overline{F}_q' = \Phi(\overline{F}_q)$. Then the circulant graphs $G_q = G_q(\overline{F}_q')$ are quasi-random.*

P r o o f. Again we may proceed as in the proofs of Theorems 5.2 and 5.3 (note that clearly the function $f(x)$ is even by (5.12)), and we may reduce the proof to showing that (5.9) holds.

It follows from (5.13) that $f(x)$ is of the form described in (ii) in Theorem 3.4 (with $\varphi_1(x) = g(x)$, $\varphi_2(x) = g(-x)$, $\alpha_1 = \alpha_2 = 1$, $\beta = t$), and (iii) also holds for $\ell = 2$. Thus we may apply Theorem 3.4 to estimate $C_2(F_q)$, and then by

$$k = \deg f = 2\deg g = 2t$$

and (5.12) we get from (3.11) that

$$C_2(F_q) < 10k\ell 4^\ell q^{1/2}(\log q)^{\ell+1} = 640tq^{1/2}(\log q)^3 = o(q).$$

Thus again we may apply Corollary 4.2 with $n = q$ and $C_2(F_q)$ and we get the result. □

Theorem 5.5. *Assume that q is prime with $q \to \infty$, let $t \in \mathbb{N}$,*

$$t > 1, \tag{5.14}$$

$$t = o\left(\frac{q^{1/2}}{(\log q)^3}\right), \tag{5.15}$$

let $g(x) \in \mathbb{F}_q[x]$ be any polynomial with $\deg g(x) = t$, *and define $f(x)$ by* (5.13). *Define the binary sequence $F_q = (f_0, f_1, \dots, f_{q-1})$ by*

$$f_i = \begin{cases} +1 & \text{if} \quad 0 \le r_q\big(f(i)\big) < q/2, \\ -1 & \text{if} \quad q/2 < r_q\big(f(i)\big) < q \end{cases}$$

(for $i = 0, 1, \dots, q-1$) where $r_q(a)$ is defined by (3.8). *Define $\overline{F}_q$ and $\overline{F}_q'$ by* (4.17), (4.18) *and* (4.19) *(with q in place of n) so that $\overline{F}_q' = \Phi(\overline{F}_q)$. Then the circulant graphs $G_q = G_q(\overline{F}_q')$ are quasi-random.*

P r o o f. Again (4.4) and (i) in Theorem 4.1 (with $\overline{F}_q$ in place of F_n) hold trivially, while (ii) (symmetry) in the theorem follows from (5.13). Thus we may reduce the proof to showing that (5.9) holds. Writing $k = \deg f = 2t$, by (5.14) we have $k = 2t \ge 3$, so that taking $\ell = 2$ inequality (3.13) holds. Thus with this k and ℓ Theorem 3.5 can be applied to estimate $C_\ell(F_q) = C_2(F_q)$, and then we obtain

$$C_2(F_q) \ll kq^{1/2}(\log q)^3 = 2tq^{1/2}(\log q)^3. \tag{5.16}$$

The equation (5.9) follows from (5.15) and (5.16) so that (by Lemma 4.2) Corollary 4.2 can be applied to complete the proof of the theorem. □

Theorem 5.6. *Assume that p, q are primes with*

$$p \to \infty, \quad q \to \infty, \tag{5.17}$$

$$p < q < 2p, \tag{5.18}$$

and let

$$m = pq. \tag{5.19}$$

Let $t \in \mathbb{N}$ with

$$t = o(m^{1/4}) \tag{5.20}$$

and $b_1, b_2, \dots, b_t$ be integers with

$$0 < b_1 < b_2 < \cdots < b_t < p/2, \tag{5.21}$$

write $d = 2t$, and let

$$f(x) = \prod_{i=1}^{t} \left(x^2 - b_i^2\right) = \prod_{i=1}^{t}(x - b_i) \prod_{i=1}^{t}(x + b_i) = a_d x^d + \cdots + a_1 x + a_0 \tag{5.22}$$

with

$$a_d = 1. \tag{5.23}$$

Define the binary sequence $F_m = (f_0, f_1, \dots, f_{m-1})$ by

$$f_i = \begin{cases} \left(\frac{f(i)}{m}\right) & \text{for } \big(f(i), m\big) = 1, \\ +1 & \text{for } \big(f(i), m\big) > 1 \end{cases}$$

(for $i = 0, 1, \dots, m-1$) where $\left(\frac{\cdots}{m}\right)$ denotes the Jacobi symbol, and define $\overline{F}_m$ and $\overline{F}'_m$ as in (4.17), (4.18) *and* (4.19) *(with m in place of n) so that $\overline{F}'_m = \Phi(\overline{F}_m)$. Then the circulant graphs $G_m = G_m(\overline{F}'_m)$ are quasi-random.*

P r o o f. Again we argue as in the proof of Theorems 5.2–5.5. (4.4) and (i) in Theorem 4.1 hold with $\overline{F}_m$ in place of F_n, while (ii) (the symmetry of F_m and $\overline{F}_m$) follows from the fact that the polynomial $f(x)$ in (5.22) is even. Thus again it suffices to show that (5.9) holds with m in place of q:

$$C_2(F_m) = o(m). \tag{5.24}$$

We will do this by using Theorem 3.7. This can be done since the polynomial $f(x)$ in (5.22) satisfies (3.15) and (3.16) by (5.18), (5.19) and (5.23), and clearly $f(x)$ as a polynomial over $\mathbb{F}_p$ and $\mathbb{F}_q$ is not the constant multiple of the square of a polynomial over $\mathbb{F}_p$ and $\mathbb{F}_q$, resp., since it has $d = \deg f$ distinct zeros in both fields. Then we obtain from (3.17) in Theorem 3.7 that

$$C_2(F_m) \ll dm^{3/4} = 2tm^{3/4}. \tag{5.25}$$

(5.24) follows from (5.20) and (5.25), thus by Lemma 4.2, Corollary 4.2 can be applied and this completes the proof of the theorem. □

6. Conclusions

In this paper we have described a principle to generate *quasi-random graphs* from *binary sequences* having certain pseudorandom properties. There are many large families of binary sequences known to possess many of these properties, and we illustrated by several examples that in many cases one can find a subfamily of the given family such that the sequences belonging to this subfamily (perhaps, after some trivial adjustments) possess *all* the pseudorandom properties needed. The most important step in this procedure is to find many sequences in the given family which possess a certain *symmetry property.*

Here we have constructed quasi-random graphs, i.e., we have showed that one of certain seven quantities related to the given graphs is asymptotically equal to its expected value and (following [5], [6]) we restricted ourselves to graphs whose edge density p is near $1/2$. One might like to extend this work to the case of any $0 < p < 1$ and also to estimate the deviation between the actual value of the graph quantity studied and its expectation. Indeed, we will carry out this extension in Part II of this paper by using the notion of "(p, α)-jumbled" graphs which was introduced by Thomason [21], [22].

Finally, we remark that the proofs of Theorems 5.1–5.6 were based on Theorems 3.1–3.5 and 3.7 but from them we used only the estimates given for the correlation of order 2 of the binary sequences studied, although Theorems 3.1–3.5 also contain good upper bounds for correlations of higher order. Is it not possible to also utilize these higher order correlation estimates, do they not imply that the quasi-random graphs generated by them possess more and/or stronger "random type" properties? In case of Theorem 3.7 already the correlation of order 4 is large, while in case of Theorem 3.6 only the correlation of order k (= the degree of the polynomial in question) can be large. Does this lead to any difference in the "random type" properties of the graphs generated by the polynomials in question? Moreover, in case of quasi-randomness the focus is on finding *many* (seven) *equivalent* ("random type") properties. One also might like to say something in the opposite direction, namely that there are graphs possessing *many independent* "random type" properties, where the word "many" means that if the number of the vertices of the graph tends to infinity then also the number of these properties tends to infinity (relatively fast). All these problems will be studied in Part III of this paper.

REFERENCES

[1] ALON, N.—KOHAYAKAWA, Y.— MAUDUIT, C.—MOREIRA, C. G.—RÖDL, V.: *Measures of pseudorandomness for finite sequences: minimal values*, Comb. Probab. Comput. **15** (2006), 1–29.

[2] ——— *Measures of pseudorandomness for finite sequences: typical values*, Proc. London Math. Soc. **95** (2007), 778–812.

[3] BOLLOBÁS, B.: *Random Graphs*. Academic Press, London, 1985.

[4] CASSAIGNE, J.—MAUDUIT, C.—SÁRKÖZY, A.: *On finite pseudorandom binary sequences VII: The measures of pseudorandomness*, Acta Arith. **103** (2002), 97–118.

[5] CHUNG, F. R. K.—GRAHAM, R. L.—WILSON, R. M.: *Quasirandom graphs*, Proc. Nat. Acad. Sci. U.S.A., **85** (1988), 969–970.

[6] ——— *Quasirandom graphs*, Combinatorica **9** (1989), 345–362.

[7] DAVIS, P. J.: *Circulant Matrices*. Wiley, New York, 1970.

[8] GOUBIN, L.—MAUDUIT, C.—SÁRKÖZY, A.: *Construction of large families of pseudorandom binary sequences*, J. Number Theory **106** (2004), 56–69.

[9] GYARMATI, K.: *On a family of pseudorandom binary sequences*, Period. Math. Hungar. **49** (2004), 45–63.

[10] ——— *On a fast version of a pseudorandom generator*. In: General Theory of Information Transfer and Combinatorics. (R. Ahlswede et al.,eds.) In: *Lecture Notes Comput. Sci. Vol. 4123*, Springer, Berlin, 2006, pp. 326–342.

[11] ——— *Measures of pseudorandomness*. In: Finite Fields and Their Applications, Character Sums and Polynomials (P. Charpin et al., eds.) In: *Radon Series on Comput. Appl. Math. Vol. 11*, De Gruyter, Berlin, 2013, pp. 43–64.

[12] GYARMATI, K.—MAUDUIT, C.—SÁRKÖZY, A.: , *Generation of further pseudorandom binary sequences, I (Blowing up a single sequence)*, Unif. Distrib. Theory **10** (2015), 35–61.

[13] LIU, H.: , *Large families of pseudorandom binary sequences and lattices by using the multiplicative inverse*, Acta Arith. **159** (2013), 123–131.

[14] LIU, H. N.—ZHAN, T.—WANG, X. Y.: *On the correlation of pseudorandom binary sequences with composite moduli*, Publ. Math. Debrecen **74** (2009), 195–214.

[15] MAUDUIT, C.—RIVAT, J.—SÁRKÖZY, A.: *Construction of pseudorandom binary sequences using additive characters*, Monatshefte Math. **141** (2004), 197–208.

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

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

[18] RIVAT, J.—SÁRKÖZY, A.: *Modular constructions of pseudorandom binary sequences with composite moduli*, Periodica Math. Hungar. **51** (2005), 75–107.

[19] SÁRKÖZY, A.: *A finite pseudorandom binary sequence*, Studia Sci. Math. Hungar. **38** (2001), 17–35.

[20] ——— *On finite pseudorandom binary sequences and their applications in cryptography*, Tatra Mt. Math. Publ. **37** (2007), 123–136.

[21] THOMASON, A.: *Pseudorandom graphs*, In: *Random Graphs'85, Poznań 1985* (M. Karonski, ed.), In: *North-Holland Math. Stud. Vol. 144.* In: *Ann. Discrete Math. Vol. 33*, North-Holland, Amsterdam, 1987. pp. 307–33.

[22] ———— *Random graphs, strongly regular graphs and pseudo-random graphs*, In: *Surveys in Combinatorics 1987* (C. Whitehead, ed.). In: *London Mat. Soc. Lecture Notes Series Vol. 123*, Cambridge Univ. Press, Cambridge, 1987, pp. 173–196.

[23] VILFRED, V.: *On circulant graphs*, In: *Graph Theory and its Applications* (R. Balakrishnan et al., eds.), (Anna University, Chennai, March 14–16, 2001) Alpha Science, pp. 34–36.

[24] WEIL, A.: *Sur les curbes algébriques et les variétés qui s'en déduisent*, Acta Sci. Ind. **1041**, Hermann, Paris, 1948.

Received July 8, 2019
Accepted October 25, 2019

József Borbély
University of Óbuda
AMK
H-8000 Székesfehérvár
Budai út 45.
HUNGARY
E-mail: borbely.jozsef@amk.uni-obuda.hu

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