Subspace Codes generated by Linear CA

Luca Mariot, Federico Mazzone

l.mariot@utwente.nl

AUTOMATA 2023 – Trieste, September 1, 2023

Network Coding and Subspace codes

- Routing packets on networks is not always the most efficient transmission method [K12]

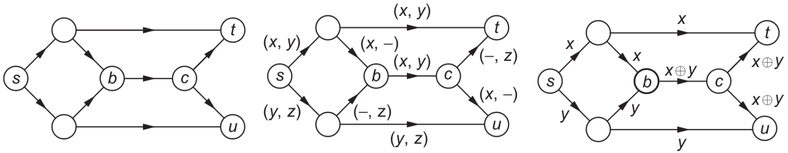
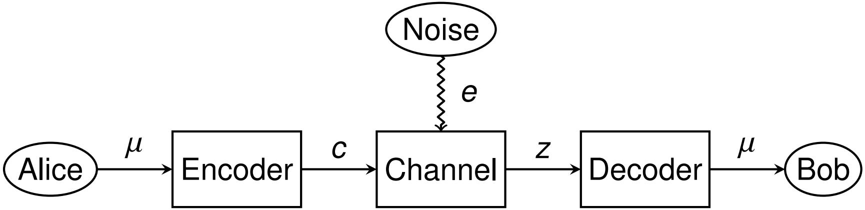


Image credits: F. R. Kschischang, *An Introduction to Network Coding*

- **Network Coding**: combine packets together as linear combinations
- **Noncoherent setting**: does not consider the underlying topology of the network (subspace codes)

Error Correction Problem



- $\mu \in \{0,1\}^k$: message
- $c \in \{0,1\}^n$: codeword ($n > k$)
- $e \in \{0,1\}^n$: error pattern
- $z = c \oplus e$ (received word)

Error-Correcting Codes

Hamming Distance (HD) of $x, y \in \{0,1\}^n$: number of positions where x and y differ

Definition

(n, d_C) Binary (unrestricted) code of length n and minimum distance d_C: subset $C \subseteq \{0,1\}^n$ such that for all $c_1, c_2 \in C$

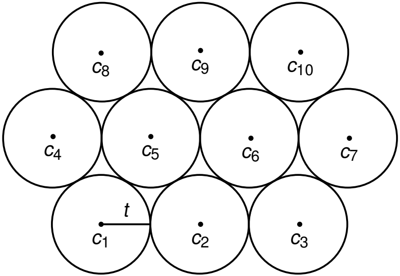
$$HD(c_1, c_2) \geq d_C$$

Example: a $(4,2)$ code $C \subseteq \{0,1\}^4$

0000	1001
0011	1010
0101	1100
0110	1111

Conflicting Requirements on Codes

- **High minimum distance d_C**
- **High number of codewords $c \in C$**



- Sphere of $c \in C \Leftrightarrow$ $S_c = \{z \in \mathbb{F}_2^n : \ d_H(z,c) \leq t\}$
- $t = \left\lfloor \frac{d-1}{2} \right\rfloor \Leftrightarrow$ Error-correction capability of C

Linear Codes

Notation:

- $\mathbb{F}_2 = \{0,1\}$: finite field of order 2
- $\mathbb{F}_2^n = \{0,1\}^n$: n-dimensional vector space over $\mathbb{F}_2$

Definition

A (n,k,d) binary linear code C: A (n,d) code C that is also a k-dimensional subspace of $\mathbb{F}_2^n$

$$g_1, \cdots, g_k \in \mathbb{F}_2^n \text{ basis of } C \Leftrightarrow G = \begin{pmatrix} g_1 \\ \vdots \\ g_k \end{pmatrix} k \times n \text{ generator matrix of } C$$

Encoding: vector-matrix multiplication

$$\mu \mapsto c = \mu G$$

Subspace Codes

- **Idea:** codewords are not vectors, but rather *vector subspaces*
- **Distance** between two subspaces:

$$d(A,B) = dim(A) + dim(B) - 2dim(A \cap B) \ .$$

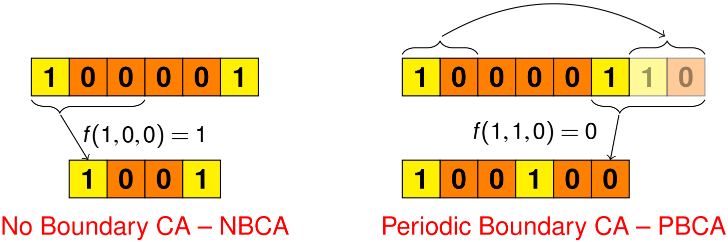
Definition ([KK08])

A subspace code $\mathcal{C}$ of parameters $[n, \ell(\mathcal{C}), \log_q |\mathcal{C}|, D(\mathcal{C})]$ is a family of subsets of $\mathbb{F}_q^n$ where $\ell(\mathcal{C}) = \max_{V \in \mathcal{C}}\{dim(V)\}$ and $D(\mathcal{C})$ is the minimum distance of $\mathcal{C}$, defined as:

$$D(\mathcal{C}) = \min_{U,V \in \mathcal{C}} \{d(U,V)\}$$

- Vectorial functions $F : \mathbb{F}_q^n \to \mathbb{F}_q^m$ with *uniform* (shift-invariant) coordinates [MPLJ19]

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



Linear CA

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

$$f(x_1, \cdots, x_d) = a_1 x_1 + \cdots + a_d x_d \ , \ a_i \in \mathbb{F}_q$$

- 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:

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

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

Sylvester Matrices

- Two linear bipermutive CA with rules $f, g : \mathbb{F}_q^d \to \mathbb{F}_q$ generate orthogonal Latin squares iff this matrix is invertible [MGLF20]:

$$
M_{F,G} = \begin{pmatrix}
a_1 & \cdots & a_d & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & a_1 & \cdots & a_d & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & a_1 & \cdots & a_d \\
b_1 & \cdots & b_d & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & b_1 & \cdots & b_d & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & b_1 & \cdots & b_d
\end{pmatrix}
$$

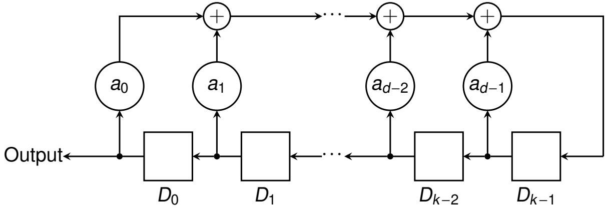
- ... but this is the **Sylvester matrix** of the two polynomials p_f, p_g, and $det(M_{F,G}) \neq 0 \Leftrightarrow \gcd(p_f, p_g) = 1$

Linear Recurring Sequences (LRS)

- Sequence $\{x_i\}_{i \in \mathbb{N}}$ satisfying the following relation [LN97]:

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

- 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 [GMP23]
- Obtain a d-dim subspace $S_f \subseteq \mathbb{F}_q^{2d}$ which is the kernel of the linear map $F : \mathbb{F}_q^{2d} \to \mathbb{F}_q^d$:

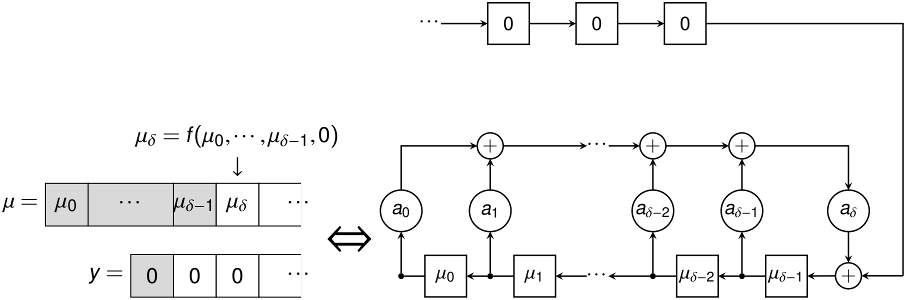
$$F(x_0, \cdots, x_{2d-1})_i = a_0 x_i + a_1 x_{i+1} + ... + a_{d-1} x_{i+d-1} + x_{i+d} \ ,$$

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!

Kernel as CA preimage computation



Kernel $\Leftrightarrow$ 0-preimage of CA [ML18]

Partial Spreads from Coprime Polynomials

Partial spread: A family $\mathcal{S}$ of subspaces of $\mathbb{F}_q^n$ with pairwise trivial intersection [C21]

Lemma

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

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

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

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 gives CA kernels that form a partial spread

Subspaces Codes from kernels of linear CA

Lemma

Let $f, g \in \mathbb{F}_q[X]$ be two polynomials, and denote by $S_{f,g}$ their Sylvester matrix. Then,

$$dim(null(S_{f,g})) = \deg(\gcd(f,g))$$

Theorem

Let $\mathcal{F}$ be a set of linear CA of length $2k$ and diameter d, $k = d-1$. Then, the minimum distance of the subspace code $\mathcal{C}_{\mathcal{F}}$ is:

$$D(\mathcal{C}_{\mathcal{F}}) = 2k - 2 \cdot \max_{\substack{F,G \in \mathcal{F} \\ F \neq G}} \left\{ \deg(\gcd(P_f, P_g)) \right\} \ , \qquad (3)$$

where P_f, P_g are the polynomials associated to F and G.

Conclusions and Future Works

Conclusions:

- **Coprime case**: optimal distance (partial spread codes [GR14])
- **trade-off**: the higher the maximum degree of the GCD, the more linear CA we can squeeze into the code

Future work:

- Characterize families with uniform degree of pairwise GCD (equidistant codes)
- Investigate decoding efficiency of CA-based subspace codes [ML18a]

References

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

[GMP23] M. Gadouleau, L. Mariot, S. Picek. Bent functions in the partial spread class generated by linear recurring sequences. Des. Codes and Cryptogr. 91(1): 63-82 (2023)

[GR14] E. Gorla, A. Ravagnani: Partial spreads in random network coding. Finite Fields Their Appl. 26: 104–115 (2014)

[KK08] R. Koetter, F.R. Kschischang: Coding for errors and erasures in random network coding. IEEE Trans. Inf. Theory 54(8): 3579–3591 (2008)

[K12] F.R. Kschischang: An introduction to network coding. In: Network Coding. Elsevier, pp. 1–37 (2012)

[LN97] R. Lidl, H. Niederreiter: Finite fields. Cambridge University Press (1997)

[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)

[MPLJ9] L. Mariot, S. Picek, A. Leporati, and D. Jakobovic. Cellular automata based S-boxes. Cryptography and Communications 11(1): 41-62 (2019)

[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)

[ML18a] L. Mariot, A. Leporati: Inversion of mutually orthogonal cellular automata. In: Proceedings of ACRI 2018, LNCS vol. 11115, pp. 364–376, Springer (2018)