國立交通大學

應用數學系

博士論文

細胞神經網絡的空間複雜度

Spatial Complexity in Some Class of Cellular Neural Networks

研究生: 張志鴻

指導教授: 林松山教授

中華民國九十七年六月

細胞神經網絡的空間複雜度

Spatial Complexity in Some Class of Cellular Neural Networks

研究生: 張志鴻　　Student:Chih-Hung Chang

指導教授: 林松山　　Advisor:Song-Sun Lin

國立交通大學
應用數學系
博士論文

A Thesis
Submitted to Department of Applied Mathematics
College of Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy
in
Applied Mathematics

June 2008

Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 七 年 六 月

Dedicated to the Memory of Grandpa,

1923 - 2008

細胞神經網絡的空間複雜度

研究生: 張志鴻　　　　　　　　　　　　　　　　　指導教授: 林松山教授

國立交通大學

應用數學系

摘要

本文旨在研究一維的多層細胞神經網絡所產生的花樣其空間複雜度以及二維非均勻空間細胞神經網路其拓樸熵的稠密性。在多層細胞神經網絡部分, 如果我們在輸出空間給予適當的符號, 輸出空間會等價於一個 sofic 移動空間。我們給出兩個物理上的不變量, 拓樸熵和動態 ζ-函數的公式。這同時給了抽象的 sofic 移動空間一個實際上可以應用的例子。除此之外, 我們發現了當考慮多層細胞神經網絡時, 拓樸熵的對稱性會被破壞掉。這也再一次證明多層細胞神經網絡和單層細胞神經網絡是兩個性質極端不同的系統。更進一步地, 當我們考慮非均勻空間的細胞神經網絡系統, 其拓樸熵會稠密的分佈在 $[0, \log 2]$ 這個封閉區間之中。換句話說, 非均勻空間的細胞神經網絡系統有著極爲豐富的物理現象蘊含在裡面。

Spatial Complexity in Some Class of Cellular Neural Networks

Student:Chih-Hung Chang Advisor:Song-Sun Lin

Department of Applied Mathematics
National Chiao Tung University

Abstract

This dissertation consists two parts. The first part investigates the complexity of the global set of output patterns for one-dimensional multi-layer cellular neural networks with input; the second part focus on the dense entropy of two-dimensional inhomogeneous cellular neural networks with/without input. For the first part, applying labeling to the output space produces a sofic shift space. Two invariants, namely spatial entropy and dynamical zeta function, can be exactly computed by studying the induced sofic shift space. This study gives sofic shift a realization through a realistic model. Furthermore, a new phenomenon, the broken of symmetry of entropy, is discovered in multi-layer cellular neural networks with input. The second part is strongly related to the learning problem (or inverse problem); the necessary and sufficient conditions for the admissibility of local patterns must be characterized. The entropy function is dense in $[0, \log 2]$ with respect to the parameter space and the radius of the interacting cells, indicating that, in some sense, such system exhibits a wide range of phenomena.

誌　　　謝

僅以此論文, 獻給未能來得及分享我的喜悅的祖父大人—張公水發先生—一位對晚輩疼愛有加的慈祥長者; 以及我那即將臨世的孩兒。

求學生涯中 50% 都在新竹的我, 經常被笑說是 "披著台中皮的新竹人"。因緣際會下, 我來到了新竹展開我的大學生活, 沒想到就此和新竹結下了不解之緣。在這十多年的求學過程中, 我經歷了清、交兩校截然不同的校風—清大的理論與交大的應用, 體會到數學這門浩瀚無邊的學問兩種不同的面貌之美; 讓我更加清楚明白的知道自己未來的方向與興趣所在。在順利獲得博士學位的同時, 我要向這一路上協助我的人致上最深的謝意。

首先我要感謝我的指導教授—林松山老師, 在我初來乍到交大時, 不但親切的帶領我認識學校的環境, 還引領我瞭解有關動態系統以及微分方程各個分支的學問。當我遇到困惑難解的瓶頸時, 老師總會適時的伸出援手, 指引我方針; 並且鼓勵我四處參加研討會, 增廣自己的見聞。更進一步的開啓了我跨領域研究的視野, 讓我清楚的看到: 原來數學不是只有理論的推衍, 它的觸角還可以延伸到許多與我們息息相關的切身問題。除了帶領我看到數學豐富的層次之外, 無論我選擇哪一條路, 老師總是給予我無限的支持與祝福, 讓我可以放心大膽地去嘗試不同的事物。嚴謹且一絲不苟的學術態度, 深入淺出的教學方法, 更是我努力學習的圭臬。

另外, 感謝靜宜大學的羅主斌老師。如果當初不是老師慷慨的與我分享下榻的旅館, 我不可能參加歐洲數學協會所舉辦關於心臟生理的 summer school, 在繁忙的會議期間, 還不厭其煩地教導我許多知識。回到台灣之後, 更不在意我那不紮實的 background, 帶領我一步步的進入生物數學的殿堂。舉凡最基礎的生理介紹、數學模型、當今生物數學家們所關心的問題, 一直到從事跨領域研究所需具備的知識與觀念, 老師都不厭其煩的講述, 更犧牲他寶貴的時間與我討論。而老師對學術研究那無止盡的熱情, 不但激勵了我的鬥志, 更在我疲憊想偷懶時, 爲我注入一劑強心針。無時無刻不在提醒著我要更加努力, 才不枉費老師的一番苦心。如果說我

今天得以略窺生物數學的門徑, 無非都是老師的功勞。

家人在背後給我的支持與鼓勵, 更是讓我在這條漫長的求學生涯中, 能夠一路走來、始終如一的重要支柱。不管家裡的經濟狀況有多艱辛, 爸爸和媽媽總不要我代爲分憂解勞, 他們只要我專心於學術研究。三不五時和弟弟的鬥嘴, 讓我忘記了生活中偶發的不愉快與研究工作上所遇到的瓶頸。我至愛的另一半—玫珍, 謝謝你在我背後無怨無悔的付出。因爲你的支持與鼓勵, 我才能毫無後顧之憂的在學業上打拼, 並且對自己的未來充滿了信心; 也謝謝我的岳父、岳母願意將他們的寶貝女兒託付與我。沒有你們大家, 就不可能成就今天的我。

感謝這幾年教導我的老師們: 石至文老師爲我開啓了微分方程的一扇窗; 李明佳老師總能用簡單的一兩句話闡述一個抽象的觀念, 在微分方程專題討論的課堂上, 一步一腳印的帶我們探討各種層面的問題與現象。謝謝靜宜大學的田慧君老師在數值技巧與程式方面的指導, 以及生活上關心的問候。

此外, 我還要謝謝伴我度過這幾年博士生涯的好朋友。班榮超學長, 除了是從大學至今的學長, 也是我研究工作上的合作伙伴。你對自我要求的高標準以及學識的淵博, 是我致力追求的目標; 在尋找問題、挖掘問題、解決問題, 乃至於論文的撰寫, 你都提供我許多寶貴的意見, 無私的交流讓我獲益良多。林吟衡學姊, 我們的師門之花; 紮實的基本功與細心的研究態度, 不啻爲我這粗心的男生樹立了一個研究人員的典範, 更親切的與我分享應對進退中該注意的小細節。謝謝同師門的學弟妹—文貴和倖綺—在研究的過程中不厭其煩的和我討論枝微末節; 同研究室的耀漢與明杰, 與你們的討論讓我收穫良多。盈智, 我從大三開始同居的室友, 也是研究路上的好伙伴, 雖然博士班後專攻的領域不同, 在心得交換中發現的共通之處讓我們更確切的知道, 數學果然是一個大家庭。怡菁、小巴、恭儉、康伶、玉雯、佳玲· · ·, 謝謝你們的陪伴, 豐富了我的博士班生活。

最後, 我想引用證嚴上人的一句話—也是我的座右銘—作爲結尾:

做就對了。

張志鴻　2008/6/16　於交大

Contents

1 Introduction

This dissertation includes two investigations. First we study the spatial complexity in multi-layer cellular neural networks, and what comes next is the elucidation of the dense property of topological entropy in inhomogeneous cellular neural networks.

The cellular neural network (CNN) proposed by Chua and Yang is a large aggregate of analogue circuits [12; 13]. The system presents itself as an array of identical cells which are all locally coupled. Many such systems have been studied as models for spatial pattern formation in biology [16; 17; 18; 26; 27], chemistry [19], physics [10], image processing and pattern recognition [11].

The complexity of the set of global patterns for one- or two-dimensional cellular neural networks has been widely discussed [3; 5; 6; 7; 8; 22; 24; 38]. However, this study is the first to explore the complexity for one-dimensional multi-layer CNN. The two-dimensional sofic and two-dimensional multi-layer CNN are discussed in other papers.

A one-dimensional multi-layer CNN system with input is realized as the following form,

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

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

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

and

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

is the output function. For $1 \leq n \leq N$, parameter $A^{(n)} = (a_{-d}^{(n)}, \cdots, a_d^{(n)})$ is called the feedback template; $B^{(n)} = (b_{-d}^{(n)}, \cdots, b_d^{(n)})$ is called the controlling template, and $z^{(n)}$ is the threshold. The quantity $x_i^{(n)}$ denotes the state of a cell C_i in the n-th layer. The stationary solutions $\bar{x} = (\bar{x}_i^{(n)})$ of (1) are essential for understanding the system, and their outputs $\bar{y}_i^{(n)} = f(\bar{x}_i^{(n)})$ are called patterns. A mosaic solution $(\bar{x}_i^{(n)})$ satisfies $|\bar{x}_i^{(n)}| > 1$ for all i, n. Hence the investigation of stationary solution of N-layer CNN is to study a N-coupled map lattice.

$$\begin{cases} x_i^{(1)} = \sum_{|k|\leq d} a_k^{(1)} y_{i+k}^{(1)} + \sum_{|k|\leq d} b_k^{(1)} u_{i+k}^{(1)} + z^{(1)}, \\ x_i^{(2)} = \sum_{|k|\leq d} a_k^{(2)} y_{i+k}^{(2)} + \sum_{|k|\leq d} b_k^{(2)} y_{i+k}^{(1)} + z^{(2)}, \\ \qquad \vdots \\ x_i^{(N)} = \sum_{|k|\leq d} a_k^{(N)} y_{i+k}^{(N)} + \sum_{|k|\leq d} b_k^{(N)} y_{i+k}^{(N-1)} + z^{(N)}. \end{cases} \tag{4}$$

One-layer CNN with input is first considered. Let

$$\mathcal{P}^{n+2} = \{(A, B, z) : A, B \in \mathcal{M}_{1\times(2d+1)}(\mathbb{R}), z \in \mathbb{R}\}, \tag{5}$$

where $n = 4d + 1$. The parameter space $\mathcal{P}^{n+2}$ can be partitioned into finite sub-regions, such that each region has the same mosaic patterns. Once the region of the parameters space is chosen, the basic set of admissible local patterns $\mathcal{B} \subseteq \{+, -\}^{\mathbb{Z}_{3\times 2}}$ is then determined. The ordering matrix of all local patterns in $\{+, -\}^{\mathbb{Z}_{3\times 2}}$ is defined. For a given basic set $\mathcal{B}$, the transition matrix $\mathbf{T}(\mathcal{B})$ is then obtained, and a shift space is induced. For simplicity, considering the case $d = 1$, i.e., each cell can only interact with their nearest neighbors. In one-dimensional one-layer CNN without input, every partition is associated with a unique set of admissible patterns $\overline{\mathcal{B}} = \mathcal{B}_{3\times 1}$ and the transition matrix $\overline{\mathbf{T}} = \mathbf{T}(\mathcal{B}_{3\times 1})$ [24]. Let

$$Y = \{(y_i)_{i\in\mathbb{Z}} |\ y_{i-1}y_iy_{i+1} \in \overline{\mathcal{B}} \text{ for all } i \in \mathbb{Z}\}, \tag{6}$$

then Y is a shift of finite type (SOFT). The number of global admissible patterns with length n and the number of periodic patterns with period m

can then be formulated from the transition matrix $\overline{\mathbf{T}}$. However, this can not be done when the basic set of admissible local patterns $\mathcal{B} = \mathcal{B}_{3\times 2}$ is derived from the one-layer CNN with input. More precisely, each pattern that is produced from the system is a coupled pattern $\frac{y_1y_2y_3}{u_1u_2u_3}$, where $y_1y_2y_3$ denotes the output pattern, and $u_1u_2u_3$ denotes the input pattern. For simplicity, rewriting the coupled pattern as $y_1y_2y_3 \diamond u_1u_2u_3$. The output space is defined as

$$Y_U = \left\{ \begin{array}{c} (\cdots y_{-1}y_0y_1\cdots) \in \{+,-\}^{\mathbb{Z}} : \text{there exists } (\cdots u_{-1}u_0u_1\cdots) \in \{+,-\}^{\mathbb{Z}} \\ \text{such that } (\cdots y_{-1}y_0y_1\cdots \diamond \cdots u_{-1}u_0u_1\cdots) \in \Sigma(\mathcal{B}) \end{array} \right\}, \tag{7}$$

where $\Sigma(\mathcal{B}) \subseteq \{+,-\}^{\mathbb{Z}_{\infty\times 2}}$ is a subshift space generated by $\mathcal{B} \subseteq \{+,-\}^{\mathbb{Z}_{3\times 2}}$. Analytical results indicate that Y_U is not a SOFT, but a sofic shift (Theorem 2.13). Under this situation, the formula of spatial entropy (entropy) $h(\mathcal{B})$ (Theorem 2.17) and dynamical zeta function (zeta function) $\zeta_\sigma(t)$ (Theorem 2.24) can be computed. Therefore, the dynamics of the mosaic solutions of multi-layer CNN are understood. Conversely, the sofic shift is realized through a realistic model.

The analysis gets more complicated in N-layer CNN, $N \geq 2$. However, once recognizing the elaborate content of one-layer CNN with input, all results for one-layer CNN with input can be extended to general case with analogous method. We like to emphasize that each layer induces a sofic shift and the N-layer coupled system induces the convolution of N-many independent sofic shifts. Hence, Section 2 studies one-layer CNN with and without input and emphases the difference. Without input, the dynamical system is subshift of finite type and then sofic when input appears. Section 3 consists those general results introduced in Section 2.

The dynamics of multi-layer CNN with input produce a phenomenon that is never seen in one-layer CNN without input. The entropy of the one-layer CNN without input has a *symmetry* about the parameters. More precisely, consider the one-dimensional CNN,

$$\frac{dx_i}{dt} = -x_i + a_l y_{i-1} + a y_i + a_r y_{i+1} + z, \tag{8}$$

and select one of the partitions of parameter space $\{(a_l, a_r) : a_l, a_r \in \mathbb{R}\} = \mathbb{R}^2$. The parameters a and z thus have 25 subregions, each with the same entropy. Furthermore,

$$h(\mathcal{B}([m,n])) = h(\mathcal{B}([n,m])), \text{ for } 0 \leq m, n \leq 4. \tag{9}$$

The details as in [24]. However, when considering multi-layer CNN with input, not only the entropy and zeta function are varied, but the symmetry of the entropy is broken even for the simplest case one-layer CNN with input. Hence, input adding for a CNN system is the main mechanism that breaks the symmetry of entropy.

Since the spatial entropy and dynamical zeta function can be formulated in a theoretical procedure, it is comprehensible to ask how complicated such system can be. In other words, how many phenomena can be observed in the system?

From the viewpoint of application aspect, most media in natural systems, including physical, biological and electronic systems, are spatially inhomogeneous [21; 33; 37; 25; 20; 14]. This motivates the study of inhomogeneous cellular networks (ICNN). A two-dimensional ICNN system is of the form,

$$\frac{dx_{i,j}}{dt} = \begin{cases} -x_{i,j} + z + \sum\limits_{|k|,|l| \leq d} a_{k,l} f(x_{i+k,j+l}) & \\ \qquad\qquad + \sum\limits_{|k|,|l| \leq d} b_{k,l} u_{i+k,j+l}, & i,j \equiv 0 \bmod m; \\ -x_{i,j} + z' + a_{0,0} f(x_{i,j}), & \text{otherwise.} \end{cases} \tag{10}$$

for some $m \in \mathbb{N}$, $i, j \in \mathbb{Z}$. Restated, the difference between CNN and ICNN is that the templates and threshold at each cell $C_{i,j}$ are spatially invariant for CNN but variant for ICNN.

It is well-known that an important class of applications is steady-state solutions, including mosaic solutions and defect solutions [11; 22; 24]. In recent years, the complexity of steady-state solutions has been extensively studied, and much attention has been paid to the complexity of the set of global patterns, with particular reference to entropy [1; 4; 2; 5; 6; 9; 10; 22; 23; 24; 28; 29; 30; 32]. To study how rich phenomena a the ICNN can achieve, it is equivalent to ask the following question.

Question 1.1. For CNN with/without input, if the radius of the interacting cells d is treated as a parameter, is $\{h(A,B,z,d)\}/\{h(A,z,d)\}$ dense in $[0,\log 2]$?

Multifractal analysis is introduced to a specified dynamical system when one of its invariant is essentially the same as an interval (See [34] for more detail), this motivates us to consider such question. However, since the well-known fact that the entropy of subshift of finite type take a family of specific values, called Perron number [32], the *dense* assumption cannot be removed. The main difficulty in solving the question is related to the fact that the admissible local patterns that are produced by CNN are very limited [22; 24]. Restated, there exists $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n\times n}}$ such that $\mathcal{U} \neq \mathcal{B}(A,z,d)/\mathcal{B}(A,B,z,d)$ for all chosen values of the parameters A,B,z,d, where $n = 2d+1$.

For example, consider the one-dimensional CNN without input, and the length of interaction $d = 1$. Figure 1 is the bifurcation diagram that relates admissible local patterns to the parameters $A = (a_l, a, a_r)$ and z; readers may reference [22; 24] for more details. First, choosing (a_l, a_r), yields a total of eight partitions, as shown in Fig. 1. Second, the $(a-1,z)$ plane has 25 regions such that the admissible local patterns will be uniquely determined once the region is chosen. For instance, if the parameters A, z are in region

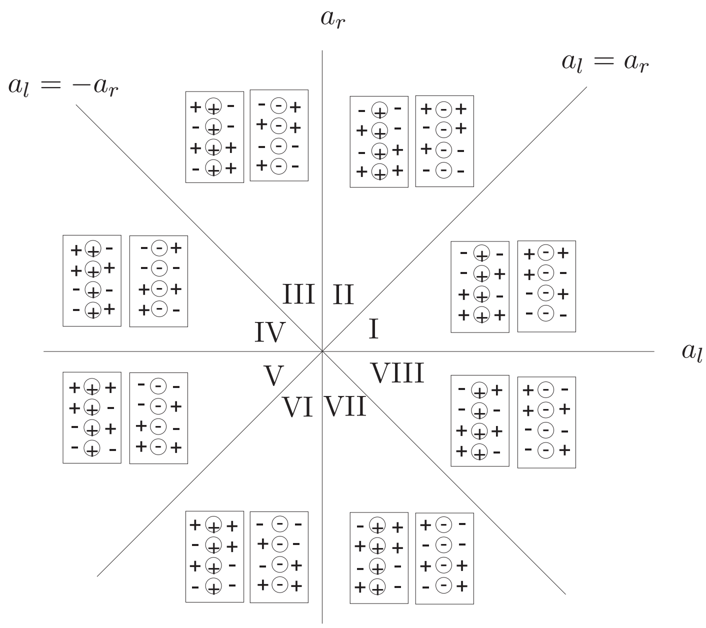


Figure 1: The bifurcation diagram of 1-D CNN

$[3, 4]$ of partition IV, the admissible local patterns are

$$\mathcal{B} = \{- \oplus +, - \oplus -, + \oplus +, + \ominus -, + \ominus +, - \ominus -, - \ominus +\}.$$

That is, "3" indicates that the three patterns with "+" in the center should be chosen from the bottom, and "4" indicates that all four patterns with "−" in the center can be chosen in IV. Thus, Figs. 1 and 2 show all admissible local patterns of 1-D CNN with $d = 1$.

However, let $\mathcal{U} \subseteq \{1, -1\}^{\mathbb{Z}_{3\times 1}}$ be the set of patterns which are listed as follows.

$$\mathcal{U} = \{- \oplus -, - \oplus +, + \oplus -, - \ominus -, - \ominus +, + \ominus -\}.$$

Notably, $\mathcal{U}$ consists of patterns that are selected from different partitions for a_l and a_r. More precisely, the patterns with "+" in the center are located in partition V such that the parameters a_l and a_r must satisfy the conditions $a_l < 0$ and $a_r > 0$. Moreover, the patterns with "−" in the center are selected

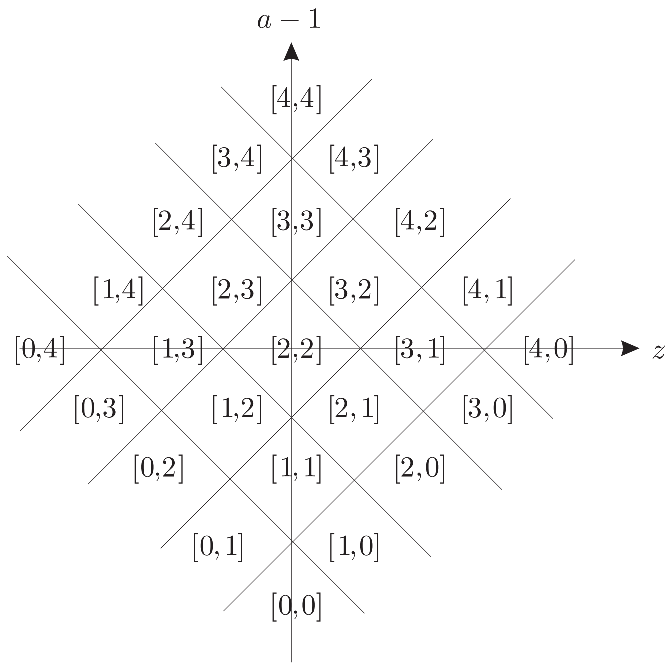


Figure 2: The partition of $a-z$ plane of 1-D CNN

from partition I, in which the associated parameters a_l, a_r must then satisfy $a_l, a_r > 0$. Accordingly, there does not exist A, z such that $\mathcal{B}(A, z) = \mathcal{U}$. Thus, some values of entropy cannot be attained for all choices of 3×1 basic sets for $d = 1$.

In a work on dense entropy, Quas and Trow [36] showed that every subshift of finite type (SFT) $\mathbf{X}$ with positive entropy has proper SFT $\mathbf{X}'$ which is a subsystem of $\mathbf{X}$ whose entropy is strictly less than the entropy of $\mathbf{X}$, but whose entropy is arbitrarily close to that of $\mathbf{X}$. However, they cannot be guaranteed to be mixing [35]. Recently, Desai [15] proved that for any $\mathbb{Z}^d$-SFT $\mathbf{R}$ of positive entropy, the SFT subsystems achieve dense entropy in $[0, h(\mathbf{R})]$. Thus, if $\mathbf{R}$ is treated as a full shift, then the SFT is dense in $[0, \log|\mathcal{A}|]$, where $\mathcal{A}$ denotes the symbols of $\mathbf{R}$, and this result can be generalized to sofic systems. Restated, given a $\mathbb{Z}^d$ sofic shift $\mathbf{T}$, the sofic shift subsystems achieve dense entropy in $[0, h(\mathbf{T})]$. However, a difficulty similar to that associated with CNN arises in solving the problem of ICNN. The difficulty is to guarantee that the patterns that would achieve the desired entropy can be produced by an ICNN system with/without input. This investigation proposes a necessary and sufficient condition for the admissibility of local patterns of ICNN, and demonstrates that suitable local patterns can be found that achieve the given $t \in [0, \log 2]$ (according to Theorem 6.11 for ICNN without input and Theorem 7.3 for the case with input). Finding these patterns solves the dense entropy problem for ICNN. About the same question to classical CNN, we have the following conjecture.

Conjecture. For any $\epsilon > 0$ and $\lambda \in [0, \log 2]$, there exists template A and threshold z such that $|h(\mathcal{B}(A, z)) - \lambda| < \epsilon$.

This dissertation is organized as follows. Section 2 describes the complexity of the global set of output patterns for one-layer CNN with input. The entropy and zeta function can be exactly computed through the induced sofic shift space. Section 3 extends all results in Section 2 to N-layer CNN, where $N \geq 2$. Section 4 lists the detail of the partition in Example 2.5. Section 5 introduces the two-dimensional ICNN model and preliminaries that constitute the background for two-dimensional CNN and extends to two-dimensional ICNN. Section 6 then presents a general theory that yields details about how ICNN relates to a shift of finite type. The solution to the dense entropy problem is also addressed. Section 7 extends the results in Section 6 to ICNN with input.

2 One-layer Cellular Neural Networks with Input

The complexity of the global set of output patterns for one-layer CNN with input is investigated in this section.

2.1 Ordering Matrix and Transition Matrix

In this section, the parameter space $\mathcal{P}^{n+2}$ as in (5) will be partitioned into finite sub-regions, such that each region has the same mosaic patterns. Once the region of the parameters space is chosen, the basic set of admissible local patterns $\mathcal{B} \subseteq \{+,-\}^{\mathbb{Z}_{3\times 2}}$ is then determined. Then, the ordering matrix of all local patterns in $\{+,-\}^{\mathbb{Z}_{3\times 2}}$ will be defined. For a given basic set $\mathcal{B}$, the transition matrix $\mathbf{T}(\mathcal{B})$ will be obtained.

2.1.1 Partition of Parameter Space

This subsection explores the relationship between the parameters of templates and the admissible local output patterns. The differential equation of CNN with input is of the form

$$\frac{dx_i}{dt} = -x_i + \sum_{|k|\leq d} a_k y_{i+k} + \sum_{|\ell|\leq d} b_\ell u_{i+\ell} + z, \tag{1}$$

where $A = [-a_d, \cdots, a, \cdots, a_d], B = [-b_d, \cdots, b, \cdots, b_d]$ are the feedback and controlling templates, respectively, $y = f(x) = \frac{1}{2}(|x+1| - |x-1|)$ is the output function, z is the threshold, and $a \equiv a_0, b \equiv b_0$.

The quantity x_i represents the state of the cell at i. The stationary solution $\bar{x} = (\bar{x}_i)$ of (1) satisfies

$$\bar{x}_i = \sum_{|k|\leq d} a_k \bar{y}_{i+k} + \sum_{|\ell|\leq d} b_\ell u_{i+\ell} + z.$$

The output $\bar{y} = (\bar{y}_i)$ is called 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., $\bar{x}_i > 1$, is

$$a - 1 + z > -(\sum_{0<|k|\le d} a_k \bar{y}_{i+k} + \sum_{|\ell|\le d} b_\ell u_{i+\ell}). \tag{2}$$

Similarly, the necessary and sufficient conditions for state "−" at cell C_i, i.e., $\bar{x}_i < -1$, is

$$a - 1 - z > \sum_{0<|k|\le d} a_k \bar{y}_{i+k} + \sum_{|\ell|\le d} b_\ell u_{i+\ell}. \tag{3}$$

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

$$\frac{y_{-d}\cdots y\cdots y_d}{u_{-d}\cdots u\cdots u_d} = Y \diamond U, \tag{4}$$

where $Y = y_{-d}\cdots y\cdots y_d$, $U = u_{-d}\cdots u\cdots u_d$. Let

$$V^n = \{v \in \mathbb{R}^n : v = (v_1, v_2, \cdots, v_n), \text{ and } |v_i| = 1, 1 \le i \le n\},$$

where $n = 4d + 1$, (2) and (3) can be rewritten in a compact form by introducing the following notation.

Denote $\alpha = (a_{-d}, \cdots, a_{-1}, a_1, \cdots, a_d)$, $\beta = (b_{-d}, \cdots, b, \cdots, b_d)$. 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

$$\mathcal{B}(+, A, B, z) = \{v \diamond w \in V^n : a - 1 + z > -(\alpha \cdot v + \beta \cdot w)\}, \tag{5}$$

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

$$\mathcal{B}(-, A, B, z) = \{v' \diamond w' \in V^n : a - 1 - z > \alpha \cdot v + \beta \cdot w\}. \tag{6}$$

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

$$\mathcal{B}(A, B, z) = (\mathcal{B}(+, A, B, z), \mathcal{B}(-, A, B, z)). \tag{7}$$

Let

$$\mathcal{P}^{n+2} = \{(A, B, z)|\ A, B \in \mathcal{M}_{1\times(2d+1)}(\mathbb{R}), z \in \mathbb{R}\}, \tag{8}$$

where $\mathcal{M}_{r\times s}(\mathbb{R})$ means a $r \times s$ real matrix. $\mathcal{P}^{n+2}$ can be partitioned so that each subregion generates the same mosaic patterns, when the controlling template $B \equiv 0$ is proved in [22]. The general results for $B \neq 0$ can be obtained similarly, so the detailed proof is omitted for simplicity.

Theorem 2.1. *There exists positive integer $K(n)$ and an unique collection of open subsets $\{P_k\}_{k=1}^{K}$ of $\mathcal{P}^{n+2}$ satisfying*

(i) $\mathcal{P}^{n+2} = \bigcup_{k=1}^{K} \overline{P}_k$.

(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 $\mathcal{P}^{n+2}$.

2.1.2 Ordering Matrix

This subsection defines the ordering matrix $\mathbb{X} = \mathbb{X}_{3\times 2}$ of all possible local patterns in $\{+, -\}^{\mathbb{Z}_{3\times 2}}$. First, the notation of the pattern with size 3×1 is considered.

Let

$$a_{00} = --, a_{01} = -+, a_{10} = +-, a_{11} = ++, \tag{9}$$

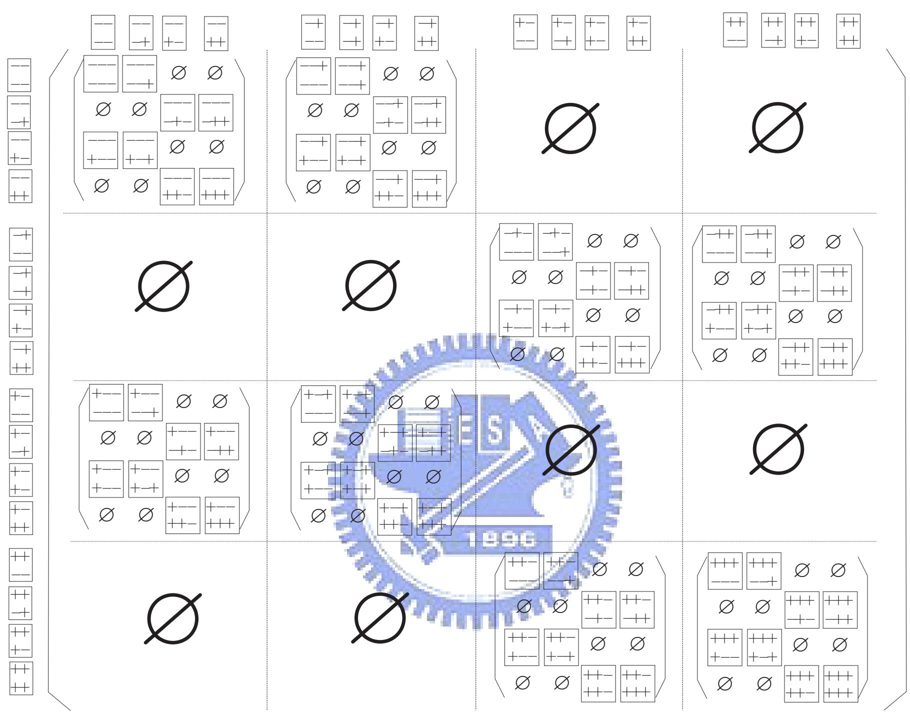


Figure 3: The ordering matrix of all local patterns in $\mathbb{Z}_{3\times 2}$

defining

$$a_{i_1 i_2} a_{i_2' i_3} = \varnothing \Leftrightarrow i_2 \neq i_2'. \tag{10}$$

If $a_{i_1 i_2} a_{i_2' i_3} \neq \varnothing$, then denoting it by $a_{i_1} a_{i_2} a_{i_3}$ and it is a pattern with size 3×1. Define

$$\mathbb{X} = \begin{bmatrix} X_{11} & X_{12} & X_{13} & X_{14} \\ X_{21} & X_{22} & X_{23} & X_{24} \\ X_{31} & X_{32} & X_{33} & X_{34} \\ X_{41} & X_{42} & X_{43} & X_{44} \end{bmatrix}, \ X_{ij} = \begin{bmatrix} x_{ij;11} & x_{ij;12} & x_{ij;13} & x_{ij;14} \\ x_{ij;21} & x_{ij;22} & x_{ij;23} & x_{ij;24} \\ x_{ij;31} & x_{ij;32} & x_{ij;33} & x_{ij;34} \\ x_{ij;41} & x_{ij;42} & x_{ij;43} & x_{ij;44} \end{bmatrix} \tag{11}$$

for $1 \leq i, j \leq 4$ as Figure 3. $x_{ij;kl}$ means the pattern $\begin{smallmatrix} a_{r_1 r_2} a_{r_2' r_3} \\ a_{s_1 s_2} a_{s_2' s_3} \end{smallmatrix}$, where

$$\begin{aligned} r_1 &= \left[\frac{i-1}{2}\right], \quad r_2 = i - 1 - 2r_1, \quad r_2' = \left[\frac{j-1}{2}\right], \quad r_3 = j - 1 - 2r_2', \\ s_1 &= \left[\frac{k-1}{2}\right], \quad s_2 = k - 1 - 2s_1, \quad s_2' = \left[\frac{l-1}{2}\right], \quad s_3 = l - 1 - 2s_2', \end{aligned} \tag{12}$$

and $[\cdot]$ is the Gauss function.

If $a_{r_1 r_2} a_{r_2' r_3} = \varnothing$ or $a_{s_1 s_2} a_{s_2' s_3} = \varnothing$, then $x_{ij;kl} = \varnothing$. Furthermore, if $x_{ij;kl} \neq \varnothing$, then it is denoted by the pattern $\begin{smallmatrix} a_{r_1} a_{r_2} a_{r_3} \\ a_{s_1} a_{s_2} a_{s_3} \end{smallmatrix}$ in $\{+, -\}^{\mathbb{Z}_{3\times 2}}$. Hence, the self-similar property appear in $\mathbb{X}$ as in Figure 3, i.e., the upper pattern of X_{ij} is the same as the lower pattern of $x_{kl;ij}$, for $1 \leq i, j, k, l \leq 4$. Once the basic set of admissible local patterns $\mathcal{B} \subseteq \{+, -\}^{\mathbb{Z}_{3\times 2}}$ is given, defining $\Sigma_{n\times 2}(\mathcal{B})$ the collection of all patterns with size $n \times 2$ generated by $\mathcal{B}$ as

$$\Sigma_{n\times 2}(\mathcal{B}) = \left\{ \begin{smallmatrix} y_1 y_2 \cdots y_n \\ u_1 u_2 \cdots u_n \end{smallmatrix} \in \{+, -\}^{\mathbb{Z}_{n\times 2}} : \begin{smallmatrix} y_{i-1} y_i y_{i+1} \\ u_{i-1} u_i u_{i+1} \end{smallmatrix} \in \mathcal{B} \text{ for all } 2 \leq i \leq n-1 \right\}. \tag{13}$$

For simplicity, rewriting $\begin{smallmatrix} y_1 y_2 \cdots y_n \\ u_1 u_2 \cdots u_n \end{smallmatrix}$ as $y_1 y_2 \cdots y_n \diamond u_1 u_2 \cdots u_n$, where $y_i, u_i \in \{+, -\}$, $1 \leq i \leq n$.

To measure the complexity of the global set of output patterns, the following subshift space in $\{+, -\}^{\mathbb{Z}}$ is considered. Defining the output space

$Y_U \equiv Y_U(\mathcal{B})$ by

$$Y_U = \left\{ \begin{array}{c} (\cdots y_{-1}y_0y_1\cdots) \in \{+,-\}^{\mathbb{Z}} : \text{there exists } (\cdots u_{-1}u_0u_1\cdots) \in \{+,-\}^{\mathbb{Z}} \\ \text{such that } (\cdots y_{-1}y_0y_1\cdots \diamond \cdots u_{-1}u_0u_1\cdots) \in \Sigma(\mathcal{B}) \end{array} \right\}, \tag{14}$$

where $\Sigma(\mathcal{B}) \subseteq \{+,-\}^{\mathbb{Z}_{\infty\times 2}}$ is a subshift space generated by $\mathcal{B} \subseteq \{+,-\}^{\mathbb{Z}_{3\times 2}}$.

2.1.3 Transition Matrix

This subsection derives the transition matrix for a given basic set $\mathcal{B}$. The transition matrix $\mathbf{T}$ is defined as

$$\mathbf{T}(\mathcal{B}) = (T_{ij}),\ 1 \le i,j \le 4, \tag{15}$$

where $T_{ij} = (t_{ij;kl}) \in \mathcal{M}_{4\times 4}(\mathbb{R})$ and

$$t_{ij;kl} = \begin{cases} 1, & \text{if } x_{ij;kl} \in \mathcal{B}; \\ 0, & \text{if } x_{ij;kl} \in \{+,-\}^{\mathbb{Z}_{3\times 2}} \setminus \mathcal{B} \text{ or } x_{ij;kl} = \varnothing, \end{cases} \tag{16}$$

where $x_{ij;kl} = \begin{smallmatrix} a_{r_1r_2}a_{r_2'r_3} \\ a_{s_1s_2}a_{s_2's_3} \end{smallmatrix}$ satisfies (12). Once $\mathbf{T}(\mathcal{B})$ is constructed, it is then rewritten as $\mathbf{T}(\mathcal{B}) = (t_{pq}) \in \mathcal{M}_{16\times 16}(\mathbb{R})$, where

$$t_{pq} = t_{ij;kl},\ \text{for } p = 4(i-1)+k, q = 4(j-1)+l. \tag{17}$$

If templates A, B and threshold z are given, then the basic set $\mathcal{B}(A,B,z)$ is obtained from Theorem 2.1. Moreover, the transition matrix $\mathbf{T}$ is immediately derived from (16) and (17).

If a set of input patterns $\mathcal{U} \equiv \{u_1u_2u_3\} \subseteq \{+,-\}^{\mathbb{Z}_{3\times 1}}$ is assigned, then the basic set of admissible local patterns is denoted by

$$\mathcal{B}((A,B,z);\mathcal{U}) = \left\{ \begin{smallmatrix} y_1y_2y_3 \\ u_1u_2u_3 \end{smallmatrix} \in \mathcal{B}(A,B,z) : u_1u_2u_3 \in \mathcal{U} \right\}. \tag{18}$$

The transition matrix for $\mathcal{U}$ is defined by $\mathbf{U} = (u_{ij}) \in \mathcal{M}_{4\times 4}(\mathbb{R})$, where

$$u_{ij} = \begin{cases} 1, & \text{if } u_1u_2u_3 \in \mathcal{U} \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

If $\widehat{\mathbf{T}}$ denotes the transition matrix of $\mathcal{B}((A,B,z);\mathcal{U})$, then the following theorem is obtained. Before the theorem is stated, two products of matrices are defined as follows.

Definition 2.2. *For any two matrices* $\mathbf{M} = (m_{ij}) \in \mathcal{M}_{k\times k}(\mathbb{R})$, $\mathbf{N} = (n_{i'j'}) \in \mathcal{M}_{\ell\times\ell}(\mathbb{R})$, *the Kronecker product (tensor product)* $\mathbf{M}\otimes\mathbf{N}$ *of* $\mathbf{M}$ *and* $\mathbf{N}$ *is defined by*

$$\mathbf{M}\otimes\mathbf{N} = (m_{ij}\mathbf{N}) \in \mathcal{M}_{k\ell\times k\ell}(\mathbb{R}). \tag{20}$$

Next, for any $\mathbf{P} = (p_{ij}), \mathbf{Q} = (q_{ij}) \in \mathcal{M}_{r\times r}(\mathbb{R})$, *the Hadamard product* $\mathbf{P}\circ\mathbf{Q}$ *of* $\mathbf{P}$ *and* $\mathbf{Q}$ *is defined by*

$$\mathbf{P}\circ\mathbf{Q} = (p_{ij}q_{ij}) \in \mathcal{M}_{r\times r}(\mathbb{R}). \tag{21}$$

Theorem 2.3. *If* (A,B,z) *and* $\mathcal{U}$ *are given, then*

$$\widehat{\mathbf{T}}(\mathcal{B}((A,B,z);\mathcal{U})) = \mathbf{T}(\mathcal{B}(A,B,z)) \circ (E_4 \otimes \mathbf{U}), \tag{22}$$

where $\circ$ *and* $\otimes$ *is the Hadamard product and Kronecker product in Definition 2.2, respectively,* $E_4 = (e_{ij}) \in \mathcal{M}_{4\times 4}(\mathbb{R})$ *with* $e_{ij} = 1$ *for all* i,j, *is the full matrix.*

Proof. Let the transition matrix $\widehat{\mathbf{T}} = (\hat{t}_{pq}) \in \mathcal{M}_{16\times 16}(\mathbb{R})$, rewriting $\hat{t}_{pq} = \hat{t}_{ij;kl}$, where

$$i = \left[\frac{p-1}{4}\right] + 1, \quad j = p - 4(i-1), \quad k = \left[\frac{q-1}{4}\right] + 1,$$

and $l = q - 4(j-1)$. By (16) and (18), $\hat{t}_{pq} = \hat{t}_{ij;kl} = t_{ij;kl} \cdot u_{kl}$ is obtained. This completes the proof. $\square$

2.1.4 Patterns Generation

This subsection introduces the patterns generation problem induced by (A, B, z). Some results of patterns generation problem induced by (A, z) must be recalled before stating the main theory [32]. The basic set of admissible local patterns $\mathcal{B}(A, z) \equiv \mathcal{B}$ is then determined once (A, z) is given. The shift space generated by $\mathcal{B}$, denoted by $\Sigma(\mathcal{B})$, is given by

$$\Sigma(\mathcal{B}) = \{y = (y_i)_{i \in \mathbb{Z}} \in \{+, -\}^{\mathbb{Z}} : y_{i-1}y_iy_{i+1} \in \mathcal{B} \text{ for all } i \in \mathbb{Z}\}. \qquad (23)$$

The shift space $\Sigma(\mathcal{B})$ is thus a subshift of finite type for all $\mathcal{B} = \mathcal{B}(A, z)$. Let $\Sigma_n(\mathcal{B})$ denote the set of n-blocks (i.e., the pattern with size $n \times 1$) in $\Sigma(\mathcal{B})$, and $\Gamma_n(\mathcal{B})$ denote the number of the set of n-blocks. The theorem follows below [32].

Theorem 2.4. *If $\mathcal{B} = \mathcal{B}(A, z)$ is given, $\mathbf{T}$ is the transition matrix induced by $\mathcal{B}$, then $\Gamma_n(\mathcal{B}) = |\mathbf{T}^{n-2}|$ for all $n \in \mathbb{N}, n \geq 3$, where $|\mathbf{T}| \equiv \sum\limits_{1 \leq i,j \leq k} |t_{ij}|$ for all $\mathbf{T} = (t_{ij}) \in \mathcal{M}_{k \times k}(\mathbb{R})$.*

Considering (A, B, z) with $B \neq 0$, let $\mathcal{B}(A, B, z) \equiv \mathcal{B}$ denote the basic set of admissible local patterns, $\Sigma_n(Y_U)$ denote the set of n-blocks in Y_U, i.e.,

$$\Sigma_n(Y_U) = \left\{ \begin{array}{c} y = (y_i)_{i=1}^n \in \{+, -\}^{\mathbb{Z}_{n \times 1}} : \exists\ u = (u_i)_{i=1}^n \in \{+, -\}^{\mathbb{Z}_{n \times 1}} \\ \text{such that } y \diamond u \in \Sigma_n(\mathcal{B}) \end{array} \right\}, \qquad (24)$$

and $\Gamma_n(Y_U)$ denote the number of n-blocks of output patterns generated by $\mathcal{B}$. Theorem 2.4 is invalid in general for deriving the precise value of $\Gamma_n(Y_U)$ for $n \in \mathbb{N}$. An example is given below. The Appendix explains the theorem in detail.

Example 2.5. Let $A = [a_l, a, a_r], B = [b_l, b, b_r]$ satisfy the following conditions.

(i) $a_l > b_l > a_r > b > b_r > 0$.

(ii) $a_l + b > a_r + b_l + b_r, a_l + b_r > b_l + b$.

(iii) $b_l + b > a_l > a_r + b + b_r$.

(iv) $a_r + b > b_l + b_r, b_l > a_r + b_r, a_r > b + b_r$.

The positions of ℓ_i^+ and ℓ_j^- on the $(a-1, z)$ plane are determined exactly as in the Appendix. Given region $R = [23, 18]$, i.e., R is bounded by $\ell_{23}^+, \ell_{24}^+, \ell_{18}^-$ and ℓ_{19}^-, and the set of input patterns is given by $\mathcal{U} = \{-+-, -++, +-+\}$. Thus, Figure 4 illustrates the basic set of admissible local patterns $\mathcal{B} = \mathcal{B}((A, B, z); \mathcal{U})$.

$$
\begin{array}{cccc}
\begin{array}{ccc} - & \ominus & - \\ - & \boxplus & - \end{array} &
\begin{array}{ccc} - & \ominus & - \\ - & \boxplus & + \end{array} &
\begin{array}{ccc} - & \ominus & - \\ + & \boxminus & + \end{array} &
\begin{array}{ccc} - & \ominus & + \\ - & \boxplus & - \end{array} \\[2ex]
\begin{array}{ccc} - & \ominus & + \\ - & \boxplus & + \end{array} &
\begin{array}{ccc} - & \ominus & + \\ + & \boxminus & + \end{array} &
\begin{array}{ccc} + & \ominus & - \\ - & \boxplus & - \end{array} &
\begin{array}{ccc} + & \ominus & - \\ - & \boxplus & + \end{array} \\[2ex]
\begin{array}{ccc} + & \oplus & + \\ + & \boxminus & + \end{array} &
\begin{array}{ccc} + & \oplus & + \\ - & \boxplus & + \end{array} &
\begin{array}{ccc} + & \oplus & + \\ - & \boxplus & - \end{array} &
\begin{array}{ccc} + & \oplus & - \\ + & \boxminus & + \end{array} \\[2ex]
\begin{array}{ccc} + & \oplus & - \\ - & \boxplus & + \end{array} &
\begin{array}{ccc} + & \oplus & - \\ - & \boxplus & - \end{array} &
\begin{array}{ccc} - & \oplus & + \\ + & \boxminus & + \end{array} &
\begin{array}{ccc} - & \oplus & + \\ - & \boxplus & + \end{array}
\end{array}
$$

Figure 4: The basic set of patterns for some templates A, B, z and input $\mathcal{U}$

According to Theorem 2.3, the transition matrix of $\mathcal{B}((A, B, z); \mathcal{U})$ is

$$
\widehat{\mathbf{T}} = \widehat{\mathbf{T}}(\mathcal{B}((A, B, z); \mathcal{U})) = \begin{pmatrix} \widehat{T}_{11} & \widehat{T}_{12} & 0 & 0 \\ 0 & 0 & 0 & \widehat{T}_{24} \\ \widehat{T}_{31} & 0 & 0 & 0 \\ 0 & 0 & \widehat{T}_{43} & \widehat{T}_{44} \end{pmatrix},
$$

where

$$\widehat{T}_{11} = \widehat{T}_{12} = \widehat{T}_{43} = \widehat{T}_{44} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \widehat{T}_{31} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \widehat{T}_{24} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

From the transition matrix, the output patterns $\{-++, ++-, +--\}$ exist. By the concept of subshift of finite type, the output pattern $-++--$ is admissible. However, there exists no $u_1u_2u_3u_4u_5 \in \Sigma_5(\mathcal{U})$ such that $-++-- \diamond u_1u_2u_3u_4u_5 \in \Sigma_5(\mathcal{B})$. This finding shows that the inner structure needs to be considered. More precisely, since $\widehat{T}_{24}\widehat{T}_{43}\widehat{T}_{31} = 0$, no input could possibly produce the output pattern $-++--$.

So far, this work has shown that $\Gamma_n(Y_U) \neq |\widehat{\mathbf{T}}^{n-2}|$ in general, since different input patterns might have the same output pattern. To overcome this difficulty, the next subsection introduces the concept of sofic shift in the symbolic dynamical system.

2.2 Definition and Background of Sofic Shifts

This subsection recalls some definitions and main results of sofic shifts. Lind and Marcus has described sofic shifts in detail [32].

Definition 2.6. *A labeled graph $\mathcal{G} = (G, \mathcal{L})$ consists of an underlying graph G with edge set $\mathcal{E}$, and the labeling $\mathcal{L} : \mathcal{E} \to \mathcal{A}$ assigns to each edge a label from the finite alphabet $\mathcal{A}$. A sofic shift is defined by $\mathbf{X} = \mathbf{X}_{\mathcal{G}}$ for some labeled graph $\mathcal{G}$.*

Definition 2.7. *A labeled graph $\mathcal{G} = (G, \mathcal{L})$ is right-resolving if, for each vertex of I of G, the edges starting from I carry different labels. In other words, $\mathcal{G}$ is right-resolving if, for each I, the restriction of $\mathcal{L}$ to $\mathcal{E}_I$ is one-to-one, where $\mathcal{E}_I$ consists of those edges starting from I.*

The following theorem shows that every sofic shift has a right-resolving presentation. The method for finding an explicit right-resolving presentation is called the subset construction method.

Subset Construction Method

Let $\mathbf{X}$ be a sofic shift over the alphabet $\mathcal{A}$ having a presentation $\mathcal{G} = (G, \mathcal{L})$ so that $\mathbf{X} = \mathbf{X}_{\mathcal{G}}$. If $\mathcal{G}$ is not right-resolving, then a new labeled graph $\mathcal{H} = (H, \mathcal{L}')$ 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. There are two cases.

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 2.8. *Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph which is not right-resolving, $\mathcal{H} = (H, \mathcal{L}')$ be a right-resolving labeled graph constructed under the subset construction method. Then $\mathbf{X}_{\mathcal{G}} = \mathbf{X}_{\mathcal{H}}$, i.e., $\mathcal{G}$ and $\mathcal{H}$ presents the same shift space.*

2.3 Entropy and Zeta Function

This subsection investigates the entropy and zeta function for the global set of output patterns using the concepts of sofic shifts.

2.3.1 Sofic Shift

This subsection shows that the output space of one-layer CNN with input is a sofic shift. For a given basic set $\mathcal{B}$, the transition matrix $\mathbf{T}$ is defined as (15), (16) and (17). Let the alphabet $\mathcal{S} = \{s_{ij}\}_{1\le i,j\le 4}$, where

$$s_{ij} = a_{r_1r_2}a_{r_2'r_3}, \tag{25}$$

$a_{r_kr_{k+1}}$ is defined in (9), and

$$r_1 = \left[\frac{i-1}{2}\right], \quad r_2 = i-1-2r_1, \quad r_2' = \left[\frac{j-1}{2}\right], \quad r_3 = j-1-2r_2'. \tag{26}$$

By (10), $s_{ij} = \varnothing$ if $r_2 \ne r_2'$. The symbolic transition matrix is defined as

$$\mathbf{S} = (s_{ij}T_{ij})_{1\le i,j\le 4} = (s_{ij}t_{ij;kl}) \in \mathcal{M}_{16\times 16}(\mathbb{R}), \tag{27}$$

where $s_{ij}t_{ij;kl} = \varnothing$ if $s_{ij} = \varnothing$ or $t_{ij;kl} = 0$. Rewriting $\mathbf{S} = (\tilde{s}_{pq})$, where $\tilde{s}_{pq} = s_{ij}t_{ij;kl}$ for

$$p = 4(i-1)+k, \quad q = 4(j-1)+l.$$

Let $G_{\mathbf{T}}$ be the underlying graph induced by $\mathbf{T}$ with edge set

$$\mathcal{E} = \{e_{pq} : t_{pq} = 1, 1 \le p, q \le 16\},$$

and the labeling $\mathcal{L} : \mathcal{E} \to \mathcal{S}$ defined by $\mathcal{L}(e_{pq}) = s_{ij}$. $\mathcal{G}_{\mathbf{S}} = (G_{\mathbf{T}}, \mathcal{L})$ is thus a labeled graph as in Figure 5. By (25), a word $s_{i_1i_2}s_{i_2i_3}$ in $\mathcal{S}^{\mathbb{Z}}$ can be defined by $s_{i_1i_2}s_{i_2i_3} = a_{r_1r_2}a_{r_2r_3}a_{r_3r_4}$. The edge shift with alphabet $\mathcal{S}$ is defined by

$$\mathbf{X}_{\mathcal{G}_{\mathbf{S}}} = \left\{ \begin{array}{c} (\cdots s_{i_{-1}i_0}s_{i_0i_1}s_{i_1i_2}\cdots) \in \mathcal{S}^{\mathbb{Z}} : \text{ there exists } (\cdots k_{-1}k_0k_1\cdots) \\ \text{such that } t_{i_ji_{j+1};k_jk_{j+1}} \ne 0 \text{ for all } j \in \mathbb{Z} \end{array} \right\}. \tag{28}$$

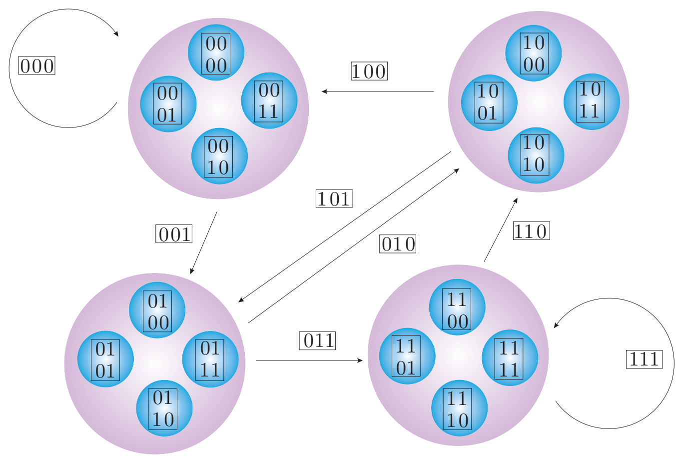


Figure 5: The labeled graph of CNN with input

The following theorem is thus obtained.

Theorem 2.9. $\mathbf{X} = \mathbf{X}_{\mathcal{G}_\mathbf{S}}$ *is a sofic shift.*

The relationship between output space Y_U and induced sofic shift $\mathbf{X}_{\mathcal{G}_\mathbf{S}}$ is then investigated.

Definition 2.10. *Let* $\mathcal{A}$,$\mathcal{U}$ *be finite alphabets,* $\mathbf{X}$ *be a shift space over* $\mathcal{A}$, $B_k(\mathbf{X})$ *denote the set of* k*-blocks that occur in points in* $\mathbf{X}$, $\Phi : B_{m+n+1}(\mathbf{X}) \to \mathcal{U}$ *be a block map. Then the map* $\phi : \mathbf{X} \to \mathcal{U}^{\mathbb{Z}}$ *defined by* $y = \phi(x)$ *with*

$$y_i = \Phi(x_{i-m} \cdots x_{i-1} x_i x_{i+1} \cdots x_{i+n}) = \Phi(x_{[i-m,i+n]})$$

is called the sliding block code with memory m *and anticipate* n *induced by* Φ.

Let $\Sigma_3(Y_U)$ be the set of 3-blocks in Y_U as in (24). A block map $\Phi : \Sigma_3(Y_U) \to \mathcal{S}$ is thus defined by

$$\Phi(y_i y_{i+1} y_{i+2}) = s_{\Psi(y_i y_{i+1})\Psi(y_{i+1} y_{i+2})}, \tag{29}$$

where $\Psi(y_iy_{i+1}) = 1 + \varphi(y_{i+1}) + 2\varphi(y_i)$ and φ is defined by $\varphi(+) = 1$ and $\varphi(-) = 0$. Then the map $\phi : Y_U \to \mathcal{S}^{\mathbb{Z}}$ defined by

$$\phi(\cdots y_{-1}y_0y_1\cdots) = (\cdots s_{i_{-1}i_0}s_{i_0i_1}s_{i_1i_2}\cdots) \tag{30}$$

with $s_{i_ki_{k+1}} = \Phi(y_{i_k}y_{i_{k+1}}y_{i_{k+2}})$ is a sliding block code from Y_U to $\mathbf{X}_{\mathcal{G}_\mathbf{S}}$.

Definition 2.11. *Let $\phi : \mathbf{X} \to \mathbf{Y}$ be a sliding block code, then ϕ is a conjugacy if ϕ is invertible. It is also called $\mathbf{X}$ is conjugate to $\mathbf{Y}$.*

Theorem 2.12 ([32])**.** *If a sliding code is one-to-one and onto, then it is a conjugacy.*

The conjugacy between Y_U and $\mathbf{X}_{\mathcal{G}_\mathbf{S}}$ can be proved now.

Theorem 2.13. *Given a basic set $\mathcal{B}$, the transition matrix $\mathbf{T}$ and the output space Y_U are then obtained. Let $\mathcal{G}_\mathbf{S} = (G_\mathbf{T}, \mathcal{L})$ be the labeled graph induced by $\mathcal{B}$, then Y_U is conjugate to $\mathbf{X}_{\mathcal{G}_\mathbf{S}}$ under the sliding block code ϕ defined in (30).*

Proof. It suffices to prove that ϕ is one-to-one and onto. If there exist $x \neq y \in Y_U$ such that $\phi(x) = \phi(y)$, without loss of generality, assuming that there is a number n such that $x_n \neq y_n$ and $x_i = y_i$ for all $i < n$. This means that at state $x_{n-1} = y_{n-1}$ and $x_{n-2} = y_{n-2}$,

$$\Psi(x_{n-1}x_n) = 1 + \varphi(x_{n-1}) + 2\varphi(x_n) \neq 1 + \varphi(y_{n-1}) + 2\varphi(y_n) = \Psi(y_{n-1}y_n).$$

Therefore,

$$\Phi(x_{n-2}x_{n-1}x_n) = s_{\Psi(x_{n-2}x_{n-1})\Psi(x_{n-1}x_n)} \neq s_{\Psi(y_{n-2}y_{n-1})\Psi(y_{n-1}y_n)} = \Phi(y_{n-2}y_{n-1}y_n).$$

This contradicts to $\phi(x) = \phi(y)$. Thus, ϕ is one-to-one.

For every $s = (\cdots s_{i_{-1}i_0}s_{i_0i_1}s_{i_1i_2}\cdots) \in \mathbf{X}_{\mathcal{G}_\mathbf{S}}$, by (28) there exists a sequence $(\cdots k_{-1}k_0k_1\cdots)$ such that $t_{i_ji_{j+1};k_jk_{j+1}} \neq 0$ for all $j \in \mathbb{Z}$. By

(16), the related pattern $\cdots x_{i_{-1}i_0;k_{-1}k_0} x_{i_0 i_1;k_0 k_1} x_{i_1 i_2;k_1 k_2}\cdots$ is admissible, i.e., $\begin{smallmatrix}\cdots a_{r_{-1}}a_{r_0}a_{r_1}\cdots \\ \cdots a_{s_{-1}}a_{s_0}a_{s_1}\cdots\end{smallmatrix}$ is admissible. By the definition of output space (14), the pattern $\cdots a_{r_{-1}}a_{r_0}a_{r_1}\cdots \in Y_U$. Moreover, by the relation in (12),

$$i_k = 1 + r_{k+1} + 2r_k, \text{ and } r_k = \begin{cases} 1, & \text{if } a_{r_0} = +; \\ 0, & \text{if } a_{r_0} = -. \end{cases}$$

Hence, there exists $\cdots a_{r_{-1}}a_{r_0}a_{r_1}\cdots \in Y_U$ such that $\phi(\cdots a_{r_{-1}}a_{r_0}a_{r_1}\cdots) = s$. This shows that ϕ is onto, and the proof is completed. $\square$

2.3.2 Entropy

Let $\Sigma_n(Y_U) \subseteq \{+,-\}^{\mathbb{Z}_{n\times 1}}$ denote the set of n-blocks in Y_U as in (24). The spatial entropy of Y_U is defined by

$$h(Y_U) = \lim_{n\to\infty} \frac{\log \Gamma_n(Y_U)}{n}, \tag{31}$$

where $\Gamma_n(Y_U)$ is the cardinal number of $\Sigma_n(Y_U)$. Example 2.5 demonstrates that Theorem 2.4 is invalid in computing the spatial entropy of Y_U. However, Theorem 2.13 shows that the output space Y_U is conjugate to the sofic shift $\mathbf{X}_{\mathcal{G}_\mathbf{S}}$. The entropy of Y_U can be computed by the following theorems show in [32].

Theorem 2.14 ([32]). *If two shift spaces* $\mathbf{X}$ *and* $\mathbf{Y}$ *are conjugate, then* $h(\mathbf{X}) = h(\mathbf{Y})$.

Theorem 2.15 ([32]). *Let* $\mathcal{G} = (G, \mathcal{L})$ *be a labeled graph. If* $\mathcal{G}$ *is right-resolving, then* $h(\mathbf{X}_\mathcal{G}) = h(\mathbf{X}_G)$.

Theorem 2.16. *For* Y_U *is given,* $\mathcal{G}_\mathbf{S}$ *is the sofic shift induced by* Y_U. *If* $\mathcal{G}_\mathbf{S}$ *is right-resolving, then* $h(Y_U) = \log \rho(\mathbf{T})$, *where* $\rho(\mathbf{T})$ *denotes the maximal eigenvalue of* $\mathbf{T}$.

Proof. Since Y_U is conjugate to $\mathbf{X}_{\mathcal{G}_\mathbf{S}}$, and $\mathcal{G}_\mathbf{S}$ is right-resolving, by Theorem 2.14, Theorem 2.15, and Perron-Frobenius theorem,

$$h(Y_U) = h(\mathbf{X}_{\mathcal{G}_\mathbf{S}}) = h(\mathbf{X}_\mathbf{T}) = \log \rho(\mathbf{T}). \tag{32}$$

This completes the proof. □

In general, $\mathcal{G}_\mathbf{S}$ induced by $Y_U(\mathcal{B})$ might not be right-resolving. However, a sofic shift $\mathcal{H} = (H, \mathcal{L}')$ which is right-resolving can be constructed and still conjugate to $Y_U(\mathcal{B})$ via the subset construction method stated in Subsection 2.2. Thus Theorem 2.16 can be extended to the general case.

Theorem 2.17. *For a given $\mathcal{B} \subseteq \{-,+\}^{\mathbb{Z}_{3\times 2}}$, let $Y_U \equiv Y_U(\mathcal{B})$ be the shift space induced by $\mathcal{B}$. Then there exists a labeled graph representation $\mathcal{H} = (H, \mathcal{L}')$ such that*

$$h(Y_U) = h(\mathbf{X}_\mathcal{H}) = \log \rho(\mathbf{H}). \tag{33}$$

Proof. For the admissible local patterns $\mathcal{B}$ given, let $\mathbf{T}$ and the labeled graph representation $\mathcal{G}_\mathbf{S}$ defined as above. If $\mathcal{G}_\mathbf{S}$ is already right-resolving, then it is done. If not, using subset construction method, there is a labeled graph representation $\mathcal{H}$ such that $\mathbf{X}_{\mathcal{G}_\mathbf{S}} = \mathbf{X}_\mathcal{H}$ and is right-resolving. Note that the underlying graph and transition matrix of $\mathcal{H}$, H and $\mathbf{H}$, are also derived. Moreover, by Theorem 2.8 and Theorem 2.13, Y_U is conjugate to $\mathbf{X}_\mathcal{H}$. Thus, by Theorem 2.16, $h(Y_U) = h(\mathbf{X}_\mathcal{H}) = \log \rho(\mathbf{H})$. □

Example 2.18 (Continued)**.** Let (A, B, z) be the same as in Example 2.5, $\mathcal{S} = \{s_{11}, s_{12}, s_{24}, s_{31}, s_{43}, s_{44}\}$, by (27), the symbolic transition matrix is

$$\mathbf{S} = \begin{pmatrix} s_{11}T_{11} & s_{12}T_{12} & 0 & 0 \\ 0 & 0 & 0 & s_{24}T_{24} \\ s_{31}T_{31} & 0 & 0 & 0 \\ 0 & 0 & s_{43}T_{43} & s_{44}T_{44} \end{pmatrix}. \tag{34}$$

Then, the labeled graph representation of $\mathcal{B}$, $\mathcal{G}_{\mathbf{S}} = (G_{\mathbf{T}}, \mathcal{L})$, is not right-resolving. For simplicity, denoting the vertex set of $G_{\mathbf{T}}$ by $\mathcal{V} = \{1, 2, \cdots, 16\}$. Using the subset construction method to construct another labeled graph $\mathcal{H} = (H, \mathcal{L}')$ as Figure 6. Theorem 2.8 shows that $\mathbf{X}_{\mathcal{G}_{\mathbf{S}}} = \mathbf{X}_{\mathcal{H}}$.

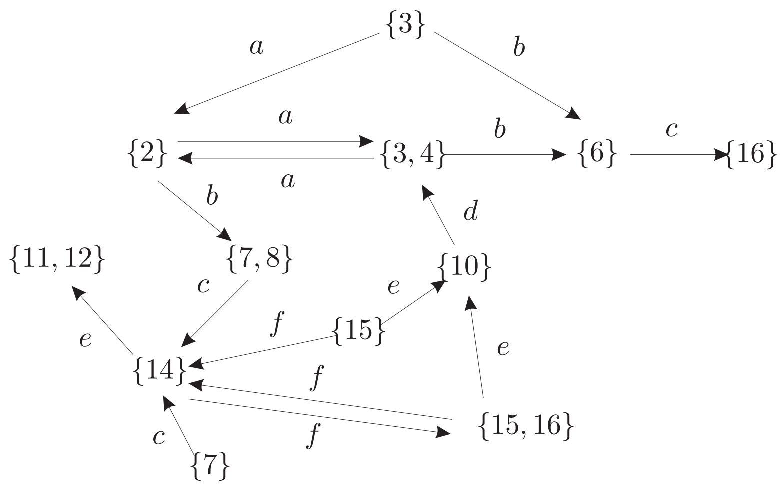


Figure 6: A right-resolving labeled graph via subset construction.

The transition matrix of $\mathbf{X}_{\mathcal{H}}$ is obtained as below.

$$\mathbf{H} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}. \tag{35}$$

Thus, $h((A,B,z);\mathcal{U}) = \log\lambda > 0$, where $\lambda \doteq 1.324718$ is the root of $f(t) = t^6 - 2t^4 + t^2 - 1$.

In the next subsection, the zeta function of Y_U will be discussed.

2.3.3 Zeta Function

Given a sofic shift Y_U with shift map σ, invariant values and invariant functions of the shift space (Y_U, σ) are interested. In the last subsection, the entropy $h(\mathcal{B})$ is studied. This subsection examines the zeta function $\zeta_\sigma(t)$ with respect to the shift map σ.

Let $p_n(\sigma) = \{y = (y_i)_{i\in\mathbb{Z}} \in Y_U | \sigma^n(y) = y\}$ be the collection of all periodic patterns of period n. The zeta function of σ is defined as

$$\zeta_\sigma(t) = \exp(\sum_{n=1}^{\infty} \frac{p_n(\sigma)}{n} t^n), \tag{36}$$

where $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ is the classical exponential function.

If a shift space $\mathbf{X}$ is a shift of finite type, then there is an edge shift $\mathbf{X}_A$ conjugate to $\mathbf{X}$. The following theorem computes the zeta function of any shift of finite type.

Theorem 2.19 ([32])**.** *Let A be a $k \times k$ nonnegative integer matrix, σ_A the associated shift map. Then*

$$\zeta_{\sigma_A}(t) = \frac{1}{\det(I_k - tA)}, \tag{37}$$

where I_k is the $k \times k$ identity matrix. Thus the zeta function of a shift of finite type is the reciprocal of a polynomial.

The following notations are needed to investigate the zeta function of sofic shift. Let $F = \{f_1, f_2, \cdots, f_m\}$ be a finite set. A permutation π of F is given below as an impression $(f_{i_1}, \cdots, f_{i_m})$, i.e., $\pi(f_\ell) = f_{i_\ell}$ for $1 \le \ell \le m$.

Definition 2.20. *A permutation π is said to be even (odd, resp.) if the number of interchanges (or transpositions) needed to generate the permutation is even (odd, resp.) Moreover, the sign of π is defined as*

$$sgn(\pi) = \begin{cases} 1, & \pi \text{ is even;} \\ -1, & \pi \text{ is odd.} \end{cases} \tag{38}$$

Let $\mathcal{H} = (H, \mathcal{L}')$ be a labeled graph which is right-resolving with r many vertices, assuming that $\mathcal{V} = \{1, 2, \cdots, r\}$. Let $\mathbf{H}_S$ denote the symbolic transition matrix of $\mathcal{H}$ and $\mathbf{H}$ the transition matrix of the underlying graph H.

For $1 \le k \le r$, constructing a labeled graph $\mathcal{H}_k$ with alphabet $\{\pm s_{ij} : s_{ij} \in \mathcal{S}\}$ as follows.

1. The vertex set of $\mathcal{H}_k$ is the set $\mathcal{V}_k$ of all subsets of $\mathcal{V}$ having k elements, i.e., $|\mathcal{V}_k| = \binom{r}{k}$. Moreover, the ordering on the states in each element of $\mathcal{V}_k$ is fixed.

2. For each $s_{ij} \in \mathcal{S}$, we denote $s_{ij}(I)$ the terminal state of s_{ij} starts at I. For $\mathcal{I} = \{I_1, \cdots, I_k\}, \mathcal{J} = \{J_1, \cdots, J_k\} \in \mathcal{V}_k$, there is an edge from $\mathcal{I}$ to $\mathcal{J}$ provided there exists $s_{ij} \in \mathcal{S}$ such that $s_{ij}(I_1), \cdots, s_{ij}(I_k)$ are well-defined and $(s_{ij}(I_1), \cdots, s_{ij}(I_k))$ is a permutation of $\mathcal{J}$. More than this, the edge is labeled as s_{ij} ($-s_{ij}$, reps.) if the permutation is even (odd, resp.) Otherwise, there is no edge with label $\pm s_{ij}$ from $\mathcal{I}$ to $\mathcal{J}$.

Definition 2.21. *Let $\mathbf{H}_{S_k}$ denote the symbolic transition matrix of $\mathcal{H}_k$, $\mathbf{H}_k$ be obtained from $\mathbf{H}_{S_k}$ by setting all the symbols in $\mathcal{S}$ equal to 1. We call $\mathbf{H}_k$ the k-th signed subset matrix of $\mathcal{H}$.*

Theorem 2.22 ([32]). *Let $\mathcal{H} = (H, \mathcal{L}')$ be a right-resolving labeled graph with r many vertices, and $\mathbf{H}_k$ be its k-th signed subset matrix. Then*

$$\zeta_{\sigma_{\mathcal{H}}}(t) = \prod_{k=1}^{r} \det(I - t\mathbf{H}_k)^{(-1)^k}, \tag{39}$$

where I is the identity matrix.

Theorem 2.23 ([32]). *If two shift spaces $\mathbf{X}$ and $\mathbf{Y}$ are conjugate, then $\zeta_{\sigma_{\mathbf{X}}}(t) = \zeta_{\sigma_{\mathbf{Y}}}(t)$.*

Therefore, the zeta function of the sofic shift $\mathbf{X}_{\mathcal{G}_{\mathbf{S}}}$ can be derived using the following theorem.

Theorem 2.24. *For a given $\mathcal{B} \subseteq \{-,+\}^{\mathbb{Z}_{3\times 2}}$, let $Y_U \equiv Y_U(\mathcal{B})$ be the shift space induced by $\mathcal{B}$ with shift map σ. Then there exists a labeled graph representation $\mathcal{H}$ such that*

$$\zeta_{\sigma}(t) = \prod_{k=1}^{n} \det(I - t\mathbf{H}_k)^{(-1)^k}, \tag{40}$$

where $\mathbf{H}_k$ is the k-th signed subset matrix of $\mathcal{H}$, and n is the cardinal number of the underlying graph H.

Proof. Let $\mathcal{G}_{\mathbf{S}}$ be the labeled graph representation of Y_U, and $\mathbf{X}_{\mathcal{G}_{\mathbf{S}}}$ be the sofic shift induced by $\mathcal{G}_{\mathbf{S}}$ with shift map $\sigma_{\mathcal{G}_{\mathbf{S}}}$. By Theorem 2.13, $\mathbf{X}_{\mathcal{G}_{\mathbf{S}}}$ is conjugate to Y_U. Thus, we have

$$\zeta_{\sigma}(t) = \zeta_{\sigma_{\mathcal{G}_{\mathbf{S}}}}(t). \tag{41}$$

If $\mathcal{G}_{\mathbf{S}}$ is right-resolving, then it is done by Theorem 2.22. Otherwise, constructing a labeled graph $\mathcal{H}$ which is right-resolving and represents the same shift space as $\mathcal{G}_{\mathbf{S}}$ via subset construction method. This completes the proof. $\square$

Example 2.25 (Continued)**.** Continuing with Example 2.5 and Example 2.18, for the investigation of zeta function, all the k-th signed subset matrix $\mathbf{H}_k$ of $\mathcal{H}$ are needed to be constructed. Therefore, $\mathbf{H}_1 = \mathbf{H}$ as in (35),

$$\mathbf{H}_2 = \begin{pmatrix} 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and $\mathbf{H}_3 = \mathbf{H}_4 = \cdots = \mathbf{H}_{12} = 0$, the zero matrix. Hence, by Theorem 2.22, $\zeta_\sigma(t) = \dfrac{(1+t)^2}{1-2t^2+t^4-t^6}$.

3 Multi-layer Cellular Neural Networks

In this section, all results in one-layer CNN with input will be extended to multi-layer CNN.

3.1 Partition of Parameter Space

As in (1), an N-layer CNN system with input is of the form,

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

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

$$u_i^{(n)} = y_i^{(n-1)} \text{ for } 2 \le n \le N, \quad u_i^{(1)} = u_i, \quad x_i(0) = x_i^0. \tag{2}$$

The feedback and controlling templates of each layer are

$$A^{(n)} = (a_{-d}^{(n)}, a_{-d+1}^{(n)}, \cdots, a_d^{(n)}) \text{ and } B^{(n)} = (b_{-d}^{(n)}, b_{-d+1}^{(n)}, \cdots, b_d^{(n)}),$$

where $1 \le n \le N$. The parameter space and the admissible local patterns of each layer can be represented by $\mathcal{P}^{(n)} = \{(A^{(n)}, B^{(n)}, z^{(n)})\}$ and $\mathcal{B}^{(n)}(A^{(n)}, B^{(n)}, z^{(n)})$, where $1 \le n \le N$. Let $A = (A^{(1)}, A^{(2)}, \cdots, A^{(N)})$, $B = (B^{(1)}, B^{(2)}, \cdots, B^{(N)})$, $z = (z^{(1)}, z^{(2)}, \cdots, z^{(N)})$, $\mathcal{P}^m = (\mathcal{P}^{(1)}, \mathcal{P}^{(2)}, \cdots, \mathcal{P}^{(N)})$, where $m = N(2d+1) - 1$, $Y^{(n)} = y_{-d}^{(n)} y_{-d+1}^{(n)} \cdots y_d^{(n)}$, where $1 \le n \le N$, and $U = u_{-d} u_{-d+1} \cdots u_d$, then

$$\mathcal{B}(A, B, z) = \left\{ \begin{array}{l} Y^{(N)} \diamond Y^{(N-1)} \diamond \cdots \diamond Y^{(1)} \diamond U : \\ Y^{(n)} \diamond Y^{(n-1)} \in \mathcal{B}^{(n)} \text{ for } 2 \le n \le N, \text{ and } Y^{(1)} \diamond U \in \mathcal{B}^{(1)} \end{array} \right\}, \tag{3}$$

where $\diamond$ is defined in (4). The generalized partition theorem of N-layer CNN then follows.

Theorem 3.1. *There exists $K(m) \in \mathbb{N}$ and unique collection of open subsets $\{P_k\}_{k=1}^{K(m)}$ of $\mathcal{P}^m$ such that*

(i) $\mathcal{P}^m = \bigcup_{k=1}^{K(m)} \bar{P}_k$.

(ii) $P_k \bigcap P_j = \varnothing$ *for* $k \neq j$.

(iii) $\mathcal{B}(A, B, z) = \mathcal{B}(\tilde{A}, \tilde{B}, \tilde{z}) \Leftrightarrow (A, B, z), (\tilde{A}, \tilde{B}, \tilde{z}) \in P_k$ *for some* k.

Proof. For simplicity, the case $N = 2$ is proved. The general case can be done analogously, the details are omitted here.

By Theorem 2.1, there exist $K_i \in \mathbb{N}$ and an unique collection of open subsets $\{P_k^{(i)}\}_{k=1}^{K_i}$ of $\mathcal{P}^{(i)}$ such that

(1) $\mathcal{P}_{(i)}^{4d+3} = \bigcup_{k=1}^{K_i} \overline{P_k^{(i)}}$.

(2) $P_k^{(i)} \bigcap P_\ell^{(i)} = \varnothing$ for $k \neq \ell$.

(3) $\mathcal{B}^{(i)}(A^{(i)}, B^{(i)}, z^{(i)}) = \mathcal{B}^{(i)}(\tilde{A}^{(i)}, \tilde{B}^{(i)}, \tilde{z}^{(i)})$ if and only if

$$(A^{(i)}, B^{(i)}, z^{(i)}),\ (\tilde{A}^{(i)}, \tilde{B}^{(i)}, \tilde{z}^{(i)}) \in P_k^{(i)} \text{ for some } k,$$

where $i = 1, 2$. Let $K' = K_1 \cdot K_2$, define $P'_k = (P_{k_1}^{(1)}, P_{k_2}^{(2)})$, where $k = (k_1 - 1)K_2 + k_2,\ 1 \leq k_1 \leq K_1, 1 \leq k_2 \leq K_2$. Let

$$P_1 = P'_i, \text{ where } i = \min\{k |\text{ there exists } Y^{(2)} \diamond Y^{(1)} \in \mathcal{B}^{(2)} \text{ and } Y^{(1)} \diamond U \in \mathcal{B}^{(1)}\}, \tag{4}$$

and

$$P_\ell = P'_i, \text{ where } i = \min_{k > k_{\ell-1}} \{k |\text{ there exists } Y^{(2)} \diamond Y^{(1)} \in \mathcal{B}^{(2)} \text{ and } Y^{(1)} \diamond U \in \mathcal{B}^{(1)}\}, \tag{5}$$

for $\ell \geq 2$ and $k_j \in \mathbb{N}$ such that $P'_{k_j} = P_j$. Then there exists an positive integer $K \leq K'$ such that $\{P_k\}_{k=1}^K$ satisfies (i), (ii), (iii), then the proof is completed. $\square$

3.2 Ordering Matrix

The ordering matrix $\mathbb{X}_{3\times N}$ of all possible local patterns in $\{+,-\}^{\mathbb{Z}_{3\times N}}$ is defined recursively as

$$\mathbb{X}_{3\times N}=\begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} & \varnothing & \varnothing \\ \varnothing & \varnothing & \mathbf{X}_{23} & \mathbf{X}_{24} \\ \mathbf{X}_{31} & \mathbf{X}_{32} & \varnothing & \varnothing \\ \varnothing & \varnothing & \mathbf{X}_{43} & \mathbf{X}_{44} \end{bmatrix}, \tag{6}$$

where

$$\mathbf{X}_{i_1j_1}=\begin{bmatrix} X_{i_1j_1;11} & X_{i_1j_1;12} & \varnothing & \varnothing \\ \varnothing & \varnothing & X_{i_1j_1;23} & X_{i_1j_1;24} \\ X_{i_1j_1;31} & X_{i_1j_1;32} & \varnothing & \varnothing \\ \varnothing & \varnothing & X_{i_1j_1;43} & X_{i_1j_1;44} \end{bmatrix}, \tag{7}$$

$$X_{i_1j_1;i_2j_2;\cdots;i_kj_k}=$$
$$\begin{bmatrix} X_{i_1j_1;i_2j_2;\cdots;i_kj_k;11} & X_{i_1j_1;i_2j_2;\cdots;i_kj_k;12} & \varnothing & \varnothing \\ \varnothing & \varnothing & X_{i_1j_1;i_2j_2;\cdots;i_kj_k;23} & X_{i_1j_1;i_2j_2;\cdots;i_kj_k;24} \\ X_{i_1j_1;i_2j_2;\cdots;i_kj_k;31} & X_{i_1j_1;i_2j_2;\cdots;i_kj_k;32} & \varnothing & \varnothing \\ \varnothing & \varnothing & X_{i_1j_1;i_2j_2;\cdots;i_kj_k;43} & X_{i_1j_1;i_2j_2;\cdots;i_kj_k;44} \end{bmatrix}, \tag{8}$$

for $1\le k\le N-2$, and

$$X_{i_1j_1;i_2j_2;\cdots;i_{N-1}j_{N-1}}=$$
$$\begin{bmatrix} x_{i_1j_1;\cdots;i_{N-1}j_{N-1};11} & x_{i_1j_1;\cdots;i_{N-1}j_{N-1};12} & \varnothing & \varnothing \\ \varnothing & \varnothing & x_{i_1j_1;\cdots;i_{N-1}j_{N-1};23} & x_{i_1j_1;\cdots;i_{N-1}j_{N-1};24} \\ x_{i_1j_1;\cdots;i_{N-1}j_{N-1};31} & x_{i_1j_1;\cdots;i_{N-1}j_{N-1};32} & \varnothing & \varnothing \\ \varnothing & \varnothing & x_{i_1j_1;\cdots;i_{N-1}j_{N-1};43} & x_{i_1j_1;\cdots;i_{N-1}j_{N-1};44} \end{bmatrix}, \tag{9}$$

where $1\le i_k,j_k\le 4$, and $1\le k\le N$. The construction contains a self-similarity property in $\mathbb{X}_{3\times N}$. As in Subsection 2.1.2, $x_{i_1j_1;i_2j_2;\cdots;i_{N-1}j_{N-1};i_Nj_N}$

means the pattern

$$(a_{r_{11}r_{12}}a_{r'_{12}r_{13}}) \diamond (a_{r_{21}r_{22}}a_{r'_{22}r_{23}}) \diamond \cdots \diamond (a_{r_{N1}r_{N2}}a_{r'_{N2}r_{N3}})$$

in $\{+,-\}^{\mathbb{Z}_{3\times N}}$, where $a_{r_{k1}r_{k2}}a_{r'_{k2}r_{k3}}$ is defined in (10), and

$$r_{k1} = \left[\frac{i_k-1}{2}\right], \quad r_{k2} = i_k - 1 - 2r_{k1}, \quad r'_{k2} = \left[\frac{j_k-1}{2}\right], \quad r_{k3} = j_k - 1 - 2r'_{k2}.$$

The pattern is $\varnothing$ if $a_{r_{k1}r_{k2}}a_{r'_{k2}r_{k3}} = \varnothing$ for some $1 \leq k \leq N$. Otherwise, it is denoted by the pattern

$$(a_{r_{11}}a_{r_{12}}a_{r_{13}}) \diamond (a_{r_{21}}a_{r_{22}}a_{r_{23}}) \diamond \cdots \diamond (a_{r_{N1}}a_{r_{N2}}a_{r_{N3}})$$

in $\{+,-\}^{\mathbb{Z}_{3\times\infty}}$.

As long as the basic set of the admissible local patterns $\mathcal{B} \subseteq \{+,-\}^{\mathbb{Z}_{3\times(N+1)}}$ is given, $\Sigma_m(\mathcal{B})$ denotes the collection of all m-blocks generated by $\mathcal{B}$. The subshift space of $\{+,-\}^{\mathbb{Z}}$ is then defined by

$$Y_U = \left\{ \begin{array}{c} Y^{(N)} = (y_i^{(N)})_{i\in\mathbb{Z}} : \text{ there exist } U, Y^{(1)}, Y^{(2)}, \cdots, Y^{(N-1)} \\ \text{such that } Y^{(N)} \diamond Y^{(N-1)} \diamond \cdots \diamond Y^{(1)} \diamond U \in \Sigma(\mathcal{B}) \end{array} \right\}, \qquad (10)$$

where $\Sigma(\mathcal{B}) \subseteq \{+,-\}^{\mathbb{Z}_{\infty\times(N+1)}}$ is generated by $\mathcal{B} \subseteq \{+,-\}^{\mathbb{Z}_{3\times(N+1)}}$.

3.3 Transition Matrix

The basic set of admissible local patterns $\mathcal{B} = \mathcal{B}(A, B, z)$ can be determined from the N-layer CNN parameters (A, B, z). Denote by T_n the transition matrix induced by $\mathcal{B}^{(n)} \subseteq \{+,-\}^{\mathbb{Z}_{3\times 2}}$, where $\mathcal{B}^{(n)}$ is the basic set of admissible local patterns in the n-th layer, and $1 \leq n \leq N$. Let $\widehat{\mathbf{T}}_N = \mathbf{T}(\mathcal{B};\mathcal{U})$ be the transition matrix induced by $\mathcal{B}$ with the set of input patterns $\mathcal{U}$. The following theorem is then obtained.

Theorem 3.2.

$$\widehat{\mathbf{T}}_N = (T_N \otimes E_{4^{N-1}}) \circ (E_4 \otimes \overline{\mathbf{T}}_{N-1}) \in \mathcal{M}_{4^{n+1}\times 4^{n+1}}(\mathbb{R}), \qquad (11)$$

where

$$\overline{\mathbf{T}}_n = (T_n \otimes E_{4^{n-1}}) \circ (E_4 \otimes \overline{\mathbf{T}}_{n-1}) \in \mathcal{M}_{4^{n+1}\times 4^{n+1}}(\mathbb{R}), \quad \textit{for } 2 \le n \le N-1, \tag{12}$$

and

$$\overline{\mathbf{T}}_1 = T_1 \circ (E_4 \otimes \mathbf{U}) \in \mathcal{M}_{16\times 16}(\mathbb{R}), \tag{13}$$

$\mathbf{U}$ *is the transition matrix of* $\mathcal{U}$*. Hence* $\overline{\mathbf{T}}_1$ *is the transition matrix given in Theorem 2.3.*

In particular, if $N = 2$*,*

$$\widehat{\mathbf{T}}_2 = (T_2 \otimes E_4) \circ (E_4 \otimes (T_1 \circ (E_4 \otimes \mathbf{U}))). \tag{14}$$

Proof. For simplicity, the case $N = 2$ is proved. For $N \ge 2$, it can be done by mathematical induction, thus is omitted.

Denoting $\widehat{\mathbf{T}}_2 = (\widehat{T}_{i_1j_1})_{1\le i_1,j_1\le 4}$ and $T_2 = (T_{i_1j_1})_{1\le i_1,j_1\le 4}$, where $\widehat{T}_{i_1j_1} \in \mathcal{M}_{16\times 16}(\mathbb{R})$ and $T_{i_1j_1} \in \mathcal{M}_{4\times 4}(\mathbb{R})$ for $1 \le i_1, j_1 \le 4$. The case $i_1 = j_1 = 1$ is proved, the others can be treated analogously.

Denoting $\widehat{T}_{11} = (\widehat{T}_{11;i_2j_2})_{1\le i_2,j_2\le 4}$, where $\widehat{T}_{11;i_2j_2} = (\hat{t}_{11;i_2j_2;i_3j_3})_{1\le i_3,j_3\le 4} \in \mathcal{M}_{4\times 4}(\mathbb{R})$, for fixed $1 \le i_2, j_2 \le 4$, and $T_{11} = (t_{11;i_2j_2})_{1\le i_2,j_2\le 4} \in \mathcal{M}_{4\times 4}(\mathbb{R})$. Since the output patterns of the first layer will be treated as the input patterns of the second layer, let $\mathcal{U}_2$ be the output patterns of the first layer coupled with input $\mathcal{U}$. By Theorem 2.3, the transition matrix of $\mathcal{U}_2$ is

$$\overline{\mathbf{T}}_1 = T_1 \circ (E_4 \otimes \mathbf{U}) \in \mathcal{M}_{16\times 16}(\mathbb{R}). \tag{15}$$

Denoting

$$\overline{\mathbf{T}}_1 = (\overline{T}_{i_2j_2})_{1\le i_2,j_2\le 4}, \quad \overline{T}_{i_2j_2} = (\bar{t}_{i_2j_2;i_3j_3})_{1\le i_3,j_3\le 4} \in \mathcal{M}_{4\times 4}(\mathbb{R}). \tag{16}$$

Then

$$\hat{t}_{11;i_2j_2;i_3j_3} = 1 \Leftrightarrow t_{11;i_2j_2} = 1 \text{ and } \bar{t}_{i_2j_2;i_3j_3} = 1, \tag{17}$$

for $1 \leq i_2, j_2, i_3, j_3 \leq 4$. That is,

$$\widehat{T}_{11} = (T_{11} \otimes E_4) \circ (T_1 \circ (E_4 \otimes \mathbf{U})). \tag{18}$$

The proof is completed. $\square$

3.4 Entropy and Zeta Function

This subsection introduces the formula for calculating entropy and zeta function of N-layer CNN. Let $\mathcal{S}^{(n)} = \{s_{ij}^{(n)}\}_{1 \leq i,j \leq 4}$ be the alphabets, and let $\mathbf{S}_n$ and $\mathbf{S}$ be the symbolic transition matrices of T_n over $\mathcal{S}^{(n)}$ and $\widehat{\mathbf{T}}_N$ for $1 \leq n \leq N$. By Theorem 2.9, $\mathbf{X}_{\mathcal{G}_{\mathbf{S}_n}}$ is a sofic shift induced by $\mathcal{B}^{(n)}$, where $\mathcal{G}_{\mathbf{S}_n}$ is the labeled graph representation of the n-th layer. Furthermore, Y_U is the output space induced by the N-layer CNN as defined in (10). The following theorem can be obtained by the same method in Theorem 2.13, so the details are omitted.

Theorem 3.3. *Y_U is conjugate to $\mathbf{X}_{\mathcal{G}_{\mathbf{S}}}$.*

The definition of convolution is given below.

Definition 3.4. *Let $\mathbf{X}, \mathbf{Y}$ be two shift spaces with graph representation $G_{\mathbf{X}} = (\mathcal{V}_{\mathbf{X}}, \mathcal{E}_{\mathbf{X}})$, $G_{\mathbf{Y}} = (\mathcal{V}_{\mathbf{Y}}, \mathcal{E}_{\mathbf{Y}})$, resp., then the convolution of $\mathbf{X}, \mathbf{Y}$, denoted by $\mathbf{X} * \mathbf{Y}$, is the shift space with underlying graph $G_{\mathbf{X}*\mathbf{Y}} = (\mathcal{V}_{\mathbf{X}*\mathbf{Y}}, \mathcal{E}_{\mathbf{X}*\mathbf{Y}})$, where*

$$\mathcal{V}_{\mathbf{X}*\mathbf{Y}} = \{f(x) \in \mathcal{E}_{\mathbf{Y}} |\ x \in \mathcal{V}_{\mathbf{X}}\} \tag{19}$$

for some $f : \mathcal{V}_{\mathbf{X}} \to \mathcal{E}_{\mathbf{Y}}$.

The convolution theorem for an N-many sofic shift is then obtained.

Theorem 3.5. *Let $\mathbf{X}_{\mathcal{G}_{\mathbf{S}}}$ be the sofic shift induced by $\mathcal{B}$, then*

$$\mathbf{X}_{\mathcal{G}_{\mathbf{S}}} = \mathbf{X}_{\mathcal{G}_{\mathbf{S}_N}} * \cdots * \mathbf{X}_{\mathcal{G}_{\mathbf{S}_2}} * \mathbf{X}_{\mathcal{G}_{\mathbf{S}_1}} \tag{20}$$

is the convolution of $\mathbf{X}_{\mathcal{G}_{\mathbf{S}_1}}, \cdots, \mathbf{X}_{\mathcal{G}_{\mathbf{S}_N}}$,

$$\widehat{\mathbf{S}}_N = (\mathbf{S}_N \otimes E_{4^{N-1}}) \circ (E_4 \otimes \overline{\mathbf{S}}_{N-1}), \tag{21}$$

where

$$\overline{\mathbf{S}}_n = (\mathbf{S}_n \otimes E_{4^{n-1}}) \circ (E_4 \otimes \overline{\mathbf{S}}_{n-1}) \in \mathcal{M}_{4^{n+1} \times 4^{n+1}}(\mathbb{R}), \quad \text{for } 2 \leq n \leq N-1, \tag{22}$$

and

$$\overline{\mathbf{S}}_1 = \mathbf{S}_1 \circ (E_4 \otimes \mathbf{U}) \in \mathcal{M}_{16}(\mathbb{R}). \tag{23}$$

Proof. This can be done using the same method used in the proof of Theorem 3.2, the details are omitted. □

Thus, the theorems for entropy and zeta function can be found via the same methods as described in the last section.

Theorem 3.6. *For a given* $\mathcal{B} \subseteq \{+,-\}^{\mathbb{Z}_{3\times(N+1)}}$*, let* $Y_U \equiv Y_U(\mathcal{B})$ *be the shift space induced by* $\mathcal{B}$*. Then there exists a labeled graph representation* $\mathcal{H} = (H, \mathcal{L}')$ *such that*

$$h(Y_U) = h(\mathbf{X}_{\mathcal{H}}) = \log \rho(\mathbf{H}), \tag{24}$$

and

$$\zeta_\sigma(t) = \prod_{k=1}^{r} \det(I - t\mathbf{H}_k)^{(-1)^k}, \tag{25}$$

where $\mathbf{H}_k$ *is the* k*-th signed subset matrix of* $\mathcal{H}$*, and* r *is the cardinal number of the underlying graph* H*.*

An example for 2-layer CNN is illustrated here.

Example 3.7. Consider (A, B, z) with $A^{(1)} = A^{(2)} \equiv \bar{A}$, $B^{(1)} = B^{(2)} \equiv \bar{B}$, $z^{(1)} = z^{(2)} \equiv \bar{z}$, and $\bar{A}$, $\bar{B}$ and $\bar{z}$ satisfy the same condition described in Example 2.5. Moreover, the set of input patterns is given by $\mathcal{U} = \{- +$

$-,-++,+-+\}$. Then $\mathcal{B}^{(1)} = \mathcal{B}(A^{(1)}, B^{(1)}, z^{(1)}; \mathcal{U})$ is consisting of the following patterns.

$$\begin{array}{ccc} - & \ominus & - \\ - & \boxplus & - \end{array} \qquad \begin{array}{ccc} - & \ominus & - \\ - & \boxplus & + \end{array} \qquad \begin{array}{ccc} - & \ominus & - \\ + & \boxminus & + \end{array} \qquad \begin{array}{ccc} - & \ominus & + \\ - & \boxplus & - \end{array}$$

$$\begin{array}{ccc} - & \ominus & + \\ - & \boxplus & + \end{array} \qquad \begin{array}{ccc} - & \ominus & + \\ + & \boxminus & + \end{array} \qquad \begin{array}{ccc} + & \ominus & - \\ - & \boxplus & - \end{array} \qquad \begin{array}{ccc} + & \ominus & - \\ - & \boxplus & + \end{array}$$

$$\begin{array}{ccc} + & \oplus & + \\ + & \boxminus & + \end{array} \qquad \begin{array}{ccc} + & \oplus & + \\ - & \boxplus & + \end{array} \qquad \begin{array}{ccc} + & \oplus & + \\ - & \boxplus & - \end{array} \qquad \begin{array}{ccc} + & \oplus & - \\ + & \boxminus & + \end{array}$$

$$\begin{array}{ccc} + & \oplus & - \\ - & \boxplus & + \end{array} \qquad \begin{array}{ccc} + & \oplus & - \\ - & \boxplus & - \end{array} \qquad \begin{array}{ccc} - & \oplus & + \\ + & \boxminus & + \end{array} \qquad \begin{array}{ccc} - & \oplus & + \\ - & \boxplus & + \end{array}$$

Denote $\mathcal{U}_2$ the output patterns of $\mathcal{B}^{(1)}$, i.e.,

$$\mathcal{U}_2 = \{---,--+,+--,+++,++-,-++\}.$$

Then $\mathcal{B}^{(2)} = \mathcal{B}(A^{(2)}, B^{(2)}, z^{(2)}; \mathcal{U}_2)$ is consisting of the following patterns.

$$\begin{array}{ccc} + & \oplus & + \\ + & \boxplus & + \end{array} \qquad \begin{array}{ccc} + & \oplus & + \\ + & \boxplus & - \end{array} \qquad \begin{array}{ccc} + & \oplus & + \\ + & \boxminus & - \end{array} \qquad \begin{array}{ccc} + & \oplus & - \\ + & \boxplus & + \end{array}$$

$$\begin{array}{ccc} + & \oplus & - \\ + & \boxplus & - \end{array} \qquad \begin{array}{ccc} + & \oplus & + \\ - & \boxplus & + \end{array} \qquad \begin{array}{ccc} + & \oplus & - \\ + & \boxminus & - \end{array} \qquad \begin{array}{ccc} - & \oplus & + \\ + & \boxplus & + \end{array}$$

$$\begin{array}{ccc} + & \oplus & + \\ - & \boxminus & + \end{array} \qquad \begin{array}{ccc} - & \oplus & + \\ + & \boxplus & - \end{array} \qquad \begin{array}{ccc} + & \oplus & + \\ - & \boxminus & - \end{array} \qquad \begin{array}{ccc} + & \oplus & - \\ - & \boxplus & + \end{array}$$

$$\begin{array}{ccc} - & \oplus & + \\ + & \boxminus & - \end{array} \qquad \begin{array}{ccc} - & \oplus & - \\ + & \boxplus & + \end{array} \qquad \begin{array}{ccc} + & \oplus & - \\ - & \boxminus & + \end{array} \qquad \begin{array}{ccc} - & \oplus & - \\ + & \boxplus & - \end{array}$$

$$\begin{array}{ccc} + & \oplus & - \\ - & \boxminus & - \end{array} \qquad \begin{array}{ccc} - & \oplus & + \\ - & \boxplus & + \end{array} \qquad \begin{array}{ccc} - & \ominus & - \\ - & \boxminus & - \end{array} \qquad \begin{array}{ccc} - & \ominus & - \\ - & \boxminus & + \end{array}$$

$$\begin{array}{ccc} - & \ominus & - \\ - & \boxplus & + \end{array} \qquad \begin{array}{ccc} - & \ominus & + \\ - & \boxminus & - \end{array} \qquad \begin{array}{ccc} - & \ominus & + \\ - & \boxminus & + \end{array} \qquad \begin{array}{ccc} - & \ominus & - \\ + & \boxminus & - \end{array}$$

$$\begin{array}{ccc} - & \ominus & + \\ - & \boxplus & + \end{array} \qquad \begin{array}{ccc} + & \ominus & - \\ - & \boxminus & - \end{array} \qquad \begin{array}{ccc} - & \ominus & - \\ + & \boxplus & - \end{array} \qquad \begin{array}{ccc} + & \ominus & - \\ - & \boxminus & + \end{array}$$

$$\begin{array}{ccc} - & \ominus & - \\ + & \boxplus & + \end{array} \qquad \begin{array}{ccc} - & \ominus & + \\ + & \boxminus & - \end{array} \qquad \begin{array}{ccc} + & \ominus & - \\ - & \boxplus & + \end{array}$$

The transition matrix $\widehat{\mathbf{T}} = \mathbf{T}((A,B,z);\mathcal{U})$ is then

$$\widehat{\mathbf{T}} = \begin{pmatrix} \widehat{\mathbf{T}}_{11} & \widehat{\mathbf{T}}_{12} & 0 & 0 \\ 0 & 0 & \widehat{\mathbf{T}}_{23} & \widehat{\mathbf{T}}_{24} \\ \widehat{\mathbf{T}}_{31} & 0 & 0 & 0 \\ 0 & 0 & \widehat{\mathbf{T}}_{43} & \widehat{\mathbf{T}}_{44} \end{pmatrix}, \tag{26}$$

where

$$\widehat{\mathbf{T}}_{11} = \widehat{\mathbf{T}}_{43} = \widehat{\mathbf{T}}_{44} = \begin{pmatrix} T_1 & T_1 & 0 & 0 \\ 0 & 0 & 0 & T_3 \\ T_2 & 0 & 0 & 0 \\ 0 & 0 & T_1 & T_1 \end{pmatrix}, \quad \widehat{\mathbf{T}}_{12} = \begin{pmatrix} T_1 & T_1 & 0 & 0 \\ 0 & 0 & 0 & T_3 \\ T_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\widehat{\mathbf{T}}_{23} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & T_1 & T_1 \end{pmatrix}, \ \widehat{\mathbf{T}}_{24} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & T_3 \\ T_2 & 0 & 0 & 0 \\ 0 & 0 & T_1 & T_1 \end{pmatrix}, \ \widehat{\mathbf{T}}_{31} = \begin{pmatrix} T_1 & T_1 & 0 & 0 \\ 0 & 0 & 0 & T_3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and

$$T_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, T_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, T_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Let $\mathcal{S} = \{s_{11}, s_{12}, s_{23}, s_{24}, s_{31}, s_{43}, s_{44}\}$, the symbolic transition matrix is

$$\mathbf{S} = \begin{pmatrix} s_{11}\widehat{\mathbf{T}}_{11} & s_{12}\widehat{\mathbf{T}}_{12} & 0 & 0 \\ 0 & 0 & s_{23}\widehat{\mathbf{T}}_{23} & s_{24}\widehat{\mathbf{T}}_{24} \\ s_{31}\widehat{\mathbf{T}}_{31} & 0 & 0 & 0 \\ 0 & 0 & s_{43}\widehat{\mathbf{T}}_{43} & s_{44}\widehat{\mathbf{T}}_{44} \end{pmatrix}, \tag{27}$$

which is not right-resolving. Using subset construction method, the spatial entropy then can be found, $h((A,B,z);\mathcal{U}) = \log\lambda$, where $\lambda \doteq 1.49676$ is a root of $f(t) = t^8 - 2t^6 + t^4 - 3t^2 - 1$. Moreover, the zeta function is $\zeta_\sigma(t) = \dfrac{(1+t+t^3)(1+t-t^3)}{1-2t^2+t^4-3t^6-t^8}$.

3.5 The Broken of Symmetry

The basic set of admissible local patterns $\mathcal{B}$ can be determined from (A, B, z). The entropy of each partition is symmetrical in one-dimensional CNN without input, i.e., where $B \equiv 0$ [24]. For example, if (A, z) is picked such that $a_l > a_r > 0$, then parameters a and z have 25 regions. Clearly,

$$h(\mathcal{B}([m,n])) = h(\mathcal{B}([n,m])), \text{ for } 1 \leq m, n \leq 4. \tag{28}$$

The symmetry is broken for the one-layer CNN with input, as shown below with an example.

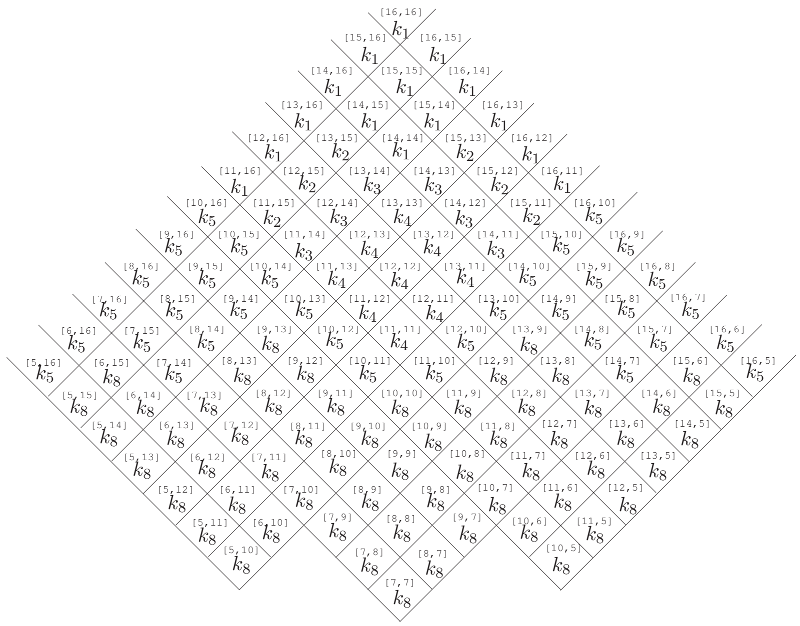
Consider

$$\frac{dx_i}{dt} = -x_i + a_l y_{i-1} + a y_i + a_r y_{i+1} + b_l u_{i-1} + b u_i + b_r u_{i+1} + z, \tag{29}$$

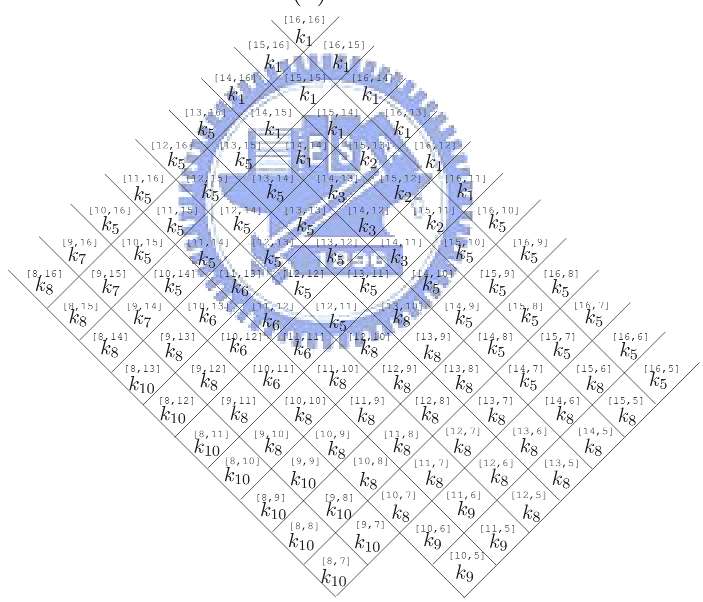
where $b_l = 0$, then the symmetry of entropy is broken, as revealed in Figure 7.

Table 1: Some maximal eigenvalues produced in one-layer CNN with input.

maximal eigenvalue	characteristic polynomial
$\lambda_1 = 2$	$t - 2$
$\lambda_2 \doteq 1.9479$	$t^5 - 2t^4 + t^3 - 2t^2 + t - 1$
$\lambda_3 \doteq 1.8832$	$t^4 - 2t^3 + t^2 - 2t + 1$
$\lambda_4 \doteq 1.8393$	$t^3 - t^2 - t + 1$
$\lambda_5 \doteq 1.7549$	$t^3 - 2t^2 + t - 1$
$\lambda_6 \doteq 1.7417$	$t^8 - 2t^7 + t^6 - t^5 + t^4 - 2t^3 + t^2 - 1$
$\lambda_7 \doteq 1.6992$	$t^5 - 2t^4 + t^3 - 2t + 1$
$\lambda_8 = g \doteq 1.618$	$t^2 - t - 1$
$\lambda_9 \doteq 1.5618$	$t^6 - 2t^5 + t^4 - t^2 + t - 1$
$\lambda_{10} \doteq 1.5289$	$t^5 - 2t^4 + t^3 - 1$



(a)



(b)

Figure 7: The effect of input patterns. The parameters a_l, a_r, b, b_r are considered as follows. (i) $a_l > a_r > b > b_r > 0$, (ii) $a_l < b + b_r$, (iii) $a_l + b_r < a_r + b$. Subfigure (a) lists regions that produce positive entropy. Those regions with positive entropy are symmetric, i.e., $h([m,n]) = h([n,m])$. However, such property would be destroyed when input patterns are given. Subfigure (b) lists the same regions as in (a) but the input patterns $\mathcal{U} = \{--, -+, +-\}$ are considered. It is seen that the symmetry is no longer hold. Herein, $k_i = \log \lambda_i$ for $1 \leq i \leq 10$ are listed in Table 1.

4 Study of an Example

This section introduces in detail the relationship between the admissible local patterns and the partition of parameter space in Example 2.5. A one-dimensional CNN with input is of the form,

$$\frac{dx_i}{dt} = -x_i + a_l y_{i-1} + a y_i + a_r y_{i+1} + b_l u_{i-1} + b u_i + b_r u_{i+1} + z, \tag{30}$$

where $A = [a_l, a, a_r], B = [b_l, b, b_r]$ represent the feedback and controlling templates, respectively; $y = f(x) = \frac{1}{2}(|x+1| - |x-1|)$ is the output function, and z is the threshold.

Consider that a_l, a_r, b_l, b, b_r satisfies the inequality in Example 2.5, i.e., the partition for the parameter space $\{(a_l, a_r, b_l, b, b_r)\}$ is chosen. For a given mosaic solution $\bar{x}$, the state at cell C_i is $+$, i.e., $\bar{x}_i > 1$, if and only if

$$a - 1 + z > -(a_l y_{i-1} + a_r y_{i+1} + b u_i + b_r u_{i+1}). \tag{31}$$

Similarly, the state at cell C_i is $-$, i.e., $\bar{x}_i < -1$, if and only if

$$a - 1 - z > a_l y_{i-1} + a_r y_{i+1} + b u_i + b_r u_{i+1}. \tag{32}$$

Let $\alpha = (a_l, a_r)$, $\beta = (b_l, b, b_r)$, $V^n = \{v = (v_i) \in \mathbb{R}^n : |v_i| = 1 \text{ for all } 1 \leq i \leq n\}$, the basic set of admissible local patterns with "+" state in the center is defined as

$$\mathcal{B}(+, A, B, z) = \{v \diamond w : a - 1 + z > -(\alpha \cdot v + \beta \cdot w)\}, \tag{33}$$

where $v \in V^2$ and $w \in V^3$. Similarly, the basic set of admissible local patterns with "$-$" state in the center is defined as

$$\mathcal{B}(-, A, B, z) = \{v' \diamond w' : a - 1 - z > \alpha \cdot v' + \beta \cdot w'\}. \tag{34}$$

Furthermore, the basic set of admissible local patterns derived from (A, B, z) is denoted by

$$\mathcal{B}(A, B, z) = (\mathcal{B}(+, A, B, z), \mathcal{B}(-, A, B, z)). \tag{35}$$

Let ℓ_i^+, ℓ_j^- denote the linear maps

$$a - 1 + z = c_i^+ \text{ and } a - 1 - z = c_j^-$$

for some $c_i^+, c_j^-, 1 \leq i, j \leq 32$, respectively. By the condition (i) $\sim$ (iv) in Example 2.5, the following relation can be obtained.

$$c_1^- < c_2^- < \cdots < c_{16}^- < 0 < c_{17}^- < c_{18}^- < \cdots < c_{32}^-, \tag{36}$$

and $c_k^- = -c_{33-k}^+, 1 \leq k \leq 32$, where

$$\begin{aligned}
c_1^- &= -a_l - a_r - b_l - b - b_r, & c_{17}^- &= a_l - a_r - b_l + b - b_r,\\
c_2^- &= -a_l - a_r - b_l - b + b_r, & c_{18}^- &= a_l - a_r - b_l + b + b_r,\\
c_3^- &= -a_l - a_r - b_l + b - b_r, & c_{19}^- &= a_l + a_r - b_l - b - b_r,\\
c_4^- &= -a_l - a_r - b_l + b + b_r, & c_{20}^- &= -a_l + a_r + b_l + b - b_r,\\
c_5^- &= -a_l + a_r - b_l - b - b_r, & c_{21}^- &= a_l + a_r - b_l - b + b_r,\\
c_6^- &= -a_l + a_r - b_l - b + b_r, & c_{22}^- &= -a_l + a_r + b_l + b + b_r,\\
c_7^- &= -a_l - a_r + b_l - b - b_r, & c_{23}^- &= a_l - a_r + b_l - b - b_r,\\
c_8^- &= -a_l - a_r + b_l - b + b_r, & c_{24}^- &= a_l - a_r + b_l - b + b_r,\\
c_9^- &= -a_l + a_r - b_l + b - b_r, & c_{25}^- &= a_l + a_r - b_l + b - b_r,\\
c_{10}^- &= -a_l + a_r - b_l + b + b_r, & c_{26}^- &= a_l + a_r - b_l + b + b_r,\\
c_{11}^- &= a_l - a_r - b_l - b - b_r, & c_{27}^- &= a_l - a_r + b_l + b - b_r,\\
c_{12}^- &= -a_l - a_r + b_l + b - b_r, & c_{28}^- &= a_l - a_r + b_l + b + b_r,\\
c_{13}^- &= a_l - a_r - b_l - b + b_r, & c_{29}^- &= a_l + a_r + b_l - b - b_r,\\
c_{14}^- &= -a_l - a_r + b_l + b + b_r, & c_{30}^- &= a_l + a_r + b_l - b + b_r,\\
c_{15}^- &= -a_l + a_r + b_l - b - b_r, & c_{31}^- &= a_l + a_r + b_l + b - b_r,\\
c_{16}^- &= -a_l + a_r + b_l - b + b_r, & c_{32}^- &= a_l + a_r + b_l + b + b_r.
\end{aligned}$$

Figure 8 depicts a bifurcation diagram of the $(a-1, z)$ parameter space. The basic set of admissible local patterns is then determined once the parameters a and z are chosen, i.e., some specified region in the bifurcation diagram is selected. More precisely, if a region $[m, n]$ in the bifurcation diagram has been chosen, then for $y \diamond u \in \mathcal{B}(+, A, B, z)$, the admissible local pattern with

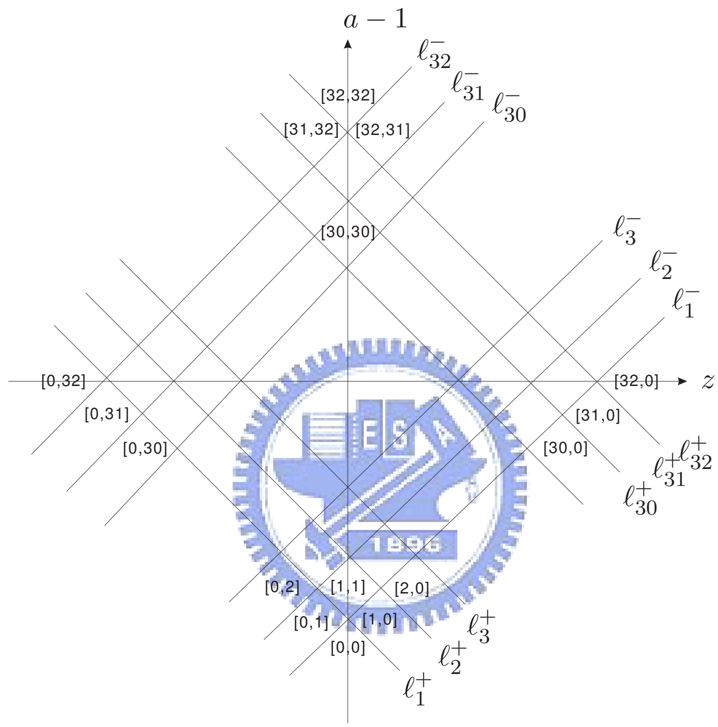


Figure 8: The $(a-1, z)$ bifurcation diagram.

$$
\begin{array}{llll}
1: \begin{array}{ccc} + & \oplus & + \\ + & \boxplus & + \end{array} &
2: \begin{array}{ccc} + & \oplus & + \\ + & \boxplus & - \end{array} &
3: \begin{array}{ccc} + & \oplus & + \\ + & \boxminus & + \end{array} &
4: \begin{array}{ccc} + & \oplus & + \\ + & \boxminus & - \end{array} \\
5: \begin{array}{ccc} + & \oplus & - \\ + & \boxplus & + \end{array} &
6: \begin{array}{ccc} + & \oplus & - \\ + & \boxplus & - \end{array} &
7: \begin{array}{ccc} + & \oplus & + \\ - & \boxplus & + \end{array} &
8: \begin{array}{ccc} + & \oplus & + \\ - & \boxplus & - \end{array} \\
9: \begin{array}{ccc} + & \oplus & - \\ + & \boxminus & + \end{array} &
10: \begin{array}{ccc} + & \oplus & - \\ + & \boxminus & - \end{array} &
11: \begin{array}{ccc} - & \oplus & + \\ + & \boxplus & + \end{array} &
12: \begin{array}{ccc} + & \oplus & + \\ - & \boxminus & + \end{array} \\
13: \begin{array}{ccc} - & \oplus & + \\ + & \boxplus & - \end{array} &
14: \begin{array}{ccc} + & \oplus & + \\ - & \boxminus & - \end{array} &
15: \begin{array}{ccc} + & \oplus & - \\ - & \boxplus & + \end{array} &
16: \begin{array}{ccc} + & \oplus & - \\ - & \boxplus & - \end{array} \\
17: \begin{array}{ccc} - & \oplus & + \\ + & \boxminus & + \end{array} &
18: \begin{array}{ccc} - & \oplus & + \\ + & \boxminus & - \end{array} &
19: \begin{array}{ccc} - & \oplus & - \\ + & \boxplus & + \end{array} &
20: \begin{array}{ccc} + & \oplus & - \\ - & \boxminus & + \end{array} \\
21: \begin{array}{ccc} - & \oplus & - \\ + & \boxplus & - \end{array} &
22: \begin{array}{ccc} + & \oplus & - \\ - & \boxminus & - \end{array} &
23: \begin{array}{ccc} - & \oplus & + \\ - & \boxplus & + \end{array} &
24: \begin{array}{ccc} - & \oplus & + \\ - & \boxplus & - \end{array} \\
25: \begin{array}{ccc} - & \oplus & - \\ + & \boxminus & + \end{array} &
26: \begin{array}{ccc} - & \oplus & - \\ + & \boxminus & - \end{array} &
27: \begin{array}{ccc} - & \oplus & + \\ - & \boxminus & + \end{array} &
28: \begin{array}{ccc} - & \oplus & + \\ - & \boxminus & - \end{array} \\
29: \begin{array}{ccc} - & \oplus & - \\ - & \boxplus & + \end{array} &
30: \begin{array}{ccc} - & \oplus & - \\ - & \boxplus & - \end{array} &
31: \begin{array}{ccc} - & \oplus & - \\ - & \boxminus & + \end{array} &
32: \begin{array}{ccc} - & \oplus & - \\ - & \boxminus & - \end{array}
\end{array}
$$

Figure 9: The order of the appearance of the patterns with state "+" in the center.

state "+" in the center, $y \diamond u$ satisfies the following inequalities.

$$a - 1 + z > -(\alpha \cdot y + \beta \cdot u), \tag{37}$$

$$c_m^+ < a - 1 + z < c_{m+1}^+. \tag{38}$$

Similarly, for $y' \diamond u' \in \mathcal{B}(-, A, B, z)$, $y' \diamond u'$ satisfies the following inequalities.

$$a - 1 - z > \alpha \cdot y' + \beta \cdot u', \tag{39}$$

$$c_n^- < a - 1 - z < c_{n+1}^-. \tag{40}$$

In other words, m patterns have the center state " + ", and n patterns have the center state " − ". The chosen partition uniquely determines the order of those patterns. Figure 9 lists the order of the patterns with "+" in the center. Figure 10 lists the order of the patterns with "−" in the center.

$$
\begin{array}{llll}
1:\begin{array}{ccc} - & \ominus & - \\ - & \boxminus & - \end{array} &
2:\begin{array}{ccc} - & \ominus & - \\ - & \boxminus & + \end{array} &
3:\begin{array}{ccc} - & \ominus & - \\ - & \boxplus & - \end{array} &
4:\begin{array}{ccc} - & \ominus & - \\ - & \boxplus & + \end{array} \\
5:\begin{array}{ccc} - & \ominus & + \\ - & \boxminus & - \end{array} &
6:\begin{array}{ccc} - & \ominus & + \\ - & \boxminus & + \end{array} &
7:\begin{array}{ccc} - & \ominus & - \\ + & \boxminus & - \end{array} &
8:\begin{array}{ccc} - & \ominus & - \\ + & \boxminus & + \end{array} \\
9:\begin{array}{ccc} - & \ominus & + \\ - & \boxplus & - \end{array} &
10:\begin{array}{ccc} - & \ominus & + \\ - & \boxplus & + \end{array} &
11:\begin{array}{ccc} + & \ominus & - \\ - & \boxminus & - \end{array} &
12:\begin{array}{ccc} - & \ominus & - \\ + & \boxplus & - \end{array} \\
13:\begin{array}{ccc} + & \ominus & - \\ - & \boxminus & + \end{array} &
14:\begin{array}{ccc} - & \ominus & - \\ + & \boxplus & + \end{array} &
15:\begin{array}{ccc} - & \ominus & + \\ + & \boxminus & - \end{array} &
16:\begin{array}{ccc} - & \ominus & + \\ + & \boxminus & + \end{array} \\
17:\begin{array}{ccc} + & \ominus & - \\ - & \boxplus & - \end{array} &
18:\begin{array}{ccc} + & \ominus & - \\ - & \boxplus & + \end{array} &
19:\begin{array}{ccc} + & \ominus & + \\ - & \boxminus & - \end{array} &
20:\begin{array}{ccc} - & \ominus & + \\ + & \boxplus & - \end{array} \\
21:\begin{array}{ccc} + & \ominus & + \\ - & \boxminus & + \end{array} &
22:\begin{array}{ccc} - & \ominus & + \\ + & \boxplus & + \end{array} &
23:\begin{array}{ccc} + & \ominus & - \\ + & \boxminus & - \end{array} &
24:\begin{array}{ccc} + & \ominus & - \\ + & \boxminus & + \end{array} \\
25:\begin{array}{ccc} + & \ominus & + \\ - & \boxplus & - \end{array} &
26:\begin{array}{ccc} + & \ominus & + \\ - & \boxplus & + \end{array} &
27:\begin{array}{ccc} + & \ominus & - \\ + & \boxplus & - \end{array} &
28:\begin{array}{ccc} + & \ominus & - \\ + & \boxplus & + \end{array} \\
29:\begin{array}{ccc} + & \ominus & + \\ + & \boxminus & - \end{array} &
30:\begin{array}{ccc} + & \ominus & + \\ + & \boxminus & + \end{array} &
31:\begin{array}{ccc} + & \ominus & + \\ + & \boxplus & - \end{array} &
32:\begin{array}{ccc} + & \ominus & + \\ + & \boxplus & + \end{array}
\end{array}
$$

Figure 10: The order of the appearance of the patterns with state "$-$" in the center.

5 Inhomogeneous Cellular Neural Networks

From now on, a two-dimensional ICNN is investigated. Recall that a two-dimensional (2-D) CNN is of the form,

$$\frac{dx_{i,j}}{dt} = -x_{i,j} + z + \sum_{|k|,|l|\leq d} a_{k,l} f(x_{i+k,j+l}) + \sum_{|k|,|l|\leq d} b_{k,l} u_{i+k,j+l}, \tag{41}$$

where $(i,j) \in \mathbb{Z}^2$, $d \in \mathbb{N}$, $f(x)$ is a piecewise-linear output function, defined by

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

$$A = [a_{k,l}] = \begin{bmatrix} a_{-d,d} & \cdots & a_{d,d} \\ \vdots & \ddots & \vdots \\ a_{-d,-d} & \cdots & a_{d,-d} \end{bmatrix} \text{ and } B = [b_{k,l}] = \begin{bmatrix} b_{-d,d} & \cdots & b_{d,d} \\ \vdots & \ddots & \vdots \\ b_{-d,-d} & \cdots & b_{d,-d} \end{bmatrix}$$

represent the feedback template and the controlling template, respectively; z denotes the biased term or threshold. The quantities $x_{i,j}$ denote the state at cell $C_{i,j}$, and $y_{i,j}$ denote the output at $C_{i,j}$. Stationary solutions $\bar{x} = (\bar{x}_{i,j})$ are essential to understand CNN, and their outputs are called *patterns*. Here we concern a specified class of output patterns called mosaic patterns. The connection between CNN with/without input and shift spaces is investigated.

The ICNN system is of the form,

$$\frac{dx_{i,j}}{dt} = \begin{cases} -x_{i,j} + z + \sum_{|k|,|l|\leq d} a_{k,l} f(x_{i+k,j+l}) \\ \qquad + \sum_{|k|,|l|\leq d} b_{k,l} u_{i+k,j+l}, & i,j \equiv 0 \bmod m; \\ -x_{i,j} + z' + a_{0,0} f(x_{i,j}), & \text{otherwise}. \end{cases} \tag{42}$$

for some $m \in \mathbb{N}$, $i,j \in \mathbb{Z}$. Restated, the difference between CNN and ICNN is that the templates and threshold at each cell $C_{i,j}$ are spatially invariant for CNN but variant for ICNN.

5.1 Separation property

Several notions for the formulation of the main results in Secs. 6 and 7 are presented in this subsection. Since the states $C_{i,j}$ with $i = k_1 m, j = k_2 m$ for $k_1, k_2 \in \mathbb{Z}$ are crucial for the study of the mosaic solutions of ICNN, these cells are the main focus in the rest of this investigation.

Definition 5.1. *Let $\bar{x} = (\bar{x}_{i,j})$ be the stationary solution of system (42). $\bar{x}$ is called a mosaic solution if $|\bar{x}_{i,j}| > 1$ for all $i, j \in \mathbb{Z}$, and is called a interior solution if $|\bar{x}_{i,j}| < 1$ for all $i, j \in \mathbb{Z}$. A defect solution $\bar{x}$ satisfies $|\bar{x}_{i,j}| > 1$ for some $(i, j) \in D$ and $|\bar{x}_{k,\ell}| < 1$ for some $(k, \ell) \notin D$, where $D \subsetneq \mathbb{Z}^2$ and $D \neq \varnothing$.*

First, considering the system (42) without input, that is, the template $B \equiv 0$. For each given mosaic solution $\bar{x}$, the output pattern at cell $C_{i,j}$ is $+$, i.e., $\bar{x}_{i,j} > 1$, if and only if

$$\sum_{\substack{|k|,|l| \leq d \\ (k,l) \neq (0,0)}} a_{k,l} \bar{y}_{i+k,j+l} + a + z - 1 > 0, \tag{43}$$

where $a_{0,0} \equiv a$. Similarly, the output pattern at cell $C_{i,j}$ is $-$, i.e., $\bar{x}_{i,j} < -1$, if and only if

$$\sum_{\substack{|k|,|l| \leq d \\ (k,l) \neq (0,0)}} a_{k,l} \bar{y}_{i+k,j+l} - a + z + 1 < 0. \tag{44}$$

(43) and (44) can be rewritten in a much more compact form by introducing the following notations.

Denote by $n = 4d^2 + 4d$. Let X^n be the n-dimensional lattice points, i.e.,

$$X^n = \{v = (v_i) \in \mathbb{R}^n : |v_i| = 1 \text{ for } 1 \leq i \leq n\}. \tag{45}$$

Then, for a given pair of template A and threshold z, the basic set of admissible local patterns with "+" state in the center is defined by

$$\mathcal{B}(+, A, z, d) = \{v \in X^n : \alpha \cdot v + a + z - 1 > 0\},$$

where "$\cdot$" is the inner product, $\alpha = (a_1, a_2, \cdots, a_n), v = (v_1, v_2, \cdots, v_n)$ are obtained from

$$\begin{bmatrix} a_{4d^2+2d} & \cdots & a_{2d^2+d+1} & \cdots & a_1 \\ & & \vdots & & \\ \vdots & \cdots & a_{i,j} & \cdots & \vdots \\ & & \vdots & & \\ a_n & \cdots & a_{2d^2+3d} & \cdots & a_{2d+1} \end{bmatrix} = \begin{bmatrix} a_{i-d,j+d} & \cdots & a_{i,j+d} & \cdots & a_{i+d,j+d} \\ & & \vdots & & \\ \vdots & \cdots & a_{i,j} & \cdots & \vdots \\ & & \vdots & & \\ a_{i-d,j-d} & \cdots & a_{i,j-d} & \cdots & a_{i+d,j-d} \end{bmatrix}$$

and

$$\begin{bmatrix} v_{4d^2+2d} & \cdots & v_{2d^2+d+1} & \cdots & v_1 \\ & & \vdots & & \\ \vdots & \cdots & v_{i,j} & \cdots & \vdots \\ & & \vdots & & \\ v_n & \cdots & v_{2d^2+3d} & \cdots & v_{2d+1} \end{bmatrix} = \begin{bmatrix} y_{i-d,j+d} & \cdots & y_{i,j+d} & \cdots & y_{i+d,j+d} \\ & & \vdots & & \\ \vdots & \cdots & y_{i,j} & \cdots & \vdots \\ & & \vdots & & \\ y_{i-d,j-d} & \cdots & y_{i,j-d} & \cdots & y_{i+d,j-d} \end{bmatrix},$$

respectively. In other words, α represents the surrounding template of A without center, and v indicates the output patterns at cell $C_{i,j}$ whose center is omitted. Similarly, the basic set of admissible local patterns with "$-$" in the center is defined by

$$\mathcal{B}(-, A, z, d) = \{v \in X^n : \alpha \cdot v - a + z + 1 < 0\}.$$

An investigation of the basic sets of admissible local patterns $\mathcal{B}(+, A, z, d)$ and $\mathcal{B}(-, A, z, d)$ are essential for the understanding of the global mosaic patterns on $\mathbb{Z}^2$ that are generated by the given (A, z). Some definitions and theorems should be stated first.

Definition 5.2. *Given $\mathcal{U} \subset X^n$, $\mathcal{U}$ is called separable if there is a hyperplane H in $\mathbb{R}^n$ such that $\mathcal{U}$ and $\mathcal{U}^c$ can be separated by H, where $\mathcal{U}^c = X^n \setminus \mathcal{U}$.*

Hsu et al. [22] investigate how the admissible local mosaic patterns $\mathcal{B}(*, A, z, d)$ relate to the parameters A, z and d in CNN systems, where $* \in \{+, -\}$.

Theorem 5.3 ([22])**.** *There exists (A, z) and d such that $\mathcal{U} = \mathcal{B}(*, A, z, d)$ for some $* \in \{+, -\}$ if and only if $\mathcal{U}$ is separable.*

Moreover, the classical theory of convex set [31] give the necessary and sufficient condition when $\mathcal{U} \subseteq X^n$ is separable.

Theorem 5.4 (Linear Separating Theorem)**.** *$\mathcal{U}$ and $\mathcal{U}^c$ can be separated by a hyperplane in $\mathbb{R}^n$ if and only if*

$$conv(\mathcal{U}) \cap conv(\mathcal{U}^c) = \varnothing, \tag{46}$$

where $conv(\mathcal{K})$ is the convex hull of $\mathcal{K}$ in $\mathbb{R}^n$.

Let $\mathbf{z} = (z, z')$ denote the thresholds, and let $\mathcal{B}(A, \mathbf{z})/\mathcal{B}(A, B, \mathbf{z})$ denote the basic set of admissible local patterns of ICNN without/with input for the given templates. The result in Theorem 5.3 still holds for ICNN systems.

Theorem 5.5. *There exists $(A, \mathbf{z})$ and d such that $\mathcal{U} = \mathcal{B}(*, A, \mathbf{z}, d)$ for some $* \in \{+, -\}$ if and only if $\mathcal{U} \subseteq X^n$ is separable.*

Proof. It suffices to show that $\mathcal{B}(+, A, \mathbf{z}, d) = \mathcal{U}$ for some $(A, \mathbf{z})$ and d if and only if $\mathcal{U}$ is separable. The proof for $* = -$ is essentially the same, thus is omitted.

First, considering the output pattern at $C_{i,j}$, where $(i, j) \neq (k_1 m, k_2 m)$ for some $k_1, k_2 \in \mathbb{Z}$. The output pattern is $+$ if and only if $a + z' - 1 > 0$, and is $-$ if and only if $a - z' - 1 > 0$. Let $a > 1$ and $z' = \frac{1}{2}(a - 1)$. The output pattern at $C_{i,j}$ can be arbitrary in such a case. It remains to show that $\mathcal{U}$ can be realized on $C_{i,j}$ for some appropriate choice of (A, z), where $i, j \equiv 0 \mod m$.

Let $S = \{\mathcal{U} \subseteq X^n | \ \mathcal{U} \text{ satisfies } (46)\}$. For each $\mathcal{U} \in S$, denoting by

$$\mathcal{A}^+(\mathcal{U}) = \{(\alpha, p) | \ \alpha \cdot v + p > 0 \text{ for all } v \in \mathcal{U}\}, \tag{47}$$

$$\mathcal{A}^-(\mathcal{U}) = \{(\alpha, q) | \ \alpha \cdot v + q < 0 \text{ for all } v \in \mathcal{U}^c\}. \tag{48}$$

Then $\mathcal{A}^+(\mathcal{U}) \cap \mathcal{A}^-(\mathcal{U}) \neq \varnothing$ if and only if $\mathcal{U}$ satisfies (46). In this case, the boundary $\partial\mathcal{A}^+(\mathcal{U})$ of $\mathcal{A}^+(\mathcal{U})$ consists of (A, B, z) such that $\alpha \cdot v + a + z - 1 = 0$, where $p = a + z - 1$.

Defining

$$\widehat{\mathcal{B}}(+, \alpha, p) = \{v : \alpha \cdot v + p > 0\}, \tag{49}$$

then $\widehat{\mathcal{B}}(+, \alpha, p) = \mathcal{U}$ for all $(\alpha, p) \in \mathcal{A}^+(\mathcal{U})$. For each $\mathcal{U} \in S$ so that there exists $(\alpha, p) \in \mathcal{A}^+(\mathcal{U})$, considering

$$z = \frac{p}{2} - k, \ \ a = 1 + \frac{p}{2} + k, \tag{50}$$

where k is chosen so that $\frac{p}{2} + k > 0$. Then $\mathcal{B}(+, A, \mathbf{z}, d) = \widehat{\mathcal{B}}(+, \alpha, p) = \mathcal{U}$, and vice verse. This completes the proof. $\square$

Next, considering system (42) with input. Given a mosaic solution $\bar{x}$, the output pattern at cell C_{ij} is $+$ if and only if

$$\sum_{\substack{|k|,|l| \leq d \\ (k,l) \neq (0,0)}} a_{k,l} \bar{y}_{i+k,j+l} + \sum_{|k|,|l| \leq d} b_{k,l} \bar{u}_{i+k,j+l} + a + z - 1 > 0. \tag{51}$$

Similarly, the output pattern at cell C_{ij} is $-$ if and only if

$$\sum_{\substack{|k|,|l| \leq d \\ (k,l) \neq (0,0)}} a_{k,l} \bar{y}_{i+k,j+l} + \sum_{|k|,|l| \leq d} b_{k,l} \bar{u}_{i+k,j+l} - a + z + 1 < 0. \tag{52}$$

It is seen from the above discussion that the basic set of admissible local patterns with "+" in the center is defined by

$$\mathcal{B}(+, A, B, z, d) = \{(v, w) \in X^n \times X^{n+1} : \alpha \cdot v + \beta \cdot w + a + z - 1 > 0\},$$

and the basic set of admissible local patterns with "−" in the center is defined by

$$\mathcal{B}(-, A, B, z, d) = \{(v, w) \in X^n \times X^{n+1} : \alpha \cdot v + \beta \cdot w - a + z + 1 < 0\}.$$

Herein, $\beta = (b_1, b_2, \cdots, b_{n+1})$ and $w = (w_1, w_2, \cdots, w_{n+1})$ are obtained from

$$\begin{bmatrix} b_{4d^2+2d+1} & \cdots & b_{2d^2+d+1} & \cdots & b_1 \\ & & \vdots & & \\ \vdots & \cdots & b_{2d^2+2d+1} & \cdots & \vdots \\ & & \vdots & & \\ b_{n+1} & \cdots & b_{2d^2+3d+1} & \cdots & b_{2d+1} \end{bmatrix} = \begin{bmatrix} b_{i-d,j+d} & \cdots & b_{i,j+d} & \cdots & b_{i+d,j+d} \\ & & \vdots & & \\ \vdots & \cdots & b_{i,j} & \cdots & \vdots \\ & & \vdots & & \\ b_{i-d,j-d} & \cdots & b_{i,j-d} & \cdots & b_{i+d,j-d} \end{bmatrix}$$

and

$$\begin{bmatrix} w_{4d^2+2d+1} & \cdots & w_{2d^2+d+1} & \cdots & w_1 \\ & & \vdots & & \\ \vdots & \cdots & w_{2d^2+2d+1} & \cdots & \vdots \\ & & \vdots & & \\ w_{n+1} & \cdots & w_{2d^2+3d+1} & \cdots & w_{2d+1} \end{bmatrix} = \begin{bmatrix} u_{i-d,j+d} & \cdots & u_{i,j+d} & \cdots & u_{i+d,j+d} \\ & & \vdots & & \\ \vdots & \cdots & u_{i,j} & \cdots & \vdots \\ & & \vdots & & \\ u_{i-d,j-d} & \cdots & u_{i,j-d} & \cdots & u_{i+d,j-d} \end{bmatrix},$$

respectively. Namely, β represents the template B and w indicates the input patterns at cell $C_{i,j}$.

Theorem 2.1 generalized Theorem 5.3 to a common case that the controlling template B is considered and can be restated as follows.

Theorem 5.6. *There exists (A, B, z) and d such that $\mathcal{U} = \mathcal{B}(*, A, B, z, d)$ for some $* \in \{+, -\}$ if and only if $\mathcal{U}$ is separable.*

Theorem 5.6 can also be applied for ICNN with input.

Theorem 5.7. *There exists* $(A, B, \mathbf{z})$ *and* d *such that* $\mathcal{U} = \mathcal{B}(*, A, B, \mathbf{z}, d)$ *for some* $* \in \{+, -\}$ *if and only if* $\mathcal{U}$ *is separable.*

Proof. This can be accomplished via analogous method as in the proof of Theorem 5.5, thus is omitted. $\square$

6 Inhomogeneous Cellular Neural Networks without Input

The dense entropy property for the ICNN without input is studied in this section. Subsection 6.1 develops the fundamental theory and gives it an application for ICNN in Subsec. 6.2.

6.1 Two-dimensional subshift of finite type

This subsection investigates the preliminaries that are necessary for the understanding of dense entropy property of ICNN without input.

Definition 6.1. *Let* $\mathbf{X} \subseteq \{1,-1\}^{\mathbb{Z}^2}$ *be a 2-dimensional shift space with finite alphabet* $\mathcal{A}(\mathbf{X}) = \{1,-1\}$.

(1) *If* $x \in \mathbf{X}$ *and* $S \subseteq \mathbb{Z}^2$*, the restriction of* x *to* S *is denoted by* $\pi_S(x)$.

(2) *Let* $\Lambda(n) = \{(p,q) : p,q \in \mathbb{Z}, 0 \leq p,q \leq n-1\}$. *An* n*-block is* $\pi_{\boldsymbol{c}+\Lambda(n)}(x)$ *for some* $\boldsymbol{c} \in \mathbb{Z}^2$, $x \in \mathbf{X}$. *The set of* n*-blocks is denoted by* $B_n(\mathbf{X})$.

(3) *A configuration on* $S \subseteq \mathbb{Z}^2$ *is a map* $E : S \to \mathcal{A}(\mathbf{X})$. *For* $x \in \mathbf{X}$, E *occurs in* x *if* $\pi_{\boldsymbol{c}+S}(x) = E$ *for some* $\boldsymbol{c} \in \mathbb{Z}^2$.

(4) *For each* $\boldsymbol{c} \in \mathbb{Z}^2$*, the shift map* $\sigma_{\boldsymbol{c}} : \mathbf{X} \to \mathbf{X}$ *is defined by* $\pi_{\boldsymbol{d}}(\sigma_{\boldsymbol{c}}(x)) = \pi_{\boldsymbol{c}+\boldsymbol{d}}(x)$ *for all* $\boldsymbol{d} \in \mathbb{Z}^2$. *Moreover, the iteration of* $\sigma_{\boldsymbol{c}}$ *is denoted by* $\sigma_{\boldsymbol{c}}^{\ell} = \sigma_{\boldsymbol{c}} \circ \sigma_{\boldsymbol{c}}^{\ell-1}$ *for all* $\ell \in \mathbb{N}$.

Denote $\pi_{\Lambda(n)}(x)$ by $\pi_n(x)$ for simplicity.

Definition 6.2. *Given* $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n\times n}}, s \in \mathbb{N}, s < n$*, the shift space*

$\mathbf{X}_s(\mathcal{U}) \subseteq \{1,-1\}^{\mathbb{Z}^2}$ *is defined by*

$$\mathbf{X}_s(\mathcal{U}) = \{x \in \{1,-1\}^{\mathbb{Z}^2} : \pi_n(\sigma^{\ell}_{(i,j)}(x)) \in \mathcal{U} \text{ for all } \ell \in \mathbb{Z}, i,j \in \{0, n-s\}\}. \tag{53}$$

Moreover, the r*-copy of* $\mathcal{U}$, $\mathcal{U}^r \subseteq \{1,-1\}^{\mathbb{Z}_{k\times k}}$, *where* $k = rn-(r-1)s$, *is defined by*

$$\mathcal{U}^r = \{v \in \{1,-1\}^{\mathbb{Z}_{k\times k}} : \exists\, x \in \mathbf{X}_s(\mathcal{U}) \text{ such that } \pi_k(x) = v\}. \tag{54}$$

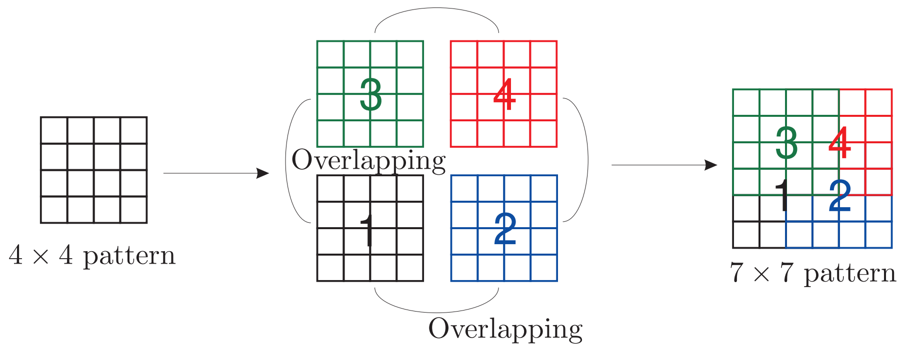
Remark 6.3. In other words, $\mathcal{U}^r$ is consisting of those patterns combined by r^2-many patterns in $\mathcal{U}$ with s-many rows/columns overlapped. For example, consider $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{4\times 4}}$ and $s=1$. $\mathcal{U}^2$ consists of those patterns with size 7×7 such that each pattern $v \in \mathcal{U}^2$ is a combination of four patterns in $\mathcal{U}$ with one-row/column overlapped. As being seen in Fig. 11, the last column on the right hand side in pattern 1 can be overlapped with the first column on the left hand side in pattern 2 if and only if these two 1×4 patterns are exactly the same. The same applies to the top row in pattern 1 and the bottom row in pattern 3.

Next, the effect of the parameter s is studied. In general, the range of s is less than n and greater than one. After constructing $\mathcal{U}^r$ from a given $\mathcal{U}$, the lemma below studies the relationship between the subshifts of finite type $\mathbf{X}_s(\mathcal{U})$ and $\mathbf{X}_s(\mathcal{U}^r)$. In addition, it reduces the complexity caused by s.

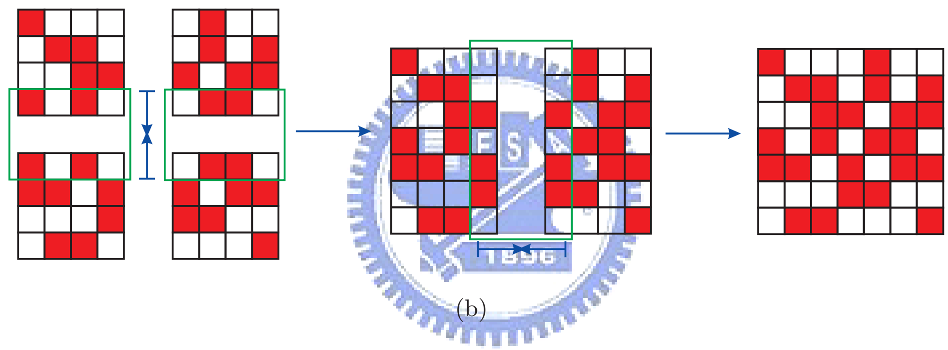
Lemma 6.4. *Given* $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n\times n}}$ *and* $r \in \mathbb{N}$, *then* $\mathbf{X}_s(\mathcal{U}) = \mathbf{X}_s(\mathcal{U}^r)$.

Proof. Since $\mathcal{U}^r$ is constructed from $\mathcal{U}$ such that each pattern in $\mathcal{U}^r$ consists of r^2-many patterns in $\mathcal{U}$ with s-many columns/rows overlapped, it is seen that $\mathbf{X}_s(\mathcal{U}^r) \subseteq \mathbf{X}_s(\mathcal{U})$. This remains to show that $\mathbf{X}_s(\mathcal{U}) \subseteq \mathbf{X}_s(\mathcal{U}^r)$.

If $x \in \mathbf{X}_s(\mathcal{U})$, then $\pi_n(\sigma^{\ell}_{(i,j)}(x)) \in \mathcal{U}$ for all $\ell \in \mathbb{Z}$, where $i,j \in \{0, n-s\}$. Definition 6.2 shows that $\pi_k(x) \in \mathcal{U}^r$, where $k = rn-(r-1)s-1$. Let



(a)



(b)

Figure 11: The construction of $\mathcal{U}^r$ for a given $\mathcal{U}$ and $r \in \mathbb{N}$. Take $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{4\times 4}}$, $r = 2$ and $s = 1$ as an example. First picking four patterns in $\mathcal{U}$, say P_1, P_2, P_3, P_4. If the patterns in the first row of P_1 are differ from the patterns in the last row of P_3, then nothing happens. Otherwise, P_1 and P_3 are combined with one-row overlapped. Repeating this process, a new pattern with size 7×7 is thus derived.

$y = \sigma_{(i,j)}(x)$ for some $i, j \in \{0, n-s\}$. Then $\pi_k(y) \in \mathcal{U}^r$ via the same argument. It can be easily check that $\pi_k(\sigma^{\ell}_{(i,j)}(x)) \in \mathcal{U}$ for all $\ell \in \mathbb{Z}$ by mathematical induction, where $i, j \in \{0, n-s\}$. Therefore, $x \in \mathbf{X}_s(\mathcal{U}^r)$ and this completes the proof. $\square$

Without loss of generality, assuming that $s \leq [\frac{n}{2}]$, where $[\cdot]$ is the Gauss function. The case where $s > [\frac{n}{2}]$ is discussed in Remark 6.8.

It is seen so far that a subshift of finite type is generated once $\mathcal{U}$ and s are given. The method that embeds a chosen set of admissible local patterns in an ICNN system is introduced.

If $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n\times n}}$ is given and n is even, then an extension of $\mathcal{U}$, denoted by $\mathcal{V} \subset \{1,-1\}^{\mathbb{Z}_{(n+1)\times(n+1)}}$, is constructed as follows. $v = (v_{(i,j)}) \in \mathcal{V}$ if and only if

(i) $v_{(i,j)} = -1$ if $i = \frac{n}{2}+1$ or $j = \frac{n}{2}+1$.

(ii) $v_{<\frac{n}{2}+1>} = u$ for some $u \in \mathcal{U}$, where $v_{<p;q>} \in \{1,-1\}^{\mathbb{Z}_{n\times n}}$ is obtained from v by deleting row p and column q, and denoted by $v_{<p>}$ if $p = q$.

Similarly, if n is odd, constructing $\mathcal{V} \subset \{1,-1\}^{\mathbb{Z}_{(n+2)\times(n+2)}}$ by $v = (v_{(i,j)}) \in \mathcal{V}$ if and only if

(i) $v_{(i,j)} = -1$ if either i or $j \in \{\frac{n+1}{2}, \frac{n+3}{2}\}$.

(ii) $v'_{<\frac{n-1}{2}>} = u$ for some $u \in \mathcal{U}$, where $v'_{<p;q>} \in \{1,-1\}^{\mathbb{Z}_{n\times n}}$ is obtained from $v_{<p;q>}$ by deleting row p and column q, and denoted by $v'_{<p>}$ if $p = q$.

More precisely, $\mathcal{U}$ is extended to $\mathcal{V}$ by adding a cross of "-1" to the center of each $u \in \mathcal{U}$. Under such extension, there is a one-to-one correspondence

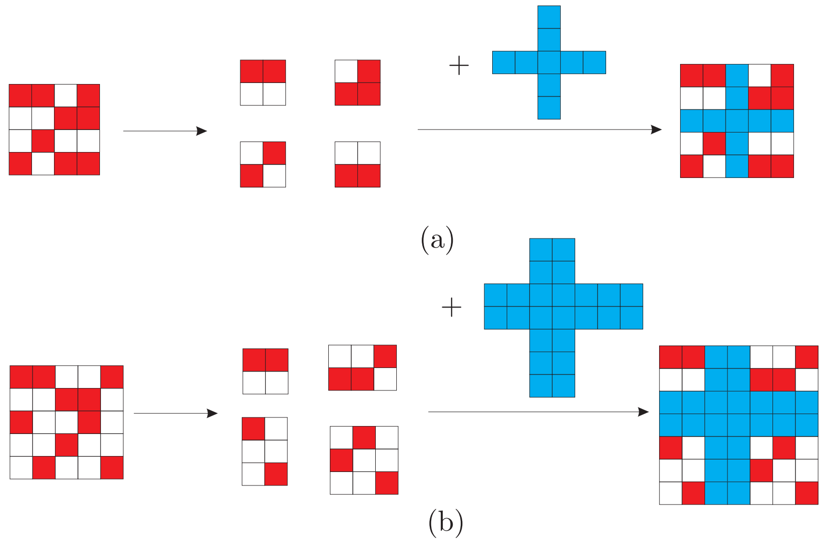


Figure 12: (a) Extend a 4×4 pattern to a 5×5 pattern by adding a cross of pattern into the center of the original one. The pattern "+" is represented by red and the pattern "−" is represented by white and blue, herein blue is used to distinguish from the original pattern. (b) Extend a 5×5 pattern to a 7×7 pattern.

between $\mathcal{U}$ and $\mathcal{V}$. Figure 12 gives two examples for the cases where n is odd and n is even, respectively.

Remark 6.5. Notably, the size of $\mathcal{V}$ is odd no matter what the size of $\mathcal{U}$ is. That is, $\mathcal{V} \subseteq \{1,-1\}^{\mathbb{Z}_{\ell \times \ell}}$ for some $\ell = 2k+1$, $k \in \mathbb{N}$.

For each $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n \times n}}$, there associates an unique $\mathcal{V} \subseteq \{1,-1\}^{\mathbb{Z}_{(n+1)\times(n+1)}}$ under the construction above. The relationship between $\mathbf{X}_s(\mathcal{U})$ and $\mathbf{X}_s(\mathcal{V})$ is investigated below. Before stating the lemma, a definition is given first.

Definition 6.6. *Let* $\mathbf{X}, \mathbf{Y}$ *be shift spaces with shift maps* $\sigma_{\mathbf{X}}$ *and* $\sigma_{\mathbf{Y}}$*, respectively. Define* $\phi : \mathbf{X} \to \mathbf{Y}$ *be a factor map from* $\mathbf{X}$ *to* $\mathbf{Y}$ *if* ϕ *is onto and* $\phi \circ \sigma_{\mathbf{X}} = \sigma_{\mathbf{Y}} \circ \phi$*.* $\mathbf{X}$ *is conjugate to* $\mathbf{Y}$*, denoted by* $\mathbf{X} \cong \mathbf{Y}$*, if* ϕ *is a factor map and one-to-one.*

A key lemma then follows.

Lemma 6.7. *Given* $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n\times n}}$*, constructing* $\mathcal{V}$ *as above, then* $\mathbf{X}_s(\mathcal{U}) \cong \mathbf{X}_s(\mathcal{V})$*.*

Proof. Define $\psi : \mathcal{V} \to \mathcal{U}$ by $\psi(v) = u$, where

$$u = \begin{cases} v_{<\frac{n}{2}+1>}, & n \text{ is even;} \\ v'_{<\frac{n-1}{2}>}, & n \text{ is odd.} \end{cases} \tag{55}$$

For simplicity, assuming n is even. The case where n is odd can be done similarly. It is easily seen that $\psi(v)$ is bijective. Furthermore, defining $\phi : \mathbf{X}_s(\mathcal{V}) \to \mathbf{X}_s(\mathcal{U})$ by $\phi(y)_{(\ell i,\ell j)+\mathbf{c}} = \psi(\pi_{n+1}(y_{(\ell q i,\ell q j)}))_{\mathbf{c}}$, where $i,j \in \{0, n-s\}$, $\ell \in \mathbb{Z}$, $q = \dfrac{n-s+1}{n-s}$, $\mathbf{c} \in \Lambda(n)$ and $y \in \mathbf{X}_s(\mathcal{V})$. In such a case, $\phi \circ \sigma_{\mathbf{X}_s(\mathcal{V})} = \sigma_{\mathbf{X}_s(\mathcal{U})} \circ \phi$ and ϕ is a conjugacy since ψ is one-to-one and onto. This completes the proof. $\square$

Remark 6.8. If $s > [\frac{n}{2}]$, let $\ell \in \mathbb{N}$ satisfies

$$\left[\frac{(\ell-1)(n-s)+s}{2}\right] < s \leq \left[\frac{\ell(n-s)+s}{2}\right]. \tag{56}$$

Then constructing $\mathcal{V}$ via the same method mentioned above so that there is a one-to-one correspondence between $\mathcal{V}$ and $\mathcal{U}^{\ell}$. Similar as above, Lemma 6.4 and Lemma 6.7 show that $\mathbf{X}_s(\mathcal{U}) \cong \mathbf{X}_s(\mathcal{V})$.

6.2 Two-dimensional inhomogeneous cellular neural networks without input

Subsection 6.1 shows that $\mathbf{X}_s(\mathcal{U}) = \mathbf{X}_s(\mathcal{U}^r)$ and $\mathbf{X}_s(\mathcal{U}) \cong \mathbf{X}_s(\mathcal{V})$, where $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n\times n}}$ is given, $\mathcal{V} \subseteq \{1,-1\}^{\mathbb{Z}_{(n+1)\times(n+1)}}$ is obtained from $\mathcal{U}$ and $r \in \mathbb{N}$. This subsection applies the theory developed in the last subsection to ICNN without input. First, the preservation of the separation property between $\mathcal{U}$ and $\mathcal{V}$ is given below.

Lemma 6.9. *Given $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n\times n}}$, then $\mathcal{U}$ is separable if and only if $\mathcal{V}$ is separable.*

Proof. For simplicity, the case where n is even is proved. It can be done similarly when n is odd.

If $\mathcal{U}$ is separable, there is a linear functional $g : \{1,-1\}^{\mathbb{Z}_{n\times n}} \to \mathbb{R}$ and $\alpha \in \mathbb{R}$ so that $g(u) < \alpha$ for all $u \in \mathcal{U}$, and $g(u) > \alpha$ for all $u \in \mathcal{U}^c$. Let $\rho = \alpha - \min\{g(u) : u \in \mathcal{U}\}$. Define $\hat{g} : \{1,-1\}^{\mathbb{Z}_{(n+1)\times(n+1)}} \to \mathbb{R}$ by

$$\hat{g}(v) = g(u) + \rho \sum_{i \text{ or } j=\frac{n}{2}+1} v_{(i,j)}, \tag{57}$$

where $u \in \{1,-1\}^{\mathbb{Z}_{n\times n}}$ is obtained from v by deleting row $(\frac{n}{2}+1)$ and column $(\frac{n}{2}+1)$. Then $\hat{g}(v) < \alpha - (2n+1)\rho$ for $v \in \mathcal{V}$ and $\hat{g}(v) > \alpha - (2n+1)\rho$ for $v \in \mathcal{V}^c$. Thus, $\mathcal{V}$ is separable.

Similarly, if $\mathcal{V}$ is separable, then so is $\mathcal{U}$. This completes the proof. $\square$

Before stating the main theorem, the following theorem is essential for the study of the mosaic solutions of ICNN.

Theorem 6.10. *Given $\mathcal{U} \subseteq \{1,-1\}^{\mathbb{Z}_{n\times n}}$, and $s \in \mathbb{N}$. If $\mathcal{U}$ is separable, then there exist $m \in \mathbb{N}$ and $(A, \mathbf{z}, d)$ for system (42) such that $\mathbf{X}(\mathcal{B}(A, \mathbf{z}, d)) \cong \mathbf{X}_s(\mathcal{U})$, where $\mathcal{B}(A, \mathbf{z}, d)$ is the admissible local patterns obtained from (42) with parameters $(A, \mathbf{z}, d)$,*

$$\mathbf{X}(\mathcal{B}(A,\mathbf{z},d)) = \left\{ \begin{array}{c} x \in \{1,-1\}^{\mathbb{Z}^2} : \pi_{\kappa(d)}(\sigma^{\ell}_{(i,j)}(x)) \in \mathcal{B}(A,\mathbf{z},d) \\ \text{for all } \ell \in \mathbb{Z}, i,j \in \{0,m\} \end{array} \right\}, \tag{58}$$

and $\kappa(d) = \{(p,q) : -d \le p,q \le d, p,q \in \mathbb{Z}\}$.

Proof. Without loss of generality, assuming that n is even and $s \le \frac{n}{2}$. Once $\mathcal{U}$ is given, constructing $\mathcal{V}$ as above. Lemma 6.7 and Lemma 6.9 indicate that $\mathbf{X}_s(\mathcal{V}) \cong \mathbf{X}_s(\mathcal{U})$ and $\mathcal{V}$ is separable. Consider $d = \frac{n}{2}$, Theorem 5.5 shows that there exists $(A, \mathbf{z}, d)$ so that $\mathcal{B}(A, \mathbf{z}, d) = \mathcal{V}$.

Let $m = 2d - s + 1$. For each $x \in \mathbf{X}(\mathcal{B}(A, \mathbf{z}, d))$, (58) implies

$$\pi_{\kappa(d)}(x_{\ell(i,j)}) \in \mathcal{B}(A, \mathbf{z}, d) \text{ for all } \ell \in \mathbb{Z}, i, j \in \{0, n - s + 1\}.$$

It is easily seen that $\mathbf{X}(\mathcal{B}(A, \mathbf{z}, d)) = \mathbf{X}_s(\mathcal{V})$. Since $\mathbf{X}_s(\mathcal{U}) \cong \mathbf{X}_s(\mathcal{V})$, the proof is completed. $\square$

When $(A, \mathbf{z}, d)$ is given, the basic set of admissible local patterns $\mathcal{B} = \mathcal{B}(A, \mathbf{z}, d)$ is immediately determined. Let $\Sigma_{p,q}(\mathbf{X}(\mathcal{B}))$ denote the set of global patterns in $\mathbf{X}(\mathcal{B})$ with size $p \times q$, and let $\Gamma_{p,q}(\mathbf{X}(\mathcal{B})) = |\Sigma_{p,q}(\mathbf{X}(\mathcal{B}))|$. The entropy of $\mathbf{X}(\mathcal{B})$ is defined by

$$h(\mathbf{X}(\mathcal{B})) \equiv \lim_{p,q \to \infty} \frac{\log \Gamma_{p,q}(\mathbf{X}(\mathcal{B}))}{pq}.$$

The existence of the limit can be found in [10].

The first main theorem of this investigation, the dense entropy property of ICNN without input, as follows.

Theorem 6.11. *For $t \in [0, \log 2]$, $\varepsilon > 0$, there exist $m \in \mathbb{N}$ and $(A, \mathbf{z}, d)$ such that $|h(\mathbf{X}(\mathcal{B}(A, \mathbf{z}, d))) - t| < \varepsilon$.*

Before proving the theorem, the following lemmas should be stated first.

Lemma 6.12. *Let $S_{n,l} \subset X^n$ be defined by*

$$S_{n,l} = \{x = (x_1, \cdots, x_n) \in X^n : x_k = -1 \text{ for all } l + 1 \le k \le n\}, \tag{59}$$

$1 \le l \le n - 1$, and $S_{n,n} = X^n$. Then $S_{n,l}$ is separable.

Proof. Define a linear functional $g : \mathbb{R}^n \to \mathbb{R}$ by

$$g(x) = \sum_{i=l+1}^{n} x_i \quad \text{for all } x = (x_i)_{i=1}^{n} \in \mathbb{R}^n. \tag{60}$$

Let $h(x) = g(x) + (n - l - 1)$. It can be easily checked that $h(x) < 0$ for all $x \in S_{n,l}$ and $h(x) > 0$ for all $x \in S^c_{n,l}$. That is, $S_{n,l}$ and $S^c_{n,l}$ can be separated by the hyperplane

$$H = \{x \in \mathbb{R}^n : g(x) = l - n + 1\}. \tag{61}$$

This completes the proof. □

Theorem 6.13. *Given $l, d \in \mathbb{N}$ and $n = 4d^2$. There exists $\mathcal{U}_{d,l} \subseteq \{1, -1\}^{\mathbb{Z}_{2d\times 2d}}$ such that $h(\mathbf{X}_d(\mathcal{U}_{d,l})) = \frac{l}{n} \log 2$ and $\mathcal{U}_{d,l}$ is separable, where $1 \leq l \leq n$.*

Proof. If $d, l \in \mathbb{N}$ is given, $n = 4d^2$ and $1 \leq l \leq n$. Define

$$T : \{1, -1\}^{\mathbb{Z}_{n\times 1}} \to \{1, -1\}^{\mathbb{Z}_{2d\times 2d}}$$

by

$$(T\nu)_{i,j} = \nu_{2d(i-1)+j} \quad \text{for all } \nu = (\nu_k) \in \{1, -1\}^{\mathbb{Z}_{n\times 1}}. \tag{62}$$

Let $S_{n,l}$ be defined as in Lemma 6.12, and let $M_{n,l}$ be defined as follows.

$$M_{n,l} \equiv \{K \in \{1, -1\}^{\mathbb{Z}_{2d\times 2d}} : \exists\, \nu \in S_{n,l} \text{ such that } K = T\nu\}.$$

Furthermore, constructing $\mathcal{U}_{d,l} \subseteq \{1, -1\}^{\mathbb{Z}_{4d\times 4d}}$ as follow. $J \in \mathcal{U}_{d,l}$ if $\pi_{(i,j)+\Lambda(2d)}(J) \in M_{n,l}$ for $i, j \in \{0, 2d\}$, where $\Lambda(n)$ is defined as Def. 6.1.

Claim. $\mathcal{U}_{d,l}$ is separable.

Let $g : \mathbb{R}^n \to \mathbb{R}$ be defined as in (60) and $\tilde{g} = g \circ T^{-1}$. For $w \in \{1, -1\}^{\mathbb{Z}_{4d\times 4d}}$, rewriting w as $w = \boxed{\begin{matrix} w_1 & w_2 \\ w_3 & w_4 \end{matrix}}$, where $w_i \in \{1, -1\}^{\mathbb{Z}_{2d\times 2d}}$ for all i. Define a linear functional $\tau : \{1, -1\}^{\mathbb{Z}_{4d\times 4d}} \to \mathbb{R}$ by $\tau(w) = \tilde{g}(w_1) + \tilde{g}(w_2) + \tilde{g}(w_3) + \tilde{g}(w_4)$ and $\tilde{\tau}(w) = \tau(w) + 4n - 4l - 1$. The above constitution confirms that $\tilde{\tau}(w) < 0$ for all $w \in \mathcal{U}_{d,l}$ and $\tilde{\tau}(w) > 0$ otherwise. That means $\mathcal{U}_{d,l}$ is separable.

Moreover, there are 2^l-many patterns for each block of $x \in \mathbf{X}_d(\mathcal{U}_{d,l})$ with size $2d \times 2d$. Therefore,

$$h(\mathbf{X}_d(\mathcal{U}_{d,l})) = \lim_{p,q\to\infty} \frac{\log \Gamma_{2dp\cdot 2dq}(\mathbf{X}_d(\mathcal{U}_{d,l}))}{2dp \cdot 2dq} = \lim_{p,q\to\infty} \frac{\log (2^l)^{pq}}{4d^2pq} = \frac{l}{n} \log 2.$$

This completes the proof. $\square$

Proof of Theorem 6.11. For $t \in [0, \log 2]$ and $\varepsilon > 0$, there exist $d, l \in \mathbb{N}$ such that $|\frac{l}{n} \log 2 - t| < \varepsilon$, where $n = 4d^2$. Theorem 6.13 indicates there is a separable set $\mathcal{U}_{d,l}$ such that $h(\mathbf{X}_d(\mathcal{U}_{d,l})) = \frac{l}{n} \log 2$. Lemma 6.7, Lemma 6.9 and Theorem 6.10 shows $h(\mathbf{X}_d(\mathcal{U}_{d,l})) = h(\mathbf{X}_d(\mathcal{V}_{d,l}))$ and there exist $m \in \mathbb{N}$ and $(A, \mathbf{z}, d)$ such that $\mathcal{B}(A, \mathbf{z}, d) = \mathcal{V}_{d,l}$. The proof is then completed. $\square$

7 Inhomogeneous Cellular Neural Networks with Input

In this section, Theorem 6.11 is extended to the case where B is not identical to zero.

Once the parameters $(A, B, \mathbf{z}, d)$ are given, the basic set of admissible local patterns is determined and denoted by

$$\mathcal{B} \equiv \mathcal{B}(A, B, \mathbf{z}, d) = \{Y \circ U\} \subseteq \{1, -1\}^{\mathbb{Z}_{(2d+1)\times(2d+1)\times 2}},$$

where $Y, U \in \{1, -1\}^{\mathbb{Z}_{(2d+1)\times(2d+1)}}$. The output pattern Y coupled with input pattern U, denoted by $Y \circ U$, is a two-layer array. Defining the output space generated by $\mathcal{B}(A, B, \mathbf{z}, d)$ as follows.

$$\mathbf{X}(\mathcal{B}) = \left\{ \begin{array}{c} y \in \{1, -1\}^{\mathbb{Z}^2} : \text{there exists } u \in \{1, -1\}^{\mathbb{Z}^2} \text{ such that} \\ \pi_{\kappa(d)}(\sigma^{\ell}_{(i,j)}(y \circ u)) \in \mathcal{B} \text{ for all } \ell \in \mathbb{Z}, i, j \in \{0, m\} \end{array} \right\}, \tag{63}$$

where $\kappa(d)$ is defined in Theorem 6.10 and

$$\pi_{\kappa(d)}(\sigma_{(i,j)}(y \circ u)) \equiv \pi_{\kappa(d)}(\sigma_{(i,j)}(y)) \circ \pi_{\kappa(d)}(\sigma_{(i,j)}(u)).$$

For $d, l \in \mathbb{N}$, let $\mathcal{U}_{d,l}$ be the same as defined in the proof of Theorem 6.13. Denote by

$$\mathcal{V}_{d,l} = \{Y \circ U : Y, U \in \mathcal{U}_{d,l}\} \subseteq \{1, -1\}^{\mathbb{Z}_{2d\times 2d\times 2}}. \tag{64}$$

Then the lemma follows.

Lemma 7.1. *$\mathcal{V}_{d,l}$ is separable.*

Proof. Let τ be the same as in the proof of Theorem 6.13. Define a linear functional $\theta : \{1, -1\}^{\mathbb{Z}_{4d\times 4d}} \times \{1, -1\}^{\mathbb{Z}_{4d\times 4d}} \to \mathbb{R}$ by $\theta(u, v) = \tau(u) + \tau(v)$ and $\widetilde{\theta}(u, v) = \theta(u, v) + 8n + 8l - 1$. It is easily to check that $\widetilde{\theta}(u \circ v) < 0$ for all $u \circ v \in \mathcal{V}_{d,l}$ and $\widetilde{\theta}(u \circ v) > 0$ otherwise. This completes the proof. $\square$

Furthermore, the entropy of the subshift space induced by $\mathcal{V}_{d,l}$ can be computed via the same method as in the proof of Theorem 6.13, thus the proof is omitted.

Theorem 7.2. $h(\mathbf{X}_d(\mathcal{V}_{d,l})) = \frac{l}{n} \log 2$.

The dense entropy property of ICNN with input then follows.

Theorem 7.3. *For* $t \in [0, \log 2]$*,* $\varepsilon > 0$*, there exist* $m \in \mathbb{N}$ *and* $(A, B, \mathbf{z}, d)$ *such that* $|h(\mathbf{X}(\mathcal{B}(A, B, \mathbf{z}, d))) - t| < \varepsilon$*.*

The proof of Theorem 7.3 can be accomplished via the same discussion in the proof of Theorem 6.11, hence is skipped.

References

[1] J.C. Ban, K.P. Chien, C.H. Hsu and S.S. Lin, *Spatial disorder of CNN-with asymmetric output funtion*, International J. of Bifurcation and Chaos 11(2001), pp. 2085-2095.

[2] J.C. Ban, C.H. Hsu and S.S. Lin, *Spatial disorder of cellular neural network-with biased term*, International J. of Bifurcation and Chaos 12(2002), pp. 525-534.

[3] J.C. Ban, W.G. Hu, S.S. Lin and Y.H. Lin, *Zeta functions for two-dimensional subshift of finite type*, preprint(2007).

[4] J.C. Ban, S.S. Lin and C.W. Shih, *Exact number of mosaic patterns in one-dimensional cellular neural networks*, International J. of Bifurcation and Chaos 11(2001), pp. 1645-1653.

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

[6] J.C. Ban, S.S. Lin and Y.H. Lin, *Patterns generation and spatial entropy in two-dimensional lattice models*, to appear in Asian J. Math.(2006).

[7] J.C. Ban, S.S. Lin and Y.H. Lin, *Three-dimensional cellular neural networks and patterns generation problems*, to appear in Internat. J. Bifur. Chaos Appl. Sci. Engrg.(2007).

[8] J.C. Ban, S.S. Lin and Y.H. Lin, *Primitivity of subshifts of finite type in two-dimensional lattice models*, preprint(2007).

[9] S.N. Chow, J. Mallet-Paret and E.S. Van Vleck, Dynamics

OF LATTICE DIFFERENTIAL EQUATIONS, *International J. of Bifurcation and Chaos* 6(1996), pp. 1605-1621.

[10] S.N. CHOW, J. MALLET-PARET AND E.S. VAN VLECK, *Pattern formation and spatial chaos in spatially discrete evolution equations*, Random Comput. Dynam. 4(1996), pp. 109-178.

[11] L.O. CHUA, *CNN: A Paradigm for Complexity*, World Scientific Series on Nonlinear Science, Series A,31. World Scietific, Singapore, 1998.

[12] L.O. CHUA AND L. YANG, *Cellular neural networks: Theory*, IEEE Trans. Circuits Syst. I Regul. Pap. 35(1988), pp. 1257-1272.

[13] L.O. CHUA AND L. YANG, *Cellular neural networks: Applications*, IEEE Trans. Circuits Syst. I Regul. Pap. 35(1988), pp. 1273-1290.

[14] P. DEBYE AND A.M. BUECHE, *Scattering by an inhomogeneous solid*, J. Applied Physics 20(2004), pp. 518-525.

[15] A. DESAI, *Subsystem Entropies for $\mathbb{Z}^d$ Sofic Systems*, Indagationes Mathematicae 17(2006), pp. 353-359.

[16] G.B. ERMENTROUT, *Stable periodic solutions to discrete and continuum arrays of weakly coupled nonlinear oscillators*, SIAM J. Appl. Math. 52(1992), pp. 1665-1687.

[17] G.B. ERMENTROUT AND N. KOPELL, *Inhibition-produced patterning in chains of coupled nonlinear oscillators*, SIAM J. Appl. Math. 54(1994), pp. 478-507.

[18] G.B. ERMENTROUT, N. KOPELL AND T. L. WILLIAMS, *On chains of oscillators forced at one end*, SIAM J. Appl. Math. 51(1991), pp. 1397-1417.

[19] T. Eveneux and J.P. Laplante, *Propagation failure in arrays of coupled bistable chemical reactors*, J. Phys. Chem. 96(1992), pp. 4931-4934.

[20] M. Fäth, S. Freisem, A.A. Menovsky, Y. Tomioka, J. Aarts and J.A. Mydosh, *Spatially inhomogeneous metal-insulator transition in doped manganites*, Science **285**(1999), pp. 1540-1542.

[21] A.E. Ferdinand, and M.E. Fisher, *Bounded and inhomogeneous ising models. I. Specific-Heat anomaly of a finite lattice*, Phys. Rev. 185(1969), pp. 832-846.

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

[23] C.H. Hsu, and T.-H. Yang, *Abundance of mosaic patterns for CNN with spatially variant templates*, Int. J. Bifurcation and Chaos 12(2002), pp. 1321-1332.

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

[25] Y.A. Kravtsov, and Y.I. Orlov, *Geometrical optics of inhomogeneous media*, Springer-Verlag Berlin Heidelberg New York, 1990.

[26] J.P. Keener, *Propagation and its failure in coupled systems of discrete excitable cells*, SIAM J. Appl. Math. 47(1987), pp. 556-572.

[27] J.P. Keener, *The effects of discrete gap junction coupling on propagation in myocardium*, J. Theor. Biol. 148(1991), pp. 49-82.

[28] S.S. Lin, and C.W. Shih, *Complete stability for standard cellular neural networks*, Int. J. Bifurcation and Chaos 9(1999), pp. 909-918.

[29] S.S. Lin, and T.S. Yang, *Spatial entropy of one-dimensional celluar neural network*, International J. of Bifurcation and Chaos 10(2000), pp. 2129-2140.

[30] S.S. Lin, and T.S. Yang, *On the spatial entropy and patterns of two-dimensional cellular neural networks*, International J. of Bifurcation and Chaos 12(2002), pp. 115-128.

[31] R. Lay, *Convex Sets and Their Applications* Wiley, NY, 1992.

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

[33] J.P. Perdew, *Density-functional approximation for the correlation energy of the inhomogeneous electron gas*, Phys. Rev. B 33(1986), pp. 8822-8824.

[34] Y. Pesin, *Dimension theory in dynamical systems*, The University of Chicago Press, 1997.

[35] A. Quas, and A. Sahin, *Entropy Gaps and Locally Maximal Entropy in $\mathbb{Z}_d$ Subshifts*, Erg. Th. and Dyn. Syst. 23(2003), pp. 1227-1245.

[36] A. Quas, and P. Trow, *Subshifts of Multi-Dimensional Shifts of Finite Type*, Erg. Th. and Dyn. Syst. 20(2000), pp. 859-874.

[37] Y. Rosenfeld, *Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing*, Phys. Rev. Lett., 63(1989), pp. 980-983.

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