



# An exploration of reversible septenary number-conserving cellular automata: a survey of known methods

Barbara Wolnik<sup>1,3</sup> · Adam Dzedzej<sup>1</sup> · Maciej Dziemiańczuk<sup>2</sup> · Aleksander Wardyn<sup>2</sup> · Bernard De Baets<sup>3</sup>

Accepted: 24 May 2023 / Published online: 20 June 2023  
© The Author(s) 2023

## Abstract

Little is known about the dynamics of  $k$ -ary (binary, ternary, quaternary, quinary, etc.) reversible number-conserving cellular automata. Here, we present some preliminary results in the case of seven states. In particular, we examine one of the most complex seven-state reversible and number-conserving rules and provide a full description of its dynamics.

**Keywords** Multi-state cellular automata · Number conservation · Reversibility

## 1 Introduction

A popular class of multi-state cellular automata (CAs) is the class of so-called  $k$ -ary CAs having  $\{0, 1, 2, \dots, k - 1\}$  as state set, for some natural number  $k$  greater than 1. Since each state is a nonnegative integer, it can, for instance, be interpreted as the number of particles occupying a given cell and, for this reason, these CAs are readily used for modeling various physical phenomena. Unfortunately, we cannot equip the state set  $\{0, 1, 2, \dots, k - 1\}$  with a nice mathematical structure as, for example, in the case of the ring  $\mathbb{Z}_k$ , to resort to methods that have been developed so far.

When physical phenomena governed by some conservation law (e.g., of mass or energy) are simulated, usually a special subclass of CAs is used: number-conserving CAs (NCCAs), i.e., CAs that preserve the sum of the states of all the cells upon every update (see, e.g., Durand et al. 2003). The second very desirable property of CAs used for modeling physical phenomena is reversibility, which ensures preservation of information. As a result, from the modeling point

of view,  $k$ -ary CAs that are both number conserving and reversible are the most interesting ones.

The dynamics of reversible  $k$ -ary NCCAs has been very little studied, even in the one-dimensional case. However, when the smallest possible radius of the neighborhood is considered, i.e., radius  $1/2$ , it is known that all reversible  $k$ -ary NCCAs can be seen as shift-identity product cellular automata (García-Ramos 2012). Of course, such CAs have a very simple dynamics and therefore their computing abilities are seriously limited. On the other hand, it has been shown that when the radius is increased to  $3/2$ , then it is possible to find a reversible  $k$ -ary NCCA that is computationally universal (Morita 2017).

The study of the dynamics of reversible  $k$ -ary NCCAs with radius 1 has gained momentum in recent years, in particular thanks to the establishment of complete lists of such CAs for  $k \in \{5, 6, 7\}$  (previously, complete lists were known only for  $k \leq 4$ ). These investigations revealed, *inter alia*, that for  $k = 7$  all reversible  $k$ -ary NCCAs (septenary NCCAs), except for shifts, have a finite order. More precisely, each of them repeats each configuration in a 60-cycle (Wolnik et al. 2022). Hence, their dynamics is definitely simpler than that of reversible  $k$ -ary NCCAs for  $k = 6$  (senary NCCAs), as most of the latter (306 out of 471) do not have a finite order. For example, this means that there is a particle interpretation of a reversible septenary NCCA, where each particle moves in a restricted space, while for reversible senary NCCAs the particles can move arbitrarily far (see Figs. 1a, b for a visual comparison of the dynamics of reversible senary and septenary NCCAs). It is worth pointing out here that Kari and Ollinger (2008) showed that, in general, the problem of

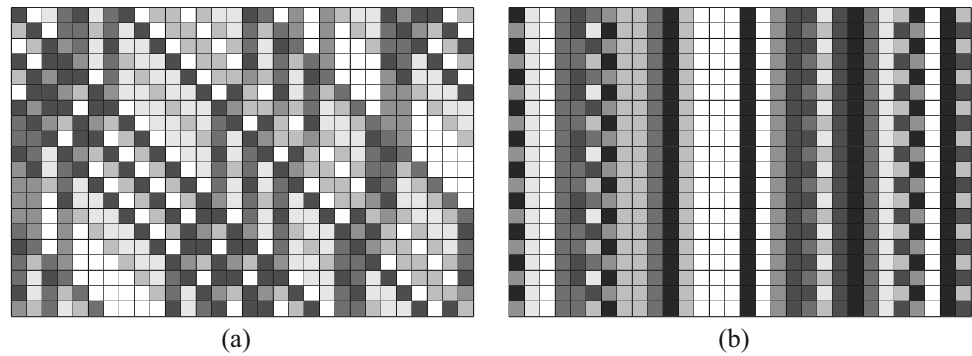
✉ Maciej Dziemiańczuk  
maciej.dziemianczuk@ug.edu.pl

<sup>1</sup> Institute of Mathematics, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, 80-308, Gdańsk, Poland

<sup>2</sup> Institute of Informatics, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, 80-308, Gdańsk, Poland

<sup>3</sup> KERMIT, Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Gent, Belgium

**Fig. 1** A sample space-time diagram of a reversible senary (a) and a reversible septenary (b) NCCA



determining if a given reversible one-dimensional cellular automata has a finite order (is periodic) is undecidable.

It is not known why the reversible septenary NCCAs with radius 1 are so simple, but intuition says that it has to do with the fact that 7 is a prime number. Indeed, if  $k$  is a composite number, then there exist methods (see, e.g., Wolnik et al. 2022) that allow to construct reversible  $k$ -ary NCCAs with radius 1 having infinite order (other than shifts). Hence, such a simple dynamics can happen only for prime  $k$ . Moreover, we observe such a simple dynamic for  $k \in \{2, 3, 5\}$  (see, for example, Wolnik et al. 2022 or Sect. 2.2). But does it happen for every prime  $k$ ? In an attempt to answer this question, we initiated the study of reversible septenary NCCAs and present our first observations.

## 2 Preliminaries

### 2.1 Basic definitions

We decided to only sketch the notations needed, as we hope that the readers are familiar with cellular automata. In the paper we consider only CAs with radius 1, even if we do not emphasize it in some places.

Let  $k$  be some natural number greater than 1 and let  $[0..k)$  denote the set  $\{0, 1, 2, \dots, k - 1\}$  (we adopt the notation of García-Ramos 2012). A one-dimensional  $k$ -ary CA with radius 1 is defined by a function  $f : [0..k)^3 \rightarrow [0..k)$ , referred as the local rule. For a natural number  $N$ , as cell space we will consider regular linear grids  $C_N = \{0, 1, \dots, N - 1\}$  with periodic boundary conditions or the infinite linear grid  $C_\infty = \mathbb{Z}$ .

A configuration on  $C_N$  is any mapping from the grid  $C_N$  to  $[0..k)$ . The set of all possible configurations on the grid  $C_N$  is denoted by  $X_N$  and is identified with  $[0..k)^N$ . The set of all configurations is denoted by  $X^*$ , i.e.,

$$X^* = \bigcup_{N=1}^{\infty} X_N = \bigcup_{N=1}^{\infty} [0..k)^N.$$

Analogously, an infinite configuration is any mapping from the grid  $C_\infty$  to  $[0..k)$ . The set of all possible infinite configurations is denoted by  $X_\infty$  and is identified with  $[0..k)^\mathbb{Z}$ .

For a given configuration  $\mathbf{x}$  (infinite or not), the value of cell  $n$  is denoted by  $x_n$  and  $\#(\mathbf{x})$  denotes the sum of the states in  $\mathbf{x}$ , i.e.,  $\#(\mathbf{x}) = \sum_{n \in C} x_n$ , where  $C$  stands for  $C_\infty$  or  $C_N$ . Additionally, in the case of  $\mathbf{x} \in X_\infty$ , we define the partial sum of  $\mathbf{x}$  between the index  $-n$  and the index  $n$  as  $\#_n(\mathbf{x}) = \sum_{i=-n}^n x_i$ .

Furthermore, by  $X_{\text{fin}}$  we denote the subset of  $X_\infty$  consisting of all infinite configurations with finite sum of states, or equivalently

$$X_{\text{fin}} = \{ \mathbf{x} \in X_\infty : \{n \in \mathbb{Z} : x_n \neq 0\} \text{ is finite} \},$$

while by  $X_{\text{per}}$  we denote the subset of  $X_\infty$  consisting of all periodic configurations, i.e.,

$$X_{\text{per}} = \{ \mathbf{x} \in X_\infty : \exists i \in \mathbb{N} \ \forall n \in \mathbb{Z} \ \mathbf{x}(n+i) = \mathbf{x}(n) \}.$$

A given local rule  $f$  induces the following three global functions:

- the global rule  $F : X^* \rightarrow X^*$ , which we identify with the cellular automaton and which is defined in the usual way: if  $\mathbf{x} \in X_N$ , then  $F(\mathbf{x}) \in X_N$  and  $F(\mathbf{x})_n = f(x_{n-1}, x_n, x_{n+1})$ , where, in view of periodic boundary conditions, all operations on the indices are performed modulo  $N$ ;
- the infinite global rule  $F_\infty : X_\infty \rightarrow X_\infty$  given for each  $\mathbf{x} \in X_\infty$  and  $n \in \mathbb{Z}$  by  $F_\infty(\mathbf{x})_n = f(x_{n-1}, x_n, x_{n+1})$ ;
- the block-mapping  $F_{\text{bl}} : \bigcup_{N=3}^{\infty} X_N \rightarrow X^*$  given for each  $N \in \mathbb{N}$  and  $\mathbf{x} \in X_{N+2}$  by

$$F_{\text{bl}}(x_0, x_1, \dots, x_{N+1}) = (f(x_0, x_1, x_2), f(x_1, x_2, x_3), \dots, f(x_{N-1}, x_N, x_{N+1})).$$

Additionally, one can consider the restriction of  $F_\infty$  to the set  $X_{\text{per}}$ , which we will denote by  $F$ . Of course  $F : X_{\text{per}} \rightarrow X_{\text{per}}$ . The differences between the definitions of  $F$ ,  $F_\infty$  and  $F_{\text{bl}}$  are shown in Fig. 2.

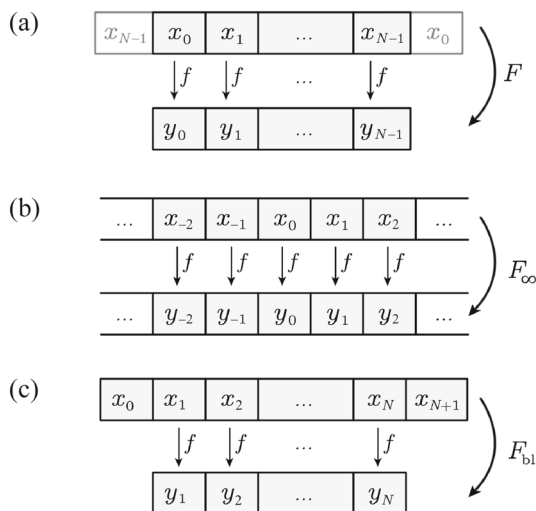


Fig. 2 Differences between the functions  $F$ ,  $F_\infty$  and  $F_{bl}$

Let us recall that each local rule  $f : [0..k]^3 \rightarrow [0..k]$  can be represented as a labelled graph with  $[0..k]^2$  as the set of vertices (the so-called de Bruijn representation). The set of edges is then defined as follows: for any  $a, b, c \in [0..k]$ , there is an edge  $abc$  from vertex  $ab$  to vertex  $bc$  and it is labelled by  $f(a, b, c)$  (e.g., Fig. 3 presents the de Bruijn representation of the elementary cellular automaton ECA110).

The de Bruijn representation allows to see configurations (infinite or not) as paths in this graph. Moreover, it enables a very simple way to read the value of the functions  $F$ ,  $F_\infty$  and  $F_{bl}$ . Indeed, for any configuration  $\mathbf{x} \in \bigcup_{N=3}^\infty X_N$ , the value  $F_{bl}(\mathbf{x})$  is given by a word consisting of labels of edges along the corresponding path  $x_0x_1x_2, x_1x_2x_3, \dots, x_{N-1}x_Nx_{N+1}$ . For any configuration  $\mathbf{x} \in X_N$  the value  $F(\mathbf{x})$  is given by a word consisting of labels of edges along the corresponding cycle  $x_{N-1}x_0x_1, x_0x_1x_2, \dots, x_{N-3}x_{N-2}x_{N-1}, x_{N-2}x_{N-1}x_0$ . Similarly, values of  $F_\infty$  are given by labels of edges along bi-infinite paths.

In our investigation, we focus on  $k$ -ary CAs that have two important (from the point of view of applications) properties. The first one is number conservation, which means that the sum of all states in any configuration remains constant throughout the evolution of the automaton.

**Definition 2.1** A global rule  $F : X^* \rightarrow X^*$  is number conserving if for all  $\mathbf{x} \in X^*$ , it holds that  $\#(F(\mathbf{x})) = \#(\mathbf{x})$ .

If a local rule  $f$  satisfies  $f(0, 0, 0) = 0$ , then it is called *legal* (or we say that 0 is a quiescent state). Obviously, if a global rule  $F$  is number conserving, then the corresponding local rule has to be legal, i.e., the following necessary condition holds (see, e.g., Boccara and Fukš 1998).

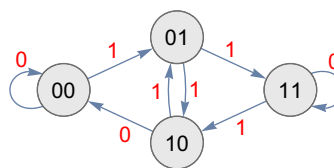


Fig. 3 The de Bruijn representation of the elementary cellular automaton ECA110

**Remark 2.2** Let  $f : [0..k]^3 \rightarrow [0..k]$  be a local rule and  $F : X^* \rightarrow X^*$  be the corresponding global rule. If  $F : X^* \rightarrow X^*$  is number conserving, then  $f(0, 0, 0) = 0$ .

In the case of the infinite grid  $C_\infty$ , the definition of the property of being number conserving is a little bit complicated.

**Definition 2.3** An infinite global rule  $F_\infty$  is number conserving if the following conditions hold:

- (1)  $f(0, 0, 0) = 0$ ,
- (2) for each non-zero infinite configuration  $\mathbf{x} \in X_\infty$ , 
$$\lim_{n \rightarrow \infty} \frac{\#_n(F_\infty(\mathbf{x}))}{\#_n(\mathbf{x})} = 1.$$

Fortunately, according to Proposition 1 in Dennunzio et al. (2013), these definitions are equivalent in the sense presented below.

**Remark 2.4** For a given local rule  $f$  the global rule  $F$  is number conserving if and only if the infinite global rule  $F_\infty$  is number conserving.

Note that for a legal local rule  $f$  the restriction of  $F_\infty$  to the set  $X_{fin}$ , which we will denote by  $F_{fin}$ , is an endomorphism, i.e.,  $F_{fin} : X_{fin} \rightarrow X_{fin}$ . Moreover, one-dimensional  $k$ -ary number-conserving global rules have a very simple characterization in terms of their local rules (see, e.g., Boccara and Fukš 1998).

**Theorem 2.5** Let  $f : [0..k]^3 \rightarrow [0..k]$  be a local rule and  $F : X^* \rightarrow X^*$  be the corresponding global rule. The global rule  $F$  is number conserving if and only if for any  $x, y, z \in [0..k]$  it holds that

$$f(x, y, z) = x + f(0, y, z) - f(0, x, y) + f(0, 0, y) - f(0, 0, x). \tag{1}$$

An equivalent formulation of the above condition is also

$$f(x, y, z) = z + f(x, y, 0) - f(y, z, 0) + f(y, 0, 0) - f(z, 0, 0). \tag{2}$$

The second property of  $k$ -ary CAs we are interested in is reversibility. Since  $X^*$  is a disjointive sum of finite sets  $X_N$ ,

bijection of  $F$  is equivalent to injectivity. We thus use this property as a definition.

**Definition 2.6** A global rule  $F : X^* \rightarrow X^*$  is reversible if  $F$  is an injection, i.e., for any  $\mathbf{x}_1, \mathbf{x}_2 \in X^*$  such that  $\mathbf{x}_1 \neq \mathbf{x}_2$ , it holds that  $F(\mathbf{x}_1) \neq F(\mathbf{x}_2)$ .

Of course, for any  $k > 1$  there are at least three global rules that are both number conserving and reversible: the identity rule  $f_{\text{Id}}(x, y, z) = y$ , the left-shift rule  $f_{\text{L}}(x, y, z) = z$  and the right-shift rule  $f_{\text{R}}(x, y, z) = x$ . We will call these rules *trivial*. For  $k < 4$ , there is no other reversible  $k$ -ary NCCA (even in a multi-dimensional case—see Wolnik and De Baets 2019, 2020). However, for  $k \geq 4$ , there are some non-trivial ones. Until advances were made recently, the complete lists of reversible  $k$ -ary NCCAs were known only for  $k \leq 4$ . First, the theory introduced in Wolnik et al. (2020) made it possible to find all quinary NCCAs ( $k = 5$ ) and then check which of them are reversible. Next, using the method described in Wolnik et al. (2022) it was possible to enumerate all reversible senary and septenary NCCAs, without listing all the NCCAs first. The cardinalities of the obtained sets are shown in Table 1 and the mentioned lists can be found in Dziemiańczuk et al. (2020).

The following definition is formulated in the language of group theory.

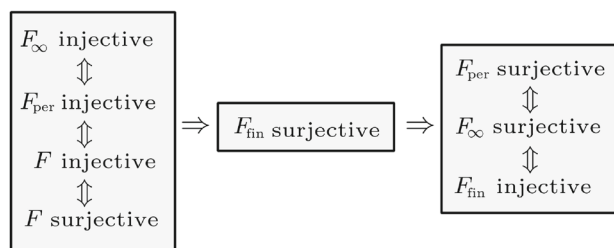
**Definition 2.7** Let  $F$  be a global rule. If there exists a natural number  $m$  such that the function  $F^m$  is the identity on  $X^*$ , then we say that  $F$  has finite order and we define the order of  $F$  as the smallest natural number with this property.

If a global rule  $F$  has a finite order, then it is not very interesting for applications as it repeats each configuration in an  $m$ -cycle, where  $m$  is the order of  $F$ . In particular, such a CA cannot be computationally universal (in any reasonable sense).

At the end of this section, in Fig. 4 we present the dependencies between the surjectivity and injectivity of  $F_\infty$ ,  $F$ ,  $F$  and  $F_{\text{fin}}$  (in the case of a legal local rule  $f$ ). Since we deal with one-dimensional CAs, all the dependencies between  $F_\infty$ ,  $F$  and  $F_{\text{fin}}$  are known (see, e.g., Hedlund 1969 or Kari 2005a), while the presented dependencies between  $F$  and  $F$  are obvious.

It is known that in the case of the infinite global rule  $F_\infty$  the concepts of injectivity, bijectivity and reversibility are equivalent (see, e.g., Kari 2005b). Thus, according to Remark 2.4, for a given local rule  $f$  the infinite global rule  $F_\infty$  is reversible and number conserving if and only if the global rule  $F$  is so.

Although in this paper we focus on the global rules  $F$  only, we present the other ones and the dependences between them, since we will use results from papers that consider infinite grids and, accordingly,  $F_\infty$ ,  $F_{\text{fin}}$  or  $F$  are used.



**Fig. 4** Implications between the injectivity and surjectivity properties of  $F_\infty$ ,  $F$ ,  $F$  and  $F_{\text{fin}}$ , in the case of a legal local rule  $f$

## 2.2 The description of reversible quinary NCCAs

Although our investigation concerns reversible septenary NCCAs, we start by recalling the description of the dynamics of reversible quinary NCCAs (since  $k = 5$  is also a prime number). The full description of all 21 such CAs is given in Wolnik et al. (2022), but we decided to present it here for two reasons: for the readers' convenience and to have the possibility to introduce the language of *swaps*.

Since the state set  $[0..5)$  is sufficiently rich, we can easily design a quinary CA that is both number conserving and reversible. Indeed, if we pick some swap  $(ab) \leftrightarrow (cd)$ , where  $a, b, c, d$  are different elements from  $[0..5)$  such that  $a + b = c + d$ , then we can define a CA of radius 1 by a simple relation: if in a given time step in a configuration there is the pattern 'ab', then in the next time step it will be replaced by the pattern 'cd', while each pattern 'cd' will be replaced by the pattern 'ab'. Such a CA is obviously number conserving (since we assume that  $a + b = c + d$ ) and reversible (since it has order 2, i.e., it is self-inverse). In the case of  $[0..5)$ , there are exactly 12 possible swaps, thus we can easily design 12 reversible quinary NCCAs using this method.

We will call two swaps  $(a_1b_1) \leftrightarrow (c_1d_1)$  and  $(a_2b_2) \leftrightarrow (c_2d_2)$  *grade-separated* if no two patterns  $(a_1b_1, c_1d_1, a_2b_2)$  and  $(c_2d_2)$  can overlap. Note that two grade-separated swaps  $(a_1b_1) \leftrightarrow (c_1d_1)$  and  $(a_2b_2) \leftrightarrow (c_2d_2)$  can coexist in the same CA. For example,  $(30) \leftrightarrow (12)$  and  $(14) \leftrightarrow (32)$  can coexist, i.e., we can define a reversible NCCA by the relation: each pattern '30', '12', '14', '32' is replaced in the next time step by the pattern '12', '30', '32', '14', respectively. Since in the case of the set  $[0..5)$  there are exactly six pairs of such swaps, we can design an additional six reversible quinary NCCAs.

It turned out that there exists no other non-trivial quinary reversible NCCA apart from the ones described above (Wolnik et al. 2022). Moreover, each of the non-trivial quinary reversible NCCAs acts as follows: each configuration is repeated every two time steps. The list of all reversible quinary NCCAs and their description in the language of swaps is given in Table 2.

**Table 1** Numbers of specific types of one-dimensional  $k$ -ary CAs with radius 1

$k$	All $k$ -ary CAs	$k$ -ary NCCAs	Reversible $k$ -ary CAs	Reversible $k$ -ary NCCAs
2	$2^{2^3} = 256$	5	6	3
3	$3^{3^3} \approx 7.6 \cdot 10^{12}$	144	1800	3
4	$4^{4^3} \approx 3.4 \cdot 10^{38}$	89 588	?	21
5	$5^{5^3} \approx 2.4 \cdot 10^{87}$	1 876 088 314	?	21
6	$6^{6^3} \approx 1.2 \cdot 10^{168}$	?	?	471
7	$7^{7^3} \approx 7.4 \cdot 10^{289}$	?	?	1669

**Table 2** The list of all reversible quinary NCCAs with radius 1

$F$	Description of $F$	$F$	Description of $F$	$F$	Description of $F$
1	The right-shift rule	8	(31) $\leftrightarrow$ (40), (32) $\leftrightarrow$ (41)	15	(04) $\leftrightarrow$ (13), (14) $\leftrightarrow$ (23)
2	(12) $\leftrightarrow$ (30)	9	The identity rule	16	(04) $\leftrightarrow$ (31)
3	(12) $\leftrightarrow$ (30), (14) $\leftrightarrow$ (32)	10	(32) $\leftrightarrow$ (41)	17	(03) $\leftrightarrow$ (12)
4	(21) $\leftrightarrow$ (30), (31) $\leftrightarrow$ (40)	11	(23) $\leftrightarrow$ (41)	18	(03) $\leftrightarrow$ (12), (04) $\leftrightarrow$ (13)
5	(21) $\leftrightarrow$ (30)	12	(14) $\leftrightarrow$ (23)	19	(03) $\leftrightarrow$ (21)
6	(13) $\leftrightarrow$ (40)	13	(14) $\leftrightarrow$ (32)	20	(03) $\leftrightarrow$ (21), (23) $\leftrightarrow$ (41)
7	(31) $\leftrightarrow$ (40)	14	(04) $\leftrightarrow$ (13)	21	The left-shift rule

The global rule of a given CA is described in terms of swaps, where a swap  $(ab) \leftrightarrow (cd)$  means that every pattern ‘ $ab$ ’ in the subsequent time step is replaced by the pattern ‘ $cd$ ’ and vice versa

**Table 3** The number of reversible septenary NCCAs with radius 1 that have order  $m$  (Table 6 in Wolnik et al. (2022))

$m$	1	2	3	4	6	12	30	60
	1	634	72	324	540	8	60	28

### 3 An exploration of the dynamics of reversible septenary NCCAs

We have been able to find all reversible NCCAs with state set [0..7] and radius 1 (Wolnik et al. 2022). It turns out that there are as many as 1669 of them. It is not possible to list all of them in this paper, but the complete list can be found in Dziemiańczuk et al. (2020). A computational study of the obtained CAs has shown that all reversible septenary NCCAs have a very limited dynamical behavior: for each global rule  $F$ , except for two shifts, there exists  $m \in \{1, 2, 3, 4, 6, 12, 30, 60\}$  such that  $F^m$  is the identity rule (see Table 3), i.e., all of them have finite order.

Additionally, 1249 reversible septenary NCCAs have an inverse CA with radius 1 and only 420 rules have an inverse CAs with radius 2. Moreover, if  $F$  is the global rule of a CA with radius 1, then its  $m$ -th power  $F^m$  can even have radius  $m$ . For non-trivial reversible septenary NCCAs, however, we have found that the actual radius of any power is at most 2 (thanks to this property, the calculation of  $F^{60}$  was possible). More general approach to determining the minimal radius of an inverse CA was studied by Czeizler (2004).

Unfortunately, the language of swaps used for quinary CAs proved to be insufficient to describe the rules with seven states: only 634 of them allow for a description using swaps only. These are exactly the ones having order 2.

In [0..7] we can consider longer pattern cycles than just swaps (one can see a swap  $(ab) \leftrightarrow (cd)$  as a pattern 2-cycle  $(ab) \rightarrow (cd) \rightarrow (ab)$ ). There are 72 new global rules that can be described by using pattern 3-cycles  $(ab) \rightarrow (cd) \rightarrow (ef) \rightarrow (ab)$ , where  $\{a, b, c, d, e, f\} \subset \{0, 1, 2, 3, 4, 5, 6\}$  and  $a + b = c + d = e + f \in \{6, 7, 8\}$ . Obviously, these global rules have order 3 and their inverse CA has radius 1, since it can be described by the reverse pattern 3-cycles (the reverse pattern 3-cycle of  $(ab) \rightarrow (cd) \rightarrow (ef) \rightarrow (ab)$  is, of course,  $(ab) \rightarrow (ef) \rightarrow (cd) \rightarrow (ab)$ ).

Considering the combination of swaps and pattern 3-cycles (taking into account grade separation), we obtain another set of 540 global rules. These rules have order 6 and their inverse rules are to be found among them (in particular the inverse is of radius 1).

There are 420 remaining rules that do not allow for a simple description as above. They need to be more complicated as their orders are 4, 12, 30 or 60. Moreover, we know that their inverses have radius 2. We set forth to create another description of these rules in order to understand them better and explain their order.

It is known that there always exists a particle representation of an NCCA, however, it is usually not unique (see, e.g., Boccara and Fukś 2006 or Pivato 2002). We decided to base our particle representation on 2-cell patterns, which



could be seen as a generalization of the descriptions in the language of swaps and pattern 3-cycles.

Consider a global rule  $F$ . We will write the arrow  $a \xrightarrow{x} b$ , where  $a, b \in [0..7]$  and  $x > 0$ , to indicate that  $F$  acts as follows: exactly  $x$  particles from state  $a$  move to the right, whenever the cell on the right is in state  $b$  (and analogously  $a \xleftarrow{x} b$ ). For example, a swap  $(ab) \leftrightarrow (cd)$ , with  $a > c$ , can be described as a pair of arrows  $a \xrightarrow{x} b$  and  $c \xleftarrow{x} d$ , where  $x = a - c$ . Similarly, for  $a < c$ , we get  $c \xrightarrow{x} d$  and  $a \xleftarrow{x} b$ , where  $x = c - a$ . The set of all arrows for the global rule  $F$  will be denoted as  $\alpha(F)$ . Note that the corresponding local rule  $f$  can be obtained by the following simple formula:

$$f(a, b, c) = b + x + y - z - u,$$

where  $\alpha(F)$  contains the arrows  $a \xrightarrow{x} b, b \xleftarrow{y} c, a \xleftarrow{z} b$  and  $b \xrightarrow{u} c$  (if  $\alpha(F)$  does not contain any of these arrows, then we put 0 as the corresponding term). This approach also allows us to introduce a partial order on the set of all reversible septenary NCCAs:  $F_1 \leq F_2$  if and only if  $\alpha(F_1) \subseteq \alpha(F_2)$ . Although this partial order is quite sparse (as there are as many as 416 maximal elements), it seems to be a decent measure of the complexity of the considered global rules.

We decided to choose one of the most complex rules to carry out a more comprehensive study and the choice fell on Rule2 (we use this name because this rule has number 2 in the dataset in Dziemiańczuk et al. (2020)), since (i) it is a maximal element in the partial order, (ii) it is one of 28 global rules having order 60 and (iii) its inverse has radius 2. Below we present the lookup table of Rule2 as the sequence of  $7^3$  values  $f(0, 0, 0), f(0, 0, 1), f(0, 0, 2), \dots, f(6, 6, 6)$  with additional spaces after each seven values for the reader's convenience:

```
0000000 1131313 2222622 1353135 4444444 1555355 2666666
0000000 1131313 0000400 1353135 2222222 1555355 0444444
0000000 1131313 2222622 1353135 0000000 1555355 2666666
2222222 1131313 0000400 1353135 6666666 1555355 0444444
0000000 1131313 2222622 1353135 4444444 1555355 2666666
4444444 1131313 2222622 1353135 6666666 1555355 2666666
4444444 1131313 2222622 1353135 4444444 1555355 2666666
```

A detailed study of Rule2 allows to explain where the order 60 originates from and fully describe its dynamics. More precisely, the dynamics of Rule2 can be described easily in terms of the rows of the following seven matrices:

$$A = \begin{bmatrix} 3 & 6 & 0 \\ 5 & 0 & 4 \\ 1 & 4 & 4 \\ 3 & 2 & 4 \\ 5 & 4 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 6 & 0 \\ 3 & 0 & 4 \\ 1 & 2 & 4 \\ 3 & 4 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 2 \\ 5 & 0 \\ 1 & 4 \end{bmatrix},$$

$$D = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}, \quad E = \begin{bmatrix} 1 & 6 \\ 3 & 4 \end{bmatrix}, \quad F = \begin{bmatrix} 2 & 4 \\ 6 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 3 & 6 \\ 5 & 4 \end{bmatrix}. \quad (3)$$

Indeed, let any initial configuration  $\mathbf{x} \in X^*$  be given. We can cut it into small pieces (with a length of at most three) according to the following procedure. First, we find and mark all positions in  $\mathbf{x}$  at which there is some row of  $A$  or  $B$ . Next, in parts of  $\mathbf{x}$  not marked at this time, we find and mark all positions at which there is some row of  $C, D, E, F$  or  $G$ . Finally, all unmarked parts of  $\mathbf{x}$  are cut into one-cell pieces. It turns out that Rule2 acts on each of the obtained pieces separately. The pieces with length one remain unaltered. The pieces with length two or three change cyclically, according to the cycle given by the appropriate matrix. Below we describe this more formally.

For  $M \in \{A, B, C, D, E, F, G\}$ , let  $M$  denote the number of rows of  $M$  and for  $i \in \{1, 2, \dots, M\}$ , let  $M_i$  denote the  $i$ -th row of  $M$ . Then in each time step, Rule2 acts as follows:

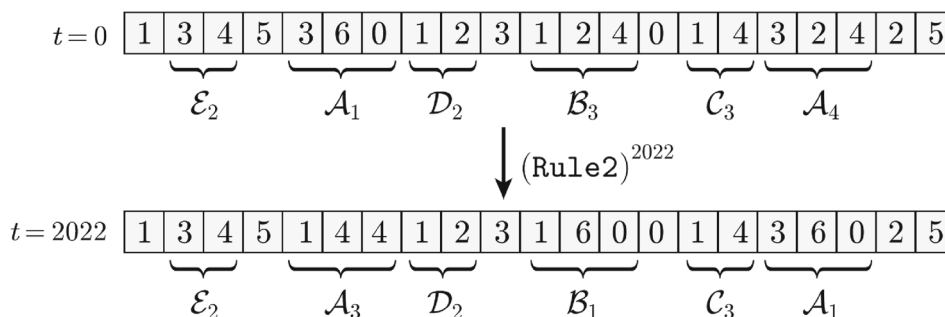
$$M_i \xrightarrow{(\text{Rule2})^t} M_{\text{mod}(i+t, M)},$$

where  $\text{mod}(t, m)$  denotes the remainder of the division of  $t$  by  $m$ . Since 60 is the least common multiple of 5, 4, 3 and 2, it is the order of Rule2. Figure 5 illustrates the above description on a sample configuration.

Following this description, it is easily to see that any power of Rule2 has radius at most 2. Indeed, the  $m$ th power of Rule2 acts as follows:  $M_i \xrightarrow{(\text{Rule2}^m)^t} M_{\text{mod}(i+mt, M)}$ . Since each row of  $M \in \{A, B, C, D, E, F, G\}$  has a length of at most three, so the radius 2 is sufficient to define Rule2<sup>m</sup>.

We are able to theoretically prove all the facts about Rule2 presented above, however, our proof is strongly based on properties of Rule2's lookup table, thus it cannot be easily adapted to other rules. In the future, more general

**Fig. 5** Illustration on a sample configuration with periodic boundary conditions of how Rule2 acts in terms of the rows of the matrices listed in (3)



methods to discover the dynamics of reversible septenary NCCAs would need to be found. Nevertheless we decided to present our current version of the proof for the readers' convenience in Appendix A.

### 4 The neighborhood scope of inverse automata

The description of the inverse rule by indicating its radius expressed as a natural number, as was done in the previous section, is very imprecise. A more detailed description uses the notion of *scope* of the neighborhood (for a given local rule  $f$ , the scope of its neighborhood means the number of consecutive cells whose states are needed to calculate the values of  $f$ ). Indeed, suppose that we consider a local rule  $f$  with radius  $r$ , i.e.,

$$F(\mathbf{x})_n = f(x_{n-r}, \dots, x_0, \dots, x_{n+r}).$$

It may happen that  $f$  does not depend on  $x_{n+r-1}$  and  $x_{n+r}$ . In such case, the scope of the neighborhood is definitely smaller than  $2r + 1$ . For example, the left-shift rule and the right-shift rule both have radius 1, but the scope of their neighborhood is not 3, but only 1. In this section we present a very useful tool introduced in Nasu (1977) for surjective legal CAs that allows to easily determine the scope of the neighborhood of the inverse automaton (see Theorem 4.9 below).

#### 4.1 Bundle graphs for reversible NCCAs

The theory presented in Nasu (1977) has been developed for the general case of an arbitrary surjective  $F_\infty$  (in particular, there is no condition on the radius). According to Fig. 4, we can use this theory in the case of a reversible  $F$ . Below we recall the definitions and facts from Nasu (1977) reformulated to the present setting, i.e., with state set  $[0..k]$  and radius equal to 1. Note that we can use this theory to study reversible NCCAs, since they are surjective (see Fig. 4) and their local rules are legal (which is additionally required for some theorems in Nasu (1977)).

For a given  $\mathbf{x} \in X_{N+2}$ , let us define the left end and the right end of  $\mathbf{x}$  as  $l(\mathbf{x}) = x_0x_1$  and  $r(\mathbf{x}) = x_Nx_{N+1}$ , respectively. Using the de Bruijn representation, we can say that the path corresponding to the configuration  $\mathbf{x}$  starts at the vertex  $x_0x_1$  and ends at the vertex  $x_Nx_{N+1}$ .

**Definition 4.1** Let  $f$  be a local rule,  $u \in [0..k]^2$  and  $\mathbf{x} \in X_N$ . The right bundle  $R_f(u, \mathbf{x})$  is the set of all  $v \in [0..k]^2$  such that there exists  $\mathbf{y} \in X_{N+2}$  with  $l(\mathbf{y}) = u$ ,  $r(\mathbf{y}) = v$  and  $F_{bl}(\mathbf{y}) = \mathbf{x}$ .

In the language of the de Bruijn representation, the right bundle  $R_f(u, \mathbf{x})$  is the set of all vertices  $v$  such that there exists a path  $\mathbf{y}$  from the vertex  $u$  to  $v$  satisfying  $F_{bl}(\mathbf{y}) = \mathbf{x}$ . To illustrate this definition, let us look at Fig. 3 showing the de Bruijn representation of the elementary cellular automaton ECA110. One can see, for example, that  $R_f(00, 1) = \{01\}$ ,  $R_f(00, 10) = \emptyset$ , and  $R_f(01, 111) = \{01, 11\}$ .

Analogously, we define the left bundle  $L_f(u, \mathbf{x})$  as the set of all vertices  $v$  such that there exists a path  $\mathbf{y}$  from the vertex  $v$  to  $u$  satisfying  $F_{bl}(\mathbf{y}) = \mathbf{x}$ .

**Definition 4.2** Let  $F$  be a local rule,  $u \in [0..k]^2$  and  $\mathbf{x} \in X_N$ . The left bundle  $L_f(u, \mathbf{x})$  is the set of all  $v \in [0..k]^2$  such that there exists  $\mathbf{y} \in X_{N+2}$  with  $l(\mathbf{y}) = v$ ,  $r(\mathbf{y}) = u$  and  $F_{bl}(\mathbf{y}) = \mathbf{x}$ .

For example, for the local rule  $f$  of ECA110, we have  $L_f(00, 1) = \emptyset$ ,  $L_f(00, 10) = \{01, 11\}$ , and  $L_f(01, 111) = \{00, 01, 10\}$ .

Note that for any  $u \in [0..k]^2$  and  $\mathbf{x} \in X_N$ , the cardinality of  $R_f(u, \mathbf{x})$  does not exceed  $k^2$  (since  $R_f(u, \mathbf{x}) \subseteq [0..k]^2$ ). Thus, we can define *maximal* right bundles as those that have the maximum cardinality among all the right bundles of  $f$ . Analogously,  $L_f(u, \mathbf{x})$  is said to be *maximal* if it has the maximum cardinality among all the left bundles of  $f$ . Let  $\mathcal{R}_f$  and  $\mathcal{L}_f$  denote the set of all maximal right bundles and all maximal left bundles, respectively.

The following theorem collects essential facts about the sets  $\mathcal{R}_f$  and  $\mathcal{L}_f$  proved in Nasu (1977).

**Theorem 4.3** Let  $F$  be a reversible  $k$ -ary CA.

- (P1) If  $R_f(u, \mathbf{x}) \in \mathcal{R}_f$ , then  $R_f(u, \mathbf{xy}) \in \mathcal{R}_f$ , for any  $y \in [0..k]$ .

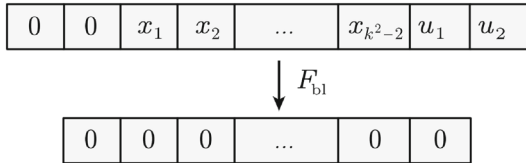


Fig. 6 The example used in the proof of Lemma 4.4

- (P2) Let  $R_f(u_1, \mathbf{x}_1), R_f(u_2, \mathbf{x}_2) \in \mathcal{R}_f$ . If  $R_f(u_1, \mathbf{x}_1) = R_f(u_2, \mathbf{x}_2)$ , then  $R_f(u_1, \mathbf{x}_1 y) = R_f(u_2, \mathbf{x}_2 y)$ , for any  $y \in [0..k)$ .
- (P3) The right bundle  $R_f(00, 0^{k^2})$  is maximal.
- (P1') If  $L_f(u, \mathbf{x}) \in \mathcal{L}_f$ , then  $L_f(u, y\mathbf{x}) \in \mathcal{L}_f$ , for any  $y \in [0..k)$ .
- (P2') Let  $L_f(u_1, \mathbf{x}_1), L_f(u_2, \mathbf{x}_2) \in \mathcal{L}_f$ . If  $L_f(u_1, \mathbf{x}_1) = L_f(u_2, \mathbf{x}_2)$ , then  $L_f(u_1, y\mathbf{x}_1) = L_f(u_2, y\mathbf{x}_2)$ , for any  $y \in [0..k)$ .
- (P3') The left bundle  $L_f(00, 0^{k^2})$  is maximal.

Note that in the above theorem, it is not required that  $F$  is number conserving. The following lemma shows the additional benefit of number conservation.

**Lemma 4.4** Let  $F$  be a reversible  $k$ -ary NCCA. Then  $R_f(00, 0^{k^2}) = R_f(00, 00)$  and  $L_f(00, 0^{k^2}) = L_f(00, 00)$ . In particular, both  $R_f(00, 00)$  and  $L_f(00, 00)$  are maximal.

**Proof** Let  $u_1 u_2 \in R_f(00, 0^{k^2})$ . Thus  $F_{bl}(00x_1x_2 \dots x_{k^2-2} u_1 u_2) = 0^{k^2}$  (see Fig. 6). We will show that all  $x_i$  have to be zero. If not, then there exists  $x_m$  with minimal index  $m$  such that  $x_m \neq 0$ . But then, according to Eq. (1), we have (if needed, we put  $x_{k^2-1} = u_1$  and  $x_{k^2} = u_2$ )

$$\begin{aligned}
 0 &= f(x_m, x_{m+1}, x_{m+2}) = x_m + f(0, x_{m+1}, x_{m+2}) \\
 &\quad - f(0, x_m, x_{m+1}) \\
 &\quad + f(0, 0, x_{m+1}) - f(0, 0, x_m) \\
 &= x_m + f(0, x_{m+1}, x_{m+2}) \\
 &\quad + f(0, 0, x_{m+1}),
 \end{aligned}$$

which means that all terms  $x_m, f(0, x_{m+1}, x_{m+2}), f(0, 0, x_{m+1})$  are zero. In particular,  $x_m = 0$  contrary to the assumption. As all  $x_i$  are zero, then both  $f(0, 0, u_1) = 0$  and  $f(0, u_1, u_2) = 0$ , so  $u_1 u_2 \in R_f(00, 00)$ . Thus  $R_f(00, 0^{k^2}) \subseteq R_f(00, 00)$ . Since  $R_f(00, 0^{k^2})$  is maximal, it holds that  $R_f(00, 0^{k^2}) = R_f(00, 00)$ . The proof of  $L_f(00, 0^{k^2}) = L_f(00, 00)$  is similar and is based on Eq. (2).  $\square$

Properties (P1) and (P2) allow to construct a directed graph  $\mathfrak{R}_f$  having  $\mathcal{R}_f$  as set of vertices and labelled arcs. Indeed, let  $R \in \mathcal{R}_f$ , i.e.  $R = R_f(u, \mathbf{x})$  for some  $u \in [0..k)^2$  and  $\mathbf{x} \in X_N$ , then for any  $y \in [0..k)$  we add an arc from  $R$  to  $R_f(u, \mathbf{x}y)$  labelled by  $y$ .

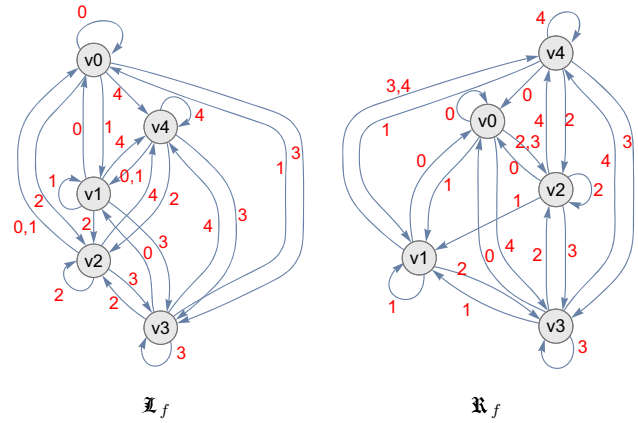


Fig. 7 Bundle graphs for the quinary NCCA with  $f$  being the local rule with index 18 listed in Table 2

Analogously, properties (P1') and (P2') allow to construct a directed graph  $\mathfrak{L}_f$  having  $\mathcal{L}_f$  as set of vertices and labelled arcs created as follows: if  $L = L_f(u, \mathbf{x}) \in \mathcal{L}_f$ , then for any  $y \in [0..k)$  we add an arc from  $L$  to  $L_f(u, y\mathbf{x}) \in \mathcal{L}_f$  labelled by  $y$ .

The following important theorem holds Nasu (1977).

**Theorem 4.5** Let  $F$  be a reversible  $k$ -ary CA. Both graphs  $\mathfrak{R}_f$  and  $\mathfrak{L}_f$  are strongly connected.

A useful consequence of the above theorem is the possibility of constructing the entire graph  $\mathfrak{R}_f$  (or  $\mathfrak{L}_f$ ) knowing at least one vertex of it. By Lemma 4.4, we know that the right bundle  $R_f(00, 00)$  is maximal, so it is one of the vertices of  $\mathfrak{R}_f$ . Moreover, the right bundle  $R_f(00, 00)$  is very simple to calculate. Thus, we can generate the entire graph  $\mathfrak{R}_f$  using Algorithm 1; the algorithm for generating the graph  $\mathfrak{L}_f$  is analogous.

Figure 7 shows the graphs  $\mathfrak{R}_f$  and  $\mathfrak{L}_f$  for the reversible quinary NCCA with  $f$  being the local rule with index 18 listed in Table 2. The vertices of these graphs are maximal (left and right, respectively) bundles and are listed below.

vertex in $\mathfrak{L}_f$	maximal left bundle	vertex in $\mathfrak{R}_f$	maximal right bundle
v0	{00, 10, 20, 30, 40}	v0	{00, 01, 02, 12, 13}
v1	{01, 11, 21, 31, 41}	v1	{03, 04, 10, 11, 14}
v2	{02, 03, 22, 32, 42}	v2	{20, 21, 22, 23, 24}
v3	{04, 12, 23, 33, 43}	v3	{30, 31, 32, 33, 34}
v4	{13, 14, 24, 34, 44}	v4	{40, 41, 42, 43, 44}

It appears that for all reversible quinary NCCAs except for the shifts (i.e., the quinary NCCAs listed in Table 2 with indices 2–20), the bundle graphs have a similar structure. Each of them has five vertices representing pairwise disjoint sets. For the right-shift rule, the left bundle graph has 25



**Algorithm 1** Generating the right bundle graph  $\mathfrak{R}_F$

**Input:** a local rule  $f$  given as a lookup table (LUT).  
**Output:** right bundle graph  $\mathfrak{R}_F$  given as  $(V, E)$ , where  $V$  is the set of vertices,  $E$  is the set of labelled arcs represented as triples  $(u, v, z)$ , where  $(u, v)$  is a directed edge from vertex  $u$  to vertex  $v$  with label  $z$ .

```

1. Set  $V = \emptyset, E = \emptyset$ .
2. Construction of the first vertex  $v_0$  (using Lemma 4.4)
   Set  $v_0 = \emptyset$ .
   for  $x, y$  in  $\{0, 1, \dots, k - 1\}$  do
     if  $f(0, 0, x) = 0$  and  $f(0, x, y) = 0$  then
       Add  $(x, y)$  to  $v_0$ .
     end if
   end for
3. Calculate the remaining vertices and edges using the BFS algorithm
   starting at vertex  $v_0$  using the procedure NEWEDGE( $v, z$ ) which
   creates the outgoing edge with label  $z$  from the vertex  $v$ .
   procedure NEWEDGE( $v, z$ ):
     Set  $u = \emptyset$ .
     for  $(a, b)$  in  $v$  do
       for  $c$  in  $[0..k)$  do
         if  $f(a, b, c) = z$  then
           Add  $(b, c)$  to  $u$ 
         end if
       end for
     end for
     if  $u \notin V$  then
       Add vertex  $u$  to  $V$ 
     end if
     Add  $(v, u, z)$  to  $E$ 
   end procedure

```

vertices (each is a singleton), while the right bundle graph has only one vertex (equal to  $[0..k)^2$ ). For the left-shift rule, of course, it is the opposite.

**4.2 Definite finite automata**

Any labelled directed graph can be considered as a finite automaton. Let us recall the definition and some useful properties of this concept (see, e.g., Nasu 1977 for details).

**Definition 4.6** A finite automaton  $A$  is a triple  $\langle \Sigma, S, T \rangle$ , where  $\Sigma$  is an input alphabet,  $S$  is a finite set of states and  $T : S \times \Sigma \rightarrow S$  is a transition function.

Any transition function  $T$  can be extended in a natural way to  $\Sigma^*$ , where  $\Sigma^*$  is the set of all finite strings of elements of  $\Sigma$ , as follows:

$$T(s, \varepsilon) = s \quad \text{and} \quad T(s, \mathbf{x}y) = T(T(s, \mathbf{x}), y),$$

for any  $s \in S, \mathbf{x} \in \Sigma^*$  and  $y \in \Sigma$  (here  $\varepsilon$  denotes the string of length zero).

In the theory of finite automata, the notion of definiteness is of importance. A finite automaton  $A = \langle \Sigma, S, T \rangle$  is 0-definite if it has only one state, while it is  $d$ -definite, with

$d \geq 1$ , if for each string  $\mathbf{x} \in \Sigma^*$  of length  $d$  and any states  $s_1$  and  $s_2$ , it holds that  $T(s_1, \mathbf{x}) = T(s_2, \mathbf{x})$ , while there are two states  $s'_1$  and  $s'_2$  and a string  $\mathbf{y} \in \Sigma^*$  of length  $d - 1$  such that  $T(s'_1, \mathbf{y}) \neq T(s'_2, \mathbf{y})$ . An automaton  $A$  is definite if it is  $d$ -definite for some  $d \geq 0$ .

**Definition 4.7** Let  $A = \langle \Sigma, S, T \rangle$  be a finite automaton. Two states  $s_1, s_2 \in S$  are said to be one-equivalent if  $T(s_1, y) = T(s_2, y)$  for each  $y \in \Sigma$ . The fact that states  $s_1$  and  $s_2$  are one-equivalent is denoted by  $s_1 \simeq s_2$ .

Note that the relation  $\simeq$  is indeed an equivalence relation. Hence, starting from any finite automaton  $A = \langle \Sigma, S, T \rangle$ , we can define the contraction of  $A$  w.r.t.  $\simeq$ , which yields the finite automaton  $A/\simeq = \langle \Sigma, S/\simeq, T' \rangle$  where  $S/\simeq = \{[s] : s \in S\}$  is the quotient set of  $S$  w.r.t.  $\simeq$  (the set of all equivalence classes) and  $T'$  acts as follows:  $T'([s], y) = [T(s, y)]$ . If  $S$  contains at least two one-equivalent states, then the resulting finite automaton has less states, so it is easier to study. We will use the following facts concerning definite automata (Perles et al. 1963) (see also Nasu 1977).

**Theorem 4.8** Let  $A = \langle \Sigma, S, T \rangle$  be a finite automaton with at least two states. If  $A$  is  $d$ -definite, for some  $d \geq 1$ , then there exist two distinct one-equivalent states in  $S$ . Moreover, the contracted automaton  $A/\simeq = \langle \Sigma, S/\simeq, T' \rangle$  is  $(d - 1)$ -definite.

According to the above theorem, for any definite automaton  $A$  with at least two states, repeatedly applying contraction, we get a sequence of simpler and simpler finite automata:

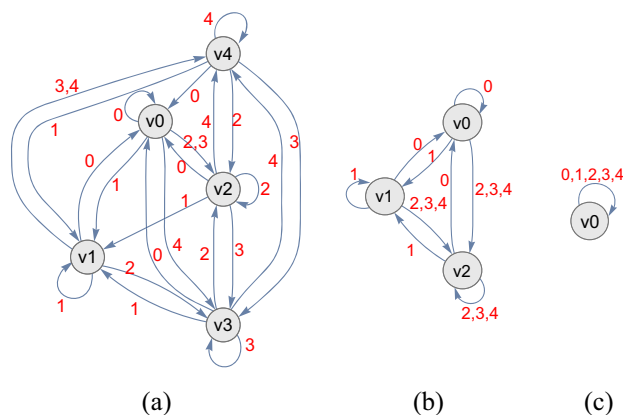
$$A_0, A_1, A_2 \dots, A_{d-1}, A_d,$$

where  $A_0 = A$ , for  $k \in \{0, 1, \dots, d - 1\}, A_{k+1} = A_k/\simeq$  and  $A_d$  has only one state. Moreover, since  $A_d$  is 0-definite, the initial finite automaton  $A$  then is  $d$ -definite.

**4.3 Definiteness of bundle graphs**

For a given  $k$ -ary global rule  $F$ , the graph  $\mathfrak{R}_f$  is a finite automaton  $\mathfrak{R}_f = \langle [0..k), \mathcal{R}_f, T \rangle$ , where the transition function  $T$  is determined by the arcs of  $\mathfrak{R}_f$  in the following way: for  $R_1, R_2 \in \mathcal{R}_f$  and  $y \in [0..k)$ , it holds that  $T(R_1, y) = R_2$  if and only if there is an arc in  $\mathfrak{R}_f$  from  $R_1$  to  $R_2$  labelled by  $y$ . The main result of Nasu (1977) states the following.

**Theorem 4.9** Let  $F$  be a reversible  $k$ -ary NCCA. There exist natural numbers  $d_r$  and  $d_l$  such that  $\mathfrak{R}_f$  is  $d_r$ -definite and  $\mathfrak{L}_f$  is  $d_l$ -definite. Moreover,  $d_r + d_l - 1$  is the least scope of the neighbourhood of the local rule  $g$  of the inverse automaton of  $F$ . More specifically,  $x_i = g(y_{i-(d_r-1)}, \dots, y_i, \dots, y_{i+(d_l-1)})$ .



**Fig. 8** **a** The right bundle graph  $\mathfrak{R}_f$  for the quinary NCCA with  $f$  being the local rule with index 18 listed in Table 2, **b** the contraction of the graph in **a**, **c** the contraction of the graph in **b**

Note that according to the above theorem, the neighborhood of the inverse automaton includes  $d_r - 1$  cells on the left, the cell itself and  $d_l - 1$  cells on the right. Hence, the radius of the inverse automaton equals  $\max(d_r, d_l)$ .

Using Theorem 4.8, we can calculate  $d_r$  and  $d_l$  for a given reversible  $k$ -ary NCCA. For example, let us start with the right bundle graph  $\mathfrak{R}_f$  given in Fig. 7. This graph has five vertices  $v_0, v_1, v_2, v_3$  and  $v_4$ . Among them, three vertices have the same set of outgoing arcs (w.r.t. the labels):  $v_2, v_3$  and  $v_4$  (see Fig. 8(a)). Therefore, we use the contraction to obtain a new graph with three vertices only, given in Fig. 8(b). These three vertices have the same set of outgoing arcs and contracting them we obtain a new graph with one vertex and one loop arc with all labels, as shown in Fig. 8(c). We applied contraction twice and therefore  $\mathfrak{R}_f$  is 2-definite.

To conclude our exploratory study of reversible septenary NCCAs, we have applied the above theorems to all 1669 rules. The results are gathered in Table 4. We were curious what the bundle graphs for reversible septenary NCCAs look like. In particular, we were interested whether their vertices are disjoint subsets of  $[0..7)$  or not. Additionally, we wanted to see if there were any dependencies between the description of reversible septenary NCCAs *via* bundle graphs and the previously considered descriptions in the terminology of patterns or arrows. As we have not observed any relationship with the number of arrows from  $\alpha(F)$  in the representation of the rule (we had expected such a dependence at the beginning of the study), we do not mention these numbers in the table.

Below we present the main results of our investigation, but, to make the description more clear, we exclude the trivial rules and focus on the remaining 1666 reversible septenary NCCAs (rows 4–7 in Table 4).

First of all, unlike in the quinary case, the bundle graphs for reversible septenary NCCAs need not have disjoint ver-

tices. However,  $\mathfrak{R}_f$  has disjoint vertices if and only if  $\mathfrak{R}_f$  is 2-definite. Moreover, all 1246 reversible septenary NCCAs that allow for a description in the language of pattern 2- and 3-cycles (and only these) have bundle graphs  $\mathfrak{R}_f$  and  $\mathfrak{K}_f$  with disjoint vertices. For this reason, both  $\mathfrak{R}_f$  and  $\mathfrak{K}_f$  are 2-definite, which agrees with the fact that their inverse automaton has radius 1, but now we have additional information that the scope of the neighborhood of their local rules is really equal to 3 (see row 4 in Table 4).

We already knew that the other 420 reversible septenary NCCAs are more complicated—in particular, their inverse automata have radius 2. It appeared that only for 72 of them the scope of the neighborhood really equals 5. Moreover, these CAs have only order 4, 12 or 60 (see row 7 in Table 4). The vast majority of reversible septenary NCCAs having an inverse automaton of radius 2 have to be, in a way, less complicated, since the scope of the neighborhood of their inverse automata is only 4. Note that these CAs have only order 4, 30 or 60 (see rows 5 and 6 in Table 4).

Our exploratory study suggests that there is hope to incorporate the property of number conservation in the theory developed in Nasu (1977) in order to make some general statements about the scope of inverse automata. The goal would be to find considerably sharper estimates for inverse automata of reversible NCCAs.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix A

Below we present the reasoning that leads to finding a description of Rule2 dynamics in terms of the matrices  $A, B, C, D, E, F, G$ . We start with the following definition.

**Definition A.1** Let  $F$  be the global rule of a reversible septenary NCCA and let  $a, b \in [0..7)$ . We will say that the string  $ab$  is a *left barrier* (resp. a *right barrier*) for  $F$ , if for each  $q \in [0..7)$  it holds that  $F(xab) = F(0ab)$  (resp.  $F(abx) = F(ab0)$ ).

Note that the barriers for a given global rule  $F$  are very easy to see in its lookup table, as long as it is presented as

**Table 4** Description of all 1669 reversible septenary NCCAs partitioned into seven classes according to values of  $d_l$  and  $d_r$

No	Defined numbers	Number of rules	Description
1	$d_l = 0, d_r = 2$	1	Left shift
2	$d_l = 1, d_r = 1$	1	Identity rule
3	$d_l = 2, d_r = 0$	1	Right shift
4	$d_l = 2, d_r = 2$	1246	Each rule has graphs $\mathfrak{L}_f$ and $\mathfrak{R}_f$ with 7 disjoint vertices. Each rule is described in the language of pattern 2-cycles and pattern 3-cycles. Each of these rules of order 2, 3 or 6
5	$d_l = 2, d_r = 3$	174	Each rule has a graph $\mathfrak{R}_f$ with 7 disjoint vertices. The vertices of the graph $\mathfrak{L}_f$ are not necessarily disjoint and there are 11, 13 or 14 vertices. Each of these rules is of order 4, 30 or 60
6	$d_l = 3, d_r = 2$	174	Each rule has a graph $\mathfrak{L}_f$ with 7 disjoint vertices. The vertices of the graph $\mathfrak{R}_f$ are not necessarily disjoint and there are 11, 13 or 14 vertices. Each of these rules is of order 4, 30 or 60
7	$d_l = 3, d_r = 3$	72	For each rule the vertices of both graphs $\mathfrak{L}_f$ and $\mathfrak{R}_f$ are not disjoint. The numbers of vertices are (11, 11), (12, 13) or (13, 12). Each of these rules is of order 4, 12 or 60

the one for Rule2. Indeed, each seven-element string represents the values:  $f(ab0), f(ab1), f(ab2), f(ab3), f(ab4), f(ab5), f(ab6)$  for some  $a, b \in [0..7]$ , thus if all elements of this string are the same, then  $ab$  is a right barrier for  $F$ . Similarly for columns: each homogeneous column indicates the existence of a left barrier. For example, Rule2 has the following barriers:

left:	10	11	12	13	14	15	16
	30	31	32	33	34	35	36
	50	51	52	53	54	55	56
right:	00	10	20	30	40	50	60
	04	14	24	34	44	54	64

The next definition highlights those parts of an initial configuration where the global rule  $F$  always behaves the same, no matter what the rest of the configuration looks like.

**Definition A.2** Let  $F$  be the global rule of a reversible septenary NCCA and let  $q_1q_2 \dots q_s \in [0..7]$  be a string of length  $s$ . We will say that  $F$  is locally number-conserving on  $q_1q_2 \dots q_s$ , if the following holds: there exists a string  $q'_1q'_2 \dots q'_s$  such that  $q_1 + q_2 + \dots + q_s = q'_1 + q'_2 + \dots + q'_s$  and as soon as the string  $q_1q_2 \dots q_s$  appears in a configuration,  $F$  always converts it to the string  $q'_1q'_2 \dots q'_s$ .

It appears that if  $ab$  and  $cd$  are a left and a right barrier for  $F$ , respectively, then  $F$  is locally number-conserving on each string starting with  $ab$  and ending with  $cd$ .

**Lemma A.3** Let  $F$  be the global rule of a reversible septenary NCCA and let  $q_1q_2 \dots q_s \in [0..7]$  be a string of length  $s \geq 2$ . If  $q_1q_2$  is a left barrier and  $q_{s-1}q_s$  is a right barrier for  $F$ , then  $F$  is locally number-conserving on  $q_1q_2 \dots q_s$ .

**Proof** If  $q_1q_2$  is a left barrier and  $q_{s-1}q_s$  is a right barrier for  $F$ , then  $F$  always converts  $q_1q_2 \dots q_s$  to  $q'_1q'_2 \dots q'_s = f(0, q_1, q_2)f(q_1, q_2, q_3) \dots f(q_{s-1}, q_s, 0)$ . Thus, for the configuration  $\mathbf{x} = q_1q_2 \dots q_s \in X_s$  we have  $F(q_1q_2 \dots q_s) = q'_1q'_2 \dots q'_s$ , which gives  $q_1 + q_2 + \dots + q_s = q'_1 + q'_2 + \dots + q'_s$ , since  $F$  is number-conserving.  $\square$

Let us note that the above lemma also holds for a very short string, for example for  $abc$  if  $ab$  is a left barrier and  $bc$  is a right barrier, or even for  $ab$  in the case when  $ab$  is both left and right barrier for  $F$ .

For a given matrix  $M$  let  $M$  denote the number of rows of  $M$  and for  $i \in \{1, 2, \dots, M\}$ , let  $M_i$  denote the  $i$ -th row of  $M$ .

**Theorem A.4** If  $M \in \{A, B\}$ , then Rule2 is locally number-conserving on each row of  $M$ . Moreover in each time step,

Rule2 acts as follows on a given row  $M_i$ :

$$M_i \xrightarrow{(\text{Rule2})^t} M_{\text{mod}(i+t,M)},$$

where  $\text{mod}(t, m)$  denotes the remainder of the division of  $t$  by  $m$ .

**Proof** Let us consider  $M = \mathcal{A}$ . First of all, let us note that Rule2 is locally number-conserving on the first row 360 (since 36 is a left barrier and 60 is a right barrier) as well as on each of the other rows. Moreover, it is easily to check that Rule2 converts  $q_1q_2q_3 = 360$  to  $q'_1q'_2q'_3 = 504, 504$  to 144, 144 to 324, 324 to 540 and 540 back to 360. According to Lemma A.3 the proof in the case of  $M = \mathcal{A}$  is finished. The case of  $M = \mathcal{B}$  is similar.  $\square$

The above theorem has a very significant impact on the form of space-time diagrams for Rule2. Indeed, let us consider some initial configuration  $\mathbf{x} \in X^*$ . If we find and mark all positions in  $\mathbf{x}$  at which there is some row of  $\mathcal{A}$  or  $\mathcal{B}$ , then Theorem A.4 describes the look of parts of the space-time diagram for  $\mathbf{x}$  below the marked parts: these will be columns of width 3 repeating some patterns in a cycle of length 5 or 4. Moreover, in the parts of  $\mathbf{x}$  not marked at this time, there is no strings identical with rows of  $\mathcal{A}$  or  $\mathcal{B}$ . Thus, to calculate  $F(\mathbf{x})$  for these parts, we do not need the values  $f(3, 6, 0)$ ,  $f(5, 0, 4)$  and so on. If so, then we can use the following lookup table (LUT\*):

---

0000000	1131313	2222622	1353135	4444444	1555355	2666666
0000000	1131313	0000*00	1353135	2222*22	1555355	*444444
0000000	1131313	2222622	1353135	0000000	1555355	2666666
2222*22	1131313	0000*00	1353135	*666666	1555355	*444444
0000000	1131313	2222622	1353135	4444444	1555355	2666666
4444*44	1131313	2222622	1353135	*666666	1555355	2666666
4444444	1131313	2222622	1353135	4444444	1555355	2666666

---

where by \* we denote all the unused values, so these places we can fill in any way, for example, keeping the seven-element strings or columns homogeneous. Note that after such a procedure, we get additional barriers for, so now the list of all barriers is as follows:

---

Left:	10	11	12	13	14	15	16
	30	31	32	33	34	35	36
	50	51	52	53	54	55	56
	24	60					
Right:	00	10	20	30	40	50	60
	04	14	24	34	44	54	64
	12	16	32	36			

---

The next theorem describes how Rule2 works on each of the unmarked segment of  $\mathbf{x}$  separately.

**Theorem A.5** Let  $\mathbf{x} \in X^*$  and let  $x_n \dots x_m$  be a segment of  $\mathbf{x}$  that contains no rows of  $\mathcal{A}$  or  $\mathcal{B}$ . On each row  $M_i$  of a matrix  $M \in \{\mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}\}$  occurring in the segment  $x_n \dots x_m$  Rule2 acts as follows:

$$M_i \xrightarrow{(\text{Rule2})^t} M_{\text{mod}(i+t,M)} \tag{4}$$

and leaves all other parts of the segment  $x_n \dots x_m$  unchanged.

**Proof** First of all, let us note that the formula given in Eq. (4) holds for  $t = 1$ . Indeed, is some row  $M_i$  of a matrix  $M \in \{\mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}\}$  occurs in the segment  $x_n \dots x_m$ , then to calculate Rule2 on  $M_i$  we can use LUT\*, since the segment  $x_n \dots x_m$  contains no rows of  $\mathcal{A}$  or  $\mathcal{B}$ . Note that each row  $M_i$  of  $\mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$  is both a left and a right barrier for LUT\*. Moreover, as one can easily check,  $M_i$  is converted to  $M_{\text{mod}(i+t,M)}$ .

Additionally, if we want to calculate Rule2 on the rest of the segment  $x_n \dots x_m$ , then we do not need the values  $f(q, a, b)$  and  $f(a, b, q)$ , where  $ab$  is any row of  $\mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$  and  $q \in [0..7)$ . Thus we can use the following lookup table:

```

0000000 11*1*1* 2222*22 *3*3*3* 4444444 *555*55 *666666
0000000 11*1*1* *555*55 *3*3*3* *555*55 *555*55 *666666
0000000 11*1*1* 2222*22 *3*3*3* *555*55 *555*55 *666666
*555*55 11*1*1* 0000*00 *3*3*3* *555*55 *555*55 *666666
0000000 11*1*1* 2222*22 *3*3*3* 4444444 *555*55 *666666
*555*55 11*1*1* 2222*22 *3*3*3* *555*55 *555*55 *666666
*555*55 11*1*1* 2222*22 *3*3*3* 4444444 *555*55 *666666

```

so, we see that  $\text{Rule}2$  acts on the rest of the segment  $x_n \dots x_m$  as the identity.

To the end, it is sufficient to show that  $F(\mathbf{x})_n \dots F(\mathbf{x})_m$  also contains no rows of  $\mathcal{A}$  or  $\mathcal{B}$ , since then the principle of mathematical induction will do the rest. But it is easily (though arduous) to check.  $\square$

## References

- Boccara N, Fukś H (1998) Cellular automaton rules conserving the number of active sites. *J Phys A Math Gen* 31(28):6007
- Boccara N, Fukś H (2006) Motion representation of one-dimensional cellular automaton rules. *Int J Mod Phys C* 17(11):1605–1611
- Czeizler E (2004) On the size of the inverse neighborhoods for one-dimensional reversible cellular automata. *Theor Comput Sci* 325(2):273–284. <https://doi.org/10.1016/j.tcs.2004.06.009>. (Theoretical Aspects of Cellular Automata)
- Dennunzio A, Formenti E, Provillard J (2013) Local rule distributions, language complexity and non-uniform cellular automata. *Theor Comput Sci* 504:38–51. <https://doi.org/10.1016/j.tcs.2012.05.013>
- Durand B, Formenti E, Róka Z (2003) Number-conserving cellular automata I: decidability. *Theor Comput Sci* 299(1):523–535
- Dziemiańczuk M, Dzedzej A, Wolnik B (2020) The complete lists of 1D reversible number-conserving cellular automata with radius one of up to 7 states. University of Gdańsk. <https://doi.org/10.34808/b8pn-1523>
- García-Ramos F (2012) Product decomposition for surjective 2-block NCCA. *Discrete Math Theor Comput Sci Proc Autom* 2011:147–158
- Hedlund GA (1969) Endomorphisms and automorphisms of the shift dynamical system. *Math Syst Theory* 3(4):320–375
- Kari J (2005) Theory of cellular automata: a survey. *Theor Comput Sci* 334(1):3–33. <https://doi.org/10.1016/j.tcs.2004.11.021>
- Kari J (2005) Reversible cellular automata. In: De Felice C, Restivo A (eds) *Developments in language theory*. Springer, Berlin, pp 57–68
- Kari J, Ollinger N (2008) Periodicity and immortality in reversible computing. In: Ochmański E, Tyszkiewicz J (eds) *Mathematical foundations of computer science 2008*. Springer, Berlin, pp 419–430
- Morita K (2017) One-dimensional universal reversible cellular automata, Chap. 11. Springer, Tokyo, pp 299–329. [https://doi.org/10.1007/978-4-431-56606-9\\_11](https://doi.org/10.1007/978-4-431-56606-9_11)
- Nasu M (1977) Local maps inducing surjective global maps of one-dimensional tessellation automata. *Math Syst Theory* 11(1):327–351
- Perles M, Rabin MO, Shamir E (1963) The theory of definite automata. *IEEE Trans Electron Comput EC* 12(3):233–243. <https://doi.org/10.1109/PGEC.1963.263534>
- Pivato M (2002) Conservation laws in cellular automata. *Nonlinearity* 15(6):1781–1793
- Wolnik B, De Baets B (2019) All binary number-conserving cellular automata based on adjacent cells are intrinsically one-dimensional. *Phys Rev E* 100:022126. <https://doi.org/10.1103/PhysRevE.100.022126>
- Wolnik B, De Baets B (2020) Ternary reversible number-conserving cellular automata are trivial. *Inf Sci* 513:180–189. <https://doi.org/10.1016/j.ins.2019.10.068>
- Wolnik B, Nenca A, Baetens JM, De Baets B (2020) A split-and-perturb decomposition of number-conserving cellular automata. *Phys D Nonlinear Phenom* 413:132645
- Wolnik B, Dziemiańczuk M, Dzedzej A, De Baets B (2022) Reversibility of number-conserving 1D cellular automata: unlocking insights into the dynamics for larger state sets. *Phys D Nonlinear Phenom* 429:133075. <https://doi.org/10.1016/j.physd.2021.133075>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.