An inequality between the measures of pseudorandomness

Katalin Gyarmati

1 Introduction

In this paper I will improve on a generalization of an inequality of Mauduit and Sárközy [6]. They introduced the following measures of pseudorandomness in [5]:

For a binary sequence

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

write

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

and, for $D = (d_1, \dots, d_k)$ with non-negative integers $0 \le d_1 < \dots < d_k$,

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

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

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

where the maximum is taken over all a, b, t such that $a \in \mathbb{Z}$, $b, t \in \mathbb{N}$ and $1 \le a + b \le a + tb \le N$, while the *correlation measure of order k of E_N* is defined as

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

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

In [6] Mauduit and Sárközy proved that for all sequences $E_N \in \{-1, +1\}^N$ we have $W(E_N) \le \sqrt{NC_2(E_N)}$. Later in [3] this inequality was generalized by me to correlation measure of any even order: If $3\ell^2 \le N$ and $E_N \in \{-1, +1\}^N$ then $W(E_N) \le 3\ell N^{1-1/(2\ell)} \left(C_{2\ell}(E_N)\right)^{1/(2\ell)}$. In the present paper I will improve on the factor 3ℓ showing that this inequality even holds with an absolute constant factor:

Theorem 1 *If $\varepsilon > 0$, $N \ge 18\ell/\varepsilon^2$, then for all $E_N \in \{-1, +1\}^N$ we have*

$$W(E_N) \le (\sqrt{2} + \varepsilon) N^{1-1/(2\ell)} C_{2\ell}(E_N)^{1/(2\ell)}.$$

Mauduit and Sárközy [6] also proved that their inequality is sharp by using probabilistic arguments. In [3] I presented an explicit construction for which the generalized inequality is sharp apart from a $\sqrt{\ell}$ factor. This construction was based on the notion of index (discrete logarithm): Denote ind n the index of n modulo p, defined as the unique integer with

$$g^{\text{ind } n} \equiv n \pmod{p},$$

and $1 \le \text{ind } n \le p - 1$, where g is a fixed primitive root modulo p. Let $\text{ind}^* n$ be the modulo m residue of ind n:

$$\text{ind}^* n \equiv \text{ind } n \pmod{m} \tag{1}$$

with $1 \le \mathrm{ind}^* n \le m$.

Construction 1 *Let* $m \mid p-1$ *and* $\mathrm{ind}^* n$ *be the function defined by* (1). *Then let the sequence* $E_{p-1} = \{e_1, \dots, e_{p-1}\}$ *be*

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

where $f(x) \in \mathbb{F}_p[x]$ *is a polynomial with the degree* k.

In Theorem 1 and 3 in [3] I gave estimates for the well-distribution measure and correlation measures of this sequence E_{p-1} if some, not too restrictive conditions hold on the polynomial $f(x)$. Then

$$W(E_{p-1}) \gg \frac{1}{\sqrt{\ell k^{\ell+1}}} p^{1-1/(2\ell)} \left(C_{2\ell}(E_{p-1})\right)^{1/(2\ell)} \tag{3}$$

follows from these theorems, where the implied constant factor is absolute.

This inspired me to consider the simplest polynomial $f(x) = x$ in Construction 1, hoping that inequality (3) holds with a factor larger than $\frac{1}{\sqrt{\ell}}$. Indeed we will study the following sequence:

Construction 2 *Let* $m \mid p-1$ *and* $\mathrm{ind}^* n$ *be the function defined by* (1). *Then let the sequence* $E_{p-1} = \{e_1, \dots, e_{p-1}\}$ *be*

$$e_n = \begin{cases} +1 & \text{if } 1 \le \mathrm{ind}^* n \le \frac{m}{2}, \\ -1 & \text{if } \frac{m}{2} < \mathrm{ind}^* n \le m. \end{cases} \tag{4}$$

For this sequence we have:

Theorem 2 *If* m *is even then the sequence in Construction 2 satisfies*

$$W(E_{p-1}) \le 36 p^{1/2} \log p \log(m+1)$$

while for odd m we have

$$W(E_{p-1}) = \frac{p-1}{m} + O(p^{1/2}\log p \log(m+1)).$$

Indeed, this is Theorem 1 in [3] in the special case when k, the degree of the polynomial is 1.

In case of the correlation measure we will give slightly better upper bound than in Theorem 3 (in the special case $k = 1$) in [3]:

Theorem 3 *If m is even then the sequence in Construction 2 satisfies:*

$$C_\ell(E_{p-1}) \le 9\ell 4^\ell p^{1/2} \log p \left(\log m\right)^\ell,$$

while for odd m we have

$$C_\ell(E_{p-1}) = \frac{p}{m^\ell} + O(5^\ell p^{1/2}\log p(\log m)^\ell).$$

It follows from Theorems 2 and 3:

Corollary 1 *For every $\varepsilon > 0$ there exist positive constants $p_0(\varepsilon)$ and $c_0(\varepsilon)$ such that if $p > p_0(\varepsilon)$ and m is an odd divisor of $p-1$ with*

$$m < c_0(\varepsilon)\frac{p^{1/(2\ell)}}{\ell\left(\log p\right)^{1+1/\ell}} \tag{5}$$

(so $\frac{p}{m^\ell} \gg 5^\ell p^{1/2}\log p\left(\log m\right)^\ell$), then

$$W(E_{p-1}) \ge (1-\varepsilon)p^{1-1/(2\ell)}\left(C_{2\ell}(E_{p-1})\right)^{1/(2\ell)}. \tag{6}$$

I remark that to make sure that condition (5) holds, first we fix an odd integer m, and after this we look for a prime number p with $m \mid p-1$ and (5). This is possible by Dirichlet's theorem on primes in arithmetic progressions.

So, indeed Theorem 1 is best possible apart from a constant factor. The interesting feature of this proof is that it is explicit, we give a sequence for which (6) holds. In the most cases there is only an existence proof for the sharpness of an inequality between pseudorandom measures.

2 Proofs of Theorem 1 and 3

Proof of Theorem 1

It follows from the definition of $W(E_N)$ that there exist $a \in \mathbb{Z}$, $b, t \in \mathbb{N}$ with $1 \leq a + b < a + tb \leq N$ such that

$$W(E_N) = \Bigg| \sum_{\substack{a+b \leq i \leq a+tb \\ i \equiv a+b \pmod b}} e_i \Bigg|. \tag{7}$$

For $0 \leq h < b$ let

$$D_h \stackrel{\text{def}}{=} \Bigg(\sum_{\substack{a+b \leq i \leq a+tb \\ i \equiv h \pmod b}} e_i \Bigg)^{2\ell} - 2\ell! \sum_{\substack{a+b \leq i_1 < \dots < i_{2\ell} \leq a+tb \\ h \equiv i_1 \equiv \dots \equiv i_{2\ell} \pmod b}} e_{i_1} \dots e_{i_{2\ell}}. \tag{8}$$

Using the multinomial theorem we get that D_h is a sum of products of the form $c \cdot e_{j_1} \dots e_{j_r}$ where $c \geq 0$. Thus D_h takes his maximum when all e_i's are $+1$ (or all e_i's are -1). So:

$$\begin{aligned} D_h &\leq \Bigg(\sum_{\substack{a+b \leq i \leq a+tb \\ i \equiv h \pmod b}} 1 \Bigg)^{2\ell} - 2\ell! \sum_{\substack{a+b \leq i_1 < \dots < i_{2\ell} \leq a+tb \\ h \equiv i_1 \equiv \dots \equiv i_{2\ell} \pmod b}} 1 \\ &\leq t^{2\ell} - (t-1)(t-2)\dots(t-2\ell) \leq t^{2\ell} - (t-2\ell)^{2\ell} \leq 4\ell^2 t^{2\ell-1}. \end{aligned}$$

By this, (7) and (8) we have

$$\begin{aligned}(W(E_N))^{2\ell} &\le \sum_{h=0}^{b-1}\Big(\sum_{\substack{a+b\le i\le a+tb\\ i\equiv h \pmod b}} e_i\Big)^{2\ell}\\ &= \sum_{h=0}^{b-1}\Big(D_h + 2\ell! \sum_{\substack{a+b\le i_1<\dots<i_{2\ell}\le a+tb\\ h\equiv i_1\equiv\dots\equiv i_{2\ell} \pmod b}} e_{i_1}\dots e_{i_{2\ell}}\Big)\\ &\le \sum_{h=0}^{b-1}\Big(4\ell^2 t^{2\ell-1} + 2\ell! \sum_{\substack{a+b\le i_1<\dots<i_{2\ell}\le a+tb\\ h\equiv i_1\equiv\dots\equiv i_{2\ell} \pmod b}} e_{i_1}\dots e_{i_{2\ell}}\Big)\\ &= 4b\ell^2 t^{2\ell-1} + 2\ell! \sum_{\substack{a+b\le i_1<\dots<i_{2\ell}\le a+tb\\ i_1\equiv\dots\equiv i_{2\ell} \pmod b}} e_{i_1}\dots e_{i_{2\ell}}.\end{aligned}$$

From this replacing i_2 by i_1+d_1, i_3 by i_1+d_2 and so on, finally $i_{2\ell}$ by $i_1+d_{2\ell-1}$ we obtain

$$\begin{aligned}(W(E_N))^{2\ell} \le 4b\ell^2 t^{2\ell-1} + 2\ell!&\\ \sum_{\substack{1\le d_1<\dots<d_{2\ell-1}\le(t-1)b\\ d_1\equiv\dots\equiv d_{2\ell-1}\equiv 0 \pmod b}} \ \sum_{i_1=a+b}^{a+tb-d_{2\ell-1}} & e_{i_1}e_{i_1+d_1}\dots e_{i_1+d_{2\ell-1}}. \end{aligned} \tag{9}$$

By the definition of the correlation measure we have

$$\Big|\sum_{i_1=a+b}^{a+tb-d_{2\ell-1}} e_{i_1}e_{i_1+d_1}\dots e_{i_1+d_{2\ell-1}}\Big| \le C_{2\ell}E_N. \tag{10}$$

By $tb\le a+tb\le N$ we have $4b\ell^2t^{2\ell-1}=4\ell^2(tb)t^{2\ell-2}\le 4\ell^2N^{2\ell-1}$, and so from (9) and (10) we obtain

$$\begin{aligned}(W(E_N))^{2\ell} &\le 4\ell^2N^{2\ell-1} + 2\ell!\frac{N^{2\ell-1}}{(2\ell-1)!}C_{2\ell}(E_N)\\ &= 2\ell\left(1+\frac{2\ell}{C_{2\ell}(E_N)}\right)N^{2\ell-1}C_{2\ell}(E_N).\end{aligned}$$

From this by the binomial theorem we get:

$$W(E_N) \le (2\ell)^{1/(2\ell)}\left(1+\frac{1}{C_{2\ell}(E_N)}\right)N^{1-1/(2\ell)}\left(C_{2\ell}(E_N)\right)^{1/(2\ell)}.$$

Kohayakawa, Mauduit, Moreira and V. Rödl [4] proved that $C_{2\ell}(E_N) > \sqrt{\frac{N}{3(2\ell+1)}}$ holds for all $E_N \in \{-1,+1\}^N$ by this and since $(2\ell)^{1/(2\ell)} \leq \sqrt{2}$ we get:

$$W(E_N) \leq \sqrt{2}\left(1+\sqrt{\frac{3(2\ell+1)}{N}}\right) N^{1-1/(2\ell)}\left(C_{2\ell}(E_N)\right)^{1/(2\ell)}.$$

If $N \geq 18\ell/\varepsilon^2 \geq 6(2\ell+1)/\varepsilon^2$ then this completes the proof of the theorem.

Proof of Theorem 3

The proof of the theorem is very similar to the proof of Theorem 1 in [2]. By the formula

$$\frac{1}{m}\sum_{\chi:\chi^m=1}\overline{\chi}^j(a)\chi(b) = \begin{cases} 1 & \text{if } m \mid \text{ind } a - \text{ind } b, \\ 0 & \text{if } m \nmid \text{ind } a - \text{ind } b, \end{cases}$$

we obtain

$$e_n = 2\sum_{\substack{1\leq i\leq m/2 \\ i\equiv \text{ind } n \pmod m}} 1 - 1 = \frac{2}{m}\sum_{1\leq i\leq m/2}\ \sum_{\chi:\chi^m=1}\overline{\chi}(n)\chi(g^i) - 1.$$

Thus

$$e_n = \frac{2}{m}\left(\sum_{1\leq i\leq m/2}\ \sum_{\chi\neq\chi_0:\chi^m=1}\overline{\chi}(n)\chi(g^j) + \frac{(-1)^m-1}{4}\right). \tag{11}$$

To prove Theorem 3, consider any $\mathcal{D} = \{d_1, d_2, \ldots, d_\ell\}$ with non-negative integers $d_1 < d_2 < \cdots < d_\ell$ and positive integer M with $M + d_\ell \leq p-1$.

Then arguing as in [7, p. 382] with m in place of $p-1$ from (11) we obtain:

$$V(E_N, M, D) = \frac{2^\ell}{m^\ell} \sum_{n=1}^{M} \prod_{j=1}^{\ell} \left(\sum_{1 \le i \le m/2} \sum_{\substack{\chi_j \ne \chi_0, \\ \chi_j^m = 1}} \overline{\chi_j}(n+d_j)\chi_j(g^i) + \frac{(-1)^m + 1}{4} \right)$$

$$= \frac{2^\ell}{m^\ell} \left(\sum_{k=0}^{\ell} \sum_{1 \le j_1 < \dots < j_k \le \ell} \left(\frac{(-1)^m + 1}{4} \right)^{\ell-k} \sum_{\substack{\chi_{j_1} \ne \chi_0, \\ \chi_{j_1}^m = 1}} \dots \sum_{\substack{\chi_{j_k} \ne \chi_0, \\ \chi_{j_k}^m = 1}} \right.$$

$$\left. \sum_{n=1}^{M} \overline{\chi}_{j_1}(n+d_{j_1}) \dots \overline{\chi}_{j_k}(n+d_{j_k}) \prod_{t=1}^{k} \left(\sum_{1 \le \ell_t \le m/2} \chi_{j_i}(g^{\ell_t}) \right) \right). \tag{12}$$

Let $S_0 = M$, $V_0 = \left(\frac{1}{2}\right)^\ell$ and for $1 \le k \le \ell$ let

$$S_k = \max_{\substack{\chi_1 \ne \chi_0, \dots, \chi_k \ne \chi_0 \\ 1 \le j_1 < \dots < j_k \le \ell}} \left| \sum_{n=1}^{M} \overline{\chi}_1(n+d_{j_1}) \dots \overline{\chi}_k(n+d_{j_k}) \right| \tag{13}$$

and

$$V_k = \sum_{1 \le j_1 < \dots < j_k \le \ell} \left(\frac{1}{2}\right)^{\ell-k} \sum_{\substack{\chi_{j_1} \ne \chi_0, \\ \chi_{j_1}^m = 1}} \dots \sum_{\substack{\chi_{j_k} \ne \chi_0, \\ \chi_{j_k}^m = 1}} \prod_{t=1}^{k} \left| \sum_{1 \le \ell_t \le m/2} \chi_{j_i}(g^{\ell_t}) \right|. \tag{14}$$

Then by the triangle-inequality, the value of $\frac{(-1)^m+1}{4}$ and (12) we obtain that if m is even then

$$|V(E_N, M, D)| \le \frac{2^\ell}{m^\ell} S_\ell V_\ell \tag{15}$$

and

$$V(E_N, M, D) = \frac{2^\ell}{m^\ell} S_0 V_0 + O\left(\frac{2^\ell}{m^\ell} \sum_{k=1}^{\ell} S_k V_k \right) \tag{16}$$

Next we give an upper bound for S_k. In order to do this we will use the following lemma:

Lemma 1 *Suppose that p is a prime, χ is a non-principal character modulo p of order z, $f \in \mathbb{F}_p[x]$ has s distinct roots in $\overline{F}_p$, and it is not a constant multiple of a z-th power of a polynomial over $\mathbb{F}_p$. Let y be a real number with $0 < y \le p$. Then for any $x \in \mathbb{R}$:*

$$\left| \sum_{x<n\le x+y} \chi(f(n)) \right| < 9sp^{1/2} \log p.$$

Poof of Lemma 1

This is a trivial consequence of Lemma 1 in [1]. Indeed, there this result is deduced from Weil's theorem, see [8].

Now let χ be a modulo p character of order m; for simplicity we will choose χ as the character uniquely defined by $\chi(g) = e\left(\frac{1}{m}\right)$.

Returning to the estimate of S_k, let $\overline{\chi_u} = \chi^{\delta_u}$ for $u = 1, 2, \ldots, \ell$, whence by $\chi_1 \ne \chi_0, \ldots, \chi_\ell \ne \chi_0$, we may take

$$1 \le \delta_u < m.$$

Thus in (13) we have

$$\left| \sum_{n=1}^{M} \overline{\chi}_1(n + d_{j_1}) \ldots \overline{\chi}_k(n + d_{j_k}) \right| = \left| \sum_{n=1}^{M} \chi^{\delta_1}(n + d_{j_1}) \ldots \chi^{\delta_\ell}(n + d_{j_k}) \right|$$
$$= \left| \sum_{n=1}^{M} \chi\left((n + d_{j_1})^{\delta_1} \ldots (n + d_{j_k})^{\delta_k}\right) \right|.$$

Since $(n + d_{j_1})^{\delta_1} \ldots (n + d_{j_k})^{\delta_k}$ is not a perfect m-th power, this sum can be estimated by Lemma 1, whence

$$S_k \le 9kp^{1/2} \log p. \tag{17}$$

By (14) we have

$$V_k = \sum_{1\le j_1\cdots\le j_k\le \ell} \left(\frac{1}{2}\right)^{\ell-k} \left(\sum_{\substack{\chi\ne\chi_0,\\ \chi^m=1}} \left| \sum_{j=1}^{[m/2]} \chi^j(g) \right| \right)^k$$

Lemma 2

$$\sum_{\substack{\chi\ne\chi_0,\\ \chi^m=1}} \left| \sum_{j=1}^{[m/2]} \chi^j(g) \right| \le \sum_{\substack{\chi\ne\chi_0,\\ \chi^m=1}} \frac{2}{|1-\chi(g)|} < 2m\log(m+1).$$

Proof of Lemma 2 This is Lemma 3 in [2] with m in place of d and $m/2$ in place of $(p-1)/2$, and it can be proved in the same way.

Using Lemma 2 we obtain

$$V_k \le \sum_{1\le j_1\cdots\le j_k\le \ell} \left(\frac{1}{2}\right)^{\ell-k} \left(2m\left(\log(m+1)\right)^k\right) = \frac{4^k}{2^\ell}\binom{\ell}{k} m^k \left(\log(m+1)\right)^k \tag{18}$$

By (15), (16), (17) and (18) we obtain that if m is even then

$$|V(E_N,M,D)| \le 9\ell 4^\ell p^{1/2}\log p\left(\log(m+1)\right)^\ell,$$

and if m is odd then

$$\begin{aligned} V(E_N,M,D) &= \frac{M}{m^\ell} + O\left(\frac{9p^{1/2}\log p}{m^\ell}\sum_{k=1}^{\ell} k\binom{\ell}{k} 4^k m^k \left(\log(m+1)\right)^k\right) \\ &= \frac{M}{m^\ell} + O\left(\frac{9\ell p^{1/2}\log p}{m^\ell}\left(4m\log(m+1)\right)^\ell\right) \\ &= \frac{M}{m^\ell} + O\left(5^\ell p^{1/2}\log p\left(\log(m+1)\right)^\ell\right), \end{aligned}$$

which completes the proof of the theorem.

References

[1] R. Ahlswede, C. Mauduit, A. Sárközy, *Large families of pseudorandom sequences of k symbols and their complexity, Part I, Part II.*, Proceedings on General Theory of Information Transfer and Combinatorics, to appear.

[2] K. Gyarmati, *On a family of pseudorandom binary sequences*, Periodica Math. Hungar., to appear.

[3] K. Gyarmati, *On a fast version of a pseudorandom generator*, General Theory of Information Transfer and Combinatorics, Conference Proceedings, to appear.

[4] Y. Kohayakawa, C. Mauduit, C. G. Moreira and V. Rödl, *Measures of pseudorandomness for finite sequences: minimum and typical values*, submitted to J. London Math. Soc.

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

[6] C. Mauduit, A. Sárközy, *On the measures of pseudorandomness of binary sequences*, Discrete Math. 271 (2003), 195-207.

[7] A. Sárközy, *A finite pseudorandom binary sequence*, Studia Scientiarum Mathematicarum Hungarica 38 (2001), 377-384.

[8] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Act. Sci. Ind. 1041, Hermann, Paris, 1948.