J Dyn Diff Equat (2016) 28:69–92
DOI 10.1007/s10884-015-9471-9

Solution Structure of Multi-layer Neural Networks with Initial Condition

Jung-Chao Ban[1] · **Chih-Hung Chang**[2]

Received: 20 May 2013 / Revised: 14 July 2015 / Published online: 25 July 2015

Abstract This paper studies the initial value problem of multi-layer cellular neural networks. We demonstrate that the mosaic solutions of such system is topologically conjugated to a new class in symbolic dynamical systems called the path set (Abram and Lagarias in Adv Appl Math 56:109–134, 2014). The topological entropies of the solution, output, and hidden spaces of a multi-layer cellular neural network with initial condition are formulated explicitly. Also, a sufficient condition for whether the mosaic solution space of a multi-layer cellular neural network is independent of initial conditions is addressed. Furthermore, two spaces exhibit identical topological entropy if and only if they are finitely equivalent.

Keywords Initial value problem · Cellular neural networks · Sofic shift · Path set

Mathematics Subject Classification Primary 37B10

1 Main Results

In the past few decades, cellular neural networks (CNNs) introduced by Chua and Yang [14,15] have been one of the most investigated paradigms for neural information processing [13]. In a wide range of applications, the CNNs are required to be completely stable, i.e., each trajectory should converge toward some stationary state. In the study of stationary solutions, the investigation of mosaic solutions is most essential in CNNs due to the learning algorithm

✉ Chih-Hung Chang
chchang@nuk.edu.tw

Jung-Chao Ban
jcban@mail.ndhu.edu.tw

[1] Department of Applied Mathematics, National Dong Hwa University, Hualien 970003, Taiwan, ROC

[2] Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung 81148, Taiwan, ROC

and training processing. More abundant output patterns make the learning algorithm more efficient.

Multi-layer cellular neural networks (MCNNs), which are the coupled systems based on CNNs, have received considerable attention and have been successfully applied to many areas such as signal propagation between neurons, image processing, pattern recognition, information technology, CMOS realization and VLSI implement [4,9,12,14,16–19,25,28,30,34–36,36]. One important reason for coupling CNNs is the simulation of the visual systems of mammals ([22,23], with each layer symbolizing a single cortex in the visual system). In [33], the authors demonstrated a sufficient condition for the complete stability of MCNNs. Recently, Ban and Chang [6] showed that for MCNNs, more layers infer more phenomena that the models are capable of. This work intends to investigate the complexity of the space consisting of mosaic solutions under the constraint of initial conditions.

A multi-layer cellular neural network is realized as

$$\begin{cases} \dfrac{d}{dt}x_i^{(n)}(t) = -x_i^{(n)}(t) + z^{(n)} + \displaystyle\sum_{k\in\mathcal{N}} (a_k^{(n)} f(x_{i+k}^{(n)}(t)) + b_k^{(n)} f(x_{i+k}^{(n-1)}(t))), \\ \qquad\qquad \vdots \\ \dfrac{d}{dt}x_i^{(2)}(t) = -x_i^{(2)}(t) + z^{(2)} + \displaystyle\sum_{k\in\mathcal{N}} (a_k^{(2)} f(x_{i+k}^{(2)}(t)) + b_k^{(2)} f(x_{i+k}^{(1)}(t))), \\ \dfrac{d}{dt}x_i^{(1)}(t) = -x_i^{(1)}(t) + z^{(1)} + \displaystyle\sum_{k\in\mathcal{N}} a_k^{(1)} f(x_{i+k}^{(1)}(t)), \end{cases} \tag{1}$$

for some integer $n \geq 2$, $i \in \mathbb{N}$, and $t \geq 0$. Herein $x_i^{(\ell)}(t) = 0$ for $1 \leq \ell \leq n$ and $t \geq 0$ provided $i \leq 0$. The so-called *neighborhood* $\mathcal{N}$ is a finite subset of integers $\mathbb{Z}$; the output function

$$f(x) = \frac{1}{2}(|x+1| - |x-1|) \tag{2}$$

is a piecewise linear map. $\mathbf{A} = [A^{(1)}, \ldots, A^{(n)}]$ and $\mathbf{B} = [B^{(2)}, \ldots, B^{(n)}]$ are the feedback and controlling templates, respectively, where $A^{(j)} = [a_k^{(j)}]_{k\in\mathcal{N}}$, $B^{(l)} = [b_k^{(l)}]_{k\in\mathcal{N}}$ for $1 \leq j \leq n, 2 \leq l \leq n$; $\mathbf{z} = [z^{(1)}, \ldots, z^{(n)}]$ is the threshold. The template $\mathbb{T}$ of (1) consists of the feedback and controlling templates and the threshold, namely, $\mathbb{T} = [\mathbf{A}, \mathbf{B}, \mathbf{z}]$.

Given $\mathbf{a} \in \mathbb{R}^n$ and $\ell \in \mathbb{N}$, the *initial value problem* (IVP) of a MCNN is investigating those solutions that satisfy (1) with initial condition $\mathbf{a}$ at coordinate ℓ. More precisely, instead of studying the space

$$\mathbf{X} = \left\{\mathbf{x}(t) = (x_i^{(j)}(t))_{i\in\mathbb{N},1\leq j\leq n} \in \mathbb{R}^{\infty\times n} : \mathbf{x}(t) \text{ satisfies (1)}\right\} \tag{3}$$

which consists of all possible solutions of (1), the IVP of a MCNN focuses on investigating the space

$$\mathbf{X}_\mathbf{a} = \left\{\mathbf{x}(t) = (x_i^{(j)}(t))_{i\in\mathbb{N},1\leq j\leq n} \in \mathbb{R}^{\infty\times n} : \mathbf{x}(t) \text{ satisfies (1)}, (x_\ell^{(j)}(0))_{j=1}^n = \mathbf{a}\right\}. \tag{4}$$

For ease of notation, we omit the time variable t and let $f(\mathbf{x}) = (f(x_i^{(j)}))_{i\in\mathbb{N},1\leq j\leq n}$, where $\mathbf{x} = (x_i^{(j)})_{i\in\mathbb{N},1\leq j\leq n} \in \mathbb{R}^{\infty\times n}$ and f is the output function defined in (2). Set

$$\mathbf{Y} = \{\mathbf{y} = (y_i^{(j)})_{i\in\mathbb{N},1\leq j\leq n} \in \mathbb{R}^{\infty\times n} : \mathbf{y} = f(\mathbf{x}), \mathbf{x} \in \mathbf{X}\} \tag{5}$$

and

$$\mathbf{Y_a} = \{\mathbf{y} = (y_i^{(j)})_{i\in\mathbb{N},1\le j\le n} \in \mathbb{R}^{\infty\times n} : \mathbf{y} = f(\mathbf{x}), \mathbf{x} \in \mathbf{X_a}\}. \tag{6}$$

For $1 \le m \le n$, define $\Phi^{(m)} : \mathbb{R}^{\infty\times n} \to \mathbb{R}^{\infty\times 1}$ by $\Phi^{(m)}(\mathbf{x}) = (x_i^{(m)})_{i\in\mathbb{N}}$, where $\mathbf{x} = (x_i^{(j)})_{i\in\mathbb{N},1\le j\le n}$. Set

$$\mathbf{Y}_\mathbf{a}^{(m)} = \{y = (y_i)_{i\in\mathbb{N}} : y = \Phi^{(m)}(\mathbf{y}), \mathbf{y} \in \mathbf{Y}\}. \tag{7}$$

We call $\mathbf{Y_a}$ the solution space, $\mathbf{Y}_\mathbf{a}^{(n)}$ the output space, and $\mathbf{Y}_\mathbf{a}^{(m)}$ the m -th hidden space of (1) with initial condition

$$\left(x_\ell^{(j)}\right)_{j=1}^{n}(0) = \mathbf{a}, \quad \ell \in \mathbb{N}, \mathbf{a} \in \mathbb{R}^n, \tag{8}$$

where $1 \le m \le n-1$.

The present paper focuses on those completely stable solutions in the space $\mathbf{X_a}$ that tend to a mosaic solutions in finite time. More specifically, there exists $T > 0$ such that, if $\mathbf{x} = (x_i^{(j)})$ in (4), then $|x_i^{(j)}(t)| > 1$ for all i, j, and $t \ge T$. To make the investigation self-contained and more readable, we recall some results about the complete stability and exponential convergence of MCNNs in the following. There are a lot of related references, those referred ones are constrained by the authors' interest.

Among wide applications of MCNNs, many of them, such as image recognition, require the complete stability, which means that every solution of (1) converges to an equilibrium solution. Meanwhile, the binary outputs of the solutions of MCNNs are also needed; more precisely, mosaic solutions are essential for these applications. For those sufficient conditions that make a one-layer CNN completely stable, many of them are followed by $|\overline{x}_i| > 1$ for all $i \in \mathbb{Z}$, where $\overline{\mathbf{x}} = (\overline{x}_i)$ is an equilibrium solution of the system. Despite of some theoretical results, numerical experiments infer that most solutions converge exponentially to the equilibria. The complete stability and exponential convergence of MCNNs is also demonstrated. Reader is referred to [5,20,21,26,31–33] and the references therein for more details. This motivates the investigation of those solutions $\mathbf{x} = (x_i^{(j)})_{i\in\mathbb{N},1\le j\le n}$ of (1) satisfying $|x_i^{(j)}(t)| > 1$ for all i, j in finite time.

The aim of this study is two fold. First we compute the spatial complexity of $\mathbf{Y_a}$ and $\mathbf{Y}_\mathbf{a}^{(m)}$, $1 \le m \le n$. We achieve this by characterizing the underlined shift spaces of them. That is, we prove that $\mathbf{Y_a}$ and $\mathbf{Y}_\mathbf{a}^{(m)}$, $1 \le m \le n$ are topologically conjugate to path sets (Theorem 1.2); then using the latest result for path sets to compute the entropy. Secondly, we present the classification result. That is, we characterize when two shift spaces in $\{\mathbf{Y_a}, \mathbf{Y}_\mathbf{a}^{(1)}, \ldots, \mathbf{Y}_\mathbf{a}^{(n)}\}$ are finitely shift equivalent. Such an equivalence relation provides the relations between the inner structures of two shift spaces. Roughly speaking, two mosaic solution spaces are finitely shift equivalent revealing that such two spaces produce "almost" the same structures and share the same complexities. Theorem 1.5 shows that the topological entropy is the complete invariant for the finite shift equivalent. This extends the classical result on sofic systems [27] to path sets. MCNNs seems a natural application of this new result. We achieve the above goals by raising a series of interesting problems.

Problem 1 How to calculate the topological entropy of $\mathbf{Y_a}$ and $\mathbf{Y}_\mathbf{a}^{(m)}$, $1 \le m \le n$?

Such a topic reveals the deep relationship with symbolic dynamical systems. We recall some recent results first. In 1-*d* CNN, it has been proved that the space of the mosaic solutions forms a 1-*d subshift of finite type* (SFT, [24]). Recently, it has also been proved that the mosaic solution of a MCNN forms a *sofic space* ([7,10,11]), which is a factor of SFT. The mosaic

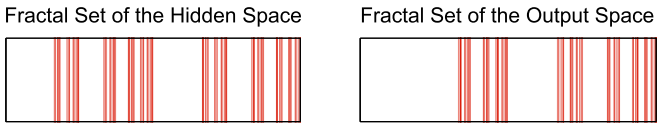


Fig. 1 The fractal set of the hidden and output spaces of the two-layer CNN in Example 2.2. Assigning the patterns -1 and $+1$ by 0 and 1 embeds $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$ in the closed interval $[0, 1]$ by the expansion $\phi(\alpha) = \Sigma_{i\geq 1}\frac{\alpha_i}{2^i}$, where $\alpha = (\alpha_i) \in \{0, 1\}^{\mathbb{N}}$. The fractal sets on the *left* and *right* hand sides are the hidden and output spaces, respectively

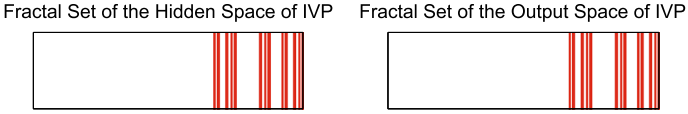


Fig. 2 The fractal set of the hidden and output spaces of IVP of the two-layer CNN in Example 2.2. It is demonstrated that $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{a}}^{(2)}$ exhibit identical topological entropy, and hence they are finitely equivalent

solution of $\mathbf{Y}_{\mathbf{a}}$ and $\mathbf{Y}_{\mathbf{a}}^{(m)}$, $1 \leq m \leq n$, indeed produce new shift spaces, called *path sets* [2]. Such shift spaces are recently defined by Abram and Lagarias, which is a generalisation of classical sofic space, and has many applications on number theory (p-adic expansions [1]). A main difference between a sofic shift and a path set is that a path set may not be invariant under the shift map $\sigma(x)_i = x_{i+1}$, where $x \in \mathcal{A}^{\mathbb{N}}$, $\mathcal{A}$ is a finite set. Reader is referred to [2] and the references therein for more details.

The difference between X and $X_{\mathbf{a}}$ is that $X_{\mathbf{a}}$ may be observed as a projection of X, and $X_{\mathbf{a}}$ is not invariant under the shift map (defined later), where $X = \mathbf{Y}, \mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}$. Namely, X is topologically conjugated to a sofic shift in symbolic dynamical systems, while $X_{\mathbf{a}}$ is topologically conjugated to a path set in general. Such essentially different topological structure motivates the investigation. Figures 1 and 2 present the output and hidden spaces with/without the constraint of initial condition. To answer Problem 1, recall that the number of distinct solution spaces of (1) without initial condition is finite, if the neighborhood $\mathcal{N}$ is fixed. More precisely, suppose $\mathcal{N} = \{-d, \ldots, -1, 0, 1, \ldots, d\}$ for some $d \in \mathbb{N}$. The following proposition demonstrates that the parameter space is partitioned into finitely equivalent subregions.

Proposition 1.1 (See [11]) *Let $\mathcal{P}_N$ be the parameter space of (1), where $N = (4d + 3)n - 2d - 1$. There is a positive integer K and unique set of open subregions $\{P_k\}_{k=1}^{K}$ satisfying*

(1) *$\mathcal{P}_n = \bigcup_{k=1}^{K} \overline{P}_k$, where $\overline{U}$ refers to the closure of U;*
(2) *$P_i \bigcap P_j = \varnothing$ if $i \neq j$;*
(3) *Templates $\mathbb{T}, \mathbb{T}' \in P_k$ for some k if and only if $\mathbf{Y}_{\mathbb{T}} = \mathbf{Y}_{\mathbb{T}'}$.*

A straightforward examination infers that Proposition 1.1 still holds for the solution spaces of IVP of (1), as it does for the output and hidden spaces.

This following result indicates that these spaces of IVP of (1) are topologically conjugated to path sets.

Theorem 1.2 *Given $\mathbf{a} \in \mathbb{R}^n$. Suppose $\mathbf{Y}_{\mathbf{a}}$, $\mathbf{Y}_{\mathbf{a}}^{(n)}$, and $\mathbf{Y}_{\mathbf{a}}^{(m)}$ are the solution, output, and hidden spaces of (1) with initial condition (8), $1 \leq m \leq n - 1$. Then $\mathbf{Y}_{\mathbf{a}}$ and $\mathbf{Y}_{\mathbf{a}}^{(m)}$ are topologically conjugated to path sets for $1 \leq m \leq n$.*

Theorem 1.3 gives an affirmative answer for Problem 1.

Theorem 1.3 *Suppose* $\mathbf{X} \in \{\mathbf{Y_a}, \mathbf{Y_a}^{(1)}, \ldots, \mathbf{Y_a}^{(n)}\}$ *and* $(\mathcal{G}, v)$ *is a reachable presentation (defined later) of* $\mathbf{X}$*. If the labeled graph* $\mathcal{G}$ *is right-resolving, then*

$$h(\mathbf{X}) = \log \lambda, \tag{9}$$

where λ *is the spectral radius of the transition matrix* $\mathbf{T}$ *of the underlying graph* G *of* $\mathcal{G}$*. Herein a labeled graph is* right-resolving *if no two edges from the one vertex carry the same label.*

Notably, every path set has a reachable right-resolving labeled graph presentation (cf. [2] Theorem 3.2). After investigating the spatial complexity of the space of IVP of MCNNs, it is natural to ask the following problem.

Problem 2 Suppose the template of a MCNN is given. (In this case, the solution space $\mathbf{Y}$ is determined.) Is the system independent of the initial condition? More specifically, is $\mathbf{Y_a}$ identical to $\mathbf{Y_b}$ up to some shifts, where $\mathbf{Y_a}$ and $\mathbf{Y_b}$ are the solution spaces with respect to two different initial conditions?

We say a MCNN is *independent of initial condition* if for any two initial conditions $\mathbf{a}$ and $\mathbf{b}$, there exists $k \in \mathbb{N}$ such that either $\sigma^k(\mathbf{Y_a}) = \mathbf{Y_b}$ or $\sigma^k(\mathbf{Y_b}) = \mathbf{Y_a}$. Roughly speaking, a MCNN is independent of initial condition if each pattern of the solution under initial condition one can be obtained by eliminating some digits of a pattern of the solution space under initial condition two.

Theorem 1.4 *Suppose* G *is the graph presentation (defined later) of the solution space of a MCNN. If* G *is irreducible, then the system is independent of initial condition. Moreover,* $h(\mathbf{X_a}) = h(\mathbf{X_b})$ *for any initial conditions* $\mathbf{a}, \mathbf{b}$*, where* $\mathbf{X} \in \{\mathbf{Y}, \mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(n)}\}$.

Furthermore, we investigate the *finite equivalence* between two path sets. Theorem 1.2 demonstrates that the hidden and output spaces of MCNNs under initial condition are topologically conjugated to path sets. That means the investigation of the topological structure of the hidden and output spaces under initial condition is equivalent to elucidating the topological structure of path sets, and finite equivalence between two path sets is an important topological relation. The formal definition of finite equivalence is referred to Definition 3.8. Roughly speaking, two path sets $\mathcal{P}_1$ and $\mathcal{P}_2$ are finitely equivalent if there exists a graph G together with vertices v_1, v_2 and two labeling $\mathcal{L}_1, \mathcal{L}_2$ such that $(\mathcal{G}_i, v_i)$ is either a right-resolving or a left-resolving presentation of $\mathcal{P}_i$ for $i = 1, 2$, where $\mathcal{G}_i = (G, \mathcal{L}_i)$. Note here that a labeled graph is called left-resolving if no two edges end at one vertex carrying the same label. Theorem 1.5 shows that identical topological entropy is a necessary and sufficient condition for two spaces of IVP of MCNNs being finitely equivalent.

Theorem 1.5 *Given* $\mathbf{a} \in \mathbb{R}^n$*. Suppose* $\mathbf{Y_a}, \mathbf{Y_a}^{(n)}$*, and* $\mathbf{Y_a}^{(m)}$ *are the solution, output, and hidden spaces of* (1) *with initial condition* (8)*,* $1 \leq m \leq n-1$ *. For* $\mathbf{X}_1, \mathbf{X}_2 \in \{\mathbf{Y_a}, \mathbf{Y_a}^{(1)}, \ldots, \mathbf{Y_a}^{(n)}\}$*, if* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *are irreducible, then* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *are finitely equivalent if and only if they admit the same topological entropy.*

The main contribution of the present paper is revealing the topological conjugacy between the initial value problems of differential systems and symbolic dynamical systems. This infers that the more properties we know about path sets, the more we know about the topological structure of the solution spaces derived from differential equations with initial conditions, and vice versa. It is essential to find an equivalent relation in systems of differential equations

so that we can characterize how many different types of dynamical phenomena are exhibited and investigate their detailed structures. To achieve this goal, an invariant quantum is needed, and topological entropy is one of the well-known invariant quantum. A classical result, which features topological entropy, in symbolic dynamical system is the so-called *finite equivalence* in Markov shift spaces. More precisely, two Markow shift spaces are finitely equivalent if and only if they have identical topological entropy. In this case, there is a covering space and two finite-to-one factors from the covering space to these two Markov shift spaces. This elucidation extends finite equivalence to path sets. In other words, two initial value problems of MCNNs are finitely equivalent if and only if they have coincident topological entropy. This result is novel in symbolic dynamics and is important for differential equations. Most important, the whole process is routinely checkable.

The rest of this work is organized as follows. Section 2 addresses two examples for our main results. The proofs of theorems are postponed to Sect. 3, and the equivalent partition of the parameter space is elucidated in Appendix.

2 Examples

To give an over all picture of the present investigation, we examine two examples in this section. The detailed discussion is revealed in Sect. 3. First example addresses the difference between the hidden and output spaces of a MCNN and the influence of the initial condition.

Example 2.1 Suppose the MCNN is given by

$$\begin{cases} \dfrac{d}{dt}x_i^{(2)}(t) = -x_i^{(2)}(t) + 2.5 + 4y_i^{(2)}(t) + 6y_{i+1}^{(2)}(t) + 3y_i^{(1)}(t) + 2y_{i+1}^{(1)}(t), \\ \dfrac{d}{dt}x_i^{(1)}(t) = -x_i^{(1)}(t) - 0.3 + 2y_i^{(1)}(t) - y_{i+1}^{(1)}(t). \end{cases} \tag{10}$$

To elucidation of the solution space, we should identify the basic set of admissible local patterns first (see the Appendix). A straightforward verification infers that the basic set of admissible local patterns is

$$\mathcal{B} = \left\{ \boxed{\begin{matrix} - & - \\ - & - \end{matrix}}, \boxed{\begin{matrix} - & - \\ - & + \end{matrix}}, \boxed{\begin{matrix} - & - \\ + & - \end{matrix}}, \boxed{\begin{matrix} + & - \\ + & - \end{matrix}}, \boxed{\begin{matrix} + & + \\ - & - \end{matrix}}, \boxed{\begin{matrix} + & + \\ - & + \end{matrix}}, \boxed{\begin{matrix} + & + \\ + & - \end{matrix}} \right\}$$

After determining the admissible local patterns, the transition matrix (defined in (12)) is used for elucidating the spatial complexity of the system. In this case, the transition matrix $\mathbf{T}$ of the solution space $\mathbf{Y}$ is

$$\mathbf{T} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix},$$

and its graph presentation is seen in Fig. 3. However, the transition matrix does not present the hidden and output spaces. To reveal the dynamics of the hidden and output spaces, we introduce the symbolic transition matrix (defined in (15)). The symbolic transition matrices of $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$ of (2.1) are

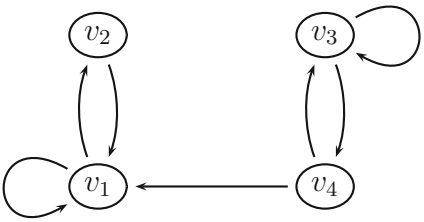


Fig. 3 The graph presentation of the solution space in Example 2.1. It consists of two irreducible subgraphs

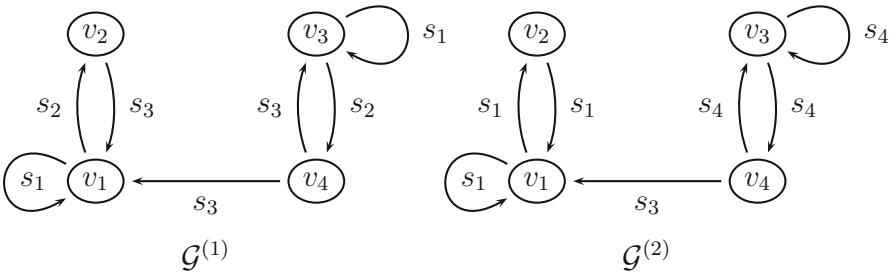


Fig. 4 The labeled graph presentation of the hidden and output spaces in Example 2.1. $\mathcal{G}^{(1)}$ represents the hidden space $\mathbf{Y}^{(1)}$ while $\mathcal{G}^{(2)}$ represents the output space $\mathbf{Y}^{(2)}$. It is seen that neither $\mathcal{G}^{(1)}$ nor $\mathcal{G}^{(2)}$ are right-resolving

$$\mathbf{S}^{(1)} = \begin{pmatrix} s_1 & s_2 & \varnothing & \varnothing \\ s_3 & \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & s_1 & s_2 \\ s_3 & \varnothing & s_3 & \varnothing \end{pmatrix} \quad \text{and} \quad \mathbf{S}^{(2)} = \begin{pmatrix} s_1 & s_1 & \varnothing & \varnothing \\ s_1 & \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & s_4 & s_4 \\ s_3 & \varnothing & s_4 & \varnothing \end{pmatrix}$$

respectively, where

$$s_1 = --, \quad s_2 = -+, \quad s_3 = +-, \quad s_4 = ++.$$

The labeled graph presentation of the hidden and output spaces are pictured in Fig. 4. Notably the topological dynamics of the solution spaces with initial vertices v_1 and v_4 are different. Suppose the initial condition $\mathbf{a}$ is given such that the graph presentation of $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{a}}^{(2)}$ are $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ with initial vertex v_1 respectively. Theorem 3.6 asserts that the path topological entropy of the hidden space is $h_P(\mathbf{Y}_{\mathbf{a}}^{(1)}) = \log g$ while the path topological entropy of the output space is $h_P(\mathbf{Y}_{\mathbf{a}}^{(2)}) = 0$. The detail discussion is elucidated in Sect. 3.

Example 2.2 demonstrates the finite equivalence of the hidden and output spaces. The upcoming example also presents a MCNN which is independent of initial condition. Furthermore, due to the labeled graph presentation of the hidden and output spaces are not right-resolving, the investigation of Example 2.2 is more complicated than the previous example.

Example 2.2 Consider the MCNN given by

$$\begin{cases} \dfrac{d}{dt}x_i^{(2)}(t) = -x_i^{(2)}(t) + 0.9 - 0.3y_i^{(2)}(t) - 1.2y_{i+1}^{(2)}(t) + 0.7y_i^{(1)}(t) + 2.3y_{i+1}^{(1)}(t), \\ \dfrac{d}{dt}x_i^{(1)}(t) = -x_i^{(1)}(t) + 0.9 + 2.9y_i^{(1)}(t) + 1.7y_{i+1}^{(1)}(t). \end{cases} \tag{11}$$

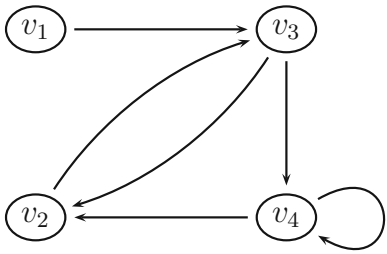


Fig. 5 The graph presentation of the solution space in Example 2.2

Then the basic set of admissible local patterns is

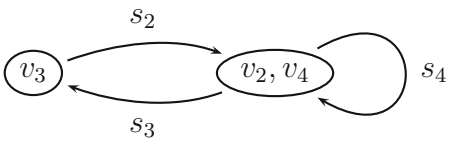
$$\mathcal{B} = \left\{ \boxed{\begin{matrix} - & + \\ - & - \end{matrix}}, \boxed{\begin{matrix} - & + \\ + & - \end{matrix}}, \boxed{\begin{matrix} + & - \\ - & + \end{matrix}}, \boxed{\begin{matrix} + & - \\ + & + \end{matrix}}, \boxed{\begin{matrix} + & + \\ - & + \end{matrix}}, \boxed{\begin{matrix} + & + \\ + & + \end{matrix}} \right\}.$$

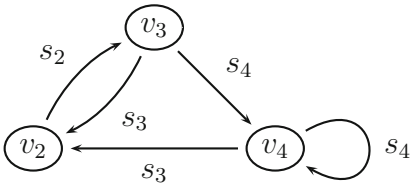
The transition matrix $\mathbf{T}$ of the solution space $\mathbf{Y}$ is

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

and its graph presentation is seen in Fig. 5.

Suppose the initial condition is given so that the initial vertex of $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{a}}^{(2)}$ is v_3. In this case, the right-resolving presentations of the hidden and output spaces after applying subset construction method (introduced in Sect. 3.2) are irreducible and are shown as follows.





It comes immediately that the transition matrices of $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{a}}^{(2)}$ are

$$\mathbf{T}^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{T}^{(2)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

respectively. It is easy to see that the MCNN is independent of initial condition and Theorem 3.6 infers that $h_P(\mathbf{Y}_{\mathbf{a}}^{(1)}) = h_P(\mathbf{Y}_{\mathbf{a}}^{(2)}) = \log g$ for all initial conditions, where $g = (1+\sqrt{5})/2$ is the golden mean. Theorem 3.9 demonstrates that $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{a}}^{(2)}$ are finitely equivalent, and their common underlying graph is seen in Fig. 6. The detailed investigation is referred to Example 3.10.

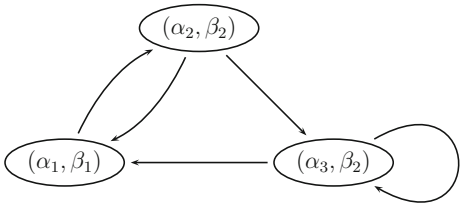


Fig. 6 The common graph presentation of the hidden and output spaces in Example 2.2

3 The Complexity of Initial Value Problems of Multi-layer Cellular Neural Networks

This section elucidates the complexity of initial value problems of MCNNs. We introduce a methodology for the computation of an indicator of the complexity. By complexity we refer to the topological entropy.

Proposition 1.1 (also known as *the separation property*) infers there are finitely many kinds of nonequivalent templates for MCNNs if we focus on mosaic solutions. (For the reader's convenience, we examine the separation property in the Appendix.) More than that, the investigation of MCNNs is equivalent to the investigation of the basic sets of admissible local patterns. Recall that we refer to $\mathbf{Y}$, $\mathbf{Y}^{(1)}$, and $\mathbf{Y}^{(2)}$ the solution, hidden, and output spaces without the constraint of the initial condition and refer to $\mathbf{Y_a}$, $\mathbf{Y_a}^{(1)}$, and $\mathbf{Y_a}^{(2)}$ those with initial condition $\mathbf{a}$. Suppose a basic set of admissible local patterns $\mathcal{B}$ is given, the solution space $\mathbf{Y}$ is then represented as

$$\mathbf{Y} = \left\{ \begin{pmatrix} y_i \\ u_i \end{pmatrix}_{i\in\mathbb{N}} : \boxed{\begin{matrix} y_i y_{i+1} \\ u_i u_{i+1} \end{matrix}} \in \mathcal{B} \text{ for } i \in \mathbb{N} \right\};$$

the output space $\mathbf{Y}^{(2)}$ and hidden space $\mathbf{Y}^{(1)}$ are

$$\mathbf{Y}^{(2)} = \left\{ (y_i)_{i\in\mathbb{N}} : \begin{pmatrix} y_i \\ u_i \end{pmatrix}_{i\in\mathbb{N}} \in \mathbf{Y} \text{ for some } (u_i)_{i\in\mathbb{N}} \right\}$$

and

$$\mathbf{Y}^{(1)} = \left\{ (u_i)_{i\in\mathbb{N}} : \begin{pmatrix} y_i \\ u_i \end{pmatrix}_{i\in\mathbb{N}} \in \mathbf{Y} \text{ for some } (y_i)_{i\in\mathbb{N}} \right\}$$

respectively.

3.1 The Topological Structure of the Solution, Hidden, and Output Spaces

First considering (23) without initial condition, substitute mosaic patterns -1 and 1 with symbols $-$ and $+$, respectively. Define the ordering matrix of $\{-, +\}^{\mathbb{Z}_{2\times 2}}$ by

$$\mathbb{X} = \begin{array}{c|cccc} & \boxed{\begin{matrix}-\\-\end{matrix}} & \boxed{\begin{matrix}-\\+\end{matrix}} & \boxed{\begin{matrix}+\\-\end{matrix}} & \boxed{\begin{matrix}+\\+\end{matrix}} \\ \hline \boxed{\begin{matrix}-\\-\end{matrix}} & \boxed{\begin{matrix}--\\--\end{matrix}} & \boxed{\begin{matrix}--\\-+\end{matrix}} & \boxed{\begin{matrix}-+\\--\end{matrix}} & \boxed{\begin{matrix}-+\\-+\end{matrix}} \\ \boxed{\begin{matrix}-\\+\end{matrix}} & \boxed{\begin{matrix}--\\+-\end{matrix}} & \boxed{\begin{matrix}--\\++\end{matrix}} & \boxed{\begin{matrix}-+\\+-\end{matrix}} & \boxed{\begin{matrix}-+\\++\end{matrix}} \\ \boxed{\begin{matrix}+\\-\end{matrix}} & \boxed{\begin{matrix}+-\\--\end{matrix}} & \boxed{\begin{matrix}+-\\-+\end{matrix}} & \boxed{\begin{matrix}++\\--\end{matrix}} & \boxed{\begin{matrix}++\\-+\end{matrix}} \\ \boxed{\begin{matrix}+\\+\end{matrix}} & \boxed{\begin{matrix}+-\\+-\end{matrix}} & \boxed{\begin{matrix}+-\\++\end{matrix}} & \boxed{\begin{matrix}++\\+-\end{matrix}} & \boxed{\begin{matrix}++\\++\end{matrix}} \end{array} = (x_{pq})_{1\le k,l\le 4}$$

We emphasize that each entry in $\mathbb{X}$ is a 2×2 pattern since $\mathcal{B}$ consists of 2×2 local patterns. Once the size of local patterns varies, there exists a corresponding ordering matrix which represents the basic set of admissible local patterns. Suppose that $\mathcal{B}$ is given. The transition matrix $\mathbf{T}\in\mathcal{M}_4(\mathbb{R})$ is a 4×4 matrix defined by

$$\mathbf{T}(k,l)=\begin{cases}1, & \text{if } x_{kl}\in\mathcal{B};\\ 0, & \text{otherwise.}\end{cases} \tag{12}$$

After determining the transition matrix, the solution space $\mathbf{Y}$ can be described via a *directed graph* as follows.

Let

$$v_1\equiv\boxed{\begin{matrix}-\\-\end{matrix}},\quad v_2\equiv\boxed{\begin{matrix}-\\+\end{matrix}},\quad v_3\equiv\boxed{\begin{matrix}+\\-\end{matrix}},\quad v_4\equiv\boxed{\begin{matrix}+\\+\end{matrix}}.$$

Then v_i is in the vertex set $\mathcal{V}$ if and only if neither the i-th row nor the i-th column of $\mathbf{T}$ are zero vectors. There exists an edge $e=(v_{i_1},v_{i_2})\in\mathcal{E}$ from v_{i_1} to v_{i_2} if and only if $\mathbf{T}(i_1,i_2)=1$. The graph $G=(\mathcal{V},\mathcal{E})$ is called a *graph presentation* of $\mathbf{T}$. A graph is called *irreducible* is its corresponding matrix is irreducible; namely, there exists a path for any two vertices. It follows immediately that the solution space $\mathbf{Y}$ is the collection of infinite paths in G. More precisely, set

$$X_G=\{(g_i)_{i\in\mathbb{N}}: g_i\in\mathcal{V},(g_i,g_{i+1})\in\mathcal{E}\text{ for } i\in\mathbb{N}\},$$

then $\mathbf{Y}=X_G$. In other words, $\mathbf{Y}$ is a *shift of finite type* (SFT) in symbolic dynamical systems (cf. [10,24]). For example, suppose the basic set of admissible local patterns $\mathcal{B}$ is given as in Example 2.1. Then the transition matrix of the solution space $\mathbf{Y}$ is

$$\mathbf{T}=\begin{pmatrix}1&1&0&0\\1&0&0&0\\0&0&1&1\\1&0&1&0\end{pmatrix},$$

and the graph presentation $G=(\mathcal{V},\mathcal{E})$ is seen in Fig. 3.

For ease of notation, denote $\boxed{\begin{matrix}y_1y_2\\u_1u_2\end{matrix}}$ by $y_1y_2\diamond u_1u_2$ and

$$\mathbf{y}\diamond\mathbf{u}\equiv\boxed{\begin{matrix}\mathbf{y}\\\mathbf{u}\end{matrix}}=\begin{matrix}y_1y_2y_3\cdots\\u_1u_2u_3\cdots\end{matrix},\quad\text{where}\quad\mathbf{y}=(y_i)_{i\in\mathbb{N}},\mathbf{u}=(u_i)_{i\in\mathbb{N}}.$$

Then we can write $\mathbf{Y}$ as

$$\mathbf{Y} = \{\mathbf{y} \diamond \mathbf{u} : \mathbf{y} \in \mathbf{Y}^{(2)}, \mathbf{u} \in \mathbf{Y}^{(1)}\}.$$

Since the solution space consists of the output and hidden spaces, the dynamical behavior of the output space $\mathbf{Y}^{(2)}$ is influenced by the hidden space $\mathbf{Y}^{(1)}$, and vice versa. For instance, a phenomenon which cannot be seen in one-layer cellular neural networks is that $\mathbf{Y}^{(1)}$ would break the symmetry of the entropy diagram of $\mathbf{Y}^{(2)}$ [8,11]. This motivates the study of the IVP for $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$.

Since the output space $\mathbf{Y}^{(2)}$ and the hidden space $\mathbf{Y}^{(1)}$ are both factors of the solution space $\mathbf{Y}$, the transition matrix $\mathbf{T}$ and its corresponding graph G are not capable of elucidating $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$. Aside from the transition matrix and directed graph introduced in the previous subsection, the so-called *symbolic transition matrix* and *labeled graph* are presented instead.

Let $\mathcal{A} = \{s_1, s_2, s_3, s_4\}$, where

$$s_1 = --, \quad s_2 = -+, \quad s_3 = +-, \quad s_4 = ++.$$

For $\alpha_1 = \alpha_{1;1}\alpha_{1;2}, \alpha_2 = \alpha_{2;1}\alpha_{2;2} \in \mathcal{A}$, denote the compound symbol $\alpha_1\alpha_2$ by $\alpha_{1;1}\alpha_{1;2}\alpha_{2;2}$ if and only if $\alpha_{1;2} = \alpha_{2;1}$. In such a manner we can define a word of arbitrary length with symbols being in $\mathcal{A}$. Suppose $G = (\mathcal{V}, \mathcal{E})$ and $\mathbf{T}$ are the graph presentation and transition matrix of $\mathbf{Y}$. Define $\mathcal{L}^{(1)}, \mathcal{L}^{(2)} : \mathcal{E} \to \mathcal{A}$ by

$$\mathcal{L}^{(1)}(e) = s_{2\tau(i-1)+\tau(j-1)+1}, \quad \tau(c) := c \bmod 2; \tag{13}$$

$$\mathcal{L}^{(2)}(e) = s_{2[(i-1)/2]+[(j-1)/2]+1}, \quad [\cdot] \text{ is the Gauss function}; \tag{14}$$

where $e = (v_i, v_j)$. These two labeling functions $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ define two labeled graphs $\mathcal{G}^{(1)} = (G, \mathcal{L}^{(1)})$ and $\mathcal{G}^{(2)} = (G, \mathcal{L}^{(2)})$, respectively. For $\ell = 1, 2$, set

$$X_{\mathcal{G}^{(\ell)}} = \{(\omega_i)_{i\in\mathbb{N}} : \mathcal{L}^{(\ell)}(e_i) = \omega_i, e_i \in \mathcal{E} \text{ for } i \in \mathbb{N}\}.$$

In [10,11], the authors demonstrated that the output space $\mathbf{Y}^{(2)}$ is topologically conjugated to $X_{\mathcal{G}^{(2)}}$ and the hidden space $\mathbf{Y}^{(1)}$ is topologically conjugated to $X_{\mathcal{G}^{(1)}}$. A space that is represented as a labeled graph is called a *sofic shift*, which is an extension of a SFT. The reader is referred to [27] for more details.

Similar to extending a directed graph to a labeled graph, the transition matrix can be extended to a symbolic transition matrix. The symbolic transition matrix $\mathbf{S}^{(\ell)}$ with respect to the labeled graph $\mathcal{G}^{(\ell)}$ is defined by

$$\mathbf{S}^{(\ell)}(p, q) = \begin{cases} \alpha_j, & \text{if } \ T^{(\ell)}(p, q) = 1 \ \text{ and } \ \mathcal{L}^{(\ell)}((v_p, v_q)) = \alpha_j; \\ \varnothing, & \text{otherwise.} \end{cases} \tag{15}$$

Herein the empty symbol $\varnothing$ means there exists no local pattern in $\mathcal{B}$ related to its corresponding entry in the ordering matrix. A labeled graph is called *right-resolving* if edges start from the same vertex carrying different labels. Similarly, a labeled graph is called *left-resolving* if edges end at the same vertex carrying different labels. It follows from this definition that the symbolic transition matrix $\mathbf{S}$ of a right-resolving labeled graph must satisfy $\mathbf{S}(p, q) \neq \mathbf{S}(p, q')$ for all p, q, q' if $\mathbf{S}(p, q) \neq \varnothing$. Conversely, suppose $\mathbf{S}(p, q) \neq \varnothing$. $\mathbf{S}(p, q) \neq \mathbf{S}(p, q')$ for all p, q, q' demonstrates that no two edges starting from the same initial state carry the same label. In other words, a labeled graph is right-resolving if and only if the multiplicity of every nonempty symbol is less than or equal to one in each row of its corresponding symbolic transition matrix.

Continuing with Example 2.1, the graph presentation G of the output space $\mathbf{Y}$ is seen in Fig. 3. The labeled graph presentation $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ of the hidden space $\mathbf{Y}^{(1)}$ and the output space $\mathbf{Y}^{(2)}$ are presented in Fig. 4; their corresponding symbolic transition matrices are

$$\mathbf{S}^{(1)} = \begin{pmatrix} s_1 & s_2 & \varnothing & \varnothing \\ s_3 & \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & s_1 & s_2 \\ s_3 & \varnothing & s_3 & \varnothing \end{pmatrix} \quad \text{and} \quad \mathbf{S}^{(2)} = \begin{pmatrix} s_1 & s_1 & \varnothing & \varnothing \\ s_1 & \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & s_4 & s_4 \\ s_3 & \varnothing & s_4 & \varnothing \end{pmatrix}$$

respectively. Notably neither $\mathcal{G}^{(1)}$ nor $\mathcal{G}^{(2)}$ are right-resolving. In [11], the authors indicated the structure of the solution, hidden, and output spaces.

Theorem 3.1 *Suppose* $\mathbf{Y}$, $\mathbf{Y}^{(1)}$, *and* $\mathbf{Y}^{(2)}$ *are the solution, hidden, and output spaces of a MCNN, respectively. Then* $\mathbf{Y}$ *is a shift of finite type, and* $\mathbf{Y}^{(1)}$ *and* $\mathbf{Y}^{(2)}$ *are sofic shifts.*

Notably a shift of finite type is also a sofic shift. In other words, sofic shifts generalize the concept of shifts of finite type.

Next we consider the topological structure of the solution, hidden, and output spaces under the influence of initial conditions. The initial condition is given by

$$\begin{pmatrix} x_k^{(2)} \\ x_k^{(1)} \end{pmatrix} = \mathbf{a} \quad \text{for some} \quad \mathbf{a} \in \mathbb{R}^2, k \in \mathbb{N}. \tag{16}$$

Without loss of generality, we may assume that $i = 1$. Recall that we focus on the mosaic solutions; namely, the space $\mathbf{X_a}$ defined in (4) consists of $\begin{pmatrix} x_i^{(2)} \\ x_i^{(1)} \end{pmatrix}_{i\in\mathbb{N}}$ satisfying $|x_i^{(\ell)}(t)| > 1$ for $i \in \mathbb{N}$, $\ell = 1, 2$, provided $t \geq T$ for some $T > 0$. In this case, the solution space $\mathbf{Y_a}$, the hidden space $\mathbf{Y}_\mathbf{a}^{(1)}$, and the output space $\mathbf{Y}_\mathbf{a}^{(2)}$ are realized as

$$\mathbf{Y_a} = \left\{ \mathbf{y} = \begin{pmatrix} y_i^{(2)} \\ y_i^{(1)} \end{pmatrix}_{i\in\mathbb{N}} : \begin{matrix} y_1^{(2)} \\ y_2^{(1)} \end{matrix} \in \left\{ \begin{matrix} -, & -, & +, & + \\ -, & +, & -, & + \end{matrix} \right\} \right\} \subseteq \{-, +\}^{\mathbb{N}_\infty \times 2}, \tag{17}$$

and

$$\mathbf{Y}_\mathbf{a}^{(\ell)} = \left\{ y = \left(y_i^{(\ell)} \right)_{i\in\mathbb{N}} : y = \Phi^{(\ell)}(\mathbf{y}), \mathbf{y} \in \mathbf{Y} \right\}, \qquad \ell = 1, 2. \tag{18}$$

More precisely, the solution space $\mathbf{Y_a}$ consists of paths in its corresponding graph G starting from a particular vertex, and so are the hidden and output spaces. Notably, $\mathbf{Y}_\mathbf{a}^{(\ell)}$ consists of labeled paths starting from a particular vertex in $\mathcal{G}^{(\ell)}$ for $\ell = 1, 2$. This concludes that, considering the IVP of (23), the solution space $\mathbf{Y_a}$, the hidden space $\mathbf{Y}_\mathbf{a}^{(1)}$, and the output space $\mathbf{Y}_\mathbf{a}^{(2)}$ are topologically conjugated to the so-called *path sets* in symbolic dynamical systems (cf. [2]).

Definition 3.2 Suppose $\mathcal{G} = (G, \mathcal{L})$ is a labeled graph with underlying directed graph $G = (\mathcal{V}, \mathcal{E})$ and labeling $\mathcal{L} : \mathcal{E} \to \mathcal{A}$. The path set (or *pointed follower set*) $\mathcal{P} = X_\mathcal{G}(v)$ is the subset of $\mathcal{A}^\mathbb{N}$ made up of the symbol sequences of successive edge labels of all possible one-sided infinite walks in $\mathcal{G}$ issuing from the distinguished vertex v. Many different $(\mathcal{G}, v)$ may give the same path set $\mathcal{P} \subset \mathcal{A}^\mathbb{N}$, and we call any such $(\mathcal{G}, v)$ a presentation of $\mathcal{P}$.

Theorem 3.1 and the above discussion derive our first main theorem that the solution, hidden, and output spaces with initial condition are path sets. In other words, we have the following theorem, which is the case that $n = 2$ in Theorem 1.2.

Theorem 3.3 *Suppose* $\mathbf{Y}_{\mathbf{a}}$, $\mathbf{Y}_{\mathbf{a}}^{(1)}$, *and* $\mathbf{Y}_{\mathbf{a}}^{(2)}$ *are the solution, hidden, and output spaces of* (23) *with initial condition* (16). *Then* $\mathbf{Y}_{\mathbf{a}}$, $\mathbf{Y}_{\mathbf{a}}^{(1)}$, *and* $\mathbf{Y}_{\mathbf{a}}^{(2)}$ *are path sets.*

3.2 The Spatial Complexity of the Solution, Hidden, and Output Spaces

The previous subsection reveals that either one of the solution, hidden, and output spaces is topologically conjugated to a path set. Following the topological structure of these spaces, it is on the topic of complexity which this subsection considers. One of the most frequently used quantum for the index of spatial complexity is the topological entropy. Notably, the topological entropy measures the growth rate of a number of patterns of an invariant closed space. Meanwhile, the solution, hidden, and output spaces under initial condition, that are demonstrated as path sets, are not invariant. By complexity, instead of the classical topological entropy, we focus on the so-called *path topological entropy*.

Definition 3.4 (*See* [2]) Suppose $\mathcal{P}$ is a path set. Let $N_n^I(\mathcal{P})$ denote the number of distinct initial blocks of length n in $\mathcal{P}$. The path topological entropy of $\mathcal{P}$ is defined by

$$h_p(\mathcal{P}) = \limsup_{n\to\infty} \frac{1}{n} \log N_n^I(\mathcal{P}). \tag{19}$$

The difference between the topological entropy and the path topological entropy is that the topological entropy considers the growth rate of *all* distinct blocks in the given space (cf. [3]). More precisely, let $N_n(P)$ denote the number of distinct blocks of length n occurring anywhere in a path set $\mathcal{P}$. The topological entropy of $\mathcal{P}$ is

$$h_{top}(\mathcal{P}) = \limsup_{n\to\infty} \frac{1}{n} \log N_n(\mathcal{P}). \tag{20}$$

In [2], the authors showed that the topological entropy of a path set coincides with its path topological entropy.

Theorem 3.5 (See [2]) *For a path set* $\mathcal{P}$,

$$h_P(\mathcal{P}) = h_{top}(\mathcal{P}). \tag{21}$$

It is known that the topological entropy of a sofic shift relates to the spectral radius of its corresponding transition matrix (cf. [27]). Theorems 3.3 and 3.5 indicate that the path topological entropy of either one of the solution, hidden, and output spaces also relates to the spectral radius of its corresponding transition matrix if the transition matrix comes from a *reachable* presentation. Herein a presentation $(\mathcal{G}, v)$ of a path set $\mathcal{P}$ is called reachable if each vertex of $\mathcal{G}$ can be reached by a directed path from v. This derives the following theorem, which is the case that $n = 2$ in Theorem 1.3.

Theorem 3.6 *Suppose* $\mathbf{X} \in \{\mathbf{Y}_{\mathbf{a}}, \mathbf{Y}_{\mathbf{a}}^{(1)}, \mathbf{Y}_{\mathbf{a}}^{(2)}\}$ *and* $(\mathcal{G}, v)$ *is a reachable presentation of* $\mathbf{X}$. *If the labeled graph* $\mathcal{G}$ *is right-resolving, then*

$$h_P(\mathbf{X}) = \log \lambda, \tag{22}$$

where λ *is the spectral radius of the transition matrix* $\mathbf{T}$ *of the underlying graph* G *of* $\mathcal{G}$.

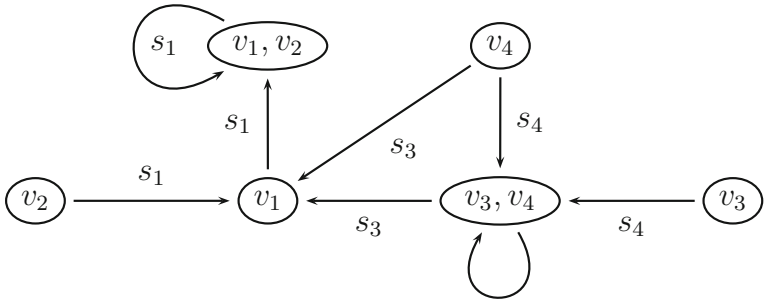


Fig. 7 The right-resolving labeled graph presentation of the output spaces in Example 2.1 after applying SCM

In general, the labeled graph of a presentation may not be right-resolving (cf. Fig. 4). Nevertheless, applying the so-called *subset construction method* (SCM) to the original labeled graph $\mathcal{G}$ derives a right-resolving labeled graph $\mathcal{H}$. Moreover, the space represented by $\mathcal{H}$ is identical to the original one, say, $X_{\mathcal{G}} = X_{\mathcal{H}}$. It is easily seen that, by applying SCM, every path set has a reachable right-resolving presentation $(\mathcal{G}, v)$. We introduce SCM to make the investigation self-contained. For the details, the reader is referred to [2,27].

Subset Construction Method Let X be a sofic shift over the alphabet $\mathcal{A}$ having a presentation $\mathcal{G} = (G, \mathcal{L})$. If $\mathcal{G}$ is not right-resolving, then a new labeled graph $\mathcal{H} = (H, \mathcal{L}_H)$ is constructed as follows. The vertices I of H are the nonempty subsets of the vertex set $\mathcal{V}(G)$ of G. If $I \in \mathcal{V}(H)$ and $a \in \mathcal{A}$, let J denote the set of terminal vertices of edges in G starting at some vertices in I and labeled a, i.e., J is the set of vertices reachable from I using the edges labeled a.

1) If $J = \varnothing$, do nothing.
2) If $J \neq \varnothing$, $J \in \mathcal{V}(H)$ and draw an edge in H from I to J labeled a.

Carrying this out for each $I \in \mathcal{V}(H)$ and each $a \in \mathcal{A}$ produces the labeled graph $\mathcal{H}$. Then, each vertex I in H has at most one edge with a given label starting at I. This implies that $\mathcal{H}$ is right-resolving.

We apply SCM to $\mathcal{G}^{(2)}$, a labeled graph presentation of the output space $\mathbf{Y}^{(2)}$ in Example 2.1, as an example. It is seen that the label on both the edges from v_1 to v_2 and from v_1 to itself are s_1, and the label on the edges starting from v_3 are s_4. Hence two new vertices $\{v_1, v_2\}$ and $\{v_3, v_4\}$ are constructed. The newly generated labeled graph $\mathcal{H}^{(2)}$ is obtained as in Fig. 7.

Recall that a MCNN (23) is *independent of initial conditions* if, for any two initial conditions $\mathbf{a}, \mathbf{b}$, there exists $k \in \mathbb{N}$ such that either $\sigma^k(\mathbf{X_a}) = \mathbf{X_b}$ or $\sigma^k(\mathbf{X_b}) = \mathbf{X_a}$, where $\mathbf{X} \in \{\mathbf{Y}, \mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}\}$. An immediate consequence of Theorem 3.5 is that a MCNN is independent of initial condition if the graph presentation G of the solution space without initial condition is irreducible. The proof is straightforward, and thus is omitted. Example 3.10 infers that an MCNN which is independent of initial condition does not relate to an irreducible graph presentation. This derives the following theorem, which is the case that $n = 2$ in Theorem 1.4.

Theorem 3.7 *Let G be the graph presentation of the solution space of a MCNN without initial condition which is obtained from the basic set of admissible local patterns. Then the prescribed MCNN is independent of initial condition if G is irreducible.*

One of the well-known applications of entropy is classifying shift spaces into equivalent classes (cf. [27]). Aside from calculating the path topological entropy of the solution, hidden, and output spaces of MCNNs, it is interesting to investigate the finer structure, say, the relationship between the hidden and output spaces. Herein we focus on the *finite shift equivalence* of spaces. Let X and Y be two shift spaces. A map $\phi : X \to Y$ is called a *factor map* if ϕ is onto. A factor map $\phi : X \to Y$ is *finite-to-one* if there exists $M \in \mathbb{N}$ such that the cardinality of $\phi^{-1}(y)$ is less than or equal to M for $y \in Y$, where $\phi^{-1}(y)$ indicates the pre-image of y. Two shift spaces X and Y are *finitely equivalent* (FE), denoted by $X \sim_{\mathcal{F}} Y$, if there exists a shift of finite type W together with finite-to-one factor maps $\phi_X : W \to X$ and $\phi_Y : W \to Y$. We say that W is a *common extension* of X and Y, and the triple (W, ϕ_X, ϕ_Y) is a finite equivalence between X and Y.

Definition 3.8 Two path sets $\mathcal{P}_1$ and $\mathcal{P}_2$ are said to be finitely equivalent if there exist an irreducible directed graph G and two factors Ψ_1, Ψ_2 such that, for $i = 1, 2$, $\Psi_i : X_G \to X_{\mathcal{G}_i}$ is finite-to-one, where $(\mathcal{G}_i, v_i)$ is a presentation of $\mathcal{P}_i$.

Notably, path sets are not shift invariant in general. (cf. [2]). Theorem 3.9 asserts that irreducible path sets are FE if and only if they are carrying the same path topological entropy. Hence we derive the following theorem, which is the case that $n = 2$ in Theorem 1.5.

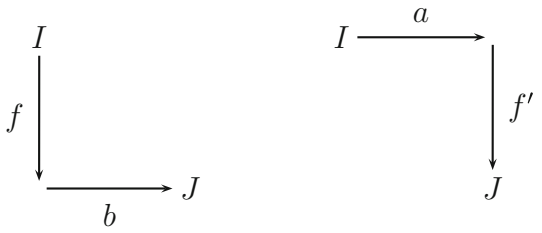
Theorem 3.9 *Suppose* $\mathbf{Y}_{\mathbf{a}}^{(1)}$ *and* $\mathbf{Y}_{\mathbf{b}}^{(2)}$ *are the hidden and output spaces of a MCNN with initial condition. If* $\mathbf{Y}_{\mathbf{a}}^{(1)}$ *and* $\mathbf{Y}_{\mathbf{b}}^{(2)}$ *are both irreducible, then* $\mathbf{Y}_{\mathbf{a}}^{(1)} \sim_{\mathcal{F}} \mathbf{Y}_{\mathbf{b}}^{(2)}$ *if and only if* $h_P(\mathbf{Y}_{\mathbf{a}}^{(1)}) = h_P(\mathbf{Y}_{\mathbf{b}}^{(2)})$.

Proof $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{b}}^{(2)}$ share the same path topological entropy coming immediately from the FE of $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{b}}^{(2)}$. It suffices to show that $h_P(\mathbf{Y}_{\mathbf{a}}^{(1)}) = h_P(\mathbf{Y}_{\mathbf{b}}^{(2)})$ infers $\mathbf{Y}_{\mathbf{a}}^{(1)} \sim_{\mathcal{F}} \mathbf{Y}_{\mathbf{b}}^{(2)}$.

Since the presentations of $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{b}}^{(2)}$ come from the same underlying graph G, it is seen that they have identical path topological entropy whenever both their presentations are right-resolving. It remains to show that at least one of the presentation is not right-resolving. The key to showing that $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{b}}^{(2)}$ are FE is constructing the underlying graph of their presentations. The construction is similar to the proof in [27, Theorem8.3.8], which we restate herein for the reader's convenience.

Applying SCM if necessary, with the abuse of notations, we assume that $\mathcal{G}^{(i)}$ is right-resolving with transition matrix $\mathbf{T}^{(i)}$ for $i = 1, 2$. In [29], Parry demonstrated that two shifts with the same topological entropy infer there exists an integral matrix that commutes with their transition matrices. More specifically, there exist nonnegative integral matrices E, F such that $\mathbf{T}^{(2)}E = E\mathbf{T}^{(1)}$ and $\mathbf{T}^{(1)}F = F\mathbf{T}^{(2)}$. We use the matrix F to construct the desired graph M. Let $A = \mathbf{T}^{(1)}$ and $B = \mathbf{T}^{(2)}$ for ease of notations. First we introduce an auxiliary graph N.

The auxiliary graph N is formed as follows. The vertex set $\mathcal{V}(N)$ of N is the disjointed union of $\mathcal{V}(G^{(1)})$ and $\mathcal{V}(G^{(2)})$, where $G^{(i)}$ is the underlying graph of $\mathcal{G}^{(i)}$ for $i = 1, 2$. There are three types of edges in N: A-edges, B-edges, and F-edges. The A-edges are simply all edges in $G^{(1)}$, now regarded as having initial and terminal states in $\mathcal{V}(N)$. The B-edges are defined similarly. The F-edges are between $G^{(1)}$ and $G^{(2)}$; more specifically, for each pair $I \in \mathcal{V}(G^{(1)})$ and $J \in \mathcal{V}(G^{(2)})$, there are $F(I, J)$ edges in N from I to J. For each path of length two in N from $I \in \mathcal{V}(G^{(1)})$ to $J \in \mathcal{V}(G^{(2)})$ is either a left-bottom path or a top-right path that are seen as follows.



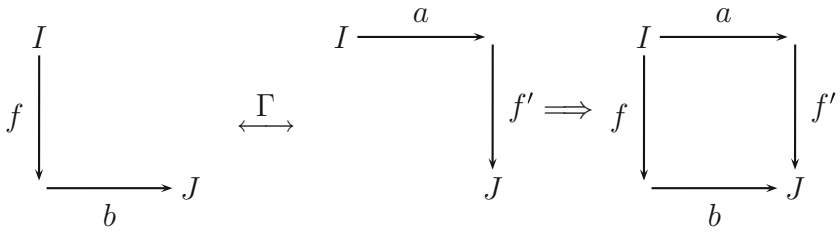
Let $LB(I, J)$ denote the set of all left-bottom paths from I to J, and $TR(I, J)$ the set of all top-right paths from I to J. The equation $(AF)(I, J) = (FB)(I, J)$ asserts that $|LB(I, J)| = |TR(I, J)|$. Hence for each I, J we can define a bijection

$$\Gamma_{IJ} : LB(I, J) \to TR(I, J).$$

We assemble all these bijections to form a bijection

$$\Gamma : \bigcup_{\substack{I\in\mathcal{V}(G^{(1)})\\ J\in\mathcal{V}(G^{(2)})}} LB(I, J) \to \bigcup_{\substack{I\in\mathcal{V}(G^{(1)})\\ J\in\mathcal{V}(G^{(2)})}} TR(I, J).$$

Each left-bottom path fb is paired by Γ to a top-right pair af', and each such pair can be put together to form a closed box as follows.



We denote the resulting box by $\square(f, a, b, f')$. Such a box is determined by its left-bottom path fb and also by its top-right path af'.

Using N to construct the required common extension using a new graph M. The vertex set $\mathcal{V}(M)$ is the set of all F-edges in N, and the edge set $\mathcal{E}(M)$ is the set of all boxes $\square(f, a, b, f')$. The initial state of $\square(f, a, b, f')$ is f and $\square(f, a, b, f')$ terminates at the vertex f'.

Define $\phi^{(1)} : \mathcal{E}(M) \to \mathcal{E}(G^{(1)})$ and $\phi^{(2)} : \mathcal{E}(M) \to \mathcal{E}(G^{(2)})$ by

$$\phi^{(1)}(\square(f, a, b, f')) = a \quad \text{and} \quad \phi^{(2)}(\square(f, a, b, f')) = b,$$

respectively. It is straightforward to verify that $\phi^{(1)}$ is left-resolving and $\phi^{(2)}$ is right-resolving. Furthermore, the edge map $\phi^{(i)} : \mathcal{E}(M) \to \mathcal{E}(G^{(i)})$ induces vertex map $\Phi^{(i)} : \mathcal{V}(M) \to \mathcal{V}(G^{(i)})$ for $i = 1, 2$. This infers there exist two vertices $\overline{v}_1, \overline{v}_2 \in \mathcal{V}(M)$ such that $(M, \mathcal{L}^{(1)} \circ \phi^{(1)}, \overline{v}_1)$ and $(M, \mathcal{L}^{(2)} \circ \phi^{(2)}, \overline{v}_2)$ are presentation of $\mathbf{Y}^{(1)}_{\mathbf{a}}$ and $\mathbf{Y}^{(2)}_{\mathbf{b}}$, respectively.

This completes the proof (Fig. 8). □

Example 3.10 Suppose the templates of a MCNN are given by the following:

$$[a^{(1)}, a^{(1)}_r, z^{(1)}] = [2.9, 1.7, 0.1]$$
$$[a^{(2)}, a^{(2)}_r, b^{(2)}, b^{(2)}_r, z^{(2)}] = [-0.3, -1.2, 0.7, 2.3, 0.9]$$

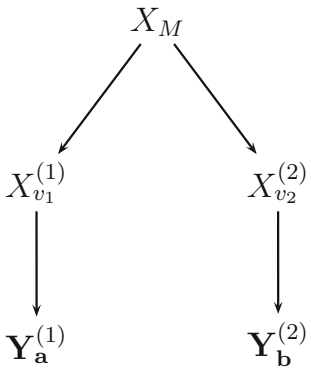


Fig. 8 If $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{b}}^{(2)}$ are finitely equivalent, then we can construct a new graph M such that M is the underlying graph of presentation of $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{b}}^{(2)}$. Herein $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{b}}^{(2)}$ are factors of $X_{v_1}^{(1)}$ and $X_{v_2}^{(2)}$, respectively, and $X^{(1)}$, $X^{(2)}$ are shifts of finite type

Then the basic set of admissible local patterns is

$$\mathcal{B} = \left\{ \begin{array}{|cc|}\hline - & + \\ - & - \\ \hline \end{array}, \begin{array}{|cc|}\hline - & + \\ + & - \\ \hline \end{array}, \begin{array}{|cc|}\hline + & - \\ - & + \\ \hline \end{array}, \begin{array}{|cc|}\hline + & - \\ + & + \\ \hline \end{array}, \begin{array}{|cc|}\hline + & + \\ - & + \\ \hline \end{array}, \begin{array}{|cc|}\hline + & + \\ + & + \\ \hline \end{array} \right\}.$$

The transition matrix $\mathbf{T}$ of the solution space $\mathbf{Y}$ is

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

and the symbolic transition matrices of the hidden and output spaces are

$$\mathbf{S^{(1)}} = \begin{pmatrix} \varnothing & \varnothing & s_1 & \varnothing \\ \varnothing & \varnothing & s_3 & \varnothing \\ \varnothing & s_2 & \varnothing & s_2 \\ \varnothing & s_4 & \varnothing & s_4 \end{pmatrix} \quad \textit{and} \quad \mathbf{S^{(2)}} = \begin{pmatrix} \varnothing & \varnothing & s_2 & \varnothing \\ \varnothing & \varnothing & s_2 & \varnothing \\ \varnothing & s_3 & \varnothing & s_4 \\ \varnothing & s_3 & \varnothing & s_4 \end{pmatrix}$$

respectively. See Fig. 9 for the graph presentation of $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$, respectively.

The right-resolving graph presentation $\mathcal{H}^{(1)}$ of $\mathbf{Y}^{(1)}$ is referred to as Fig. 10, and the symbolic transition matrix is

$$\overline{\mathbf{S}}^{(1)} = \begin{pmatrix} \varnothing & \varnothing & s_1 & \varnothing & \varnothing \\ \varnothing & \varnothing & s_3 & \varnothing & \varnothing \\ \varnothing & \varnothing & \varnothing & \varnothing & s_2 \\ \varnothing & \varnothing & \varnothing & \varnothing & s_4 \\ \varnothing & \varnothing & s_3 & \varnothing & s_4 \end{pmatrix}.$$

It is easy to see that $h_P(\mathbf{Y}) = h_P(\mathbf{Y}^{(1)}) = h_P(\mathbf{Y}^{(2)}) = \log g$ for all initial conditions, where $g = (1+\sqrt{5})/2$ is the golden mean. Moreover, the MCNN is independent of initial condition.

Let

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

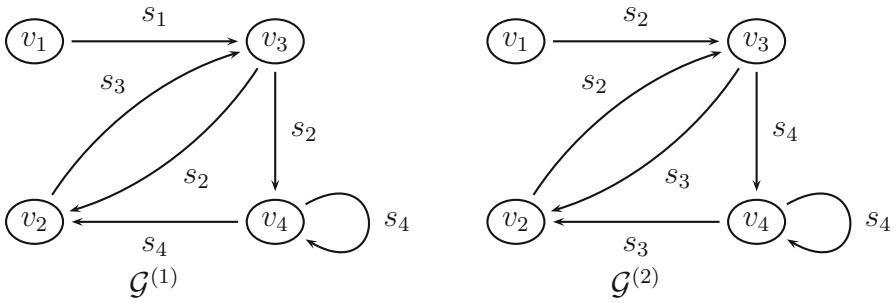


Fig. 9 The graph presentation of the hidden and output spaces of Example 3.10. It can be seen that $\mathcal{G}^{(2)}$ is right-resolving while $\mathcal{G}^{(1)}$ is not

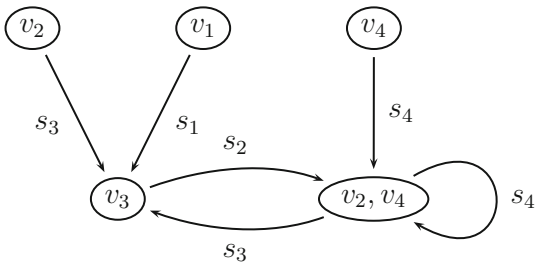


Fig. 10 The labeled graph obtained by applying SCM to $\mathcal{G}^{(1)}$ in Example 3.10

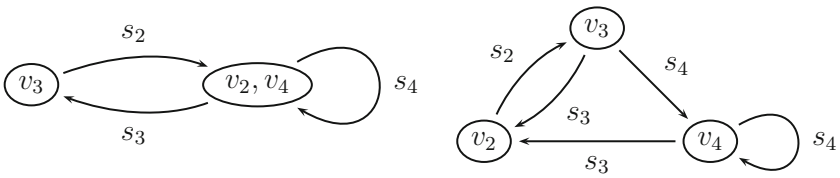


Fig. 11 The graph presentation of the hidden and output spaces in Example 3.10. The *left one* presents $\mathbf{Y}^{(1)}$ and the *right one* presents $\mathbf{Y}^{(2)}$

Then $\mathbf{H^{(1)}}E = E\mathbf{T^{(2)}}$ and $\mathbf{T^{(2)}}F = F\mathbf{H^{(1)}}$, where $\mathbf{H^{(1)}}$ is the transition matrix of $\mathcal{H}^{(1)}$.

Suppose the initial condition is given so that the presentations of $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{a}}^{(2)}$ are $(\mathcal{H}^{(1)}, v_3)$ and $(\mathcal{G}^{(2)}, v_3)$, respectively. In this case, the presentations of the hidden and output spaces are irreducible and are shown as in Fig. 11.

With an abuse of notation we denote the transition matrix and symbolic transition matrix of $\mathbf{Y}_{\mathbf{a}}^{(i)}$ by $\mathbf{T^{(i)}}$ and $\mathbf{S}^{(i)}$ respectively for $i = 1, 2$. Then

$$\mathbf{S}^{(1)} = \begin{pmatrix} \varnothing & s_2 \\ s_3 & s_4 \end{pmatrix} \quad \text{and} \quad \mathbf{S}^{(2)} = \begin{pmatrix} \varnothing & s_2 & \varnothing \\ s_3 & \varnothing & s_4 \\ s_3 & \varnothing & s_4 \end{pmatrix}$$

respectively. Define

$$\mathbf{F} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

It comes immediately that $\mathbf{S}^{(2)}\mathbf{F} = \mathbf{F}\mathbf{S}^{(1)}$ and therefore $\mathbf{T}^{(2)}\mathbf{F} = \mathbf{F}\mathbf{T}^{(1)}$. Theorem 3.9 indicates that $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{a}}^{(2)}$ are FE.

We use $\mathbf{F}$ to construct the common graph presentation of $\mathbf{Y}_{\mathbf{a}}^{(1)}$ and $\mathbf{Y}_{\mathbf{a}}^{(2)}$. The vertex set of the common graph M consists of $\mathbf{F}$-edges as follows.

$$(\alpha_1, \beta_1), \quad , (\alpha_2, \beta_2), \quad , (\alpha_3, \beta_2),$$

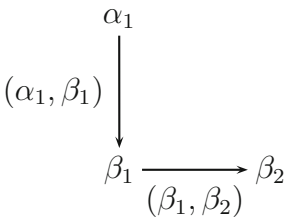
where α_i and β_j are vertices of, with the abuse of notations, $\mathcal{G}^{(2)}$ and $\mathcal{H}^{(1)}$ respectively for $1 \le i \le 3, 1 \le j \le 2$. Next we need to define the edges of M. Fix states $\alpha_i \in \mathcal{G}^{(2)}$ and $\beta_{j'} \in \mathcal{H}^{(1)}$. For each edge $\alpha_i \to \alpha_{i'}$ in $\mathcal{G}^{(2)}$, there is an edge from (α_i, β_j) to $(\alpha_{i'}, \beta_{j'})$ if and only if $(\alpha_{i'}, \beta_j)$ is an $\mathbf{F}$-edge and $\beta_j \to \beta_{j'}$ is an edge of $\mathcal{H}^{(1)}$. For example, consider $i = 1$ and $j' = 1$. Then $i' = 2$ and $j = 2$ is the only choice. Since there is no such $\mathbf{F}$-edge (α_1, β_2), there is no edge generated. Instead, consider $i = 1$ and $j' = 2$. Then $i' = 2$ and $j = 1, 2$. Since (α_1, β_2) is not an $\mathbf{F}$-edge, we conclude that

$$(\alpha_1, \beta_1) \quad \to \quad (\alpha_2, \beta_2)$$

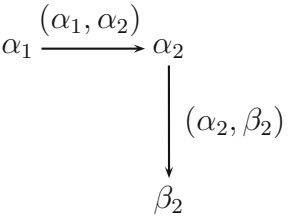
is an edge in M. Repeating this procedure derives the edges of M are as follows.

$$\begin{array}{lll} (\alpha_1, \beta_1) & \to & (\alpha_2, \beta_2) \\ (\alpha_2, \beta_2) & \to & (\alpha_1, \beta_1) \quad (\alpha_3, \beta_2) \\ (\alpha_3, \beta_2) & \to & (\alpha_1, \beta_1) \quad (\alpha_3, \beta_2) \end{array}$$

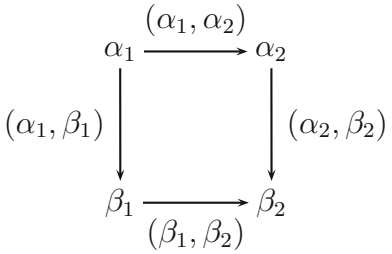
Notably the above procedure is actually the same as making a choice Γ of bijections as is presented in the proof of Theorem 3.9, and that the graph determined by such transitions is isomorphic to the graph M constructed there. Consider α_1 and β_2 for example. The only left-bottom path from α_1 to β_2 is



On the other hand, the only top-right path from α_1 to β_2 is



This gives the part $\Gamma_{\alpha_1,\beta_2}$ of the bijection Γ, and together these define the graph. The edge $(\alpha_1, \beta_1) \to (\alpha_2, \beta_2)$ in M corresponds to the following box.



See Fig. 6 for the underlying graph M.

Acknowledgments The authors wish to express their gratitude to the anonymous referees for their careful reading and useful suggestions, which make significant improvements to this work. Ban is partially supported by the National Science Council, ROC (Contract No. NSC 102-2628-M-259-001-MY3). Chang is grateful for the partial support of the National Science Council, ROC (Contract No. NSC 104-2115-M-390-004-).

Appendix: Equivalent Relation for Templates

The essential study of (1) is investigating two-layer cellular neural networks with the nearest neighborhood; namely, $n = 2$ and $\mathcal{N} = \{-1, 0, 1\}$. For the clarification of the investigation, we assume that $a^{(1)}_{-1} = a^{(2)}_{-1} = 0$. Under such condition, (1) is represented as

$$\begin{cases} \dfrac{d}{dt}x^{(2)}_i(t) = -x^{(2)}_i(t) + z^{(2)} + a^{(2)}y^{(2)}_i(t) + a^{(2)}_r y^{(2)}_{i+1}(t) + b^{(2)}y^{(1)}_i(t) + b^{(2)}_r y^{(1)}_{i+1}(t), \\ \dfrac{d}{dt}x^{(1)}_i(t) = -x^{(1)}_i(t) + z^{(1)} + a^{(1)}y^{(1)}_i(t) + a^{(1)}_r y^{(1)}_{i+1}(t), \end{cases} \tag{23}$$

where $i \in \mathbb{N}$, $t \geq 0$, and $y^{(j)}_i = f(x^{(j)}_i)$ for $j = 1, 2$. Since, for all i, j, $|x^{(j)}_i(t)| > 1$ provided t is large enough, the output $y^{(j)}_i(t)$ is either 1 or -1 after finite time. Hence we omit the time factor in the following discussion.

Suppose $\mathbf{y} = \begin{pmatrix} y^{(2)}_1 y^{(2)}_2 y^{(2)}_3 \cdots \\ y^{(1)}_1 y^{(1)}_2 y^{(1)}_3 \cdots \end{pmatrix} \in \mathbf{Y}$ is a mosaic pattern. For $i \in \mathbb{N}$, $y^{(1)}_i = 1$ if and only if $x^{(1)}_i > 1$. This derives

$$a^{(1)} + z^{(1)} - 1 > -a^{(1)}_r y^{(1)}_{i+1}. \tag{24}$$

Similarly, $y^{(1)}_i = -1$ if and only if $x^{(1)}_i < -1$. This implies $y^{(1)}_i = -1$ if and only if

$$a^{(1)} - z^{(1)} - 1 > a^{(1)}_r y^{(1)}_{i+1}. \tag{25}$$

The same argument asserts

$$a^{(2)} + z^{(2)} - 1 > -a^{(2)}_r y^{(2)}_{i+1} - (b^{(2)}y^{(1)}_i + b^{(2)}_r y^{(1)}_{i+1}) \tag{26}$$

and

$$a^{(2)} - z^{(2)} - 1 > a^{(2)}_r y^{(2)}_{i+1} + (b^{(2)}y^{(1)}_i + b^{(2)}_r y^{(1)}_{i+1}) \tag{27}$$

are the necessary and sufficient condition for $y^{(2)}_i = -1$ and $y^{(2)}_i = 1$, respectively. Define $\xi_1 : \{-1, 1\} \to \mathbb{R}$ and $\xi_2 : \{-1, 1\}^{\mathbb{Z}_{3\times 1}} \to \mathbb{R}$ by

$$\xi_1(w) = a^{(1)}_r w, \qquad \xi_2(w_1, w_2, w_3) = a^{(2)}_r w_1 + b^{(2)} w_2 + b^{(2)}_r w_3.$$

Set

$$\mathcal{B}^{(1)} = \left\{ \boxed{y^{(1)}y_r^{(1)}} : y^{(1)}, y_r^{(1)} \in \{-1, 1\} \text{ satisfy } (24), (25) \right\},$$

$$\mathcal{B}^{(2)} = \left\{ \boxed{\begin{matrix} y^{(2)}y_r^{(2)} \\ y^{(1)}y_r^{(1)} \end{matrix}} : y^{(2)}, y_r^{(2)}, y^{(1)}, y_r^{(1)} \in \{-1, 1\} \text{ satisfy } (26), (27) \right\}.$$

That is,

$$\boxed{y^{(1)}y_r^{(1)}} \in \mathcal{B}^{(1)} \Leftrightarrow \begin{cases} a^{(1)} + z^{(1)} - 1 > -\xi_1(y_r^{(1)}), & \text{if } y^{(1)} = 1; \\ a^{(1)} - z^{(1)} - 1 > \xi_1(y_r^{(1)}), & \text{if } y^{(1)} = -1. \end{cases}$$

$$\boxed{\begin{matrix} y^{(2)}y_r^{(2)} \\ y^{(1)}y_r^{(1)} \end{matrix}} \in \mathcal{B}^{(2)} \Leftrightarrow \begin{cases} a^{(2)} + z^{(2)} - 1 > -\xi_2(y_r^{(2)}, y^{(1)}, y_r^{(1)}), & \text{if } y^{(2)} = 1; \\ a^{(2)} - z^{(2)} - 1 > \xi_2(y_r^{(2)}, y^{(1)}, y_r^{(1)}), & \text{if } y^{(2)} = -1. \end{cases}$$

Since two-layer cellular neural networks are locally coupled systems, $\mathcal{B}^{(1)}$ and $\mathcal{B}^{(2)}$ represents the basic sets of admissible local patterns of the first and second layer of (23), respectively. The set of admissible local patterns $\mathcal{B}$ of (23) is then

$$\mathcal{B} \equiv (\mathcal{B}^{(1)}, \mathcal{B}^{(2)}) := \left\{ \boxed{\begin{matrix} yy_r \\ uu_r \end{matrix}} : \boxed{\begin{matrix} yy_r \\ uu_r \end{matrix}} \in \mathcal{B}^{(2)} \text{ and } \boxed{uu_r} \in \mathcal{B}^{(1)} \right\}$$

This is to say, the investigation of the equivalent relation for the templates is identical to the discussion of the equivalent relation for the basic sets of admissible local patterns. The fact that $y^{(1)}, y_r^{(1)} \in \{-1, 1\}$ indicates $a^{(1)} + z^{(1)} - 1 = -\xi_1(y_r^{(1)})$ and $a^{(1)} + z^{(1)} - 1 = \xi_1(y_r^{(1)})$ partition $a^{(1)}$-$z^{(1)}$ plane into 9 regions. More precisely, $a^{(1)}$-$z^{(1)}$ plane is partitioned by

$$a^{(1)} + z^{(1)} - 1 = a_r^{(1)}, \quad a^{(1)} + z^{(1)} - 1 > -a_r^{(1)},$$

and

$$a^{(1)} - z^{(1)} - 1 = a_r^{(1)}, \quad a^{(1)} - z^{(1)} - 1 = -a_r^{(1)}.$$

Encode these nine regions as $[p, q]$ for $0 \le p, q \le 2$, then a pair of parameters $(a^{(1)}, z^{(1)}) \in [p, q]$ infers that $(a^{(1)}, z^{(1)})$ satisfies m inequalities in (24) and n inequalities in (25). The relative positions of

$$a^{(1)} + z^{(1)} - 1 = a_r^{(1)} \quad \text{and} \quad a^{(1)} + z^{(1)} - 1 > -a_r^{(1)}$$

and the relative positions of

$$a^{(1)} - z^{(1)} - 1 = a_r^{(1)} \quad \text{and} \quad a^{(1)} - z^{(1)} - 1 = -a_r^{(1)}$$

remain to be determined. The "order" of lines $a^{(1)} + z^{(1)} - 1 = (-1)^\ell \xi_1(y_r^{(1)})$, $\ell = 0, 1$, in the plane come from the sign of $a_r^{(1)}$, this demonstrates that the parameter space $\{(a^{(1)}, a_r^{(1)}, z^{(1)})\}$ is partitioned into $2 \times 9 = 18$ equivalent regions. (Notably, the order of lines $a^{(1)} - z^{(1)} - 1 = \xi_1(y_r^{(1)})$ and $a^{(1)} - z^{(1)} - 1 = -\xi_1(y_r^{(1)})$ is determined seamlessly once the order of $a^{(1)} + z^{(1)} - 1 = (-1)^\ell \xi_1(y_r^{(1)})$ is given.) Namely, any two sets of parameters located in the same region determine the identical basic set of admissible local patterns $\mathcal{B}^{(1)}$. See Fig. 12a.

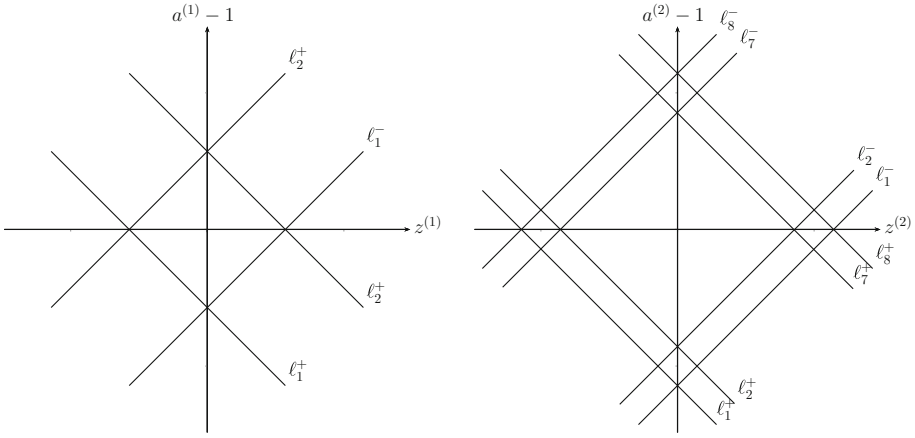


Fig. 12 The partition of $a^{(1)}$-$z^{(1)}$ and $a^{(2)}$-$z^{(2)}$ planes. In (a), ℓ_i^+ and ℓ_i^-, $i = 1, 2$, that represent the *lines* in (24) and (25) have partitioned the $a^{(1)}$-$z^{(1)}$ plane into nine regions. In (b), ℓ_i^+ and ℓ_i^-, represent the *lines* in (26) and (27), ave partitioned the $a^{(2)}$-$z^{(2)}$ plane into nine regions, where $1 \le i \le 8$

In an analogous manner, $y^{(2)}, y_r^{(2)}, u^{(2)}, u_r^{(2)} \in \{-1, 1\}$ indicates that $a^{(2)} + z^{(2)} - 1 > -\xi_2(y_r^{(2)}, u^{(2)}, u_r^{(2)})$ and $a^{(2)} + z^{(2)} - 1 > \xi_2(y_r^{(2)}, u^{(2)}, u_r^{(2)})$ partition $a^{(2)}$-$z^{(2)}$ plane into 81 regions. Encode these regions as $[p, q]$ for $0 \le p, q \le 8$, then a pair of parameters $(a^{(2)}, z^{(2)}) \in [p, q]$ infers that $(a^{(2)}, z^{(2)})$ satisfies m inequalities in (26) and n inequalities in (27). Furthermore, the order of $a^{(2)} + z^{(2)} - 1 = \xi_2(y_r^{(2)}, u^{(2)}, u_r^{(2)})$ can be uniquely determined according to the following procedures.

(1) The signs of $a_r^{(2)}, b^{(2)}, b_r^{(2)}$.
(2) The magnitude of $a_r^{(2)}, b^{(2)}, b_r^{(2)}$.
(3) The competition between the parameter with the largest magnitude and the others. In other words, suppose $m_1 > m_2 > m_3$ represent $|a_r^{(2)}|, |b^{(2)}|, |b_r^{(2)}|$. We need to determine whether $m_1 > m_2 + m_3$ or $m_1 < m_2 + m_3$.

This partitions the parameter space $\{(a^{(2)}, a_r^{(2)}, b^{(2)}, b_r^{(2)}, z^{(2)})\}$ into $8 \times 6 \times 2 \times 81 = 7776$ regions and each region is associated with a basic set of admissible local patterns (cf. Fig. 12b).

The above discussion demonstrates the following proposition.

Proposition A.1 *Let* $\mathcal{P}_8 = \{(a^{(1)}, a_r^{(1)}, a^{(2)}, a_r^{(2)}, b^{(2)}, b_r^{(2)}, z^{(1)}, z^{(2)})\}$ *be the parameter space of* (23)*. There is a positive integer* K *and unique set of open subregions* $\{P_k\}_{k=1}^K$ *satisfying*

(1) $\mathcal{P}_8 = \bigcup_{k=1}^K \overline{P}_k$*, where* $\overline{U}$ *refers to the closure of* U*;*
(2) $P_i \bigcap P_j = \varnothing$ *if* $i \neq j$*;*
(3) *Templates* $\mathbb{T}, \mathbb{T}' \in P_k$ *for some* k *if and only if* $\mathbf{Y}_{\mathbb{T}} = \mathbf{Y}_{\mathbb{T}'}$.

Example A.2 Suppose $a_r^{(1)} < 0$ and $a_r^{(2)}, b^{(2)}, b_r^{(2)}$ are all positive. Moreover, choose $a_r^{(2)}, b^{(2)}, b_r^{(2)}$ such that

$$a_r^{(2)} > b^{(2)} > b_r^{(2)} \quad \text{and} \quad a_r^{(2)} > b^{(2)} + b_r^{(2)}.$$

For instance, $a_r^{(1)} = -1$, $a_r^{(2)} = 6$, $b^{(2)} = 3$, and $b_r^{(2)} = 2$. Then the position of each line is settled. We number the partitions of $a^{(1)}$-$z^{(1)}$ and $a^{(2)}$-$z^{(2)}$ planes by a pair $[m_\ell, n_\ell]$, $\ell = 1, 2$, where m_ℓ, n_ℓ illustrate how many inequalities

$$a^{(\ell)} + z^{(\ell)} - 1 > -\xi_\ell(\cdot) \quad \text{and} \quad a^{(\ell)} - z^{(\ell)} - 1 > \xi_\ell(\cdot)$$

are satisfied, respectively. Thus $0 \le m_1, n_1 \le 2$ and $0 \le m_2, n_2 \le 8$. Pick $[m_1, n_1] = [1, 2]$ and $[m_2, n_2] = [6, 4]$, for instance, $a^{(1)} = 2$, $z^{(1)} = -0.3$, $a^{(2)} = 4$, and $z^{(2)} = 2.5$. It is easy to check that the basic set of admissible local patterns is

$$\mathcal{B} = \left\{ \boxed{\begin{matrix}-&-\\-&-\end{matrix}}, \quad \boxed{\begin{matrix}-&-\\-&+\end{matrix}}, \quad \boxed{\begin{matrix}-&-\\+&-\end{matrix}}, \quad \boxed{\begin{matrix}+&-\\+&-\end{matrix}}, \quad \boxed{\begin{matrix}+&+\\-&-\end{matrix}}, \quad \boxed{\begin{matrix}+&+\\-&+\end{matrix}}, \quad \boxed{\begin{matrix}+&+\\+&-\end{matrix}} \right\}$$

References

1. Abram, W., Lagarias, J.C.: p-Adic path set fractals and arithemetics. J. Fractal Geom. **1**, 45–81 (2014)
2. Abram, W., Lagarias, J.C.: Path sets in one-sided symbolic dynamics. Adv. Appl. Math. **56**, 109–134 (2014)
3. Adler, R.L., Marcus, B.: Topological entropy and equivalence of dynamical systems. Mem. Am. Math. Soc., vol. 219. AMS: Providence, RI (1979)
4. Arena, P., Baglio, S., Fortuna, L., Manganaro, G.: Self-organization in a two-layer CNN. IEEE Trans. Circuits Syst. I Fundam. Theory Appl. **45**, 157–162 (1998)
5. Arik, S., Tavsanoglu, V.: Equilibrium analysis of non-symmetric cnns. Int. J. Circ. Theor. Appl. **24**, 269–274 (1996)
6. Ban, J.-C., Chang, C.-H.: On the monotonicity of entropy for multi-layer cellular neural networks. Int. J. Bifur. Chaos Appl. Sci. Eng. **19**, 3657–3670 (2009)
7. Ban, J.-C., Chang, C.-H.: On the structure of multi-layer cellular neural networks. Part II: The complexity between two layers (2012, under review)
8. Ban, J.-C., Chang, C.-H.: Diamond in multi-layer cellular neural networks. Appl. Math. Comput. **222**, 1–12 (2013)
9. Ban, J.-C., Chang, C.-H.: The learning problem of multi-layer neural networks. Neural Netw. **46**, 116–123 (2013)
10. Ban, J.-C., Chang, C.-H., Lin, S.-S.: The structure of multi-layer cellular neural networks. J. Differ. Equ. **252**, 4563–4597 (2012)
11. Ban, J.-C., Chang, C.-H., Lin, S.-S., Lin, Y.-H.: Spatial complexity in multi-layer cellular neural networks. J. Differ. Equ. **246**, 552–580 (2009)
12. Carmona, R., Jimenez-Garrido, F., Dominguez-Castro, R., Espejo, S., Rodriguez-Vazquez, A.: CMOS realization of a 2-layer cnn universal machine chip. In: Proceedings of the 2002 7th IEEE International Workshop on Cellular Neural Networks and Their Applications, 2002, (CNNA 2002), pp. 444–451 (2002)
13. Chua, L.O.: Cnn: A Paradigm for Complexity, World Scientific Series on Nonlinear Science, Series A, vol. 31. World Scietific, Singapore (1998)
14. Chua, L.O., Yang, L.: Cellular neural networks: applications. IEEE Trans. Circuits Syst. **35**, 1273–1290 (1988)
15. Chua, L.O., Yang, L.: Cellular neural networks: theory. IEEE Trans. Circuits Syst. **35**, 1257–1272 (1988)
16. Chua, L.O., Shi, B.E.: Multiple layer cellular neural networks: a tutorial. In: Deprettere, E.F., van der Veen, A.-J. (eds.) Algorithms and Parallel VLSI Architectures, pp. 137–168. Elsevier, Amsterdam (1991)
17. Chua, L.O., Roska, T.: Cellular Neural Networks and Visual Computing. Cambridge University Press, Cambridge, MA (2002)
18. Crounse, K.R., Chua, L.O.: Methods for image processing and pattern formation in cellular neural networks: a tutorial. IEEE Trans. Circuits Syst. **42**, 583–601 (1995)
19. Crounse, K.R., Roska, T., Chua, L.O.: Image halftoning with cellular neural networks. IEEE Trans. Circuits Syst. **40**, 267–283 (1993)
20. Forti, M., Tesi, A.: New conditions for global stability of neural networks with application to linear and quadratic programming problems. IEEE Trans. Circuits Syst. I. Funda. Theory Appl. **42**, 354–366 (1995)
21. Forti, M., Tesi, A.: A new method to analyze complete stability of PWL cellular neural networks. Int. J. Bifur. Chaos Appl. Sci. Eng. **11**, 655–676 (2001)

22. Fukushima, K.: Artificial vision by multi-layered neural networks: neocognitron and its advances. Neural Netw. **37**, 103–119 (2013)
23. Fukushima, K.: Training multi-layered neural network neocognitron. Neural Netw. **40**, 18–31 (2013)
24. Juang, J., Lin, S.-S.: Cellular neural networks: mosaic pattern and spatial chaos. SIAM J. Appl. Math. **60**, 891–915 (2000)
25. Li, X.: Analysis of complete stability for discrete-time cellular neural networks with piecewise linear output functions. Neural Comput. **21**, 1434–1458 (2009)
26. Liang, X.-B., Si, J.: Global exponential stability of neural networks with globally lipschitz continuous activations and its application to linear variational inequality problem. IEEE Trans. Neural Netw. **12**, 349–359 (2001)
27. Lind, D., Marcus, B.: An Introduction to Symbolic Dynamics and Coding. Cambridge University Press, Cambridge, MA (1995)
28. Murugesh, V.: Image processing applications via time-multiplexing cellular neural network simulator with numerical integration algorithms. Int. J. Comput. Math. **87**, 840–848 (2010)
29. Parry, W.: A finitary classification of topological markov chains and sofic systems. Bull. Lond. Math. Soc. **9**, 86–92 (1977)
30. Peng, J., Zhang, D., Liao, X.: A digital image encryption algorithm based on hyper-chaotic cellular neural network. Fund. Inform. **90**, 269–282 (2009)
31. Savaci, F.A., Vandewalle, Joos, On the stability analysis of cellular neural networks. In: Proceedings, Second International Workshop on Cellular Neural Networks and their Applications, 1992, CNNA-92. IEEE, 1992, pp. 240–245 (1992)
32. Takahashi, N., Chua, L.O.: On the complete stability of nonsymmetric cellular neural networks. IEEE Trans. Circuits-I. **45**, 754–758 (1998)
33. Török, L., Roska, T.: Stability of multi-layer cellular neural/nonlinear networks. Int. J. Bifur. Chaos Appl. Sci. Eng. **14**, 3567–3586 (2004)
34. Xavier-de Souza, S., Yalcin, M.E., Suykens, J.A.K., Vandewalle, J.: Toward CNN chip-specific robustness. IEEE Trans. Circuits Syst. I Reg. Papers **51**, 892–902 (2004)
35. Yang, Z., Nishio, Y., Ushida, A.: A two layer cnn in image processing applications. In: Proceedings of the 2001 International Symposium on Nonlinear Theory and Its Applicationsm, pp. 67–70 (2001)
36. Yang, Z., Nishio, Y., Ushida, A.: Image processing of two-layer CNNs — applications and their stability. IEICE Trans. Fundam. **E85–A**, 2052–2060 (2002)