On difference matrices and generalised Rudin–Shapiro sequences

Thomas Stoll

(joint work with Irène Marcovici and Pierre-Adrien Tahay)

UDT 2021
24 February, 2021

Plan

1. Original problem: Golay – Rudin – Shapiro
2. Small-scale and large-scale correlations
3. Large-scale correlations on prime and squarefree alphabets
4. Difference matrices
5. Large-scale correlations on arbitrary alphabets
6. Small-scale correlations on arbitrary alphabets
7. Higher dimensions

Plan

1. Original problem: Golay – Rudin – Shapiro
2. Small-scale and large-scale correlations
3. Large-scale correlations on prime and squarefree alphabets
4. Difference matrices
5. Large-scale correlations on arbitrary alphabets
6. Small-scale correlations on arbitrary alphabets
7. Higher dimensions

Original problem

Given a sequence $\mathbf{a} = (a_n)$ with $a_n \in \{-1, +1\}$. Consider

$$S_N(\mathbf{a}) = \sup_{x \in [0,1]} \left| \sum_{0 \le n < N} a_n e^{2\pi i n x} \right|.$$

Original problem

Given a sequence $\mathbf{a} = (a_n)$ with $a_n \in \{-1, +1\}$. Consider

$$S_N(\mathbf{a}) = \sup_{x \in [0,1]} \left| \sum_{0 \le n < N} a_n e^{2\pi i n x} \right|.$$

- Parseval:

$$\sqrt{N} \ll S_N(\mathbf{a}) \ll N.$$

Original problem

Given a sequence $\mathbf{a} = (a_n)$ with $a_n \in \{-1, +1\}$. Consider

$$S_N(\mathbf{a}) = \sup_{x \in [0,1]} \left| \sum_{0 \le n < N} a_n e^{2\pi i n x} \right|.$$

- Parseval:
$$\sqrt{N} \ll S_N(\mathbf{a}) \ll N.$$

- For almost all $\mathbf{a}$:
$$S_N(\mathbf{a}) \ll \sqrt{N \log N}.$$

Original problem

Given a sequence $\mathbf{a} = (a_n)$ with $a_n \in \{-1, +1\}$. Consider

$$S_N(\mathbf{a}) = \sup_{x \in [0,1]} \left| \sum_{0 \le n < N} a_n e^{2\pi i n x} \right|.$$

- Parseval:
$$\sqrt{N} \ll S_N(\mathbf{a}) \ll N.$$

- For almost all $\mathbf{a}$:
$$S_N(\mathbf{a}) \ll \sqrt{N \log N}.$$

Problem

Construct a deterministic sequence $\mathbf{a}$ with the "root N" property.

Shapiro's approach

Let

$$P_1(z) = Q_1(z) = 1$$

and for $n \geq 1$,

$$P_{2^{n+1}}(z) = P_{2^n}(z) + z^{2^n} Q_{2^n}(z),$$
$$Q_{2^{n+1}}(z) = P_{2^n}(z) - z^{2^n} Q_{2^n}(z).$$

Shapiro's approach

Let

$$P_1(z) = Q_1(z) = 1$$

and for $n \geq 1$,

$$P_{2^{n+1}}(z) = P_{2^n}(z) + z^{2^n} Q_{2^n}(z),$$
$$Q_{2^{n+1}}(z) = P_{2^n}(z) - z^{2^n} Q_{2^n}(z).$$

By the parallelogram rule,

$$|P_{2^{n+1}}(z)|^2 + |Q_{2^{n+1}}(z)|^2 = 2\,(|P_{2^n}(z)|^2 + |Q_{2^n}(z)|^2),$$

and by induction,

$$|P_{2^{n+1}}(z)| \leq \sqrt{2} \cdot 2^{n/2}, \quad |Q_{2^{n+1}}(z)| \leq \sqrt{2} \cdot 2^{n/2} \qquad (|z| = 1).$$

(Golay–) Rudin–Shapiro

This leads to the definition of the (Golay–) Rudin–Shapiro sequence:

$$P_\infty(z) = \sum_{n\geq 0} r_n z^n.$$

(Golay–) Rudin–Shapiro sequence $\mathbf{r} = (r_n)$:

$$S_N(\mathbf{r}) = \sup_{x\in[0,1]} \left| \sum_{0\leq n<N} r_n e^{2\pi i n x} \right| \leq (2+\sqrt{2})\sqrt{N}.$$

(Golay–) Rudin–Shapiro

This leads to the definition of the (Golay–) Rudin–Shapiro sequence:

$$P_\infty(z) = \sum_{n\geq 0} r_n z^n.$$

(Golay–) Rudin–Shapiro sequence $\mathbf{r} = (r_n)$:

$$S_N(\mathbf{r}) = \sup_{x\in[0,1]} \left| \sum_{0\leq n<N} r_n e^{2\pi i n x} \right| \leq (2+\sqrt{2})\sqrt{N}.$$

- *Golay:* Statistic multislit spectrometry and its application to the panoramic display of infrared spectra, *J. Optical Soc. America* **41** (1951) 468–472.
- *Rudin:* Some theorems on Fourier coefficients, *Proc. Amer. Math. Soc.* **10** (1959), 855–859.
- *Shapiro:* Extremal Problems for Polynomials and Power Series, Thesis (M.S.), Massachusetts Institute of Technology, Department of Mathematics, 1951.

Many works: Allouche & Mendès France (1985), Allouche & Liardet (1991), Balister (2019+), Brillhart (1973), Brillhart & Carlitz (1970), Brillhart & Morton (1978), Brillhart & Erdős & Morton (1983), Doche & Habsieger (2004), Mauduit & Rivat (2015, 2018), Mendès France & Tenenbaum (1981), Montgomery (2017), Müllner (2018), Queffélec (1987), Rodgers (2017), Saffari (1986) etc.

Many works: Allouche & Mendès France (1985), Allouche & Liardet (1991), Balister (2019+), Brillhart (1973), Brillhart & Carlitz (1970), Brillhart & Morton (1978), Brillhart & Erdős & Morton (1983), Doche & Habsieger (2004), Mauduit & Rivat (2015, 2018), Mendès France & Tenenbaum (1981), Montgomery (2017), Müllner (2018), Queffélec (1987), Rodgers (2017), Saffari (1986) etc.

- **Recursion:**

$$\begin{aligned} r_0 &= 1, \\ r_{2n} &= r_n, && n \geq 1, \\ r_{2n+1} &= (-1)^n r_n, && n \geq 0. \end{aligned}$$

Many works: Allouche & Mendès France (1985), Allouche & Liardet (1991), Balister (2019+), Brillhart (1973), Brillhart & Carlitz (1970), Brillhart & Morton (1978), Brillhart & Erdős & Morton (1983), Doche & Habsieger (2004), Mauduit & Rivat (2015, 2018), Mendès France & Tenenbaum (1981), Montgomery (2017), Müllner (2018), Queffélec (1987), Rodgers (2017), Saffari (1986) etc.

- **Recursion:**

$$\begin{aligned} r_0 &= 1, \\ r_{2n} &= r_n, && n \geq 1, \\ r_{2n+1} &= (-1)^n r_n, && n \geq 0. \end{aligned}$$

- **Binary digital expansion:**

$$r_n = (-1)^{\#e_{11}(n)},$$

where $e_{11}(n)$ denotes the number of (possibly overlapping) blocks "11" in the base 2 expansion of n.

Many works: Allouche & Mendès France (1985), Allouche & Liardet (1991), Balister (2019+), Brillhart (1973), Brillhart & Carlitz (1970), Brillhart & Morton (1978), Brillhart & Erdős & Morton (1983), Doche & Habsieger (2004), Mauduit & Rivat (2015, 2018), Mendès France & Tenenbaum (1981), Montgomery (2017), Müllner (2018), Queffélec (1987), Rodgers (2017), Saffari (1986) etc.

- **Recursion:**

$$\begin{aligned} r_0 &= 1, \\ r_{2n} &= r_n, && n \geq 1, \\ r_{2n+1} &= (-1)^n r_n, && n \geq 0. \end{aligned}$$

- **Binary digital expansion:**

$$r_n = (-1)^{\#e_{11}(n)},$$

where $e_{11}(n)$ denotes the number of (possibly overlapping) blocks "11" in the base 2 expansion of n.

$$187 = (10111011)_2, \quad \#e_{11}(n) = 3, \quad r_{187} = -1.$$

Plan

1. Original problem: Golay – Rudin – Shapiro
2. Small-scale and large-scale correlations
3. Large-scale correlations on prime and squarefree alphabets
4. Difference matrices
5. Large-scale correlations on arbitrary alphabets
6. Small-scale correlations on arbitrary alphabets
7. Higher dimensions

Theorem (Mauduit-Sárközy (1998))

$$\left| \sum_{n \leq N} r_n r_{n+d} \right| \leq 2d + \frac{4d}{\log 2} \log \frac{2N}{d}.$$

Theorem (Mauduit-Sárközy (1998))

$$\left| \sum_{n \le N} r_n r_{n+d} \right| \le 2d + \frac{4d}{\log 2} \log \frac{2N}{d}.$$

For $d = o(N)$, still

$$\sum_{n \le N} r_n r_{n+d} = o(N).$$

Theorem (Mauduit-Sárközy (1998))

$$\left|\sum_{n\le N} r_n r_{n+d}\right| \le 2d + \frac{4d}{\log 2}\log\frac{2N}{d}.$$

For $d = o(N)$, still

$$\sum_{n\le N} r_n r_{n+d} = o(N).$$

Question

Find a *purely combinatorial* extension to general alphabets.

Generalisations: Allouche & Bousquet-Mélou (1994), Rider (1966), M. Queffélec (1987), Allouche & Liardet (1991) etc.

Let $\mathbf{a} = a_0, a_1, \ldots$ be an infinite sequence over $\{0, 1, \ldots, k-1\}$. For an integral vector $(i_1, i_2, \ldots, i_m)$ with $0 \leq i_1 < i_2 < \cdots < i_m$ define the discrete correlation coefficient of order m by

$$\delta(i_1, i_2, \ldots, i_m) = \begin{cases} 0, & \text{if } a_{i_1} = a_{i_2} = \cdots = a_{i_m}; \\ 1, & \text{otherwise.} \end{cases}$$

Let $\mathbf{a} = a_0, a_1, \ldots$ be an infinite sequence over $\{0, 1, \ldots, k-1\}$. For an integral vector $(i_1, i_2, \ldots, i_m)$ with $0 \leq i_1 < i_2 < \cdots < i_m$ define the discrete correlation coefficient of order m by

$$\delta(i_1, i_2, \ldots, i_m) = \begin{cases} 0, & \text{if } a_{i_1} = a_{i_2} = \cdots = a_{i_m}; \\ 1, & \text{otherwise.} \end{cases}$$

Theorem (Grant, Shallit, S. (2009))

We have

$$\liminf_{d\to\infty} \left(\liminf_{N\to\infty} \frac{1}{N} \sum_{n<N} \delta(n, n+d) \right) \leq 1 - \frac{1}{k}.$$

For $\mathbf{d} = (d_1, d_2, \ldots, d_m)$ with $0 \leq d_1 < d_2 < \cdots < d_m$ define

$$C_{\mathbf{d}} = \liminf_{N\to\infty} \frac{1}{N} \sum_{n<N} \delta(n + d_1, n + d_2, \ldots, n + d_m).$$

The vector $\mathbf{d}$ is *normalized*, if $d_1 = 0$. Let $\|\cdot\|$ be a norm on the finite dimensional vector space $\mathbb{R}^m$.

For $\mathbf{d} = (d_1, d_2, \ldots, d_m)$ with $0 \leq d_1 < d_2 < \cdots < d_m$ define

$$C_{\mathbf{d}} = \liminf_{N\to\infty} \frac{1}{N} \sum_{n<N} \delta(n + d_1, n + d_2, \ldots, n + d_m).$$

The vector $\mathbf{d}$ is *normalized*, if $d_1 = 0$. Let $\|\cdot\|$ be a norm on the finite dimensional vector space $\mathbb{R}^m$.

Theorem (Grant, Shallit, S. (2009))

For any $m \geq 2$ and any norm $\|\cdot\|$, we have

$$\lim_{\lambda\to\infty} \left(\inf\{C_{\mathbf{d}} : \mathbf{d} \in \mathbb{N}^m,\ \mathbf{d} \textit{ normalized},\ \|\mathbf{d}\| \geq \lambda\}\right) \leq 1 - \frac{1}{k^{m-1}}.$$

Plan

1. Original problem: Golay – Rudin – Shapiro
2. Small-scale and large-scale correlations
3. Large-scale correlations on prime and squarefree alphabets
4. Difference matrices
5. Large-scale correlations on arbitrary alphabets
6. Small-scale correlations on arbitrary alphabets
7. Higher dimensions

Definition: Let

$$f : \{0, 1, \ldots, k-1\} \times \mathbb{Z} \to \mathbb{Z}, \qquad (j, n) \mapsto f(j, n)$$

be a function which is periodic in n with period k and such that for all $u, i \in \mathbb{Z}$ with $0 \leq u < u + i \leq k - 1$ we have

$$\{(f(u+i, n) - f(u, n)) \mod k : 0 \leq n \leq k-1\} = \{ 0, 1, \ldots, k-1 \}.$$

Definition: Let

$$f : \{0, 1, \ldots, k-1\} \times \mathbb{Z} \to \mathbb{Z}, \qquad (j, n) \mapsto f(j, n)$$

be a function which is periodic in n with period k and such that for all $u, i \in \mathbb{Z}$ with $0 \leq u < u + i \leq k - 1$ we have

$$\{(f(u+i, n) - f(u, n)) \mod k : 0 \leq n \leq k-1\} = \{ 0, 1, \ldots, k-1 \}.$$

Then $\mathbf{u} = (u_n)$ over the alphabet $\{0, 1, \ldots, k-1\}$ is a generalised Rudin-Shapiro sequence if there exists $(a(n))$ such that $u_n = a(n) \mod k$ and

$$a(nk + j) = a(n) + f(j, n), \qquad 0 \leq j \leq k-1, \quad n \geq 1.$$

(1) **M. Queffélec (1987):** Let

$$f(j,n) = j \cdot (n \bmod k).$$

Then

$$f(u+i,n) - f(u,n) \equiv in \pmod{k},$$

and $\{in : \ 0 \le n \le k-1\}$ runs through all residue classes **provided** k **is prime.**

In particular, for $k = 2$ and

$$f(j,n) = \begin{cases} 1, & \text{if } j = 1,\ n \equiv 1 \pmod 2; \\ 0, & \text{otherwise} \end{cases}$$

we get the the Rudin-Shapiro sequence over the alphabet $\{0,1\}$.

(2) For $k = 2$ and appropriate initial conditions, we get sequences which count any fixed block of size two. For example,

$$f(1,0) = 1, \qquad f(0,0) = f(1,1) = f(0,1) = 0,$$

then $(a(n))$ counts the number of subblocks (01) in the binary expansion of n.

(2) For $k = 2$ and appropriate initial conditions, we get sequences which count any fixed block of size two. For example,

$$f(1,0) = 1, \qquad f(0,0) = f(1,1) = f(0,1) = 0,$$

then $(a(n))$ counts the number of subblocks (01) in the binary expansion of n.

(3) For $k = 3$ set

$$f(j,n) = \begin{cases} 1, & \text{if } j \equiv n \pmod 3; \\ 0, & \text{otherwise.} \end{cases}$$

Here, the sequence $(a(n))$ counts the number of subblocks (00), (11) and (22) in the ternary expansion of integers.

Theorem (Grant, Shallit, S. (2009))

Let $\mathbf{u} = u_0, u_1, \ldots$ *be a generalised Rudin-Shapiro sequence over* $\{0, 1, \ldots, k-1\}$ ***with*** k ***prime****. Moreover, let* $0 \leq d_1 < d_2$*. Then, as* $N \to \infty$*, we have*

$$\sum_{n<N} \delta(n+d_1, n+d_2) = \left(1 - \frac{1}{k}\right) N + O_k\left((d_2 - d_1) \log \frac{N}{d_2 - d_1} + d_2\right).$$

Theorem (Grant, Shallit, S. (2009))

Let $\mathbf{u} = u_0, u_1, \ldots$ *be a generalised Rudin-Shapiro sequence over* $\{0, 1, \ldots, k-1\}$ ***with*** k ***prime****. Moreover, let* $0 \leq d_1 < d_2$*. Then, as* $N \to \infty$*, we have*

$$\sum_{n<N} \delta(n+d_1, n+d_2) = \left(1 - \frac{1}{k}\right) N + O_k\left((d_2 - d_1) \log \frac{N}{d_2 - d_1} + d_2\right).$$

If $d_2 = o(N)$ then

$$\sum_{n<N} \delta(n+d_1, n+d_2) \sim \left(1 - \frac{1}{k}\right) N.$$

Extension to k squarefree:

Let $k = p_1 p_2 \cdots p_\ell$ be a product of pairwise distinct primes. Define $\mathbf{u} = (u_n)$ by

$$u_n = a(n) \bmod k,$$

where $a(n)$ is defined by

$$a(n) = a_1(n) + p_1\, a_2(n) + p_1 p_2\, a_3(n) + \cdots + p_1 p_2 \cdots p_{\ell-1}\, a_\ell(n).$$

Herein, $(a_i(n))$ satisfies the recursive relation

$$a_i(p_i n + j) = a_i(n) + g_i(j, n), \qquad 1 \le i \le \ell,$$

for $n \ge 1$ and $0 \le j \le p_i - 1$ (the functions g_i are admissible functions in the previous sense).

Theorem (Grant, Shallit, S. (2009))

Let $k = p_1 p_2 \cdots p_\ell$ ***be squarefree*** *($\ell \geq 2$) and denote by $\mathbf{u} = u_0, u_1, \ldots$ a generalised Rudin-Shapiro sequence over $\{0, 1, \ldots, k-1\}$. Moreover, let $0 \leq d_1 < d_2$ and $0 < \gamma < 1$. Then, as $N \to \infty$, we have*

$$\sum_{n<N} \delta(n + d_1, n + d_2)$$

$$= \left(1 - \frac{1}{k}\right) N + O_k\Big((d_2 - d_1) N^{1-\gamma/\ell} + (d_2 - d_1) N^{1-\gamma} \log \frac{N^{\gamma/\ell}}{d_2 - d_1} + N^\gamma + d_1\Big).$$

Theorem (Grant, Shallit, S. (2009))

Let $k = p_1 p_2 \cdots p_\ell$ ***be squarefree*** *($\ell \geq 2$) and denote by $\mathbf{u} = u_0, u_1, \ldots$ a generalised Rudin-Shapiro sequence over $\{0, 1, \ldots, k-1\}$. Moreover, let $0 \leq d_1 < d_2$ and $0 < \gamma < 1$. Then, as $N \to \infty$, we have*

$$\sum_{n<N} \delta(n + d_1, n + d_2)$$
$$= \left(1 - \frac{1}{k}\right) N + O_k\Big((d_2 - d_1)N^{1-\gamma/\ell} + (d_2 - d_1)N^{1-\gamma} \log \frac{N^{\gamma/\ell}}{d_2 - d_1} + N^\gamma + d_1\Big).$$

If $d_2 = o(N^{\gamma/\ell})$ then

$$\sum_{n<N} \delta(n + d_1, n + d_2) \sim \left(1 - \frac{1}{k}\right) N.$$

Plan

1. Original problem: Golay – Rudin – Shapiro
2. Small-scale and large-scale correlations
3. Large-scale correlations on prime and squarefree alphabets
4. Difference matrices
5. Large-scale correlations on arbitrary alphabets
6. Small-scale correlations on arbitrary alphabets
7. Higher dimensions

Let $k \geq 1$ be an integer, set $\Sigma_k = \{0, \ldots, k-1\}$. Let G be a finite abelian group (e.g., $G = \mathbb{Z}_k$).

Let $k \geq 1$ be an integer, set $\Sigma_k = \{0, \ldots, k-1\}$. Let G be a finite abelian group (e.g., $G = \mathbb{Z}_k$).

Definition

Let $f : \Sigma_k \times \Sigma_k \to G$ be such that $f(0,0) = 0$.

The sequence $\mathbf{u} = (u_n) \in G^{\mathbb{N}}$ is called **block-additive** in base k with **weight function** f if for $n = \sum_{i \geq 0} x_i k^i$, we have

$$u_n = \sum_{i \geq 0} f(x_i, x_{i+1}).$$

Let $k \geq 1$ be an integer, set $\Sigma_k = \{0, \ldots, k-1\}$. Let G be a finite abelian group (e.g., $G = \mathbb{Z}_k$).

Definition

Let $f : \Sigma_k \times \Sigma_k \to G$ be such that $f(0,0) = 0$.

The sequence $\mathbf{u} = (u_n) \in G^{\mathbb{N}}$ is called **block-additive** in base k with **weight function** f if for $n = \sum_{i \geq 0} x_i k^i$, we have

$$u_n = \sum_{i \geq 0} f(x_i, x_{i+1}).$$

- Each block-additive function is k-automatic.

Let $k \geq 1$ be an integer, set $\Sigma_k = \{0, \ldots, k-1\}$. Let G be a finite abelian group (e.g., $G = \mathbb{Z}_k$).

Definition

Let $f : \Sigma_k \times \Sigma_k \to G$ be such that $f(0,0) = 0$.

The sequence $\mathbf{u} = (u_n) \in G^{\mathbb{N}}$ is called **block-additive** in base k with **weight function** f if for $n = \sum_{i \geq 0} x_i k^i$, we have

$$u_n = \sum_{i \geq 0} f(x_i, x_{i+1}).$$

- Each block-additive function is k-automatic.
- The (Golay–) Rudin–Shapiro sequence is block-additive.

Difference condition

f fulfills the difference condition if for all $i \neq j$, and for all $g \in G$,

$$\operatorname{card}\left\{h \in \Sigma_k : f(i,h) - f(j,h) = g\right\} = \frac{k}{|G|}.$$

Difference condition

f fulfills the difference condition if for all $i \neq j$, and for all $g \in G$,

$$\operatorname{card}\left\{h \in \Sigma_k : f(i,h) - f(j,h) = g\right\} = \frac{k}{|G|}.$$

The matrix of the weights is a **difference matrix**: the difference of any 2 lines in the matrix contains each element equally often.

Difference condition

f fulfills the difference condition if for all $i \neq j$, and for all $g \in G$,

$$\operatorname{card}\left\{h \in \Sigma_k : f(i,h) - f(j,h) = g\right\} = \frac{k}{|G|}.$$

The matrix of the weights is a **difference matrix**: the difference of any 2 lines in the matrix contains each element equally often.

Example: Generalised Rudin-Shapiro for $p = 2$; $p = 3$; $p = 5$

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

Difference condition

f fulfills the difference condition if for all $i \neq j$, and for all $g \in G$,

$$\operatorname{card}\left\{h \in \Sigma_k : f(i,h) - f(j,h) = g\right\} = \frac{k}{|G|}.$$

The matrix of the weights is a **difference matrix**: the difference of any 2 lines in the matrix contains each element equally often.

Example: Generalised Rudin-Shapiro for $p = 2$; $p = 3$; $p = 5$

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 1 & 3 \\ 0 & 3 & 1 & 4 & 2 \\ 0 & 4 & 3 & 2 & 1 \end{pmatrix}$$

The set of difference matrices

- D. Jungnickel, On difference matrices, resolvable transversal designs and generalized Hadamard matrices, Math. Z. (1979).
- A. S. Hedayat, N. J. A. Sloane, J. Stufken, *Orthogonal arrays*, Springer Series in Statistics, Springer-Verlag, 1999.
- P. H. J. Lampio, *Classification of difference matrices and complex Hadamard matrices,* PhD thesis, Aalto University, 2015.

The set of difference matrices

- D. Jungnickel, On difference matrices, resolvable transversal designs and generalized Hadamard matrices, Math. Z. (1979).
- A. S. Hedayat, N. J. A. Sloane, J. Stufken, *Orthogonal arrays*, Springer Series in Statistics, Springer-Verlag, 1999.
- P. H. J. Lampio, *Classification of difference matrices and complex Hadamard matrices,* PhD thesis, Aalto University, 2015.

Let $D(r, c, G)$ denote the set of all difference matrices of size $r \times c$ with entries in the group G.

The set of difference matrices

- D. Jungnickel, On difference matrices, resolvable transversal designs and generalized Hadamard matrices, Math. Z. (1979).
- A. S. Hedayat, N. J. A. Sloane, J. Stufken, *Orthogonal arrays*, Springer Series in Statistics, Springer-Verlag, 1999.
- P. H. J. Lampio, *Classification of difference matrices and complex Hadamard matrices,* PhD thesis, Aalto University, 2015.

Let $D(r, c, G)$ denote the set of all difference matrices of size $r \times c$ with entries in the group G.

- For an even integer $k \geq 4$, the set $D(k, k, \mathbb{Z}_k)$ is empty.

The set of difference matrices

- D. Jungnickel, On difference matrices, resolvable transversal designs and generalized Hadamard matrices, Math. Z. (1979).
- A. S. Hedayat, N. J. A. Sloane, J. Stufken, *Orthogonal arrays*, Springer Series in Statistics, Springer-Verlag, 1999.
- P. H. J. Lampio, *Classification of difference matrices and complex Hadamard matrices,* PhD thesis, Aalto University, 2015.

Let $D(r, c, G)$ denote the set of all difference matrices of size $r \times c$ with entries in the group G.

- For an even integer $k \geq 4$, the set $D(k, k, \mathbb{Z}_k)$ is empty.
- The set $D(4, 4, \mathbb{Z}_2 \times \mathbb{Z}_2)$ is non-empty.

$$M = \begin{pmatrix} (0,0) & (0,0) & (0,0) & (0,0) \\ (0,0) & (0,1) & (1,0) & (1,1) \\ (0,0) & (1,0) & (1,1) & (0,1) \\ (0,0) & (1,1) & (0,1) & (1,0) \end{pmatrix}$$

is an element of this set. (Hedayat et al. (1999))

Proposition (Hedayat et al. (1999))

For any prime p and any integers k, n with $k \geq n \geq 1$, there exists an abelian group G of order p^n such that the set $D(p^k, p^k, G)$ is non-empty.

- Present the elements of $\mathbb{F}_{p^k}$ as

$$\beta_0 + \beta_1 x + \cdots + \beta_{n-1} x^{n-1} + \cdots + \beta_{k-1} x^{k-1}, \qquad \beta_0, \ldots, \beta_{k-1} \in \mathbb{Z}_p.$$

- Regard $\mathbb{F}_{p^n}$ as an additive subgroup of $\mathbb{F}_{p^k}$ with elements

$$\beta_0 + \beta_1 x + \cdots + \beta_{n-1} x^{n-1}.$$

- Let D^* be the multiplication table of $\mathbb{F}_{p^k}$ and let $\phi : \mathbb{F}_{p^k} \to \mathbb{F}_{p^n}$ be the map

$$\beta_0 + \beta_1 x + \cdots + \beta_{k-1} x^{k-1} \quad \mapsto \quad \beta_0 + \beta_1 x + \cdots + \beta_{n-1} x^{n-1}.$$

- Apply ϕ to each element of the table D^*, then D (the new table) is a difference matrix of $D(p^k, p^k, \mathbb{F}_{p^n})$.

In $D(9, 9, \mathbb{Z}_3)$, there are two *equivalence classes* of difference matrices. A representative of each equivalence class is:

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\
0 & 0 & 0 & 2 & 2 & 2 & 1 & 1 & 1 \\
0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\
0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \\
0 & 1 & 2 & 2 & 0 & 1 & 1 & 2 & 0 \\
0 & 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 \\
0 & 2 & 1 & 1 & 0 & 2 & 2 & 1 & 0 \\
0 & 2 & 1 & 2 & 1 & 0 & 1 & 0 & 2
\end{pmatrix}, \qquad
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\
0 & 0 & 0 & 2 & 2 & 2 & 1 & 1 & 1 \\
0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\
0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \\
0 & 1 & 2 & 2 & 0 & 1 & 1 & 2 & 0 \\
0 & 2 & 1 & 0 & 2 & 1 & 1 & 0 & 2 \\
0 & 2 & 1 & 1 & 0 & 2 & 0 & 2 & 1 \\
0 & 2 & 1 & 2 & 1 & 0 & 2 & 1 & 0
\end{pmatrix}$$

Lampio / Östergård (2011).

In $D(9, 15, \mathbb{Z}_3)$, there are 5 classes, e.g.

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 2 & 2 & 2 & 1 & 1 & 1 & 2 & 2 \\
0 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 0 & 2 & 1 & 2 & 2 & 1 & 1 \\
0 & 0 & 1 & 2 & 2 & 0 & 2 & 1 & 2 & 1 & 2 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 2 & 2 & 2 & 0 & 2 & 1 & 1 & 1 & 0 & 2 & 1 & 0 \\
0 & 1 & 1 & 2 & 0 & 2 & 2 & 0 & 1 & 0 & 0 & 1 & 1 & 2 & 2 \\
0 & 1 & 2 & 2 & 1 & 0 & 1 & 2 & 0 & 1 & 0 & 2 & 0 & 2 & 1 \\
0 & 2 & 1 & 1 & 2 & 1 & 0 & 0 & 2 & 1 & 0 & 2 & 0 & 1 & 2
\end{pmatrix},$$

Lampio / Östergård (2011).

In $D(9, 15, \mathbb{Z}_3)$, there are 5 classes, e.g.

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 2 & 2 & 2 & 1 & 1 & 1 & 2 & 2 \\
0 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 0 & 2 & 1 & 2 & 2 & 1 & 1 \\
0 & 0 & 1 & 2 & 2 & 0 & 2 & 1 & 2 & 1 & 2 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 2 & 2 & 2 & 0 & 2 & 1 & 1 & 1 & 0 & 2 & 1 & 0 \\
0 & 1 & 1 & 2 & 0 & 2 & 2 & 0 & 1 & 0 & 0 & 1 & 1 & 2 & 2 \\
0 & 1 & 2 & 2 & 1 & 0 & 1 & 2 & 0 & 1 & 0 & 2 & 0 & 2 & 1 \\
0 & 2 & 1 & 1 & 2 & 1 & 0 & 0 & 2 & 1 & 0 & 2 & 0 & 1 & 2
\end{pmatrix},$$

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 2 & 2 & 2 & 1 & 1 & 1 & 2 & 2 \\
0 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 0 & 2 & 1 & 2 & 2 & 1 & 1 \\
0 & 0 & 1 & 2 & 2 & 0 & 2 & 1 & 2 & 1 & 2 & 0 & 1 & 0 & 1 \\
0 & 0 & 2 & 2 & 1 & 1 & 2 & 2 & 1 & 0 & 2 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 2 & 2 & 2 & 1 & 2 & 0 & 1 & 1 & 0 & 0 & 1 & 2 \\
0 & 1 & 2 & 2 & 0 & 2 & 0 & 1 & 1 & 2 & 1 & 2 & 1 & 0 & 0 \\
0 & 2 & 1 & 1 & 2 & 0 & 0 & 2 & 1 & 1 & 0 & 2 & 2 & 1 & 0
\end{pmatrix}, \quad \text{etc.}$$

Plan

1. Original problem: Golay – Rudin – Shapiro
2. Small-scale and large-scale correlations
3. Large-scale correlations on prime and squarefree alphabets
4. Difference matrices
5. Large-scale correlations on arbitrary alphabets
6. Small-scale correlations on arbitrary alphabets
7. Higher dimensions

Theorem (Tahay, 2020)

Let p be a prime number and $\ell \geq 1$. Let M be a difference matrix in $D(p^\ell, p^\ell, \mathbb{Z}_p^\ell)$ and let $\mathbf{u} = (u_n)$ be the Rudin–Shapiro sequence associated to M. Moreover, let $0 \leq d_1 < d_2$. Then, as $N \to \infty$, we have

$$\sum_{n<N} \delta(n+d_1, n+d_2) = N\left(1 - \frac{1}{p^\ell}\right) + O_{p,\ell}\left((d_2 - d_1)\log\frac{N}{d_2 - d_1} + d_2\right).$$

Theorem (Tahay, 2020)

Let p be a prime number and $\ell \geq 1$. Let M be a difference matrix in $D(p^\ell, p^\ell, \mathbb{Z}_p^\ell)$ and let $\mathbf{u} = (u_n)$ be the Rudin–Shapiro sequence associated to M. Moreover, let $0 \leq d_1 < d_2$. Then, as $N \to \infty$, we have

$$\sum_{n<N} \delta(n+d_1, n+d_2) = N\left(1 - \frac{1}{p^\ell}\right) + O_{p,\ell}\left((d_2 - d_1)\log\frac{N}{d_2 - d_1} + d_2\right).$$

Example: Consider $D(4, 4, \mathbb{Z}_2 \times \mathbb{Z}_2)$ and

$$M = \begin{pmatrix} (0,0) & (0,0) & (0,0) & (0,0) \\ (0,0) & (0,1) & (1,0) & (1,1) \\ (0,0) & (1,0) & (1,1) & (0,1) \\ (0,0) & (1,1) & (0,1) & (1,0) \end{pmatrix}.$$

Theorem (Tahay, 2020)

Let p be a prime number and $\ell \geq 1$. Let M be a difference matrix in $D(p^\ell, p^\ell, \mathbb{Z}_p^\ell)$ and let $\mathbf{u} = (u_n)$ be the Rudin–Shapiro sequence associated to M. Moreover, let $0 \leq d_1 < d_2$. Then, as $N \to \infty$, we have

$$\sum_{n<N} \delta(n+d_1, n+d_2) = N\left(1 - \frac{1}{p^\ell}\right) + O_{p,\ell}\left((d_2 - d_1)\log\frac{N}{d_2 - d_1} + d_2\right).$$

Example: Consider $D(4, 4, \mathbb{Z}_2 \times \mathbb{Z}_2)$ and

$$M = \begin{pmatrix} (0,0) & (0,0) & (0,0) & (0,0) \\ (0,0) & (0,1) & (1,0) & (1,1) \\ (0,0) & (1,0) & (1,1) & (0,1) \\ (0,0) & (1,1) & (0,1) & (1,0) \end{pmatrix}.$$

After recoding $(0,0) \mapsto 0$, $(0,1) \mapsto 1$, $(1,0) \mapsto 2$ and $(1,1) \mapsto 3$, we get

$$\mathbf{u} = 0, 0, 0, 0, 0, 1, 2, 3, 0, 2, 3, 1, 0, 3, 1, 2, 0, 0, 0, 0, 1, 0, 3, 2, 2, 0, 1, 3, \ldots$$

and

$$\sum_{n<N} \delta(n+d_1, n+d_2) = \frac{3}{4}N + O\left((d_2 - d_1)\log\frac{N}{d_2 - d_1} + d_2\right).$$

Theorem (Tahay, 2020)

Let $\ell \geq 2$ and $k = p_1^{k_1} \cdots p_\ell^{k_\ell}$. For each i consider a difference matrix M_i of $D(p_i^{k_i}, p_i^{k_i}, \mathbb{Z}_{p_i}^{k_i})$. We associate a function

$$g^i(j,n) = (g_1^i(j,n), \ldots, g_{k_i}^i(j,n))$$

and a sequence $a^i(n) = (a_1^i(n), \ldots, a_{k_i}^i(n))$ as before. Let

$$\hat{a}(n) = (a^1(n) \bmod p_1, \ldots, a^\ell(n) \bmod p_\ell).$$

Then, as $N \to \infty$, we have

$$\sum_{n<N} \delta(n+d_1, n+d_2) = N\left(1 - \frac{1}{k}\right) + O_k\left(\left((d_2-d_1)\log\frac{N^{\frac{1}{\ell}}}{d_2-d_1} + d_2\right) N^{\frac{\ell-1}{\ell}}\right).$$

If $d_2 = o(N^{\frac{1}{\ell}})$ then

$$\sum_{n<N} \delta(n+d_1, n+d_2) \sim N\left(1 - \frac{1}{k}\right).$$

Plan

1. Original problem: Golay – Rudin – Shapiro
2. Small-scale and large-scale correlations
3. Large-scale correlations on prime and squarefree alphabets
4. Difference matrices
5. Large-scale correlations on arbitrary alphabets
6. Small-scale correlations on arbitrary alphabets
7. Higher dimensions

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

Question: Given a group G, a difference matrix M and the associated generalised Rudin–Shapiro sequence. Let d be **fixed**.

For all fixed $(i,j) \in G^2$, as $N \to \infty$, do we have

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{N}{|G|^2} + O(\log N)?$$

Example: generalised Rudin-Shapiro with $p = 3$

0 0 0 0 1 2 0 2 1 0 0 0 1 2 0 2 1 0 0 0 0 2 0 1 1 0 2 0 0 0 0 1 2 0 2 1 1 1 1
2 0 1 0 2 1 2 2 2 1 2 0 0 2 1 0 0 0 0 1 2...

We fix the distance d, and write n and $n+d$ in base k.

Denote by c_n the index of the most significant digit where two numbers differ (we write the expansion from left to the right, in the *unusual* way).

$$\begin{aligned} [n]_k &= x_0 \quad x_1 \quad \cdots \quad x_{c_n} \quad x_{c_n+1} \quad x_{c_n+2} \quad \cdots \\ [n+d]_k &= y_0 \quad y_1 \quad \cdots \quad y_{c_n} \quad x_{c_n+1} \quad x_{c_n+2} \quad \cdots \end{aligned}$$

We fix the distance d, and write n and $n+d$ in base k.

Denote by c_n the index of the most significant digit where two numbers differ (we write the expansion from left to the right, in the *unusual* way).

$$\begin{aligned} [n]_k &= x_0 \quad x_1 \quad \cdots \quad x_{c_n} \quad x_{c_n+1} \quad x_{c_n+2} \quad \cdots \\ [n+d]_k &= y_0 \quad y_1 \quad \cdots \quad y_{c_n} \quad x_{c_n+1} \quad x_{c_n+2} \quad \cdots \end{aligned}$$

We call **fibre** of n the set of the k elements:

$$\begin{aligned} \mathcal{F}(n) = \{ & x_0 \quad x_1 \quad \cdots \quad x_{c_n} \quad 0 \quad x_{c_n+2} \quad x_{c_n+3} \quad \cdots , \\ & x_0 \quad x_1 \quad \cdots \quad x_{c_n} \quad 1 \quad x_{c_n+2} \quad x_{c_n+3} \quad \cdots , \\ & x_0 \quad x_1 \quad \cdots \quad x_{c_n} \quad 2 \quad x_{c_n+2} \quad x_{c_n+3} \quad \cdots , \\ & \qquad\qquad\qquad\qquad \vdots \\ & x_0 \quad x_1 \quad \cdots \quad x_{c_n} \quad k-1 \quad x_{c_n+2} \quad x_{c_n+3} \quad \cdots \}. \end{aligned}$$

We consider n_i, the ith element of the fibre,

$$\begin{aligned} [n_i]_k &= x_0 \quad x_1 \quad \cdots \quad x_{c_n} \quad i \quad x_{c_n+2} \quad x_{c_n+3} \quad \cdots \\ [n_i + d]_k &= y_0 \quad y_1 \quad \cdots \quad y_{c_n} \quad i \quad x_{c_n+2} \quad x_{c_n+3} \quad \cdots \end{aligned}$$

We consider n_i, the ith element of the fibre,

$$\begin{aligned}
[n_i]_k &= x_0 \quad x_1 \quad \cdots \quad x_{c_n} \quad i \quad x_{c_n+2} \quad x_{c_n+3} \quad \cdots \\
[n_i + d]_k &= y_0 \quad y_1 \quad \cdots \quad y_{c_n} \quad i \quad x_{c_n+2} \quad x_{c_n+3} \quad \cdots
\end{aligned}$$

We have, when i varies,

$$\begin{aligned}
u_{n_i} &= C_1 + f(x_{c_n}, i) + f(i, x_{c_n+2}) \\
u_{n_i+d} &= C_2 + f(y_{c_n}, i) + f(i, x_{c_n+2}).
\end{aligned}$$

We consider n_i, the ith element of the fibre,

$$\begin{array}{rccccccccc} [n_i]_k = & x_0 & x_1 & \cdots & x_{c_n} & i & x_{c_n+2} & x_{c_n+3} & \cdots \\ [n_i+d]_k = & y_0 & y_1 & \cdots & y_{c_n} & i & x_{c_n+2} & x_{c_n+3} & \cdots \end{array}$$

We have, when i varies,

$$\begin{aligned} u_{n_i} &= C_1 + f(x_{c_n}, i) + f(i, x_{c_n+2}) \\ u_{n_i+d} &= C_2 + f(y_{c_n}, i) + f(i, x_{c_n+2}). \end{aligned}$$

Therefore

$$u_{n_i+d} - u_{n_i} = f(y_{c_n}, i) - f(x_{c_n}, i) + (C_2 - C_1).$$

We consider n_i, the ith element of the fibre,

$$\begin{array}{rccccccccc} [n_i]_k = & x_0 & x_1 & \cdots & x_{c_n} & i & x_{c_n+2} & x_{c_n+3} & \cdots \\ [n_i+d]_k = & y_0 & y_1 & \cdots & y_{c_n} & i & x_{c_n+2} & x_{c_n+3} & \cdots \end{array}$$

We have, when i varies,

$$\begin{aligned} u_{n_i} &= C_1 + f(x_{c_n}, i) + f(i, x_{c_n+2}) \\ u_{n_i+d} &= C_2 + f(y_{c_n}, i) + f(i, x_{c_n+2}). \end{aligned}$$

Therefore

$$u_{n_i+d} - u_{n_i} = f(y_{c_n}, i) - f(x_{c_n}, i) + (C_2 - C_1).$$

By the difference condition, when i varies, $u_{n_i+d} - u_{n_i}$ hits each element of G equally often.

We consider n_i, the ith element of the fibre,

$$\begin{array}{rccccccc} [n_i]_k = & x_0 & x_1 & \cdots & x_{c_n} & i & x_{c_n+2} & x_{c_n+3} & \cdots \\ [n_i+d]_k = & y_0 & y_1 & \cdots & y_{c_n} & i & x_{c_n+2} & x_{c_n+3} & \cdots \end{array}$$

We have, when i varies,

$$\begin{aligned} u_{n_i} &= C_1 + f(x_{c_n}, i) + f(i, x_{c_n+2}) \\ u_{n_i+d} &= C_2 + f(y_{c_n}, i) + f(i, x_{c_n+2}). \end{aligned}$$

Therefore

$$u_{n_i+d} - u_{n_i} = f(y_{c_n}, i) - f(x_{c_n}, i) + (C_2 - C_1).$$

By the difference condition, when i varies, $u_{n_i+d} - u_{n_i}$ hits each element of G equally often.

This means, in a fibre $\mathcal{F}(n)$ we have $u_{n+d} = u_n + g$ for each fixed $g \in G$ equally often.

If we fix N large with respect to d, then for most of the integers $n \in [\![0, N-1]\!]$, we have $\mathcal{F}(n) \subset [\![0, N-1]\!]$.

$$[N]_k = a_0\ a_1\ \cdots\ a_{\ell_N-1}\ a_{\ell_N}\ 0\ 0\ \cdots$$

If we fix N large with respect to d, then for most of the integers $n \in [\![0, N-1]\!]$, we have $\mathcal{F}(n) \subset [\![0, N-1]\!]$.

$$[N]_k = a_0\ a_1\ \cdots\ a_{\ell_N-1}\ a_{\ell_N}\ 0\ 0\ \cdots$$

$$[n]_k = \underbrace{x_0\ x_1\ \cdots\ x_{\ell_N-2}}_{\gamma < k^{\ell_N-1}-d}\ \alpha\ \underbrace{a'_{\ell_N}}_{<a_{\ell_N}}\ 0\ 0\ \cdots$$

$$[n+d]_k = x'_0\ x'_1\ \cdots\ x'_{\ell_N-2}\ \alpha\ a'_{\ell_N}\ 0\ 0\ \cdots$$

If we fix N large with respect to d, then for most of the integers $n \in [\![0, N-1]\!]$, we have $\mathcal{F}(n) \subset [\![0, N-1]\!]$.

$$[N]_k = a_0\ a_1\ \cdots\ a_{\ell_N-1}\ a_{\ell_N}\ 0\ 0\ \cdots$$

$$[n]_k = \underbrace{x_0\ x_1\ \cdots\ x_{\ell_N-2}}_{\gamma < k^{\ell_N-1}-d}\ \alpha\ \underbrace{a'_{\ell_N}}_{<a_{\ell_N}}\ 0\ 0\ \cdots$$

$$[n+d]_k = x'_0\ x'_1\ \cdots\ x'_{\ell_N-2}\ \alpha\ a'_{\ell_N}\ 0\ 0\ \cdots$$

$$[n]_k = \underbrace{x_0\ x_1\ \cdots\ x_{\ell_N-3}}_{\gamma < k^{\ell_N-2}-d}\ \alpha\ \underbrace{a'_{\ell_N-1}}_{<a_{\ell_N-1}}\ a_{\ell_N}\ 0\ 0\ \cdots$$

$$[n+d]_k = x'_0\ x'_1\ \cdots\ x'_{\ell_N-3}\ \alpha\ a'_{\ell_N-1}\ a_{\ell_N}\ 0\ 0\ \cdots$$

Therefore, the integers $n \in [\![0, N-1]\!]$ such that $\mathcal{F}(n) \subset [\![0, N-1]\!]$ is

$$> \; N - d\,k^2 - d\,k\,\sigma_k(N),$$

where $\sigma_k(N)$ is the sum of digits in base k of N.

Theorem (Marcovici/S./Tahay (2021+))

For all $g \in G$,

$$\operatorname{card}\left\{n \in [\![0, N-1]\!] : u_{n+d} - u_n = g\right\} \geq \frac{N}{|G|} - \frac{dk^2}{|G|}\left(1 + \frac{\log N}{\log k}\right).$$

Therefore

$$\left|\frac{1}{N}\operatorname{card}\left\{n \in [\![0, N-1]\!] : u_{n+d} - u_n = g\right\} - \frac{1}{|G|}\right| \leq \frac{d\,k}{N}\left(1 + \frac{\log N}{\log k}\right).$$

Theorem (Marcovici/S./Tahay (2020+))

For all $(i,j) \in G^2$,

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{1}{|G|^2}.$$

Theorem (Marcovici/S./Tahay (2020+))

For all $(i,j) \in G^2$,

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{card}\left\{n \in [\![0, N-1]\!] : (u_n, u_{n+d}) = (i,j)\right\} = \frac{1}{|G|^2}.$$

Idea of the proof: Let $n \in [\![0, k^{2N+1}-1]\!]$ such that $x_{N+1} = \alpha$, i.e. $n = m_1\, k^{N+1} + \alpha\, k^N + m_2$, with $m_1, m_2 \in [\![0, k^N - 1]\!]$,

$$(u_n, u_{n+d}) = (u_{km_1+\alpha}, u_{km_1+\alpha}) + (u_{\alpha k^N + m_2}, u_{\alpha k^N + m_2 + d}).$$

- For m_1 being chosen i.i.d, the distribution of $u_{km_1+\alpha}$ tends to the uniform distribution on G.
- For m_2 being chosen i.i.d, the distribution of $u_{\alpha k^N + m_2 + d} - u_{\alpha k^N + m_2}$ tends to the uniform distribution on G.

Plan

1. Original problem: Golay – Rudin – Shapiro
2. Small-scale and large-scale correlations
3. Large-scale correlations on prime and squarefree alphabets
4. Difference matrices
5. Large-scale correlations on arbitrary alphabets
6. Small-scale correlations on arbitrary alphabets
7. Higher dimensions

Higher dimensions

Work by Frank-Priebe (2003, 2004), Barbé & von Haeseler (2005), Ben-Abraham & Amit David (2019).

Higher dimensions

Work by Frank-Priebe (2003, 2004), Barbé & von Haeseler (2005), Ben-Abraham & Amit David (2019).

In dimension t, we define $\mathbf{u} = (u_{n_1,\dots,n_t})_{(n_1,\dots,n_t)\in\mathbb{N}^t} \in G^{\mathbb{N}^t}$ via the weight function $f : \Sigma_k^t \times \Sigma_k^t \to G$ with $f(0,0) = 0$ and

$$u_{n_1,\dots,n_t} = \sum_{i\in\mathbb{N}} f\left(\begin{pmatrix} x_i^1 \\ \vdots \\ x_i^t \end{pmatrix}, \begin{pmatrix} x_{i+1}^1 \\ \vdots \\ x_{i+1}^t \end{pmatrix}\right) = \sum_{i\in\mathbb{N}} f(x_i, x_{i+1}).$$

Higher dimensions

Work by Frank-Priebe (2003, 2004), Barbé & von Haeseler (2005), Ben-Abraham & Amit David (2019).

In dimension t, we define $\mathbf{u} = (u_{n_1,\dots,n_t})_{(n_1,\dots,n_t)\in\mathbb{N}^t} \in G^{\mathbb{N}^t}$ via the weight function $f : \Sigma_k^t \times \Sigma_k^t \to G$ with $f(0,0) = 0$ and

$$u_{n_1,\dots,n_t} = \sum_{i\in\mathbb{N}} f\left(\begin{pmatrix} x_i^1 \\ \vdots \\ x_i^t \end{pmatrix}, \begin{pmatrix} x_{i+1}^1 \\ \vdots \\ x_{i+1}^t \end{pmatrix}\right) = \sum_{i\in\mathbb{N}} f(x_i, x_{i+1}).$$

Difference condition

f fulfills the difference condition if for all $i \neq j$, and for all $g \in G$,

$$\operatorname{card}\left\{h \in \Sigma_k^t : f(i,h) - f(j,h) = g\right\} = \frac{k}{|G|}.$$

Analogous results for (u_n, u_{n+d}), for a fixed vector $d \in \mathbb{N}^t$, $d \neq 0$.

Analogous results for (u_n, u_{n+d}), for a fixed vector $d \in \mathbb{N}^t$, $d \neq 0$.
Examples for $k = 2$, $t = 2$, $G = \mathbb{Z}_2$.

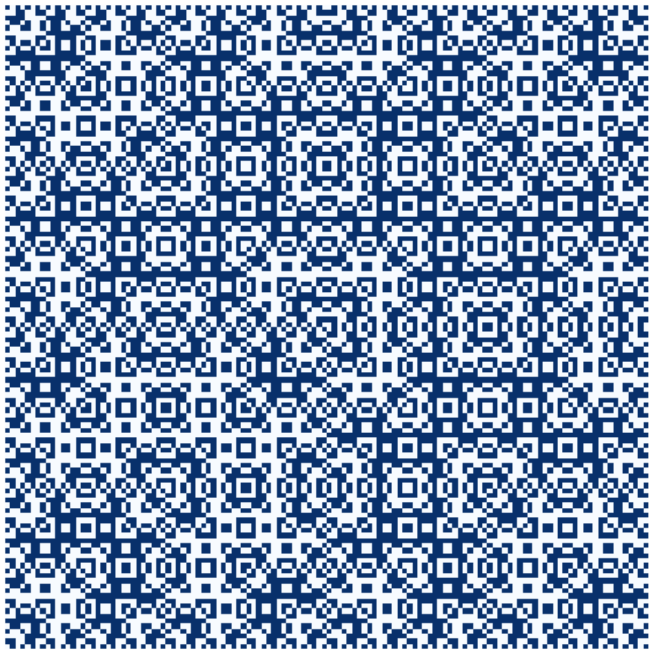
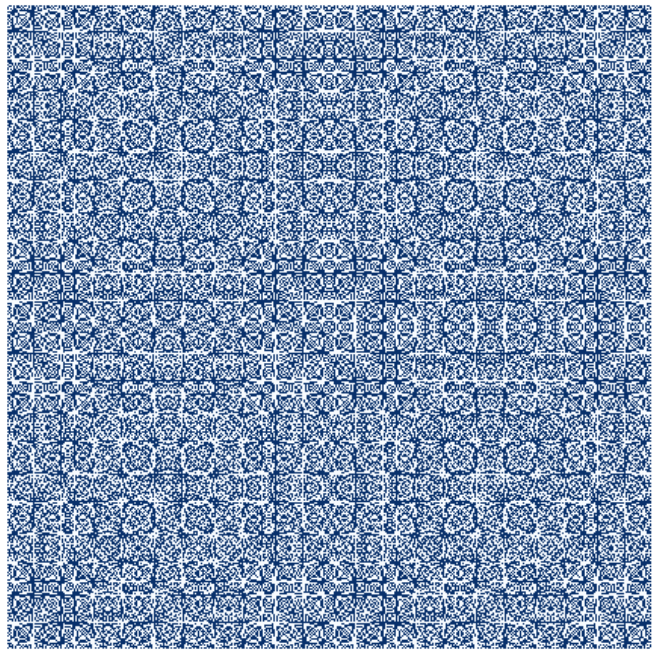
Example

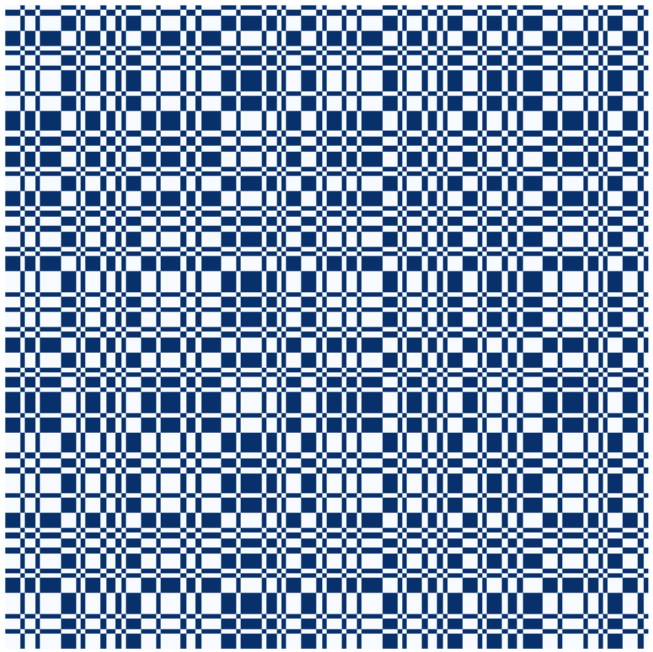
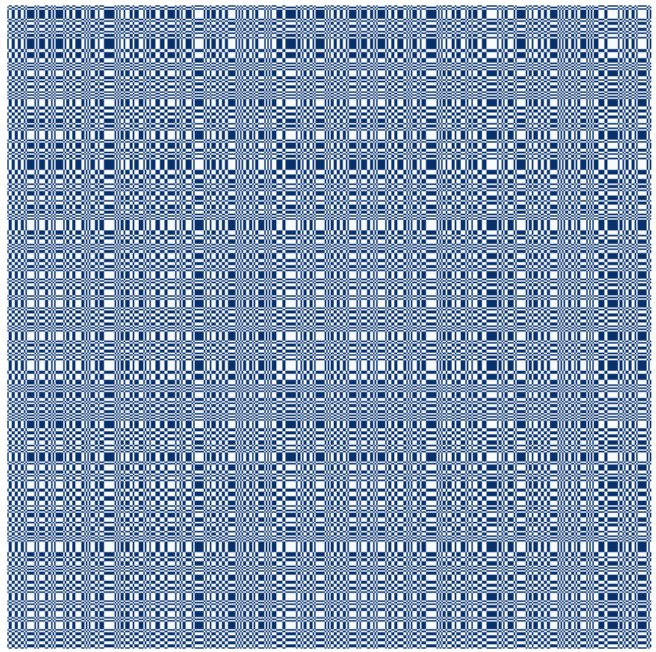
Matrix of weights (lexicographic order on Σ_2^2):

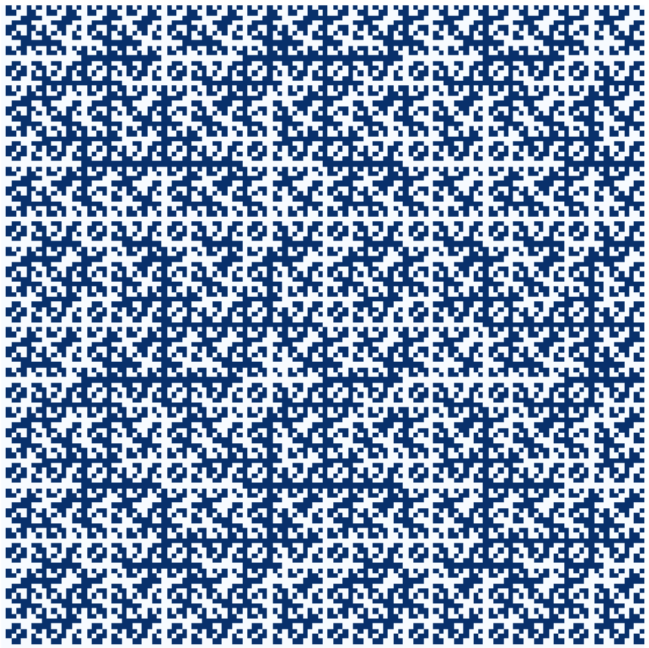
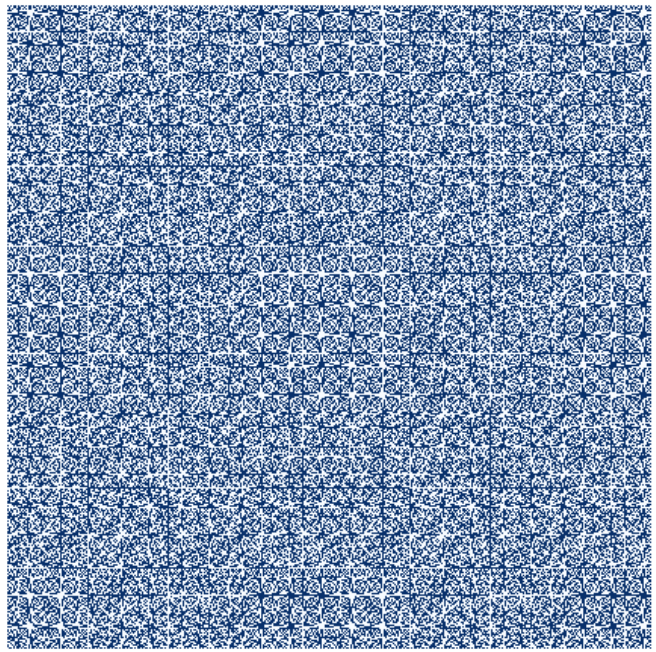
$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

$$\begin{array}{lcccccccccccccc} [436]_2 & = & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & \cdots & & \\ [48]_2 & = & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & & \\ \hline u_{436,48} & \equiv & 0 + & 1 + & 1 + & 1 + & 0 + & 1 + & 1 + & 0 + & 1 + & 0 + & & \cdots & \equiv 0 & \pmod 2 \end{array}$$

7	1	0	1	0	0	1	0	1
6	0	0	1	1	1	1	0	0
5	1	1	1	1	1	1	1	1
4	0	1	1	0	0	1	1	0
3	1	0	0	1	0	1	1	0
2	0	0	0	0	1	1	1	1
1	1	1	0	0	1	1	0	0
0	0	1	0	1	0	1	0	1
n_2 / n_1	0	1	2	3	4	5	6	7

Matrix	Terms in $[\![0, 2^7 - 1]\!]^2$	Terms in $[\![0, 2^{10} - 1]\!]^2$
$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$		

Matrix	Terms in $[\![0, 2^7 - 1]\!]^2$	Terms in $[\![0, 2^{10} - 1]\!]^2$
$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$		

Matrix	Terms in $[\![0, 2^7 - 1]\!]^2$	Terms in $[\![0, 2^{10} - 1]\!]^2$
$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$		

Matrix	Terms in $[\![0, 2^7 - 1]\!]^2$	Terms in $[\![0, 2^{10} - 1]\!]^2$
$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$