UNIVERSITY OF TWENTE.

Connections between Latin squares, Cellular Automata and Coprime Polynomials

Luca Mariot

l.mariot@utwente.nl

Algebra Seminars University of Guadalajara – October 13, 2023

Summary

Part 1: Cellular Automata and Mutually Orthogonal Latin Squares

Part 2: Bent Functions from CA

Part 3: A Simplified Construction with Linear Recurring Sequences

Conclusions

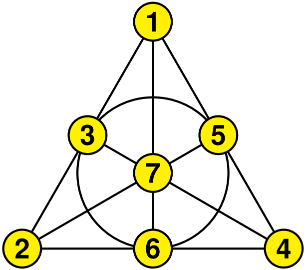
Part 1: Cellular Automata and Mutually Orthogonal Latin Squares

What is a Combinatorial Design (CD)?

- A collection $\mathcal{A}$ of subsets (or **blocks**) of a finite set X satisfying particular **balancedness** properties

- Example: the **Fano Plane**

$$X = \{1,2,3,4,5,6,7\}$$
$$\mathcal{A} = \{123, 145, 167, 246, 257, 347, 356\}$$

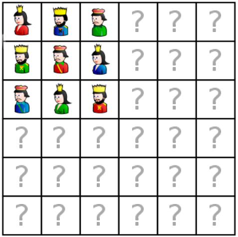


- Each block in $\mathcal{A}$ has 3 elements and each pair of distinct points in X occurs in exactly 1 block
- $\Rightarrow$ $(7,3,1)$-BIBD (**Balanced Incomplete Block Design**)

Euler’s 36 Officers Problem

« *A very curious question [...] revolves around arranging 36 officers to be drawn from 6 different ranks and also from 6 different regiments so that they are ranged in a square so that in each line (both horizontal and vertical) there are 6 officers of different ranks and different regiments.* »

L. Euler, *Sur une nouvelle espèce de quarrés magiques*, 1782



Latin Squares

Definition

A *Latin square* of order N is a $N \times N$ matrix L such that every row and every column are permutations of $[N] = \{1, \cdots, N\}$

1	3	4	2
4	2	1	3
2	4	3	1
3	1	2	4

Orthogonal Latin Squares (OLS)

Definition

Two Latin squares L_1 and L_2 of order N are *orthogonal* if their superposition yields all the pairs $(x, y) \in [N] \times [N]$.

1	3	4	2
4	2	1	3
2	4	3	1
3	1	2	4

(a) L_1

1	4	2	3
3	2	4	1
4	1	3	2
2	3	4	1

(b) L_2

1, 1	3, 4	4, 2	2, 3
4, 3	2, 2	1, 4	3, 1
2, 4	4, 1	3, 3	1, 2
3, 2	1, 3	2, 1	4, 4

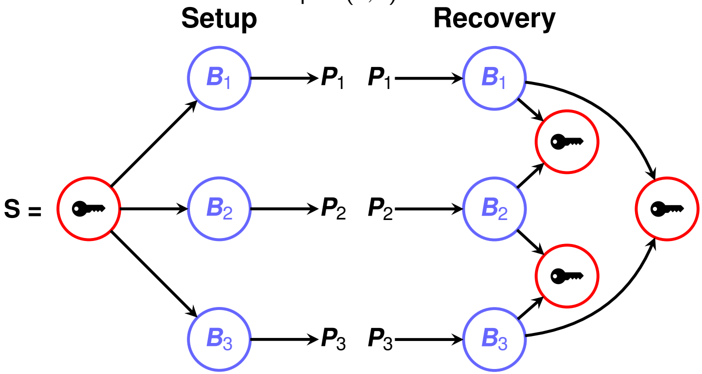
(c) (L_1, L_2)

n pairwise orthogonal Latin squares are denoted as *n*-MOLS (**Mutually Orthogonal Latin Squares**)

A Cryptographic Application of n-MOLS

(k, n) **Threshold Secret Sharing Scheme**: a **dealer** shares a **secret** S among n **players** so that at least k players out of n are required to recover S

Example: $(2,3)$–scheme

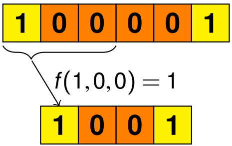


Remark: $(2, n)$–scheme $\Leftrightarrow$ set of n-MOLS

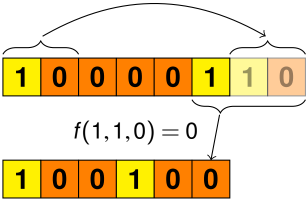
Cellular Automata

- Vectorial functions $F : \mathbb{F}_q^n \to \mathbb{F}_q^m$ with *uniform* (shift-invariant) coordinates

Example: $q = 2$, $n = 6$, $d = 3$, $f(s_i, s_{i+1}, s_{i+2}) = s_i \oplus s_{i+1} \oplus s_{i+2}$



No Boundary CA – NBCA



Periodic Boundary CA – PBCA

- Each cell updates its state $s \in \{0, 1\}$ by evaluating a local rule $f : \{0, 1\}^d \to \{0, 1\}$ on itself and the $d - 1$ cells on its right

Mutually Orthogonal Latin Squares (MOLS)

Definition

A *Latin square* is a $n \times n$ matrix where all rows and columns are permutations of $[n] = \{1, \cdots, n\}$. Two Latin squares are *orthogonal* if their superposition yields all the pairs $(x, y) \in [n] \times [n]$.

1	3	4	2
4	2	1	3
2	4	3	1
3	1	2	4

1	4	2	3
3	2	4	1
4	1	3	2
2	3	1	4

1 1	3 4	4 2	2 3
4 3	2 2	1 4	3 1
2 4	4 1	3 3	1 2
3 2	1 3	2 1	4 4

- k-**MOLS**: set of k pairwise orthogonal Latin squares

Latin Squares through Bipermutive CA (1/2)

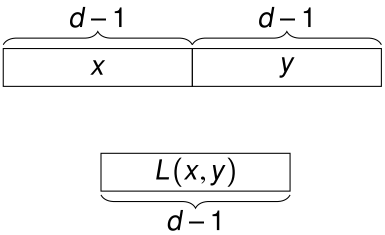
- Bipermutive CA: local rule f is defined as

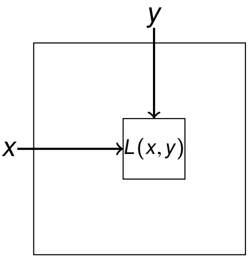
$$f(x_1, \cdots, x_d) = x_1 + \varphi(x_2, \cdots, x_{d-1}) + x_d$$

- $\varphi : \mathbb{F}_q^{d-2} \to \mathbb{F}_q$: generating function of f [LM13]

Lemma ([MFL16])

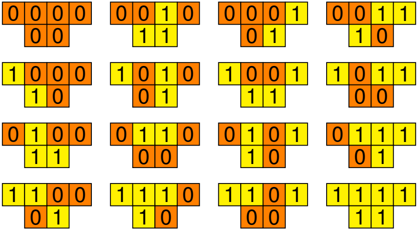
A (no-boundary) CA $F : \mathbb{F}_q^{2(d-1)} \to \mathbb{F}_q^d$ with bipermutive rule $f : \mathbb{F}_q^d \to \mathbb{F}_q$ generates a Latin square of order $N = q^{d-1}$





Latin Squares through Bipermutive CA (2/2)

- Example: CA $F : \mathbb{F}_2^4 \to \mathbb{F}_2^2$, $f(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3$ (Rule 150)
- Encoding: $00 \mapsto 1, 10 \mapsto 2, 01 \mapsto 3, 11 \mapsto 4$



(a) Rule 150 on 4 bits

1	4	3	2
2	3	4	1
4	1	2	3
3	2	1	4

(b) Latin square L_{150}

Linear CA

- Local rule: *linear combination* of the neighborhood cells

$$f(x_1,\cdots,x_d) = a_1x_1 + \cdots + a_dx_d \ , \ a_i \in \mathbb{F}_q$$

- Associated polynomial:

$$f \mapsto p_f(X) = a_1 + a_2X + \cdots + a_dX^{d-1}$$

- $(n-d+1)\times n$ transition matrix:

$$M_F = \begin{pmatrix} a_1 & \cdots & a_d & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & a_1 & \cdots & a_d & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & a_1 & \cdots & a_d \end{pmatrix}, \quad x \mapsto M_F x^\top$$

- **Remark:** a linear rule is bipermutive iff $a_1, a_d \neq 0$

MOLS from Linear Bipermutive CA (LBCA)

Theorem ([MGLF20])

A set of t linear bipermutive CA $F_1, \ldots F_t : \mathbb{F}_q^{2(d-1)} \to \mathbb{F}_q^{d-1}$ generates a family of t-MOLS of order $N = q^{d-1}$ if and only if their associated polynomials are pairwise coprime

1	4	3	2
2	3	4	1
4	1	2	3
3	2	1	4

(a) Rule 150

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

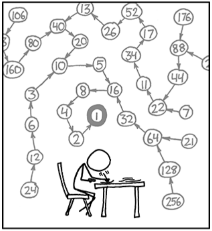
(b) Rule 90

1 1	4 2	3 3	2 4
2 2	3 1	4 4	1 3
4 3	1 4	2 1	3 2
3 4	2 3	1 2	4 1

(c) Superposition

Figure: $P_{150}(X) = 1 + X + X^2$, $P_{90}(X) = 1 + X^2$ (coprime)

Counting MOLS from linear CA



THE COLLATZ CONJECTURE STATES THAT IF YOU PICK A NUMBER, AND IF IT'S EVEN DIVIDE IT BY TWO AND IF IT'S ODD MULTIPLY IT BY THREE AND ADD ONE, AND YOU REPEAT THIS PROCEDURE LONG ENOUGH, EVENTUALLY YOUR FRIENDS WILL STOP CALLING TO SEE IF YOU WANT TO HANG OUT.

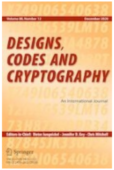
S `https://xkcd.com/710/`

- Number of coprime polynomials over $\mathbb{F}_2$ of degree n and nonzero constant term:

$$a(n) = 4^{n-1} + a(n-1) = \frac{4^{n-1} - 1}{3}$$
$$= 0, 1, 5, 21, 85, \ldots$$

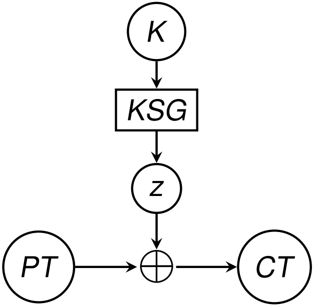
- Corresponds to OEIS A002450
- Generalized to any finite field, along with size of largest family of pairwise coprime polynomials, in:

L. Mariot, M. Gadouleau, E. Formenti, and A. Leporati. Mutually orthogonal latin squares based on cellular automata. Des. Codes Cryptogr. 88(2):391–411 (2020)

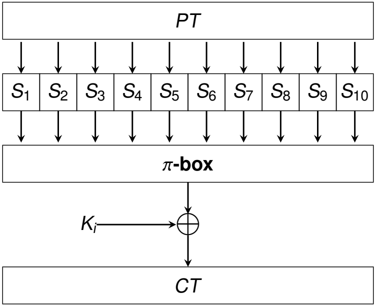


Part 2: Bent functions from CA

Boolean Functions in Symmetric Ciphers



(a) Stream cipher



(b) Block cipher

Boolean functions $f : \{0,1\}^n \to \{0,1\}$ are used in [C21]

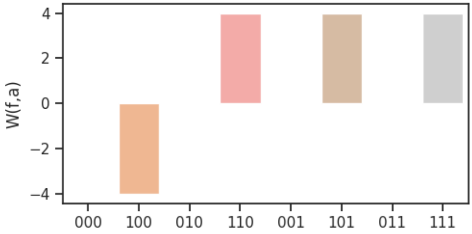
- **Stream ciphers**, to design the *keystream generator* (KSG)
- **Block ciphers**, as the coordinate functions of *S-boxes* (S_i)

Boolean Functions - Basic Representations

- Truth table: a 2^n-bit vector Ω_f specifying $f(x)$ for all $x \in \{0,1\}^n$

(x_1, x_2, x_3)	000	100	010	110	001	101	011	111
Ω_f	0	1	1	0	1	0	1	0

- Walsh Transform: correlation with linear functions $a \cdot x$, $W(f,a) = \sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus a \cdot x}$ for all $a \in \{0,1\}^n$

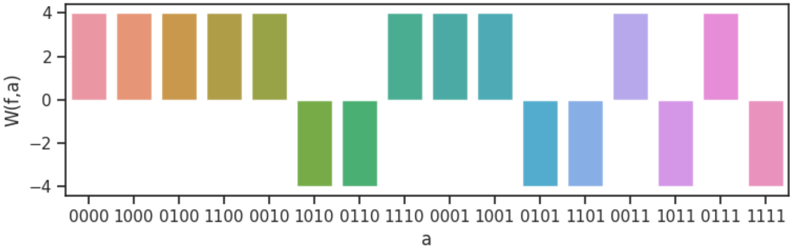


Bent Functions

- *Parseval's Relation*, valid on any Boolean function:

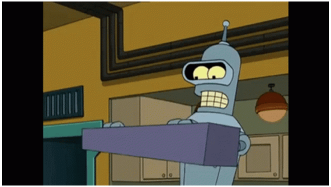
$$\sum_{a \in \{0,1\}^n} [W(f,a)]^2 = 2^{2n} \quad \text{for all} \quad f : \{0,1\}^n \to \{0,1\}$$

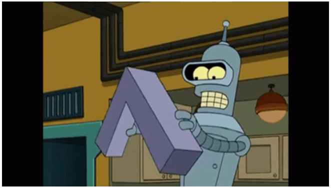
- Bent functions: $W(f,a) = \pm 2^{\frac{n}{2}}$ for all $a \in \{0,1\}^n$
  - Reach the highest possible *nonlinearity*
  - Exist only for n even and they are *unbalanced*



Example: $f(x_1, x_2, x_3, x_4) = x_1 x_3 + x_1 x_4 + x_2 x_4$

Intuition behind the name "bent"





- **Nonlinearity** of f: minimum Hamming distance of the truth table of f from all linear functions
- "Bent" functions are the farthest from linear ("straight") ones
- Related to the covering radius of **Reed-Muller codes**

Constructions of Bent Functions

Given $n = 2m$:

- **Maiorana-McFarland** [M73]): $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is defined as

$$f(x,y) = x \cdot \pi(y) \oplus g(y)$$

  where:

  - $\pi : \mathbb{F}_2^m \to \mathbb{F}_2^m$ permutation of $\mathbb{F}_2^m$
  - $g : \mathbb{F}_2^m \to \mathbb{F}_2$ any m-variable Boolean function

- **Partial spreads** [D74]: $f \in \mathcal{PS}^-$ ($f \in \mathcal{PS}^+$) is defined as

$$supp(f) = \bigcup_{S \in \mathcal{S}} (S \setminus \{\underline{0}\}) \quad \left(supp(f) = \bigcup_{S \in \mathcal{S}} S \right) ,$$

  with $\mathcal{S}$ a family of 2^{m-1} $(+1)$ m-dimensional subspaces of $\mathbb{F}_2^n$ with pairwise trivial intersection

Hadamard Matrices

- Hadamard Matrix: a $n \times n$ matrix with ± 1 entries and s.t. $H \cdot H^\top = I_n$

$$H = \begin{pmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{pmatrix}, \; n = 4$$

- Necessary condition: $n = 1, 2$ or $n = 4k$
- Hadamard Conjecture: a Hadamard matrix exists for every $n = 4k$

Hadamard Matrices and Bent Functions

Theorem (Dillon, 1974 [D74])

Given $f : \{0,1\}^n \to \{0,1\}$ and $\hat{f}(x) = (-1)^{f(x)}$. Define the $2^n \times 2^n$ matrix H for all $x, y \in \{0,1\}^n$ as:

$$H(x,y) = \hat{f}(x \oplus y)$$

Then, f is a bent function if and only if H is a Hadamard matrix.

Example: $f(x_1, x_2) = x_1 x_2$

x_1	x_2	$x_1 x_2$
0	0	0
1	0	0
0	1	0
1	1	1

$$H = \begin{pmatrix} + & + & + & - \\ + & + & - & + \\ + & - & + & + \\ - & + & + & + \end{pmatrix}$$

Hadamard Matrices from MOLS

Orthogonal Array $OA(t,N)$ for t MOLS of order N: $N^2 \times (t+2)$ matrix where each Latin square is "linearized" as a column

L_{90} $(1+X^2)$

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

L_{150} $(1+X+X^2)$

1	4	3	2
2	3	4	1
4	1	2	3
3	2	1	4

$\Rightarrow$

x	y	L_{90}	L_{150}
1	1	1	1
1	2	2	4
1	3	3	3
1	4	4	2
2	1	2	2
2	2	1	3
2	3	4	4
2	4	3	1
3	1	3	4
3	2	4	1
3	3	1	2
3	4	2	3
4	1	4	3
4	2	3	2
4	3	2	1
4	4	1	4

Theorem (Bush, 1973 [B73])

Given t MOLS of order $N = 2t$, there exists a $4t^2 \times 4t^2$ symmetric Hadamard matrix H

Construction:

- Put – only in (i,j) where $i \neq j$ and there is a column k in the OA s.t the rows i and j have the same symbol
- Put $+$ everywhere else

Bent Functions from any MOLS?

- **Remark**: Not all t-MOLS sets give rise to a Hadamard matrix with the $\hat{f}(x \oplus y)$ structure required for a bent function!
- Smallest counterexample: $n = 6$, $t = 2^{\frac{n-2}{2}} = 4$, $N = 2t = 8$

1	2	3	4	5	8	6	7
2	1	4	3	8	5	7	6
3	4	1	2	7	6	8	5
4	3	2	1	6	7	5	8
5	8	7	6	1	2	4	3
8	5	6	7	2	1	3	4
6	7	8	5	4	3	1	2
7	6	5	8	3	4	2	1

(a) L_1

1	2	3	4	5	6	7	8
3	4	1	2	8	7	6	5
5	6	8	7	1	2	4	3
8	7	5	6	3	4	2	1
4	3	2	1	7	8	5	6
2	1	4	3	6	5	8	7
6	5	7	8	2	1	3	4
7	8	5	6	4	3	1	2

(b) L_2

1	2	3	4	5	6	7	8
4	3	2	1	7	8	5	6
8	7	5	6	3	4	2	1
6	5	7	8	2	1	3	4
7	8	6	5	4	3	1	2
5	6	8	7	1	2	4	3
3	4	1	2	8	7	6	5
2	1	4	3	6	5	8	7

(c) L_3

1	2	3	4	5	6	7	8
5	6	8	7	1	2	4	3
4	3	2	1	7	8	5	6
7	8	6	5	4	3	1	2
8	7	5	6	3	4	2	1
3	4	1	2	8	7	6	5
2	1	4	3	6	5	8	7
6	5	7	8	2	1	3	4

(d) L_4

- The resulting 64×64 Hadamard matrix does not give a bent function

From Linear CA to Bent Functions

- **Question**: Are MOLS arising from linear CA suitable for constructing bent functions?
- We consider only CA over $\mathbb{F}_q$ with $q = 2^l$, $l \in \mathbb{N}$
- The order of the Hadamard matrix must be $4t^2 = 2^n$
- We need t coprime polynomials of degree $b = d - 1$:

$$2^{lb} = 2t \Leftrightarrow lb = 1 + \log_2 t$$

- Since both l and b are integers, $t = 2^w$ for $w \in \mathbb{N}$

From Linear CA to Bent Functions

Theorem

Let H be the Hadamard matrix of order $2^{2(w+1)}$ defined by the t LBCA $F_1, \cdots F_t : \mathbb{F}_q^{2b} \to \mathbb{F}_q^b$, and define $f : \mathbb{F}_2^n \to \mathbb{F}_2$, $n = 2(w+1)$ as:

$$f(x) = \begin{cases} 0 \ , & \textit{if } x = 0 \\ 1 \ , & \textit{if } x \neq 0 \textit{ and } \exists k \in \{1, \cdots, t\} \textit{ s.t. } F_k(x) = 0 \\ 0 \ , & \textit{otherwise} \end{cases} ,$$

Then, it holds that:

$$H(x,y) = \hat{f}(x \oplus y)$$

and thus f is a bent function

Remark: The linearity of the CA is crucial to grant this result (and costed us our first reject! [GMP20])

Example

$p_f(X) = 1 + X^2$

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

$p_g(X) = 1 + X + X^2$

1	4	3	2
2	3	4	1
4	1	2	3
3	2	1	4

$$A = \left\{ \begin{array}{cc} L_1 & L_2 \\ \hline 1 & 1 \\ 2 & 4 \\ 3 & 3 \\ 4 & 2 \\ 2 & 2 \\ 1 & 3 \\ 4 & 4 \\ 3 & 1 \\ 3 & 4 \\ 4 & 1 \\ 1 & 2 \\ 2 & 3 \\ 4 & 3 \\ 3 & 2 \\ 2 & 1 \\ 1 & 4 \end{array} \right.$$

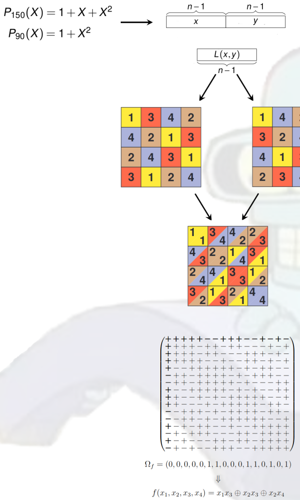
$$H = \begin{pmatrix}
+ & + & + & + & + & - & - & + & + & + & - & - & + & - & + & - \\
+ & + & + & + & - & + & + & - & + & + & - & - & - & + & - & + \\
+ & + & + & + & - & + & + & - & - & - & + & + & + & - & + & - \\
+ & + & + & + & + & - & - & + & - & - & + & + & - & + & - & + \\
+ & - & - & + & + & + & + & + & + & - & + & - & + & + & - & - \\
- & + & + & - & + & + & + & + & - & + & - & + & + & + & - & - \\
- & + & + & - & + & + & + & + & + & - & + & - & - & - & + & + \\
+ & - & - & + & + & + & + & + & - & + & - & + & - & - & + & + \\
+ & + & - & - & + & - & + & - & + & + & + & + & + & - & - & + \\
+ & + & - & - & - & + & - & + & + & + & + & + & - & + & + & - \\
- & - & + & + & + & - & + & - & + & + & + & + & - & + & + & - \\
- & - & + & + & - & + & - & + & + & + & + & + & + & - & - & + \\
+ & - & + & - & + & + & - & - & + & - & - & + & + & + & + & + \\
- & + & - & + & + & + & - & - & - & + & + & - & + & + & + & + \\
+ & - & + & - & - & - & + & + & - & + & + & - & + & + & + & + \\
- & + & - & + & - & - & + & + & + & - & - & + & + & + & + & +
\end{pmatrix}$$

$$\Omega_f = (0,0,0,0,0,1,1,0,0,0,1,1,0,1,0,1)$$

$$\Downarrow$$

$$f(x_1, x_2, x_3, x_4) = x_1 x_3 \oplus x_2 x_3 \oplus x_2 x_4$$

Figure 3: Example of bent function of $n = 4$ variables generated by the $t = 2$ MOLS of order $2t = 4$ defined by the LBCA with rule 90 and 150, respectively. The two Latin squares are represented on the left in the OA form. The first row and the first column of the Hadamard matrix H coincide with the polarity truth table of the function.



Combinatorial questions addressed in [GMP20]:

- **Existence:** for even n, does a large enough family of coprime polynomials exist?
- **Counting:** how many families of this kind exist (= number of CA-based bent functions)?

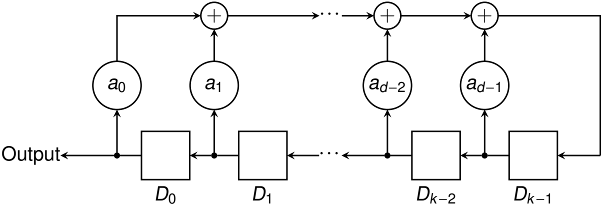
Part 3: A Simplified Construction with Linear Recurring Sequences

Linear Recurring Sequences (LRS)

- Sequence $\{x_i\}_{i\in\mathbb{N}}$ satisfying the following relation:

$$a_0 x_i + a_1 x_{i+1} + ... + a_{d-1} x_{i+d-1} = x_{i+d}$$

- Computed by a *Linear Feedback Shift Register* (LFSR):



- Feedback polynomial:

$$f(X) = a_0 + a_1 X + \cdots a_{d-1} X^{d-1} + X^d$$

Linear map associated to a LRS

- Take the *projection* of all sequences satisfying the LRS defined by $f(X)$ onto their first $2d$ coordinates
- Obtain a d-dim subspace $S_f \subseteq \mathbb{F}_q^{2d}$ which is the kernel of the linear map $F : \mathbb{F}_q^{2d} \to \mathbb{F}_q^d$:

$$F(x_0, \cdots, x_{2d-1})_i = a_0 x_i + a_1 x_{i+1} + \ldots + a_{d-1} x_{i+d-1} + x_{i+d} \ ,$$

associated matrix:

$$M_F = \begin{pmatrix} a_0 & \cdots & a_{d-1} & 1 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & a_0 & \cdots & a_{d-1} & 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & a_0 & \cdots & a_{d-1} & 1 \end{pmatrix}$$

- ... but this is *exactly* the global rule of a linear CA!

Partial Spreads from Coprime Polynomials

Lemma ([GMP23])

Given $f, g \in \mathbb{F}_q[X]$ *over* $\mathbb{F}_q$ *of degree* $d \geq 1$*, defined as:*

$$f(X) = a_0 + a_1 X + \cdots + a_{d-1} X^{d-1} + X^d \ , \tag{1}$$

$$g(X) = b_0 + b_1 X + \cdots + b_{d-1} X^{d-1} + X^d \ , \tag{2}$$

Then, the kernels of $F, G : \mathbb{F}_q^{2d} \to \mathbb{F}_q^d$ *have trivial intersection if and only if* $\gcd(f, g) = 1$

Consequence: a family of t pairwise coprime polynomials defines a partial spread

Equivalence check

For degree $b = 1$, actually nothing new:

Lemma ([GMP23])

Our construction coincides with the class $\mathcal{PS}_{ap}$ when $b = 1$.

For degree $b = 2$:

- Computed the ranks of the associated Hadamard matrices in binary form to check equivalence
- **1st Finding**: none of our functions are equivalent to Maiorana-McFarland ones
- **2nd Finding**: many of our functions are not even equivalent to $\mathcal{PS}_{ap}$ ones

Conclusions

Recap and Open Problems

Remarkable findings:

- (Complicated!) construction of bent functions via CA, Latin Squares and Hadamard matrices [GMP20]
- Simplification based on kernels of LRS subspaces [GMP23]
- Resulting bent functions coincide with $\mathcal{PS}_{ap}$ for degree $b = 1$
- For $b = 2$, many functions are not in $\mathcal{PS}_{ap}$

Open problems:

- Are functions from polynomials of degree $b = 2$ *really* new?
- Implementation of CA-based bent functions via LFSR [ML18]

References

[B73] K. Bush: Construction of symmetric Hadamard matrices. In: A survey of combinatorial theory, pp. 81–83. Elsevier (1973)

[C21] C. Carlet: Boolean functions for cryptography and coding theory. Cambridge University Press (2021)

[D74] J.F. Dillon.: Elementary Hadamard difference sets. Ph.D. thesis (1974)

[GMP23] M. Gadouleau, L. Mariot, S. Picek. Bent functions in the partial spread class generated by linear recurring sequences. Des. Codes and Cryptogr. (2022) DOI: `https://doi.org/10.1007/s10623-022-01097-1`

[GMP20] M. Gadouleau, L. Mariot, S. Picek: Bent Functions from Cellular Automata. IACR Cryptol. ePrint Arch. 2020: 1272 (2020)

[LM13] A. Leporati, L. Mariot: 1-Resiliency of Bipermutive Cellular Automata Rules. In: Proceedings of AUTOMATA 2013: 110-123 (2013)

[MGLF20] L. Mariot, M. Gadouleau, E. Formenti, and A. Leporati. Mutually orthogonal latin squares based on cellular automata. Des. Codes Cryptogr. 88(2):391–411 (2020)

[MFL16] L. Mariot, E. Formenti, A. Leporati: Constructing Orthogonal Latin Squares from Linear Cellular Automata. In: Exploratory papers of AUTOMATA 2016 (2016)

[ML18] L. Mariot, A. Leporati: A cryptographic and coding-theoretic perspective on the global rules of cellular automata. Nat. Comput. 17(3):487–498 (2018)

[M73] R. L. McFarland. A family of difference sets in non-cyclic groups. J. Comb. Theory, Ser. A 15(1):1–10 (1973)

[M16] S. Mesnager: Bent Functions – Fundamentals and Results. Springer (2016)