On the structure of multi-layer cellular neural networks

Jung-Chao Ban [a,*,1], Chih-Hung Chang [b,2], Song-Sun Lin [c,3]

[a] *Department of Applied Mathematics, National Dong Hwa University, Hualien 97401, Taiwan, ROC*
[b] *Department of Applied Mathematics, Feng Chia University, Taichung 40724, Taiwan, ROC*
[c] *Department of Applied Mathematics, National Chiao-Tung University, Hsinchu 30050, Taiwan, ROC*

ARTICLE INFO

ABSTRACT

Let $\mathbf{Y} \subseteq \{-1, 1\}^{\mathbb{Z}_{\infty \times n}}$ be the mosaic solution space of an n-layer cellular neural network. We decouple $\mathbf{Y}$ into n subspaces, say $Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)}$, and give a necessary and sufficient condition for the existence of factor maps between them. In such a case, $Y^{(i)}$ is a sofic shift for $1 \leqslant i \leqslant n$. This investigation is equivalent to study the existence of factor maps between two sofic shifts. Moreover, we investigate whether $Y^{(i)}$ and $Y^{(j)}$ are topological conjugate, strongly shift equivalent, shift equivalent, or finitely equivalent via the well-developed theory in symbolic dynamical systems. This clarifies, in a multi-layer cellular neural network, each layer's structure. As an extension, we can decouple $\mathbf{Y}$ into arbitrary k-subspaces, where $2 \leqslant k \leqslant n$, and demonstrates each subspace's structure.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Multi-layer cellular neural networks (MCNNs) are large aggregates of analogue circuits presenting themselves as arrays of identical cells which are locally coupled. MCNNs have been widely applied in

studying the signal propagation between neurons, and in image processing, pattern recognition and information technology [1–11]. A One-dimensional MCNN is realized in the following form.

$$\frac{dx_i^{(\ell)}}{dt} = -x_i^{(\ell)} + \sum_{|k|\leqslant d} a_k^{(\ell)} y_{i+k}^{(\ell)} + \sum_{|k|\leqslant d} b_k^{(\ell)} u_{i+k}^{(\ell)} + z^{(\ell)}, \tag{1}$$

for some $d \in \mathbb{N}$, $1 \leqslant \ell \leqslant n \in \mathbb{N}$, $i \in \mathbb{Z}$, where

$$u_i^{(\ell)} = y_i^{(\ell-1)} \quad \text{for } 2 \leqslant \ell \leqslant n, \qquad u_i^{(1)} = u_i, \qquad x_i(0) = x_i^0, \tag{2}$$

and

$$y = f(x) = \frac{1}{2}\big(|x+1| - |x-1|\big) \tag{3}$$

is the output function. For $1 \leqslant \ell \leqslant n$, $A^{(\ell)} = (a_{-d}^{(\ell)}, \ldots, a_d^{(\ell)})$ is called the feedback template, $B^{(\ell)} = (b_{-d}^{(\ell)}, \ldots, b_d^{(\ell)})$ is called the controlling template, and $z^{(\ell)}$ is the threshold. The quantity $x_i^{(\ell)}$ denotes the state of a cell C_i in the ℓth layer. The stationary solutions $\bar{x} = (\bar{x}_i^{(\ell)})$ of (1) are essential for understanding the system, and their outputs $\bar{y}_i^{(\ell)} = f(\bar{x}_i^{(\ell)})$ are called *output patterns*. Among the stationary solutions, the mosaic solutions are crucial for studying the complexity of (1)[12–18]. A *mosaic solution* $(\bar{x}_i^{(\ell)})$ satisfies $|\bar{x}_i^{(\ell)}| > 1$ for all i, ℓ and the output of a mosaic solution is called a mosaic output pattern. Mosaic solutions are asymptotically stable for one-layer CNNs without input [15,19]. Despite a lack of rigorous proof, some criteria and numerical experiments have been proposed to assert the stability of mosaic solutions for MCNNs [20–22]. Aside from mosaic solutions, MCNNs also exhibit periodic patterns and limit cycles, readers are referred to [23–26] and references therein. In a MCNN system, the "status" of each cell is taken as an input for a cell in the next layer except for those cells in the last layer. The results that can be recorded are the output of the cells in the last layer. Since the phenomena that can be observed are only the output patterns of the nth layer, the nth layer of (1) is called the *output layer*, while the other $n-1$ layers are called *hidden layers*.

Since mosaic solutions are stable, it is essential to characterize the structure of mosaic solutions of (1). Juang and Lin [14] and Ban et al. [27] investigated mosaic solutions systematically, and characterized the complexity of mosaic patterns via topological entropy. Shih [19] elucidated how the boundary condition influence the spatial complexity of the global patterns. In the present study, a pattern stands for a stationary solution for (1). Since the feedback and controlling templates are spatially invariant, the global pattern formation is thus completely determined by the so-called admissible local patterns. Hence, investigation of admissible local patterns is essential for studying the complexity of global patterns. The difficulty stems from the fact that the set of admissible local patterns is constrained by the differential equation (1). Suppose

$$\mathcal{B} \equiv \mathcal{B}\big(A^{(1)}, \ldots, A^{(n)}, B^{(1)}, \ldots, B^{(n)}, z^{(1)}, \ldots, z^{(n)}\big)$$

$$= \left\{ \begin{array}{c} \boxed{\begin{array}{c} y_{-d}^{(n)} \cdots y_{-1}^{(n)} y_0^{(n)} y_1^{(n)} \cdots y_d^{(n)} \\ \vdots \\ y_{-d}^{(2)} \cdots y_{-1}^{(2)} y_0^{(2)} y_1^{(2)} \cdots y_d^{(2)} \\ y_{-d}^{(1)} \cdots y_{-1}^{(1)} y_0^{(1)} y_1^{(1)} \cdots y_d^{(1)} \end{array}} \end{array} \right\} \subseteq \{-1,1\}^{\mathbb{Z}(2d+1)\times n}$$

is a basic set of admissible local patterns. The predicament is that there exists a subset of $\{-1,1\}^{\mathbb{Z}(2d+1)\times n}$ that cannot be realized via MCNNs. Such a constraint arises from the so-called *linear*

separation property. Hsu et al. [28] demonstrated that, for one-layer CNNs without input, the parameter space can be divided into a finite number of partitions such that any two sets of parameters in the same partition admit the same basic set of admissible local patterns. This property remains true for MCNNs [29,27]. Proposition 3.1 gives a brief introduction to the procedure for determining the partitions of the parameter space of simplified two-layer CNNs.

Suppose **Y** is the solution space of a MCNN. For $\ell = 1, 2, \ldots, n$, let

$$Y^{(\ell)} = \left\{ \cdots y_{-1}^{(\ell)} y_0^{(\ell)} y_1^{(\ell)} \cdots \right\}$$

be the space which consists of patterns in the ℓth layer of **Y**. Then $Y^{(n)}$ is called the *output space* and $Y^{(\ell)}$ is called the (ℓth) *hidden space* for $\ell = 1, 2, \ldots, n-1$. There is a canonical projection $\phi^{(\ell)} : \mathbf{Y} \to Y^{(\ell)}$ for each ℓ. It is natural to ask whether there exists a relation between $Y^{(i)}$ and $Y^{(j)}$ for $1 \leqslant i \neq j \leqslant n$. Take $n = 2$ for instance; the existence of map connecting $Y^{(1)}$ and $Y^{(2)}$ that commutes with $\phi^{(1)}$ and $\phi^{(2)}$ means the decoupling of the solution space **Y**. More precisely, if there exists $\pi_{12} : Y^{(1)} \to Y^{(2)}$ such that $\pi_{12} \circ \phi^{(1)} = \phi^{(2)}$. Then π_{12} enables the investigation of structures between the output space and hidden space.

Ban et al. [27] demonstrated that the output space $Y^{(n)}$ is a one-dimensional sofic shift. An analogous argument asserts that each hidden space is also a sofic shift. To study the existence of $\pi_{ij} : Y^{(i)} \to Y^{(j)}$ for some i, j is equivalent to illustrate the existence between two sofic shifts. This elucidation gives a systematic strategy for determining whether there exists a map between $Y^{(i)}$ and $Y^{(j)}$ via well-developed theory in symbolic dynamical systems. Readers are referred to [30,31] for more details.

To elucidate the relation between $Y^{(i)}$ and $Y^{(j)}$ in a MCNN, we start with a simplified two-layer CNN as (12). Suppose $\mathcal{B}$ is a basic set of admissible local patterns induced by a particular partition of the parameter space. Ban and Lin [32] propose a methodology for the study of the complexity of global patterns via assigning each local pattern its order and defining a so-called *ordering matrix*. After defining an ordering matrix $\mathbb{X}$, $\mathcal{B}$ derives a transition matrix T and has a graph representation G. Applying appropriate labeling $\mathcal{L}^{(1)}, \mathcal{L}^{(2)}$ on G there associates labeled graph $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ such that $\mathcal{G}^{(i)}$ is a labeled graph representation of $Y^{(i)}$ for $i = 1, 2$. If the topological entropies of $Y^{(1)}$ and $Y^{(2)}$ are different, then either $\mathcal{G}^{(1)}$ or $\mathcal{G}^{(2)}$ has a *graph diamond*. Proposition 3.34 illustrates a sufficient condition for the existence of a graph diamond.

Once $Y^{(1)}$ and $Y^{(2)}$ admit the same topological entropy, $Y^{(1)}$ and $Y^{(2)}$ are finitely equivalent and there is a *triangular structure* between them (cf. Fig. 1 for example). For the case that $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are both right-resolving, that is, any two different edges with the same initial state carrying different labels, $Y^{(1)}$ is topological conjugate to $Y^{(2)}$ if and only if $S^{(1)} P = P S^{(2)}$ for some permutation matrix P (Theorem 3.11), where $S^{(i)}$ is the symbolic transition matrix of $\mathcal{G}^{(i)}$ for $i = 1, 2$. Moreover, as one can see in Fig. 1, the solution space **Y** is at the top of the triangle. In this case **Y** is called a *common extension* of $Y^{(1)}$ and $Y^{(2)}$.

If $\mathcal{G}^{(i)}$ is not right-resolving for some i, there exists $\mathcal{G}_{Y^{(i)}} = (G_{Y^{(i)}}, \mathcal{L}_{Y^{(i)}})$ which is right-resolving and is a labeled graph representation of $Y^{(i)}$. In this case **Y** is no longer a common extension of $Y^{(1)}$ and $Y^{(2)}$. Since $h(Y^{(1)}) = h(Y^{(2)})$, there exists an integral matrix F such that $T_{G_{Y^{(1)}}} F = F T_{G_{Y^{(2)}}}$. Using F we can construct a common extension W of $Y^{(1)}$ and $Y^{(2)}$ (Theorem 3.12, cf. Figs. 3 and 4). Before studying the map between $Y^{(1)}$ and $Y^{(2)}$, we focus on the relation between $W^{(1)} \equiv X_{G_{Y^{(1)}}}$ and $W^{(2)} \equiv X_{G_{Y^{(2)}}}$, that is, the shift spaces generated from the underlying graph $G_{Y^{(1)}}$ and $G_{Y^{(2)}}$.

The transition matrices $T_{G_{Y^{(1)}}}$ and $T_{G_{Y^{(2)}}}$ are significant for studying the existence of factor map between $W^{(1)}$ and $W^{(2)}$. Williams' Classification Theorem [33,34] indicates that $W^{(1)}$ is conjugate to $W^{(2)}$ if and only if their transition matrices $T_{W^{(1)}}, T_{W^{(2)}}$ are *strongly shift equivalent*. Suppose $T_{W^{(1)}}, T_{W^{(2)}}$ are strongly shift equivalent and $\pi : W^{(1)} \to W^{(2)}$ is the conjugacy map. It still does not guarantee that there is a map connecting $Y^{(1)}$ and $Y^{(2)}$. To determine a map between $Y^{(1)}$ and $Y^{(2)}$ which still preserves the structure of a tree diagram, it is sufficient to find a map from $Y^{(i)}$ to $W^{(i)}$ for some i. Proposition 2.8 and Theorem 3.19 propose that this is equivalent to verifying the *synchronizing words* on $\mathcal{G}^{(i)}$. If words on $\mathcal{G}^{(1)}$ are synchronized, then there exists a factor map $\bar{\pi} : Y^{(1)} \to Y^{(2)}$. If words on $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are synchronized respectively, then $Y^{(1)}$ is conjugate to $Y^{(2)}$.

A weaker equivalent class is that $T_{W^{(1)}}$ and $T_{W^{(2)}}$ are *shift equivalent*. That is, there exists $N \in \mathbb{N}$ such that $T_{W^{(1)}}^N$ and $T_{W^{(2)}}^N$ are strongly shift equivalent. The difference between strong shift equivalence and shift equivalence is that shift equivalence of $W^{(1)}$ and $W^{(2)}$ does not imply the conjugacy of $W^{(1)}$ and $W^{(2)}$ [35,36]. Moreover, strong shift equivalence is much harder to verify than shift equivalence since $W^{(1)}$ and $W^{(2)}$ are shift equivalent if and only if they admit isomorphic dimension groups (Theorem 3.27). In general, if there exists a factor-like matrix F such that $T_{G_{Y^{(1)}}} F = F T_{G_{Y^{(2)}}}$, then there exists a factor map between $W^{(1)}$ and $W^{(2)}$.

The investigation of two-layer CNNs can be extended to decoupling a MCNN into k subspaces. This leads to how we decouple the solution space $\mathbf{Y}$. Suppose $\mathbf{Y}$ is the solution space of a three-layer CNN. If we decouple $\mathbf{Y}$ into two spaces, for example, $Y^{(2)}$ is the output space and $Y^{(1)}$ is the hidden space consisting of the global patterns of the first two layers. This reduces to the case that $n = 2$ but with a change of labels. Suppose $\mathbf{Y}$ induces a output space $Y^{(3)}$ and two hidden spaces $Y^{(1)}$ and $Y^{(2)}$. If $h(Y^{(i)}) \neq h(Y^{(j)})$ for some $i \neq j$, then $\mathcal{G}^{(\ell)}$ has a graph diamond for some ℓ, where $\mathcal{G}^{(\ell)}$ is a labeled graph representation of $Y^{(\ell)}$, $\ell = 1, 2, 3$. Suppose $Y^{(\ell)}$ admits the same topological entropy for $\ell = 1, 2, 3$. The above discussion infers that $Y^{(2)}$ and $Y^{(3)}$ form a triangle diagram, and $Y^{(1)}$ and $Y^{(2)}$ also form a triangle diagram. Assume $W_1^{(1)}$ is a common extension of $Y^{(1)}, Y^{(2)}$, and $W_2^{(1)}$ is a common extension of $Y^{(2)}, Y^{(3)}$. Using *fiber product* we get a common extension W of $Y^{(1)}$ and $Y^{(3)}$ (Theorem 4.3). More precisely, such extension W is a common extension of $Y^{(\ell)}$ for $\ell = 1, 2, 3$. See Figs. 8, 9, and 10. The triangular structure illustrates the relation between each of three layers and helps for examining whether there exist factor maps connecting some layers.

The investigation of the existence of a map between $Y^{(\ell)}, Y^{(\ell+1)}$ is similar to the discussion above for $\ell = 1, 2$. To determine whether there is a factor map connecting $Y^{(1)}$ and $Y^{(3)}$, the existence of factor maps, for instance, from $Y^{(1)}$ to $Y^{(2)}$ and from $Y^{(2)}$ to $Y^{(3)}$ obviously indicates a factor map from $Y^{(1)}$ to $Y^{(3)}$. If there is no factor map between $Y^{(\ell)}$ and $Y^{(\ell+1)}$ for some $1 \leqslant \ell \leqslant 2$, we have the following cases:

(i) The equivalent relation between $W^{(1)}, W^{(2)}$, and $W^{(3)}$, where $W^{(\ell)}$ is a shift of finite type such that $Y^{(\ell)}$ is a factor of $W^{(\ell)}$.
(ii) The equivalent relation between $W_1^{(1)}$ and $W_2^{(1)}$.
(iii) The existence of factor maps between $W_1^{(1)}, W_2^{(1)}$ and $W^{(\ell)}, W^{(\ell+1)}$ for $\ell = 1, 2$.

For any of the above cases, it leaves the verification of synchronizing words of $Y^{(\ell)}$ for $\ell = 1, 2, 3$.

In our previous work [27], we focused on the complexity of the output space of a MCNN and the recurrent formula of the transition matrix. Aside from showing that the output space is sofic, we also observed that the topological entropy diagram is asymmetric. A further study [37] indicates that the symmetry of the diagram comes from the conjugacy of two systems while the asymmetry is caused by the existence of diamonds. This elucidation, based on the results in [27], examines the finer structure in MCNNs and provides some sufficient conditions to decouple the mosaic solution spaces of MCNNs.

Since, for example, in a 2-layer CNN every mosaic solution space $Y^{(\ell)}$ is a sofic shift for $\ell = 1, 2$, this investigation considers whether there exists a factor map between two sofic shifts. As an application, the existence of a factor map can be used for the examination of the rate of loss during data transmission. Meanwhile, from the viewpoint of thermodynamics, suppose μ is an equilibrium measure on $Y^{(1)}$, a factor map $\pi : Y^{(1)} \to Y^{(2)}$ helps for the understanding of properties of the equilibrium measures on $Y^{(2)}$.

The above discussion can be extended to general multi-layer neural networks with only slight modification. The rest of this study is organized as follows. A brief introduction of symbolic dynamical systems and MCNNs is given in Section 2. Section 3 analyzes simplified two-layer cellular neural networks. Section 4 studies simplified n-layer cellular neural networks for $n \geqslant 3$ and Section 5 extends the results to general multi-layer cellular neural networks. Some discussion and conclusions are given in Section 6.

2. Preliminary

For reader's convenience, we recall some definitions and known results for symbolic dynamical systems and MCNNs that are needed in the present elucidation. We refer the reader to [27,14,31] and the references therein for more details.

2.1. Symbolic dynamical systems

Let $\mathcal{A} = \{0, 1, \ldots, n-1\}$ be a finite alphabet with cardinality $|\mathcal{A}| = n$. The full $\mathcal{A}$-shift $\mathcal{A}^{\mathbb{Z}}$ is the collection of all bi-infinite sequences of symbols from $\mathcal{A}$. More precisely,

$$\mathcal{A}^{\mathbb{Z}} = \left\{ \alpha = (\alpha_i)_{i \in \mathbb{Z}} : \alpha_i \in \mathcal{A} \text{ for all } i \in \mathbb{Z} \right\}.$$

The shift map σ on the full shift $\mathcal{A}^{\mathbb{Z}}$ is defined by

$$\sigma(\alpha)_i = \alpha_{i+1} \quad \text{for } i \in \mathbb{Z}.$$

A *shift space* X is a subset of $\mathcal{A}^{\mathbb{Z}}$ such that $\sigma(X) \subseteq X$.

Definition 2.1. For each $m \in \mathbb{N}$, let

$$\mathcal{A}_m = \{w_0 w_1 \cdots w_{k-1} : w_i \in \mathcal{A}, \ 1 \leqslant k \leqslant m\}$$

and let $\mathcal{A}_0$ denote the empty set. If X is a shift space and there exist $L \geqslant 0$ and $\mathcal{F} \subseteq \mathcal{A}_L$ such that

$$X = \left\{ (\alpha_i)_{i \in \mathbb{Z}} : \alpha_i \alpha_{i+1} \cdots \alpha_{i+k-1} \notin \mathcal{F} \text{ for } k \leqslant N, \ i \in \mathbb{Z} \right\}$$

then we say that X is a *shift of finite type* (SFT).

A SFT can be constructed via a finite, directed graph by considering the collection of all bi-infinite walks on the graph. We recall some definitions first.

Definition 2.2. A (directed) *graph* $G = (\mathcal{V}, \mathcal{E})$ consists of a finite set $\mathcal{V} = \mathcal{V}(G)$ of *vertices* (or *states*) together with a finite set $\mathcal{E} = \mathcal{E}(G)$ of *edges*. Each edge $e \in \mathcal{E}$ starts at an initial state $i(e)$ and terminates at a terminal state $t(e)$. Sometimes we also denote an edge e by $e = (i(e), t(e))$ to emphasize the initial and terminal states of e.

Let G and H be graphs. A *homomorphism* $(\partial\Phi, \Phi) : G \to H$ consists of a pair of maps $\partial\Phi : \mathcal{V}(G) \to \mathcal{V}(H)$ and $\Phi : \mathcal{E}(G) \to \mathcal{E}(H)$ such that $i(\Phi(e)) = \partial\Phi(i(e))$ and $t(\Phi(e)) = \partial\Phi(t(e))$ for all $e \in \mathcal{E}(G)$. A homomorphism is an *isomorphism* if both $\partial\Phi$ and Φ are one-to-one and onto.

Without loss of generality, we assume that, for any two vertices in a graph, there is at most one corresponding (directed) edge. Let $G = (\mathcal{V}, \mathcal{E})$ be a graph. The *transition matrix* T_G of G, indexed by $\mathcal{V}$, is an incidence matrix defined by $T_G(I, J) = 1$ if and only if $(I, J) \in \mathcal{E}$. On the other hand, suppose T is an $n \times n$ incidence matrix, then the graph of T is the graph $G = G_T$ with vertex set $\mathcal{V} = \{1, 2, \ldots, n\}$, and with $T(I, J)$ edge from vertex I to vertex J. It follows immediately from the definitions that

$$T = T_{G_T} \quad \text{and} \quad G \cong G_{T_G}.$$

Each graph G with corresponding transition matrix T gives rise to a SFT.

Definition 2.3. Let $G = (\mathcal{V}, \mathcal{E})$ be a graph with transition matrix T. The edge shift X_G or X_T is the shift space over the alphabet $\mathcal{A} = \mathcal{E}$ specified by

$$X_G = X_T = \left\{ \xi = (\xi_j)_{j \in \mathbb{Z}} \in \mathcal{E}^{\mathbb{Z}} : \xi_j, \xi_{j+1} \in \mathcal{E} \text{ such that } i(\xi_{j+1}) = t(\xi_j) \text{ for } j \in \mathbb{Z} \right\}. \tag{4}$$

Theorem 2.4. *(See [31, Proposition 2.2.6].) Suppose G is a graph. Then X_G is a shift of finite type.*

Sometimes certain edges of G can never appear in X_G, and such edges are inessential for the edge shift. A vertex $I \in \mathcal{V}$ is called *stranded* if either no edges start at I or no edges terminate at I. We say that a graph is *essential* if no vertex of the graph is stranded. The following proposition demonstrates that we can focus the discussion on those essential graphs.

Proposition 2.5. *(See [31, Proposition 2.2.10].) If G is a graph, then there is a unique subgraph H of G such that H is essential and $X_H = X_G$.*

A graph G is *irreducible* if for every ordered pair of vertices I and J there is a path in G starting at I and terminating at J, herein a path $\pi = v_1 v_2 \ldots v_m$ on a graph G we mean a finite sequence of vertices from G such that $(v_i, v_{i+1}) \in \mathcal{E}$ for $i = 1, 2, \ldots, m-1$. It can be verified that an essential graph is irreducible if and only if its edge shift is irreducible.

Suppose we label the edges in a graph with symbols from an alphabet $\mathcal{S}$. Every bi-infinite walk on the graph yields a point in the full shift $\mathcal{S}^{\mathbb{Z}}$ by reading the labels of its edges, and the set of all such points is called a *sofic shift*.

Definition 2.6. Suppose $G = (\mathcal{V}, \mathcal{E})$ is a directed graph, and $\mathcal{S}$ is a finite alphabet. A labeled graph $\mathcal{G}$ is a pair $(G, \mathcal{L})$ with graph G and the labeling $\mathcal{L} : \mathcal{E} \to \mathcal{S}$ assigns to each edge e of G a label $\mathcal{L}(e) \in \mathcal{S}$. The underlying graph of $\mathcal{G}$ is G.

Let $\mathcal{G} = (G, \mathcal{L}_G)$ and $\mathcal{H} = (H, \mathcal{L}_H)$ be labeled graphs. A *labeled-graph homomorphism* from $\mathcal{G}$ to $\mathcal{H}$ is a graph homomorphism $(\partial \Phi, \Phi) : G \to H$ such that $\mathcal{L}_H(\Phi(e)) = \mathcal{L}_G(e)$ for all $e \in \mathcal{E}(G)$. A labeled-graph homomorphism is actually a *labeled-graph isomorphism* if $(\partial \Phi, \Phi)$ is an isomorphism.

Definition 2.7. Suppose $\mathcal{G} = (G, \mathcal{L})$ is a labeled graph. The shift space $X_{\mathcal{G}}$ is called a *sofic shift*. Moreover, we say that $\mathcal{G}$ is *right-resolving* if $\mathcal{L}((v, w)) \neq \mathcal{L}((v, w'))$ for $v \in \mathcal{V}$ and $(v, w), (v, w') \in \mathcal{E}$.

It is seen that a SFT is also a sofic shift. Indeed, sofic shifts is an extension of SFTs.

Theorem 2.8. *(See [31, Proposition 3.1.6 & Theorem 3.2.1].) A shift space is sofic if and only if it is a factor of a SFT. Furthermore, a sofic shift is a SFT if and only if it has a presentation $(G, \mathcal{L})$ such that $\mathcal{L}_\infty$ is a conjugacy.*

A quantity that describes the complexity of a system is *topological entropy*. Suppose X is a shift space. Denote $\Gamma_k(X)$ the cardinality of the collection of words of length k. The topological entropy of X is then defined by

$$h(X) = \lim_{k \to \infty} \frac{\Gamma_k(X)}{k}.$$

Let X be a SFT with transition matrix T. Perron–Frobenius Theorem indicates that $h(X) = \log \rho(T)$, where $\rho(T)$ is the spectral radius of T. Nevertheless, if X is a sofic shift which is not right-resolving, then $\log \rho(T)$ might no longer be the topological entropy of X. Instead, we need to find X a right-resolving presentation via the so-called *subset construction method*.

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.

Theorem 2.9. *(See [31, Theorem 3.3.2].) Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph which is not right-resolving, and let $\mathcal{H} = (H, \mathcal{L}_H)$ be a right-resolving labeled graph constructed under the subset construction method. Then $X_{\mathcal{G}} = X_{\mathcal{H}}$.*

It is verified that $h(X) = h(X_{\mathcal{H}}) = \log \rho(T_{\mathcal{H}})$.

2.2. Multi-layer cellular neural networks

The fundamental part of a MCNN is one-layer cellular neural networks with inputs:

$$\frac{dx_i}{dt} = -x_i + \sum_{|k| \leqslant d} a_k y_{i+k} + \sum_{|\ell| \leqslant d} b_\ell u_{i+\ell} + z,$$

where $A = [-a_d, \ldots, a, \ldots, a_d]$, $B = [-b_d, \ldots, b, \ldots, b_d]$ are the feedback and controlling templates, respectively, and z is the threshold. The quantity x_i represents the state of the cell at i for $i \in \mathbb{Z}$. The output $\bar{y} = (\bar{y}_i) = (f(\bar{x}_i))$ of a stationary solution $(\bar{x}_i)$ is called a output pattern. A mosaic solution $\bar{x}$ satisfies $|\bar{x}_i| > 1$ and its corresponding pattern $\bar{y}$ is called a mosaic output pattern. Consider the mosaic solution $\bar{x}$, the necessary and sufficient conditions for state "+" at cell C_i, i.e., $\bar{y}_i = 1$, is

$$a - 1 + z > -\left(\sum_{0 < |k| \leqslant d} a_k \bar{y}_{i+k} + \sum_{|\ell| \leqslant d} b_\ell u_{i+\ell} \right), \qquad (5)$$

where $a = a_0$. Similarly, the necessary and sufficient conditions for state "−" at cell C_i, i.e., $\bar{y}_i = -1$, is

$$a - 1 - z > \sum_{0 < |k| \leqslant d} a_k \bar{y}_{i+k} + \sum_{|\ell| \leqslant d} b_\ell u_{i+\ell}. \qquad (6)$$

For simplicity, denoting $\bar{y}_i$ by y_i and rewriting the output patterns $y_{-d} \cdots y \cdots y_d$ coupled with input $u_{-d} \cdots u \cdots u_d$ as

$$\boxed{\begin{array}{c} y_{-d} \cdots y_{-1} y_0 y_1 \cdots y_d \\ u_{-d} \cdots u_{-1} u_0 u_1 \cdots u_d \end{array}} \equiv y_{-d} \cdots y_d \diamond u_{-d} \cdots u_d \in \{-, +\}^{\mathbb{Z}_{(2d+1) \times 2}}. \qquad (7)$$

Let

$$V^n = \left\{ v \in \mathbb{R}^n \colon v = (v_1, v_2, \ldots, v_n) \text{ and } |v_i| = 1, 1 \leqslant i \leqslant n \right\},$$

where $n = 4d + 1$, (5) and (6) can be rewritten in a compact form by introducing the following notation.

Denote $\alpha = (a_{-d}, \ldots, a_{-1}, a_1, \ldots, a_d)$, $\beta = (b_{-d}, \ldots, b, \ldots, b_d)$. Then, α can be used to represent A', the surrounding template of A without center, and β can be used to represent the template B. The basic set of admissible local patterns with "+" state in the center is defined as

$$\mathcal{B}(+, A, B, z) = \{v \diamond w \in V^n : a - 1 + z > -(\alpha \cdot v + \beta \cdot w)\}, \tag{8}$$

where $\cdot$ is the inner product in Euclidean space. Similarly, the basic set of admissible local patterns with "−" state in the center is defined as

$$\mathcal{B}(-, A, B, z) = \{v' \diamond w' \in V^n : a - 1 - z > \alpha \cdot v + \beta \cdot w\}. \tag{9}$$

Furthermore, the admissible local patterns induced by (A, B, z) can be denoted by

$$\mathcal{B}(A, B, z) = \big(\mathcal{B}(+, A, B, z), \mathcal{B}(-, A, B, z)\big). \tag{10}$$

The following theorem infers that the parameter space can be partitioned into finite equivalent subregions.

Theorem 2.10. *(See [27, Theorem 2.1].) Let*

$$P^{n+2} = \big\{(A, B, z) \,\big|\, A, B \in \mathcal{M}_{1 \times (2d+1)}(\mathbb{R}), \ z \in \mathbb{R}\big\},$$

where $\mathcal{M}_{r \times s}(\mathbb{R})$ means a $r \times s$ real matrix. There exists a positive integer $K(n)$ and the unique collection of open subsets $\{P_k\}_{k=1}^K$ of P^{n+2} satisfying

(i) $P^{n+2} = \bigcup_{k=1}^K \overline{P}_k$.
(ii) $P_k \cap P_\ell = \varnothing$ for all $k \neq \ell$.
(iii) $\mathcal{B}(A, B, z) = \mathcal{B}(\widetilde{A}, \widetilde{B}, \widetilde{z})$ if and only if $(A, B, z), (\widetilde{A}, \widetilde{B}, \widetilde{z}) \in P_k$ for some k.

Here $\overline{P}$ is the closure of P in P^{n+2}.

Suppose $\mathcal{B}$ is a basic set of admissible local patterns induced from a CNN. The output space $Y_U = Y_U(\mathcal{B})$ is defined by

$$Y_U = \left\{ \begin{array}{l} (\cdots y_{-1} y_0 y_1 \cdots) \in \{+, -\}^{\mathbb{Z}} : \text{there exists } (\cdots u_{-1} u_0 u_1 \cdots) \in \{+, -\}^{\mathbb{Z}} \\ \text{such that } (y_{i-d} \cdots y_i \cdots y_{i+d} \diamond u_{i-d} \cdots u_i \cdots u_{i+d}) \in \mathcal{B} \text{ for } i \in \mathbb{Z} \end{array} \right\}.$$

Theorem 2.11. *(See [27, Theorem 2.13].) Y_U is a sofic shift.*

In [27], the authors demonstrate that Theorems 2.10 and 2.11 still hold for MCNNs.

3. Two-layer cellular neural networks

This section comprises the fundamental part of this elucidation. A two-layer cellular neural network is realized as

$$\begin{cases} \dfrac{dx_i^{(1)}}{dt} = -x_i^{(1)} + \displaystyle\sum_{|k|\leqslant d} a_k^{(1)} y_{i+k}^{(1)} + \sum_{|k|\leqslant d} b_k^{(1)} u_{i+k}^{(1)} + z^{(1)}, \\[4mm] \dfrac{dx_i^{(2)}}{dt} = -x_i^{(2)} + \displaystyle\sum_{|k|\leqslant d} a_k^{(2)} y_{i+k}^{(2)} + \sum_{|k|\leqslant d} b_k^{(2)} u_{i+k}^{(2)} + z^{(2)}, \end{cases} \tag{11}$$

where $d \in \mathbb{N}$ and $u_i^{(2)} = y_i^{(1)}$ for $i \in \mathbb{Z}$. Suppose a partition of the parameter space is determined, that is, the templates

$$A^{(\ell)} = \big(a_{-d}^{(\ell)}, \dots, a_d^{(\ell)}\big), \qquad B^{(\ell)} = \big(b_{-d}^{(\ell)}, \dots, b_d^{(\ell)}\big), \qquad z^{(\ell)}, \quad \ell = 1, 2$$

are given. A stationary solution $\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(2)} \\ \mathbf{x}^{(1)} \end{pmatrix} = \begin{pmatrix} x_i^{(2)} \\ x_i^{(1)} \end{pmatrix}_{i\in\mathbb{Z}}$ is called *mosaic* if $|x_i^{(\ell)}| > 1$ for $n = 1, 2$ and $i \in \mathbb{Z}$. The output $\mathbf{y} = \begin{pmatrix} \mathbf{y}^{(2)} \\ \mathbf{y}^{(1)} \end{pmatrix} = \begin{pmatrix} y_i^{(2)} \\ y_i^{(1)} \end{pmatrix}_{i\in\mathbb{Z}}$ of a mosaic solution $\mathbf{x}$ is called a mosaic pattern. To clarify the discussion, we consider a simplified two-layer cellular neural network.

3.1. Simplified two-layer cellular neural networks

Let $d = 1$, $B^{(1)} = (0, 0, 0)$, and $a_{-1}^{(1)} = a_{-1}^{(2)} = 0$. If the other parameters are all nonzero, (11) is reduced as a simplified two-layer cellular neural network,

$$\begin{cases} \dfrac{dx_i^{(1)}}{dt} = -x_i^{(1)} + a^{(1)} y_i^{(1)} + a_r^{(1)} y_{i+1}^{(1)} + z^{(1)}, \\[4mm] \dfrac{dx_i^{(2)}}{dt} = -x_i^{(2)} + a^{(2)} y_i^{(2)} + a_r^{(2)} y_{i+1}^{(2)} + b^{(2)} u_i^{(2)} + b_r^{(2)} u_{i+1}^{(2)} + z^{(2)}. \end{cases} \tag{12}$$

Suppose $\mathbf{y} = \begin{pmatrix} \cdots y_{-1}^{(2)} y_0^{(2)} y_1^{(2)} \cdots \\ \cdots y_{-1}^{(1)} y_0^{(1)} y_1^{(1)} \cdots \end{pmatrix}$ is a mosaic pattern. For $i \in \mathbb{Z}$, $y_i^{(1)} = 1$ if and only if $x_i^{(1)} > 1$. This derives

$$a^{(1)} + z^{(1)} - 1 > -a_r^{(1)} y_{i+1}^{(1)}. \tag{13}$$

Similarly, $y_i^{(1)} = -1$ if and only if $x_i^{(1)} < -1$. This implies $y_i^{(1)} = -1$ if and only if

$$a^{(1)} - z^{(1)} - 1 > a_r^{(1)} y_{i+1}^{(1)}. \tag{14}$$

The same argument asserts

$$a^{(2)} + z^{(2)} - 1 > -a_r^{(2)} y_{i+1}^{(2)} - \big(b^{(2)} u_i^{(2)} + b_r^{(2)} u_{i+1}^{(2)}\big), \tag{15}$$

and

$$a^{(2)} - z^{(2)} - 1 > a_r^{(2)} y_{i+1}^{(2)} + \big(b^{(2)} u_i^{(2)} + b_r^{(2)} u_{i+1}^{(2)}\big) \tag{16}$$

are the necessary and sufficient condition for $y_i^{(2)} = -1$ and $y_i^{(2)} = 1$, respectively. Note that the quantity $u_i^{(2)}$ in (15) and (16) satisfies $|u_i^{(2)}| = 1$ for each i. Define $\xi_1 : \{-1, 1\} \to \mathbb{R}$ and $\xi_2 : \{-1, 1\}^{\mathbb{Z}_{3\times1}} \to \mathbb{R}$ by

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

Set

$$
\mathcal{B}^{(1)} = \left\{ \boxed{y^{(1)} y_r^{(1)}} : \ y^{(1)}, y_r^{(1)} \in \{-1, 1\} \text{ satisfy } (13), (14) \right\},
$$

$$
\mathcal{B}^{(2)} = \left\{ \boxed{\begin{array}{c} y^{(2)} y_r^{(2)} \\ u^{(2)} u_r^{(2)} \end{array}} : \ y^{(2)}, y_r^{(2)}, u^{(2)}, u_r^{(2)} \in \{-1, 1\} \text{ satisfy } (15), (16) \right\}.
$$

That is,

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

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

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 (12), respectively. The set of admissible local patterns $\mathcal{B}$ of (12) is then

$$
\mathcal{B} = \left\{ \boxed{\begin{array}{c} yy_r \\ uu_r \end{array}} : \ \boxed{\begin{array}{c} yy_r \\ uu_r \end{array}} \in \mathcal{B}^{(2)} \text{ and } \boxed{uu_r} \in \mathcal{B}^{(1)} \right\}.
$$

Since we only consider mosaic patterns, $y^{(1)}, y^{(2)}, y_r^{(1)}, y_r^{(2)}, u^{(2)}, u_r^{(2)} \in \{-1, 1\}$. This 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, and $a^{(2)} + z^{(2)} - 1 > -\xi_2(y_r^{(2)}, u^{(2)}, u_r^{(2)})$ 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. The "order" of lines $a^{(1)} + z^{(1)} - 1 = (-1)^\ell \xi_1(y_r^{(1)})$, $\ell = 0, 1$, come from the sign of $a_r^{(1)}$. Thus the parameter space $\{(a^{(1)}, a_r^{(1)}, z^{(1)})\}$ is partitioned into $2 \times 9 = 18$ regions. Similarly, the order of $a^{(2)} + z^{(2)} - 1 > (-1)^\ell \xi_2(y_r^{(2)}, u^{(2)}, u_r^{(2)})$ can be uniquely determined according to the following procedures.

(i) The signs of $a_r^{(2)}, b^{(2)}, b_r^{(2)}$.
(ii) The magnitude of $a_r^{(2)}, b^{(2)}, b_r^{(2)}$.
(iii) 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. Each region associates a basic set of admissible local patterns.

The above discussion indicates the following proposition.

Proposition 3.1. *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 (12). There exist 139 968 regions in $\mathcal{P}$ such that any two set of templates that locate in the same region infer the same basic set of admissible local patterns. Conversely, suppose $\mathcal{B} \subseteq \{-1, 1\}^{\mathbb{Z}_{2 \times 2}}$ comes from a simplified two-layer cellular neural networks. Then there exists a partition that admits $\mathcal{B}$ as its basic set of admissible local patterns.*

Example 3.2. 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. Numbered the partitions of $a^{(1)} - z^{(1)}$ and $a^{(2)} - z^{(2)}$ planes by a pair $[m_\ell, n_\ell]$, 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 $0 \leqslant m_1, n_1 \leqslant 2$ and $0 \leqslant m_2, n_2 \leqslant 8$. 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 $z^{(1)} = 2.5$. It is easy to chuck that the basic set of admissible local patterns is

$$\mathcal{B} = \left\{ \begin{array}{c} -- \\ -- \end{array}, \begin{array}{c} -- \\ -+ \end{array}, \begin{array}{c} -- \\ +- \end{array}, \begin{array}{c} +- \\ +- \end{array}, \begin{array}{c} ++ \\ -- \end{array}, \begin{array}{c} ++ \\ -+ \end{array}, \begin{array}{c} ++ \\ +- \end{array} \right\}.$$

3.1.1. Ordering matrix, transition matrix and graph

The previous section demonstrates that each partition of the parameter space associates with a collection of local patterns that allow for generalization of global patterns. Hence the basic set of admissible local patterns plays an essential role for investigating two-layer cellular neural networks. This section studies the structure of admissible local patterns through defining the ordering for each pattern. Substitute mosaic patterns -1 and 1 as symbols $-$ and $+$, respectively. Define the ordering matrix of $\{-, +\}^{\mathbb{Z}_{2 \times 2}}$ by

$$\mathbb{X} = \begin{pmatrix} \begin{array}{cccc} \begin{array}{c}--\\--\end{array} & \begin{array}{c}--\\-+\end{array} & \begin{array}{c}-+\\--\end{array} & \begin{array}{c}-+\\-+\end{array} \\ \begin{array}{c}--\\+-\end{array} & \begin{array}{c}--\\++\end{array} & \begin{array}{c}-+\\+-\end{array} & \begin{array}{c}-+\\++\end{array} \\ \begin{array}{c}+-\\--\end{array} & \begin{array}{c}+-\\-+\end{array} & \begin{array}{c}++\\--\end{array} & \begin{array}{c}++\\-+\end{array} \\ \begin{array}{c}+-\\+-\end{array} & \begin{array}{c}+-\\++\end{array} & \begin{array}{c}++\\+-\end{array} & \begin{array}{c}++\\++\end{array} \end{array} \end{pmatrix} = (x_{pq})_{1 \leqslant p, q \leqslant 4}.$$

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 $T \equiv T(\mathcal{B}) \in \mathcal{M}_4(\mathbb{R})$ is a 4×4 matrix defined by

$$T(p, q) = \begin{cases} 1, & \text{if } x_{pq} \in \mathcal{B}; \\ 0, & \text{otherwise.} \end{cases}$$

Let $\mathcal{V} = \{0, 1, 2, 3\}$, where

$$0 \equiv \begin{array}{c}-\\-\end{array}, \qquad 1 \equiv \begin{array}{c}-\\+\end{array}, \qquad 2 \equiv \begin{array}{c}+\\-\end{array}, \qquad 3 \equiv \begin{array}{c}+\\+\end{array}.$$

There exists an edge $e \in \mathcal{E}$ if and only if $T(i(e) + 1, t(e) + 1) = 1$. This infers $G_T = (\mathcal{V}, \mathcal{E})$ is a graph representation of T.

Let $\mathbf{Y} \subseteq \{-, +\}^{\mathbb{Z}_{\infty \times 2}}$ be the solution space of (12). That is,

$$\mathbf{Y} = \left\{ \begin{pmatrix} y_i \\ u_i \end{pmatrix}_{i \in \mathbb{Z}} : \begin{array}{c} y_i y_{i+1} \\ u_i u_{i+1} \end{array} \in \mathcal{B} \text{ for } i \in \mathbb{Z} \right\}.$$

It comes immediately from Theorem 2.4 that **Y** is a shift of finite type.

For ease of notation, denote $\boxed{\begin{smallmatrix} y_1 y_2 \\ u_1 u_2 \end{smallmatrix}}$ by $y_1 y_2 \diamond u_1 u_2$ and

$$\mathbf{y} \diamond \mathbf{u} \equiv \boxed{\begin{smallmatrix} \mathbf{y} \\ \mathbf{u} \end{smallmatrix}} = \frac{\cdots y_{-2} y_{-1} y_0 y_1 y_2 \cdots}{\cdots u_{-2} u_{-1} u_0 u_1 u_2 \cdots}, \quad \text{where } \mathbf{y} = (y_i)_{i \in \mathbb{Z}}, \ \mathbf{u} = (u_i)_{i \in \mathbb{Z}}.$$

Then we can write **Y** as

$$\mathbf{Y} = \{ \mathbf{y} \diamond \mathbf{u} \colon y_i y_{i+1} \diamond u_i u_{i+1} \in \mathcal{B} \text{ for } i \in \mathbb{Z} \}.$$

Define $\phi^{(1)}, \phi^{(2)} \colon \mathbf{Y} \to \{-, +\}^{\mathbb{Z}}$ by

$$\phi^{(1)}(\mathbf{y} \diamond \mathbf{u}) = \mathbf{u} \quad \text{and} \quad \phi^{(2)}(\mathbf{y} \diamond \mathbf{u}) = \mathbf{y}. \tag{17}$$

Set $Y^{(\ell)} = \phi^{(\ell)}(\mathbf{Y})$ for $\ell = 1, 2$. $Y^{(1)}$ is called the *hidden space*, and $Y^{(2)}$ is called the *output space*. Obviously the dynamical behavior of the output space $Y^{(2)}$ is influenced by the hidden space $Y^{(1)}$. For instance, a phenomenon which cannot be seen in one-layer cellular neural networks is that $Y^{(1)}$ would break the symmetry of the entropy diagram of $Y^{(2)}$ [37,27]. This motivates the study of the relation between $Y^{(1)}$ and $Y^{(2)}$.

Since $Y^{(1)}$ and $Y^{(2)}$ are derived from the same system **Y**, it is natural to ask the following questions:

Question 1. Is there a relation between $Y^{(1)}$ and $Y^{(2)}$? For instance, can we define a factor map π from $Y^{(1)}$ to $Y^{(2)}$ such that $\phi^{(2)} = \pi \circ \phi^{(1)}$?

Question 2. Suppose that $\pi \colon Y^{(1)} \to Y^{(2)}$ exists.

 (i) Under what conditions does π preserve topological entropy?
(ii) Does π help for demonstrating the structure of $Y^{(1)}$ and $Y^{(2)}$?

The structure of $Y^{(1)}$ and $Y^{(2)}$ are essential for these questions. We introduce the labeled graph first.

3.1.2. Labeled graph and symbolic transition matrix

Assume that $G_T = (\mathcal{V}, \mathcal{E})$ is the graph representation of (12). Let $\mathcal{A} = \{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$, where

$$\alpha_0 = --, \qquad \alpha_1 = -+, \qquad \alpha_2 = +- \quad \text{and} \quad \alpha_3 = ++.$$

Define $\mathcal{L}^{(1)}, \mathcal{L}^{(2)} \colon \mathcal{E} \to \mathcal{A}$ by

$$\mathcal{L}^{(1)}(e) = \alpha_{2\tau(i(e)) + \tau(t(e))}, \quad \text{where } \tau(c) := c \mod 2, \tag{18}$$

$$\mathcal{L}^{(2)}(e) = \alpha_{2[i(e)/2] + [t(e)/2]}, \quad \text{where } [\cdot] \text{ is the Gauss function.} \tag{19}$$

These two labeling $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ define two labeled graphs $\mathcal{G}^{(1)} = (G_T, \mathcal{L}^{(1)})$ and $\mathcal{G}^{(2)} = (G_T, \mathcal{L}^{(2)})$, respectively. Ban et al. [27] demonstrated that the output space is a sofic shift with graph representation $\mathcal{G}^{(2)}$. With a small modification we also get that $Y^{(1)}$ is a sofic shift with graph representation $\mathcal{G}^{(1)}$.

Theorem 3.3. *Suppose* $\mathbf{Y} \subset \{-, +\}^{\mathbb{Z}_{\infty} \times 2}$ *is the solution space of a two-layer cellular neural network. Then* $Y^{(\ell)}$ *is a sofic shift and* $Y^{(\ell)} = X_{\mathcal{G}^{(\ell)}}$ *for* $\ell = 1, 2$.

It is worth emphasizing that $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ share the same underlying graph, but have different labeling. Moreover, it is seen that $(\mathcal{L}^{(\ell)})_\infty$ is conjugate to $\phi^{(\ell)}$, where $(\mathcal{L}^{(\ell)})_\infty : X_{G_T} \to X_{\mathcal{G}^{(\ell)}}$ is defined by $(\mathcal{L}^{(\ell)})_\infty(\xi)_j = \mathcal{L}^{(\ell)}((\xi_j, \xi_{j+1}))$ for $j \in \mathbb{Z}$. Since $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}$ are two different labeled graphs that share the same underlying graph, their transition matrices $T^{(1)}, T^{(2)}$ make no difference. This motivates the introduction of a symbolic transition matrix. The symbolic transition matrix $S^{(\ell)}$ of $\mathcal{G}^{(\ell)}$ is defined by

$$S^{(\ell)}(p,q) = \begin{cases} \alpha_j, & \text{if } T^{(\ell)}(p,q) = 1 \text{ and } \mathcal{L}^{(\ell)}((p,q)) = \alpha_j \text{ for some } j; \\ \varnothing, & \text{otherwise.} \end{cases} \tag{20}$$

Herein $\varnothing$ 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. It follows from this definition that the symbolic transition matrix S of a right-resolving labeled graph must satisfy $S(p,q) \neq S(p,q')$ for all p, q, q' if $S(p,q) \neq \varnothing$. Conversely, suppose $S(p,q) \neq \varnothing$. $S(p,q) \neq S(p,q')$ for all q' demonstrates that no two edges starting from the same initial state carry the same label. This derives the following lemma.

Lemma 3.4. *Suppose $\mathcal{G}$ is a labeled graph and S is its symbolic transition matrix. Then $\mathcal{G}$ is right-resolving if and only if, for each p, q, $S(p,q) \neq S(p,q')$ for all q' provided $S(p,q) \neq \varnothing$.*

Example 3.5. Continue with Example 3.2. The symbolic transition matrices for $Y^{(1)}$ and $Y^{(2)}$ are

$$S^{(1)} = \begin{pmatrix} \alpha_0 & \alpha_1 & \varnothing & \varnothing \\ \alpha_2 & \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & \alpha_0 & \alpha_1 \\ \alpha_2 & \varnothing & \alpha_2 & \varnothing \end{pmatrix} \quad \text{and} \quad S^{(2)} = \begin{pmatrix} \alpha_0 & \alpha_0 & \varnothing & \varnothing \\ \alpha_0 & \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & \alpha_3 & \alpha_3 \\ \alpha_2 & \varnothing & \alpha_3 & \varnothing \end{pmatrix},$$

respectively. $S^{(1)}$ and $S^{(2)}$ infer that $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ make no difference except the labeling. Moreover, neither $\mathcal{G}^{(1)}$ nor $\mathcal{G}^{(2)}$ is right-resolving.

3.2. Ordering matrix and labeled graph for general two-layer cellular neural networks

For a general two-layer cellular neural network, suppose the admissible local pattern of (11) is of the form

$$\boxed{\begin{array}{l} y_{-d} \cdots y_{-1} y_0 y_1 \cdots y_d \\ u_{-d} \cdots u_{-1} u_0 u_1 \cdots u_d \end{array}} \equiv y_{-d} \cdots y_d \diamond u_{-d} \cdots u_d \in \{-, +\}^{\mathbb{Z}_{(2d+1) \times 2}}.$$

Note that the entry is allowed to be left empty if its corresponding parameter is zero. For example, if $a_{-d}^{(2)} = 0$, then the local pattern is of the form

$$\boxed{\begin{array}{l} y_{-d+1} \cdots y_{-1} y_0 y_1 \cdots y_d \\ u_{-d} u_{-d+1} \cdots u_{-1} u_0 u_1 \cdots u_d \end{array}} \equiv y_{-d+1} \cdots y_d \diamond u_{-d} \cdots u_d$$

and vice versa.

For simplicity, we assume that none of the entries in the basic pattern are empty. Define $\chi : \{-, +\} \to \{0, 1\}$ by $\chi(-) = 0$ and $\chi(+) = 1$. Dividing the basic pattern into two components such that each component is of the form $y_{-d} \cdots y_{d-1} \diamond u_{-d} \cdots u_{d-1}$. Let $\varrho : y_{-d} \cdots y_{d-1} \diamond u_{-d} \cdots u_{d-1} \to \mathbb{Z}$ be defined by

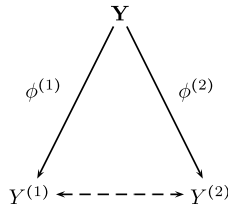**Fig. 1.** Since $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ are both right-resolving, $\phi^{(1)}$ and $\phi^{(2)}$ are finite-to-one factors and **Y** is a common extension space of $Y^{(1)}$ and $Y^{(2)}$.

$$\varrho(y_{-d} \cdots y_{d-1} \diamond u_{-d} \cdots u_{d-1}) = \chi(y_{-d}) + 2\chi(y_{-d+1}) + \cdots + 2^{2d-1}\chi(y_{d-1})$$
$$+ 2^{2d}\chi(u_{-d}) + \cdots + 2^{4d}\chi(u_d).$$

The ordering matrix is then a $2^{4d+1} \times 2^{4d+1}$ matrix which is indexed by $\varrho(y_{-d} \cdots y_{d-1} \diamond u_{-d} \cdots u_{d-1})$.

Suppose a basic set of admissible local patterns of (11) is determined. In a similar way as before, we can define two labeled graphs sharing the same underlying graph with different labeling. Some examples are considered later.

3.3. Classification of hidden and output spaces

This section investigates whether there is a relation between $Y^{(1)}$ and $Y^{(2)}$. It turns out that topological entropy provides some evidence for the existence of the map $\pi : Y^{(1)} \to Y^{(2)}$ (or $\pi' : Y^{(2)} \to Y^{(1)}$). For the rest of this section, we assume that $h(Y^{(1)}) = h(Y^{(2)})$ unless otherwise stated, where $h(X)$ indicates the topological entropy of X. To clarify the discussion, we first consider the case where both $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are right-resolving labeled graphs. If either $\mathcal{G}^{(1)}$ or $\mathcal{G}^{(2)}$ is not right-resolving, we need to construct a right-resolving labeled graph that still presents the same sofic shift.

3.3.1. Two right-resolving labeled graphs

Suppose $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are both right-resolving. First we consider a relation between two shift spaces called *finite equivalence*.

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 $|\phi^{-1}(y)| \leqslant M$ for $y \in Y$. Two shift spaces X and Y are *finitely equivalent*, 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.

Proposition 3.6. *If $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are both right-resolving, then $Y^{(1)}$ and $Y^{(2)}$ are finitely equivalent.*

Proof. An analogous argument as in the proof of Theorem 3.3 implies $\mathbf{Y} = X_{G_T}$, hence **Y** is a shift of finite type. $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are both right-resolving indicates that $\phi^{(1)} : \mathbf{Y} \to Y^{(1)}$ and $\phi^{(2)} : \mathbf{Y} \to Y^{(2)}$ are both finite-to-one factor maps. Hence **Y** is a common extension of $Y^{(1)}$ and $Y^{(2)}$. See Fig. 1. $\quad\square$

It is well known that finite equivalence is an equivalent relation. The following proposition shows that conjugacy between two spaces demonstrates finite shift equivalence.

Proposition 3.7. *If X and Y are conjugate, then X is finitely shift equivalent to Y.*

Proof. Suppose that W is a shift of finite type with a finite-to-one factor map $\phi : W \to Y$. Let $\psi : X \to Y$ be a conjugacy. Define $\phi' : W \to Y$ by $\phi' = \psi \circ \phi$. It comes immediately that ϕ' is also a finite-to-one factor from W to Y and derive the desired result. $\quad\square$
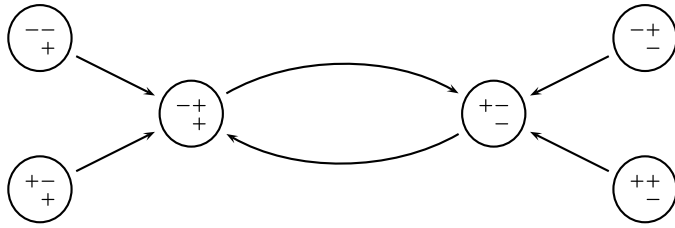
Fig. 2. The underlying graph G_T of Example 3.8.

Example 3.8. Consider two-layer cellular neural network (11) with $d = 1, a_{-1}^{(1)} = 0, B^{(1)} = (0, 0, 0)$ and $b_{-1}^{(2)} = 0$. Then the ordering matrix is defined as

$$\mathbb{X}_{3\times2} = \begin{pmatrix} \end{pmatrix}.$$

Set $\mathbb{X}_{3\times2} = (x_{pq})_{1\leqslant p,q\leqslant8}$. For each nonempty pattern $x_{pq} = \mathbf{y}_{pq} \diamond \mathbf{u}_{pq}$, we define $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ as

$$\mathcal{L}^{(1)}(\mathbf{y}_{pq} \diamond \mathbf{u}_{pq}) = \mathbf{u}_{pq}, \qquad \mathcal{L}^{(2)}(\mathbf{y}_{pq} \diamond \mathbf{u}_{pq}) = \mathbf{y}_{pq}.$$

Let the parameters be

$$A^{(1)} = (4, 1), \qquad A^{(2)} = (-1, -3, -2.5),$$
$$B^{(2)} = (-2, -2.6), \qquad (z^{(1)}, z^{(2)}.) = (2, 0).$$

The basic set of admissible local patterns $\mathcal{B}$ then consists of the following patterns:

Fig. 2 illustrates the graph representation of **Y**. A straightforward examination indicates that $(\mathbf{Y}, \phi^{(1)}, \phi^{(2)})$ is a finite equivalence between $Y^{(1)}$ and $Y^{(2)}$.

Definition 3.9. A graph G is irreducible if for any two vertices v_1, v_2 of G there is a path starts from v_1 and terminates at v_2. A labeled graph $\mathcal{G} = (G, \mathcal{L})$ is irreducible if its underlying graph is irreducible. Moreover, a sofic shift Y is irreducible if Y has an irreducible labeled graph representation.

Definition 3.10. A graph G is essential if for every vertex v there are edges e_1, e_2 such that $i(e_1) = v$ and $t(e_2) = v$.

Proposition 3.7 motivates the determination of the conjugacy between $Y^{(1)}$ and $Y^{(2)}$. A necessary and sufficient condition thus follows for the case where $Y^{(1)}$ and $Y^{(2)}$ are irreducible sofic shifts.

Theorem 3.11. *Suppose* $Y^{(1)}, Y^{(2)}$ *are irreducible sofic shifts. If* G_T *is an essential graph, then* $Y^{(1)} \cong Y^{(2)}$ *if and only if* $P S^{(1)} = S^{(2)} P$, *where*

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

Proof. Without loss of generality, we may assume that the vertex set of G_T is $\mathcal{V} = \{0, 1, 2, 3\}$. Suppose $P S^{(1)} = S^{(2)} P$. Define $(\partial \Psi, \Psi) : \mathcal{G}^{(1)} \to \mathcal{G}^{(2)}$ by

$$
\partial \Psi(v) = \begin{cases} 3 - v, & v = 1, 2; \\ v, & \text{otherwise,} \end{cases}
$$

$$
\Psi(e) = e', \quad \text{where } i(e') = \partial \Psi(i(e)), t(e') = \partial \Psi(t(e)) \tag{21}
$$

for all $v \in \mathcal{V}, e \in \mathcal{E}$. Then $\partial \Psi$ is one-to-one and onto. $P S^{(1)} = S^{(2)} P$ asserts $P T^{(1)} = T^{(2)} P$ and

$$
T^{(1)}(p, q) = T^{(2)}\big(1 + \partial \Psi(p - 1), 1 + \partial \Psi(q - 1)\big), \quad 1 \leqslant p, q \leqslant 4. \tag{22}
$$

That is, an edge $e \in \mathcal{E}$ infers another edge e' with $i(e') = \partial \Psi(i(e))$ and $t(e') = \partial \Psi(t(e))$. Hence Ψ is well defined, one-to-one and onto. This shows that $(\partial \Psi, \Psi)$ is a graph isomorphism. Moreover, it can be verified that

$$
\big[\partial \Psi(v)/2\big] = \tau(v), \quad \text{for } v = 0, 1, 2, 3.
$$

For each $e \in \mathcal{E}$,

$$
\mathcal{L}^{(2)}\big(\Psi(e)\big) = \alpha_{2[\partial \Psi(i(e))/2] + [\partial \Psi(t(e))/2]} = \alpha_{2\tau(i(e)) + \tau(t(e))} = \mathcal{L}^{(1)}(e).
$$

This demonstrates $(\partial \Psi, \Psi) : \mathcal{G}^{(1)} \to \mathcal{G}^{(2)}$ is a labeled-graph isomorphism. In other words, $Y^{(1)} \cong Y^{(2)}$.

Conversely, suppose that $Y^{(1)} \cong Y^{(2)}$. Since $Y^{(1)}, Y^{(2)}$ are irreducible and $T^{(1)}$ and $T^{(2)}$ are $0 - 1$ matrices, there exists a conjugacy matrix D such that $D S^{(1)} = S^{(2)} D$ and D is also a 0–1 matrix. Similar as above, D defines a graph isomorphism $(\partial \Psi, \Psi) : \mathcal{G}^{(1)} \to \mathcal{G}^{(2)}$. If $\partial \Psi(0) \neq 0$, it comes immediately from $D S^{(1)} = S^{(2)} D$ that $\partial \Psi(0) = 1$ and $\partial \Psi(2) = 0$. Assume that $\partial \Psi(1) = 2$ and $\partial \Psi(3) = 3$. It is easily seen that $\mathcal{L}^{(2)}(\Psi(2, i)) \neq \mathcal{L}^{(1)}(2, i)$ for all $0 \leqslant i \leqslant 3$, which gets a contradiction to $D S^{(1)} = S^{(2)} D$. Another contradiction occurs for the case that $\partial \Psi(1) = 3, \partial \Psi(3) = 2$. This forces $\partial \Psi(0) = 0$. Repeating the same procedures derive $\partial \Psi(1) = 2, \partial \Psi(2) = 1$, and $\partial \Psi(3) = 3$. In other words, $D = P$.

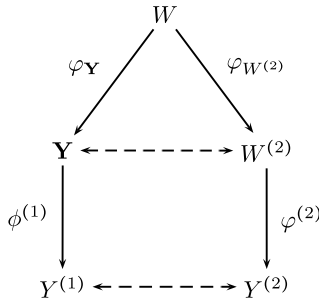This completes the proof. $\quad \square$

Fig. 3. For the case that $\mathcal{G}^{(1)}$ is right-resolving while $\mathcal{G}^{(2)}$ is not, we construct $W^{(2)} \equiv X_{\mathcal{G}_{Y^{(2)}}}$ such that $\varphi^{(2)} : W^{(2)} \to Y^{(2)}$ is finite-to-one, where $\mathcal{G}_{Y^{(2)}}$ is a right-resolving labeled graph representation of $Y^{(2)}$. After constructing W such that there exist two finite-to-one factors $\varphi_{\mathbf{Y}} : W \to \mathbf{Y}, \varphi_{W^{(2)}} : W \to W^{(2)}$, we get a common extension of $Y^{(1)}$ and $Y^{(2)}$.

3.3.2. The case that some labeled graphs are not right-resolving

When $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ are both right-resolving, the common extension for $Y^{(1)}$ and $Y^{(2)}$ is the original space $\mathbf{Y}$. However, $\mathbf{Y}$ is no longer $Y^{(1)}, Y^{(2)}$'s common extension if either $\mathcal{G}^{(1)}$ or $\mathcal{G}^{(2)}$ is not right-resolving. For this reason, we need to construct a real common extension W.

Theorem 3.12. *Suppose either $\mathcal{G}^{(1)}$ or $\mathcal{G}^{(2)}$ is not right-resolving. If $h(Y^{(1)}) = h(Y^{(2)})$, then there exists finite equivalence $(W, \phi_{W^{(1)}}, \phi_{W^{(2)}})$ between $Y^{(1)}$ and $Y^{(2)}$. Moreover, there exists an integral matrix F such that $F T_{G^{(1)}} = T_{G^{(2)}} F$.*

Proof. The proof is similar as the proof in [31]. We sketch the process for reader's convenience. For the simplicity, we assume that the underlying graph G_T is essential. Since ϕ_X is right-resolving and ϕ_Y is not, the labeling $\mathcal{L}_X$ on the original graph representation G_T of W is right-resolving and $\mathcal{L}_Y$ on G_T is not. Using subset construction method we get a new graph, say G_Y, and a new labeling, we still denote by $\mathcal{L}_Y$ for the simplification of notation, which is right-resolving. Moreover, the right-resolving labeled graph $\mathcal{G}_Y = (G_Y, \mathcal{L}_Y)$ still represents Y. Notice that the vertex set $\mathcal{V}(G_T)$ of G_T is a proper subset of the vertex set $\mathcal{V}(G_Y)$. Let A_{G_Y} be the transition matrix of G_Y with the same order as T, that is, the first four rows of A_{G_Y} represent the same vertices as T.

Constructing a graph $M = (\mathcal{V}(M), \mathcal{E}(M))$ as follows. Suppose $v_1, v_2 \in \mathcal{V}(G_Y)$ and there is an edge in G_Y starts at v_1 and terminates at v_2, that is, $A_{G_Y}(v_1, v_2) = 1$. If v_1 is a vertex of G_T, then $\{v_1, v_2\}$ is a vertex of M. For $\{v_1, v_2\}, \{v_3, v_4\} \in \mathcal{V}(G_Y)$, there is an edge starts from $\{v_1, v_2\}$ and stops at $\{v_3, v_4\}$ if and only if $v_3 \in v_2$. Note that $v_2 \in \mathcal{V}(G_Y)$ is a subset of $\mathcal{V}(G_T)$. It is easily seen that the graph M is well defined.

Define $\Phi_{G_T} : \mathcal{V}(M) \to \mathcal{V}(G_T)$ by $\Phi_{G_T}(\{v_1, v_2\}) = v_1$. Observe that if $\{v_1, v_2\}$ is followed by $\{v_2, v_3\}$ in M, then v_1 is followed by v_1 in G_T. Hence Φ_{G_T} is the vertex map of a graph homomorphism from M to G_T and induces $\phi_W \equiv (\Phi_{G_T})_\infty : \widetilde{W} \equiv X_M \to W$. It comes immediately that ϕ_W is a right-resolving factor map.

Similarly, let $\Phi_{G_Y} : \mathcal{V}(M) \to \mathcal{V}(G_Y)$ by $\Phi_{G_Y}(\{v_1, v_2\}) = v_2$. This indicates $\phi_{W_Y} \equiv (\Phi_{G_Y})_\infty$ is a left-resolving factor map from $\widetilde{W}$ to $W_Y \equiv X_{G_Y}$. Since G_Y comes from applying subset construction to G_T, $X_{\mathcal{G}_{Y=(G_Y, \mathcal{L}_Y)}} = Y$.

This completes the proof. $\quad\square$

Remark 3.13. We mention that Proposition 3.6 can be treated as a special case of Theorem 3.12. Let W and $W^{(2)}$ in Fig. 3 be $\mathbf{Y}$ at the same time. Then $\varphi_{\mathbf{Y}}, \varphi_{W^{(2)}}$ and the dash line between $\mathbf{Y}$ and $W^{(2)}$ are the identity map. Additionally, $\varphi^{(2)} = \phi^{(2)}$. Fig. 3 is then simplified as Fig. 1.

It is natural to ask whether, in Fig. 4, the dashed line between $W^{(1)}$ and $W^{(2)}$ can be replaced with a solid line? To answer this question, we introduce the so-called *factor-like* matrix.
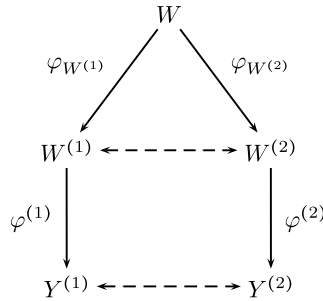
Fig. 4. If neither $\mathcal{G}^{(1)}$ or $\mathcal{G}^{(2)}$ is not right-resolving, we construct $W^{(1)} \equiv X_{\mathcal{G}_{Y^{(1)}}}$ and $W^{(2)} \equiv X_{\mathcal{G}_{Y^{(2)}}}$ such that $\varphi^{(1)} : W^{(1)} \rightarrow Y^{(1)}$ and $\varphi^{(2)} : W^{(2)} \rightarrow Y^{(2)}$ are both finite-to-one, where $\mathcal{G}_{Y^{(1)}}$ and $\mathcal{G}_{Y^{(2)}}$ are right-resolving labeled graph representation of $(G, \mathcal{L}^{(1)})$ and $(G, \mathcal{L}^{(2)})$, respectively.

Definition 3.14. Given two integral matrices $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{n \times n}$. Suppose $F \in \mathbb{R}^{m \times n}$ is an integral matrix satisfies $FA = BF$. F is called factor-like if there is at most one 1's in each row of F.

Proposition 3.15. *Under the same assumption of Theorem* 3.12. *If F is factor-like, then there exists $\pi : W^{(1)} \rightarrow W^{(2)}$ which preserves topological entropy.*

Proof. Define $\pi : W^{(1)} \rightarrow W^{(2)}$ by $\pi(x) = y$ if $F(x_i, y_i) = 1$ for all $i \in \mathbb{Z}$. Since F is factor-like, π is well defined and injective. It follows immediately that π preserves topological entropy since $h(W^{(1)}) = h(Y^{(1)}) = h(Y^{(2)}) = h(W^{(2)})$. $\square$

Example 3.16. Consider (11) with $d = 1$. Suppose the parameters are set up as

$$\left(a_{-1}^{(2)}, a_0^{(2)}, a_1^{(2)}\right) = (-4, -3, -2), \qquad \left(b_{-1}^{(2)}, b_0^{(2)}, b_1^{(2)}, z^{(2)}\right) = (0, 1, 6.5, 2.8),$$

and

$$\left(a_{-1}^{(1)}, a_0^{(1)}, a_1^{(1)}, z^{(1)}\right) = (0, 1.7, 0.3), \qquad \left(b_{-1}^{(1)}, b_0^{(1)}, b_1^{(1)}, z^{(1)}\right) = (0, 0, 0, 0.2).$$

Then the basic set of admissible local patterns are the following:

$$\boxed{\begin{matrix} + & - & - \\ - & - \end{matrix}} \quad \boxed{\begin{matrix} + & - & - \\ + & - \end{matrix}} \quad \boxed{\begin{matrix} + & - & + \\ - & - \end{matrix}} \quad \boxed{\begin{matrix} + & - & + \\ + & - \end{matrix}} \quad \boxed{\begin{matrix} - & + & - \\ - & + \end{matrix}} \quad \boxed{\begin{matrix} - & + & - \\ + & + \end{matrix}}$$

$$\boxed{\begin{matrix} - & + & + \\ - & + \end{matrix}} \quad \boxed{\begin{matrix} - & + & + \\ + & + \end{matrix}} \quad \boxed{\begin{matrix} + & + & - \\ - & + \end{matrix}} \quad \boxed{\begin{matrix} + & + & - \\ + & + \end{matrix}} \quad \boxed{\begin{matrix} + & + & + \\ + & + \end{matrix}}.$$
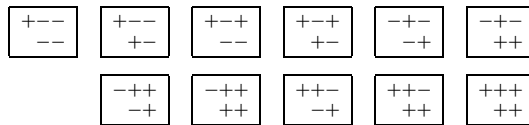
The transition matrix of the underlying graph G_T is

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

Fig. 5. The graph representation of the common extension in Example 3.16.

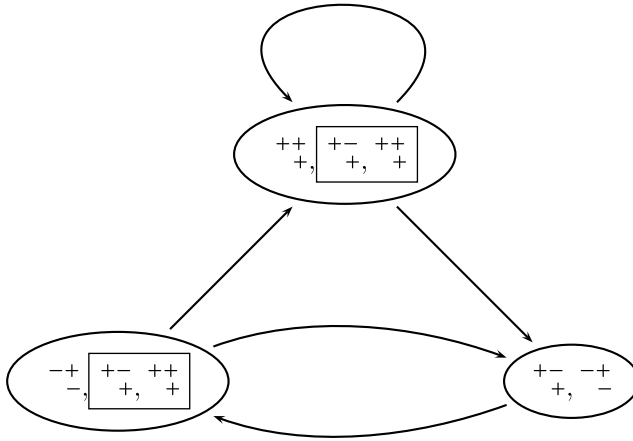It follows that the vertices of the essential subgraph G' is $\{-+\diamond-, +-\diamond+, ++\diamond+\} \equiv \{x_1, x_2, x_3\}$ with transition matrix

$$T' = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

The transition matrix of $G_{Y^{(1)}}$, indexed by $\{-+\diamond-, \{+-\diamond+, ++\diamond+\}\} \equiv \{y_1, y_2\}$, is

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

Similar as above example, we construct M with $\mathcal{V}(M) = \{x_1 y_2, x_2 y_1, x_3 y_2\}$ and $\mathcal{E}(M) = \{(x_1 y_2, x_2 y_1),$ $(x_1 y_2, x_3 y_2), (x_2 y_1, x_1 y_2), (x_3 y_2, x_2 y_1), (x_3 y_2, x_3 y_2)\}$. Then $W = X_M$ is a common extension space of $Y^{(1)}$ and $Y^{(2)}$ (see Fig. 5). Moreover,

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

satisfies $T'F = FH$ is factor-like. Hence there exists a factor map $\pi : \mathbf{Y} \to W^{(1)}$.

Proposition 3.15 asserts whether or not the dashed line between $W^{(1)}$ and $W^{(2)}$ in Fig. 4 can be replaced by a solid line. It is natural to ask when the bottom dashed line would also be solid. The following theorem indicates an affirmative answer.

Theorem 3.17. *Along with the same assumption of Theorem* 3.12. *Suppose F is factor-like and* $Y^{(1)}, Y^{(2)}$ *are shifts of finite type, then there exists* $\bar{\pi} : Y^{(1)} \to Y^{(2)}$ *which preserves topological entropy.*

Proof. Suppose $Y^{(1)}, Y^{(2)}$ are shifts of finite type. Theorem 2.8 derives $\phi^{(1)}$ and $\phi^{(2)}$ are both conjugacy. Since F is factor-like, Proposition 3.15 indicates that there exists $\pi : W^{(1)} \to W^{(2)}$ which preserves topological entropy. Define $\bar{\pi} : Y^{(1)} \to Y^{(2)}$ by $\bar{\pi}(x) = (\phi^{(2)} \circ \pi \circ (\phi^{(1)})^{-1})(x)$. The proof is done. $\quad\square$

Aside from verifying the conjugacy of a map, the synchronization of paths in a labeled graph provided a necessary and sufficient determination of the conjugacy. Suppose X is a shift space, w is a *word* in X if there exist $N \in \mathbb{N}$ and $x \in X$ such that $w = x_j x_{j+1} \ldots x_{j+N-1}$ for some $j \in \mathbb{Z}$.

Definition 3.18. Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph and let w be a word of $X_{\mathcal{G}}$. We say that w is a synchronizing word for $\mathcal{G}$ if all paths in G presenting w terminate at the same vertex.

Theorem 3.19. *(See [31, Theorem 3.4.17].) Suppose that* **S** *is a sofic shift with right-resolving labeled graph* $\mathcal{G}$. *If there exists* $N \in \mathbb{N}$ *such that all words of* **S** *of length* N *are synchronizing for* $\mathcal{G}$, *then* **S** *is an* N-*step shift of finite type.*

Conversely, suppose **S** *is an irreducible sofic shift and its minimal right-resolving presentation* $\mathcal{G}$ *has* n *vertices. If* **S** *is a shift of finite type, then it must be* $(n^2 - n)$-*step.*

The next theorem follows immediately from Theorem 3.19.

Theorem 3.20. *Along with the same assumption of Theorem* 3.12. *If* F *is factor-like from* $T_{G_{Y^{(1)}}}$ *to* $T_{G_{Y^{(2)}}}$ *and all words of* $Y^{(1)}$ *of length* N *are synchronizing for some* $N \in \mathbb{N}$, *then there exists a factor map* $\pi : Y^{(1)} \to Y^{(2)}$ *which preserves entropy.*

Example 3.21. In Example 3.16, the essential graph for **Y** is
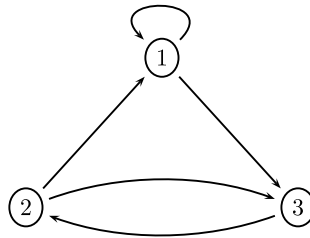
where

$$1 \equiv ++\diamond+, \qquad 2 \equiv -+\diamond-, \qquad 3 \equiv +-\diamond+.$$

Recall that $\mathcal{L}^{(1)}(11) = \mathcal{L}^{(1)}(13) = ++$ indicates $\mathcal{G}^{(1)}$ is not right-resolving. The minimal right-resolving presentation of $Y^{(1)}$ is

It is easy to see that $Y^{(1)}$ is a shift of finite type via Theorems 3.19 and 2.8 demonstrates $\varphi^{(1)} : W^{(1)} \to Y^{(1)}$ is a conjugacy. The conjugacy of $\phi^{(2)}$ can be derived analogously. Let $\bar{\pi} = \phi^{(1)} \circ \pi \circ (\phi^{(2)})^{-1}$, we have a factor map $\bar{\pi} : Y^{(2)} \to Y^{(1)}$.

Since, in Example 3.16, there exists a factor map $\bar{\pi}$ from $Y^{(2)}$ to $Y^{(1)}$, it is natural to ask whether $\bar{\pi}$ is invertible. That is, is $\bar{\pi}$ actually a conjugacy? To answer this question, we introduce a definition first.

Definition 3.22. Let A and B be nonnegative integral matrices. An *elementary equivalence* from A to B is a pair (R, S) of rectangular nonnegative matrices satisfying $A = RS$ and $B = SR$. In this case we write $(R, S) : A \approx B$.

A *strong shift equivalence* from A to B is a sequence of ℓ elementary equivalences

$$(R_1, S_1)\colon A = A_0 \approx A_1, \qquad (R_2, S_2)\colon A_1 \approx A_2, \qquad \ldots, \qquad (R_\ell, S_\ell)\colon A_{\ell-1} \approx A_\ell = B$$

for some ℓ. In this case we say that A is strong shift equivalent to B and write $A \sim_{\mathcal{F}_{SS}} B$.

We now lay the groundwork for Williams' criterion.

Theorem 3.23 *(Classification Theorem). Suppose A and B are nonnegative integral matrices. X_A and X_B are conjugate if and only if A and B are strong shift equivalent.*

Example 3.24. Continued from Example 3.16. Set

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

Then $(E, F)\colon H \approx T'$. Classification Theorem indicates $W^{(1)}$ is conjugate to $\mathbf{Y}$. Example 3.21 shows that $\phi^{(2)}$ and $\varphi^{(1)}$ are conjugacies. Therefore $Y^{(1)}$ is conjugate to $Y^{(2)}$.

Williams' Classification Theorem asserts the determination of conjugacy between, in our cases, $W^{(1)}$ and $W^{(2)}$. It is still hard to find a strong shift equivalence between $T_{G_{Y^{(1)}}}$ and $T_{G_{Y^{(2)}}}$. What we find instead is a weaker relation called *shift equivalence*.

Definition 3.25. Let A and B be nonnegative integral matrices. A shift equivalence from A to B is a pair (R, S) of rectangular nonnegative integral matrices satisfying

$$AR = RB, \qquad SA = BS, \qquad A^\ell = RS, \qquad B^\ell = SR$$

for some $\ell \in \mathbb{N}$. In this case we say that A is *shift equivalent* to B and write $A \sim_{\mathcal{F}_S} B$.

It follows directly that $A \sim_{\mathcal{F}_{SS}} B$ implies $A \sim_{\mathcal{F}_S} B$. A necessary and sufficient condition for the shift equivalence between A and B is the isomorphism between dimension triples of A and B. Suppose A is an $n \times n$ nonnegative integral matrix. The *eventual range* $\mathcal{R}_A$ of A is defined by

$$\mathcal{R}_A = \bigcap_{k=1}^{\infty} \mathbb{Q}^n A^k,$$

where $\mathbb{Q}^n$ is the n-dimensional rational space.

Definition 3.26. Let A be an $n \times n$ nonnegative integral matrix. The dimension group of A is

$$\Delta_A = \left\{ \mathbf{v} \in \mathcal{R}_A \colon \mathbf{v} A^k \in \mathbb{Z}^n \text{ for some } k \geqslant 0 \right\}.$$

The dimension group automorphism δ_A of A is the restriction of A to Δ_A such that $\delta_A(\mathbf{v}) = \mathbf{v} A$ for $\mathbf{v} \in \Delta_A$. We call (Δ_A, δ_A) the dimension pair of A. Moreover, we define the dimension semigroup of A to be

$$\Delta_A^+ = \left\{ \mathbf{v} \in \mathcal{R}_A \colon \mathbf{v} A^k \in \left(\mathbb{Z}^+\right)^n \text{ for some } k \geqslant 0 \right\}.$$

We call $(\Delta_A, \Delta_A^+, \delta_A)$ the dimension triple of A.

Theorem 3.27. *(See [31, Theorem 7.5.8].) Let A and B be nonnegative integral matrices. $A \sim_{\mathcal{F}_S} B$ if and only if $(\Delta_A, \Delta_A^+, \delta_A)$ is group isomorphic to $(\Delta_B, \Delta_B^+, \delta_B)$. That is, there exists a group isomorphism $\theta : \Delta_A \to \Delta_B$ satisfying $\theta \circ \delta_A = \delta_B \circ \theta$ and $\theta(\Delta_A^+) = \Delta_B^+$.*

Instead of demonstrating a strong shift equivalence between two matrices, it is much easier to determine whether their dimension groups are isomorphic to one another. Since T' and H in Example 3.16 are strong shift equivalent (cf. Example 3.24), they must admit isomorphic dimension groups. The following example gives an examination.

Example 3.28. Continued from Example 3.16. Examine that the eigenvalues of T' are $\frac{1 \pm \sqrt{5}}{2}, 0$ and its eventual range $\mathcal{R}_{T'} = \mathbb{Q}(1\,1\,1) \oplus \mathbb{Q}(0\,1\,1)$, and the eigenvalues of H are $\frac{1 \pm \sqrt{5}}{2}$ and its eventual range $\mathcal{R}_H = \mathbb{Q}^2$. The dimension groups of T' and H are

$$\Delta_{T'} = \{t_1 u_1 + t_2 u_2 : t_1, t_2 \in \mathbb{Z}\} \quad \text{and} \quad \Delta_H = \{s_1 v_1 + s_2 v_2 : s_1, s_2 \in \mathbb{Z}\},$$

where $u_1 = (1\,1\,1), u_2 = (0\,1\,1), v_1 = (0\ 1)$ and $v_2 = (1\,0)$. Moreover,

$$\Delta_{T'}^+ = \{t_1 u_1 + t_2 u_2 : t_1 \geqslant 0,\ t_1 + t_2 \geqslant 0\}$$

and

$$\Delta_H^+ = \{s_1 v_1 + s_2 v_2 : s_1 \geqslant 0,\ s_1 + s_2 \geqslant 0\}.$$

Define $\theta : \Delta_{T'} \to \Delta_H$ as $\theta(t_1 u_1 + t_2 u_2) = t_1 v_1 + t_2 v_2$. It follows θ is an isomorphism and $\theta \circ \delta_{T'} = \delta_H \circ \theta$, $\theta(\Delta_{T'}^+) = \Delta_H^+$. Theorem 3.27 asserts $T' \sim_{\mathcal{F}_S} H$ which is a necessary condition for $T' \sim_{\mathcal{F}_{SS}} H$.

For any two nonnegative integral matrices A, B, it is easier to find their Jordan forms $J(A), J(B)$ than to demonstrate the group isomorphism between their dimension triples. Although $J(A) = J(B)$ is not equivalent to $A \sim_{\mathcal{F}_S} B$, it still provides a necessary condition.

Definition 3.29. Let A be an $n \times n$ integral matrix. The invertible part $A^\times$ of A is the linear transformation obtained by restricting A to its eventual range. That is, $A^\times : \mathcal{R}_A \to \mathcal{R}_A$ is defined by $A^\times(\mathbf{v}) = \mathbf{v}A$.

Theorem 3.30. *(See [31, Theorem 7.4.10].) Suppose A and B are nonnegative integral matrices. If $A \sim_{\mathcal{F}_S} B$, then $J^\times(A) = J^\times(B)$, where $J^\times(M)$ is the Jordan form of $M^\times$.*

Example 3.31. Consider (11) with $d = 1$. Suppose the parameters are set up as

$$\left(a_{-1}^{(2)}, a_0^{(2)}, a_1^{(2)}\right) = (5.1, 8.4, 1), \qquad \left(b_{-1}^{(2)}, b_0^{(2)}, b_1^{(2)}, z^{(2)}\right) = (0, 3, 3.5, -1.4),$$

and

$$\left(a_{-1}^{(1)}, a_0^{(1)}, a_1^{(1)}, z^{(1)}\right) = (0, 1.7, 0.3), \qquad \left(b_{-1}^{(1)}, b_0^{(1)}, b_1^{(1)}, z^{(1)}\right) = (0, 0, 0, 0, 2).$$

Then the basic set of admissible local patterns are the following:
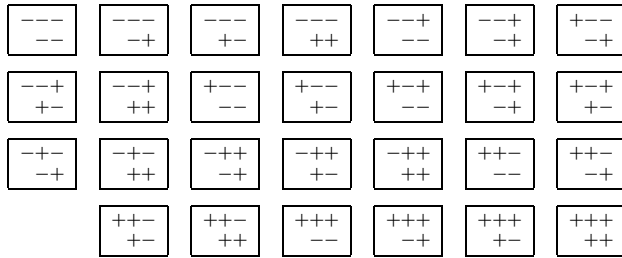
The transition matrix of the underlying graph G_T is

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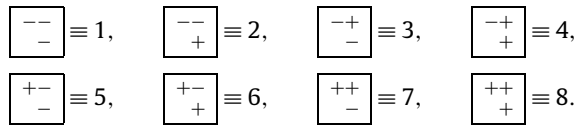0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Obviously $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are not right-resolving. For simplicity, we number the vertices of G_T by

$$\boxed{\begin{smallmatrix} - & - \\ & - \end{smallmatrix}} \equiv 1, \quad \boxed{\begin{smallmatrix} - & - \\ & + \end{smallmatrix}} \equiv 2, \quad \boxed{\begin{smallmatrix} - & + \\ & - \end{smallmatrix}} \equiv 3, \quad \boxed{\begin{smallmatrix} - & + \\ & + \end{smallmatrix}} \equiv 4,$$

$$\boxed{\begin{smallmatrix} + & - \\ & - \end{smallmatrix}} \equiv 5, \quad \boxed{\begin{smallmatrix} + & - \\ & + \end{smallmatrix}} \equiv 6, \quad \boxed{\begin{smallmatrix} + & + \\ & - \end{smallmatrix}} \equiv 7, \quad \boxed{\begin{smallmatrix} + & + \\ & + \end{smallmatrix}} \equiv 8.$$

Applying subset construction to $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$, the transition matrix $T_{G_{Y(2)}}$, indexed by

$$\{1, 3, 6, 8, \boxed{1, 2}, \boxed{3, 4}, \boxed{5, 6}, \boxed{7, 8}\} \equiv \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\},$$

is

$$T_{G_{Y(2)}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Similarly, the transition matrix $T_{G_{Y(1)}}$, indexed by

$$\{\boxed{1, 3}, \boxed{1, 3, 5, 7}, \boxed{2, 4}, \boxed{2, 4, 6, 8}\} \equiv \{y_1, y_2, y_3, y_4\},$$

is

$$T_{G_{Y(1)}} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

Since the Jordan form of $T_{G_{Y^{(1)}}}, T_{G_{Y^{(2)}}}$ are

$$J(T_{G_{Y^{(1)}}}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad J(T_{G_{Y^{(2)}}}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

respectively, we have

$$J^{\times}(T_{G_{Y^{(1)}}}) = (2) \neq \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = J^{\times}(T_{G_{Y^{(2)}}}).$$

Theorem 3.30 asserts $Y^{(1)}$ and $Y^{(2)}$ are not shift equivalent, but rather finitely equivalent. Let

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

Then $T_{G_{Y^{(2)}}} F = F T_{G_{Y^{(1)}}}$ and the graph M is constructed as in Fig. 6.

Remark 3.32. Suppose $h(Y^{(1)}) = h(Y^{(2)})$. Let $W^{(i)}$ be a shift of finite type with finite-to-one factor map $\varphi^{(i)} : W^{(i)} \to Y^{(i)}$ for $i = 1, 2$. We classify $W^{(1)}$ and $W^{(2)}$ into three classes: $W^{(1)} \sim_{\mathcal{F}_{SS}} W^{(2)}$, $W^{(1)} \sim_{\mathcal{F}_S} W^{(2)}$, and $W^{(1)} \sim_{\mathcal{F}} W^{(2)}$. Moreover,

$$W^{(1)} \sim_{\mathcal{F}_{SS}} W^{(2)} \quad \Rightarrow \quad W^{(1)} \sim_{\mathcal{F}_S} W^{(2)} \quad \Rightarrow \quad W^{(1)} \sim_{\mathcal{F}} W^{(2)}.$$

See Fig. 7.

3.4. Existence of diamonds

For a two-layer cellular neural network, we classify the hidden space and the output space into three classes whenever $h(Y^{(1)}) = h(Y^{(2)})$. However, numerical results show that mostly $h(Y^{(1)}) \neq h(Y^{(2)})$ [37]. Suppose $h(Y^{(1)}) \neq h(Y^{(2)})$, then either $\phi^{(1)}$ or $\phi^{(2)}$ has a *diamond*.

Definition 3.33. Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph. A *graph diamond* for $\mathcal{L}$ is a pair of distinct paths in G having the same $\mathcal{L}$-label, the same initial state, and the same terminal state.

Set $\phi = \mathcal{L}_\infty : X_G \to X_G$. A diamond for ϕ is a pair of distinct points in X_G differing in only finitely many coordinates with the same image under ϕ.

It is easily seen that if G is essential, then ϕ has a diamond if and only if $\mathcal{L}$ has a graph diamond. Moreover, the transition matrix gives a sufficient criterion for the existence of diamond.
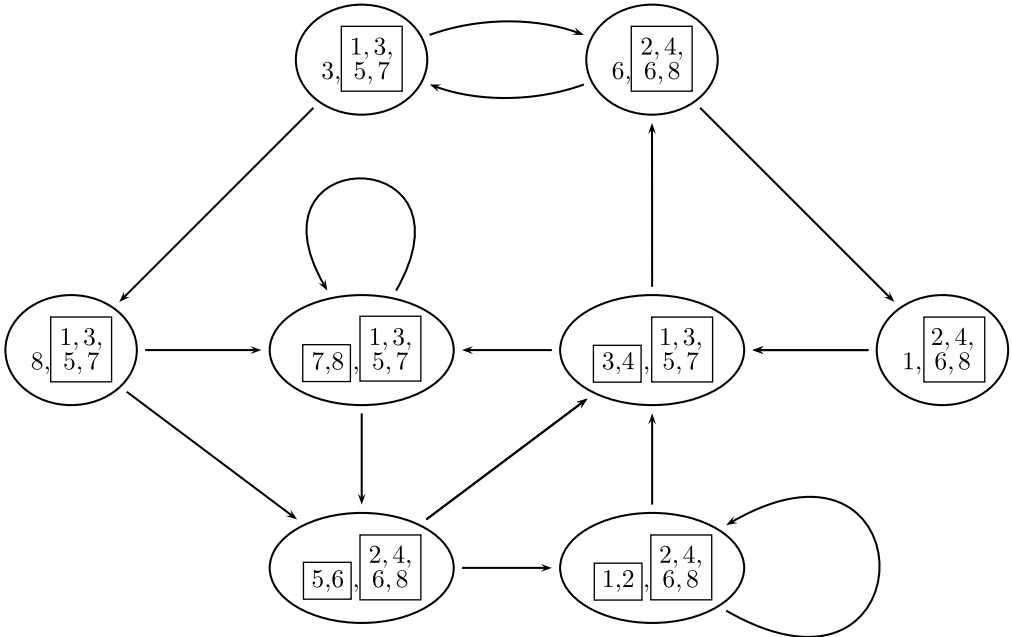
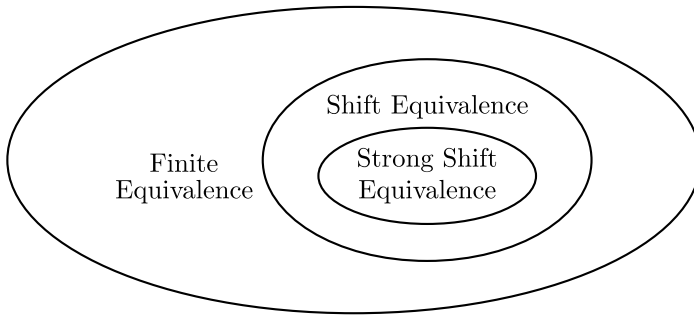**Fig. 6.** The graph representation of the common extension of $Y^{(1)}$ and $Y^{(2)}$ in Example 3.31.

Fig. 7. The relation between three classification.

Proposition 3.34. *If $\phi^{(k)}$ has a diamond, $k = 1, 2$, then there exists $n \in \mathbb{N}$ such that $(T^{(k)})^n(p, p) > 2$ for some p.*

Proof. Since G_T is essential, $\phi^{(k)}$ has a diamond indicates that there are two paths ω, τ in G_T with the same initial and terminate vertices such that $\mathcal{L}^{(k)}(\omega) = \mathcal{L}^{(k)}(\tau)$. The irreducibility of G_T asserts a path υ in G with $i(\upsilon) = t(\omega)$ and $t(\upsilon) = i(\omega)$. Moreover, $\omega\upsilon, \tau\upsilon$ are two loops in G implies $(T^{(k)})^{|\omega\upsilon|}(1 + i(\omega), 1 + i(\omega)) \geqslant 2$. This completes the proof. $\quad\square$

Example 3.35. Suppose $\mathcal{A} = \{0, 1\}$. Let $X \subset \mathcal{A}^{\mathbb{Z}}$ be the set of binary sequences so that between any two 1's there are even number of 0's. A labeled graph representation $\mathcal{G}$ for X as follows:
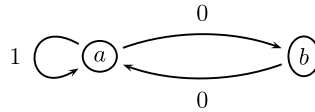
The transition matrix for X is

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

It is seen that $T^3(1,1) = 3$. Nevertheless, $\mathcal{G}$ has no graph diamond since it is right-resolving.

If either $\phi^{(1)}$ or $\phi^{(2)}$ has a diamond, those methods we use to determine whether there is a relation between $Y^{(1)}$ and $Y^{(2)}$ fail. A separate investigation to investigate such cases is in the preparation stages.

4. Decoupling the solution spaces of simplified multi-layer cellular neural networks

This section extends the results in the previous section to simplified n-layer cellular neural networks for $n \geqslant 2$. We consider the simplified n-layer cellular neural networks as follows:

$$\begin{cases} \dfrac{dx_i^{(1)}}{dt} = -x_i^{(1)} + a^{(1)} y_i^{(1)} + a_r^{(1)} y_{i+1}^{(1)} + z^{(1)}, \\ \dfrac{dx_i^{(2)}}{dt} = -x_i^{(2)} + a^{(2)} y_i^{(2)} + a_r^{(2)} y_{i+1}^{(2)} + b^{(2)} u_i^{(2)} + b_r^{(2)} u_{i+1}^{(2)} + z^{(2)}, \\ \quad \vdots \\ \dfrac{dx_i^{(n)}}{dt} = -x_i^{(n)} + a^{(n)} y_i^{(n)} + a_r^{(n)} y_{i+1}^{(n)} + b^{(n)} u_i^{(n)} + b_r^{(n)} u_{i+1}^{(n)} + z^{(n)}, \end{cases}$$

where $u_i^{(k)} = y_i^{(k-1)}$ for $i \in \mathbb{Z}$ and $2 \leqslant k \leqslant n$. In this case, the ordering matrix $\mathbb{X}_n$ is indexed by $\{- \diamond - \diamond \cdots \diamond -, - \diamond - \diamond \cdots \diamond +, \ldots, + \diamond + \diamond \cdots \diamond +\}$. Suppose $\mathcal{B} \subseteq \{-,+\}^{\mathbb{Z}_{2 \times n}}$ is a basic set of admissible local patterns determined by a partition of parameter space and $\mathbf{Y} \subseteq \{-,+\}^{\mathbb{Z}_{\infty \times n}}$ presents the solution space. The system then induces a output space $Y^{(n)}$ and $n-1$ hidden spaces $Y^{(i)}$ for $i = 1, 2, \ldots, n-1$.

Let $\mathcal{A} = \{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$, where

$$\alpha_0 = --, \qquad \alpha_1 = -+, \qquad \alpha_2 = +-, \quad \text{and} \quad \alpha_3 = ++.$$

For $i = 1, 2, \ldots, n$, define the labeling $\mathcal{L}^{(i)} : \mathcal{B} \to \mathcal{A}$ by

$$\mathcal{L}^{(i)}\big(y^{(n)} y_r^{(n)} \diamond y^{(n-1)} y_r^{(n-1)} \diamond \cdots \diamond y^{(1)} y_r^{(1)}\big) = y^{(i)} y_r^{(i)},$$

where

$$y^{(n)} y_r^{(n)} \diamond y^{(n-1)} y_r^{(n-1)} \diamond \cdots \diamond y^{(1)} y_r^{(1)} \equiv \begin{bmatrix} y^{(n)} y_r^{(n)} \\ y^{(n-1)} y_r^{(n-1)} \\ \vdots \\ y^{(1)} y_r^{(1)} \end{bmatrix}.$$
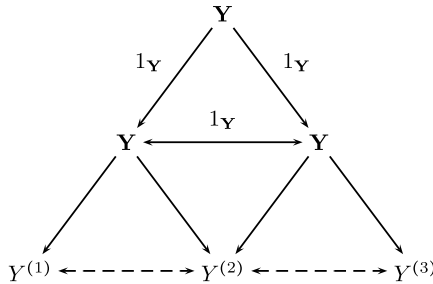
Fig. 8. For three-layer cellular neural networks, suppose that $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}$, and $\mathcal{G}^{(3)}$ are right-resolving labeled graphs. A natural common extension of $Y^{(1)}, Y^{(2)}$, and $Y^{(3)}$ is the original space $\mathbf{Y}$. Notice that herein $1_{\mathbf{Y}}$ means identity map.

Suppose T is the transition matrix induced by $\mathcal{B}$ and G_T is the graph representation of T. Theorem 3.3 indicates $Y^{(i)}$ is a sofic shift with labeled graph representation $\mathcal{G}^{(i)} = (G_T, \mathcal{L}^{(i)})$ and $Y^{(i)} = X_{\mathcal{G}^{(i)}}$ for $i = 1, 2, \ldots, n$.

Section 2 studies, for the case $n = 2$, whether there is a factor map between $Y^{(1)}$ and $Y^{(2)}$. This section considers the following question.

Question 3. Suppose that $1 \leqslant i < j \leqslant n$, does a factor map exist between $Y^{(i)}$ and $Y^{(j)}$?

For the rest of this investigation, we assume that any two spaces that derived from solution space $\mathbf{Y}$ admit the same topological entropy unless otherwise stated.

4.1. Simplified three-layer cellular neural networks

First we consider the case $n = 3$. Let $\mathcal{V} = \{0, 1, \ldots, 7\}$ be the vertex set of G_T, where

$$0 = -\diamond - \diamond -, \qquad 1 = -\diamond - \diamond +, \qquad \ldots, \qquad 6 = +\diamond + \diamond -, \qquad 7 = +\diamond + \diamond +.$$

The labeling $\mathcal{L}^{(1)}, \mathcal{L}^{(2)}, \mathcal{L}^{(3)} : \mathcal{E} \to \mathcal{A}$ can be expressed explicitly by

$$\mathcal{L}^{(1)}(e) = \alpha_{2\tau(i(e)) + \tau(t(e))}, \quad \text{where} \quad \tau(c) := c \mod 2, \tag{23}$$

$$\mathcal{L}^{(2)}(e) = \alpha_{2\tau([i(e)/2]) + \tau([t(e)/2])}, \quad \text{where } [\cdot] \text{ is the Gauss function,} \tag{24}$$

$$\mathcal{L}^{(3)}(e) = \alpha_{2[i(e)/4] + [t(e)/4]}. \tag{25}$$

Suppose $\mathcal{G}^{(i)}$ is right-resolving for all i. The following proposition is derived via an analogous method to that in the proof of Proposition 3.6.

Proposition 4.1. *Suppose that $\mathcal{G}^{(i)}$ is right-resolving for $1 \leqslant i \leqslant 3$. Then $Y^{(1)}, Y^{(2)}$, and $Y^{(3)}$ are finitely equivalent, and $\mathbf{Y}$ is a common extension of $Y^{(1)}, Y^{(2)}$, and $Y^{(3)}$.*

Proof. Since $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}$ are right-resolving, the shift of finite type $\mathbf{Y}$ is a common extension of $Y^{(1)}$ and $Y^{(2)}$. Similarly, $\mathcal{G}^{(2)}, \mathcal{G}^{(3)}$ are right-resolving asserts $\mathbf{Y}$ is also a common extension of $Y^{(2)}$ and $Y^{(3)}$. Therefore, $\mathbf{Y}$ is a natural common extension of $Y^{(1)}, Y^{(2)}$, and $Y^{(3)}$. See Fig. 8. $\square$

Proposition 4.2. *Under the assumption of Proposition 4.1. Suppose $S^{(1)}, S^{(2)}$ and $S^{(3)}$ is the symbolic transition matrix of $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}$ and $\mathcal{G}^{(3)}$, respectively. Let*

$$\mathbb{P}_{3,2} = \left\{ (P \otimes I_2) \operatorname{diag}(C_p)_{1 \leqslant p \leqslant 4} \colon C_p = I_2, J_2, \text{ for all } p \right\},$$

$$\mathbb{P}_{2,1} = \left\{ K^{-1} \big((P \otimes I_2) \operatorname{diag}(C_p)_{1 \leqslant p \leqslant 4} \big) K \colon C_p = I_2, J_2, \text{ for all } p \right\},$$

$$\mathbb{P}_{3,1} = \left\{ L^{-1} \big((P \otimes I_2) \operatorname{diag}(C_p)_{1 \leqslant p \leqslant 4} \big) L \colon C_p = I_2, J_2, \text{ for all } p \right\},$$

where

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad J_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$K = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and $\otimes$ is the Kronecker product. If G is an essential graph, then $Y^{(i)} \cong Y^{(j)}$ if and only if $S^{(j)} P_{j,i} = P_{j,i} S^{(i)}$ for some $P_{j,i} \in \mathbb{P}_{j,i}$, where $1 \leqslant i < j \leqslant 3$.

Proof. We show that $Y^{(2)} \cong Y^{(3)}$ if and only if either $P_{3,2} S^{(2)} = S^{(3)} P_{3,2}$ for some $P_{3,2} \in \mathbb{P}_{3,2}$. The proof of $Y^{(1)} \cong Y^{(2)}$ and $Y^{(1)} \cong Y^{(3)}$ can be done analogously.

Observe that $\mathcal{L}^{(2)}(v, v') = \mathcal{L}^{(3)}(v + 1, v' + 1)$ for $v, v' \in \{0, 2, 4, 6\} \subset \mathcal{V}$. Let $V_k = \{2k, 2k + 1\}$, $k = 0, 1, 2, 3$. Set $\widetilde{G} = (\widetilde{\mathcal{V}}, \widetilde{\mathcal{E}})$ with $\widetilde{\mathcal{V}} = \{V_0, V_1, V_2, V_3\}$ and $(V, V') \in \widetilde{\mathcal{E}}$ if there exist $v \in V$, $v' \in V'$ such that $(v, v') \in \mathcal{E}$. In other words, $\widetilde{G}$ is obtained from G by bundling vertices carrying the same labels under $\mathcal{L}^{(2)}$ and $\mathcal{L}^{(3)}$. Theorem 3.11 demonstrates $X_{\widetilde{G}^{(2)}} \cong X_{\widetilde{G}^{(3)}}$ if and only if $PS_{\widetilde{G}^{(2)}} = S_{\widetilde{G}^{(3)}} P$. It can be seen that $Y^{(2)} \cong Y^{(3)}$ if and only if $X_{\widetilde{G}^{(2)}} \cong X_{\widetilde{G}^{(3)}}$ and there are conjugacies in V_k, $k = 0, 1, 2, 3$. Therefore, $Y^{(2)} \cong Y^{(3)}$ if and only if $P_{3,2} S^{(2)} = S^{(3)} P_{3,2}$ for some $P_{3,2} \in \mathbb{P}_{3,2}$. □

The solution space **Y** is a natural extension of $Y^{(i)}$ provided $\mathcal{G}^{(i)}$ is right-resolving for $i = 1, 2, 3$. If $\mathcal{G}^{(i)}$ is not right-resolving for some i, we construct a real common extension via analogous argument as in the proof of Theorem 3.12.

Theorem 4.3. *Suppose $\mathcal{G}^{(i)}$ is not right-resolving for some i. There exists a common extension shift of finite type W of $Y^{(1)}, Y^{(2)}, Y^{(3)}$ and $\phi_{Y^{(i)}} \colon W \to Y^{(i)}$ such that $\phi_{Y^{(i)}}$ is a finite-to-one factor for $i = 1, 2, 3$.*

Proof. Suppose $\mathcal{G}^{(1)}$ is not right-resolving and $\mathcal{G}^{(2)}, \mathcal{G}^{(3)}$ are right-resolving. Let $(G_{Y^{(1)}}, \mathcal{L}^{(1)})$ be a minimal right-resolving labeled graph obtained by applying subset construction on $(G_T, \mathcal{L}^{(1)})$. Note that we still denote the labeling of $G_{Y^{(1)}}$ by $\mathcal{L}^{(1)}$ for simplicity. The proof of Theorem 3.12 demonstrates that there is a common extension space W of $W^{(1)}$ and $\mathbf{Y}$, and finite-to-one factors $\widetilde{\varphi}_1 \colon W \to W^{(1)}$, $\widetilde{\varphi}_2 \colon W \to \mathbf{Y}$. $(G_{Y^{(1)}}, \mathcal{L}^{(1)})$ and $(G_T, \mathcal{L}^{(2)})$ are right-resolving imply $\phi^{(1)} \colon W^{(1)} \to Y^{(1)}$ and $\phi^{(2)} \colon \mathbf{Y} \to Y^{(2)}$ are both finite-to-one. Hence $\phi^{(i)} \circ \widetilde{\varphi}_i \colon \widetilde{W} \to Y^{(i)}$ is also a finite-to-one factor for $i = 1, 2$. Similarly, we derive $\phi^{(3)} \circ \widetilde{\varphi}_2 \colon \widetilde{W} \to Y^{(3)}$ is a finite-to-one factor map. Therefore the shift of finite type W is a common extension space of $Y^{(1)}, Y^{(2)}$, and $Y^{(3)}$. See Fig. 9.

The other cases can be done by applying the construction in the proof of Theorem 3.12 repeatedly. See Fig. 10. □

The finite equivalence between $Y^{(1)}, Y^{(2)}$ and $Y^{(3)}$ induces a dendrogram as, for example, in Fig. 10. It is natural to ask which dashed lines could be solid lines. As we elucidate in the last section, this
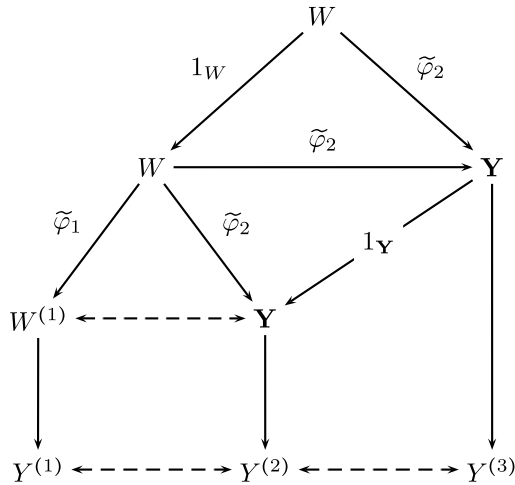
Fig. 9. If $\mathcal{L}^{(1)}$ is not right-resolving, applying subset construction on $(G, \mathcal{L}^{(1)})$ induces a right-resolving labeled graph $(G_{Y^{(1)}}, \mathcal{L}^{(1)})$. There exists a common extension space W of $W^{(1)}$ and $\mathbf{Y}$. Since $\mathcal{G}^{(2)}, \mathcal{G}^{(3)}$ are right-resolving, W is a common extension space of $Y^{(i)}$ for $i = 1, 2, 3$.
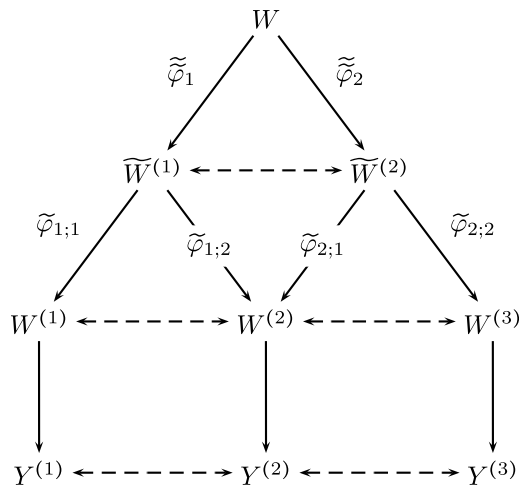

Fig. 10. The common extension space of $Y^{(i)}$ for none of $\mathcal{L}^{(i)}$ is right-resolving.

depends on whether there exists a factor-like matrix which commutes with the transition matrices of these spaces. This derives the following proposition.

Proposition 4.4. *Under the same assumption of Theorem 4.3. Suppose F_{ij} is an integral matrix which satisfies $T_{G_{Y^{(i)}}} F_{ij} = F_{ij} T_{G_{Y^{(j)}}}$ for $1 \leqslant i \neq j \leqslant 3$. If F_{ij} is factor-like for some i, j, then there exists a factor map $\pi_{ij} : W^{(i)} \to W^{(j)}$ that preserves topological entropy.*

Suppose there exists a factor map π_{ij} from $W^{(i)}$ to $W^{(j)}$ for some $i \neq j$. The question then follows of whether there exists a factor map $\bar{\pi}_{ij}$ from $Y^{(i)}$ to $Y^{(j)}$. Similar to the discussion in the last section, $\bar{\pi}_{ij}$ exists if $\varphi^{(i)} : W^{(i)} \to Y^{(i)}$ has an inverse. Theorem 2.8 demonstrates that a necessary and sufficient

condition for the existence of $(\varphi^{(i)})^{-1}$ is that $Y^{(i)}$ is a shift of finite type. Theorem 3.19 asserts that an equivalent condition is to verify synchronizing words. This indicates the following proposition.

Proposition 4.5. *Under the same assumption of Theorem 4.3 and Proposition 4.4. Adding that there exist $N_i \in \mathbb{N}$ such that all words of $Y^{(i)}$ of length N_i are synchronizing for $\mathcal{G}_{Y^{(i)}}$. Then there exists a factor map from $Y^{(i)}$ to $Y^{(j)}$ which preserves topological entropy.*

The previous section classifies hidden and output spaces into three types: strong shift equivalence, shift equivalence, and finite equivalence. The same classification can be applied here to classify $W^{(1)}$, $W^{(2)}$, and $W^{(3)}$. This leads to the following theorem.

Theorem 4.6. *Suppose $Y^{(1)}$, $Y^{(2)}$, and $Y^{(3)}$ are derived from a simplified three-layer cellular neural networks with labeled graph representation $\mathcal{G}^{(1)}$, $\mathcal{G}^{(2)}$, and $\mathcal{G}^{(3)}$. Then*

(1) *If $\mathcal{G}^{(i)}$ is right-resolving and $Y^{(i)}$ is a shift of finite type for $i = 1, 2, 3$, then there exist factor maps $\bar{\pi}_{ij} : Y^{(i)} \to Y^{(j)}$ for $1 \leqslant i, j \leqslant 3$.*

(2) *If $\mathcal{G}^{(i)}$ is not right-resolving for some i, let $\mathcal{G}_{Y^{(i)}}$ be a right-resolving labeled graph representation of $Y^{(i)}$. Suppose $Y^{(i)}$ is a shift of finite type for $i = 1, 2, 3$.*

    (i) *If $T_{G_{Y^{(i)}}} \sim_{\mathcal{F}_{SS}} T_{G_{Y^{(j)}}}$ for some i, j, then $Y^{(i)} \cong Y^{(j)}$.*

    (ii) *If $T_{G_{Y^{(i)}}} \sim_{\mathcal{F}_S} T_{G_{Y^{(j)}}}$ for some i, j, then there exists a factor map between $Y^{(i)}$ and $Y^{(j)}$ provided there is a factor-like matrix F which commutes with $T_{G_{Y^{(i)}}}$ and $T_{G_{Y^{(j)}}}$.*

    (iii) *Otherwise, $Y^{(i)}$ is strictly finitely equivalent to $Y^{(j)}$.*

4.2. Simplified multi-layer cellular neural networks

In general, consider a simplified n-layer cellular neural network for $n \geqslant 3$. The solution space **Y** derives n sofic shifts $Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)}$ with labeled graph representation $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(n)}$.

Proposition 4.7. *Suppose that $\mathcal{G}^{(i)}$ is right-resolving for $1 \leqslant i \leqslant n$. Then $\phi^{(i)} : \mathbf{Y} \to Y^{(i)}$ is a finite-to-one factor map for each i.*

Suppose that the transition matrix T is indexed by

$$- \diamond - \diamond \cdots \diamond -, \qquad - \diamond - \diamond \cdots \diamond +, \qquad , \ldots, \qquad + \diamond + \diamond \cdots \diamond +.$$

Let $\mathbb{P}_{2,1} = \{P\}$, where P is as defined in Theorem 3.11. For $1 \leqslant i < j \leqslant n$ and $(i, j) \neq (n - 1, n)$, denote by $K_{j,i}$ the permutation matrix that bundles those vertices carrying the same label under $\mathcal{L}^{(i)}$ and $\mathcal{L}^{(j)}$. Take $n = 3$ for instance. Then $K_{2,1} = K$ and $K_{3,1} = L$ are defined in Proposition 4.2.

The following proposition extends Proposition 4.2 to the general case through the application of mathematical induction. For brevity, we omit the proof.

Proposition 4.8. *Under the assumption of Proposition 4.7. Suppose $S^{(i)}$ is the symbolic transition matrix of $(G, \mathcal{L}^{(i)})$ for $i = 1, 2, \ldots, n$. Let*

$$\mathbb{P}_{n,n-1} = \left\{ (P_{n-1,n-2} \otimes I_2) \cdot (C_p)_{1 \leqslant p \leqslant 2^{n-1}} \colon P_{n-1,n-2} \in \mathbb{P}_{n-1,n-2}, \ C_p \in \{I_2, J_2\} \text{ for all } p \right\} \qquad (26)$$

and

$$\mathbb{P}_{j,i} = \left\{ K_{j,i}^{-1} P_{n,n-1} K_{j,i} \colon P_{n,n-1} \in \mathbb{P}_{n,n-1} \right\}, \quad 1 \leqslant i < j \leqslant n, \ (i, j) \neq (n - 1, n). \qquad (27)$$

Then $Y^{(i)} \cong Y^{(j)}$ if and only if $S^{(j)} P_{j,i} = P_{j,i} S^{(i)}$ for some $P_{j,i} \in \mathbb{P}_{j,i}$, where $1 \leqslant i < j \leqslant n$.
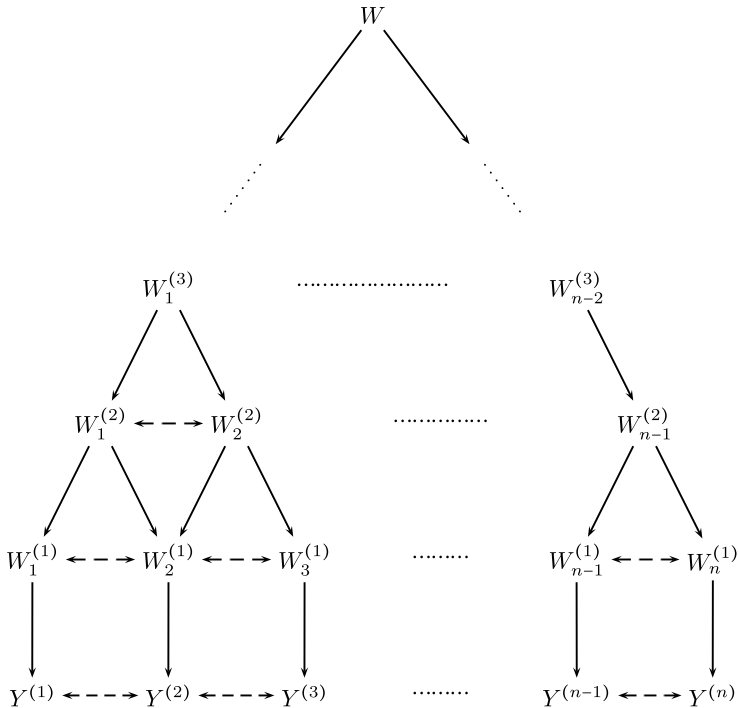
Fig. 11. A common extension space of $Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)}$ can be obtained iteratively.

Suppose $\mathcal{G}^{(i)}$ is not right-resolving for some i. After applying subset construction we still have the following theorem.

Theorem 4.9. *Suppose $\mathcal{G}^{(i)}$ is not right-resolving for some i. There exists a common extension shift of finite type W of $Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)}$ and $\widetilde{\phi}_{Y^{(i)}} : W \to Y^{(i)}$ such that $\widetilde{\phi}_{Y^{(i)}}$ is a finite-to-one factor for each i.*

Proof. Let $(G_{Y^{(i)}}, \mathcal{L}^{(i)})$ be the minimal right-resolving graph of $(G_T, \mathcal{L}^{(i)})$ for $i = 1, 2, \ldots, n$. Note that $G_{Y^{(i)}} = G_T$ for those i's such that $(G, \mathcal{L}^{(i)})$ is itself a right-resolving graph. Following the steps in the proof of Theorem 3.12 we construct a common extension space $W_i^{(1)}$ of $Y^{(i)}$ and $Y^{(i+1)}$ for $i = 1, 2, \ldots, n-1$. The proof of Theorem 4.3 derive a common extension space $W_i^{(2)}$ of $W_i^{(1)}$ and $W_{i+1}^{(1)}$ for $i = 1, 2, \ldots, n-2$. Moreover, $W_i^{(2)}$ is a common extension space of $Y^{(i)}, Y^{(i+1)}$ and $Y^{(i+2)}$ for $i = 1, 2, \ldots, n-2$. Repeating the same process we get a common extension space W of $Y^{(1)}, Y^{(2)}, \ldots, Y^{(k)}$. See Fig. 11. $\quad\square$

To elucidate the existence of a map between $Y^{(i)}$ and $Y^{(j)}$ and the relation between $W^{(i)}$ and $W^{(j)}$ for some i, j, we can follow the flow chart described in Theorem 4.6. This informs whether $Y^{(i)}$ can connect with $Y^{(j)}$. The next section extends the study to general multi-layer cellular neural networks.

5. Structures in multi-layer cellular neural networks

The propositions and theorems we obtain in previous sections still hold for general multi-layer neural networks. In this section, we only state the results instead of giving their explicit proofs.

5.1. Two-layer cellular neural networks

First we consider a two-layer cellular neural network whose basic set of admissible local patterns are of size $k \times 2$, $k \geqslant 3$. For simplicity, we assume that each parameter is nonzero.

Suppose that $k = 3$, we embed our problem in the case that $k = 4$. In other words, suppose $\mathcal{B} \subseteq \{-, +\}^{\mathbb{Z}_{3 \times 2}}$. Let

$$\overline{\mathcal{B}} = \left\{ \begin{array}{|cccc|} y_0^{(2)} & y_1^{(2)} & y_2^{(2)} & y_3^{(2)} \\ y_0^{(1)} & y_1^{(1)} & y_2^{(1)} & y_3^{(1)} \end{array} : \begin{array}{|ccc|} y_0^{(2)} & y_1^{(2)} & y_2^{(2)} \\ y_0^{(1)} & y_1^{(1)} & y_2^{(1)} \end{array}, \begin{array}{|ccc|} y_1^{(2)} & y_2^{(2)} & y_3^{(2)} \\ y_1^{(1)} & y_2^{(1)} & y_3^{(1)} \end{array} \in \mathcal{B} \right\}.$$

Set up the order for each pattern $\varrho : \{-, +\}^{\mathbb{Z}_{2 \times 2}} \to \{1, 2, \ldots, 16\}$ by

$$\varrho(y_1 y_2 \diamond u_1 u_2) = 8\chi(y_1) + 4\chi(y_2) + 2\chi(u_1) + \chi(u_2) + 1,$$

where $\chi : \{-, +\} \to \{0, 1\}$ is given by $\chi(-) = 0$ and $\chi(+) = 1$. This defines an ordering matrix whose elements consist of 4×2 patterns.

Let $Y^{(1)}$ and $Y^{(2)}$ be the output space and hidden space extracted from the solution space $\mathbf{Y}$. Define two labeling

$$\mathcal{L}^{(1)}(y_0 y_1 y_2 y_3 \diamond u_0 u_1 u_2 u_3) = u_0 u_1 u_2 u_3,$$

$$\mathcal{L}^{(2)}(y_0 y_1 y_2 y_3 \diamond u_0 u_1 u_2 u_3) = y_0 y_1 y_2 y_3.$$

Then $\mathcal{G}^{(i)} = (G_T, \mathcal{L}^{(i)})$ is the labeled graph representation of $Y^{(i)}$ for $i = 1, 2$.

Theorem 5.1. Suppose $Y^{(1)}$ is the hidden space and $Y^{(2)}$ is the output space. Then $\mathcal{G}^{(i)}$ is the labeled graph representation of $Y^{(i)}$ and $Y^{(i)} = X_{\mathcal{G}^{(i)}}$ for $i = 1, 2$.

Define $\eta : \{-, +\}^{\mathbb{Z}_{2 \times 2}} \to \{-, +\}^{\mathbb{Z}_{2 \times 2}}$ and $P_4 \in \mathbb{R}^{16 \times 16}$ by

$$\eta(y_1 y_2 \diamond u_1 u_2) = u_1 u_2 \diamond y_1 y_2,$$

$$P_4(i, j) = \begin{cases} 1, & \chi(y_1 y_2 \diamond u_1 u_2) = i \text{ and } \chi \circ \eta(y_1 y_2 \diamond u_1 u_2) = j; \\ 0, & \text{otherwise.} \end{cases}$$

Let $S^{(1)}$ and $S^{(2)}$ be the symbolic transition matrices of $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$, respectively. The following theorems are derived via an analogous method to the proof of the theorems in Section 3.3, and thus we omit the proof.

Theorem 5.2. If $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are both right-resolving, then $Y^{(1)}$ and $Y^{(2)}$ are finitely equivalent. Moreover, $Y^{(1)} \cong Y^{(2)}$ if and only if $S^{(1)} P_4 = P_4 S^{(2)}$.

Theorem 5.3. Suppose either $\mathcal{G}^{(1)}$ or $\mathcal{G}^{(2)}$ is not right-resolving. Let $(G_{Y^{(1)}}, \mathcal{L}^{(1)})$, $(G_{Y^{(2)}}, \mathcal{L}^{(2)})$ be the minimal right-resolving graph representation of $Y^{(1)}$ and $Y^{(2)}$, respectively.

(a) There exists finite equivalence $(W, \widetilde{\phi}_{Y^{(1)}}, \widetilde{\phi}_{Y^{(2)}})$ between $Y^{(1)}$ and $Y^{(2)}$.
(b) If there exist E, F such that $T_{G_{Y^{(1)}}} = EF$, $T_{G_{Y^{(2)}}} = FE$, where $T_{G_{Y^{(1)}}}, T_{G_{Y^{(2)}}}$ is the transition matrices of $G_{Y^{(1)}}$ and $G_{Y^{(2)}}$, respectively. Then $W^{(1)} \equiv X_{G_{Y^{(1)}}}$ is conjugate to $W^{(2)} \equiv X_{G_{Y^{(2)}}}$.
(c) $T_{G_{Y^{(1)}}} \sim_{\mathcal{F}_S} T_{G_{Y^{(2)}}}$ if and only if $(\Delta_{T_{G_{Y^{(1)}}}}, \Delta^+_{T_{G_{Y^{(1)}}}}, \delta_{T_{G_{Y^{(1)}}}})$ is isomorphic to $(\Delta_{T_{G_{Y^{(2)}}}}, \Delta^+_{T_{G_{Y^{(2)}}}}, \delta_{T_{G_{Y^{(2)}}}})$.
(d) If $h(Y^{(1)}) = h(Y^{(2)})$, then there exists integral matrices E, F such that $T_{G_{Y^{(1)}}} E = E T_{G_{Y^{(2)}}}$ and $T_{G_{Y^{(2)}}} F = F T_{G_{Y^{(1)}}}$.

(e) *If E is factor-like, then there exists a factor map $\pi_{12} : W^{(1)} \to W^{(2)}$, where and similarly, if F is factor-like, there exists a factor map $\pi_{21} : W^{(2)} \to W^{(1)}$.*

(f) *Suppose, for instance, there exists a factor map $\pi_{12} : W^{(1)} \to W^{(2)}$. If $Y^{(1)}$ is a shift of finite type, then there exists $\bar{\pi}_{12} : Y^{(1)} \to Y^{(2)}$.*

Suppose $k \geqslant 3$. Without loss of generality, we assume that $k = 2\ell$ for some $\ell \in \mathbb{N}$. If k is odd, then we extend the size of the basic pattern to $k + 1$ as above. Hence the basic set of admissible local patterns consists of patterns of size $2\ell \times 2$. Set up the order for each pattern by

$$\varrho(y_1 y_2 \cdots y_{2\ell} \diamond u_1 u_2 \cdots u_{2\ell}) = 1 + \sum_{i=1}^{2\ell} \left(2^{2\ell-i}\chi(u_i) + 2^{4\ell-i}\chi(y_i)\right).$$

Similar as above, let $Y^{(1)}$ and $Y^{(2)}$ be the output space and hidden space extracted from the solution space **Y**. Define two labeling

$$\mathcal{L}^{(1)}(y_0 y_1 \cdots y_{k-1} \diamond u_0 u_1 \cdots u_{k-1}) = u_0 u_1 \cdots u_{k-1},$$
$$\mathcal{L}^{(2)}(y_0 y_1 \cdots y_{k-1} \diamond u_0 u_1 \cdots u_{k-1}) = y_0 y_1 \cdots y_{k-1}.$$

Then $\mathcal{G}^{(i)} = (G_T, \mathcal{L}^{(i)})$ is the labeled graph representation of $Y^{(i)}$ and $Y^{(i)} = X_{\mathcal{G}^{(i)}}$ for $i = 1, 2$. Define $\bar{\eta} : \{-, +\}^{\mathbb{Z}_{k \times 2}} \to \{-, +\}^{\mathbb{Z}_{k \times 2}}$ and $P_k \in \mathbb{R}^{2^{2k} \times 2^{2k}}$ by

$$\bar{\eta}(y_0 y_1 \cdots y_{k-1} \diamond u_0 u_1 \cdots u_{k-1}) = u_0 u_1 \cdots u_{k-1} \diamond y_0 y_1 \cdots y_{k-1},$$
$$P_k(i, j) = \begin{cases} 1, & \varrho(y_0 y_1 \cdots y_{k-1} \diamond u_0 u_1 \cdots u_{k-1}) = i \text{ and} \\ & \varrho \circ \bar{\eta}(y_0 y_1 \cdots y_{k-1} \diamond u_0 u_1 \cdots u_{k-1}) = j; \\ 0, & \text{otherwise.} \end{cases}$$

Let $S^{(1)}$ and $S^{(2)}$ be the symbolic transition matrices of $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$, respectively. We then have the following theorem.

Theorem 5.4. *If $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$ are both right-resolving, then $Y^{(1)}$ and $Y^{(2)}$ are finitely equivalent. Moreover, $Y^{(1)} \cong Y^{(2)}$ if and only if $S^{(1)} P_k = P_k S^{(2)}$.*

The classification of $Y^{(1)}$ and $Y^{(2)}$ is similar as that in Theorem 5.3. We omit therefore the description.

5.2. Multi-layer cellular neural networks

This section considers multi-layer cellular neural networks whose basic patterns are of size $k \times n$. The foregoing elucidation infers that, without loss of generality, we may assume that $k = 2\ell$ for some $\ell \in \mathbb{N}$. Suppose the basic patterns are ordered by $\chi : \{-, +\}^{\mathbb{Z}_{2\ell \times n}} \to \{1, 2, \ldots, 2^{2n\ell}\}$ and P_k is defined analogously as above. The solution space **Y** induces $Y^{(i)}$ for $i = 1, 2, \ldots, n$. Similarly, we have labeling $\mathcal{L}^{(i)}$ for $i = 1, 2, \ldots, n$. This leads us to the following theorem.

Proposition 5.5. *Suppose $S^{(i)}$ is the symbolic transition matrix of $\mathcal{G}^{(i)} = (G, \mathcal{L}^{(i)})$ and $\mathcal{G}^{(i)}$ is right-resolving for $i = 1, 2, \ldots, n$. Let*

$$\mathbb{P}_{n,n-1;k} = \big\{(P_{n-1,n-2} \otimes I_2) \cdot (C_p)_{1 \leqslant p \leqslant 2^{n-1}} : P_{n-1,n-2} \in \mathbb{P}_{n-1,n-2;k}, $$
$$C_p \text{ is an } \ell \times \ell \text{ permutation matrix}\big\}$$

and

$$\mathbb{P}_{j,i;k} = \left\{ K_{j,i;k}^{-1} P_{n,n-1;k} K_{j,i;k} \colon P_{n,n-1;k} \in \mathbb{P}_{n,n-1;k} \right\}, \quad 1 \leqslant i < j \leqslant n, (i,j) \neq (n-1,n),$$

where $K_{j,i;k}$ is the permutation which bundles those vertices carrying the same label under $\mathcal{L}^{(i)}$ and $\mathcal{L}^{(j)}$ and $\mathbb{P}_{2,1;k} = \{P_k\}$. Then $Y^{(i)} \cong Y^{(j)}$ if and only if $S^{(j)} P_{j,i;k} = P_{j,i;k} S^{(i)}$ for some $P_{j,i;k} \in \mathbb{P}_{j,i;k}$, where $1 \leqslant i < j \leqslant n$.

To discuss the relation between $Y(i)$ and $Y^{(j)}$ for $i \neq j$, we follow the flow chart as in Theorem 4.6 (see Fig. 11). Then, the main results follow.

Theorem 5.6. *Suppose $Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)}$ are extracted from an n-layer cellular neural network with labeled graph representation $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(n)}$. If $h(Y^{(1)}) = h(Y^{(2)}) = \cdots = h(Y^{(n)})$, then*

1. *If $\mathcal{G}^{(i)}$ is right-resolving and $Y^{(i)}$ is a shift of finite type for all i, then there exist factor maps $\bar{\pi}_{ij} : Y^{(i)} \to Y^{(j)}$ for $1 \leqslant i, j \leqslant n$.*
2. *If $\mathcal{G}^{(i)}$ is not right-resolving for some i, let $\mathcal{G}_{Y^{(i)}}$ be a right-resolving labeled graph representation of $Y^{(i)}$. Suppose $Y^{(i)}$ is a shift of finite type for some i.*
   (i) *If $T_{G_{Y^{(i)}}} \sim_{\mathcal{F}_{SS}} T_{G_{Y^{(j)}}}$ for some i, j, then $Y^{(i)} \cong Y^{(j)}$.*
   (ii) *If $T_{G_{Y^{(i)}}} \sim_{\mathcal{F}_{S}} T_{G_{Y^{(j)}}}$ for some i, j, then there exists a factor map between $Y^{(i)}$ and $Y^{(j)}$ provided there is a factor-like matrix F which commutes with $T_{G_{Y^{(i)}}}$ and $T_{G_{Y^{(j)}}}$.*
   (iii) *Otherwise, $Y^{(i)}$ is strictly finitely equivalent to $Y^{(j)}$.*

6. Discussion

This elucidation investigates the relations between subspaces of the solution space of a multi-layer cellular neural network. A small modification of the above procedure allows for decoupling the solution space of an n-layer cellular neural network into arbitrary k subspaces for $2 \leqslant k \leqslant n$. The existence of a factor map between two subspaces depends on whether there exists a factor map between their covering spaces. Note that a covering space of a sofic shift is a shift of finite type. In other words, to classify the subspaces of a solution space is equivalent to the classification of subshifts of finite type induced by multi-layer cellular neural networks. It is known that shift equivalence cannot conclusively establish the conjugacy between two arbitrary subshifts of finite type [35,36]. We conjecture that shift equivalence implies that two subshifts of finite type induced from a multi-layer cellular neural network are conjugate.

Conjecture 1. *Suppose $\mathbf{Y}$ is the solution space of a multi-layer neural network. Let $Y^{(1)}, Y^{(2)}$ be two subshifts of finite type such that $(\mathbf{Y}, \phi^{(1)}, \phi^{(2)})$ is a finite equivalence between $Y^{(1)}$ and $Y^{(2)}$. Then $Y^{(1)}$ is conjugate to $Y^{(2)}$ if and only if $Y^{(1)}$ and $Y^{(2)}$ are shift equivalent.*

Acknowledgments

References

[1] P. Arena, S. Baglio, L. Fortuna, G. Manganaro, Self-organization in a two-layer CNN, IEEE Trans. Circuits Syst. I. Fundam. Theory Appl. 45 (1998) 157–162.
[2] L.O. Chua, L. Yang, Cellular neural networks: Theory, IEEE Trans. Circuits Syst. 35 (1988) 1257–1272.
[3] L.O. Chua, L. Yang, Cellular neural networks: Applications, IEEE Trans. Circuits Syst. 35 (1988) 1273–1290.
[4] L.O. Chua, T. Roska, Cellular Neural Networks and Visual Computing, Cambridge University Press, 2002.

[5] K.R. Crounse, L.O. Chua, Methods for image processing and pattern formation in cellular neural networks: A tutorial, IEEE Trans. Circuits Syst. 42 (1995) 583–601.

[6] K.R. Crounse, T. Roska, L.O. Chua, Image halftoning with cellular neural networks, IEEE Trans. Circuits Syst. 40 (1993) 267–283.

[7] S. Xavier-de Souza, M. Yalcin, J. Suykens, J. Vandewalle, Toward CNN chip-specific robustness, IEEE Trans. Circuits Syst. I. Regul. Pap. 51 (2004) 892–902.

[8] Z. Yang, Y. Nishio, A. Ushida, Image processing of two-layer CNNs – Applications and their stability, IEICE Trans. Fundam. E85-A (2002) 2052–2060.

[9] G. Li, L. Min, H. Zang, Color edge detections based on cellular neural network, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 18 (2008) 1231–1242.

[10] V. Murugesh, Image processing applications via time-multiplexing cellular neural network simulator with numerical integration algorithms, Int. J. Comput. Math. 87 (2010) 840–848.

[11] J. Peng, D. Zhang, X. Liao, A digital image encryption algorithm based on hyper-chaotic cellular neural network, Fund. Inform. 90 (2009) 269–282.

[12] S.-N. Chow, J. Mallet-Paret, E.V. Vleck, Pattern formation and spatial chaos in spatially discrete evolution equations, Random Comput. Dynam. 4 (1996) 109–178.

[13] S.-N. Chow, W. Shen, Dynamics in a discrete Nagumo equation spatial topological chaos, SIAM J. Appl. Math. 55 (1995) 1764–1781.

[14] J. Juang, S.-S. Lin, Cellular neural networks: Mosaic pattern and spatial chaos, SIAM J. Appl. Math. 60 (2000) 891–915.

[15] S.-S. Lin, C.-W. Shih, Complete stability for standard cellular neural networks, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 9 (1999) 909–918.

[16] Y. Ke, The mosaic patterns of CNN with symmetric feedback template, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 18 (2008) 375–390.

[17] Y.-Q. Ke, C.-F. Miao, Existence analysis of stationary solutions for RTD-based cellular neural networks, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 20 (2010) 2123–2136.

[18] X. Li, Analysis of complete stability for discrete-time cellular neural networks with piecewise linear output functions, Neural Comput. 21 (2009) 1434–1458.

[19] C.-W. Shih, Influence of boundary conditions on pattern formation and spatial chaos in lattice systems, SIAM J. Appl. Math. 61 (2000) 335–368.

[20] M. Itoh, L.O. Chua, Structurally stable two-cell cellular neural networks, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 14 (2004) 2579–2653.

[21] M. Gilli, M. Biey, P. Checco, Equilibrium analysis of cellular neural networks, IEEE Trans. Circuits Syst. I. Regul. Pap. 51 (2004) 903–912.

[22] X.M. Li, L.H. Huang, On the complete stability of cellular neural networks with external inputs and bias, Acta Math. Appl. Sin. 26 (2003) 475–486.

[23] X. Li, Z. Yuan, Existence of periodic solutions and closed invariant curves in a class of discrete-time cellular neural networks, Phys. D 238 (2009) 1658–1667.

[24] T. Nishio, Y. Nishio, Periodic pattern formation and its applications in cellular neural networks, IEEE Trans. Circuits Syst. I. Regul. Pap. 55 (2008) 2736–2742.

[25] N. Takahashi, K. Ishitobi, T. Nishi, Sufficient conditions for one-dimensional cellular neural networks to perform connected component detection, Nonlinear Anal. Real World Appl. 11 (2010) 4202–4213.

[26] X. Vilasís-Cardona, M. Vinyoles-Serra, Two neuron CNNs: Search for limit cycles, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 20 (2010) 1137–1173.

[27] J.-C. Ban, C.-H. Chang, S.-S. Lin, Y.-H. Lin, Spatial complexity in multi-layer cellular neural networks, J. Differential Equations 246 (2009) 552–580.

[28] C.-H. Hsu, J. Juang, S.-S. Lin, W.-W. Lin, Cellular neural networks: Local patterns for general template, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 10 (2000) 1645–1659.

[29] J.-C. Ban, C.-H. Chang, On the monotonicity of entropy for multi-layer cellular neural networks, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 19 (2009) 3657–3670.

[30] B. Kitchens, Symbolic Dynamics. One-Sided, Two-Sided and Countable State Markov Shifts, Springer-Verlag, New York, 1998.

[31] D. Lind, B. Marcus, An Introduction to Symbolic Dynamics and Coding, Cambridge University Press, Cambridge, 1995.

[32] J.-C. Ban, S.-S. Lin, Patterns generation and transition matrices in multi-dimensional lattice models, Discrete Contin. Dyn. Syst. 13 (2005) 637–658.

[33] R.F. Williams, Classification of subshifts of finite type, Ann. of Math. 99 (1974) 380–381 (Erratum).

[34] R.F. Williams, Classification of subshifts of finite type, Ann. of Math. 98 (1973) 120–153.

[35] K.H. Kim, F.W. Roush, The Williams conjecture is false for irreducible subshifts, Ann. of Math. 149 (1999) 545–558.

[36] K.H. Kim, F.W. Roush, Williams's conjecture is false for reducible subshifts, J. Amer. Math. Soc. 5 (1992) 213–215.

[37] J.-C. Ban, C.-H. Chang, S.-S. Lin, Diamond in multi-layer cellular neural networks, submitted for publication, 2011.