A new construction for bent functions based on cellular automata, Latin squares and linear recurring sequences

Luca Mariot
l.mariot@utwente.nl

Joint work with Maximilien Gadouleau and Stjepan Picek

DIS Lunch Talk – February 17, 2023
Introduction: an (unlucky) paper


Published August 13, 2022 [GMP23]
Luca Mariot  
A new construction for bent functions based on cellular automata, Latin squares and linear recurring sequences
Boolean Functions in Symmetric Ciphers

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

- **Stream ciphers**, to design the *keystream generator* (KSG)
- **Block ciphers**, as the coordinate functions of *S-boxes* ($S_i$)
Truth table: a $2^n$-bit vector $\Omega_f$ specifying $f(x)$ for all $x \in \{0, 1\}^n$

$\Omega_f$

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

- **Parseval’s Relation**, valid on any Boolean function:
  \[ \sum_{a \in \{0, 1\}^n} [W(f, a)]^2 = 2^{2n} \text{ for all } f : \{0, 1\}^n \rightarrow \{0, 1\} \]

- **Bent functions**: \( W(f, a) = \pm 2^{n/2} \) 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]: \( f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2 \) is defined as
  \[
  f(x, y) = x \cdot \pi(y) \oplus g(y)
  \]

  where:
  - \( \pi : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^m \) permutation of \( \mathbb{F}_2^m \)
  - \( g : \mathbb{F}_2^m \rightarrow \mathbb{F}_2 \) any \( m \)-variable Boolean function

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

  \[
  \text{supp}(f) = \bigcup_{S \in \mathcal{S}} (S \setminus \{0\}) \quad \text{or} \quad \text{supp}(f) = \bigcup_{S \in \mathcal{S}} S
  \]

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

  with pairwise trivial intersection
Part 1: Cellular Automata and Mutually Orthogonal Latin Squares
Vectorial functions $F : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^m$ 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

Concrete example: the $\chi$ transformation in Keccak [BDPV11]:

$q = 2$, $d = 3$, $\chi(x_i, x_{i+1}, x_{i+2}) = x_i \oplus (1 \oplus x_{i+1})x_{i+2}$
Mutually Orthogonal Latin Squares (MOLS)

Definition

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

$k$-MOLS: set of $k$ pairwise orthogonal Latin squares
Bipermutive CA: local rule $f$ is defined as

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

$\varphi: \mathbb{F}_q^{d-2} \rightarrow \mathbb{F}_q$: generating function of $f$

Lemma ([M16])

A (no-boundary) CA $F: \mathbb{F}_q^{2(d-1)} \rightarrow \mathbb{F}_q^d$ with bipermutive rule $f: \mathbb{F}_q^d \rightarrow \mathbb{F}_q$ generates a Latin square of order $N = q^{d-1}$
Example: CA $F : \mathbb{F}_2^4 \rightarrow \mathbb{F}_2^2$, $f(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3$ (Rule 150)

Encoding: $00 \leftrightarrow 1, 10 \leftrightarrow 2, 01 \leftrightarrow 3, 11 \leftrightarrow 4$

(a) Rule 150 on 4 bits

(b) Latin square $L_{150}$
Local rule: *linear combination* of the neighborhood cells

\[ f(x_1, \ldots, x_d) = a_1 x_1 + \cdots + a_d x_d, \quad a_i \in \mathbb{F}_q \]

Associated polynomial:

\[ f \mapsto p_f(X) = a_1 + a_2 X + \cdots + a_d X^{d-1} \]

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

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

**Remark:** a linear rule is bipermutative iff \(a_1, a_d \neq 0\)
MOLS from Linear Bipermutive CA (LBCA)

Theorem ([MGLF20])

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

Figure: $P_{150}(X) = 1 + X + X^2$, $P_{90}(X) = 1 + X^2$ (coprime)
Part 2: The Complicated Construction
Hadamard Matrices

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

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

- **Necessary condition**: $n = 1, 2$ or $n = 4k$

- **Hadamard Conjecture**: a Hadamard matrix exists for every $n = 4k$
Theorem (Dillon, 1974 [D74])

Given \( f : \{0, 1\}^n \rightarrow \{0, 1\} \) 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: \( f(x_1, x_2) = x_1x_2 \)

\[
\begin{array}{c|c|c|c|c|c|c|c|c}
  x_1 & x_2 & x_1x_2 \\
\hline
  0 & 0 & 0 \\
  1 & 0 & 0 \\
  0 & 1 & 0 \\
  1 & 1 & 1 \\
\end{array}
\]

\[
H = \begin{pmatrix}
  + & + & + & - \\
  + & + & - & + \\
  + & - & + & + \\
  - & + & + & + \\
\end{pmatrix}
\]
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

$\begin{array}{cccc}
L_{90} (1 + X^2) \\
1 & 2 & 3 & 4 \\
2 & 1 & 4 & 3 \\
3 & 4 & 1 & 2 \\
4 & 3 & 2 & 1 \\
\end{array}$

$\begin{array}{cccc}
L_{150} (1 + X + X^2) \\
1 & 4 & 3 & 2 \\
2 & 3 & 4 & 1 \\
4 & 1 & 2 & 3 \\
3 & 2 & 1 & 4 \\
\end{array}$

$\begin{array}{cccc}
x & y & L_{90} & L_{150} \\
1 & 1 & 1 & 1 \\
1 & 2 & 2 & 4 \\
1 & 3 & 3 & 3 \\
1 & 4 & 4 & 2 \\
2 & 1 & 2 & 2 \\
2 & 3 & 4 & 4 \\
2 & 4 & 3 & 1 \\
3 & 1 & 3 & 4 \\
3 & 2 & 4 & 1 \\
3 & 3 & 1 & 2 \\
3 & 4 & 2 & 3 \\
4 & 1 & 4 & 3 \\
4 & 2 & 3 & 2 \\
4 & 3 & 2 & 1 \\
4 & 4 & 1 & 4 \\
\end{array}$

**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
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 \( b = d - 1 \):
  \[
  2^{lb} = 2t \iff lb = 1 + \log_2 t
  \]
- Since both \( l \) and \( b \) are integers, \( t = 2^w \) for \( w \in \mathbb{N} \)
Let \( H \) be the Hadamard matrix of order \( 2^{2(w+1)} \) defined by the \( t \) LBCA \( F_1, \cdots F_t : \mathbb{F}_q^{2^b} \rightarrow \mathbb{F}_q^b \), and define \( f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2 \), \( n = 2(w+1) \) as:

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

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!)
Example

\[ p_f(X) = 1 + X^2 \]

\[ p_g(X) = 1 + X + X^2 \]

\[ L_1 \]
\[ \begin{array}{cccc}
1 & 1 \\
2 & 4 \\
3 & 3 \\
4 & 2 \\
1 & 3 \\
4 & 4 \\
3 & 1 \\
3 & 4 \\
4 & 1 \\
1 & 2 \\
2 & 3 \\
4 & 3 \\
3 & 2 \\
2 & 1 \\
1 & 4 \\
\end{array} \]

\[ L_2 \]

\[ A = \]

\[ H = \]

\[ \Omega_f = (0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1) \]

\[ f(x_1, x_2, x_3, x_4) = x_1x_3 \oplus x_2x_3 \oplus x_2x_4 \]

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.
Existence and Counting

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
Reviewers Feedback

- First attempt: BFA, reject (incomplete proof)
- Second attempt: CCDS, reject (complicated construction, no guarantee the obtained bent functions are new)
- Third attempt: DESI, major revision

Strictest (and most enthusiastic!) review:

This paper must be published in some form! :) It has the potential of becoming a major reference on bent functions because it identifies a new source of partial spreads large enough to give bent functions!

But not in its present form which buries and obscures their great new result in many pages of material on subjects which, in retrospect, are unnecessary for a lucid exposition of their new results. It may be in order to briefly mention how they were led to their theorem via the CA, Latin square, MOLS and Bush
Linear Recurring Sequences (LRS)

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

\[
a_0 x_i + a_1 x_{i+1} + \ldots + a_{d-1} x_{i+d-1} = x_{i+d}
\]

- 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 $F : \mathbb{F}_q^{2d} \rightarrow \mathbb{F}_q^d$:

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

associated matrix:

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

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

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

\begin{align*}
  f(X) &= a_0 + a_1 X + \cdots + a_{d-1} X^{d-1} + X^d, \\  g(X) &= b_0 + b_1 X + \cdots + b_{d-1} X^{d-1} + X^d,
\end{align*}

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

Consequence: a family of $t$ pairwise coprime polynomials defines a partial spread
For degree $b = 1$, actually nothing new:

**Lemma**

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
Remarkable findings:

- (Complicated!) construction of bent functions via CA, Latin Squares and Hadamard matrices
- Simplification based on kernels of LRS subspaces
- 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]
Luca Mariot
A new construction for bent functions based on cellular automata, Latin squares and linear recurring sequences