

# Latin Hypercubes and Cellular Automata

Maximilien Gadouleau<sup>1</sup> and Luca Mariot<sup>2</sup>

<sup>1</sup>Department of Computer Science, Durham University, South Road, Durham  
DH1 3LE, United Kingdom ,

[m.r.gadouleau@durham.ac.uk](mailto:m.r.gadouleau@durham.ac.uk)

<sup>2</sup>Cyber Security Research Group, Delft University of Technology,, Mekelweg 2,  
Delft, The Netherlands ,

[l.mariot@tudelft.nl](mailto:l.mariot@tudelft.nl)

November 26, 2020

## Abstract

Latin squares and hypercubes are combinatorial designs with several applications in statistics, cryptography and coding theory. In this paper, we generalize a construction of Latin squares based on bipermutive cellular automata (CA) to the case of Latin hypercubes of dimension  $k > 2$ . In particular, we prove that linear bipermutive CA (LBCA) yielding Latin hypercubes of dimension  $k > 2$  are defined by sequences of invertible Toeplitz matrices with partially overlapping coefficients, which can be described by a specific kind of regular de Bruijn graph induced by the support of the determinant function. Further, we derive the number of  $k$ -dimensional Latin hypercubes generated by LBCA by counting the number of paths of length  $k - 3$  on this de Bruijn graph.

**Keywords** Latin squares · Latin hypercubes · cellular automata · bipermutivity · Toeplitz matrices · de Bruijn graphs

## 1 Introduction

Several cryptographic protocols with information-theoretic security guarantees can be defined in terms of *Combinatorial Designs*. One such example are  $(k, n)$  threshold *Secret Sharing Schemes* (SSS), where a *dealer* wants to share a secret information  $S$  among a set of  $n$  *players* by giving to each of them a *share*  $B$ , so that at least  $k$  players are required to recover  $S$  by combining their shares. On the other hand, any subset of  $k - 1$  or less players do not gain any information on the secret by pooling the respective shares. Such a protocol corresponds to a set of  $n$  *Mutually Orthogonal Latin Squares* (MOLS) when  $k = 2$  players are required to reconstruct the secret, while it is equivalent to a set of  $n$  *Mutually Orthogonal Latin Hypercubes* of dimension  $k$  (MOLH) when  $k > 2$  [10]. Indeed, the schemes proposed by Shamir [9] and Blakley's [1] can be thought as particular instances of  $(k, n)$  threshold SSS, where polynomials and hyperplanes over finite fields are respectively used to represent an underlying set of MOLH in a compact way.

A recent research thread considers the design of secret sharing schemes by means of *Cellular Automata* (CA) [6], the goal being twofold. First, a CA-based architecture could be useful for efficient hardware-oriented implementations of threshold SSS, by leveraging the massive parallelism of the CA models. This approach already turned out to be interesting in another cryptographic application, namely the design of *S-boxes* based on CA [8]. Second, the locality of the CA model can be used to define access structures that are more constrained than the classic  $(k, n)$  threshold, such as the *consecutive access structure* of the CA-based scheme proposed in [6], which finds application in certain distributed cryptographic protocols, as discussed by the authors of [4]. These access structures eventually become *cyclic*, and the maximum number of players allowed is related to the period of spatially periodic preimages in surjective CA [7].

Given the equivalence between  $(k, n)$  threshold SSS and families of MOLH, a possible way to tackle these goals is to study how to construct the latter using cellular automata. Recently, the authors of [5] showed a construction of MOLS based on *linear bijective CA* (LBCA), thereby addressing the case of  $(2, n)$  threshold SSS based on CA. Generalizing this construction to a higher threshold  $k$  entails two steps: first, one needs to characterize which CA generate Latin hypercubes, since contrarily to the Latin squares case not all LBCA define Latin hypercubes of dimension  $k > 2$ . The next step is to define subsets of such CA whose Latin hypercubes are  $k$ -wise orthogonal, thus obtaining sets of *MOLH*.

The aim of this paper is to address the first step of this generalization, namely the characterization and enumeration of Latin hypercubes based on LBCA. In particular, we first prove that LBCA which generate *Latin cubes* are defined by local rules whose central coefficients compose the border of an invertible *Toeplitz matrix*. This allows us to determine the number of Latin cubes generated by LBCA by counting the number of invertible Toeplitz matrices. Next, we generalize this result to dimension  $k > 3$ , remarking that in this case the local rule of an LBCA can be defined by a *sequence* of  $k - 2$  invertible Toeplitz matrices, where adjacent matrices share the coefficients respectively of the upper and lower triangulars. We finally show that this overlapping relation can be described by a regular de Bruijn graph, and that paths of length  $k - 2$  on this graph corresponds to LBCA generating  $k$ -dimensional Latin hypercubes.

The rest of this paper is organized as follows. Section 2 covers the necessary background on Latin hypercubes and cellular automata used in the paper, and recalls the basic results about Latin squares generated by CA. Section 3 presents the characterization of Latin cubes and hypercubes generated by LBCA defined by invertible Toeplitz matrices. Section 4 proves the regularity of the de Bruijn graph associated to invertible Toeplitz matrices and derives the number of  $k$ -dimensional Latin hypercubes generated by LBCA. Section 5 recaps the main contributions of the paper and discusses some directions for future research.

## 2 Preliminary Definitions and Results

The main combinatorial objects of interest of this paper are *Latin hypercubes*, which we define as follows:

**Definition 1.** *Let  $X$  be a finite set of  $N \in \mathbb{N}$  elements. A Latin hypercube of dimension  $k \in \mathbb{N}$  and order  $N$  over  $X$  is a  $k$ -dimensional array with entries from  $X$  such that, by fixing any subset  $i_1, \dots, i_{k-1}$  of  $k - 1$  coordinates, the remaining*

coordinate  $i_k$  yields a permutation over  $X$ , that is, each element in  $X$  appears exactly once on the  $i_k$  coordinate.

Remark that when  $k = 2$  one obtains the definition of *Latin square* of order  $N$ , i.e. a square matrix where each row and each column is a permutation of  $X$ . Usually, Latin squares and hypercubes are defined over the set  $X = [N] = \{1, \dots, N\}$  of the first  $N$  positive integers. In this work, we consider the case where  $X$  is the vector space  $\mathbb{F}_q^b$ , with  $\mathbb{F}_q$  being the finite field of  $q$  elements. Hence, the order  $N$  of the hypercube will be  $q^b$ , i.e. the number of all  $q$ -ary vectors of length  $b$ .

Next, we define the model of *cellular automaton* used in the rest of this paper:

**Definition 2.** Let  $d, n \in \mathbb{N}$ , with  $d \leq n$ , and let  $f : \mathbb{F}_q^d \rightarrow \mathbb{F}_q$  be a function of  $d$  variables over the finite field  $\mathbb{F}_q$ , with  $q$  being a power of a prime. The Cellular Automaton (CA in the following) of length  $n$  and local rule  $f$  over the alphabet  $\mathbb{F}_q$  is the vectorial function  $F : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^{n-d+1}$  defined for all vectors  $x = (x_1, \dots, x_n) \in \mathbb{F}_q^n$  as:

$$F(x_1, \dots, x_n) = (f(x_1, \dots, x_d), \dots, f(x_{n-d+1}, \dots, x_n)) . \quad (1)$$

In other words, each coordinate function  $f_i : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$  for  $i \in [n-d+1]$  of a CA  $F$  corresponds to its local rule  $f$  applied to the neighborhood  $\{x_i, \dots, x_{i+d-1}\}$ . For our examples, in the following we will mainly consider the case  $q = 2$ , where a CA boils down to a particular type of *vectorial Boolean function*  $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^{n-d+1}$ . This corresponds to the definition of *no-boundary cellular automaton* (NBCA) studied in [8] as a model for cryptographic S-boxes.

The last preliminary definition we need is that of hypercube associated to a CA. In what follows, we assume that the vectors in  $\mathbb{F}_q^b$  are totally ordered, and that there is a monotone and one-to-one mapping  $\Psi : [N] \rightarrow \mathbb{F}_q^b$ , in order to associate sets of integer coordinates to sets of  $q$ -ary vectors.

**Definition 3.** Let  $b, k \in \mathbb{N}$ , with  $d = b(k-1) + 1$ . Moreover, let  $F : \mathbb{F}_q^{bk} \rightarrow \mathbb{F}_q^b$  be the CA of length  $bk$  defined by a local rule  $f : \mathbb{F}_q^d \rightarrow \mathbb{F}_q$ . Then, the hypercube associated to  $F$  of order  $N = q^b$  is the  $k$ -dimensional array  $\mathcal{H}_F$  where for all vectors of coordinates  $(i_1, \dots, i_k) \in [N]^k$  the corresponding entry is defined as:

$$\mathcal{H}_F(i_1, \dots, i_k) = \Psi^{-1}(F(\Psi(i_1) || \Psi(i_2) || \dots || \Psi(i_k))) , \quad (2)$$

where the input  $\Psi(i_1) || \Psi(i_2) || \dots || \Psi(i_k)$  denotes the concatenation of the binary vectors  $\Psi(i_1), \dots, \Psi(i_k) \in \mathbb{F}_q^b$ .

Thus, by Definition 3 the hypercube associated to a CA  $F$  of length  $bk$  and local rule  $f$  of  $b(k-1) + 1$  is constructed by splitting the input vector of  $F$  in  $k$  blocks of size  $b$ , which are used to index the coordinates of the hypercube, while the output vector represents the entry to be put at those coordinates. Figure 1 depicts an example of 3-dimensional hypercube  $\mathcal{H}_F$  of block size  $b = 2$  over  $\mathbb{F}_2$  (thus, of order  $2^2 = 4$ ) associated to the CA  $F : \mathbb{F}_2^6 \rightarrow \mathbb{F}_2^2$  which is defined by the local rule  $f(x_1, \dots, x_5) = x_1 \oplus x_3 \oplus x_5$ . In this case one can see that  $\mathcal{H}_F$  is indeed a *Latin cube*, i.e. a Latin hypercube of dimension 3.

We now give the formal statement of the problem investigated in this paper, namely the construction and enumeration of Latin hypercubes with CA:

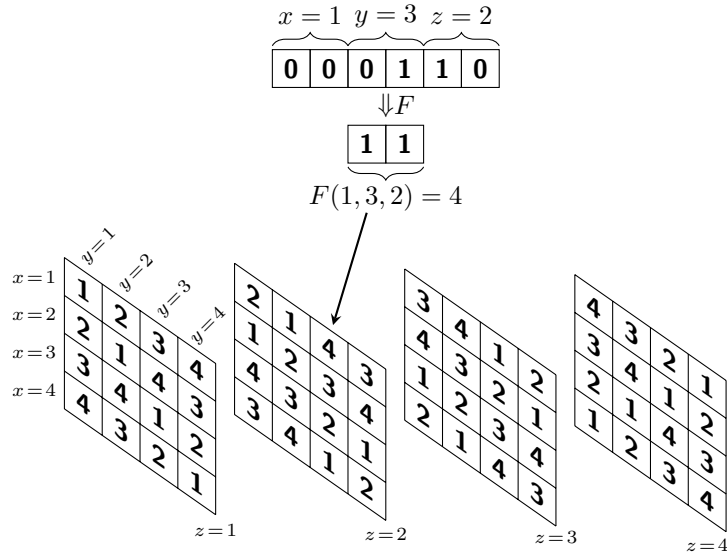


Figure 1: Latin cube of order 4 generated by a CA  $F : \mathbb{F}_2^6 \rightarrow \mathbb{F}_2^2$  defined by local rule  $f(x_1, \dots, x_5) = x_1 \oplus x_3 \oplus x_5$ . The encoding used is  $00 \mapsto 1, 10 \mapsto 2, 01 \mapsto 3, 11 \mapsto 4$ .

**Problem 1.** Let  $F : \mathbb{F}_q^{bk} \rightarrow \mathbb{F}_q^b$  be a CA equipped with a local rule  $f : \mathbb{F}_2^d \rightarrow \mathbb{F}_2$  where  $d = b(k - 1) + 1$ . When is the hypercube  $\mathcal{H}_F$  associated to  $F$  a Latin hypercube? How many local rules  $f : \mathbb{F}_q^{b(k-1)+1} \rightarrow \mathbb{F}_q$  induce a CA such that the resulting  $\mathcal{H}_F$  is a Latin hypercube?

Problem 1 requires studying under which conditions the local rule  $f$  induces a permutation between any of the  $k$  blocks of  $b$  consecutive cells used to index the coordinates of  $\mathcal{H}_F$  and the output CA configuration, when all remaining blocks are fixed. We first start by addressing the extremal cases of the leftmost and rightmost blocks in CA defined by *bipermutive* local rules.

A function  $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$  of  $n \geq 2$  variables is *bipermutive* if there exists a function  $g : \mathbb{F}_q^{n-2} \rightarrow \mathbb{F}_q$  such that

$$f(x_1, \dots, x_n) = x_1 \oplus g(x_2, \dots, x_{n-1}) \oplus x_n \tag{3}$$

for all  $x = (x_1, \dots, x_n) \in \mathbb{F}_q^n$ , where  $\oplus$  corresponds to the sum operation over  $\mathbb{F}_q$ . When  $q = 2$ , Equation (3) basically amounts to the XOR of the leftmost and rightmost input variables with the result of function  $g$  computed on the central  $n - 2$  variables. A proof of the following result can be found in [5]:

**Lemma 1.** Let  $F : \mathbb{F}_q^{bk} \rightarrow \mathbb{F}_q^b$  be a CA with bipermutive rule  $f : \mathbb{F}_q^d \rightarrow \mathbb{F}_q$ , where  $d = b(k - 1) + 1$ . Then, the restriction  $F|_{\tilde{x}} : \mathbb{F}_q^b \rightarrow \mathbb{F}_q^b$  obtained by fixing either the rightmost or leftmost  $b(k - 1)$  variables in the CA input to  $\tilde{x} \in \mathbb{F}_q^{b(k-1)}$  is a permutation over  $\mathbb{F}_q^b$  for all  $\tilde{x} \in \mathbb{F}_q^{b(k-1)}$ .

Thus, by Lemma 1 bipermutivity of the local rule  $f$  is a sufficient condition for verifying the Latin hypercube property on the leftmost and rightmost coordinate

of the hypercube  $\mathcal{H}_F$ . This also means, in turn, that for dimension  $k = 2$  (that is, when there are no blocks in the middle between the leftmost and the rightmost one) bipermutivity is a sufficient condition to ensure that  $\mathcal{H}_F$  is a Latin square:

**Corollary 1.** *Let  $F : \mathbb{F}_q^{2b} \rightarrow \mathbb{F}_q^b$  be a CA with bipermutive rule  $f : \mathbb{F}_q^d \rightarrow \mathbb{F}_q$  of diameter  $d = b + 1$ . Then, the hypercube  $\mathcal{H}_F$  associated to  $F$  is a Latin square of order  $q^b$ .*

From Corollary 1, one also gets the following straightforward counting result:

**Corollary 2.** *Let  $b \in \mathbb{N}$ . Then, the number of Latin squares of order  $q^b$  generated by bipermutive CA corresponds to the number of bipermutive local rules of  $b + 1$  variables over  $\mathbb{F}_q$ , which is  $q^{q^{b-1}}$ .*

Hence, Corollaries 1 and 2 solve Problem 1 for dimension  $k = 2$ , in the particular case of bipermutive CA. In what follows, we shall solve Problem 1 in Theorems 2 and 5 for any dimension  $k > 2$ .

### 3 CA-based Latin Hypercubes from Toeplitz Matrices

#### 3.1 Latin Cubes

Remark that for dimension  $k > 2$  bipermutivity is not enough. As a matter of fact, Lemma 1 requires the  $b(k - 1)$  variables to be adjacent. To verify the Latin hypercube property for a middle coordinate  $1 < i < k$  one needs to fix all variables on the left and on the right except for the “hole” represented by the block of  $b$  bits associated to coordinate  $i$ . As a consequence, it is necessary to characterize a proper subset of bipermutive local rules that generate Latin hypercubes when used as local rules of CA.

We begin by addressing the case of *Latin cubes*, that is with dimension  $k = 3$ . Referring to Problem 1, we have a CA  $F : \mathbb{F}_q^{3b} \rightarrow \mathbb{F}_q^b$  that maps configurations of  $3b$  cells in vectors of  $b$  cells, defined by a bipermutive local rule  $f : \mathbb{F}_q^d \rightarrow \mathbb{F}_q$  of diameter  $d = 2b + 1$ . Since the permutation between the output CA configuration and the blocks  $x_{[1,b]}$  and  $x_{[2b+1,3b]}$  is already granted by Lemma 1, we only need to consider the middle block  $x_{[b+1,2b]}$ .

In what follows, we will also make the additional assumption that, beside being bipermutive, the local rule is *linear*: In other words, there exist a binary vector  $a = (a_1, a_2, \dots, a_{d-1}, a_d) \in \mathbb{F}_q^d$  such that

$$f(x_1, x_2, \dots, x_{d-1}, x_d) = a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_{d-1}x_{d-1} \oplus a_dx_d, \quad (4)$$

where sum and product are considered over  $\mathbb{F}_q$ . Notice that a linear rule defined as in (4) is bipermutive if and only if both  $a_1$  and  $a_d$  are not null. In particular, from now on we will assume that  $a_1 = a_d = 1$ , and we will define a linear bipermutive rule by means of the vector  $(a_2, \dots, a_{d-1}) \in \mathbb{F}_q^{d-2}$  defining the  $d - 2$  central coefficients. Additionally, we will refer to a CA defined by such a rule as a *LBCA (Linear Bipermutive CA)*.

For all  $x \in \mathbb{F}_q^{3b}$ , let  $y = F(x) \in \mathbb{F}_q^b$  be the result of the CA applied to vector  $x$ . Since the local rule  $f$  is linear, one can express  $y$  as a system of  $b$  linear

equations and  $3b$  variables:

$$\begin{cases} y_1 &= x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_{2b} x_{2b} \oplus x_{2b+1} \\ y_2 &= x_2 \oplus a_2 x_3 \oplus \cdots \oplus a_{2b} x_{2b+1} \oplus x_{2b+2} \\ &\vdots \\ y_b &= x_b \oplus a_2 x_{b+1} \oplus \cdots \oplus a_{2b} x_{3b-1} \oplus x_{3b} \end{cases} \quad (5)$$

Suppose now that we fix the  $2b$  variables  $x_1, \dots, x_b$  and  $x_{2b+1}, \dots, x_{3b}$  respectively to the values  $\tilde{x}_1, \dots, \tilde{x}_b$  and  $\tilde{x}_{2b+1}, \dots, \tilde{x}_{3b}$ . This actually amounts to fixing the leftmost and the rightmost coordinates in the cube  $\mathcal{H}_F$  associated to  $F$ . Moreover, the system (5) becomes a system of  $b$  linear equations and  $b$  variables corresponding to the block  $x_{[b+1, 2b]}$ , since the remaining  $2b$  variables have been set to constant values. In order to ensure that there is a permutation between  $x_{[b+1, 2b]}$  and  $y$ , it means that the matrix of coefficients  $a_i$  multiplying the vector  $x_{[b+1, 2b]}$  in (5) must be invertible:

$$M_F = \begin{pmatrix} a_{b+1} & a_{b+2} & \cdots & a_{2b} \\ a_b & a_{b+1} & \cdots & a_{2b-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & \cdots & a_{b+1} \end{pmatrix} \quad (6)$$

Remark that the matrix in Equation (6) is a *Toeplitz matrix*, where the first row of coefficients  $a_{b+1}, \dots, a_{2b}$  is shifted to the right while the coefficients  $a_b, \dots, a_2$  progressively enter from the left. In particular, the matrix is completely characterized by the shifts of the central coefficients  $a_2, \dots, a_b, \dots, a_{2b}$  of the CA local rule  $f$ . To summarize, we obtained the following result:

**Lemma 2.** *Let  $F : \mathbb{F}_q^{3b} \rightarrow \mathbb{F}_q^b$  be a LBCA with rule  $f : \mathbb{F}_q^{2b+1} \rightarrow \mathbb{F}_q$  defined for all  $x \in \mathbb{F}_q^{2b+1}$  as*

$$f(x_1, \dots, x_{2b+1}) = x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_{2b} x_{2b} \oplus x_{2b+1} .$$

*Then, the hypercube  $\mathcal{H}_F$  associated to  $F$  is a Latin cube of order  $q^b$  if and only if the Toeplitz matrix  $M_F$  defined by the coefficients  $a_2, \dots, a_{2b} \in \mathbb{F}_q$  is invertible.*

The authors of [3] showed that the number of nonsingular  $n \times n$  Toeplitz matrices is  $q^{2(n-1)}(q-1)$ . Hence, we have the following counting result for Latin cubes:

**Theorem 1.** *Let  $b \in \mathbb{N}$ . Then, the number of LBCA  $F : \mathbb{F}_q^{3b} \rightarrow \mathbb{F}_q^b$  whose associated hypercube  $\mathcal{H}_F$  is a Latin cube is  $q^{2(b-1)}(q-1)$ .*

### 3.2 Latin Hypercubes of Dimension $k > 3$

We now generalize the investigation to Latin hypercubes of any dimension  $k > 3$ . In this case, the LBCA  $F : \mathbb{F}_q^{bk} \rightarrow \mathbb{F}_q^b$  is defined by a rule  $f : \mathbb{F}_q^{b(k-1)+1} \rightarrow \mathbb{F}_q$  of the form:

$$f(x_1, \dots, x_{b(k-1)+1}) = x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_{b(k-1)} x_{b(k-1)} \oplus x_{b(k-1)+1} . \quad (7)$$

Hence, the values of the output vector  $y = F(x) \in \mathbb{F}_q^b$  will be determined by a system analogous to (5):

$$\begin{cases} y_1 = x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_{b(k-1)} x_{b(k-1)} \oplus x_{b(k-1)+1} \\ y_2 = x_2 \oplus a_2 x_3 \oplus \cdots \oplus a_{b(k-1)} x_{b(k-1)+1} \oplus x_{b(k-1)+2} \\ \vdots \\ y_b = x_b \oplus a_2 x_{b+1} \oplus \cdots \oplus a_{b(k-1)} x_{bk-1} \oplus x_{bk} \end{cases} \quad (8)$$

The  $k$ -dimensional hypercube  $\mathcal{H}_F$  associated to  $F$  will be a Latin hypercube only if there is a permutation between any of the central  $k-2$  blocks of  $b$  cells when all the others are fixed to a constant value and  $y$  (the leftmost and rightmost cases already being granted by bipermutivity). Similarly to the three-dimensional case where we had only one central block, this means that all of the following Toeplitz matrices must be invertible for all  $i \in [k-2] = \{1, \dots, k-2\}$ :

$$M_{F,i} = \begin{pmatrix} a_{bi+1} & a_{bi+2} & \cdots & a_{b(i+1)-1} \\ a_{bi} & a_{bi+1} & \cdots & a_{b(i+1)-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{b(i-1)+2} & a_{b(i-1)+3} & \cdots & a_{bi+1} \end{pmatrix} \quad (9)$$

where  $M_{F,1}$  is associated to the permutation on the second block,  $M_{F,2}$  to the permutation on the third, and so on until  $M_{F,k-2}$ , which is associated to the permutation on the  $(k-1)$ -th block. We thus get the following characterization of LBCA that generate  $k$ -dimensional Latin hypercubes:

**Theorem 2.** *Let  $F : \mathbb{F}_2^{bk} \rightarrow \mathbb{F}_2^b$  be a CA with local rule  $f : \mathbb{F}_2^{b(k-1)+1} \rightarrow \mathbb{F}_2$  defined as in Equation (7). Then, the  $k$ -dimensional hypercube  $\mathcal{H}_F$  of order  $q^b$  associated to  $F$  is a Latin hypercube if and only if the Toeplitz matrix  $M_{F,i}$  in (9) is invertible for all  $i \in [k-2]$ .*

Since we settled the first part of Problem 1 in the context of LBCA, we now focus on the counting question, i.e. what is the number  $L_{b,k}$  of LBCA  $F : \mathbb{F}_q^{bk} \rightarrow \mathbb{F}_q^b$  that generate  $k$ -dimensional Latin hypercubes of order  $q^b$ . In other words, we want to count in how many ways the coefficients  $a_2, \dots, a_{b(k-1)} \in \mathbb{F}_2$  of the local rule  $f$  in (7) can be chosen so that the resulting hypercube  $\mathcal{H}_F$  is a Latin hypercube.

**Remark 1.** *For dimensions  $k = 2$  and  $k = 3$ , we settled the counting question of Problem 1 respectively in Corollaries 2 and 1. For dimension  $k > 3$ , remark that one cannot choose independently the coefficients defining the  $k-2$  invertible Toeplitz matrices. Indeed, consider the two adjacent matrices  $M_{F,i}$  and  $M_{F,i+1}$ :*

$$\begin{pmatrix} a_{bi+1} & a_{bi+2} & \cdots & a_{b(i+1)} \\ a_{bi} & a_{bi+1} & \cdots & a_{b(i+1)-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{b(i-1)+2} & a_{b(i-1)+3} & \cdots & a_{bi+1} \end{pmatrix} \begin{pmatrix} a_{b(i+1)+1} & a_{b(i+1)+2} & \cdots & a_{b(i+2)} \\ a_{b(i+1)} & a_{b(i+1)+1} & \cdots & a_{b(i+2)-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{bi+2} & a_{bi+3} & \cdots & a_{b(i+1)+1} \end{pmatrix}$$

*One can notice that the coefficients  $a_{bi+2}, \dots, a_{b(i+1)}$  overlap between the two matrices: in particular, they occur respectively above the main diagonal of  $M_{F,i}$  and below the main diagonal of  $M_{F,i+1}$ . As a consequence  $L_{b,k,q}$  for  $k > 3$  is lower than  $[(q-1)q^{2(b-1)}]^{k-2}$ , which is the number of ways one can choose a set of  $k-2$  invertible Toeplitz matrices of size  $b \times b$  over  $\mathbb{F}_q$  with repetitions.*

We now model the problem of determining  $L_{b,k,q}$  in terms of *determinant* functions. Let  $M_a$  be the Toeplitz matrix defined by the vector of  $2b - 1$  binary coefficients  $a = (a_2, \dots, a_{2b}) \in \mathbb{F}_q^{2b-1}$ , and let us define  $\det : \mathbb{F}_q^{2b-1} \rightarrow \mathbb{F}_q$  as the function of  $2b - 1$  variables that associates to each vector  $a \in \mathbb{F}_q^{2b-1}$  the determinant of the matrix  $M_a$ . Thus, we have that the support set  $\text{supp}(\det) = \{a \in \mathbb{F}_q^{2b-1} : \det(a) \neq 0\}$  contains the vectors  $a$  that define all  $b \times b$  non-singular Toeplitz matrices. By Theorem 1 it follows that the cardinality of the support is  $|\text{supp}(\det)| = q^{2(b-1)}(q - 1)$ .

Recall that by Remark 1 the elements of two adjacent invertible Toeplitz matrices overlap respectively on the upper and lower triangular parts. Hence, in order to construct a Latin hypercube of dimension  $k$ , one can choose from the support of the determinant function  $\det$  a *sequence* of  $k - 2$  vectors so that each pair of adjacent vectors overlap respectively on the last and the first  $b - 1$  coordinates. We now formalize this reasoning in terms of *de Bruijn graphs*. Let  $s \in \mathbb{N}$  and  $A$  be a finite alphabet, with  $A^*$  denoting the free monoid of words over  $A$ . Given  $u, v \in A^*$  such that  $|u| \geq s$  and  $|v| \geq s$ , we define the *s-fusion operator*  $\odot$  as in [11], that is  $u \odot v = z$  if and only if there exists  $x \in A^s$  and  $u_0, v_0 \in A^*$  such that  $u = u_0x$ ,  $v = xv_0$ , and  $z = u_0xv_0$ . In other words  $z$  is obtained by *overlapping* the right part of  $u$  and the left part of  $v$  of length  $s$ . Setting  $A = \mathbb{F}_q$  and  $s = b - 1$ , we obtain the *s-fusion operator* for our case of overlapping vectors  $u, v \in \text{supp}(\det)$  of length  $2b - 1$ . The set of overlapping relations can be conveniently described using the de Bruijn graph associated to the determinant function, which we formally define below:

**Definition 4.** Let  $\det : \mathbb{F}_q^{2b-1} \rightarrow \mathbb{F}_q$  be the determinant function for  $b \times b$  Toeplitz matrices over  $\mathbb{F}_q$ . The de Bruijn graph associated to  $\det$  is the directed graph  $G_{\det} = (V, E)$  where the set of vertices is  $V = \text{supp}(\det)$ , while an ordered pair of vertices  $(v_1, v_2)$  belongs to the set of edges  $E$  if and only if there exists  $z \in \mathbb{F}_q^{2b-1}$  such that  $z = v_1 \odot v_2$ , where  $\odot$  denotes the *s-fusion operator* with  $s = b - 1$ .

**Example 1.** Let  $b = 2, k = 5$  and  $q = 2$ . In this case, the local rule  $f : \mathbb{F}_2^3 \rightarrow \mathbb{F}_2$  has coefficients  $a_1$  and  $a_9$  set to 1, while the central coefficients  $(a_{2i}, a_{2i+1}, a_{2i+2})$  define the following Toeplitz matrix for  $i \in \{1, 2, 3\}$ :

$$M_{F,i} = \begin{pmatrix} a_{2i+1} & a_{2i+2} \\ a_{2i} & a_{2i+1} \end{pmatrix} \tag{10}$$

In particular  $M_{F,1}, M_{F,2}$  and  $M_{F,3}$  will be the matrices respectively associated to the second, third and fourth block. It is easily seen that the determinant of  $M_{F,i}$  is  $a_{2i+1} \oplus a_{2i}a_{2i+2}$ . Figure 2 reports the truth table and the de Bruijn graph of the determinant function  $\det : \mathbb{F}_2^3 \rightarrow \mathbb{F}_2$ . Rows in bold in the table correspond to the vectors of the support of  $\det$ , which in turn are the vertices of the de Bruijn graph. Following Figure 2, one can construct the local rule  $f$  by choosing a sequence of  $k - 2 = 3$  vectors from the support of  $\det(a_2, a_3, a_4)$  such that both the first and the second and the second and the third overlap respectively on the last and the first bit. This actually amounts to finding a path of length 2 on the de Bruijn graph. An example could be the sequence  $(0, 1, 0) - (0, 1, 1) - (1, 0, 1)$ . In particular,  $(0, 1, 0) = (a_2, a_3, a_4)$  is the vector defining the Toeplitz matrix of the second block. Similarly, the vector  $(0, 1, 1) = (a_4, a_5, a_6)$  defines the Toeplitz matrix of the third block while  $(1, 0, 1) = (a_6, a_7, a_8)$  defines the Toeplitz matrix



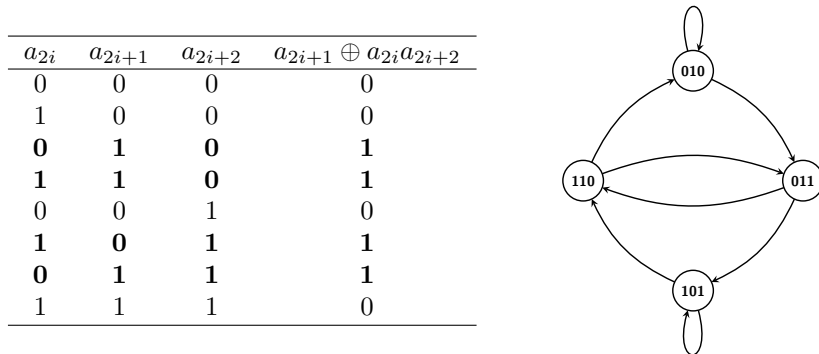


Figure 2: Truth table (left) and de Bruijn graph (right) of  $a_{2i+1} \oplus a_{2i}a_{2i+2}$ .

of the fourth block. Consequently, the local rule  $f : \mathbb{F}_2^9 \rightarrow \mathbb{F}_2$  is defined as  $f(x_1, \dots, x_9) = x_1 \oplus x_3 \oplus x_5 \oplus x_6 \oplus x_8 \oplus x_9$ , and by Theorem 2 the corresponding LBCA  $F : \mathbb{F}_2^{10} \rightarrow \mathbb{F}_2^2$  generates a 5-dimensional Latin hypercube of order 4.

## 4 Counting Sequences of Invertible Matrices

As we said above, the de Bruijn graph of the determinant function summarizes all the overlap relations between vectors of its support, which represent invertible Toeplitz matrices. Thus, we derived that Latin hypercubes of dimension  $k$  generated by LBCA correspond to *paths of length  $k - 3$*  over this graph:

**Theorem 3.** *Let  $b, k \in \mathbb{N}$ . Then,  $L_{b,k}$  equals the number of paths of length  $k - 3$  over the de Bruijn graph  $G_{det}$  of the determinant function  $det : \mathbb{F}_q^{2b-1} \rightarrow \mathbb{F}_q$ .*

Hence, in order to count the number of  $k$ -dimensional Latin hypercubes generated by LBCA, we need to look more closely at the properties of the de Bruijn graph  $G_{det}$  associated to the determinant of Toeplitz matrices. In particular, counting the number of paths of length  $k - 3$  over  $G_{det}$  requires characterizing the indegrees and outdegrees of its vertices. Looking at Figure 2 one can see that each vertex has two ingoing and two outgoing edges, hence the resulting de Bruijn graph  $G_{det}$  for  $b = 2$  and  $q = 2$  is *2-regular*. In the remainder of this section we will show that a regularity property holds in general for every  $b \geq 2$  and  $q$  power of a prime. We first need the following result:

**Theorem 4.** *Denote by  $T(n)$  the set of  $n \times n$  Toeplitz matrices over  $\mathbb{F}_q$ . Let  $A \in T(n)$  be strictly lower triangular, then there are exactly  $(q - 1)q^{n-1}$  upper triangular matrices  $B \in T(n)$  such that  $A + B$  is nonsingular.*

*Proof.* The result is clear if  $A$  is the all-zero matrix, therefore we assume henceforth that  $A$  is nonzero. We shall use the results by Daykin [2] on *persymmetric* (Hankel) matrices. Clearly, results on persymmetric matrices can be restated in terms of Toeplitz matrices, since the former are the transpose of the latter. Let  $H(b)$  be the set of  $n \times n$  persymmetric matrices. For any  $A \in H(b)$  with  $a_{i,j} = 0$  if  $i + j \geq b + 1$ , denote the number of matrices  $B \in H(b)$  with  $b_{i,j} = 0$  if  $i + j \leq n$  such that  $A + B$  is nonsingular as  $N(A)$ . We thus show that  $N(A) = (q - 1)q^{n-1}$ .

For any  $M \in H(n)$  and any  $m \leq n$ , we denote the matrix in  $H(m)$  consisting of the first  $m$  rows and  $m$  columns of  $M$  as  $M[m]$ . Let  $P \in H(m)$  be a nonzero matrix. Let  $R(P)$  be the number of nonsingular matrices  $Q \in H(2m)$  such that  $P = Q[m]$ , and for all  $m$  and  $i \leq m$ , let  $T(m : i) = (q - 1)^2 q^{2m-i-2}$  if  $i < m$  and  $T(m : i) = (q - 1)q^{m-1}$  if  $i = m$ . Theorem 3 in [2] gives the following formulas for  $R(P)$  and  $R'(P)$ :

$$R(P) = \begin{cases} \frac{T(2m-v:m-v)}{(q-1)q^{m-v-1}} & \text{if } v < m \\ q \sum_{i=1}^m T(m : i) & \text{otherwise} \end{cases}, \quad R'(P) = \begin{cases} \frac{T(2m-v:m-v)}{(q-1)q^{m-v-1}} & \text{if } v < m \\ q \sum_{i=1}^m T(m : i) & \text{otherwise} \end{cases}$$

We now show that for all nonzero  $P \in H(m)$ ,  $R(P) = (q - 1)q^{2m-1}$ . We prove the claim for  $R(P)$ . If  $v < m$ , we have

$$R(P) = \frac{T(2m-v : m-v)}{(q-1)q^{m-v-1}} = \frac{(q-1)^2 q^{2(2m-v)-(m-v)-2}}{(q-1)q^{m-v-1}} = (q-1)q^{2m-1}.$$

If  $v = m$  we have

$$\begin{aligned} R(P) &= q \sum_{i=1}^m T(m : i) = q(q-1) \left\{ (q-1)q^{m-1} \sum_{i=1}^{m-1} q^{m-1-i} + q^{m-1} \right\} = \\ &= q^m (q-1) \{ (q^{m-1} - 1) + 1 \} = (q-1)q^{2m-1}. \end{aligned}$$

With a similar argument, one can also show that  $R'(P) = (q - 1)q^{2m-1}$ . Now, let  $m = \lceil n/2 \rceil$  and  $P = A[m]$ . If  $n = 2m$  is even, then  $N(A) = R(P) = (q - 1)q^{n-1}$ . If  $n = 2m - 1$  is odd then  $P_{m,m} = 0$ , thus consider the matrices  $P^a$ , obtained by setting the value  $P_{m,m}^a = a$ , for all  $a \in \mathbb{F}_q$ . Then  $N(A) = \sum_{a \in \mathbb{F}_q} R(P^a) = (q - 1)q^{n-1}$ .  $\square$

Theorem 4 thus states that by fixing the leftmost  $b - 1$  entries of the vector  $(a_2, \dots, a_{2b})$  that defines a Toeplitz matrix  $A$ , one can complete the remaining  $b$  ones in  $(q - 1)q^{b-1}$  different ways so that the resulting Toeplitz matrix is invertible. This brings us to the following corollary:

**Corollary 3.** *Let  $\det : \mathbb{F}_q^{2b-1} \rightarrow \mathbb{F}_q$  be the determinant function associated to the set  $T(b)$ . Then, for any vector  $\tilde{a} \in \mathbb{F}_q^{b-1}$ , the restriction  $\det|_{\tilde{a}} : \mathbb{F}_q^b \rightarrow \mathbb{F}_q$  obtained by fixing either the leftmost or the rightmost  $b - 1$  coordinates to  $\tilde{a}$  is balanced, that is  $|\text{supp}(\det|_{\tilde{a}})| = (q - 1)q^{b-1}$ .*

We now show that this corollary implies the regularity of  $G_{\det}$ .

**Lemma 3.** *For any  $b \geq 2$  the de Bruijn graph  $G_{\det}$  of the determinant function  $\det$  is  $(q - 1)q^{b-1}$ -regular.*

*Proof.* As a preliminary remark, observe that by Theorem 1 the number of vertices in  $G_{\det}$  is  $|V| = (q - 1)q^{2(b-1)}$ . We prove only that the outdegree of each vertex is  $(q - 1)q^{b-1}$ , the indegree case following from a symmetrical reasoning. Let us fix the first  $b - 1$  coordinates  $a_2, \dots, a_b$  of  $\det$  to a vector  $\tilde{a} \in \mathbb{F}_q^{b-1}$ . Since the restriction  $\det|_{\tilde{a}}$  of  $b$  variables induced by  $\tilde{a}$  is balanced, it means that there is a set  $_{\tilde{a}}V = \{v \in V : (v_1, \dots, v_{b-1}) = \tilde{a}\}$  of  $(q - 1)q^{b-1}$  vertices in  $G_{\det}$  that begins by  $\tilde{a}$ . Each vertex in  $G_{\det}$  that ends by  $\tilde{a}$  has an outgoing degree of  $q(q - 1)^{b-1}$ , since it overlaps with all vertices in  $_{\tilde{a}}V$ . Let  $V_{\tilde{a}} = \{v \in V : (v_{b+1}, \dots, v_{2b-1}) = \tilde{a}\}$

be the set of all such vertices. By Corollary 3 the restriction  $det|_{\tilde{a}}$  obtained by fixing the *last*  $b - 1$  coordinates to  $\tilde{a}$  is also balanced. Thus, the cardinality of  $V_{\tilde{a}}$  is also  $(q - 1)q^{b-1}$ , meaning that there are  $(q - 1)q^{b-1}$  vertices ending by  $\tilde{a}$  that have outdegree  $(q - 1)q^{b-1}$ . Since this property holds for any vector  $\tilde{a} \in \mathbb{F}_q^{b-1}$ , it follows that there are  $q^{b-1} \cdot (q - 1)q^{b-1} = (q - 1)q^{2(b-1)} = |V|$  distinct vertices with outdegree  $(q - 1)q^{b-1}$ .  $\square$

Using Lemma 3, we can now determine what is the number of  $k$ -dimensional Latin hypercubes of order  $q^b$  generated by LBCA:

**Theorem 5.** *Let  $b, k \in \mathbb{N}$ , with  $k \geq 3$ . Then, the number of  $k$ -dimensional Latin hypercubes of order  $q^b$  generated by LBCA  $F : \mathbb{F}_q^{bk} \rightarrow \mathbb{F}_q^b$  with local rules  $f : \mathbb{F}_q^{b(k-1)+1} \rightarrow \mathbb{F}_q$  is  $L_{b,k,q} = (q - 1)^{k-2}q^{(k-1)(b-1)}$ .*

*Proof.* By Theorem 3  $L_{b,k,q}$  equals the number of paths of length  $k - 3$  over the de Bruijn graph  $G_{det}$ . We shall prove the result by induction on  $k$ .

For  $k = 3$ , the number of paths of length 0 over  $G_{det}$  obviously coincides with the number of vertices, which is  $(q - 1)q^{2(b-1)}$  by Corollary 1.

Assume now that  $k > 3$ , and let us consider  $L_{b,k+1,q}$ . Clearly, the paths of length  $k + 1$  are constructed by adding a new edge to all paths of length  $k$ , which by induction hypothesis are  $(q - 1)^{k-2}q^{(k-1)(b-1)}$ . Since  $G_{det}$  is  $(q - 1)q^{b-1}$ -regular, we thus have

$$L_{b,k+1,q} = (q - 1)^{k-2}q^{(k-1)(b-1)} \cdot (q - 1)q^{(b-1)} = (q - 1)^{k-1}q^{k(b-1)} .$$

$\square$

## 5 Conclusions

In this paper, we addressed the construction of Latin hypercubes generated by LBCA over the finite field  $\mathbb{F}_q$ , thereby taking a first step towards the generalization of the results in [5] about CA-based mutually orthogonal Latin squares. More precisely, we generalized the block construction of [5] to dimension  $k > 2$ , showing that the permutation property between any of the central  $k - 2$  blocks of the CA and the final configuration is related to the invertibility of the Toeplitz matrices defined by the central coefficients of the local rule. Moreover, we observed that the Toeplitz matrices associated to adjacent blocks share the coefficients respectively on the upper and lower triangulars, a property that can be described by the de Bruijn graph of the determinant function. We finally derived the number  $L_{b,k,q}$  of LBCA generating  $k$ -dimensional Latin hypercubes of order  $q^b$  by counting the number of paths of length  $k - 3$  over this de Bruijn graph, which we proved to be  $(q - 1)q^{b-1}$ -regular. The resulting formula shows that  $L_{b,k,q}$  is exponential both in the dimension and the block size of the hypercube, indicating that the family of Latin hypercubes generated by LBCA is quite large. We plan to study this rich structure of Latin hypercubes in future research, in particular by characterizing its mutually orthogonal subsets. Also, another interesting direction for future investigation is to assess whether Latin hypercubes can be generated by other types of CA, besides the bipermutive ones considered here.

### Acknowledgments.

This work has been partially supported by COST Action IC1405, Reversible Computation – Extending the Horizons of Computing.

### References

- [1] Blakley, G.R.: Safeguarding cryptographic keys. In: Managing Requirements Knowledge, International Workshop on. pp. 313–317 (1979)
- [2] Daykin, D.: Distribution of bordered persymmetric matrices in a finite field. *J. Reine Angew. Math.(Crelles J.)* **203**, 47–54 (1960)
- [3] García-Armas, M., Ghorpade, S.R., Ram, S.: Relatively prime polynomials and nonsingular Hankel matrices over finite fields. *J. Comb. Theory Ser. A* **118**(3), 819–828 (2011)
- [4] Herranz, J., Sáez, G.: Secret sharing schemes for  $(k, n)$ -consecutive access structures. In: CANS. LNCS, vol. 11124, pp. 463–480. Springer (2018)
- [5] Mariot, L., Gadouleau, M., Formenti, E., Leporati, A.: Mutually orthogonal latin squares based on cellular automata. *Des. Codes Cryptogr.* **88**(2), 391–411 (2020)
- [6] Mariot, L., Leporati, A.: Sharing secrets by computing preimages of bipermutive cellular automata. In: ACRI 2014. Proceedings. pp. 417–426 (2014)
- [7] Mariot, L., Leporati, A., Dennunzio, A., Formenti, E.: Computing the periods of preimages in surjective cellular automata. *Nat. Comput.* **16**(3), 367–381 (2017)
- [8] Mariot, L., Picek, S., Leporati, A., Jakobovic, D.: Cellular automata based s-boxes. *Cryptography and Communications* **11**(1), 41–62 (2019)
- [9] Shamir, A.: How to share a secret. *Commun. ACM* **22**(11), 612–613 (1979)
- [10] Stinson, D.R.: Combinatorial designs - constructions and analysis. Springer (2004)
- [11] Sutner, K.: De Bruijn graphs and linear cellular automata. *Complex Systems* **5**(1), 19–30 (1991)