A note to the paper "On a fast version of a pseudorandom generator"

Katalin Gyarmati*

Abstract

A large family of pseudorandom sequences is constructed in [7] by using the modulo m residues of the discrete logarithm. Here I prove sharper bounds for the correlation measures than in [7] using only one further assumption on the polynomial used in the construction, namely we will assume that it has no multiple roots. The proofs are similar as in [7], only minor changes are made to obtain this sharper result.

2000 AMS Mathematics Subject Classification: 11K45.

List of keywords and phrases: pseudorandom, index, discrete logarithm, correlation.

*Research partially supported by Hungarian Scientific Research Grants OTKA T043623, T043631 and T049693.

1 Introduction

C. Mauduit and A. Sárközy [5, pp. 367-370] introduced the following measures of pseudorandomness of binary sequences:

For a finite binary sequence $E_N = \{e_1, e_2, \ldots, e_N\} \in \{-1, +1\}^N$ write

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

and, for $D = (d_1, \ldots, d_k)$ with non-negative integers $d_1 < \cdots < d_k$,

$$V(E_N, M, D) = \sum_{n=1}^{M} e_{n+d_1} e_{n+d_2}, \ldots e_{n+d_k}.$$

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

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

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

$$C_k(E_N) = \max_{M,D} |V(E_N, M, D)| = \max_{M,D} \left| \sum_{n=1}^{M} e_{n+d_1} e_{n+d_2}, \ldots e_{n+d_k} \right|,$$

where the maximum is taken over all $D = (d_1, d_2, \ldots, d_k)$ and M such that $M + d_k \leq N$.

A sequence E_N is considered as a "good" pseudorandom sequence if each of these measures $W(E_N)$, $C_k(E_N)$ (at least for small k) is "small" in terms of N (in particular, all are $o(N)$ as $N \longrightarrow \infty$). Indeed, it was proved in [2, Theorem 1, 2] 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 Alon, Kohayakawa, Mauduit, Moreira and Rödl [1] improved on these bounds.

Numerous binary sequences have been tested for pseudorandomness by J. Cassaigne, S. Ferenczi, C. Mauduit, J. Rivat and A. Sárközy, and large families of pseudorandom binary sequences have been constructed first by L. Goubin, C. Mauduit, A. Sárközy [4]. I gave further construction of this type in [6], [7].

Let p be an odd prime and g be a fixed primitive root modulo p, and for $(n,p)=1$ ind n denotes the index or discrete logarithm of n modulo p. Thus ind n is defined as the unique integer with

$$g^{\mathrm{ind}\, n} \equiv n \pmod p, \tag{1}$$

and $1 \le \mathrm{ind}\, n \le p-1$. Using the discrete logarithm I introduced a new large family of pseudorandom sequences with strong pseudorandom properties in [6]. However the sequences in this family can be generated very slowly, so in [7] I slightly modified the construction so that the new family can be generated much faster. In this paper I will improve on results in [7]. Recently the discrete logarithm is used more and more frequently in cryptography. Chen Li and Xiao [3] generalized the pseudorandom constructions based on the notion of index using elliptic curves. Throughout this paper we will use the following:

Notation *Let p be an odd prime, g be a fixed primitive root modulo p. Define* ind n *by* (1). *Let $f(X) \in \mathbb{F}_p[X]$ be a polynomial of degree $k \ge 1$ which has no multiple roots. Moreover, let*

$$m \mid p-1$$

with $m \in \mathbb{N}$, and let x be coprime with m: $(x,m)=1$.

The crucial idea of the new, faster construction defined in [7] was to reduce $\mathrm{ind}\, n$ modulo m:

Construction 1 *Let* $\mathrm{ind}^* n$ *denote the following function: For all* $1 \le n \le p-1$, *let*

$$\mathrm{ind}\, n \equiv x \cdot \mathrm{ind}^* n \pmod m \text{ and } 1 \le \mathrm{ind}^* n \le m$$

($\mathrm{ind}^* n$ *exists since* $(x,m)=1$.*) Define the sequence* $E_{p-1} = \{e_1, \dots, e_{p-1}\}$ *by*

$$e_n = \begin{cases} +1 & \textit{if } 1 \le \mathrm{ind}^* f(n) \le \frac{m}{2}, \\ -1 & \textit{if } \frac{m}{2} < \mathrm{ind}^* f(n) \le m \textit{ or } p \mid f(n). \end{cases} \tag{2}$$

Note that this construction also generalizes the Legendre symbol construction described in [4] and [5]. Indeed in the special case $m=2$, $x=1$ the sequence e_n defined in (2) becomes

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

In [7], I proved that this construction has good pseudorandom properties: each of the measures $W(E_{p-1})$, $C_k(E_{p-1})$ is less than $p^{1/2}(\log p)^c$ under certain conditions on the polynomial f. However in Theorem 3 in [7] for the correlation measure we obtained an upper bound which is optimal only for large m. Indeed, there the following was proved:

Theorem A *Suppose that m is even, or m is odd with $2m \mid p-1$, and at least one of the following 4 conditions holds:*

a) *f is irreducible;*

b) *If f has the factorization $f = \varphi_1^{\alpha_1}\varphi_2^{\alpha_2}\dots\varphi_u^{\alpha_u}$ over $\mathbb{F}_p$ where $\alpha_i \in \mathbb{N}$ and the φ_i's are irreducible over $\mathbb{F}_p$, then there exists a β such that exactly one or two of φ_i's have the degree β;*

c) *$\ell = 2$;*

d) *$(4\ell)^k < p$ or $(4k)^\ell < p$.*

Then

$$C_\ell(E_{p-1}) < 9k\ell 4^\ell p^{1/2}(\log p)^{\ell+1} + \frac{\ell! k^{\ell(\ell+1)}}{m^\ell}p. \tag{3}$$

If m is odd, then it is proved in section 4 in [7] that

$$C_\ell(E_{p-1}) \gg \frac{p}{m^\ell},$$

thus the second term in (3) can not be dropped completely. If m is even, I will improve on Theorem A under the further assumption that $f(x)$ has no multiple roots.

Theorem 1 *Suppose that m is even, and at least one of the conditions a), b), c) and d) holds in Theorem A. Moreover we suppose that all conditions assumed in the Notation hold, in particular, $f(X) \in \mathbb{F}_p[x]$ has no multiple roots. Then*

$$C_\ell(E_{p-1}) < 9k\ell 4^\ell p^{1/2}(\log p)^{\ell+1}. \tag{4}$$

(4) is considerably sharper than (3) if

$$m^{2\ell+\varepsilon} \ll p.$$

Then the second term is much larger than the first term in (3). In particular, for $m = O(1)$ the second term is $\gg p$, so (3) becomes trivial, while our Theorem 1 still gives good upper bound.

2 Proof of Theorem 1.

Consider any $\mathcal{D} = \{d_1, d_2, \dots, d_\ell\}$ with non-negative integers $d_1 < d_2 < \dots < d_\ell$ and positive integers M with $M + d_\ell \le p - 1$. Then arguing as in the proof of Theorem 2 in [7] (formulas (11) and (12)) we have

$$\begin{aligned} |V(E_N, M, D)| &\le 9k\ell 4^\ell p^{1/2} (\log p)^{\ell+1} \\ &\quad + \frac{2^\ell}{m^\ell} \sum_{\substack{1 \le \delta_1, \dots, \delta_\ell < m, \\ f^{\delta_1}(n+d_1) \cdots f^{\delta_\ell}(n+d_\ell) \text{ is} \\ \text{a perfect } m\text{-th power}}} (p-1) \left| \prod_{j=1}^{\ell} \left(\sum_{l_j=1}^{m/2} \chi^{\delta_j}(g^{x\ell_j}) \right) \right| \\ &= 9k\ell 4^\ell p^{1/2} (\log p)^{\ell+1} + \sum\nolimits_2 . \end{aligned} \tag{5}$$

We will prove that $\sum_2$ is empty, which follows from the following lemma:

Lemma 1 *Suppose that the conditions of Theorem 1 hold. Then if $1 \le \delta_1, \dots, \delta_\ell \le m-1$, and $f^{\delta_1}(n+d_1) \cdots f^{\delta_\ell}(n+d_\ell)$ is a perfect m-th power, then*

$$m \mid \delta_1,\ \delta_2, \dots, \delta_\ell.$$

Indeed, this is a a sharpened version of Lemma 4 in [7]. Assuming the further condition that $f(x)$ has no multiple roots in $\overline{\mathbb{F}}_p$ we will be able to prove this stronger result.

If $\sum_2$ is empty then by (5) we have

$$|V(E_N, M, D)| \leq 9k\ell 4^\ell p^{1/2} (\log p)^{\ell+1},$$

which proves Theorem 1.

Thus it remains to prove Lemma 1. We will need the following definition and lemma:

Definition 1 *Let $\mathcal{A}$ and $\mathcal{B}$ be multi-sets of the elements of $\mathbb{Z}_p$. If $\mathcal{A}+\mathcal{B}$ represents every element of $\mathbb{Z}_p$ with multiplicity divisible by m, i.e., for all $c \in \mathbb{Z}_p$, the number of solutions of*

$$a+b=c \quad a \in \mathcal{A},\ b \in \mathcal{B}$$

(the a's and b's are counted with their multiplicities) is divisible by m, then the sum $\mathcal{A}+\mathcal{B}$ is said to have property P.

Lemma 2 *Let $\mathcal{A} = \{a_1, a_2, \dots, a_r\}$, $\mathcal{D} = \{d_1, d_2, \dots, d_\ell\} \subseteq \mathbb{Z}_p$. If one of the following two conditions holds:*
(i) $\min\{r, \ell\} \leq 2$ and $\max\{r, \ell\} \leq p-1$,
(ii) $(4\ell)^r \leq p$ or $(4r)^\ell \leq p$,
then there exist $c_1, \dots, c_\ell \in \mathbb{Z}_p$ and a permutation $(q_1, \dots, q_\ell)$ of $(d_1, \dots, d_\ell)$ such that for all $1 \leq i \leq \ell$

$$a+d=c_i \qquad a \in \mathcal{A},\ d \in \mathcal{D}$$

has at least one solution, and the number of solutions is less than or equal to i. Moreover for all solutions $a \in \mathcal{A}$, $d \in \mathcal{D}$ we have $d \in \{q_1, q_2 \dots, q_i\}$, and $d = q_i$, $a = c_i - q_i$ is always a solution.

Proof of Lemma 2

This is Lemma 5 in [7].

Now we return to the proof of Lemma 1. The following equivalence relation was defined in [4] and also used in [6] and [7]: We will say that the polynomials $\varphi(X), \psi(X) \in \mathbb{F}_p[X]$ are equivalent, $\varphi \sim \psi$, if there is an $a \in \mathbb{F}_p$ such that $\psi(X) = \varphi(X+a)$. Clearly, this is an equivalence relation.

Write $f(X)$ as the product of irreducible polynomials over $\mathbb{F}_p$. Let us group these factors so that in each group the equivalent irreducible factors are collected. Consider a typical group $\varphi(X+a_1), \dots, \varphi(X+a_r)$. Then $f(X)$ is of the form $f(X) = \varphi(X+a_1)\cdots\varphi(X+a_r)g(X)$ where $g(X)$ has no irreducible factors equivalent with any $\varphi(X+a_i)$ $(1 \le i \le r)$.

Let $h(X) = f^{\delta_1}(X+d_1)\cdots f^{\delta_\ell}(X+d_\ell)$ be a perfect m-th power where $1 \le \delta_1, \dots, \delta_\ell < m$. Then writing $h(x)$ as the product of irreducible polynomials over $\mathbb{F}_p$, all the polynomials $\varphi(X+a_i+d_j)$ with $1 \le i \le r$, $1 \le j \le \ell$ occur amongst the factors. All these polynomials are equivalent, and no other irreducible factor belonging to this equivalence class will occur amongst the irreducible factors of $h(X)$.

Since distinct irreducible polynomials cannot have a common zero, each of the zeros of $h(X)$ is of multiplicity divisible by m, if and only if in each group, formed by equivalent irreducible factors $\varphi(X+a_i+d_j)$ of $h(X)$, every polynomial of form $\varphi(X+c)$ occurs with multiplicity divisible by m. In other words, writing $\mathcal{A} = \{a_1, \dots, a_r\}$,

$\mathcal{D} = \{d_1, \dots, d_1, \dots, d_\ell, \dots, d_\ell\}$ where d_i has the multiplicity δ_i in $\mathcal{D}$ ($h(X) = f^{\delta_1}(X + d_1) \cdots f^{\delta_\ell}(X + d_\ell)$ is a perfect m-th power), then for each group $\mathcal{A} + \mathcal{D}$ must possess property P.

Let $\mathcal{D}'$ be the simple set version of $\mathcal{D}$, more exactly, let $\mathcal{D}' = \{d_1, \dots, d_\ell\}$. $\mathcal{A}$ and $\mathcal{D}'$ satisfy the conditions of Lemma 2. So by Lemma 2 for the sets $\mathcal{A}$ and $\mathcal{D}'$ we have the following: There exist $c_1, \dots, c_\ell \in \mathbb{Z}_p$ and a permutation $(q_1, \dots, q_\ell) = (d_{j_1}, \dots, d_{j_\ell})$ of $(d_1, \dots, d_\ell)$ such that if

$$a + d = c_i \qquad a \in \mathcal{A},\ d \in \mathcal{D}',$$

then we have

$$d \in \{q_1, \dots, q_i\} = \{d_{j_1}, \dots, d_{j_i}\}$$

and $d = q_i$, $a = c_i - q_i$ is a solution. Here $(j_1, \dots, j_\ell)$ is a permutation of $(1, \dots, \ell)$. Define the ρ_i's by $\rho_i = \delta_{j_i}$ (so $(\rho_1, \dots, \rho_\ell) = (\delta_{j_1}, \dots, \delta_{j_\ell})$ is the same permutation of $(\delta_1, \dots, \delta_\ell)$ as the permutation $(q_1, \dots, q_\ell) = (d_{j_1}, \dots, d_{j_\ell})$ of $(d_1, \dots, d_\ell)$). Returning to the multi-set case, using this notation we get that the number of the solutions

$$a + d = c_i \qquad a \in \mathcal{A},\ d \in \mathcal{D}$$

is of the form

$$\varepsilon_{i,1}\rho_1 + \varepsilon_{i,2}\rho_2 + \cdots + \varepsilon_{i,i}\rho_i$$

where $\varepsilon_{i,j} \in \{0, 1\}$ for $1 \le j \le i$ and $\varepsilon_{i,i} = 1$. (We study the number of solutions by multiplicity since $\mathcal{D}$ is a multi-set).

Since $\mathcal{A} + \mathcal{D}$ possesses property $\mathcal{P}$ for all $1 \le i \le \ell$ we have

$$m \mid \varepsilon_{i,1}\rho_1 + \varepsilon_{i,2}\rho_2 + \cdots + \varepsilon_{i,i}\rho_i. \tag{6}$$

By induction on i we will prove that

$$m \mid \rho_i. \tag{7}$$

Indeed, for $i = 1$ by (6) and $\varepsilon_{1,1} = 1$ we get $m \mid \rho_1$. We will prove that if (7) holds for $i \leq j-1$, then it also holds for $i = j$.

By the induction hypothesis we have

$$m \mid \rho_1,\ m \mid \rho_2,\ \ldots,\ m \mid \rho_{j-1}.$$

Using this and (6) for $i = j$ we get:

$$m \mid \rho_j,$$

which was to be proved.

References

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

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

[3] Z. Chen, S. Li, G. Xiao, *Construction of Pseudo-random Binary Sequences from Elliptic Curves by Using Discrete Logarithm*, Sequences and Their Applications, SETA 2006, Lecture Notes in Computer Science 4086 Springer 2006.

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

[5] C. Mauduit, A. Sárközy, *On finite pseudorandom binary sequences I: Measures of pseudorandomness, the Legendre symbol*, Acta Arith. 82 (1997), 365-377.

[6] K. Gyarmati, *On a family of pseudorandom binary sequences*, Periodica Math. Hungar. 49 (2004), 45-63.

[7] K. Gyarmati, *On a fast version of a pseudorandom generator*, General Theory of Information Transfer and Combinatorics, Lecture Notes in Computer Science 4123, Springer, Berlin / Heidelberg 2006, 326-342.