On the cross-combined measure of families of binary lattices and sequences

Katalin Gyarmati

Eötvös Loránd University, Faculty of Siences,
Department of Algebra and Number Theory,
Hungary, Budapest

gykati@cs.elte.hu

Measures of pseudorandomness

The constructive and quantitative study of pseudorandomness stared by the work of Mauduit and Sárközy. They introduced the following measures:

Definition

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

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

and

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

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 $C_\ell(E_N)$ are "small"."

E_N has strong pseudorandom properties if

$$W(E_N),\ C_\ell(E_N) \ll N^{1-\varepsilon}.$$

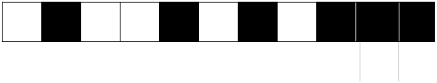
But in the best constructions we have:

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

Hoffstein and Lieman studied the following construction: Let $f(x) \in \mathbb{F}_p[x]$ be a polynomial of degree k. Consider the sequence

$$E_p(f) = \left(\left(\frac{f(1)}{p} \right), \left(\frac{f(2)}{p} \right), \ldots, \left(\frac{f(p)}{p} \right) \right) \quad \text{where now} \quad \left(\frac{0}{p} \right) \overset{\text{def}}{=} 1.$$

For example, for $p = 7$, $f(x) = x^2 + 1$ this sequence can be illustrated by



■ = +1

□ = −1

Hoffstein and Lieman didn't prove anything on the pseudorandom properties of this sequence.

$$E_p(f) = \left(\left(\frac{f(1)}{p} \right), \left(\frac{f(2)}{p} \right), \dots, \left(\frac{f(p)}{p} \right) \right)$$

In 2004 Goubin, Mauduit and Sárközy proved that under some not too restrictive conditions on the prime p or the polynomial $f(x)$ we have

$$W(E_p(f)),\ \ C_\ell(E_p(f)) \ll p^{1/2} \log p$$

Since then several other constructions have been given, but still this is one of the best: it has strong pseudorandom properties and its implementation is relatively easy and fast.

I remark that several important a posteriori tests (indicated by a package of the National Institute of Standards and Technology) were checked by Rivat, Mauduit and Sárközy. This work was continued by Mérai, Rivat and Sárközy.

Large families of binary sequences

Let $\mathcal{P}_{\leq K}$ denote the set of monic polynomials $f(x) \in \mathbb{F}_p[x]$ of degree k where $1 \leq k \leq K$.

For $f \in \mathcal{P}_{\leq K}$ define

$$E_p(f) = \left(\left(\frac{f(1)}{p} \right), \left(\frac{f(2)}{p} \right), \ldots, \left(\frac{f(p)}{p} \right) \right).$$

Let $\mathcal{F}_{\leq K, Legendre}$ denote the following large family of binary sequences:

$$\mathcal{F}_{\leq K, Legendre} = \{ E_p(f) : \ f \in \mathcal{P}_{\leq K} \}$$

By the result of Goubin, Mauduit and Sárközy many individual sequences belonging to this family have strong pseudorandom properties.

In many applications it is not enough to know that the family contains many sequences with strong pseudorandom properties it is much more important to know that the given family has a “rich”, “complex” structure.

For example, there are many independent sequences in it.

Earlier family measures

Ahlswede, Khachatrian, Mauduit and Sárközy: family complexity

Based on the well know Hamming distance: distance minimum.

Bérczes, Ködmön, Pethő, Tóth and others: avalanche effect and collision free.

Mauduit, Sárközy and Gy.: cross-correlation measure.

Now I will focus on the cross-correlation measure.

For sequences $E_N^{(1)}, \dots, E_N^{(\ell)} \in \{-1,+1\}^N$ write

$$E_N^{(i)} = (e_1^{(i)}, e_2^{(i)}, \dots e_N^{(i)}) \qquad \text{for } i = 1, 2, \dots, \ell$$

and then define

$$\widetilde{C}(E_N^{(1)}, E_N^{(2)}, \dots, E_N^{(\ell)}) \stackrel{\text{def}}{=} \max_{M, d_1 < d_2 < \dots < d_\ell}{}^{*} \left| \sum_{n=1}^{M} e_{n+d_1}^{(1)} \cdots e_{n+d_\ell}^{(\ell)} \right| .$$

where if the sequences $E_N^{(i)}$ and $E_N^{(j)}$ (for some $1 \le i < j \le \ell$) are identical then in the maximum $d_i = d_j$ is not allowed.

Let $\mathcal{F}$ be a large families of binary sequences. Define the cross-correlation measure by

$$\Phi_\ell(\mathcal{F}) \stackrel{\text{def}}{=} \max_{E_N^{(1)}, \dots, E_N^{(\ell)} \in \mathcal{F}} \widetilde{C}(E_N^{(1)}, E_N^{(2)}, \dots, E_N^{(\ell)})$$

Mauduit, Sárközy, Gy.: If the cross-correlation measure is small then the sequences in the family are highly independent.

For $f \in \mathcal{P}_{\leq K}$ define

$$E_p(f) = \left(\left(\frac{f(1)}{p} \right), \left(\frac{f(2)}{p} \right), \ldots, \left(\frac{f(p)}{p} \right) \right).$$

Let $\mathcal{F}_{\leq K, Legendre}$ denote the following large family of binary sequences:

$$\mathcal{F}_{\leq K, Legendre} = \{ E_p(f) : \ f \in \mathcal{P}_{\leq K} \}$$

Unfortunately $\mathcal{F}_{\leq K, Legendre}$ has a very bad cross-correlation measure.

$\mathcal{F}_{\leq K, Legendre}$ has a very bad cross-correlation measure: Consider the polynomials $f(x) = x, x+1$ and $x(x+1)$. Let

$$E_p^{(1)} \overset{\text{def}}{=} E_p(x) = \left(\left(\frac{1}{p}\right), \left(\frac{2}{p}\right), \dots, \left(\frac{p}{p}\right)\right) \qquad \text{so } e_n^{(1)} = \left(\frac{n}{p}\right)$$

$$E_p^{(2)} \overset{\text{def}}{=} E_p(x+1) = \left(\left(\frac{2}{p}\right), \left(\frac{3}{p}\right), \dots, \left(\frac{p+1}{p}\right)\right) \text{ so } e_n^{(2)} = \left(\frac{n+1}{p}\right)$$

$$E_p^{(3)} \overset{\text{def}}{=} E_p(x(x+1)) = \left(\left(\frac{1 \cdot 2}{p}\right), \left(\frac{2 \cdot 3}{p}\right), \dots, \left(\frac{p \cdot (p+1)}{p}\right)\right)$$

$$\text{so } e_n^{(3)} = \left(\frac{n(n+1)}{p}\right)$$ Then

$$\Phi_3(\mathcal{F}_{\leq K, Legendre}) \geq \widetilde{C}(E_p^{(1)}, E_p^{(2)}, E_p^{(3)}) \geq \left| \sum_{n=1}^{p} e_n^{(1)} e_n^{(2)} e_n^{(3)} \right|$$

$$= \left| \sum_{n=1}^{p} \left(\frac{n}{p}\right) \left(\frac{n+1}{p}\right) \left(\frac{n(n+1)}{p}\right) \right| = p - 2.$$

Construction (Mauduit, Sárközy, Gy.)

$$\mathcal{F}_{\leq K, Legendre, irreducible} = \Big\{ E_p(f) : \ f \in \mathcal{P}_{\leq K} \ \text{ and } \\ f(x) = x^k + a_{k-2}x^{k-2} + \cdots + a_0 \text{ is irreducible } \Big\}$$

This family has strong cross-correlation measure:

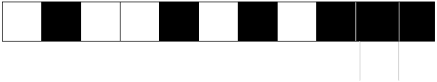
$$\Phi_\ell(\mathcal{F}_{\leq K, Legendre, irreducible}) \ll \ell K p^{1/2} \log p.$$

Problem: It is very difficult and slow to generate one variable irreducible polynomials over $\mathbb{F}_p$...

Solution: Consider the multidimensional case!

Binary lattices

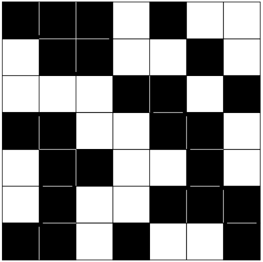
Binary sequences can be illustrated in the following way:



■ = $+1$

□ = -1

Binary lattices are the multidimensional extensions of sequences. Two dimensional binary lattices can be illustrated the following way:



First Hubert, Mauduit and Sárközy studied the multidimensional case, they introduced the following precise definitions (here I present only the two dimensional case):

$$I_N^2 = \{\mathbf{x} = (x_1, x_2) : \ 0 \le x_1 \le N-1, \ 0 \le x_2 \le N-1\}$$

Then a function

$$\eta : I_N^2 \to \{-1, +1\}$$

called as a two dimensional binary lattice.

$B \subseteq I_N^2$ is a box-lattice if it is of the form

$$B = \{\mathbf{x} = (x_1 u_1, x_2 u_2) : \ 0 \le x_1 < t_1, \ 0 \le x_2 < t_2\}$$

The combined pseudorandom measure of η of order ℓ is

$$Q_\ell(\eta) \overset{\text{def}}{=} \max_{B, d_1, \dots, d_\ell} \left| \sum_{x \in B} \eta(x + d_1) \cdots \eta(x + d_\ell) \right|$$

Analogously to the one dimensional cross-correlation measure, I defined the cross-combined measure for binary lattices:

For lattices $\eta_1, \eta_2, \dots, \eta_\ell : I_N^2 \rightarrow$ write define

$$\widetilde{Q}(\eta_1, \eta_2, \dots, \eta_\ell) \overset{\text{def}}{=} \max_{B, \mathbf{d_1} < \mathbf{d_2} < \cdots < \mathbf{d}_\ell}^{*} \left| \sum_{\mathbf{x} \in B} \eta_1(\mathbf{x} + \mathbf{d_1}) \cdots \eta_\ell(\mathbf{x} + \mathbf{d}_\ell) \right|$$

where if the lattices η_i and η_j (for some $1 \leq i < j \leq \ell$) are identical then in the maximum $\mathbf{d_i} = \mathbf{d_j}$ is not allowed.

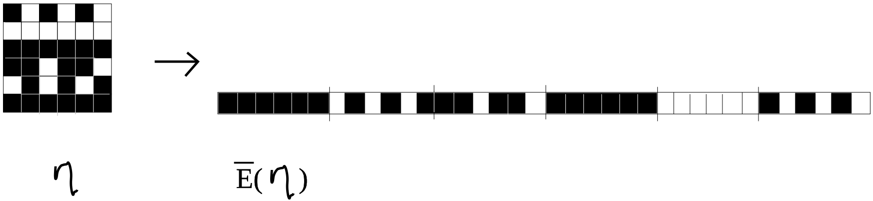
Let $\mathcal{G}$ be a large families of binary lattices. Define the cross-combined measure by

$$\Phi_\ell(\mathcal{G}) \overset{\text{def}}{=} \max_{\eta_1, \dots, \eta_\ell \in \mathcal{G}} \widetilde{Q}(\eta_1, \eta_2, \dots, \eta_\ell)$$

I studied the connections of the cross-combined measure with other family measures of lattices defined earlier.

Each earlier one-dimensional results can be easily extended to the multidimensional case.

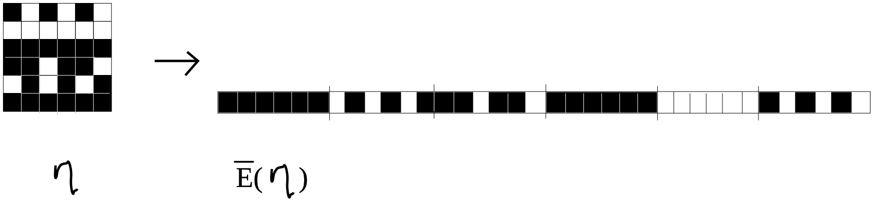
With Mauduit and Sárközy we studied connections between lattices and sequences. Starting out from a binary lattice η by taking the first (from the bottom) row of the lattice then we continue by the second row, then the third row follows etc. we get a binary sequence $E(\eta)$



Gyarmati, Mauduit, Sárközy: There exist binary lattices which have very large pseudorandom measures, but the corresponding binary sequence has strong pseudorandom properties.

Gy.: If the lattice has strong pseudorandom properties, the corresponding binary sequence also has:

$$C_\ell(\overline{E}(\eta)) \leq (\ell + 2) Q_\ell(\eta).$$



Let $\mathcal{G}$ be a family of two dimensional binary lattices. Then we define $\mathcal{F} = (\overline{E}(\mathcal{G}))$ of binary sequences by $\overline{E}(\mathcal{G}) \overset{\text{def}}{=} \{\overline{E}(\eta) : \ \eta \in \mathcal{G}\}$.

Theorem (Gy.): $\Phi_\ell(\overline{E}(\mathcal{G})) \leq (\ell + 2)\Phi_\ell(\mathcal{G})$.

This theorem gives a new way of construction of a family of binary sequences with strong cross-combined measure: First we generate a family of binary lattices $\mathcal{G}$ with strong cross-combined measure and from this we easily generate the family $\overline{E}(\mathcal{G})$ of binary sequences which also has strong cross-correlation measure.

I proved for more earlier constructions of family of binary lattices that they have optimal or nearly optimal cross-combined measure.

Theorem (Gy.)

Let γ denote the quadratic character of $\mathbb{F}_{p^2}$ and let v_1, v_2 be a basis of the vectorspace $\mathbb{F}_{p^2}$ over $\mathbb{F}_p$. Let $G_{\leq K,quadratic,irreducible}$ denote the family of all binary lattices $\eta : \ I_p^2 \rightarrow \{-1, +1\}$ defined by

$$\eta((x_1, x_2)) = \gamma(f(x_1 v_1 + x_2 v_2))$$

where the used one variable irreducible polynomial f is the form $f(x) = x^k + a_{k-2}x^{k-2} + \cdots + a_0 \in \mathbb{F}_{p^2}[x]$ with degree $k \leq K$. Then

$$\Phi_\ell(G_{\leq K,quadratic,irreducible}) \ll \ell K p \log p.$$

Thus this family has optimal cross-combined measure. Using the previous theorem about the connections of lattices and sequences we also get the following:

$$\Phi_\ell(\overline{E}(G_{\leq K,quadratic,irreducible})) \ll \ell^2 K p \log p.$$

The difficulties of the generation of lattices from the family $G_{\leq K,quadratic,irreducible}$ or
sequences from the family $\overline{E}(G_{\leq K,quadratic,irreducible})$ is that they are based on one-variable irreducible polynomials over $\mathbb{F}_{p^2}$.

The generation of one-variable irreducible polynomials over finite fields is complicated and slow.

This problem can be avoided by using two-variable polynomials at the price we have slightly weaker pseudorandom measures than the optimal bounds.

Theorem (Gy.)

Let $G_{\leq K, Legendre, irreducible}$ denote the family of all binary lattices $\eta: \ I_p^2 \to \{-1,+1\}$ defined by

$$\eta((x_1,x_2)) = \left(\frac{f(x_1,x_2)}{p}\right)$$

where the used two variable irreducible polynomial $f \in \mathbb{F}_p[x_1,x_2]$ is not of the form $f(x_1,x_2) = \varphi(\alpha x_1 + \beta x_2)$ with degree $k \leq K$. Then

$$\Phi_\ell(G_{\leq K, Legendre, irreducible}) \ll \ell K p^{3/2} \log p.$$

Then we also have

Theorem (Gy.)

$$\Phi_\ell(\overline{E}(G_{\leq K, Legendre, irreducible})) \ll \ell K p^{3/2} \log p.$$

Is it difficult to generate two variable irreducible polynomials over $\mathbb{F}_p$?

It can be easy! For example, by the analogue of the Schönemann-Eisenstein criteria the polynomials

$$f(x_1, x_2) = x_1^k + x_1 x_2 g(x_1, x_2) + x_2 h(x_2)$$

(where $\deg g \leq k-5$ and $\deg h = k-2$, the coefficient of x^{k-3} in $h(x)$ is 0, $x \nmid h(x)$) are irreducible. Using only polynomials of this form in the definitions of the binary lattices

$$\eta((x_1, x_2)) = \left(\frac{f(x_1, x_2)}{p} \right)$$

we may define a family

$$\mathcal{G}_{\leq K, Sch.-Eis., Legendre} \quad (\subseteq G_{\leq K, Legendre, irreducible}).$$

Then by $\mathcal{G}_{\leq K, Sch.-Eis., Legendre} \subseteq G_{\leq K, Legendre, irreducible}$ we have

$\Phi_\ell(G_{\leq K, Sch.-Eis, irreducible}) \leq \Phi_\ell(G_{\leq K, Legendre, irreducible}) \ll \ell K p^{3/2} \log p$.

and thus

$$\Phi_\ell(\overline{E}(G_{\leq K, Sch.-Eis., irreducible})) \ll \ell K p^{3/2} \log p.$$

Thus the family $G_{\leq K, Sch.-Eis., Legendre}$ (based on polynomials of form $f(x_1, x_2) = x_1^k + x_1 x_2 g(x_1, x_2) + x_2 h(x_2))$ has nearly optimal cross-combined measure, clearly it is very large (it contains more than $p^{K(K-.1)/2}$ lattices).

The binary lattices from this family can be generated easily and fast. Clearly the same holds for the family of sequences $\overline{E}(G_{\leq K, Sch.-Eis., irreducible})$.

Thank you for your attention!