Permitted and Forbidden Sets in Discrete-Time Linear Threshold Recurrent Neural Networks

Zhang Yi, Lei Zhang, Jiali Yu, and Kok Kiong Tan

***Abstract*—The concepts of permitted and forbidden sets enable a new perspective of the memory in neural networks. Such concepts exhibit interesting dynamics in recurrent neural networks. This paper studies the basic theories of permitted and forbidden sets of the linear threshold discrete-time recurrent neural networks. The linear threshold transfer function has been regarded as an adequate transfer function for recurrent neural networks. Networks with this transfer function form a class of hybrid analog and digital networks which are especially useful for perceptual computations. Networks in discrete time can directly provide algorithms for efficient implementation in digital hardware. The main contribution of this paper is to establish foundations of permitted and forbidden sets. Necessary and sufficient conditions for the linear threshold discrete-time recurrent neural networks are obtained for complete convergence, existence of permitted and forbidden sets, as well as conditionally multiattractivity, respectively. Simulation studies explore some possible interesting practical applications.**

***Index Terms*—Complete convergence, discrete-time recurrent neural networks, forbidden set, linear threshold, multiattractivity, permitted set.**

I. INTRODUCTION

PERMITTED and forbidden sets of linear threshold recurrent neural networks were first proposed in [1] and further studied in [2] and [3]. If a set of neurons can be coactivated at a stable equilibrium point by some input, then it is a permitted set. Otherwise, it is a forbidden set. Such concepts enable a new perspective of the memory in neural networks. Permitted sets can be regarded as memories stored in the synaptic connections of a network, since by applying an input to the network, some permitted sets can be selected. From a mathematical point of view, the concepts of permitted and forbidden sets provide new dynamics of recurrent neural networks; they deeply describe some interesting dynamical properties. It is believed that permitted and forbidden sets of neural networks can have potential practical applications. In [3], permitted and forbidden sets are successfully applied to solve the problem of group selection by using recurrent neural networks with lateral inhibition.

So far, permitted and forbidden sets have been proposed for an important class of continuous-time recurrent neural networks with linear threshold transfer functions. The linear threshold transfer function is an unbounded function with a binary pattern. Neural networks with this transfer function form a class of hybrid analog and digital networks. The coexistence of analog filtering with logical constraints on neural activation represents a form of hybrid analog–digital computation that may be especially appropriate for perceptual tasks [3]. With the unsaturation property of the transfer function, in [4], it has been argued that this transfer function is more appropriate for recurrent neural networks, because cortical neurons rarely operate close to saturation, despite a strong recurrent excitation. The linear threshold transfer function has been used for modeling many cortical networks [5], [6]. Linear threshold recurrent neural networks (henceforth to be called LT networks) have found many applications, such as associative memory [7], winner-takes-all [8], group selection [3], etc. In [9], LT networks are used to implement a competitive layer model (CLM) for feature binding. The LT networks can be built in silicon. In [1], an efficient silicon design is demonstrated for LT networks and the coexistence of analog amplification and digital selection in network circuits is discussed.

The linear threshold transfer function is essentially nonlinear, and many complex dynamic properties may exist in networks endowed with the transfer function. In [10], some interesting boundedness conditions using local inhibition are reported. Multistability of both continuous- and discrete-time LT networks is studied in [11] and [12]. Multistability is an important type of dynamics in recurrent neural networks. For more results on this topic, we refer to [13]–[19].

This paper studies the permitted and forbidden sets of discrete-time linear threshold recurrent neural networks. Discrete-time recurrent neural networks have been widely studied by many authors; see, for example, [12] and [20]–[26]. Such networks provide direct computational algorithms and they can be easily implemented on digital hardware. Discrete-time networks possess advantages for direct computer simulations over digital simulation of the continuous-time neural network model. Generally speaking, dynamics of a discrete-time networks is different from the corresponding continuous-time networks. For example, under the same conditions, a continuous-time network may be stable, while the corresponding discrete-time network may not. More results on continuous time networks can be found in [8], [11], [14], and [27]. The study of the dynamics of discrete-time neural networks is of prime interest. This paper es-

Manuscript received October 15, 2007; revised July 15, 2008 and December 28, 2008; accepted January 19, 2009. First published May 05, 2009; current version published June 03, 2009. This work was supported by the Chinese 863 High-Tech Program under Grants 2007AA01Z321 and 2008AA01Z11901Z1.

Z. Yi is with the College of Computer Science, Sichuan University, Chengdu 610065, China (e-mail: zhangyi@scu.edu.cn).

L. Zhang is with the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong (e-mail: lzhang@csc.cuhk.edu.hk).

J. Yu is with the School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: yujiali@uestc.edu.cn).

K. K. Tan is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576, Singapore (e-mail: kktan@nus.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNN.2009.2014373

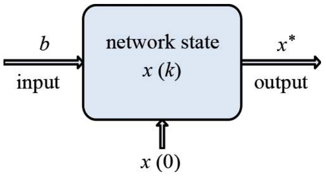

Fig. 1. Relationship between input and output of the network.

tablishes basic theories of permitted and forbidden sets for discrete-time LT networks. The main contribution from this paper is the establishment of necessary and sufficient conditions, respectively, for complete convergence, existence of permitted and forbidden sets, as well as conditional multiattractivity. Simulation studies explore some possible interesting practical applications of the established theory.

The rest of this paper is organized as follows. Preliminaries are given in Section II. Section III studies the complete convergence of the LT discrete-time networks. The theory of permitted and forbidden sets is presented in Section IV. Section V studies multiattractivity. Simulations are given in Section VI to further illustrate the theory. Finally, conclusions are given in Section VII.

II. PRELIMINARIES

The discrete-time linear threshold recurrent neural network model can be described by

$$x(k+1) = [b + Wx(k)]^{+} \tag{1}$$

for $k \geq 0$, where $[s]^{+} = \max\{s, 0\}$ is a rectification nonlinearity, $x(k) \in \mathbb{R}^n$ denotes the state of the network at time k, $W = (W_{ij})_{n\times n}$ is the synaptic weight matrix which is assumed to be symmetric, and $b \in \mathbb{R}^n$ denotes the external input.

A vector $x \in \mathbb{R}^n$ is said to be nonnegative, denoted by $x \geq 0$, if each element of x is nonnegative. A vector $x \in \mathbb{R}^n$ is said to be positive, denoted by $x > 0$, if each element of x is positive.

Denote the nonnegative orthant of the space $\mathbb{R}^n$ by

$$\mathbb{R}^n_{+} = \{x \,|\, x \in \mathbb{R}^n, x \geq 0\}.$$

It is easy to see that any trajectory of the network (1) starting in $\mathbb{R}^n_{+}$ stays in $\mathbb{R}^n_{+}$ indefinitely, i.e., $\mathbb{R}^n_{+}$ is an invariant set of the network (1). In this paper, it is assumed that each trajectory of (1) is restricted to start from a point in the nonnegative orthant $\mathbb{R}^n_{+}$.

A point $x^* \in \mathbb{R}^n_{+}$ is called an equilibrium point of the network (1), if it satisfies

$$x^* = [b + Wx^*]^{+}.$$

An output of the network in response to an input is defined to be some equilibrium point. There are two ways to look at the inputs in recurrent neural networks. One is to fix an external input b and to take initial vectors as network inputs. The other is to fix an initial vector $x(0)$ and to take external inputs as network inputs. In this paper, the latter way is adopted (see Fig. 1 for an intuitive illustration).

Next, the definition of Lyapunov stability for an equilibrium point is given.

Definition 1: An equilibrium x^* is called stable, if given any constant $\epsilon > 0$, there exists a constant $\delta > 0$ such that $||x(0) - x^*|| \leq \delta$ implies that

$$||x(k) - x^*|| \leq \epsilon$$

for all $k \geq 0$. An equilibrium point is called unstable if it is not stable.

The concept of Lyapunov stability is associated with some equilibrium points. The stable and unstable concepts describe some local properties related to equilibrium points. If an equilibrium point is stable, any trajectories starting from the points closed to the equilibrium point will stay close to it. If an equilibrium point is unstable, then for whatever small neighborhood of the equilibrium point, there must exist at least one point in the neighborhood such that the trajectory starting from it will not stay close to the equilibrium point. An equilibrium point can be stable or unstable. However, in practice, only stable equilibrium points can be observed.

Definition 2: Let $x \in \mathbb{R}^n$ be a vector, and let $P \subseteq \{1, 2, \ldots, n\}$ be an index set. Then, x_P is said to be a subvector of x, if x_P can be constructed from x simply by removing from x all components not indexed by P.

Definition 3: Let M be an $n \times n$ matrix, and let $P \subseteq \{1, 2, \ldots, n\}$ be an index set. The matrix M_P is said to be a submatrix of M if the matrix M_P can be constructed from M simply by removing from M all rows and columns not indexed by P.

Given two index sets $P \subseteq \{1, 2, \ldots, n\}$ and $Z = \{1, 2, \ldots, n\} \setminus P$, network (1) can be rewritten as

$$\begin{cases} x_P(k+1) = [b_P + W_P \cdot x_P(k) + W_{PZ} \cdot x_Z(k)]^{+} \\ x_Z(k+1) = [b_Z + W_{ZP} \cdot x_P(k) + W_Z \cdot x_Z(k)]^{+} \end{cases} \tag{2}$$

for $k \geq 0$, where x_P and x_Z are subvectors of x, b_P and b_Z are subvectors of b, W_P and W_Z are submatrices of W constructed by removing from W all rows and columns not indexed, respectively, by P and Z, W_{PZ} is a matrix constructed from W by removing from W all rows not indexed by P and all columns not indexed by Z, and W_{ZP} is a matrix constructed from W by removing from W all rows not indexed by Z and all columns not indexed by P. Clearly, W_{PZ} and W_{ZP} are not submatrices of W.

Definition 4: A matrix M is called copositive if $x^T Mx \geq 0$ for any $x \in \mathbb{R}^n_{+}$, and $x^T Mx = 0$ only if $x = 0$.

Lemma 1 [2]: A matrix M is copositive if and only if all positive eigenvectors of all submatrices of M have positive eigenvalues.

Lemma 2: A matrix M is copositive if and only if there exists a constant $\alpha > 0$ such that

$$x^T Mx \geq \alpha x^T x, \qquad x \in \mathbb{R}^n_{+}.$$

Proof: Define a compact set

$$U = \left\{x \,\middle|\, x \in \mathbb{R}^n_{+}, x^T x = 1\right\} \subset \mathbb{R}^n_{+}.$$

Taking

$$\alpha = \min_{x\in U} \left\{x^T Mx\right\}$$

since U is compact, there exists a vector $x^* \in U$ such that

$$x^{*T} M x^* = \min_{x \in U} \{x^T M x\} = \alpha.$$

Since M is copositive, it follows that $\alpha > 0$. The result now follows and the proof is complete. ■

Next, some properties of the linear threshold function will be presented. These properties play important roles in the dynamical properties of the network.

Lemma 3: It holds that

$$\left([u]^+ - v\right)\left(u - [u]^+\right) \geq 0$$

for any $u \in \mathbb{R}$ and $v \in \mathbb{R}_+$.

Proof: If $u \geq 0$, then $[u]^+ = u$, and

$$\left([u]^+ - v\right)\left(u - [u]^+\right) = 0.$$

If $u < 0$, then $[u]^+ = 0$, and

$$\left([u]^+ - v\right)\left(u - [u]^+\right) = -v \cdot u \geq 0.$$

The result now follows and the proof is complete. ■

Lemma 4: It holds that

$$\left|[u]^+ - [v]^+\right| \leq |u - v|$$

for any $u, v \in \mathbb{R}$.

Proof: The proof is trivial. ■

III. COMPLETE CONVERGENCE

This section studies the convergence of the network (1). An interesting question to address is whether each trajectory will converge to an equilibrium point. This dynamic property of the network is very important in many practical applications. If each trajectory converges to an equilibrium, the network is called completely convergent. Complete convergence of a neural network implies the network has bounded dynamics, the absence of limit cycles, and the absence of chaos. The concept of complete convergence of a network is different from the concept of Lyapunov stability of an equilibrium point. The former considers the convergence of each trajectory, while the latter is concerned with whether all trajectories starting from a small neighborhood of an equilibrium point stay close to this equilibrium point. Many results on Lyapunov stability of neural networks have been reported in recent years because there exists well-established Lyapunov stability theory in mathematics. However, for complete convergence of neural networks, due to the lack of analysis tools in mathematics in this aspect, rigorous results are seldom reported. The most popular method for studying complete stability is to construct energy functions. It should be noted here that energy functions are usually different from Lyapunov functions [14]. Generally, Lyapunov functions require some positiveness while energy functions do not have such a restriction.

To establish the theory of complete convergence of the network, we construct the following energy function:

$$E(x) = \frac{1}{2} x^T (I - W) x - x^T b \tag{3}$$

for $x \in \mathbb{R}^n_+$. This energy function has also been used to study the convergence of continuous-time recurrent neural networks; see [2]. In [12], we constructed another energy function which contains some integral parts for studying the convergence of discrete-time neural networks. The energy function of (3) is simpler since it is a quadratic function while the energy function in [12] is not.

Lemma 5: If $I - W$ is copositive, then the energy function E is lower bounded in $\mathbb{R}^n_+$.

Proof: Since $I - W$ is copositive, by Lemma 2, there must exist a constant $\alpha > 0$ such that

$$\begin{aligned} E(x) &\geq \frac{\alpha}{2} x^T x - x^T b \\ &= \frac{\alpha}{2} \cdot \left[x - \frac{b}{\alpha}\right]^T \cdot \left[x - \frac{b}{\alpha}\right] - \frac{b^T b}{2\alpha} \\ &\geq -\frac{b^T b}{2\alpha} \end{aligned}$$

for all $x \in \mathbb{R}^n_+$. The proof is complete. ■

Define a set

$$U = \left\{ \begin{bmatrix} u_1 & 0 & \cdots & 0 \\ 0 & u_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & u_n \end{bmatrix} \middle| u_i = 0 \text{ or } 1, 1 \leq i \leq n \right\}.$$

Clearly, each element of U is an $n \times n$ matrix and the set U has total 2^n elements. Using this notation, a representation of each trajectory of the network can be obtained.

Lemma 6: Given any $x(0) \in \mathbb{R}^n_+$, the trajectory of the network (1) starting from $x(0)$ can be represented by

$$\begin{aligned} x(k+1) = &\left(\prod_{j=0}^{k} U(k-j) W\right) \cdot x(0) \\ &+ \sum_{i=0}^{k-1} \left(\prod_{j=0}^{i} U(i-j) W\right) b + U(k) b \end{aligned}$$

for $k \geq 0$, where each $U(i) \in U$. Moreover, every component of $x(k+1)$ must be a polynomial of the elements of W with order no more than k.

Proof: At each step k, there exists a matrix $U(k) \in U$ such that

$$x(k+1) = U(k)\left[W x(k) + b\right]$$

for $k \geq 0$. With iterations, the trajectory representation follows.

Since the set U has 2^n elements, from the trajectory representation, each component of $x(k+1)$ must be a polynomial of the elements of W with order of at most k. The proof is complete. ■

Definition 5: The network (1) is said to be completely convergent, if each trajectory of the network converges to an equilibrium point.

Theorem 1: Suppose that the matrix $I + W$ is positive definite and the matrix $I - W$ is copositive, then the network (1) is completely convergent for any input b and any initial conditions.

Proof: Denote

$$\theta(k+1) = x(k+1) - x(k)$$

for $k \geq 0$. It can be calculated that

$$\begin{aligned} &E(x(k+1)) - E(x(k)) \\ &\quad = -\frac{1}{2}\theta^T(k+1)\,(I+W)\,\theta(k+1) \\ &\qquad -\theta^T(k+1)\,[Wx(k)+b-x(k+1)]\,. \end{aligned}$$

By Lemma 3, we have

$$\theta^T(k+1)\cdot[Wx(k)+b-x(k+1)] \geq 0$$

for $k \geq 0$. Then

$$\begin{aligned} &E(x(k+1)) - E(x(k)) \\ &\quad \leq -\frac{1}{2}\theta^T(k+1)\,(I+W)\,\theta(k+1) \\ &\quad \leq -\eta\|x(k+1)-x(k)\|^2 \end{aligned} \tag{4}$$

for $k \geq 0$, where $2\eta(\eta > 0)$ is the smallest eigenvalue of the matrix $I+W$. Then

$$\|x(k+1)-x(k)\|^2 \leq \frac{1}{\eta}\,[E(x(k)) - E(x(k+1))] \tag{5}$$

for $k \geq 0$.

By (4), $E(x(k))$ is monotone decreasing. Since $I-W$ is copositive, by Lemma 5, the energy function E is lower bounded. Hence, the sequence $\{E(x(k))\}$ is a monotone and bounded sequence. A limit must exist, i.e., there exists a constant E_0 such that

$$\lim_{k\longrightarrow+\infty} E(x(k)) = E_0.$$

By (5), it follows that

$$\sum_{k=0}^{+\infty}\|x(k+1)-x(k)\|^2 \leq \frac{1}{\eta}\,[E(x(0))-E_0] < +\infty.$$

Thus

$$\lim_{k\longrightarrow+\infty}\|x(k+1)-x(k)\| = 0. \tag{6}$$

Using Lemma 6, $\|x(k+1)-x(k)\|$ must be a polynomial of the elements of W with an order no more than k. Inferring from this fact together with (6), there must exist constants $C > 0$ and $0 < \beta < 1$ such that

$$\|x(k+1)-x(k)\| \leq C\cdot\beta^k$$

for $k \geq 0$. Given any $\epsilon > 0$, there exists a constant $K > 0$ such that

$$\frac{C\cdot\beta^K}{1-\beta} \leq \epsilon.$$

Given any $k_1 > k_2 \geq K$, it follows that

$$\begin{aligned} \|x(k_1)-x(k_2)\| &= \left\|\sum_{r=k_2}^{k_1-1}[x(r+1)-x(r)]\right\| \\ &\leq \sum_{r=k_2}^{k_1-1}\|x(r+1)-x(r)\| \\ &\leq C\cdot\sum_{r=k_2}^{k_1-1}\beta^r \\ &\leq \frac{C\cdot\beta^K}{1-\beta} \\ &\leq \epsilon. \end{aligned}$$

This shows that the sequence $\{x(k)\}$ is a *Cauchy sequence*. By the *Cauchy convergence principle*, there must exist a vector $x^* \in \mathbb{R}^n_+$ such that

$$\lim_{k\longrightarrow+\infty} x(k) = x^*.$$

It is easy to see that x^* is an equilibrium point of network (1) and the proof is complete. ■

Theorem 2: Suppose that the matrix $I+W$ is positive definite. Then, the matrix $I-W$ is copositive if and only if for all constant input b the network is completely convergent.

Proof: If the matrix $I-W$ is copositive, using Theorem 1, for any input b the network is completely convergent.

Next, suppose that the network is completely convergent for any input b; we will show that $I-W$ is copositive by using the method of counter proof. Suppose that $I-W$ is not copositive. Define a compact set

$$V = \left\{x\,\middle|\,x\in\mathbb{R}^n_+, x^Tx = 1\right\}.$$

Denote

$$\alpha = \min_{x\in V}\left\{x^T\,(I-W)\,x\right\}.$$

Since $I-W$ is not copositive, clearly, $\alpha \leq 0$. Let x^* be the minimum of $Q = x^T(I-W)x$ on V. Then, $x^* = \operatorname{argmin}_{x\in\mathbb{R}^n_+} G$, where $G = Q - \alpha(x^Tx - 1)$. Define

$$\begin{cases} P = \{i\,|x_i^* > 0, 1 \leq i \leq n\} \\ Z = \{i\,|x_i^* = 0, 1 \leq i \leq n\}\,. \end{cases}$$

Then

$$\begin{cases} \left.\dfrac{\partial G}{\partial x_P}\right|_{x=x^*} = 0 \\ \left.\dfrac{\partial G}{\partial x_Z}\right|_{x=x^*} \geq 0. \end{cases}$$

It follows that

$$\begin{cases} W_P\cdot x_P^* = (1-\alpha)\,x_P^* \\ W_{ZP}\cdot x_P^* \leq 0. \end{cases}$$

Choosing an input as

$$\begin{cases} b_P = x_P^* \\ b_Z = 0 \end{cases}$$

it is easy to check that the trajectory of (1) with initial condition

$$\begin{cases} x_P(0) = x_P^* \\ x_Z(0) = 0 \end{cases}$$

satisfies that

$$\begin{cases} x_P(k) = x_P^*\cdot\sum\limits_{r=0}^{k}(1-\alpha)^r \\ x_Z(k) = 0 \end{cases}$$

for $k \geq 0$. Since $\alpha \leq 0$, then

$$x_P(k) = x_P^*\cdot\sum_{r=0}^{k}(1-\alpha)^r \longrightarrow +\infty$$

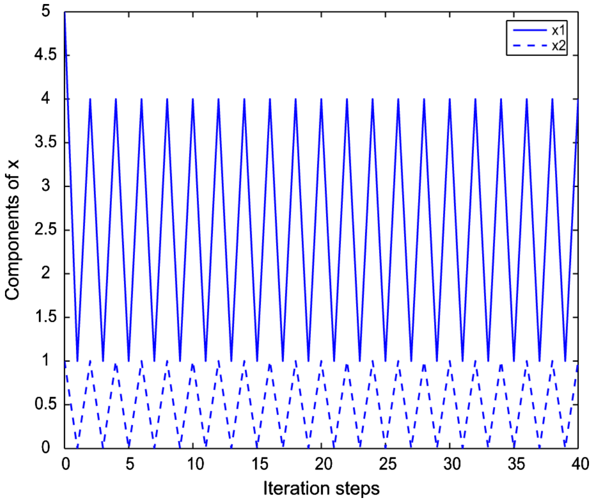

Fig. 2. A periodic trajectory of network (7). It illustrates that the copositive of the matrix $I - W$ is not sufficient for complete convergence.

as $k \longrightarrow +\infty$. Clearly, the trajectory is not convergent. This is a contradiction and it proves that the matrix $I - W$ must be copositive. The proof is complete. ■

Requiring the matrix $I + W$ to be positive definite is an important condition. This is different from the continuous-time LT networks [2]. For continuous-time LT networks, if W is symmetric, then the copositive of $I - W$ is sufficient for complete convergence [2]. However, for discrete-time LT network of (1), copositive of $I - W$ is not sufficient to guarantee complete convergence. This can be illustrated by the following simple 2-D network:

$$x(k+1) = \left[\begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} x(k) + \begin{pmatrix} 4 \\ 4 \end{pmatrix}\right]^+ \tag{7}$$

for $k \geq 0$. Clearly, the synaptic weight matrix

$$W = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$$

is symmetric and $I - W$ is copositive. However, the matrix $I + W$ is not positive definite. The network is not completely convergent. In fact, the network has many periodic trajectories. For example

$$\left\{\begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}.$$

See Fig. 2 for illustration. Generally, discrete-time networks have different dynamics from their corresponding continuous-time networks.

IV. PERMITTED AND FORBIDDEN SETS

The concepts of permitted and forbidden sets were first proposed in [1]. These concepts are closely related to Lyapunov stability of equilibrium points and the network input b. A set of neurons is said to be permitted if the neurons can be coactivated at a stable equilibrium point by some input b. A set of neurons is said to be forbidden if they are not permitted, i.e., the neurons cannot be coactivated at any stable equilibrium point regardless of what the input b is. The mathematical definitions for permitted and forbidden sets can be given as follows.

Definition 6: A set of neurons with index set P is said to be permitted if there exists an input b such that the network (1) has a stable equilibrium point x^* with the property

$$\begin{cases} x_i^* > 0, & i \in P \\ x_j^* = 0, & j \notin P. \end{cases}$$

A set of neurons is said to be forbidden if it is not permitted.

Theorem 3: A set of neurons with index set P is permitted if and only if each eigenvalue λ of W_P satisfies $|\lambda| \leq 1$.

Proof: Suppose the set of neurons with index set P is permitted. Then, there exists an input b such that the network has a stable equilibrium x^* with the property

$$\begin{cases} x_i^* > 0, & i \in P \\ x_j^* = 0, & j \in Z = \{1, \ldots, n\} \setminus P. \end{cases}$$

Suppose there exists an eigenvalue λ of W_P with eigenvector v satisfying $|\lambda| > 1$, then we will show that this will lead to a contradiction.

Since x^* is stable, given any

$$0 < \epsilon < \min\left\{\frac{x_i^*}{\|W_P\| + \|W_{PZ}\|}\middle| i \in P\right\}$$

there must exist a constant $\delta > 0$ such that $\|x(0) - x^*\| < \delta$ implies that

$$\begin{cases} \|x_P(k) - x_P^*\| \leq \epsilon \\ \|x_Z(k)\| \leq \epsilon \end{cases}$$

and

$$\begin{cases} \left|v^T\left(x_P(k) - x_P^*\right)\right| \leq \epsilon \\ \left|v^T W_{PZ} \cdot x_Z(k)\right| \leq \epsilon \end{cases}$$

for all $k \geq 0$. Then, from (2), it follows that

$$x_P(k+1) - x_P^* = W_P \cdot (x_P(k) - x_P^*) + W_{PZ} \cdot x_Z(k)$$

for all $k \geq 0$. By iteration, it gives that

$$x_P(k) - x_P^* = W_P^k \cdot (x_P(0) - x_P^*) + \sum_{j=0}^{k-1} W_P^{k-1-j} \cdot W_{PZ} \cdot x_Z(j)$$

for all $k \geq 0$. Then

$$\begin{aligned} &v^T\left(x_P(k) - x_P^*\right) \\ &\quad = \lambda^k \cdot v^T\left(x_P(0) - x_P^*\right) + \sum_{j=0}^{k-1} \lambda^{k-1-j} \cdot v^T W_{PZ} \cdot x_Z(j) \\ &\quad = \lambda^k \cdot \left[v^T(x_P(0) - x_P^*) + \sum_{j=0}^{k-1}\left(\frac{1}{\lambda}\right)^{j+1} \cdot \left(v^T W_{PZ} \cdot x_Z(j)\right)\right] \end{aligned}$$

for all $k \geq 0$. Since $|\lambda| > 1$ and $\left|v^T W_{PZ} \cdot x_Z(k)\right| \leq \epsilon$, the power series

$$\sum_{j=0}^{+\infty}\left(\frac{1}{\lambda}\right)^{j+1} \cdot \left(v^T W_{PZ} \cdot x_Z(j)\right)$$

must converge to some constant c. It must exist $x(0)$ such that

$$\begin{cases} \|x(0)-x^*\| < \delta \\ v^T\left(x_P(0)-x_P^*\right)+c \neq 0. \end{cases}$$

Then, it holds that

$$\begin{aligned} &\lim_{k\longrightarrow+\infty}\left|v^T\left(x_P(k)-x_P^*\right)\right| \\ &\quad=\lim_{k\longrightarrow+\infty}|\lambda|^k\cdot\left|v^T\left(x_P(0)-x_P^*\right)+c\right| \\ &\quad=+\infty. \end{aligned}$$

This contradicts that x^* is stable. Thus, each eigenvalue λ of W_P must satisfy $|\lambda| \le 1$.

Next, suppose that each eigenvalue λ of W_P satisfies $|\lambda| \le 1$. We will prove that the set of neurons with index set P is permitted.

Define a point $x^* \in \mathbb{R}_+^n$ by

$$\begin{cases} x_i^* = 1, & i \in P \\ x_j^* = 0, & j \in Z. \end{cases}$$

Selecting an input b as

$$b_i = \begin{cases} 1-\sum\limits_{j\in P} w_{ij}, & i \in P \\ -1-\sum\limits_{j\in P} w_{ij}, & i \in Z \end{cases}$$

it can be checked that x^* is an equilibrium point of network (1). Next, we will prove that x^* is stable.

Since x^* is an equilibrium point, it holds that

$$\begin{cases} x_P^* = b_P + W_P x_P^* \\ b_Z + W_{ZP}x_P^* = -1_Z < 0 \end{cases}$$

where 1_Z is a vector with each element as 1. Then, network (1) can be rewritten as

$$\begin{cases} x_P(k+1) = \left[x_P^* + W_P\cdot(x_P(k)-x_P^*)+W_{PZ}\cdot x_Z(k)\right]^+ \\ x_Z(k+1) = \left[-1_Z+W_{ZP}\cdot(x_P(k)-x_P^*)+W_Z\cdot x_Z(k)\right]^+ \end{cases} \tag{8}$$

for $k \ge 0$.

Given a constant ϵ such that

$$0 < \epsilon \le \min\left\{\frac{1}{\|W_P\|+1}, \frac{1}{\|W_{PZ}\|+1}, \frac{1}{\|W_Z\|+1}\right\}$$

define a constant δ by

$$\delta = \frac{\epsilon}{2}\cdot\min\left\{\frac{1}{\|W_P\|+1}, \frac{1}{\|W_{PZ}\|+1}\right\} < \frac{\epsilon}{2}.$$

Define a neighborhood B_δ of x^* by

$$B_\delta = \left\{\, x \in \mathbb{R}_+^n \,\middle|\, |x_i - x_i^*| \le \delta (i \in P); |x_j| \le \delta (j \in Z)\right\}.$$

Define another neighborhood B_ϵ of x^* by

$$B_\epsilon = \left\{\, x \in \mathbb{R}_+^n \,\middle|\, |x_i - x_i^*| \le \epsilon (i \in P); x_j = 0 (j \in Z)\right\}.$$

We will prove by mathematical induction that for any $x(0) \in B_\delta$, it holds that $x(k) \in B_\epsilon$ for all $k \ge 1$.

By $x(0) \in B_\delta$, we have

$$\begin{aligned} x_Z(1) &= \left[-1_Z + W_{ZP}\cdot(x_P(0)-x_P^*)+W_Z x_Z(0)\right]^+ \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} x_P(1) &= \left[x_P^* + W_P\cdot(x_P(0)-x_P^*)+W_{PZ}\cdot x_Z(0)\right]^+ \\ &= x_P^* + W_P\cdot(x_P(0)-x_P^*)+W_{PZ}\cdot x_Z(0). \end{aligned}$$

Then

$$\begin{aligned} &\|x_P(1)-x_P^*\| \\ &\quad\le \|W_P\|\cdot\|x_P(0)-x_P^*\| + \|W_{PZ}\|\cdot\|x_Z(0)\| \\ &\quad\le \frac{\epsilon}{2}+\frac{\epsilon}{2} \\ &\quad= \epsilon. \end{aligned}$$

This proves that $x(1) \in B_\epsilon$. Next, suppose $x(k) \in B_\epsilon$, then we will prove that $x(k+1) \in B_\epsilon$. In fact, by $x(k) \in B_\epsilon$, it follows that

$$\begin{aligned} x_Z(k+1) &= \left[b_Z + W_{ZP}\cdot x_P(k)\right]^+ \\ &= \left[-1_Z + W_{ZP}\cdot(x_P(k)-x_P^*)\right]^+ \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} x_P(k+1) &= \left[x_P^* + W_P\cdot(x_P(k)-x_P^*)\right]^+ \\ &= x_P^* + W_P\cdot(x_P(k)-x_P^*) \end{aligned}$$

then

$$\|x_P(k+1)-x_P^*\| \le \max\{|\lambda|\}\cdot\|x_P(k)-x_P^*\| \le \epsilon. \tag{9}$$

The above shows that $x(k+1) \in B_\epsilon$.

Thus, any trajectory starting from a point in B_δ satisfies

$$\begin{cases} x_P(k+1)-x_P^* = W_P\cdot[x_P(k)-x_P^*] \\ x_Z(k+1) = 0 \end{cases} \tag{10}$$

for $k \ge 0$. Since each eigenvalue λ of W_P satisfies $|\lambda| \le 1$, clearly, x^* is stable, and so the set of neurons with the index set P is permitted. The proof is complete. ■

Theorem 4: A set of neurons with the index set F is forbidden if and only if there exists an eigenvalue of W_F with its absolute value larger than one.

Proof: The result follows directly from the definition of forbidden set and Theorem 3. ■

To further study the properties of permitted and forbidden sets, we need the following lemma, which is often called the interlacing lemma.

Lemma 7 [28]: Let M be a symmetric $n \times n$ matrix with real eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ and let M_s be an $(n-1)\times(n-1)$ submatrix of M. Then the eigenvalues $\eta_1 \le \eta_2 \le \cdots \le \eta_{n-1}$ of M_s interlace the eigenvalues of M, i.e., $\lambda_1 \le \eta_1 \le \lambda_2 \le \cdots \le \eta_{n-1} \le \lambda_n$.

Theorem 5: A forbidden set exists in network (1) if and only if there exists an eigenvalue of W such that its absolute value is larger than one.

Proof: If an eigenvalue λ of W satisfies that $|\lambda| > 1$, then, by Theorem 4, the set of neurons with the index set $\{1, 2, \ldots, n\}$ must be forbidden.

Suppose there exists a set of neurons where index set F is forbidden, then, by Theorem 4, there must exist an eigenvalue λ_f of W_F such that $|\lambda_f| > 1$. Using Lemma 7, matrix W must have an eigenvalue λ such that $|\lambda| > 1$. The proof is complete. ■

Theorem 6: Any subset of a permitted set is permitted. Any superset of a forbidden set is forbidden.

Proof: The result directly follows from Lemma 7, Theorem 3, and Theorem 4. ■

V. MULTIATTRACTIVITY

Define the set of equilibrium points of the network (1) by

$$\mathcal{E} = \left\{ x \,\middle|\, x = [Wx + b]^+, x \in \mathbb{R}^n \right\}.$$

The set of stable equilibrium points, denoted by S_e, is a subset of the set $\mathcal{E}$, i.e., $S_e \subset \mathcal{E}$. The set S_e can be connected or disconnected. If the set S_e is connected, we say the network possesses continuous attractors. If the set S_e is disconnected, we say the network possesses multiattractive property. Both dynamical properties are very interesting. In this section, we study the conditional multiattractivity, i.e., the disconnection of set S_e by selecting some input b.

The set of equilibrium points $\mathcal{E}$ is closely related to the set of minimum points of the energy function

$$E(x) = \frac{1}{2} x^T (I - W) x - x^T b, \qquad x \in \mathbb{R}^n_+.$$

Denote the set of minimum points of the energy function E on the nonnegative orthant $\mathbb{R}^n_+$ by

$$\mathcal{M} = \left\{ x \,\middle|\, x = \operatorname{argmin}_{x \in \mathbb{R}^n_+} E(x) \right\}.$$

Next, we study the relationships among the three sets $\mathcal{E}, S_e$, and $\mathcal{M}$.

Lemma 8: It holds that $\mathcal{M} \subseteq \mathcal{E}$.

Proof: Given any $x^* \in \mathcal{M}$, define two index sets by

$$\begin{cases} P = \{ i \,|\, x_i^* > 0, 1 \le i \le n \} \\ Z = \{ i \,|\, x_i^* = 0, 1 \le i \le n \}. \end{cases}$$

Then

$$\begin{cases} \left.\dfrac{\partial E}{\partial x_P}\right|_{x = x^*} = 0 \\ \left.\dfrac{\partial E}{\partial x_Z}\right|_{x = x^*} \ge 0. \end{cases}$$

It gives that

$$\begin{cases} x_P^* = W_P \cdot x_P^* + b_P \\ W_{ZP} \cdot x_P^* + b_Z \le 0. \end{cases}$$

Clearly, it follows that

$$x^* = [Wx^* + b]^+$$

i.e., $x^* \in \mathcal{E}$. Thus, $\mathcal{M} \subseteq \mathcal{E}$. The proof is complete. ■

Lemma 9: Suppose that each eigenvalue λ of W satisfies $|\lambda| \le 1$. Then, $S_e = \mathcal{E} = \mathcal{M}$. Moreover, the set of stable equilibrium points S_e is connected.

Proof: We first prove that $\mathcal{E} = \mathcal{M}$. By Lemma 8, $\mathcal{M} \subseteq \mathcal{E}$, and it is sufficient to prove that $\mathcal{E} \subseteq \mathcal{M}$. Given any $x^* \in \mathcal{E}$, i.e.,

$$x^* = [Wx^* + b]^+$$

we will prove that $E(x^*) \le E(x)$ for any $x \in \mathbb{R}^n_+$. Denote

$$\theta = x^* - x.$$

It can be calculated that

$$\begin{aligned} &E(x^*) - E(x) \\ &\quad = -\frac{1}{2}\theta^T (I - W) \theta + \theta^T [x^* - Wx^* - b] \\ &\quad = -\frac{1}{2}\theta^T (I - W) \theta + \theta^T \cdot \Big[[Wx^* + b]^+ - [Wx^* + b] \Big]. \end{aligned}$$

By Lemma 3

$$\theta^T \cdot \Big[[Wx^* + b]^+ - [Wx^* + b] \Big] \le 0.$$

Then

$$E(x^*) - E(x) \le -\frac{1}{2}\theta^T (I - W) \theta.$$

Since each eigenvalue λ of W satisfies $|\lambda| \le 1$, then $I - W$ is positive semidefinite, and it follows that $E(x^*) \le E(x)$ for $x \in \mathbb{R}^n_+$. This shows that x^* is a minimum point of E, i.e., $x^* \in \mathcal{M}$, and so $\mathcal{E} \subseteq \mathcal{M}$. Using Lemma 8, it follows that $\mathcal{M} = \mathcal{E}$.

Next, we prove that $S_e = \mathcal{E}$. Since $S_e \subseteq \mathcal{E}$, it is sufficient to prove that $\mathcal{E} \subseteq S_e$. Given any $x^* \in \mathcal{E}$, we will show that x^* is a stable equilibrium point, i.e., $x^* \in S_e$. From (1), it follows that

$$x(k+1) - x^* = [b + Wx(k)]^+ - [b + Wx^*]^+$$

for $k \ge 0$. Then, by Lemma 4

$$\begin{aligned} \|x(k+1) - x^*\|^2 &\le \|W[x(k) - x^*]\|^2 \\ &= [x(k) - x^*]^T W^2 [x(k) - x^*] \end{aligned}$$

for $k \ge 0$. Let $\lambda_i (i = 1, \ldots, n)$ be the eigenvalues of W, then $\lambda_i^2 (i = 1, \ldots, n)$ must be the eigenvalues of W^2. By the condition $|\lambda_i| \le 1 (i = 1, \ldots, n)$, $\lambda_i^2 \le 1 (i = 1, \ldots, n)$. Thus

$$[x(k) - x^*]^T W^2 [x(k) - x^*] \le \|x(k) - x^*\|^2$$

for $k \ge 0$. Then

$$\begin{aligned} \|x(k+1) - x^*\| &\le \|x(k) - x^*\| \\ &\le \|x(0) - x^*\| \end{aligned}$$

for $k \ge 0$. By Definition 1 for stable equilibrium points, clearly, x^* is stable, i.e., $x^* \in S_e$. This shows that $\mathcal{E} \subseteq S_e$, and so $\mathcal{E} = S_e$.

Thus, we have proved that $S_e = \mathcal{E} = \mathcal{M}$. Since the matrix $I - W$ is positive semidefinite, then the energy function E must be a convex function; using this fact, it is easy to show that $\mathcal{M}$ is a connected set. Thus, S_e is also a connected set. The proof is complete. ■

Definition 7: The network is said to be conditionally multiattractive if there exists an input b such that the set of stable equilibrium points is disconnected.

Lemma 10: Suppose $I+W$ is positive definite and $I-W$ is copositive. Then, given any constant $c \in (0,+\infty)$, the set

$$D_c = \left\{x \,\middle|\, x \in \mathbb{R}_+^n, E(x) \le c\right\}$$

must be a bounded invariant set, and

$$\lim_{c \longrightarrow +\infty} D_c = \mathbb{R}_+^n.$$

Moreover, given any index set Z, the components $b_i (i \in Z)$ of an input b can be constructed so that any trajectory starting in D_c satisfies

$$x_i(k) = 0, \qquad i \in Z$$

for all $k \ge 1$.

Proof: Given any $x(0) \in D_c$, i.e., $E(x(0)) \le c$, since $I+W$ is positive definite, from the proof of Theorem 1, it follows that

$$E(x(k+1)) \le E(x(k))$$

for all $k \ge 1$. Then, $E(x(k)) \le E(x(0)) \le c$ for $k \ge 0$. This implies that $x(k) \in D_c$ for $k \ge 0$, thus, the set D_c is an invariant set.

Since $I-W$ is copositive, by Lemma 2, there exists a constant $\alpha > 0$ such that

$$E(x) \ge \alpha x^T x - b^T x, \qquad x \in \mathbb{R}_+^n.$$

Then, $E(x) \longrightarrow +\infty$ as $\|x\| \longrightarrow +\infty$. This implies that D_c must be a bounded set, and

$$\lim_{c \longrightarrow +\infty} D_c = \mathbb{R}_+^n.$$

Choose a constant $\Pi > 0$ such that

$$\|x\| \le \Pi, \qquad \text{for } x \in D_c.$$

Taking

$$b_i = -\Pi\|W\| - 1, \qquad i \in Z$$

then

$$\sum_{j=1}^{n} w_{ij} x_j(k) + b_i \le 0, \qquad i \in Z.$$

It follows that $x_i(k) = 0 (i \in Z)$ for all $k \ge 1$. The proof is complete. ■

Theorem 7: Suppose that $I+W$ is positive definite and $I-W$ is copositive. Then, the network (1) is conditionally multiattractive if and only if W has an eigenvalue with its absolute value larger than one.

Proof: Suppose the network is conditionally multiattractive, then we will prove that matrix W must have an eigenvalue with its absolute value large than one. If this is not true, then each eigenvalue λ of W satisfies that $|\lambda| \le 1$. By Lemma 9, the set S_e of the stable equilibrium points of the network is a connected set for any b. This contradicts the presumption that the network is conditionally multiattractive.

Next, suppose that W has an eigenvalue with its absolute value larger than one, then we will prove the network is conditionally multiattractive.

Clearly, $I-W$ is not positive semidefinite, i.e., it has at least one negative eigenvalue. By assumption, $I-W$ is copositive, then any single neuron is permitted. Thus, $I-W$ has at least one nonnegative eigenvalue. Using Lemma 7, there must exist two neurons with index set $F = \{f_1, f_2\}$ such that the submatrix $(I-W)_F$ has one negative eigenvalue and one nonnegative eigenvalue. Clearly, the two neurons with the index set F are forbidden. Denote an index set $Z = \{1, 2, \ldots, n\} \setminus F$. Given any sufficiently large constant $c \in (0,+\infty)$, by using Lemma 10, there exists a bounded invariant set $D_c \subset \mathbb{R}_+^n$. Moreover, components $b_i (i \in Z)$ of an input b can be constructed so that any trajectory starting in D_c satisfies

$$x_i(k) = 0, \qquad i \in Z$$

for all $k \ge 1$. Then, in the invariant set D_c, network (1) can be rewritten as

$$\begin{cases} x_{f_1}(k+1) = [b_{f_1} + w_{f_1 f_1} \cdot x_{f_1}(k) + w_{f_1 f_2} \cdot x_{f_2}(k)]^+ \\ x_{f_2}(k+1) = [b_{f_2} + w_{f_2 f_1} \cdot x_{f_1}(k) + w_{f_2 f_2} \cdot x_{f_2}(k)]^+ \\ x_i(k+1) = 0, \qquad i \in Z \end{cases} \tag{11}$$

for $k \ge 0$.

Let S_e^c be the set of stable equilibrium points of network (11) in D_c. Given any $x^* \in S_e^c$, it must hold that $x_i^* = 0 (i \in Z)$. Moreover, since the set of neurons with index set F is forbidden, $x_{f_1}^*$ or $x_{f_2}^*$ cannot be both positive, i.e., one of them must be zero. Next, we construct b_{f_1} and b_{f_2} so that S_e^c is a disconnected set.

By the condition that $I-W$ is copositive, it is easy to see that $1 - w_{ii} > 0 (i = 1, \ldots, n)$. By $I+W$ being positive definite, it follows that $1 + w_{ii} > 0 (i = 1, \ldots, n)$. Thus, it holds that

$$|w_{ii}| < 1, \qquad i = 1, \ldots, n. \tag{12}$$

Let $\lambda_f < 0$ be the negative eigenvalue of $(I-W)_F$ and $v = [v_{f_1}, v_{f_2}]^T$ be the corresponding eigenvector, then

$$\begin{cases} (1 - w_{f_1 f_1}) \cdot v_{f_1} - w_{f_1 f_2} \cdot v_{f_2} = \lambda_f \cdot v_{f_1} \\ -w_{f_2 f_1} \cdot v_{f_1} + (1 - w_{f_2 f_2}) \cdot v_{f_2} = \lambda_f \cdot v_{f_2}. \end{cases} \tag{13}$$

By Lemma 1, any negative eigenvalue of $(I-W)_F$ cannot have positive eigenvector; it must hold that $v_{f_1} \cdot v_{f_2} < 0$. Without loss of generality, assume that $v_{f_1} > 0$ and $v_{f_2} < 0$. Then, it follows from (12) and (13) that $w_{f_1 f_2} = w_{f_2 f_1} < 0$.

Taking $b_{f_1} = -v_{f_2}$ and $b_{f_2} = v_{f_1}$, define a point $x^\dagger \in \mathbb{R}_+^n$ as

$$\begin{cases} x_{f_1}^\dagger = \dfrac{b_{f_1}}{1 - w_{f_1 f_1}} > 0 \\ x_i^\dagger = 0, \qquad i \ne f_1 \end{cases}$$

and define another point $x^\ddagger$ as

$$\begin{cases} x_{f_2}^\ddagger = \dfrac{b_{f_2}}{1 - w_{f_2 f_2}} > 0 \\ x_i^\ddagger = 0, \qquad i \ne f_2. \end{cases}$$

It can be checked that network (11) has only $x^\dagger$ and $x^\ddagger$ as its equilibrium points on the boundary of $\mathbb{R}^n_+$. Moreover, by (13), it holds that

$$\begin{cases} b_{f_2} + w_{f_2 f_1} \cdot x^\dagger_{f_1} < 0 \\ b_{f_1} + w_{f_1 f_2} \cdot x^\ddagger_{f_2} < 0. \end{cases} \tag{14}$$

Rewrite (11) as

$$\begin{cases} x_{f_1}(k+1) = \Big[x^\dagger_{f_1} + w_{f_1 f_1} \cdot \left(x_{f_1}(k) - x^\dagger_{f_1}\right) \\ \qquad\qquad + w_{f_1 f_2} \cdot x_{f_2}(k)\Big]^+ \\ x_{f_2}(k+1) = \Big[b_{f_2} + w_{f_2 f_1} \cdot x^\dagger_{f_1} + w_{f_2 f_2} \cdot x_{f_2}(k) \\ \qquad\qquad + w_{f_2 f_1} \cdot \left(x_{f_1}(k) - x^\dagger_{f_1}\right)\Big]^+ \\ x_i(k+1) = 0, \qquad i \in Z \end{cases} \tag{15}$$

for $k \geq 0$. Define a neighborhood of $x^\dagger$ by

$$H^\dagger = \left\{ x \in \mathbb{R}^n_+ \,\middle|\, \begin{array}{l} \left|x_{f_1} - x^\dagger_{f_1}\right| \leq \Delta^\dagger_1, |x_{f_2}| \leq \Delta^\dagger_2 \\ x_i = 0,\ i \neq f_1, i \neq f_2 \end{array} \right\}$$

where

$$\begin{cases} \Delta^\dagger_1 = \dfrac{1}{2} \cdot \min\left\{ \dfrac{x^\dagger_{f_1}}{|w_{f_1 f_1}|}, \dfrac{\left|b_{f_2} + w_{f_2 f_1} \cdot x^\dagger_{f_1}\right|}{|w_{f_2 f_1}|} \right\} \\ \Delta^\dagger_2 = \dfrac{1}{2} \cdot \min\left\{ \dfrac{x^\dagger_{f_1}}{|w_{f_1 f_2}|}, \dfrac{\left|b_{f_2} + w_{f_2 f_1} \cdot x^\dagger_{f_1}\right|}{|w_{f_2 f_2}|} \right\}. \end{cases}$$

Given any $x(0) \in H^\dagger$, it is easy to see that the trajectory of (15) starting from $x(0)$ satisfies

$$\begin{cases} x_{f_1}(k) = x^\dagger_{f_1} + w^{k-1}_{f_1 f_1} \cdot \left[x_{f_1}(1) - x^\dagger_{f_1}\right] \\ x_i(k) = 0, \qquad i \neq f_1 \end{cases}$$

for $k \geq 1$. Clearly, the trajectory converges to the equilibrium point $x^\dagger$ and thus $x^\dagger$ is stable, i.e., $x^\dagger \in S^c_e$.

Rewrite (11) as

$$\begin{cases} x_{f_1}(k+1) = \Big[b_{f_1} + w_{f_1 f_2} \cdot x^\ddagger_{f_2} + w_{f_1 f_1} \cdot x_{f_1}(k) \\ \qquad\qquad + w_{f_1 f_2} \cdot \left(x_{f_2}(k) - x^\ddagger_{f_2}\right)\Big]^+ \\ x_{f_2}(k+1) = \Big[x^\ddagger_{f_2} + w_{f_2 f_2} \cdot \left(x_{f_2}(k) - x^\ddagger_{f_2}\right) \\ \qquad\qquad + w_{f_2 f_1} \cdot x_{f_1}(k)\Big]^+ \\ x_i(k+1) = 0, \qquad i \in Z \end{cases} \tag{16}$$

for $k \geq 0$. Define

$$H^\ddagger = \left\{ x \in \mathbb{R}^n_+ \,\middle|\, \begin{array}{l} \left|x_{f_2} - x^\ddagger_{f_2}\right| \leq \Delta^\ddagger_1, |x_{f_1}| \leq \Delta^\ddagger_2 \\ x_i = 0, \quad i \neq f_1, i \neq f_2 \end{array} \right\}$$

where

$$\begin{cases} \Delta^\ddagger_1 = \dfrac{1}{2} \cdot \min\left\{ \dfrac{x^\ddagger_{f_2}}{|w_{f_2 f_2}|}, \dfrac{\left|b_{f_1} + w_{f_1 f_2} \cdot x^\ddagger_{f_2}\right|}{|w_{f_1 f_2}|} \right\} \\ \Delta^\ddagger_2 = \dfrac{1}{2} \cdot \min\left\{ \dfrac{x^\ddagger_{f_2}}{|w_{f_2 f_1}|}, \dfrac{\left|b_{f_1} + w_{f_1 f_2} \cdot x^\ddagger_{f_2}\right|}{|w_{f_1 f_1}|} \right\}. \end{cases}$$

Given any $x(0) \in H^\ddagger$, it is easy to see that the trajectory of (16) starting from $x(0)$ satisfies that

$$\begin{cases} x_{f_2}(k) = x^\ddagger_{f_2} + w^{k-1}_{f_2 f_2} \cdot \left[x_{f_2}(1) - x^\ddagger_{f_2}\right] \\ x_i(k) = 0, \qquad i \neq f_2 \end{cases}$$

for $k \geq 1$. Clearly, the trajectory converges to the equilibrium point $x^\ddagger$ and thus $x^\ddagger$ is stable, i.e., $x^\ddagger \in S^c_e$.

The above shows that $S^c_e = \{x^\dagger, x^\ddagger\}$. Since $\lim_{c \longrightarrow +\infty} D_c = \mathbb{R}^n_+$, then

$$S_e = \left\{x^\dagger, x^\ddagger\right\}.$$

Clearly, S_e is a disconnected set. The proof is complete. ■

VI. SIMULATION STUDY

In this section, we will give some examples to further illustrate the theories established in above sections. Let us first consider the following 2-D network:

$$x(k+1) = \left[b + \begin{pmatrix} 0.8 & -0.5 \\ -0.5 & 0.8 \end{pmatrix} x(k)\right]^+ \tag{17}$$

for $k \geq 0$. Denote

$$W = \begin{pmatrix} 0.8 & -0.5 \\ -0.5 & 0.8 \end{pmatrix}.$$

Then

$$I - W = \begin{pmatrix} 0.2 & 0.5 \\ 0.5 & 0.2 \end{pmatrix} \quad I + W = \begin{pmatrix} 1.8 & -0.5 \\ -0.5 & 1.8 \end{pmatrix}.$$

It is easy to check that $I + W$ has eigenvalues 2.3 and 1.3. Thus, it is positive definite. Clearly, $I - W$ is copositive since each element is positive. However, $I - W$ is not positive semidefinite because it has -0.3 and 0.7 as its eigenvalues. By Theorem 7, this network is conditionally multiattractive, i.e., the set of its stable equilibrium points can be disconnected by some input b. Fig. 3 shows the conditionally multiattractivity of network (1) by choosing an input as $b = [0.7071, 0.7071]^T$. The network has three equilibrium points, two of them located on the boundary of $\mathbb{R}^2_+$ are stable, and the other one is unstable.

The theory of permitted and forbidden sets of network (1) can be used for group selection [3]. Given m groups of total n neurons, define a group membership of the ath $(1 \leq a \leq m)$ group by [3]

$$\xi^a_i = \begin{cases} 1, & \text{if the } i\text{th neuron is in the } a\text{th group} \\ 0, & \text{otherwise} \end{cases}$$

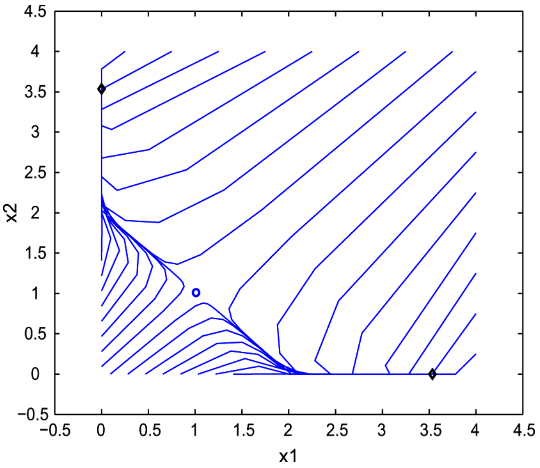

Fig. 3. Conditional multiattractivity of (17). The two diamonds are stable equilibrium points. The circle is the unstable equilibrium point. Clearly, the set S_e is disconnected.

for $i = 1, \ldots, n$. Define the inhibitory synaptic connectivity between neuron i and neuron j by

$$J_{ij} = \prod_{a=1}^{m} \left(1 - \xi_i^a \xi_j^a\right). \tag{18}$$

The connection matrix W of the network is defined by

$$W = \alpha I - \beta J \tag{19}$$

where $1 > \alpha > 0$ and $\beta > 0$ are some parameters, $J = (J_{ij})_{n\times n}$, and I represents the $n \times n$ identity matrix. The parameters α and β are selected so that Theorem 1 and Theorem 3 are satisfied, then network (1) can be used for group selection. To select a group, first choose an initial $x(0)$, then provide an input b to the network; the trajectory starting from $x(0)$ will converge to a stable equilibrium point x^*, and the permitted set thus indicates a selected group ξ^a. The relations between the selected group ξ^a, the stable equilibrium x^*, and the input b can be simply calculated out as

$$(\xi^a)^T \cdot x^* = \frac{1}{1-\alpha} \cdot (\xi^a)^T \cdot b. \tag{20}$$

Let us consider an example of the ring networks [1]–[3] for illustration. Suppose there are 15 neurons to be arranged in a ring structure and every set of five contiguous neurons is in a group. Then, we have total 15 groups. Taking $\alpha = 0.8$ and $\beta = 0.3$, it can be checked that each 5×5 submatrix of W has eigenvalue 0.8. Thus, by Theorem 3, each of the five contiguous neurons, i.e., each group forms a permitted set. The network can select a group by randomly providing an input b. Fig. 4 shows that the network selects a group by a randomly selected input as

$$\begin{aligned} b = \Big[&0.4662, 0.9138, 0.2286, 0.8620, 0.6566,\\ &0.8912, 0.4881, 0.9926, 0.3733, 0.5314,\\ &0.1813, 0.5019, 0.4222, 0.6604, 0.6737\Big]^T. \end{aligned}$$

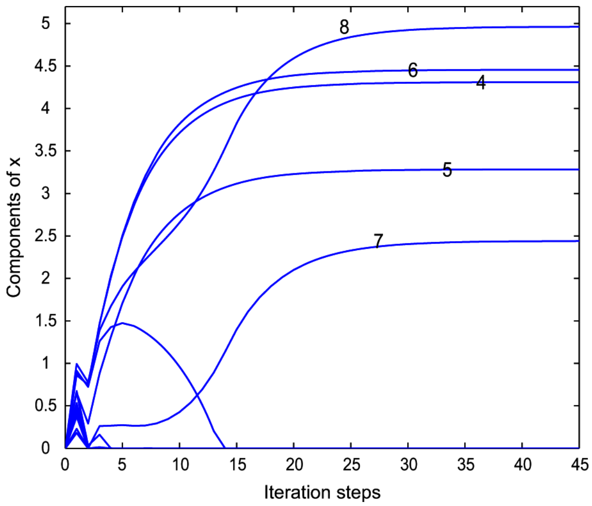

Fig. 4. Group selection by the ring network. The components of the network with positive value form a permitted set and the set indicates a group. The selected group neurons are those with the index set $P = \{4, 5, 6, 7, 8\}$.

The selected group neurons are those with index set $P = \{4, 5, 6, 7, 8\}$, and the output positive values are

$$\begin{cases} x_4 = 4.3102 \\ x_5 = 3.2831 \\ x_6 = 4.4559 \\ x_7 = 2.4405 \\ x_8 = 4.9630. \end{cases}$$

Next, we illustrate the analog–digital computation of network (1) and further interpret its storing capability of patterns. When the network converges to an equilibrium point, each neuron can be either active or inactive and thus the network forms binary computation; at the same time, each activated neuron carries an analog value which endows the network analog computation. The coexistence of analog and digital computation of the network may be quite useful for some pattern recognition problems. Let us consider an example for illustration. Suppose that we have four patterns of binary images in Fig. 5(a) with letters L, o, v, e, respectively. Each image is 56×62 pixels. The network is constructed by using 56×62 neurons. Then, four patterns of binary images can be described by group memberships ξ^L, ξ^o, ξ^v, and ξ^e. Using (18) to compute the matrices J, taking $\alpha = 0.99963$ and $\beta = 0.00055496$, and computing W by (19) (conditions of Theorem 1 and Theorem 3 are satisfied), then the four patterns are stored as permitted sets in the network. It should be noted that the four patterns are binary images and the network does not carry any gray information of the images, thus, the stored information contains actually the structure information of the patterns. However, the network can be used to extract gray images if the input to the network are gray images. In Fig. 5(b), each gray image is used as input of the network. By providing each one to the network, the network converges to an equilibrium point and the permitted set indicates an extracted pattern; the activated neurons contains analog values and these values represent the gray information. This fact can be further confirmed by (20). Fig. 5(c) shows the output gray images of the network.

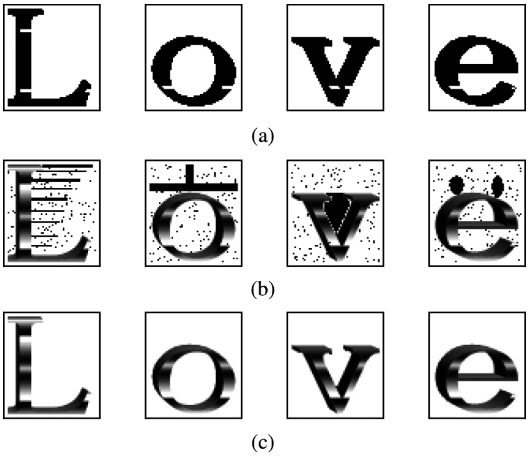

Fig. 5. The analog–digital computation of the network. (a) The patterns of binary images to be stored as permitted sets of the network. (b) The input gray images which contain structure information of the stored patterns plus some extra information and with 0.1 salt&pepper noise. (c) The outputs of the network. The gray information of the images is from the analog values of the activated neurons.

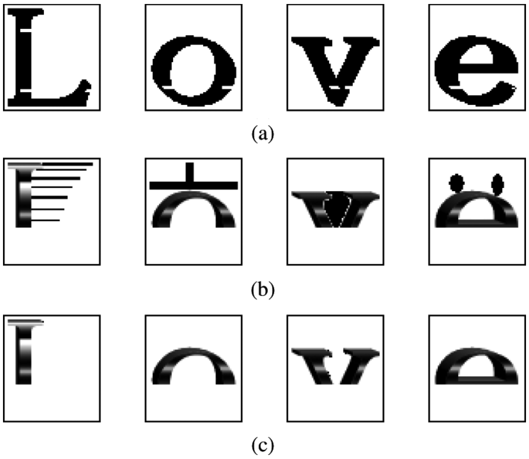

Fig. 6. All subsets of each pattern are also stored as permitted sets in the network. If the input contains part structure information of a pattern, the network can only extract part from the input. (a) The patterns of binary images to be stored as permitted sets of the network. (b) The input gray images which contain half structure information of the stored pattern. (c) The outputs of the network. The network can extract only half part from the inputs.

An important feature should be noted here: in order to extract a complete pattern from the input, the input image should contains all the structure information of the corresponding stored pattern. According to Theorem 4, any subset of a permitted set is also a permitted set, thus, if the network stores a pattern as permitted set then all of its subpatterns are also stored in the network. Any part information of the stored pattern can be extracted. If an input contains part information of the stored pattern, the network can only extract part of the associated pattern information. This fact can be easily seen from (20). Fig. 6 shows such computation. The computation characteristics of the network can be expected to find applications in data mining, data analysis, etc.

The storing and extracting methods of network (1) in this paper are different from that of associative memory implemented by some earlier models [20], in which the associative memory is implemented by encoding inputs as initial conditions to the networks and the memories are stored as stable attractors. By providing an input to the network, it converges to an attractor which represents the extracted memory in response to that input. However, the attractor loses all the information of the input. The theory of permitted and forbidden sets developed in this paper provides new perspective of the memory in neural networks. Network (1) stores memories as permitted sets, and each output of the network depends on the input b; such dependence is explicitly represented by (20).

VII. CONCLUSION

Some basic theories of permitted and forbidden sets of linear threshold discrete-time recurrent neural networks have been studied in this paper. Necessary and sufficient conditions are established for complete convergence, existence of permitted and forbidden sets, as well as conditionally multiattractivity. The concepts of permitted and forbidden sets enable a new perspective of memory in recurrent neural networks. Some applications have demonstrated the importance of these concepts. By establishing the foundations of permitted and forbidden sets, it is believed that more and interesting applications will be found.

REFERENCES

[1] R. Hahnloser, R. Sarpeshkar, M. Mahowald, R. Douglas, and H. S. Seung, "Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit," *Nature*, vol. 405, pp. 947–951, 2000.

[2] R. H. Hahnloser, H. S. Seung, and J. J. Slotine, "Permitted and forbidden sets in symmetric threshold-linear networks," *Neural Comput.*, vol. 15, pp. 621–638, 2003.

[3] X. Xie, R. H. Hahnloser, and H. S. Seung, "Selectively grouping neurons in recurrent networks of lateral inhibition," *Neural Comput.*, vol. 14, pp. 2627–2646, 2002.

[4] R. Dougleas, C. Koch, M. Mahowald, K. Matin, and H. Suarez, "Recurrent excitation in neocortical circuits," *Science*, vol. 269, pp. 981–985, 1995.

[5] R. Ben-Yishai, R. L. Bar-Or, and H. Sompolinsky, "Theory of orientation tuning in visual cortex," *Proc. Nat. Acad. Sci.*, vol. 92, pp. 3844–3848, 1995.

[6] B. Blumenfeld, D. Bibitchkov, and M. Tsodyks, "Neural network model of the primary visual cortex: From functional architecture to lateral connectivity and back," *J. Comput. Neurosci.*, vol. 20, pp. 219–241, 2006.

[7] H. J. Tang, K. C. Tan, and E. J. Teoh, "Dynamics analysis and analog associative memory of networks with LT neurons," *IEEE Trans. Neural Netw.*, vol. 17, no. 2, pp. 409–418, Mar. 2006.

[8] R. Hahnloser, "On the piecewise analysis of linear threshold neural networks," *Neural Netw.*, vol. 11, pp. 691–697, 1998.

[9] H. Wersing, J. J. Steil, and H. Ritter, "A competitive layer model for feature binding and sensory segmentation," *Neural Comput.*, vol. 13, pp. 357–387, 2001.

[10] H. Wersing, W. J. Beyn, and H. Ritter, "Dynamical stability conditions for recurrent neural networks with unsaturating piecewise linear transfer functions," *Neural Comput.*, vol. 13, pp. 1811–1825, 2001.

[11] Z. Yi, K. K. Tan, and T. H. Lee, "Multistability analysis for recurrent neural networks with unsaturating piecewise linear transfer functions," *Neural Comput.*, vol. 15, pp. 639–662, 2003.

[12] Z. Yi and K. K. Tan, "Multistability of discrete-time recurrent neural networks with unsaturating piecewise linear activation functions," *IEEE Trans. Neural Netw.*, vol. 15, no. 2, pp. 329–336, Mar. 2004.

[13] H. J. Tang, K. C. Tan, and Z. Yi, *Neural Networks: Computational Models and Applications*. Heidelberg, Germany: Springer-Verlag, 2007.

[14] Z. Yi and K. K. Tan, *Convergence Analysis of Recurrent Neural Networks*. Norwell, MA: Kluwer, 2004.

[15] Z. Yi and K. K. Tan, "Dynamical stability conditions for Lotka-Volterra recurrent neural networks with delays," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 66, Jul. 2002, 011910.

[16] Z. G. Zeng and J. Wang, "Complete stability of cellular neural networks with time-varying delays," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 4, pp. 944–955, Apr. 2006.

[17] Z. G. Zeng and J. Wang, "Multiperiodicity and exponential attractivity evoked by periodic external inputs in delayed cellular neural networks," *Neural Comput.*, vol. 18, pp. 848–870, 2006.

[18] Z. G. Zeng and J. Wang, "Multiperiodicity of discret-time delayed neural networks evoked by periodic external inputs," *IEEE Trans. Neural Netw.*, vol. 17, no. 5, pp. 1141–1151, Sep. 2006.

[19] L. Zhang, Z. Yi, and J. Yu, "Multiperiodicity and attractivity of delayed recurrent neural networks with unsaturating piecewise linear transfer functions," *IEEE Trans. Neural Netw.*, vol. 19, no. 1, pp. 158–167, Jan. 2008.

[20] J. J. Hopfield, "Neural networks and physical systems with emergent collective computatinal abilities," *Proc. Nat. Acad. Sci. USA*, vol. 79, pp. 2554–2558, 1982.

[21] S. Hui and S. H. Żak, "Dynamical analysis of the brain-state-in-a-box (BSB) neural models," *IEEE Trans. Neural Netw.*, vol. 3, no. 1, pp. 86–94, Jan. 1992.

[22] L. Jin, P. N. Nikiforuk, and M. M. Gupta, "Absolute stability conditions for discrete-time recurrent neural networks," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 954–964, Nov. 1994.

[23] A. N. Michel, J. Si, and G. Yen, "Analysis and synthesis of a class of discrete-time neural networks described on hypercubes," *IEEE Trans. Neural Netw.*, vol. 2, no. 1, pp. 32–46, Jan. 1991.

[24] M. J. Pérez-Ilzarbe, "Convergence analysis of a discrete-time recurrent neural networks to perform quadratic real optimization with bound constraints," *IEEE Trans. Neural Networks*, vol. 9, no. 6, pp. 1344–1351, 1998.

[25] R. Perfetti, "A synthesis procedure for brain-state-in-a-box neural networks," *IEEE Trans. Neural Netw.*, vol. 6, no. 5, pp. 1071–1080, Sep. 1995.

[26] J. Si and A. N. Michel, "Analysis and synthesis of a class of discrete-time neural networks with multilevel threshold neurons," *IEEE Trans. Neural Netw.*, vol. 6, no. 1, pp. 105–116, Jan. 1995.

[27] S. Hu and J. Wang, "Global stability of a class of continuous-time recurrent neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 9, pp. 1334–1347, Sep. 2002.

[28] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

Zhang Yi received the Ph.D. degree in mathematics from the Institute of Mathematics, The Chinese Academy of Science, Beijing, China, in 1994.

Currently, he is a Professor at the Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu, China. He is the coauthor of the books: *Convergence Analysis of Recurrent Neural Networks* (Norwell, MA: Kluwer, 2004) and *Neural Networks: Computational Models and Applications* (Heidelberg, Germany: Springer-Verlag, 2007). His current research interests include neural networks, data mining, and machine learning.

Lei Zhang received the B.S. and M.S. degrees in mathematics from University of Electronic Science and Technology of China, Chengdu, China, in 2002 and 2005, respectively, and the Ph.D. degree in computer science from University of Electronic Science and Technology of China, Chengdu, China, in 2008.

Currently, she is a Postdoctoral Research Fellow at the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong. Her research interests include theory and applications of neural networks.

Jiali Yu received the M.S. degree in mathematics from University of Electronic Science and Technology of China, Chengdu, China, in 2003. She is currently working towards the Ph.D. degree in computer science at the School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China.

Her current research interests include neural networks and continuous attractors.

Kok Kiong Tan received the Ph.D. degree in control engineering from the National University of Singapore, Singapore, in 1995.

He was a Research Fellow at SIMTech, where he was involved in managing industrial projects. Currently, he is an Associate Professor at the National University of Singapore. His current research interests are in precision motion control and instrumentation, advanced process control and autotuning, and general industrial automation.