Ternary reversible number-conserving cellular automata are trivial

Barbara Wolnik [a,*], Bernard De Baets [b]

[a] *Institute of Mathematics, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Gdańsk 80-308, Poland*
[b] *KERMIT, Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Coupure links 653, Gent B-9000, Belgium*

ARTICLE INFO

ABSTRACT

We introduce a novel method to study the reversibility of *d*-dimensional number-conserving multi-state cellular automata with the von Neumann neighborhood. We apply this method to ternary such cellular automata, for which, up to now, nothing was known about their reversibility. It turns out that they are all trivial: the only reversible such cellular automata are shifts that are intrinsically 1-dimensional.

© 2019 Elsevier Inc. All rights reserved.

1. Introduction

The reversibility of cellular automata (CAs) has received significant attention since John von Neumann and Stanisław Ulam introduced these discrete dynamical systems in the late 1940s. This is due to the fact that many researchers have further developed the theory of CAs to use them as a suitable computational tool for the simulation of physical systems. Thus, as reversibility is one of the fundamental laws of physics at the microscopic scale, especially in quantum mechanics, CAs should preferabley exhibit the same property.

The first studies on the reversibility of CAs are due to Hedlund [19] and Richardson [29]. In particular, they proved that a CA is reversible if and only if it is a injection. In the literature one can now find a number of results on this topic. Among them are results regarding infinite CAs as well as finite CAs with various boundary conditions. However, the studies on the reversibility of CAs concentrate on 1-dimensional ones and linear CAs with $\mathbb{Z}_m$ as state set.

When it comes to 1-dimensional CAs, it is safe to say that the problem has been completely resolved. Indeed, an effective way to determine reversibility of infinite 1-dimensional CAs was shown by Amoroso and Patt [1] and also by Di Gregorio and Trautteur [12] and Sutner [33]. For finite CAs, an algorithm to decide the reversibility was developed by Bhattacharjee and Das [3]. Moreover, there are many results concerning specific Elementary CAs (ECAs) (see, for example, [8,9,16,30]). However, it is worth emphasizing that existing tools do not allow to enumerate all reversible 1-dimensional CAs even in the case of few states (of course, except in the case of ECAs), because it is impossible to check all CAs with a given state set, due to the inhibitive cardinality of the search space.

---

* Corresponding author..
E-mail address: barbara.wolnik@mat.ug.edu.pl (B. Wolnik).

While 1-dimensional reversible CAs have been studied in detail and from many different points of view, the same cannot be said about 2- or higher-dimensional ones. This is a consequence of the negative result of Kari [22] stating that there is no algorithm that can decide whether or not an arbitrary 2-dimensional CA is reversible.

Another class of CAs (1- or higher-dimensional) for which researchers have found methods to determine reversibility, is the class of linear CAs (LCAs). The reason that LCAs have been extensively studied, especially with $\mathbb{Z}_m$ as state set, is that they are particularly amenable to theoretical analysis. Firstly, the linearity of local rules opens the door to the realm of linear algebra and one can use powerful tools like transition matrices, since an LCA is reversible if and only if its transition matrix is reversible (see, for example, [7,10,26,32]). Secondly, the fact that the state set is a finite commutative ring allows to use formal power series representations to state necessary and sufficient conditions for the reversibility of an LCA in terms of the coefficients of its local rule. This method has been initiated by Itô et al. [21] and intensively developed by other researchers (see, for example, [11,25,37]). Thus, also in the case of LCAs with the state set $\mathbb{Z}_m$, we can say that the problem of reversibility has been solved completely. In particular, LCAs have been thoroughly examined in terms of ergodic theory: it was shown that a reversible LCA is either a Bernoulli automorphism or non-ergodic [6]. More information about the reversibility problem of these types of CAs can be found in [23,24,27].

However, the transition matrix tool has fundamental limitations, since the matrix size depends on the number of cells, so, it is useless when we deal with infinite CAs and for finite CAs, every grid size has to be considered separately. Moreover, applying matrix theory for a ring $\mathbb{Z}_m$, which does not have to be a field, one has to be very careful (see [2]). Also, the linearity assumption as well as the use of a finite ring as state set are often not acceptable. For example, in particle motion modeling, one would like two plus one to be three instead of zero as in the case of the field $\mathbb{Z}_3$.

The aim of this paper is to present a method that allows to study the reversibility of another type of CAs: number-conserving cellular automata (NCCAs). NCCAs represent a particularly interesting class of CAs that have the special feature of preserving the sum of the states upon every update of the states (see, for example, [13]). As state set, we will consider $\{0, 1, \ldots, q_*\}$, where q_* is a given natural number. CAs with such a state set will be referred to as multi-state CAs. As we will deal with d-dimensional CAs, we have to define the cell neighborhood. We opt for the closed unit ball in the d-dimensional Manhattan distance metric, *i.e.*, the von Neumann neighborhood, the most popular neighborhood used in modeling physical phenomena (for example, it is "the most common in overland flow models" [5]). Unfortunately, studying multidimensional CAs with this kind of neighborhood is very complicated, because the von Neumann neighborhood is not a Cartesian product of one-dimensional neighborhoods (in contrast to the Moore neighborhood).

So far, there are only a few results concerning the reversibility of multi-state NCCAs. Schranko and De Oliveira [31] performed numerical experiments involving many rules to finally conjecture that the class of 1-dimensional such CAs is too restrictive to be computationally universal. This was proved in the case of radius 1/2 CAs [18], but it does not hold in general as shown by Morita [28]. Boccara and Fukś [4] found all 144 1-dimensional ternary NCCAs, but only three among them are reversible: the identity rule and two shift rules. Imai et al. [20] listed all reversible 1-dimensional NCCAs with radius 1 and up to four states through an exhaustive search and they also constructed some with five states. In the case of two dimensions, the situation does not look better. Dzedzej et al. [15], using a characterization of number-conserving local rules given in [35], found that for the state set $\{0, 1\}$ or $\{0, 1, 2\}$, *i.e.*, when we deal with binary or ternary 2-dimensional NCCAs with the von Neumann neighborhood, there are only trivial reversible ones, namely, the identity and the shifts (in each of the four possible directions). However, recently all four-state reversible 2-dimensional NCCAs with the von Neumann neighborhood have been enumerated [14], which until then was considered an intricate task because of two main reasons. Firstly, as we mentioned earlier, there is no algorithm to decide whether or not a given 2-dimensional CA is reversible. Secondly, there were no tools available to enumerate all four-state 2-dimensional NCCAs, so it was not possible to use the method of exhaustive search. The method used in [14] is based on the split-and-perturb decomposition of the local rule of an NCCA with the von Neumann neighborhood, proved in [36]. The same novel approach to study number conservation of CAs was used to find all binary d-dimensional NCCAs with the von Neumann neighborhood in [34], proving that for any given d there are exactly $4d + 1$ such cellular automata: the identity rule and the shift and traffic rules in each of the $2d$ possible directions. Thus, if we ask for reversible binary NCCAs in d dimensions, then only the identity rule and the shift rules remain, since the traffic rules are not reversible.

In this paper, we want to combine the ideas presented in [14] and [34] to create a method that allows to enumerate all reversible ternary d-dimensional NCCAs with the von Neumann neighborhood. It is known that the behavior of a CA can be quite different depending on whether we are considering finite, periodic or infinite configurations (see, for example, [23]). In our investigations, we adhere to the weakest version of CA reversibility: we consider a given CA on a fixed finite d-dimensional grid (with periodic boundary conditions) and we require its global function to be an injection. We show that there are only trivial reversible ternary NCCAs, namely, the identity and the shifts (in each of $2d$ possible directions). The proof of this fact is given in the case of a finite grid. However, our results also valid in the case of periodic and infinite CAs, as we discuss at the end of Section 3.

This paper is organized as follows. Section 2 contains preliminaries and a description of the main idea of the split-and-perturb decomposition of a number-conserving local rule. The formulation and the proof of the main theorem are given in Section 3. Conclusions, conjectures and open problems are discussed in Section 4.
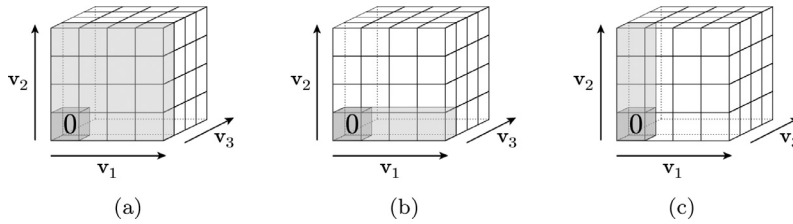
Fig. 1. Examples of subgrids: (a) [$\mathbf{v}_1$, $\mathbf{v}_2$], (b) [$\mathbf{0}$, $\mathbf{v}_1$], (c) [$\mathbf{v}_2$, -$\mathbf{v}_2$].

2. Preliminaries

In this section we recall some standard definitions about CAs and describe the approach to study number conservation presented in [36]: the split-and-perturb decomposition of a number-conserving local rule. We adhere to the notations introduced there, where also some explanations and examples are given, so, we present only a short description without details.

Let us fix the dimension $d \geq 1$. We consider the cellular space as a finite d-dimensional grid:

$$\mathcal{C} = \{0, 1, \ldots, n_1 - 1\} \times \{0, 1, \ldots, n_2 - 1\} \times \ldots \times \{0, 1, \ldots, n_d - 1\},$$

where the natural numbers $n_1, n_2, \ldots, n_d$ are greater than three. With this notation, each cell $\mathbf{i} \in \mathcal{C}$ is a d-tuple $(i_1, \ldots, i_d)$, where $i_k \in \{0, 1, \ldots, n_k - 1\}$.

For each $k \in \{1, 2, \ldots, d\}$, we define the vector $\mathbf{v}_k = (0, 0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^d$, where the kth component equals 1 and the other ones are 0. Let us denote the set of all directions as V_+, *i.e.*,

$$V_+ = \{\mathbf{v}_1, \text{-}\mathbf{v}_1, \mathbf{v}_2, \text{-}\mathbf{v}_2, \ldots, \mathbf{v}_d, \text{-}\mathbf{v}_d\}.$$

Additionally, let $\mathbf{0} = (0, 0, \ldots, 0) \in \mathbb{R}^d$ and $V = V_+ \cup \{\mathbf{0}\}$. We consider periodic boundary conditions, *i.e.*, the cell adjacent to the cell $\mathbf{i} = (i_1, \ldots, i_d)$ in the direction $\mathbf{v}_k$ is equal to $(i_1, \ldots, i_{k-1}, i_k + 1, i_{k+1}, \ldots, i_d)$, where the sum is understood modulo n_k. As a consequence, each cell in $\mathcal{C}$ has exactly $2d$ adjacent cells: one in each of the directions from V_+.

For each cell $\mathbf{i} \in \mathcal{C}$, its von Neumann neighborhood $P(\mathbf{i})$ consists of the cell itself and its $2d$ adjacent cells:

$$P(\mathbf{i}) = \mathbf{i} + V := \{\mathbf{i} + \mathbf{v} \mid \mathbf{v} \in V\}. \tag{1}$$

We use the following notation: for $\mathbf{i} \in \mathcal{C}$ and $\mathbf{v} \in V$, $\mathbf{i} + \mathbf{v} \in \mathcal{C}$ is nothing else but the cell adjacent to $\mathbf{i}$ in direction $\mathbf{v}$ if $\mathbf{v} \in V_+$, or $\mathbf{i}$ itself if $\mathbf{v} = \mathbf{0}$.

In the remainder of this paper, we consider only CAs with the von Neumann neighborhood, even if we do not always emphasize it.

If $\mathbf{u}$, $\mathbf{w}$ are any vectors from V (we exclude the case $\mathbf{u} = \mathbf{w} = \mathbf{0}$), then by [$\mathbf{u}$, $\mathbf{w}$] we denote a subset of the grid $\mathcal{C}$ defined as follows:

$$[\mathbf{u}, \mathbf{w}] = \{\mathbf{0} + k_1\mathbf{u} + k_2\mathbf{w} \mid k_1, k_2 \in \mathbb{Z}\}, \tag{2}$$

i.e., [$\mathbf{u}$, $\mathbf{w}$] is a 2-dimensional subgrid (or 1-dimensional, if the vectors $\mathbf{u}$ and $\mathbf{w}$ are linearly dependent) containing $\mathbf{0}$ and spanned by $\mathbf{u}$ and $\mathbf{w}$ (see Fig. 1).

Now, we define $\mathbf{\Omega}$ as a set containing all pairs $\{\mathbf{0}, \mathbf{v}\}$, where $\mathbf{v} \in V_+$, and all pairs $\{\mathbf{u}, \mathbf{v}\}$ for, $\mathbf{u}, \mathbf{v} \in V_+$ for which $\mathbf{u} \neq \mathbf{v}$ and $\mathbf{u} \neq \text{-}\mathbf{v}$. Because the pairs $\{\mathbf{u}, \mathbf{v}\}$ and $\{\mathbf{v}, \mathbf{u}\}$ are equal, the set $\mathbf{\Omega}$ contains exactly $2d^2$ elements. For a given pair $\{\mathbf{u}, \mathbf{v}\} \in \mathbf{\Omega}$, the pair $\{\text{-}\mathbf{u}, \text{-}\mathbf{v}\}$ is called the *matching pair*. If from each two matching pairs, we choose exactly one, then we get a set of *independent pairs*, which we denote by $\mathbf{\Lambda}$. We can construct $\mathbf{\Lambda}$ in 2^{d^2} ways, for example we can choose

$$\mathbf{\Lambda} = \bigcup_{i=1}^{d} \left\{\{\mathbf{0}, \mathbf{v}_i\}\right\} \cup \bigcup_{i=1}^{d-1} \bigcup_{j=i+1}^{d} \left\{\{\mathbf{v}_i, \mathbf{v}_j\}, \{\text{-}\mathbf{v}_i, \mathbf{v}_j\}\right\}, \tag{3}$$

but it always holds that $|\mathbf{\Lambda}| = d^2$.

In this paper, we consider ternary CAs, *i.e.*, CAs whose state set is $Q = \{0, 1, 2\}$. These CAs appear in a natural way when one wants to go beyond binary arithmetic or binary logic. A very good motivation for studying ternary CAs can be found, for example, in [17]. Unfortunately, a small step from "binary" to "ternary" entails an amazingly large increase in the number of possible local rules. For example, in the one-dimensional case there are only 256 binary CAs, while there are as many as $3^{27} \approx 7.6 \cdot 10^{12}$ ternary CAs. For this reason, there are almost no results concerning ternary CAs.

By a *configuration*, we mean any function from the grid $\mathcal{C}$ to $\{0, 1, 2\}$ and the set of all configurations is denoted by $X = \{0, 1, 2\}^{\mathcal{C}}$. We consider also *configurations in a wider sense*: mappings from the grid $\mathcal{C}$ to $\mathbb{R}$. The set of all configurations in a wider sense is, of course, a superset of X and is denoted by $\widetilde{X}$ (we will simply write configuration also for elements from $\widetilde{X}$, unless confusion is possible). The value of cell $\mathbf{i}$ in a configuration $\mathbf{x} \in \widetilde{X}$ is denoted by $\mathbf{x}(\mathbf{i})$ and we define the sum of the states in $\mathbf{x}$ as $\sigma(\mathbf{x}) = \sum_{\mathbf{i} \in \mathcal{C}} \mathbf{x}(\mathbf{i})$.

We need to denote a particular kind of subsets of X: for a given subgrid $[\mathbf{u}, \mathbf{w}]$, we denote the set of all configurations in X that are zero outside $[\mathbf{u}, \mathbf{w}]$ by $X_{[\mathbf{u},\mathbf{w}]}$. Furthermore, by X_{bin} we denote the set of all binary configurations, *i.e.*, the set $\{0, 1\}^{\mathcal{C}}$.

Let us recall that the neighborhood of each cell is labeled by the elements of the set V (see Eq. (1)). Thus it is natural to define a *neighborhood configuration* as a function $N: V \to \{0, 1, 2\}$ and we denote the set of all possible neighborhood configurations by $\mathcal{N}$. Some particular neighborhood configurations will be very important in the remainder of this paper. Firstly, for any $\mathbf{w} \in V$ and $q \in \{0, 1, 2\}$, we define a *monomer* $M_{\mathbf{w}:q}$, which takes the value q in the direction $\mathbf{w}$ and zero in the other directions, *i.e.*, for any $\mathbf{v} \in V$, it holds that

$$M_{\mathbf{w}:q}(\mathbf{v}) = \begin{cases} q, & \text{if } \mathbf{v} = \mathbf{w} \\ 0, & \text{if } \mathbf{v} \neq \mathbf{w}. \end{cases}$$

Note that if $q = 0$, then the neighborhood configuration $M_{\mathbf{w}:q}$ is a zero function, irrespective of $\mathbf{w}$, and we denote it also by M_0. Secondly, if $\{\mathbf{u}, \mathbf{w}\} \in \mathbf{\Omega}$ and $p, q \in \{0, 1, 2\}$, we introduce a *dimer* $D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}$ as the neighborhood configuration that takes the value p in the direction $\mathbf{u}$, the value q in the direction $\mathbf{w}$ and zero in the other directions, *i.e.*, for any $\mathbf{v} \in V$ it holds that

$$D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}(\mathbf{v}) = \begin{cases} p, & \text{if } \mathbf{v} = \mathbf{u} \\ q, & \text{if } \mathbf{v} = \mathbf{w} \\ 0, & \text{otherwise}. \end{cases}$$

Note that $D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}$ equals $D_{\substack{\mathbf{w}:q \\ \mathbf{u}:p}}$, and if $p = 0$ or $q = 0$, then $D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}$ is a monomer. As the pairs $\{\mathbf{u}, \mathbf{w}\}$ and $\{-\mathbf{u}, -\mathbf{w}\}$ are called matching, we also use this term to refer to the dimers $D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}$ and $D_{\substack{-\mathbf{u}:q \\ -\mathbf{w}:p}}$.

If the set of independent pairs $\mathbf{\Lambda}$ is chosen, we define the set of *generating dimers* as

$$\mathcal{D}_{\mathbf{\Lambda}} = \left\{ D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}} \,\middle|\, \{\mathbf{u}, \mathbf{w}\} \in \mathbf{\Lambda}, \ p, q \in \{1, 2\} \right\}.$$

It is easy to see that $\mathcal{D}_{\mathbf{\Lambda}}$ contains exactly $4d^2$ elements.

If $\mathbf{x} \in X$ and $\mathbf{i} \in \mathcal{C}$ are given, then $N_{\mathbf{x},\mathbf{i}}$ denotes the neighborhood configuration given by the von Neumann neighborhood of the cell $\mathbf{i}$ in the configuration $\mathbf{x}$, thus for $\mathbf{v} \in V$ we have $N_{\mathbf{x},\mathbf{i}}(\mathbf{v}) = \mathbf{x}(\mathbf{i} + \mathbf{v})$.

Any function $f : \mathcal{N} \to \mathbb{R}$ is called a *local function*. If additionally $f(\mathcal{N}) \subseteq \{0, 1, 2\}$, then we call f a *local rule*. Each local function f induces a *global function* $A_f : X \to \tilde{X}$ defined for $\mathbf{x} \in X$ and $\mathbf{i} \in \mathcal{C}$ as follows

$$A_f(\mathbf{x})(\mathbf{i}) = f(N_{\mathbf{x},\mathbf{i}}).$$

If f is a local rule, then we call A_f a global rule and then $A_f : X \to X$. If a global function A_f conserves the sum of states, *i.e.*, for each $\mathbf{x} \in X$ it holds that $\sigma(A_f(\mathbf{x})) = \sigma(\mathbf{x})$, then the local function f is called *number-conserving*.

For example, if $\mathbf{v}$ is any vector from V, then the local rule $f_{\mathbf{v}}$ given by $f_{\mathbf{v}}(N) = N(\mathbf{v})$ is number-conserving. Indeed, $A_{f_{\mathbf{v}}}$ is a shift in the direction $-\mathbf{v}$, which means that in each consecutive updating step, the new state of the cell $\mathbf{i}$ is equal to the state of the cell $\mathbf{i} + \mathbf{v}$ in the previous step. Note that if $\mathbf{v} = \mathbf{0}$, then $A_{f_{\mathbf{0}}}$ is simply the identity on X and therefore $f_{\mathbf{0}}$ is also called the identity rule. However, sometimes it is convenient to regard the identity rule as a shift by the zero vector.

Usually, a d-dimensional CA is a quadruple $(\mathcal{C}, P, Q, f)$, where $\mathcal{C}$ is a grid, P is a function returning the neighborhood for each cell, Q is a state set and f is a local rule. However, here we assume that the grid $\mathcal{C}$ is fixed (note that it is not too small, as we require that all $n_1, n_2, \ldots, n_d$ are greater than three) and use only the von Neumann neighborhood and the state set $Q = \{0, 1, 2\}$. Thus, for the sake of simplicity, we will identify a CA with its global rule and also with its local rule.

Next, we describe the approach to study number conservation we intend to rely on. It turns out that in the case of d-dimensional CAs with the von Neumann neighborhood, any number-conserving local rule can be decomposed into two easy-to-describe parts: a *split function* and a *perturbation*. Both of these objects were introduced in [36]. The theory presented there is developed for the general case of an arbitrary state set $Q \subseteq \mathbb{R}$. In this paper we consider the state set $Q = \{0, 1, 2\}$ and for this reason the definitions and facts from [36] can be reformulated. However, we do it without detailed explanations, since in [34] this theory was described once again for the simplest state set, namely $\{0, 1\}$. We start with the definition of split functions that, roughly speaking, act as follows: each state splits according to its recipe irrespective of the states of its neighbors.

Definition 2.1. (Split functions) A local function h is called a *split function* if it satisfies:

(S1) $h(M_0) = 0$, moreover, $h(M_{\mathbf{v}:q}) \in \{0, 1, 2\}$, for any $\mathbf{v} \in V$ and $q \in \{1, 2\}$;

(S2) $\sum_{\mathbf{v} \in V} h(M_{\mathbf{v}:1}) = 1$ and $\sum_{\mathbf{v} \in V} h(M_{\mathbf{v}:2}) = 2$;

(S3) $h(N) = \sum_{\mathbf{v} \in V} h(M_{\mathbf{v}:N(\mathbf{v})})$, for any $N \in \mathcal{N}$.

The set of all split functions is denoted by $\mathcal{S}$. Each split function is number-conserving (see [36]) and is defined by the values it takes on monomers, because its value on any neighborhood configuration is given by (S3). For example, if h is a split function, then for any $\{\mathbf{u}, \mathbf{w}\} \in \mathbf{\Omega}$ and p, $q \in \{0, 1, 2\}$, it holds that

$$h\left(D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}\right) = h(M_{\mathbf{u}:p}) + h(M_{\mathbf{w}:q}). \tag{4}$$

It is worth emphasizing that for every split function h there are at most three monomers on which h does not equal zero. Indeed, according to (S1) and (S2), there exists exactly one vector $\mathbf{v}_0 \in V$ such that $h\left(M_{\mathbf{v}_0:1}\right) \neq 0$ (and then, of course, $h\left(M_{\mathbf{v}_0:1}\right) = 1$). Moreover, either there is exactly one vector $\mathbf{u}_0 \in V$ such that $h\left(M_{\mathbf{u}_0:2}\right) \neq 0$ (then $h\left(M_{\mathbf{u}_0:2}\right) = 2$ and we say that the state 2 is not disintegrating) or there are exactly two different vectors $\mathbf{u}_0, \mathbf{w}_0 \in V$ such that $h\left(M_{\mathbf{u}_0:2}\right) \neq 0$ and $h\left(M_{\mathbf{w}_0:2}\right) \neq 0$ (then $h\left(M_{\mathbf{u}_0:2}\right) = h\left(M_{\mathbf{w}_0:2}\right) = 1$ and we say that the state 2 is disintegrating). If $\mathbf{0} + \mathbf{v}_0,\ \mathbf{0} + \mathbf{u}_0 \in [\mathbf{u}, \mathbf{w}]$ (or $\mathbf{0} + \mathbf{v}_0,\ \mathbf{0} + \mathbf{u}_0$, $\mathbf{0} + \mathbf{w}_0 \in [\mathbf{v}, \mathbf{u}]$ in the case when the state 2 is disintegrating), for some subgrid $[\mathbf{u}, \mathbf{w}]$, then we say that the split function h *acts inside* $[\mathbf{u}, \mathbf{w}]$.

Next, we introduce the definition of a perturbation. Similarly as in the case of a split function, we use the fact that $Q = \{0, 1, 2\}$ to present a simpler version of the definitions and facts given in [36]. Since the set of all perturbations is a linear space, it is sufficient to define basic perturbations that form a basis of this space. Note that they depend on the choice of $\mathbf{\Lambda}$.

Definition 2.2. (Perturbations) Let the set of independent pairs $\mathbf{\Lambda}$ be chosen. For a given $\{\mathbf{u}, \mathbf{w}\} \in \mathbf{\Lambda}$ and p, $q \in \{1, 2\}$, let $g_{\mathbf{u},\mathbf{w}}^{p,q}$ be the local function that satisfies the following three conditions:

(P1) $g_{\mathbf{u},\mathbf{w}}^{p,q}(M) = 0$, for any monomer M;

(P2) $g_{\mathbf{u},\mathbf{w}}^{p,q}\left(D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}\right) = 1$, while $g_{\mathbf{u},\mathbf{w}}^{p,q}(D) = 0$ for any other dimer D from $\mathcal{D}_{\mathbf{\Lambda}}$;

(P3) for any $N \in \mathcal{N}$, it holds that

$$g_{\mathbf{u},\mathbf{w}}^{p,q}(N) = \sum_{\{\mathbf{v},\mathbf{z}\} \in \mathbf{\Lambda}} \left[g_{\mathbf{u},\mathbf{w}}^{p,q}\left(D_{\substack{\mathbf{v}:N(\mathbf{v}) \\ \mathbf{z}:N(\mathbf{z})}}\right) - g_{\mathbf{u},\mathbf{w}}^{p,q}\left(D_{\substack{\mathbf{v}:N(-\mathbf{z}) \\ \mathbf{z}:N(-\mathbf{v})}}\right) \right]. \tag{5}$$

Let $\mathcal{P}$ be the linear space generated by

$$\left\{ g_{\mathbf{u},\mathbf{w}}^{p,q} \mid \{\mathbf{u}, \mathbf{w}\} \in \mathbf{\Lambda},\ p, q \in \{1, 2\} \right\}.$$

The linear space $\mathcal{P}$ has dimension $4d^2$ and its elements are called perturbations.

As we see, a perturbation is defined by the values it takes on generating dimers, because its value on any neighborhood configuration is given by Eq. (5). The next lemma shows the relationship between the values of a perturbation on matching dimers, proved in [36].

Lemma 2.3. *If a local function $g : \mathcal{N} \to \mathbb{R}$ is a perturbation, then for any $\{\mathbf{u}, \mathbf{w}\} \in \mathbf{\Omega}$ and any p, $q \in \{0, 1, 2\}$, it holds that*

$$g\left(D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}\right) = -g\left(D_{\substack{-\mathbf{u}:q \\ -\mathbf{w}:p}}\right).$$

$\square$

The main result of [36], reformulated in the case $Q = \{0, 1, 2\}$, is presented in the following theorem.

Theorem 2.4. (The decomposition theorem) *Let the set of independent pairs $\mathbf{\Lambda}$ be chosen. The local rule f of a ternary CA with the von Neumann neighborhood is number-conserving if and only if it has the form*

$$\begin{aligned} f = h &+ \sum_{\{\mathbf{u},\mathbf{w}\} \in \mathbf{\Lambda}} a_{\mathbf{u},\mathbf{w}}^{1,1} \cdot g_{\mathbf{u},\mathbf{w}}^{1,1} + \sum_{\{\mathbf{u},\mathbf{w}\} \in \mathbf{\Lambda}} a_{\mathbf{u},\mathbf{w}}^{1,2} \cdot g_{\mathbf{u},\mathbf{w}}^{1,2} \\ &+ \sum_{\{\mathbf{u},\mathbf{w}\} \in \mathbf{\Lambda}} a_{\mathbf{u},\mathbf{w}}^{2,1} \cdot g_{\mathbf{u},\mathbf{w}}^{2,1} + \sum_{\{\mathbf{u},\mathbf{w}\} \in \mathbf{\Lambda}} a_{\mathbf{u},\mathbf{w}}^{2,2} \cdot g_{\mathbf{u},\mathbf{w}}^{2,2}, \end{aligned} \tag{6}$$

where $h \in \mathcal{S}$ and the coefficients $a_{\mathbf{u},\mathbf{w}}^{p,q}$ are chosen such that the values of f are in $\{0, 1, 2\}$.

$\square$

The decomposition of a number-conserving local rule into a split function h and a perturbation g (the remaining part of the right-hand side of Eq. (6)) is unique. Hence, when dealing with a number-conserving local rule f, we will denote its split function and its perturbation by h_f and g_f, respectively.

Remark 2.5. According to the definition of the basic perturbations, if N is a binary neighborhood configuration, then

$$f(N) = h(N) + \sum_{\{\mathbf{u},\mathbf{v}\} \in \mathbf{\Lambda}} a_{\mathbf{u},\mathbf{v}}^{1,1} \cdot g_{\mathbf{u},\mathbf{v}}^{1,1}(N),$$

since the remaining terms are zero.

As mentioned in the introduction, we consider finite grids only for our convenience, as in this case there is no problem with the definition of the sum of all states in a configuration. However, according to Durand et al. [13], Theorem 2.4 is also valid for infinite grids.

Next, we define the second property of CAs we are interested in: reversibility.

Definition 2.6. Let f be the local rule of some d-dimensional CA. We say that this CA is reversible if the function A_f is an injection.

Note that we use a weaker version of reversibility than is commonly done, since we require that A_f is injective only on one grid with a given size $n_1 \times n_2 \times \ldots \times n_d$, while usually injectivity is required for all possible finite grids (a stronger version) or for all infinite configurations (the strongest version). It turns out that even if we weaken the definition of reversibility so much, the main theorem of our paper (Theorem 3.7) remains valid.

In the remainder of this paper, we will use the abbreviation d-RNCCA to denote a d-dimensional CA with the von Neumann neighborhood that is both reversible and number-conserving.

For $d = 1$, all number-conserving ternary CAs were found by Boccara and Fukś [4], and among them there are only three 1-RNCCAs: the identity rule, the shift-left rule and the shift-right rule (if we assume that the considered grid has at least four cells). Furthermore, it was shown by Dzedzej et al. [15] that there are only five 2-RNCCAs: the identity rule and four shift rules in each of the possible directions (right, left, up and down). That result, admittedly, was shown for the stronger version of reversibility, but its proof can be easily obtained also for the weaker version with the assumption that the considered grid is at least 4×4.

We summarize these results in the following theorem.

Theorem 2.7. If $d \leq 2$, then any d-RNCCA is a shift rule i.e., its local rule equals $f_{\mathbf{v}}$ for some $\mathbf{v} \in V$.

So, any 2-RNCCA acts as a 1-dimensional CA and increasing the dimension from 1 to 2 does not result in the emergence of a new kind of RNCCA. In the next section, we prove that the same holds true for any d.

3. The main theorem

In this section, we prove that all d-RNCCAs are trivial, as they act as 1-dimensional ones. Since this fact has already been proved for $d \leq 2$, we assume that $d \geq 3$. We start with a few auxiliary facts.

Lemma 3.1. If a split function h acts inside a subgrid $[\mathbf{u}, \mathbf{w}]$, then for any $\mathbf{x} \in X_{[\mathbf{u},\mathbf{w}]}$ and for any $\mathbf{i} \notin [\mathbf{u}, \mathbf{w}]$, it holds that $h(N_{\mathbf{x},\mathbf{i}}) = 0$.

Proof. Let a split function h be given and assume that the state 2 is disintegrating. There exists exactly one vector $\mathbf{v}_0 \in V$ such that $h(M_{\mathbf{v}_0:1}) \neq 0$ and exactly two different vectors $\mathbf{u}_0, \mathbf{w}_0 \in V$ such that $h(M_{\mathbf{u}_0:2}) \neq 0$ and $h(M_{\mathbf{w}_0:2}) \neq 0$. Let $\mathbf{u}, \mathbf{w} \in V$ be such that h acts inside $[\mathbf{u}, \mathbf{w}]$. Note that if $\mathbf{i} \notin [\mathbf{u}, \mathbf{w}]$, then $P(\mathbf{i}) \cap [\mathbf{u}, \mathbf{w}]$ is either empty or contains exactly one cell. In the first case, $N_{\mathbf{x},\mathbf{i}} = M_0$, so $h(N_{\mathbf{x},\mathbf{i}}) = 0$ according to (S1). In the second case, $N_{\mathbf{x},\mathbf{i}} = M_{\mathbf{v}:p}$, for some $p \in \{0, 1, 2\}$ and $\mathbf{v} \in V_+$, which is orthogonal to both $\mathbf{u}$ and $\mathbf{w}$. In particular, $\mathbf{v}$ has to be orthogonal to $\mathbf{v}_0$, $\mathbf{u}_0$ and $\mathbf{w}_0$, which means that $\mathbf{v}$ is not equal to any of them. As a consequence, we get that $h(N_{\mathbf{x},\mathbf{i}}) = h(M_{\mathbf{v}:p}) = 0$. The proof in the case when the state 2 is not disintegrating is analogous. □

Lemma 3.2. Let f be the local rule of a ternary NCCA. If h_f acts inside a subgrid $[\mathbf{u}, \mathbf{w}]$, then $A_f(X_{[\mathbf{u},\mathbf{w}]}) \subseteq X_{[\mathbf{u},\mathbf{w}]}$.

Proof. Let $\{\mathbf{u}, \mathbf{w}\} \in \Lambda$, $\mathbf{x} \in X_{[\mathbf{u},\mathbf{w}]}$ and $\mathbf{i} \notin [\mathbf{u}, \mathbf{w}]$. We just need to show that $f(N_{\mathbf{x},\mathbf{i}}) = 0$. From Lemma 3.1 we know that $h_f(N_{\mathbf{x},\mathbf{i}}) = 0$. Moreover, $N_{\mathbf{x},\mathbf{i}}$ is a monomer, so $g_f(N_{\mathbf{x},\mathbf{i}}) = 0$, according to the definition of a perturbation. Thus, $f(N_{\mathbf{x},\mathbf{i}}) = h_f(N_{\mathbf{x},\mathbf{i}}) + g_f(N_{\mathbf{x},\mathbf{i}}) = 0$, which concludes the proof. □

If a global rule A_f satisfies $A_f(X_{[\mathbf{u},\mathbf{w}]}) \subseteq X_{[\mathbf{u},\mathbf{w}]}$ for some subgrid $[\mathbf{u}, \mathbf{w}]$, then we can define a new cellular automaton $A_{[\mathbf{u},\mathbf{w}]}$ (2-dimensional or 1-dimensional) acting on the subgrid $[\mathbf{u}, \mathbf{w}]$. Indeed, let Φ be the natural isomorphism between the set $X_{[\mathbf{u},\mathbf{w}]}$ and the set of all configurations on $[\mathbf{u}, \mathbf{w}]$ (the set $\{0, 1, 2\}^{[\mathbf{u},\mathbf{w}]}$). We define $A_{[\mathbf{u},\mathbf{w}]}$ as the CA for which the following diagram commutes.

$$
\begin{array}{ccc}
X_{[\mathbf{u},\mathbf{w}]} & \xrightarrow{\;A_f\;} & X_{[\mathbf{u},\mathbf{w}]} \\
\Phi \downarrow & & \downarrow \Phi \\
\{0, 1, 2\}^{[\mathbf{u},\mathbf{w}]} & \xrightarrow{A_{[\mathbf{u},\mathbf{w}]}} & \{0, 1, 2\}^{[\mathbf{u},\mathbf{w}]}
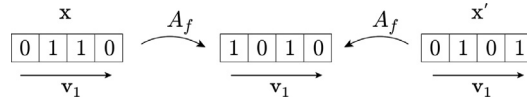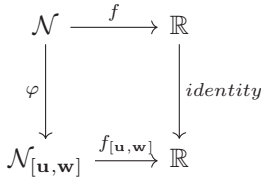\end{array}
$$

Fig. 2. Two configurations **x** and **x**′ having the same successor in the case $a_{\mathbf{0},\mathbf{v}_1}^{1,1} = -1$.

We will denote the local rule of $A_{[\mathbf{u},\mathbf{w}]}$ by $f_{[\mathbf{u},\mathbf{w}]}$. The isomorphism Φ induces an isomorphism φ between $\mathcal{N}$ and the domain $\mathcal{N}_{[\mathbf{u},\mathbf{w}]}$ of $f_{[\mathbf{u},\mathbf{w}]}$, for which we have:

$$
\begin{array}{ccc}
\mathcal{N} & \xrightarrow{\ f\ } & \mathbb{R} \\
\varphi \downarrow & & \downarrow identity \\
\mathcal{N}_{[\mathbf{u},\mathbf{w}]} & \xrightarrow{\ f_{[\mathbf{u},\mathbf{w}]}\ } & \mathbb{R}
\end{array}
$$

This means, in particular, that $f_{[\mathbf{u},\mathbf{w}]}\big(\varphi(N_{\mathbf{x},\mathbf{i}})\big) = f(N_{\mathbf{x},\mathbf{i}})$, for any $\mathbf{x} \in X_{[\mathbf{u},\mathbf{w}]}$ and $\mathbf{i} \in \mathcal{C}$. Note that if f is number-conserving, then $f_{[\mathbf{u},\mathbf{w}]}$ is also number-conserving and its split function and perturbation will be denoted by $h_{[\mathbf{u},\mathbf{w}]}$ and $g_{[\mathbf{u},\mathbf{w}]}$, respectively.

Corollary 3.3. *Let f be the local rule of a ternary number-conserving CA. If h_f acts inside a subgrid $[\mathbf{u}, \mathbf{w}]$, then $A_{[\mathbf{u},\mathbf{w}]}$ is a shift on $[\mathbf{u}, \mathbf{w}]$.*

Proof. From Lemma 3.2 we know that $A_f\big(X_{[\mathbf{u},\mathbf{w}]}\big) \subseteq X_{[\mathbf{u},\mathbf{w}]}$, so the CA $A_{[\mathbf{u},\mathbf{w}]}$ is well-defined, number-conserving and reversible, *i.e.*, $A_{[\mathbf{u},\mathbf{w}]}$ is a 2- or 1-RNCCA. From Theorem 2.7 we get that $A_{[\mathbf{u},\mathbf{w}]}$ is a shift. □

Lemma 3.4. *Let f be the local rule of a ternary d-RNCCA. If $h_f\big(M_{\mathbf{0}:1}\big) = 1$, then $h_f\big(M_{\mathbf{0}:2}\big) = 2$.*

Proof. According to (S1) and (S2), either there are two different vectors $\mathbf{u}_0, \mathbf{w}_0 \in V$ such that $h_f\big(M_{\mathbf{u}_0:2}\big) = h_f\big(M_{\mathbf{w}_0:2}\big) = 1$ or there is one vector $\mathbf{u}_0 \in V$ such that $h_f\big(M_{\mathbf{u}_0:2}\big) = 2$. In the first case, h_f acts inside $[\mathbf{u}_0, \mathbf{w}_0]$, which implies that $A_{[\mathbf{u}_0,\mathbf{w}_0]}$ is a shift on $[\mathbf{u}_0, \mathbf{w}_0]$ (see Corollary 3.3). Since $h_f\big(M_{\mathbf{0}:1}\big) = 1$, it has to be the shift by the zero vector, but then $h_f\big(M_{\mathbf{0}:2}\big) = 2$. The second case can be treated similarly. □

The next lemma states that if $h_f\big(M_{\mathbf{v}_1:1}\big) = 1$, then the global rule A_f considered on binary configurations only acts as a shift.

Lemma 3.5. *Let f be the local rule of a ternary d-RNCCA. If $h_f\big(M_{\mathbf{v}_1:1}\big) = 1$, then $A_f|_{X_{bin}}$ is the shift by the vector -$\mathbf{v}_1$.*

Proof. To simplify our reasoning, let us choose $\mathbf{\Lambda}$ as in Eq. (3). According to Remark 2.5, all we need to do is argue that for any $\{\mathbf{u}, \mathbf{w}\} \in \mathbf{\Lambda}$ it holds that $a_{\mathbf{u},\mathbf{w}}^{1,1} = 0$.

First, let $\{\mathbf{u}, \mathbf{v}\} \in \mathbf{\Lambda}$ be such that none of $\mathbf{u}$, -$\mathbf{u}$, $\mathbf{v}$, -$\mathbf{v}$ is equal to $\mathbf{v}_1$. Since

$$
f\left(D_{\substack{\mathbf{u}:1 \\ \mathbf{v}:1}}\right) = h_f\left(D_{\substack{\mathbf{u}:1 \\ \mathbf{v}:1}}\right) + g_f\left(D_{\substack{\mathbf{u}:1 \\ \mathbf{v}:1}}\right) = a_{\mathbf{u},\mathbf{v}}^{1,1},
$$

while from Lemma 2.3 it follows that

$$
f\left(D_{\substack{-\mathbf{u}:1 \\ -\mathbf{v}:1}}\right) = h_f\left(D_{\substack{-\mathbf{u}:1 \\ -\mathbf{v}:1}}\right) + g_f\left(D_{\substack{-\mathbf{u}:1 \\ -\mathbf{v}:1}}\right) = -a_{\mathbf{u},\mathbf{v}}^{1,1},
$$

we get that $a_{\mathbf{u},\mathbf{v}}^{1,1} = 0$, as both $f\left(D_{\substack{\mathbf{u}:1 \\ \mathbf{v}:1}}\right)$ and $f\left(D_{\substack{-\mathbf{u}:1 \\ -\mathbf{v}:1}}\right)$ are non-negative.

Next, we consider the pair $\{\mathbf{0}, \mathbf{v}_1\}$. As

$$
f\left(D_{\substack{\mathbf{0}:1 \\ \mathbf{v}_1:1}}\right) = 1 + a_{\mathbf{0},\mathbf{v}_1}^{1,1} \quad \text{and} \quad f\left(D_{\substack{\mathbf{0}:1 \\ -\mathbf{v}_1:1}}\right) = -a_{\mathbf{0},\mathbf{v}_1}^{1,1}, \tag{7}
$$

we conclude that $a_{\mathbf{0},\mathbf{v}_1}^{1,1} = 0$ or $a_{\mathbf{0},\mathbf{v}_1}^{1,1} = -1$. To reject the latter possibility, let us consider the following initial configuration $\mathbf{x} \in X_{bin} \cap X_{[\mathbf{0},\mathbf{v}_1]}$: the string 0110 in an arbitrary place of the subgrid $[\mathbf{0}, \mathbf{v}_1]$ and zeros in all other cells. If $a_{\mathbf{0},\mathbf{v}_1}^{1,1}$ were equal to -1, then the string 0110 would be updated to 1010, while the other cells do not change their state (see Eq. (7)). Thus, we would get the same successor as for $\mathbf{x}'$ being the modification of the configuration $\mathbf{x}$ having 0101 instead of 0110 (see Fig. 2). This would contradict the reversibility of the global rule A_f.
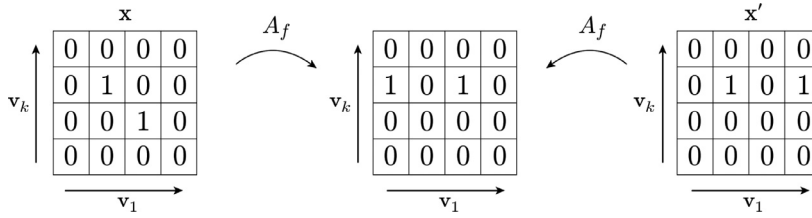
Fig. 3. Two configurations **x** and **x′** having the same successor in the case $a_{\mathbf{v}_1,\mathbf{v}_k}^{1,1} = -1$.

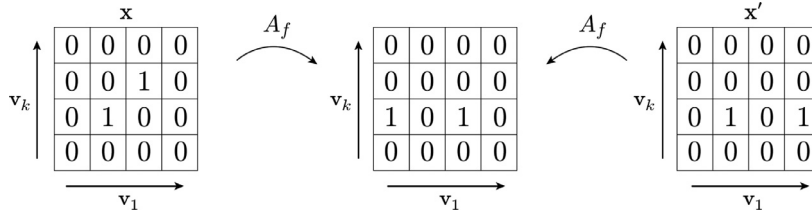**Fig. 4.** Two configurations **x** and **x′** having the same successor in the case $a_{-\mathbf{v}_1,\mathbf{v}_k}^{1,1} = 1$.

Finally, let us consider any $\mathbf{v}_k \in \{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_d\}$. As $\{\mathbf{v}_1, \mathbf{v}_k\} \in \Lambda$, it holds that

$$f\left(D_{\substack{\mathbf{v}_1:1 \\ \mathbf{v}_k:1}}\right) = 1 + a_{\mathbf{v}_1,\mathbf{v}_k}^{1,1} \quad \text{and} \quad f\left(D_{\substack{-\mathbf{v}_1:1 \\ -\mathbf{v}_k:1}}\right) = -a_{\mathbf{v}_1,\mathbf{v}_k}^{1,1}. \tag{8}$$

Thus $a_{\mathbf{v}_1,\mathbf{v}_k}^{1,1} = 0$ or $a_{\mathbf{v}_1,\mathbf{v}_k}^{1,1} = -1$. To see that the latter possibility is impossible, it is sufficient to consider the following initial configurations $\mathbf{x}, \mathbf{x}' \in X_{bin} \cap X_{[\mathbf{v}_1,\mathbf{v}_k]}$: the states of cells laying in the subgrid $\mathbf{v}_1, \mathbf{v}_k$ are shown in Fig. 3, while the other cells have state 0. If $a_{\mathbf{v}_1,\mathbf{v}_k}^{1,1}$ were equal to -1, then these configurations would have the same successor.

Similarly, for $\{-\mathbf{v}_1, \mathbf{v}_k\} \in \Lambda$ we have

$$f\left(D_{\substack{-\mathbf{v}_1:1 \\ \mathbf{v}_k:1}}\right) = a_{-\mathbf{v}_1,\mathbf{v}_k}^{1,1} \quad \text{and} \quad f\left(D_{\substack{\mathbf{v}_1:1 \\ -\mathbf{v}_k:1}}\right) = 1 - a_{-\mathbf{v}_1,\mathbf{v}_k}^{1,1}, \tag{9}$$

which gives $a_{-\mathbf{v}_1,\mathbf{v}_k}^{1,1} = 0$ or $a_{-\mathbf{v}_1,\mathbf{v}_k}^{1,1} = 1$, but Fig. 4 shows that if $a_{-\mathbf{v}_1,\mathbf{v}_k}^{1,1}$ were 1, then we would loose reversibility. □

Corollary 3.6. *Let f be the local rule of a ternary d-RNCCA. If $h_f(M_{\mathbf{v}_1:1}) = 1$, then the state 2 is not disintegrating, moreover, $h_f(M_{\mathbf{v}_1:2}) = 2$.*

Proof. Let us consider the subset X' of X consisting of all configurations **x** satisfying $\sigma(\mathbf{x}) = 2$. Thus $X' = X_1 \cup X_2$, where X_1 contains all configurations having two states 1, while the other states are 0, and X_2 is the set of all configurations with one state 2, while the others are 0. Since f is number-conserving and reversible, it holds that $A_f : X' \to X'$ is an injection. According to Lemma 3.5, $A_f(X_1) = X_1$, thus, in view of injectivity, it has to hold that $A_f(X_2) = X_2$, which means that the state 2 is not disintegrating. Thus there is a vector $\mathbf{u}_0 \in V$ such that $h_f(M_{\mathbf{u}_0:2}) = 2$. If so, h_f acts inside $[\mathbf{v}_1, \mathbf{u}_0]$ and from Corollary 3.3 it follows that $A_{[\mathbf{v}_1,\mathbf{u}_0]}$ is a shift. In particular, $\mathbf{u}_0 = \mathbf{v}_1$. □

Now we formulate and prove the main result of this paper.

Theorem 3.7. *Let $d \geq 1$. Any d-RNCCA is a shift, i.e., its local rule equals $f_{\mathbf{v}}$ for some $\mathbf{v} \in V$.*

Proof. Let us choose Λ as in Eq. (3). First, we focus on the case that the split function h_f satisfies $h_f(M_{\mathbf{0}:1}) = 1$ and we show that then f has to be the identity rule. From Lemma 3.4 it follows that $h_f = f_{\mathbf{0}}$. Thus, all we need to do is prove that $g_f = 0$, or, equivalently, that all coefficients $a_{\mathbf{u},\mathbf{w}}^{p,q}$ in Eq. (6) equal 0.

Then, let $\{\mathbf{u}, \mathbf{w}\} \in \Lambda$ and $p, q \in \{1, 2\}$. Since $h_f(M_{\mathbf{0}:1}) = 1$, Lemma 3.4 implies that $h_f(M_{\mathbf{0}:2}) = 2$, which guarantees that h_f acts inside $[\mathbf{u}, \mathbf{w}]$. From Lemma 3.2 we know that $A_{[\mathbf{u},\mathbf{w}]}$ is a ternary 2- or 1-RNCCA on the subgrid $[\mathbf{u}, \mathbf{w}]$. According to Theorem 2.7, $A_{[\mathbf{u},\mathbf{w}]}$ is a shift, so, its perturbation $g_{[\mathbf{u},\mathbf{w}]}$ is a zero function. This yields, in particular, that $a_{\mathbf{u},\mathbf{w}}^{p,q} = g_f\left(D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}\right) = g_{[\mathbf{u},\mathbf{w}]}\left(\varphi\left(D_{\substack{\mathbf{u}:p \\ \mathbf{w}:q}}\right)\right) = 0$.

Now, we consider the case that $h_f(M_{\mathbf{0}:1}) = 0$, so there exists $\mathbf{w} \in V_+$ such that $h_f(M_{\mathbf{w}:1}) = 1$. As the proof is the same for any direction $\mathbf{w} \in V_+$, we will deal with $\mathbf{v}_1$ only.

From Corollary 3.6 we know that $h_f(M_{\mathbf{v}_1:2}) = 2$, so $h_f = f_{\mathbf{v}_1}$ and again it is sufficient to prove that for any $\{\mathbf{u}, \mathbf{w}\} \in \Lambda$ and $p, q \in \{1, 2\}$ the coefficient $a_{\mathbf{u},\mathbf{w}}^{p,q}$ in Eq. (2.4) is equal to zero.

First, let us consider the case when one of the vectors **u**, **w** equals $\mathbf{v}_1$ or $-\mathbf{v}_1$. Then h_f acts inside $[\mathbf{u}, \mathbf{w}]$ and, similarly as above, we argue that $a_{\mathbf{u},\mathbf{w}}^{p,q} = 0$.

It remains to consider the case when $\mathbf{u},\ \mathbf{w} \in V \backslash \{\mathbf{v}_1,\ -\mathbf{v}_1\}$. Then $h_f\left(D_{\mathbf{w}:q}^{\mathbf{u}:p}\right) = h_f\left(D_{-\mathbf{w}:p}^{-\mathbf{u}:q}\right) = 0$, so $f\left(D_{\mathbf{w}:q}^{\mathbf{u}:p}\right) = a_{\mathbf{u},\mathbf{w}}^{p,q}$, while $f\left(D_{-\mathbf{w}:p}^{-\mathbf{u}:q}\right) = -a_{\mathbf{u},\mathbf{w}}^{p,q}$, which gives $a_{\mathbf{u},\mathbf{w}}^{p,q} = 0$. $\quad\square$

Although Theorem 3.7 concerns a weaker version of reversibility of CAs, where we require only the injectivity of the global rule on one fixed grid, it is still valid if we consider the stronger version (the global rule is injective on each finite grid) or if we consider the strongest version (the global rule is injective on infinite configurations). Indeed, the structure of the proof of Theorem 3.7 is the following: using some configurations that are non-zero only on a very small neighborhood (contained in a $4 \times 4 \times \ldots \times 4$ cube), we show that every non-shift NCCA cannot be reversible. Thus we end up only with shifts that are, of course, reversible irrespective of the version of the definition, we use.

4. Conclusions and open problems

The previous approaches to study CA reversibility concerned either 1-dimensional CAs or linear CAs with $\mathbb{Z}_m$ as state set. In this paper, we presented a novel method that is suitable for multi-state number-conserving CAs. The method is based on the split-and-perturb decomposition of a number-conserving local rule introduced in [36]. In contrast to other studies on the reversibility problem, where the authors use the fact that the local rule is linear or the state set is a ring, our approach does not assume any *nice structure* of the state set of the considered CAs and replaces the linearity of the local rule by a much weaker assumption: number conservation.

Here, we made a step forward in the study of reversibility of multi-state number-conserving CAs and enumerated all ternary d-dimensional RNCCAs to show that three states are too few to enable the existence of nontrivial reversible number-conserving CAs with the von Neumann neighborhood in any dimension. Although this result concerns ternary CAs only, it opens a future research direction to be explored. Actually, we think that, using techniques like the ones in the paper, other related questions in the field could also be analyzed.

First of all, future work may focus on determining whether or not the following conjecture presented in [14] is true.

Conjecture 4.1. *There are exactly* $(2d + 1)(6d + 1)$ *four-state d-dimensional RNCCAs with the von Neumann neighborhood.*

The existing results confirm this conjecture only for $d = 1$ and $d = 2$ (see [14]), but a detailed analysis of their proofs makes us hopeful of proving it for any dimension.

The second conjecture connected with the results of [14] concerns a lack of possibility of the appearance of a new kind of four-state RNCCAs when the dimension of a grid increases.

Conjecture 4.2. *Let $d > 2$. All four-state d-dimensional RNCCAs with the von Neumann neighborhood are intrinsically 2-dimensional.*

It can be suspected that the problem discussed in the conjecture below can be generalized to larger number of states, *i.e.*, to larger q_*. It seems that there exists a function $dim : \mathbb{N} \to \mathbb{N}$ such that all d-dimensional RNCCAs with the state set $\{0, 1, \ldots, q_*\}$ (and with the von Neumann neighborhood) are intrinsically at most $dim(q_*)$-dimensional. This hypothesis is formulated in the following open problem.

Question 4.3. Let $q_* \in \mathbb{N}$. Does there exist a dimension $dim(q_*)$ such that increasing the dimension of a grid beyond $dim(q_*)$ does not result in the appearance of a new kind of RNCCA with the state set $\{0, 1, \ldots, q_*\}$? If so, what can be said about the dependence between q_* and $dim(q_*)$?

The last open problem we would like to present is connected with the negative result of Kari [22]. In general, there is no bound on the size of the grid that is needed to verify non-injectivity of 2- or higher-dimensional CAs. However, for what concerns ternary number-conserving CAs, we only needed to consider grids of sizes $4 \times 4 \times \ldots \times 4$. The same is observed in the case of 2-dimensional four-state number-conserving CAs (see [14]). Hence, it is natural to ask whether the reason for this is the small state set or the assumed property of number conservation. Generally, the question is whether such bounds exist for any $q_* \in \mathbb{N}$.

Question 4.4. Let $q_* \in \mathbb{N}$. Does there exist a bound on the size of d-dimensional grids that are sufficient to verify the injectivity of a number-conserving CA with the von Neumann neighborhood and the state set $\{0, 1, \ldots, q_*\}$?

Because the results on d-dimensional multi-state RNCCAs are limited, any answer to the above problems will be appreciated by the scientific community.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ins.2019.10.068

References

[1] S. Amoroso, Y. Patt, Decision procedures for surjectivity and injectivity of parallel maps for tessellation structures, J. Comput. Syst. Sci. 6 (1972) 448–464.
[2] A. Augustynowicz, J.M. Baetens, B. De Baets, A. Dzedzej, A. Nenca, B. Wolnik, A note on the reversibility of 2D cellular automata on hexagonal grids, J. Cell. Autom. 13 (2018) 521–525.
[3] K. Bhattacharjee, S. Das, Reversibility of d-state finite cellular automata, J. Cell. Autom. 11 (2016) 213–245.
[4] N. Boccara, H. Fukś, Number-conserving cellular automaton rules, Fundamenta Informaticae 52 (2002) 1–13.
[5] D. Caviedes-Voulliéme, J. Fernández-Pato, C. Hinz, Cellular automata and finite volume solvers converge for 2D shallow flow modelling for hydrological modelling, J. Hydrol. 563 (2018) 411–417.
[6] C.-H. Chang, H. Chang, On the bernoulli automorphism of the reversible linear cellular automata, Inf. Sci. 345 (2016) 217–225.
[7] A.K. Das, A. Ganguly, A. Dasgupta, S. Bhawmik, P.P. Chaudhuri, Efficient characterisation of cellular automata, IEE Proc. E - Comput. Dig. Tech. 137 (1990) 81–87.
[8] A.M. del Rey, A note on the reversibility of elementary cellular automaton 150 with periodic boundary conditions, Roman. J. Inf. Sci. Technol. 16 (2013) 365–372.
[9] A.M. del Rey, G.R. Sánchez, On the reversibility of 150 wolfram cellular automata, Int. J. Modern Phys. C 17 (2006) 975–983.
[10] A.M. del Rey, G.R. Sánchez, Reversibility of linear cellular automata, Appl. Math. Comput. 217 (2011) 8360–8366, doi:10.1016/j.amc.2011.03.033.
[11] A. Dennunzio, E. Formenti, L. Manzoni, L. Margara, A.E. Porreca, On the dynamical behaviour of linear higher-order cellular automata and its decidability, Inf. Sci. 486 (2019) 73–87, doi:10.1016/j.ins.2019.02.023.
[12] S. Di Gregorio, G. Trautteur, On reversibility in cellular automata, J. Comput. Syst. Sci. 11 (1975) 382–391.
[13] B. Durand, E. Formenti, Z. Róka, Number-conserving cellular automata i: decidability, Theor. Comput. Sci. 299 (2003) 523–535, doi:10.1016/S0304-3975(02)00534-0.
[14] A. Dzedzej, B. Wolnik, M. Dziemiańczuk, A. Nenca, J.M. Baetens, B. De Baets, A two-layer representation of four-state reversible number-conserving 2D cellular automata, J. Stat. Mech. (2019) 073202, https://doi.org/10.1088/1742-5468/ab25df.
[15] A. Dzedzej, B. Wolnik, A. Nenca, J.M. Baetens, B. De Baets, Efficient enumeration of three-state two-dimensional number-conserving cellular automata, accepted to Information and Computation, 2019b.
[16] L.H. Encinas, A.M. del Rey, Inverse rules of ECA with rule number 150, Appl. Math. Comput. 189 (2007) 1782–1786.
[17] H. Fukś, R. Procyk, Explorations of ternary cellular automata and ternary density classification problems, Acta Physica Polonica B 12 (2019) 75–89.
[18] F. García-Ramos, Product decomposition for surjective 2-block NCCA, Discret. Math. Theor. Comput. Sci. (2011) 147–158.
[19] G. Hedlund, Endomorphisms and automorphisms of shift dynamical systems, Math. Syst. Theory 3 (1969) 320–375.
[20] K. Imai, B. Martin, R. Saito, On radius 1 nontrivial reversible and number-conserving cellular automata, in: A. Adamatzky (Ed.), Reversibility and Universality, Springer, 2018, pp. 269–277.
[21] M. Itô, N. Ôsato, M. Nasu, Linear cellular automata over Z_m, J. Comput. Syst. Sci. 27 (1983) 125–140, doi:10.1016/0022-0000(83)90033-8.
[22] J. Kari, Reversibility of 2D cellular automata is undecidable, Physica D Nonlinear Phenomena 45 (1990) 379–385.
[23] J. Kari, Theory of cellular automata: a survey, Theor. Comput. Sci. 334 (2005) 3–33, doi:10.1016/j.tcs.2004.11.021.
[24] J. Kari, Reversible cellular automata: from fundamental classical results to recent developments, New Gener. Comput. 36 (2018) 145–172, doi:10.1007/s00354-018-0034-6.
[25] G. Manzini, L. Margara, Invertible linear cellular automata over Z_m: algorithmic and dynamical aspects, J. Comput. Syst. Sci. 56 (1998) 60–67, doi:10.1006/jcss.1997.1535.
[26] J.C.S.T. Mora, Matrix methods and local properties of reversible one-dimensional cellular automata, J. Physics A Math. Gen. 35 (2002) 5563–5573, doi:10.1088/0305-4470/35/27/303.
[27] K. Morita, Reversible Cellular Automata, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 231–257.
[28] K. Morita, Theory of Reversible Computing, Springer, 2017.
[29] D. Richardson, Tessellation with local transformations, J. Comput. Syst. Sci. 6 (1972) 373–388.
[30] P. Sarkar, R. Barua, The set of reversible 90/150 cellular automata is regular, Discret. Appl. Math. 84 (1998) 199–213.
[31] A. Schranko, P.P.B. De Oliveira, Derivation and representation of one-dimensional, reversible, number-conserving cellular automata rules, J. Cell. Autom. 6 (2011) 77–89.
[32] D.H. Serrano, A.M. del Rey, A closed formula for the inverse of a reversible cellular automaton with $(2R + 1)$-cyclic rule, Appl. Math. Comput. 357 (2019) 23–34, doi:10.1016/j.amc.2019.03.060.
[33] K. Sutner, De Bruijn graphs and linear cellular automata, Comp. Syst. 5 (1991) 19–30.
[34] B. Wolnik, B. De Baets, 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.
[35] B. Wolnik, A. Dzedzej, J.M. Baetens, B. De Baets, Number-conserving cellular automata with a von Neumann neighborhood of range one, J. Phys. A Math. Theor. 50 (2017) 435101.
[36] B. Wolnik, A. Nenca, J.M. Baetens, B. De Baets, A split-and-perturb decomposition of number-conserving cellular automata, 2019.arXiv:1901.05067
[37] B. Yang, C. Wang, A. Xiang, Reversibility of general 1d linear cellular automata over the binary field Z_2 under null boundary conditions, Inf. Sci. 324 (2015) 23–31.