Foundations of Implementing the Competitive Layer Model by Lotka–Volterra Recurrent Neural Networks

Zhang Yi, *Member, IEEE*

***Abstract*—The competitive layer model (CLM) can be described by an optimization problem. The problem can be further formulated by an energy function, called the CLM energy function, in the subspace of nonnegative orthant. The set of minimum points of the CLM energy function forms the set of solutions of the CLM problem. Solving the CLM problem means to find out such solutions. Recurrent neural networks (RNNs) can be used to implement the CLM to solve the CLM problem. The key point is to make the set of minimum points of the CLM energy function just correspond to the set of stable attractors of the recurrent neural networks. This paper proposes to use Lotka–Volterra RNNs (LV RNNs) to implement the CLM. The contribution of this paper is to establish foundations of implementing the CLM by LV RNNs. The contribution mainly contains three parts. The first part is on the CLM energy function. Necessary and sufficient conditions for minimum points of the CLM energy function are established by detailed study. The second part is on the convergence of the proposed model of the LV RNNs. It is proven that interesting trajectories are convergent. The third part is the most important. It proves that the set of stable attractors of the proposed LV RNN just equals the set of minimum points of the CLM energy function in the nonnegative orthant. Thus, the LV RNNs can be used to solve the problem of the CLM. It is believed that by establishing such basic rigorous theories, more and interesting applications of the CLM can be found.**

***Index Terms*—Competitive layer model (CLM), convergence, energy functions, Lotka–Volterra recurrent neural networks (LV RNNs), minimum points, stable attractors.**

I. Introduction

The competitive layer model (CLM) is formed by arranging neurons in several layers. The neurons are connected recurrently. It possesses a competitive mechanism between the layers. This model has a good biological background. It is believed that in real biological neural networks the competition exists in different layers of neurons. The CLM was first proposed by Ritter [1] as a model for spatial feature binding. It is known from neurophysiology and neuroanatomy that information processing streams in the visual system are divergent and located in separate brain areas. Feature dimensions such as color, motion, location, and object identity are processed along different pathways. Understanding the neural mechanisms of feature binding is very important both in neuroscience and practical applications. Since feature binding may provide one of the basic sensory information processing principles [2], there is great interest in using similar mechanisms for pattern recognition applications like image segmentation and object recognition [3].

Mathematically, the CLM can be described by an optimization problem. This optimization problem can be further formulated by an energy function, called the CLM energy function, in some subspace of nonnegative orthant. The CLM energy function provides an energy-based approach for feature binding. By merging contextual constraints into an energy function, an energy-based approach can directly control the desired groupings. Spin models [4]–[6] and relaxation labeling (RL) [7], [8] are well-known energy-based approaches of feature binding. RL is a standard technique in pattern recognition and computer vision fields [9] which has found applications in a variety of different problems [10]. However, these models share certain drawbacks regarding their biological plausibility, since they either require iterative discrete cluster update procedures or complex normalizing nonlinearities [3]. In [11], the relation of the CLM to RL with regard to feature binding and labeling problems are discussed. It shows that feature binding with the CLM can be considered in the RL framework.

The set of minimum points of the CLM energy function forms the set of solutions of the CLM problem. Solving the CLM problem requires to find out the minimum points of the CLM energy function. Since the structure of the CLM is a recurrently connected network model, some recurrent neural networks (RNNs) can be used to solve the CLM problem, i.e., implementing the CLM by RNNs. By making the attractors of the RNNs correspond to the minimum points of the CLM energy function, the CLM problem can be then solved by running of the RNNs. In [3], linear threshold (LT) RNNs are used to implement the CLM for feature binding and sensory segmentation. LT RNNs possess many good dynamical properties and have been extensively studied by many authors; see, for example, [12]–[15]. The LT RNNs have found many applications, such as associative memory [16], winner-take-all [17], group selection [18], etc. For more dynamical properties of LT RNNs, we refer to [19]–[23]. In [24], a learning mechanism is presented for the CLM of LT RNNs.

Nevertheless, the rigorous theories of using RNNs to implement the CLM have not been established so far. The key point of using RNNs to implement the CLM is to prove that the set of minimum points of the CLM energy function just equals the set of attractors of the RNNs. This is not easy; it requires to completely understand the CLM energy function as well as the corresponding dynamics of the RNNs.

Manuscript received August 22, 2008; revised April 24, 2009 and September 19, 2009; accepted December 13, 2009. First published February 05, 2010; current version published February 26, 2010. This work was supported by the National Science Foundation of China under Grant 60970013.

The author is with the Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu 610065, China (e-mail: zhangyi@scu.edu.cn; zyi.cilab@gmail.com).

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.2009.2039758

This paper proposes to use Lotka–Volterra (LV) RNNs to implement the CLM. The model of LV RNNs was first proposed in [25]. It was derived from the conventional membrane dynamics of neurons with a sigmoid response function. LV RNNs possess good dynamical properties and have found many potential applications in winner-take-all, winner-share-all, and k-winner-take-all problems [26], [27]. In [28], some properties of permitted and forbidden sets of LV RNNs are reported. More general dynamic properties can be found in [29] and [30]. Some hardware implementations of LV RNNs are now available [27]. The main contribution of this paper is to establish the important basic theory: the set of stable attractors of LV RNNs just equals the set of minimum points of the CLM energy function in the nonnegative orthant, by rigorous mathematical analysis. To establish such basic theory, it requires detailed study of the CLM energy function as well as the corresponding LV RNNs. Three main steps will be carried out for this task. The first step is to establish necessary and sufficient conditions for minimum points of the CLM energy function. The second step is to establish a convergence property of the trajectories of LV RNNs. The trajectories will be divided into two classes: interesting trajectories and uninteresting trajectories. Uninteresting trajectories are those which converge to the set of unstable equilibrium points. Such trajectories can be ignored in practice since the convergence can be easily destroyed by some small disturbance. It will be shown that each interesting trajectory is convergent under some simple conditions. In the third step, it will be proved that the set of minimum points of the CLM energy function in the nonnegative orthant just equals the set of stable attractors of LV RNNs. Thus, LV RNNs can be used to solve the problem of the CLM.

This paper is organized as follows. In Section II, a brief description of the CLM is presented. In Section III, detailed study for the CLM energy function is given. The implementation of using LV RNNs for the CLM is studied in Section IV. Four subsections are contained in this section. Some discussions are presented in Section V to further illustrate the theories. Finally, conclusions are given in Section VI.

II. Competitive Layer Model

The structure of the CLM is a model of RNNs. Fig. 1 shows the CLM network architecture. It contains a set of L layers and with each layer there are N neurons. Thus, the CLM contains a total of $N \times L$ neurons. Neurons in each layer are connected with each other and the connection weights are assumed to be independent of any layers. Between different layers, only those neurons that are arranged in a row are connected. Neurons in each layer are required to be cooperative, while neurons in each row are required to be competitive. The competition between the neurons in each row can be implemented by using lateral inhibitions. Each neuron has an external input which is also independent of the layer.

Throughout this paper, symbols $\alpha, \beta, \gamma, \eta, 1 \leq \alpha, \beta, \gamma, \eta \leq L$, called layer indexes, will be used to index the layers of the CLM, and the symbols i, j, k, $r, 1 \leq i, j, k, r \leq N$, called row indexes, will be used to index the rows of the CLM. Thus, the index $i\alpha$ indicates that a neuron is located at ith row of αth layer. Let $x_{i\alpha}$ be the activity of the neuron located at $i\alpha$, w_{ij} be the connection weight of neuron $i\alpha$ and neuron $j\alpha$, and note that each w_{ij} is independent of the layer index α and $w_{ii} > 0$. The external input for neuron $i\alpha$ is denoted by h_i which is also independent of the layer index. Throughout this paper, it is assumed that $h_i > 0, i = 1, \ldots, N$. The CLM can be described by the following optimization problem:

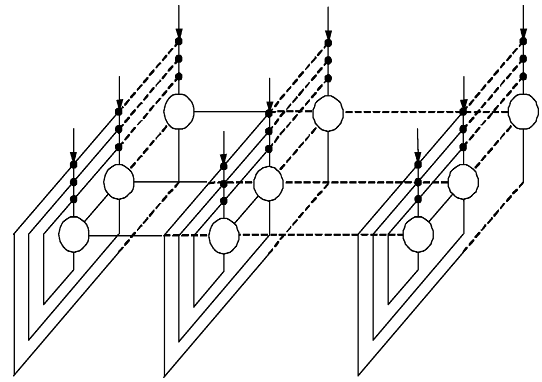


Fig. 1. Structure of the CLM.

$$\begin{cases} \min & -\frac{1}{2} \sum_{\alpha=1}^{L} \sum_{i,j=1}^{N} w_{ij} x_{i\alpha} x_{j\alpha} \\ \text{s.t.} & x_{i\alpha} \geq 0, \qquad 1 \leq i \leq N, \qquad 1 \leq \alpha \leq L \\ & \sum_{\alpha=1}^{L} x_{i\alpha} = h_i, \qquad 1 \leq i \leq N. \end{cases} \tag{1}$$

Solving the CLM problem means to find out the optimization points in the space $\mathbb{R}^{NL}$.

It is quite difficult to directly solve the optimization problem of (1), since it contains two classes of constraints. Denote the nonnegative orthant of $\mathbb{R}^{NL}$ by

$$\mathbb{R}_{+}^{NL} = \left\{ x \,\middle|\, x \in \mathbb{R}^{NL}, x_i \geq 0 (1 \leq i \leq NL) \right\}.$$

Then, the CLM can be rewritten as another equivalent optimization problem

$$\begin{cases} \min & \frac{C}{2} \sum_{i=1}^{N} \left(\sum_{\beta=1}^{L} x_{i\beta} - h_i \right)^2 - \frac{1}{2} \sum_{\alpha=1}^{L} \sum_{i,j=1}^{N} w_{ij} x_{i\alpha} x_{j\alpha} \\ \text{s.t.} & x \in \mathbb{R}_{+}^{NL} \end{cases} \tag{2}$$

where $C > 0$ is some sufficiently large constant. Thus, solving the CLM problem (1) is then equivalent to find out the optimization points of (2) in the space $\mathbb{R}_{+}^{NL}$.

Generally, the optimization problem contains many optimization points. Finding out such points is not an easy task. In this paper, we will use a class of LV RNNs to calculate the optimization points of (2).

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

$$\mathbb{R}_{+}^{N} = \left\{ y \,\middle|\, y \in \mathbb{R}^{N}, y_i \geq 0 (1 \leq i \leq N) \right\}.$$

Then, any vector $x \in \mathbb{R}_{+}^{NL}$ can be written as

$$x = \left(x_1^T, \ldots, x_L^T \right)^T$$

where

$$x_\alpha = (x_{1\alpha}, \ldots, x_{N\alpha})^T \in \mathbb{R}_{+}^{N}, \qquad 1 \leq \alpha \leq L.$$

Each x_α is called a layer of x.

In this paper, denote by $W = (w_{ij})_{N\times N}$, I the identity matrix, and

$$\overline{h} = \max_{1\le i\le N}\{h_i\} \quad \underline{h} = \min_{1\le i\le N}\{h_i\}.$$

III. The CLM Energy Function

The CLM energy function is defined by

$$E(x) = \frac{C}{2}\sum_{i=1}^{N}\left(\sum_{\beta=1}^{L} x_{i\beta} - h_i\right)^2 - \frac{1}{2}\sum_{\alpha=1}^{L}\sum_{i,j=1}^{N} w_{ij}x_{i\alpha}x_{j\alpha} \tag{3}$$

for $x \in \mathbb{R}_+^{NL}$; see [1], [3], [11], [20], and [24]. Solving of the CLM problem requires to find out the minimum points of the CLM energy function $E(x)$ in the nonnegative orthant $\mathbb{R}_+^{NL}$. Thus, it is important and interesting to completely understand the conditions for a point in $\mathbb{R}_+^{NL}$ to be a minimum point of $E(x)$. It also plays crucial role to design a model of RNNs for implementing the CLM.

Denote the set of minimum points of the energy function $E(x)$ in the nonnegative orthant $\mathbb{R}_+^{NL}$ by

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

This section will give detailed discussion on the set $\mathcal{M}$. We will address the problems: under what conditions a point in $\mathbb{R}_+^{NL}$ can be a minimum point? And, if a point is a minimum point, then what conditions must be satisfied? To completely address these problems, necessary and sufficient conditions for a point in $\mathbb{R}_+^{NL}$ to be a minimum point will be established.

Let us first consider a simple case to get some intuitive illustration for the CLM energy function. Consider the CLM network with two layers and each layer contains only one neuron. Then, the energy function becomes

$$E(x) = \frac{C}{2}(x_{11}+x_{12}-h)^2 - \frac{w}{2}\left(x_{11}^2 + x_{12}^2\right)$$

for $x = (x_{11}, x_{12})^T \in \mathbb{R}_+^2$. If $C > w > 0$, it is a saddle surface. Fig. 2 shows the saddle surface of the energy function with $C = 2, w = 1$, and $h = 1$. There are four possible minimum points of the energy function: $(0,0)^T$, $(0,2)^T$, $(2,0)^T$, and $(2/3, 2/3)^T$. However, $(2/3,2/3)^T$ is a saddle point, $(0,0)^T$ is a maximum point, and only $(0,2)^T$ and $(2,0)^T$ are minimum points.

The CLM energy function $E(x)$ is indeed a saddle hypersurface. In fact, the Hessian matrix of the $E(x)$ can be easily calculated out as

$$H_m = \begin{bmatrix} (CI-W) & CI & \ldots & CI \\ CI & (CI-W) & \ldots & CI \\ \vdots & \vdots & \ldots & \vdots \\ CI & CI & \ldots & (CI-W) \end{bmatrix}_{NL\times NL}.$$

Since $w_{ii} > 0, i = 1, \ldots, N$, clearly, the matrix W must have at least one positive eigenvalue. Let λ be a positive eigenvalue of W, then, it can be easily checked that $NC - \lambda$ and $-\lambda$ are eigenvalues of the matrix H_m. The Hessian matrix H_m has both positive and negative eigenvalues if C is sufficiently large, say, $C > \lambda/N$. Thus, the CLM energy function has a saddle point

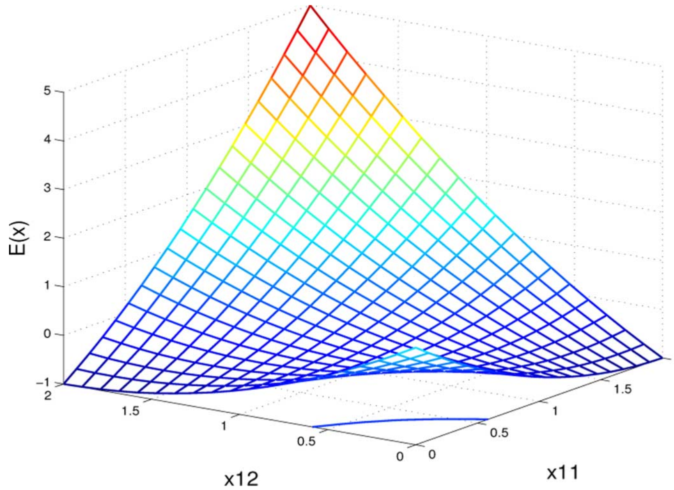


Fig. 2. Energy function of the CLM with two layers and each layer contains one neuron. It is a saddle surface.

if $C > \lambda/N$. In the $\mathbb{R}_+^{NL}$ space, the CLM energy function may have many possible minimum points, however, some of them are not really minimum points. Because each minimum point of the CLM energy function can be looked as a solution of the CLM problem, it is desired to establish necessary and sufficient conditions for the minimum points in $\mathbb{R}_+^{NL}$ space.

The concept of minimum points is a local property of the energy function. If x^* is minimum point, it requires that $E(x) \ge E(x^*)$ holds for all x in some small neighborhood of x^* in $\mathbb{R}_+^{NL}$. To explore the properties of minimum points of the CLM energy function, let us give another representation of the energy function. Since $E(x)$ is a quadratic function, given any $x^* \in \mathbb{R}_+^{NL}$, it holds that

$$\begin{aligned} E(x) - E(x^*) &= \left(\nabla E(x)|_{x=x^*}\right)^T (x - x^*) \\ &\quad + \frac{1}{2}(x-x^*)^T \left(\nabla^2 E(x)\big|_{x=x^*}\right)(x-x^*) \\ &= \sum_{i=1}^{N}\sum_{\gamma=1}^{L}\left[C\left(\sum_{\eta=1}^{L} x_{i\eta}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\gamma}^*\right](x_{i\gamma} - x_{i\gamma}^*) \\ &\quad + \frac{C}{2}\sum_{\gamma=1}^{L}\sum_{\eta=1}^{L}\left(x_\gamma - x_\gamma^*\right)^T\left(x_\eta - x_\eta^*\right) \\ &\quad - \frac{1}{2}\sum_{\gamma=1}^{L}\left(x_\gamma - x_\gamma^*\right)^T W \left(x_\gamma - x_\gamma^*\right) \end{aligned} \tag{4}$$

for $x \in \mathbb{R}_+^{NL}$. This expression could be very useful in the next for analyzing the properties of the CLM energy function $E(x)$.

To derive necessary and sufficient conditions of minimum points of the CLM energy function $E(x)$, several steps will be carried out in advance. Let us first give an outline of the main steps. In Lemma 1, some necessary conditions for minimum points will be given. These conditions will be used in Lemma 2 to prove that all the minimum points must be contained in some bounded range of $\mathbb{R}_+^{NL}$. Using Lemmas 1 and 2, it will be proved in Lemma 3 that each row of a minimum point must contain nonzero elements. Moreover, by Lemma 4, it will be shown in Lemma 5 that in each row of a minimum point, it cannot have more than one nonzero element. Thus, there exists one and only one nonzero element in each row of any minimum point. Finally, necessary and sufficient conditions for minimum points will be established in Theorem 1.

Lemma 1: If $x^* \in \mathcal{M}$, then it holds that

$$\begin{cases} C\left(\sum_{\beta=1}^{L} x_{i\beta}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\gamma}^* = 0, & \text{if } x_{i\gamma}^* > 0 \\ C\left(\sum_{\beta=1}^{L} x_{i\beta}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\gamma}^* \geq 0, & \text{if } x_{i\gamma}^* = 0 \end{cases}$$

for $1 \leq i \leq N$ and $1 \leq \gamma \leq L$.

Proof: Since $x^* \in \mathcal{M}$, by (4), it must hold that

$$\begin{cases} \left.\dfrac{\partial E}{\partial x_{i\gamma}}\right|_{x=x^*} = 0, & \text{if } x_{i\gamma}^* > 0 \\ \left.\dfrac{\partial E}{\partial x_{i\gamma}}\right|_{x=x^*} \geq 0, & \text{if } x_{i\gamma}^* = 0 \end{cases}$$

for $1 \leq i \leq N$ and $1 \leq \gamma \leq L$. The result now follows and the proof is complete. ■

Next, we show that all minimum points are located in some range if C is sufficiently large.

Lemma 2: Suppose that

$$C > \sum_{i,j=1}^{N} |w_{ij}|.$$

If $x^* \in \mathcal{M}$, then

$$x_{j\gamma}^* \leq \frac{C\overline{h}}{C - \sum_{i,k=1}^{N} |w_{ik}|}$$

for $1 \leq j \leq N$ and $1 \leq \gamma \leq L$.

Proof: Let

$$x_{i\alpha}^* = \max\left\{x_{j\gamma}^* \middle| 1 \leq j \leq N; 1 \leq \gamma \leq L\right\}.$$

If $x_{i\alpha}^* = 0$, clearly Lemma 2 holds. If $x_{i\alpha}^* > 0$, using Lemma 1, it follows that

$$C\left(\sum_{\beta=1}^{L} x_{i\beta}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\alpha}^* = 0.$$

Since

$$\begin{aligned} C&\left(\sum_{\beta=1}^{L} x_{i\beta}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\alpha}^* \\ &\geq C\left(x_{i\alpha}^* - \overline{h}\right) - x_{i\alpha}^* \sum_{k,j=1}^{N} |w_{kj}| \\ &= x_{i\alpha}^* \left(C - \sum_{k,j=1}^{N} |w_{kj}|\right) - C\overline{h} \end{aligned}$$

then

$$x_{i\alpha}^* \leq \frac{C\overline{h}}{C - \sum_{k,j=1}^{N} |w_{kj}|}.$$

The result now follows and the proof is complete. ■

Lemma 3: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}} + 1\right) \sum_{i,j=1}^{N} |w_{ij}|.$$

If $x^* \in \mathcal{M}$, then, given any row index i, there must exist at least one layer index α such that $x_{i\alpha}^* > 0$.

Proof: Otherwise, suppose there exists some row index i such that $x_{i\gamma}^* = 0$ for all $\gamma = 1, \ldots, L$. By Lemma 1, it follows that

$$C\left(\sum_{\beta=1}^{L} x_{i\beta}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\gamma}^* \geq 0.$$

However, using Lemma 2, it gives that

$$\begin{aligned} C&\left(\sum_{\beta=1}^{L} x_{i\beta}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\gamma}^* \\ &= -Ch_i - \sum_{j=1}^{N} w_{ij}x_{j\gamma}^* \\ &\leq -C\underline{h} + \sum_{k,j=1}^{N} |w_{kj}|\, x_{j\gamma}^* \\ &\leq -C\underline{h} + \sum_{k,j=1}^{N} |w_{kj}| \cdot \frac{C\overline{h}}{C - \sum_{k,j=1}^{N} |w_{kj}|} \\ &< 0. \end{aligned}$$

This is a contradiction. It implies that given any row index i, there must exist a layer index α such that $x_{i\alpha}^* > 0$. The proof is complete. ■

Lemma 4: Suppose that $x^* \in \mathbb{R}_+^{NL}$. If there exists a row index i with two different layer indexes α and β such that

$$\begin{cases} C\left(\sum_{\gamma=1}^{L} x_{i\gamma}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\alpha}^* = 0 \\ C\left(\sum_{\gamma=1}^{L} x_{i\gamma}^* - h_i\right) - \sum_{j=1}^{N} w_{ij}x_{j\beta}^* = 0 \end{cases}$$

and $x_{i\alpha}^* > 0$, then $x^* \notin \mathcal{M}$.

Proof: Define a vector $v^\dagger \in \mathbb{R}_+^{NL}$ by

$$v_{j\gamma}^\dagger = \begin{cases} -1, & \text{if } j\gamma = i\alpha \\ 1, & \text{if } j\gamma = i\beta \\ 0, & \text{otherwise} \end{cases}$$

for $1 \le j \le N$ and $1 \le \gamma \le L$. Then, for each sufficiently small constant $\epsilon \in (0, x^*_{i\alpha})$, it holds that $x^* + v^{\dagger}\epsilon \in \mathbb{R}^{NL}_{+}$, and it gives from (4) that

$$\begin{aligned} &E\left(x^* + v^{\dagger}\epsilon\right) - E\left(x^*\right) \\ &= C\epsilon^2 v^{\dagger}_{i\alpha} v^{\dagger}_{i\beta} + \frac{(C - w_{ii})\,\epsilon^2}{2}\left({v^{\dagger}}^2_{i\alpha} + {v^{\dagger}}^2_{i\beta}\right) \\ &\quad + \left[C\left(\sum_{\gamma=1}^{L} x^*_{i\gamma} - h_i\right) - \sum_{j=1}^{N} w_{ij} x^*_{j\alpha}\right] v^{\dagger}_{i\alpha}\epsilon \\ &\quad + \left[C\left(\sum_{\gamma=1}^{L} x^*_{i\gamma} - h_i\right) - \sum_{j=1}^{N} w_{ij} x^*_{j\beta}\right] v^{\dagger}_{i\beta}\epsilon \\ &= -w_{ii}\epsilon^2 \\ &< 0. \end{aligned}$$

This implies that x^* cannot be a minimum point of $E(x)$, i.e., $x^* \notin \mathcal{M}$. The proof is complete. ■

Lemma 5: Suppose that $x^* \in \mathbb{R}^{NL}_{+}$. If there exists a row index i with two different layer indexes α and β such that

$$x^*_{i\alpha} > 0 \quad x^*_{i\beta} > 0$$

then $x^* \notin \mathcal{M}$.

Proof: Otherwise, suppose that $x^* \in \mathcal{M}$, then, by Lemma 1, it follows that

$$\begin{cases} C\left(\sum\limits_{\gamma=1}^{L} x^*_{i\gamma} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\alpha} = 0 \\ C\left(\sum\limits_{\gamma=1}^{L} x^*_{i\gamma} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\beta} = 0. \end{cases}$$

Using Lemma 4, it gives that $x^* \notin \mathcal{M}$. This is a contradiction and the proof is complete. ■

Theorem 1: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}} + 1\right) \sum_{i,j=1}^{N} |w_{ij}|. \tag{5}$$

Let $x^* \in \mathbb{R}^{NL}_{+}$. Then, $x^* \in \mathcal{M}$, if and only if given any row index $i (1 \le i \le N)$, there exists one and only one layer index $\alpha (1 \le \alpha \le L)$ such that

$$x^*_{i\gamma} \begin{cases} > 0, & \text{if } \gamma = \alpha \\ = 0, & \text{if } \gamma \ne \alpha \end{cases} \tag{6}$$

and

$$\begin{cases} C\left(x^*_{i\alpha} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\alpha} = 0 \\ C\left(x^*_{i\alpha} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\gamma} > 0, \quad \text{for } \gamma \ne \alpha. \end{cases} \tag{7}$$

Proof: (Necessity) Suppose that $x^* \in \mathcal{M}$, then we will prove that (6) and (7) are true.

Given any row index i, by (5), using Lemmas 3 and 5, there exists one and only one layer index $\alpha, 1 \le \alpha \le L$, such that (6) holds. Next, we prove that (7) is true. By Lemma 1, it holds that

$$\begin{cases} C\left(\sum\limits_{\eta=1}^{L} x^*_{i\eta} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\gamma} = 0, & \text{if } \gamma = \alpha \\ C\left(\sum\limits_{\eta=1}^{L} x^*_{i\eta} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\gamma} \ge 0, & \text{if } \gamma \ne \alpha \end{cases} \tag{8}$$

for $1 \le \gamma \le L$. We claim for $\gamma \ne \alpha, 1 \le \gamma \le L$, that

$$C\left(\sum_{\eta=1}^{L} x^*_{i\eta} - h_i\right) - \sum_{j=1}^{N} w_{ij} x^*_{j\gamma} > 0. \tag{9}$$

If this is not true, then there must exist a layer index $\beta \ne \alpha$ such that

$$\begin{cases} C\left(\sum\limits_{\eta=1}^{L} x^*_{i\eta} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\alpha} = 0 \\ C\left(\sum\limits_{\eta=1}^{L} x^*_{i\eta} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\beta} = 0. \end{cases}$$

By Lemma 4, it gives that $x^* \notin \mathcal{M}$, which is a contradiction. This implies that (9) is true. Thus, from (9) and (8), it holds that

$$\begin{cases} C\left(\sum\limits_{\eta=1}^{L} x^*_{i\eta} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\gamma} = 0, & \text{if } \gamma = \alpha \\ C\left(\sum\limits_{\eta=1}^{L} x^*_{i\eta} - h_i\right) - \sum\limits_{j=1}^{N} w_{ij} x^*_{j\gamma} > 0, & \text{if } \gamma \ne \alpha \end{cases} \tag{10}$$

for $1 \le \gamma \le L$. Now, (7) follows by substituting (6) into (10). The proof for necessary part is complete.

(Sufficiency) Given any row index i, suppose that there exists one and only one layer index α such that (6) and (7) hold, then we will prove that $x^* \in \mathcal{M}$.

Denote

$$\pi_{i\gamma} = C\left(x^*_{i\alpha} - h_i\right) - \sum_{j=1}^{N} w_{ij} x^*_{j\gamma}$$

for $1 \le i \le N, \gamma \ne \alpha$, and $1 \le \gamma \le L$. Clearly, each $\pi_{i\gamma} > 0$. Define a neighborhood of x^* by

$$D = \left\{ x \in \mathbb{R}^{NL}_{+} \middle| \, |x_{i\alpha} - x^*_{i\alpha}| \le \min_{1 \le \gamma \le L, \gamma \ne \alpha} \left\{\frac{\pi_{i\gamma}}{C}\right\} (1 \le i \le N) \right\}.$$

Then

$$\begin{aligned} &E(x) - E(x^*) \\ &= \sum_{i=1}^{N} \sum_{\gamma=1, \gamma\ne\alpha}^{L} \pi_{i\gamma} x_{i\gamma} + \frac{C}{2} \sum_{i=1}^{N}\sum_{\gamma=1}^{L}\sum_{\eta=1}^{L} \left(x_{i\eta} - x^*_{i\eta}\right) \\ &\quad \times \left(x_{i\gamma} - x^*_{i\gamma}\right) - \frac{1}{2}\sum_{\gamma=1}^{L} \left(x_{\gamma} - x^*_{\gamma}\right)^T W \left(x_{\gamma} - x^*_{\gamma}\right) \\ &= \sum_{i=1}^{N} \sum_{\gamma=1, \gamma\ne\alpha}^{L} \left[\pi_{i\gamma} + C\left(x_{i\alpha} - x^*_{i\alpha}\right)\right] x_{i\gamma} \end{aligned}$$

$$
\begin{aligned}
&+\frac{C}{2}\sum_{i=1}^{N}\left[\left(\sum_{\gamma=1,\gamma\neq\alpha}^{L}x_{i\gamma}\right)^{2}+\left(x_{i\alpha}-x_{i\alpha}^{*}\right)^{2}\right]\\
&-\frac{1}{2}\sum_{\gamma=1}^{L}\left(x_{\gamma}-x_{\gamma}^{*}\right)^{T}W\left(x_{\gamma}-x_{\gamma}^{*}\right)\\
\geq&\sum_{i=1}^{N}\sum_{\gamma=1,\gamma\neq\alpha}^{L}\left[\pi_{i\gamma}+C\left(x_{i\alpha}-x_{i\alpha}^{*}\right)\right]x_{i\gamma}\\
&+\frac{C}{2}\sum_{i=1}^{N}\left[\sum_{\gamma=1,\gamma\neq\alpha}^{L}x_{i\gamma}^{2}+\left(x_{i\alpha}-x_{i\alpha}^{*}\right)^{2}\right]\\
&-\frac{1}{2}\sum_{\gamma=1}^{L}\left(x_{\gamma}-x_{\gamma}^{*}\right)^{T}W\left(x_{\gamma}-x_{\gamma}^{*}\right)\\
=&\sum_{i=1}^{N}\sum_{\gamma=1,\gamma\neq\alpha}^{L}\left[\pi_{i\gamma}+C\left(x_{i\alpha}-x_{i\alpha}^{*}\right)\right]x_{i\gamma}\\
&+\frac{1}{2}\sum_{\gamma=1}^{L}\left(x_{\gamma}-x_{\gamma}^{*}\right)^{T}\left(CI-W\right)\left(x_{\gamma}-x_{\gamma}^{*}\right)\\
\geq&\frac{1}{2}\sum_{\gamma=1}^{L}\left(x_{\gamma}-x_{\gamma}^{*}\right)^{T}\left(CI-W\right)\left(x_{\gamma}-x_{\gamma}^{*}\right)
\end{aligned}
$$

for all $x \in D$. By (5), the matrix $(CI - W)$ must be positive definite. Thus, it holds that

$$E(x) - E(x^*) \geq 0$$

for all $x \in D$. This shows that $x^* \in \mathcal{M}$ and the proof is complete. ■

Theorem 1 establishes the necessary and sufficient conditions for minimum points of the CLM energy function $E(x)$. The conditions (5)–(7) are crucial for using LV RNNs to implement the CLM in Section IV.

IV. Implementing the CLM by LV Networks

Solving the problem of the CLM requires to find out the minimum points of the CLM energy function $E(x)$. Implementing the CLM by RNNs means to design some RNNs so that trajectories of a designed RNN can converge to minimum points of the CLM energy function. Generally, a point in $\mathbb{R}_+^{NL}$ is called a stable attractor if it attracts all trajectories starting from the points around it. If an RNN is designed so that the set of stable attractors just equals the set of the minimum points of the CLM energy function, the RNN could be looked as a perfect one. In last section, we have derived necessary and sufficient conditions for minimum points of the CLM energy function $E(x)$. If the constant C is sufficiently large, i.e., (5) holds, then a point $x^* \in \mathbb{R}_+^{NL}$ is a minimum point of $E(x)$ if and only if it satisfies both (6) and (7). Thus, to implement the CLM, it is sufficient to design an RNN so that each stable attractor of the network satisfies these conditions. The model of LV RNNs possesses many good and interesting properties. We will show in this section that the CLM can be implemented by using LV RNNs.

A. The Model of LV Networks

The proposed model of LV RNNs that will be used to implement the CLM can be described by

$$\dot{x}_{i\alpha}(t) = x_{i\alpha}(t)\left[C\left(h_i - \sum_{\beta=1}^{L} x_{i\beta}(t)\right) + \sum_{j=1}^{N} w_{ij}x_{j\alpha}(t)\right] \tag{11}$$

for $t \geq 0$ and $1 \leq i \leq N, 1 \leq \alpha \leq L$. Where $x(t) \in \mathbb{R}^{NL}$ denotes the state of the network at time t, it can be written as

$$x(t) = \left(x_1^T(t), \ldots, x_L^T(t)\right)^T$$

with

$$x_\alpha(t) = \left(x_{1\alpha}(t), \ldots, x_{N\alpha}(t)\right)^T \in \mathbb{R}^N, \qquad 1 \leq \alpha \leq L.$$

Each vector $x_\alpha(t) \in \mathbb{R}^N$ is called a layer of the state $x(t)$.

Trajectories of the network (11) are expected to converge to minimum points of the CLM energy function $E(x)$ in $\mathbb{R}_+^{NL}$ space. It is a basic requirement that any trajectory of (11) starting from a point in $\mathbb{R}_+^{NL}$ remains in $\mathbb{R}_+^{NL}$ for ever. This basic requirement can be guaranteed by Lemma 6.

Lemma 6: Given any $x(0) \in \mathbb{R}_+^{NL}$, it holds that

$$x_{i\alpha}(t)\begin{cases} > 0, & \text{if } x_{i\alpha}(0) > 0 \\ = 0, & \text{if } x_{i\alpha}(0) = 0 \end{cases}$$

for all $t \geq 0$ and $1 \leq i \leq N; 1 \leq \alpha \leq L$.

Proof: Denote by

$$v_{i\alpha}(t) = C\left(h_i - \sum_{\beta=1}^{L} x_{i\beta}(t)\right) + \sum_{j=1}^{N} w_{ij}x_{j\alpha}(t).$$

Then, from (11), it follows that

$$x_{i\alpha}(t) = x_{i\alpha}(0)\exp\left(\int_0^t v_{i\alpha}(s)ds\right)\begin{cases} > 0, & \text{if } x_{i\alpha}(0) > 0 \\ = 0, & \text{if } x_{i\alpha}(0) = 0 \end{cases}$$

for all $t \geq 0$ and $1 \leq i \leq N; 1 \leq \alpha \leq L$. This completes the proof. ■

Because it is expected to use the network (11) to look for minimum points of the CLM energy function $E(x)$ in $\mathbb{R}_+^{NL}$, and each trajectory of the network (11) starting from a point in $\mathbb{R}_+^{NL}$ remains in the $\mathbb{R}_+^{NL}$ for all the time, in the next part of this paper, it is restricted only to consider the trajectories staring from points in $\mathbb{R}_+^{NL}$.

B. Equilibrium Points Analysis

This section studies the properties and distributions of the equilibrium points of the network (11) in $\mathbb{R}_+^{NL}$ space. Equilibrium points of the network (11) are special vectors in $\mathbb{R}_+^{NL}$ space, and any trajectory starting from an equilibrium point stays at the point forever.

Definition 1: Suppose that $x^* \in \mathbb{R}_+^{NL}$. If it holds that

$$x_{i\alpha}^*\left[C\left(h_i - \sum_{\beta=1}^{L} x_{i\beta}^*\right) + \sum_{j=1}^{N} w_{ij}x_{j\alpha}^*\right] = 0 \tag{12}$$

for $1 \le i \le N$ and $1 \le \alpha \le L$, then x^* is called an equilibrium point of network (11).

The equilibrium points are distributed in the space $\mathbb{R}_+^{NL}$. In this paper, denoted by $\mathcal{E}$ is the set of all the equilibrium points of network (11) in $\mathbb{R}_+^{NL}$.

The linearization of network (11) in a neighborhood of an equilibrium point x^* plays important role in stability analysis. Next, the expression of the linearization is given.

Lemma 7: Let x^* be an equilibrium, then the linearization of network (11) at x^* is given by

$$\begin{aligned}\frac{d\left[x_{i\alpha}(t)-x_{i\alpha}^*\right]}{dt} &= \left[x_{i\alpha}(t)-x_{i\alpha}^*\right]\left[C\left(h_i-\sum_{\beta=1}^{L}x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}x_{j\alpha}^*\right]\\ &\quad + x_{i\alpha}^*\left[-C\sum_{\beta=1}^{L}\left(x_{i\beta}(t)-x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}\left(x_{j\alpha}(t)-x_{j\alpha}^*\right)\right]\end{aligned} \tag{13}$$

for $t \ge 0$ and $1 \le i \le N, 1 \le \alpha \le L$.

Proof: From (11), it follows that

$$\begin{aligned}\frac{d\left[x_{i\alpha}(t)-x_{i\alpha}^*\right]}{dt} &= \left[x_{i\alpha}(t)-x_{i\alpha}^*\right]\\ &\quad\times\left[C\left(h_i-\sum_{\beta=1}^{N}x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}x_{j\alpha}^*\right.\\ &\qquad\left.-C\sum_{\beta=1}^{L}\left(x_{i\beta}(t)-x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}\left(x_{j\alpha}(t)-x_{j\alpha}^*\right)\right]\\ &\quad+x_{i\alpha}^*\left[C\left(h_i-\sum_{\beta=1}^{N}x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}x_{j\alpha}^*\right]\\ &\quad+x_{i\alpha}^*\left[-C\sum_{\beta=1}^{L}\left(x_{i\beta}(t)-x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}\left(x_{j\alpha}(t)-x_{j\alpha}^*\right)\right]\\ &= \left[x_{i\alpha}(t)-x_{i\alpha}^*\right]\left[C\left(h_i-\sum_{\beta=1}^{N}x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}x_{j\alpha}^*\right]\\ &\quad+x_{i\alpha}^*\left[-C\sum_{\beta=1}^{L}\left(x_{i\beta}(t)-x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}\left(x_{j\alpha}(t)-x_{j\alpha}^*\right)\right]\\ &\quad+o\left(x(t)-x^*\right).\end{aligned}$$

The result now follows by removing the higher order term. The proof is complete. ■

Next, we show that under some conditions, all the equilibrium points are located in a bounded set.

Denote

$$\Pi = \frac{C\overline{h}}{C-\sum_{i,j=1}^{N}|w_{ij}|}.$$

Lemma 8: Suppose that

$$C > \sum_{i,j=1}^{N}|w_{ij}|.$$

Define a set by

$$D = \left\{x \in \mathbb{R}_+^{NL} \,|\, x_{j\gamma} \le \Pi, (1 \le j \le N, 1 \le \gamma \le L)\right\}.$$

Then, it holds that $\mathcal{E} \subset D$.

Proof: Given any $x^* \in \mathcal{E}$, suppose that

$$x_{i\alpha}^* = \max\left\{x_{j\gamma}^* \middle| 1 \le j \le N; 1 \le \gamma \le L\right\}.$$

If $x_{i\alpha}^* = 0$, clearly, $x^* \in D$. If $x_{i\alpha}^* > 0$, from (12), it must hold that

$$C\left(h_i-\sum_{\beta=1}^{L}x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}x_{j\alpha}^* = 0.$$

Since

$$\begin{aligned}C\left(h_i-\sum_{\beta=1}^{L}x_{i\beta}^*\right)+\sum_{j=1}^{N}w_{ij}x_{j\alpha}^* &\le C\left(\overline{h}-x_{i\alpha}^*\right)+x_{i\alpha}^*\sum_{k,j=1}^{N}|w_{kj}|\\ &= C\overline{h}-x_{i\alpha}^*\left(C-\sum_{k,j=1}^{N}|w_{kj}|\right)\end{aligned}$$

it gives that $x_{i\alpha}^* \le \Pi$. Clearly, $x^* \in D$. The proof is complete. ■

We are especially interested in such equilibrium points that possess some stability properties.

Definition 2: Suppose that x^* is an equilibrium point of network (11). The equilibrium point x^* is called stable, if given any small constant $\epsilon > 0$, there exists a constant $\delta > 0$ such that

$$|x_{i\alpha}(0)-x_{i\alpha}^*| \le \delta, \qquad 1 \le i \le N, \qquad 1 \le \alpha \le L$$

imply that

$$|x_{i\alpha}(t)-x_{i\alpha}^*| \le \epsilon, \qquad 1 \le i \le N, \qquad 1 \le \alpha \le L$$

for all $t \ge 0$. The equilibrium point x^* is called unstable, if it is not stable.

Denote by $\mathcal{E}_s$ the set of stable equilibrium points, and denote by $\mathcal{E}_u$ the set of unstable equilibrium points. Clearly, it holds that

$$\mathcal{E} = \mathcal{E}_s \cup \mathcal{E}_u \quad \mathcal{E}_s \cap \mathcal{E}_u = \phi$$

where ϕ denote the empty set.

Lemma 9: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}}+1\right)\sum_{i,j=1}^{N}|w_{ij}|.$$

If $x^* \in \mathcal{E}_s$, then, given any row index $i, 1 \le i \le N$, there must exist at least one layer index α such that $x_{i\alpha}^* > 0$.

Proof: Otherwise, suppose the opposite is true, then there exists a row index i such that $x^*_{i\alpha} = 0$ for all $1 \le \alpha \le L$. By using (13) of Lemmas 7 and 8, it follows that

$$\begin{aligned}\dot{x}_{i\alpha}(t) &= x_{i\alpha}(t)\left[C\left(h_i - \sum_{\beta=1}^{L} x^*_{i\beta}\right) + \sum_{j=1}^{N} w_{ij}x^*_{j\alpha}\right]\\ &\ge x_{i\alpha}(t)\left[C\underline{h} - \Pi\sum_{k,j=1}^{N}|w_{kj}|\right]\end{aligned}$$

for $t \ge 0$. Since by condition

$$C\underline{h} - \Pi\sum_{k,j=1}^{N}|w_{kj}| > 0$$

then it follows for $x_{i\alpha}(0) \ne 0$ that

$$\begin{aligned}x_{i\alpha}(t) &\ge x_{i\alpha}(0)\exp\left\{\left[C\underline{h} - \Pi\sum_{k,j=1}^{N}|w_{kj}|\right]t\right\}\\ &\to +\infty\end{aligned}$$

as $t \to +\infty$. This contradicts that $x^* \in \mathcal{E}_s$. Thus, there must exist a layer index α such that $x^*_{i\alpha} > 0$. The proof is complete. ∎

Lemma 10: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}} + 1\right)\sum_{i,j=1}^{N}|w_{ij}|.$$

If $x^* \in \mathcal{E}_s$, then, given any row index $i, 1 \le i \le N$, there must exist one and only one layer index α such that

$$\begin{cases} x^*_{i\alpha} > 0 \\ x^*_{i\gamma} = 0, \quad \text{for } \gamma \ne \alpha. \end{cases}$$

Proof: Given any row index i, by Lemma 9, there exists at least one layer index α such that $x^*_{i\alpha} > 0$. Suppose there exists another layer index $\beta(\alpha \ne \beta)$ such that $x^*_{i\beta} > 0$. We will show that x^* is unstable. In order to prove x^* is unstable, it is sufficient to prove the following 2-D system:

$$\begin{cases}\dot{x}_{i\alpha} = x_{i\alpha}\left[C\left(h_i - x_{i\alpha} - x_{i\beta} - \sum\limits_{\gamma\ne\alpha,\beta}^{L} x^*_{i\gamma}\right)\right.\\ \qquad\qquad \left. + w_{ii}x_{i\alpha} + \sum\limits_{j\ne i}^{N} w_{ij}x^*_{j\alpha}\right]\\ \dot{x}_{i\beta} = x_{i\beta}\left[C\left(h_i - x_{i\alpha} - x_{i\beta} - \sum\limits_{\gamma\ne\alpha,\beta}^{L} x^*_{i\gamma}\right)\right.\\ \qquad\qquad \left. + w_{ii}x_{i\beta} + \sum\limits_{j\ne i}^{N} w_{ij}x^*_{j\beta}\right]\end{cases}$$

is unstable at $(x^*_{i\alpha}, x^*_{i\beta})$. The Jacobian matrix of the above system at $(x^*_{i\alpha}, x^*_{i\beta})$ is given by

$$J = \begin{bmatrix} x^*_{i\alpha}(-C + w_{ii}) & -Cx^*_{i\alpha} \\ -Cx^*_{i\beta} & x^*_{i\beta}(-C + w_{ii}) \end{bmatrix}.$$

Let λ_1 and λ_2 be the eigenvalues of the matrix J, then

$$\begin{aligned}\lambda_1\lambda_2 &= |J|\\ &= x^*_{i\alpha}x^*_{i\beta}(-C + w_{ii})^2 - x^*_{i\alpha}x^*_{i\beta}C^2\\ &= x^*_{i\alpha}x^*_{i\beta}w_{ii}(w_{ii} - 2C)\\ &< 0.\end{aligned}$$

This shows that either λ_1 or λ_2 must be positive, thus $(x^*_{i\alpha}, x^*_{i\beta})$ is unstable. The proof is complete. ∎

Lemma 11: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}} + 1\right)\sum_{i,j=1}^{N}|w_{ij}|.$$

Then, the set of stable equilibrium points $\mathcal{E}_s$ must be isolated.

Proof: Let x^* and $x^\dagger$ be two stable equilibrium points. Two cases will be considered next.

Case I. Given any row index i, there exists a layer index α such that

$$x^*_{i\alpha} > 0 \quad x^\dagger_{i\alpha} = 0.$$

Using Lemma 10 and (12), it follows that

$$x^*_{i\alpha} = h_i + \frac{\sum\limits_{j=1}^{N} w_{ij}x^*_{j\alpha}}{C}.$$

By Lemma 8, it follows that

$$\begin{aligned}\|x^* - x^\dagger\| &\ge |x^*_{i\alpha} - x^\dagger_{i\alpha}|\\ &\ge \underline{h} - \frac{\sum\limits_{j=1}^{N}|w_{ij}|x^*_{j\alpha}}{C}\\ &\ge \frac{C - \left(\frac{\overline{h}}{\underline{h}} + 1\right)\sum\limits_{i,j=1}^{N}|w_{ij}|}{\underline{h}\left(C - \sum\limits_{i,j=1}^{N}|w_{ij}|\right)}\\ &> 0.\end{aligned}$$

Thus, there is a distance between x^* and $x^\dagger$.

Case II. Given any row index i, there exists a layer index α such that

$$x^*_{i\alpha} > 0 \quad x^\dagger_{i\alpha} > 0.$$

By using Lemma 10 and (12), it holds that

$$\begin{cases} C(h_i - x^*_{i\alpha}) + \sum\limits_{j=1}^{N} w_{ij}x^*_{j\alpha} = 0\\ C\left(h_i - x^\dagger_{i\alpha}\right) + \sum\limits_{j=1}^{N} w_{ij}x^\dagger_{j\alpha} = 0.\end{cases}$$

Then

$$-C\left[x^*_{i\alpha} - x^\dagger_{i\alpha}\right] + \sum_{j=1}^{N} w_{ij}\left[x^*_{j\alpha} - x^\dagger_{j\alpha}\right] = 0.$$

It follows that

$$C\|x^{*}-x^{\dagger}\| \leq \sum_{i,j=1}^{N} |w_{ij}| \cdot \|x^{*}-x^{\dagger}\|.$$

Clearly, $x^{*} = x^{\dagger}$.

Thus, x^{*} and $x^{\dagger}$ either are equal or have a distance. This shows that $\mathcal{E}_s$ must be isolated. The proof is complete. ■

C. Convergence of Trajectories

In order to use the LV network (11) to implement the CLM, one must address the problem of whether each trajectory of the network is convergent. This section studies the convergence of the trajectories of network (11).

The set of equilibrium points $\mathcal{E}$ of network (11) is composed of two disjoint sets $\mathcal{E}_s$ and $\mathcal{E}_u$. The set $\mathcal{E}_u$ is composed of the unstable equilibrium points, i.e., each element of $\mathcal{E}_u$ is an unstable equilibrium point. Given any neighborhood of $\mathcal{E}_u$, no matter how small it is, there must exist points in the neighborhood such that the trajectories starting from them will diverge. However, there may still exist some trajectories that converge to the set $\mathcal{E}_u$. A trajectory $x(t)$ is said to converge to $\mathcal{E}_u$, if

$$\lim_{t\to+\infty} \inf_{x^{\dagger}\in\mathcal{E}_u} \|x(t)-x^{\dagger}\| = 0.$$

It is well known that in practical applications only stable equilibrium points can be observed. The above convergence can be easily destroyed by some small disturbances. Thus, from the application point of view, trajectories converging to the set $\mathcal{E}_u$ are not interesting and can be ignored.

Definition 3: A trajectory is called uninteresting, if it converges to a set of unstable equilibrium points. A trajectory is called interesting if it is not uninteresting.

Next, we will address the problem of whether each interesting trajectory is convergent.

Definition 4: Given any $x_0 \in \mathbb{R}_{+}^{NL}$, let $x(t,x_0)$ be a trajectory starting from x_0. A point $x^{\dagger} \in \mathbb{R}_{+}^{NL}$ is called an ω-limit point of $x(t,x_0)$, if there exists a time sequence $\{t_k\}$ with $t_k \to +\infty$ as $k \to +\infty$, such that

$$\lim_{k\to+\infty} x(t_k, x_0) = x^{\dagger}.$$

Denote by $\Omega(x_0)$ the set of all ω-limit points of $x(t,x_0)$ and call $\Omega(x_0)$ the ω-limit set of $x(t,x_0)$.

Lemma 12 [29]: If $x(t,x_0)$ is bounded, then $\Omega(x_0)$ must be a connected set.

Lemma 13: Each trajectory of (11) is bounded.

Proof: Denote

$$\tilde{\Pi} = \max\left\{\|x(0)\|, \frac{C\overline{h}}{C-\sum_{i,j=1}^{N}|w_{ij}|}\right\}+1.$$

It will be proven that

$$x_{i\alpha}(t) < \tilde{\Pi}, \qquad 1\leq i\leq N, \qquad 1\leq\alpha\leq L$$

for all $t \geq 0$. Otherwise, there must exist a row index i and a layer index α and time $t_1 > 0$ such that $\dot{x}_{i\alpha}(t_1) \geq 0$, and

$$x_{i\alpha}(t_1) = \tilde{\Pi} \quad x_{j\beta}(t)\begin{cases} < \tilde{\Pi}, & 0\leq t<t_1, j\beta = i\alpha \\ \leq \tilde{\Pi}, & 0\leq t\leq t_1, j\beta \neq i\alpha.\end{cases}$$

However, from (11), it gives that

$$\begin{aligned}\dot{x}_{i\alpha}(t_1) &\leq x_{i\alpha}(t_1)\left[C\left(h_i - x_{i\alpha}(t_1)\right) + \sum_{j=1}^{N} w_{ij}x_{j\alpha}(t_1)\right] \\ &\leq x_{i\alpha}(t_1)\left[C\left(\overline{h}-\tilde{\Pi}\right) + \tilde{\Pi}\sum_{k,j=1}^{N}|w_{kj}|\right] \\ &< 0.\end{aligned}$$

This is a contradiction, thus, each trajectory of (11) must be bounded. The proof is complete. ■

Lemma 14: Along each trajectory of network (11), the CLM energy function $E(x)$ satisfies that

$$\begin{cases} \dot{E}(x(t)) \leq 0 \\ \dot{E}(x(t)) = 0, \quad \text{if and only if } \dot{x}(t) = 0\end{cases}$$

for all $t \geq 0$.

Proof: It follows from (3) and (11) that

$$\begin{aligned}\dot{E}(t) &= -\sum_{i=1}^{N}\sum_{\alpha=1}^{L}\left[C\left(h_i - \sum_{\beta=1}^{L}x_{i\beta}(t)\right) + \sum_{j=1}^{N}w_{ij}x_{j\alpha}(t)\right] \\ &\quad \times \dot{x}_{i\alpha}(t) \\ &= -\sum_{i=1}^{N}\sum_{\alpha=1}^{L}\left[C\left(h_i - \sum_{\beta=1}^{L}x_{i\beta}(t)\right) + \sum_{j=1}^{N}w_{ij}x_{j\alpha}(t)\right]^2 \\ &\quad \times x_{i\alpha}(t) \\ &\leq 0\end{aligned}$$

for $t \geq 0$. Clearly, $\dot{E}(t) = 0$ if and only if

$$C\left(h_i - \sum_{\beta=1}^{L}x_{i\beta}(t)\right) + \sum_{j=1}^{N}w_{ij}x_{j\alpha}(t) = 0$$

or $x_{i\alpha}(t) = 0$ for $t \geq 0$ and $1 \leq i \leq N, 1 \leq \alpha \leq L$. By (11), it gives that $\dot{x}(t) = 0$ for $t \geq 0$. The proof is complete. ■

Theorem 2: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}}+1\right)\sum_{i,j=1}^{N}|w_{ij}|.$$

Then, each interesting trajectory of the network (11) converges to an equilibrium point.

Proof: By Lemma 13, $x(t,x_0)$ is bounded. Then, there exists a bounded closed set $S \subset \mathbb{R}_{+}^{NL}$ such that $x(t,x_0) \subset S$ for all $t \geq 0$. Let $\Omega(x_0)$ be the ω-limit set of $x(t,x_0)$. Clearly, $\Omega(x_0) \neq \phi$, and $\Omega(x_0) \subset S$. By Lemma 14, $\dot{E}(x(t,x_0)) \leq 0$ for $t \geq 0$, and $x(t,x_0) \subset S$, then $E(x(t,x_0))$ is a monotone decrease bounded function, thus, there must exist a constant E_0 such that

$$\lim_{t\to+\infty} E(x(t,x_0)) = E_0.$$

For any $y \in \Omega(x_0)$, let $x(t,y)$ be the trajectory of (11) passing through y. Clearly, $x(t,y) \in \Omega(x_0)$ for every $t \geq 0$. Since E is continuous, then $E(x(t,y)) = E_0$ for all $t \geq 0$. Moreover

$$\dot{E}(x(t,y)) = 0, \qquad t \geq 0.$$

Then, it follows that $\dot{x}(t,y) = 0$ for all $t \geq 0$. This implies that $x(t,y) \equiv y$ is an equilibrium point, i.e., $y \in \mathcal{E}$. Since y is arbitrarily chosen in $\Omega(x_0)$, then all the points of $\Omega(x_0)$ are equilibrium points, i.e., $\Omega(x_0) \subset \mathcal{E}$. Moreover, $\Omega(x_0) \subset \mathcal{E}_s$, otherwise, $x(t,x_0)$ must be an uninteresting trajectory. Since by Lemma 11, the equilibrium points in $\mathcal{E}_s$ are isolated, then the points of $\Omega(x_0)$ are also isolated. By Lemma 12, every ω-limit set is connected, then $\Omega(x_0)$ must contain one equilibrium point x^* only. Thus, $x(t,x_0)$ converges to x^*. This completes the proof. ■

D. The Set of Stable Attractors

Definition 5: Let x^* be a stable equilibrium point. It is called a stable attractor, if there exists some small neighborhood B of x^* such that $x(0) \in B$ implies that

$$\lim_{t\to+\infty} x_{i\alpha}(t) = x^*_{i\alpha}$$

for $1 \leq i \leq N$ and $1 \leq \alpha \leq L$.

A stable attractor is an equilibrium point with the following properties: 1) it should be stable in the sense of Definition 2, and 2) it must attract trajectories around it. In this paper, we denote by S_A the set of all stable attractors of (11). Clearly, $S_A \subseteq \mathcal{E}_s$. In this section, the main purpose is to prove that $\mathcal{M} = S_A$.

Lemma 15: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}} + 1\right) \sum_{i,j=1}^{N} |w_{ij}|.$$

Then, it holds that $S_A \subseteq \mathcal{M}$.

Proof: For any $x^* \in S_A \subseteq \mathcal{E}_s$, it will be proved that $x^* \in \mathcal{M}$.

Given any row index $i (1 \leq i \leq N)$, since $x^* \in \mathcal{E}_s$, by Lemma 10, there must exist one and only one layer index α such that

$$\begin{cases} x^*_{i\alpha} > 0 \\ x^*_{i\gamma} = 0, \quad \text{for } \gamma \neq \alpha. \end{cases} \tag{14}$$

Since x^* is an equilibrium point, from (12), it holds that

$$C\left(h_i - x^*_{i\alpha}\right) + \sum_{j=1}^{N} w_{ij} x^*_{j\alpha} = 0. \tag{15}$$

Using Lemma 7, it gives for $\gamma \neq \alpha$ that

$$\dot{x}_{i\gamma}(t) = x_{i\gamma}(t)\left[C\left(h_i - x^*_{i\alpha}\right) + \sum_{j=1}^{N} w_{ij} x^*_{j\gamma}\right]$$

for $t \geq 0$. Then

$$x_{i\gamma}(t) = x_{i\gamma}(0) \exp\left\{\left[C\left(h_i - x^*_{i\alpha}\right) + \sum_{j=1}^{N} w_{ij} x^*_{j\gamma}\right] t\right\}$$

for $t \geq 0$ and $\gamma \neq \alpha$. Since x^* is a stable attractor, it must hold that

$$C\left(h_i - x^*_{i\alpha}\right) + \sum_{j=1}^{N} w_{ij} x^*_{j\gamma} < 0, \qquad \text{for } \gamma \neq \alpha. \tag{16}$$

By (14)–(16), using Theorem 1, clearly, $x^* \in \mathcal{M}$. The proof is complete. ■

Lemma 16: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}} + 1\right) \sum_{i,j=1}^{N} |w_{ij}|.$$

Then, it holds that $\mathcal{M} \subseteq S_A$.

Proof: Let $x^* \in \mathcal{M}$. By Theorem 1, given any row index $i (1 \leq i \leq N)$, there exists one and only one layer index $\alpha (1 \leq \alpha \leq L)$ such that

$$\begin{cases} x^*_{i\alpha} > 0 \\ x^*_{i\gamma} = 0, \quad \text{for } \gamma \neq \alpha \end{cases} \tag{17}$$

and

$$\begin{cases} C\left(x^*_{i\alpha} - h_i\right) - \sum_{j=1}^{N} w_{ij} x^*_{j\alpha} = 0 \\ C\left(x^*_{i\alpha} - h_i\right) - \sum_{j=1}^{N} w_{ij} x^*_{j\gamma} > 0, \quad \text{for } \gamma \neq \alpha. \end{cases} \tag{18}$$

Denote two index sets by

$$\begin{cases} P = \{i\alpha | x^*_{i\alpha} > 0; 1 \leq i \leq N; 1 \leq \alpha \leq L\} \\ Z = \{i\gamma | x^*_{i\gamma} = 0; 1 \leq i \leq N; 1 \leq \gamma \leq L\}. \end{cases}$$

Define

$$\begin{cases} \|x_Z\| = \sqrt{\sum_{i\gamma \in Z} x^2_{i\gamma}} \\ \|x_P - x^*_P\| = \sqrt{\sum_{i\alpha \in P} \left(x_{i\alpha} - x^*_{i\alpha}\right)^2}. \end{cases}$$

Given $i\gamma \in Z$, by Lemma 7 and (17), it follows that

$$\dot{x}_{i\gamma}(t) = x_{i\gamma}(t)\left[C\left(h_i - x^*_{i\alpha}\right) + \sum_{j=1}^{N} w_{ij} x^*_{j\gamma}\right]$$

for $t \geq 0$. That is

$$x_{i\gamma}(t) = x_{i\gamma}(0) \exp\left(\left[C\left(h_i - x^*_{i\alpha}\right) + \sum_{j=1}^{N} w_{ij} x^*_{j\gamma}\right] t\right)$$

for $t \geq 0$. Denote

$$\delta \triangleq -2 \min_{i\gamma \in Z} \left\{ C\left(h_i - x_{i\alpha}^*\right) + \sum_{j=1}^{N} w_{ij} x_{j\gamma}^* \right\}.$$

By (18), clearly, $\delta > 0$. Then

$$\|x_Z(t)\|^2 \leq \|x_Z(0)\|^2 e^{-\delta t} \tag{19}$$

for $t \geq 0$.

Given $i\alpha \in P$, by Lemma 7, it follows that

$$\begin{aligned} \frac{d\left[x_{i\alpha}(t) - x_{i\alpha}^*\right]}{dt} &= x_{i\alpha}^* \Bigg[-C\left(x_{i\alpha}(t) - x_{i\alpha}^*\right) - C \sum_{\beta=1, \beta\neq\alpha, i\beta\in Z}^{L} x_{i\beta}(t) \\ &\quad + \sum_{j=1, j\alpha\in P}^{N} w_{ij}\left(x_{j\alpha}(t) - x_{j\alpha}^*\right) + \sum_{j=1, j\alpha\in Z}^{N} w_{ij} x_{j\alpha}(t) \Bigg] \end{aligned} \tag{20}$$

for $t \geq 0$.

Define the following Lyapunov function:

$$V(t) = \frac{1}{2} \sum_{i\alpha\in P} \frac{\left[x_{i\alpha}(t) - x_{i\alpha}^*\right]^2}{x_{i\alpha}^*}$$

for $t \geq 0$. Clearly

$$\begin{aligned} &\frac{1}{2\max_{i\alpha\in P}\left\{x_{i\alpha}^*\right\}} \|x_P(t) - x_P^*\|^2 \\ &\leq V(t) \\ &\leq \frac{1}{2\min_{i\alpha\in P}\left\{x_{i\alpha}^*\right\}} \|x_P(t) - x_P^*\|^2 \end{aligned}$$

for $t \geq 0$.

Denote by

$$w = \max\left\{ |w_{ij}| \,|\, 1 \leq i, j \leq N \right\}.$$

From (20) and (19), by calculating, it gives that

$$\begin{aligned} &\dot{V}(t) \\ &= \sum_{i\alpha\in P} \frac{\left[x_{i\alpha}(t) - x_{i\alpha}^*\right]}{x_{i\alpha}^*} \frac{d\left[x_{i\alpha}(t) - x_{i\alpha}^*\right]}{dt} \\ &\leq -(C-w)\|x_P(t) - x_P^*\|^2 \\ &\quad + (C+w)\|x_P(t) - x_P^*\| \cdot \|x_Z(t)\| \\ &\leq -\frac{C-w}{2}\|x_P(t) - x_P^*\|^2 + \frac{(C+w)^2}{2(C-w)}\|x_Z(t)\|^2 \\ &\leq -\frac{C-w}{2}\|x_P(t) - x_P^*\|^2 + \frac{(C+w)^2}{2(C-w)}\|x_Z(0)\|^2 e^{-\delta t} \\ &\leq -(C-w)\min_{i\alpha\in P}\left\{x_{i\alpha}^*\right\} \cdot V(t) + \frac{(C+w)^2}{2(C-w)}\|x_Z(0)\|^2 e^{-\delta t} \\ &\leq -\epsilon \cdot V(t) + \frac{(C+w)^2}{2(C-w)}\|x_Z(0)\|^2 e^{-\delta t} \end{aligned}$$

for $t \geq 0$, where

$$\epsilon = \min\left\{ (C-w)\min_{i\alpha\in P}\left\{x_{i\alpha}^*\right\}, \frac{\delta}{2} \right\} > 0.$$

Then

$$\begin{aligned} V(t) &\leq V(0)e^{-\epsilon t} + \frac{(C+w)^2 e^{-\epsilon t}}{2(C-w)} \cdot \|x_Z(0)\|^2 \int_0^t e^{(\epsilon-\delta)s} ds \\ &\leq \frac{1}{2\min_{i\alpha\in P}\{x_{i\alpha}^*\}}\|x_P(0) - x_P^*\|^2 e^{-\epsilon t} \\ &\quad + \frac{(C+w)^2}{2(\delta-\epsilon)(C-w)}\|x_Z(0)\|^2 e^{-\epsilon t} \end{aligned}$$

for $t \geq 0$. Thus

$$\begin{aligned} \|x_P(t) - x_P^*\|^2 \leq &\frac{\max_{i\alpha\in P}\{x_{i\alpha}^*\}}{\min_{i\alpha\in P}\{x_{i\alpha}^*\}}\|x_P(0) - x_P^*\|^2 e^{-\epsilon t} \\ &+ \frac{\max_{i\alpha\in P}\{x_{i\alpha}^*\}(C+w)^2}{(\delta-\epsilon)(C-w)}\|x_Z(0)\|^2 e^{-\epsilon t} \end{aligned}$$

for $t \geq 0$. By Definition 2, clearly, x^* is a stable attractor. Thus, $\mathcal{M} \subseteq S_A$. The proof is complete. ■

Theorem 3: Suppose that

$$C > \left(\frac{\overline{h}}{\underline{h}} + 1\right) \sum_{i,j=1}^{N} |w_{ij}|.$$

Then, it holds that $\mathcal{M} = S_A$, i.e., the set of the minimum points of the CLM energy function equals the set of the stable attractors of the LV network (11).

Proof: The result follows directly from Lemma 15 and Lemma 16. ■

V. Discussions

In this section, some discussions will be carried out to further illustrate the theory. We will further interpret the concept of attractors and show the group binding capability of the proposed network.

A. The Attractors

To further understand the attractors of the LV RNN (11), let us consider the case where $L = 2$, $N = 2$, $w_{ij} = w > 0, i, j = 1, \ldots, 2,$, and $h_i = h > 0, i = 1, 2$. The network can be written as

$$\begin{cases} \dot{x}_{11} = x_{11}\left[C\left(h - x_{11} - x_{12}\right) + w\left(x_{11} + x_{21}\right)\right] \\ \dot{x}_{21} = x_{21}\left[C\left(h - x_{21} - x_{22}\right) + w\left(x_{11} + x_{21}\right)\right] \\ \dot{x}_{12} = x_{12}\left[C\left(h - x_{11} - x_{12}\right) + w\left(x_{12} + x_{22}\right)\right] \\ \dot{x}_{22} = x_{22}\left[C\left(h - x_{21} - x_{22}\right) + w\left(x_{12} + x_{22}\right)\right] \end{cases} \tag{21}$$

for $t \geq 0$. The corresponding CLM energy function is as follows:

$$\begin{aligned} E = &\frac{C}{2}\left[\left(x_{11} + x_{12} - h\right)^2 + \left(x_{21} + x_{22} - h\right)^2\right] \\ &- \frac{w}{2}\left[\left(x_{11} + x_{21}\right)^2 + \left(x_{12} + x_{22}\right)^2\right]. \end{aligned} \tag{22}$$

Suppose that $C > 4w$. Using Theorem 1, it can be calculated that the set of minimum points of the CLM energy function (22) is

$$\mathcal{M} = \left\{ \begin{pmatrix} \frac{Ch}{C-2w} \\ \frac{Ch}{C-2w} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \frac{Ch}{C-2w} \\ \frac{Ch}{C-2w} \end{pmatrix} \right\}.$$

From (21), solving the equilibrium equation, it gives that

$$\mathcal{E}_s = \left\{ \begin{pmatrix} \frac{Ch}{C-2w} \\ \frac{Ch}{C-2w} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \frac{Ch}{C-2w} \\ \frac{Ch}{C-2w} \end{pmatrix} \right\}$$

and

$$\mathcal{E}_u = \left\{ \begin{pmatrix} \frac{Ch}{2(C-w)} + \kappa \\ \frac{Ch}{2(C-w)} - \kappa \\ \frac{Ch}{2(C-w)} - \kappa \\ \frac{Ch}{2(C-w)} + \kappa \end{pmatrix} \middle| \, |\kappa| \le \frac{Ch}{2(C-w)}, \kappa \in \mathbb{R} \right\} \cup \{0\}.$$

The set of stable attractors $S_A = \mathcal{E}_s$ is a disconnected set; it contains two equilibrium points. Clearly, $\mathcal{M} = S_A$, which is consistent with Theorem 3.

The set of unstable equilibrium points $\mathcal{E}_u \setminus \{0\}$ is a connected set. It forms a continuous unstable attractor. Clearly, $\mathcal{E}_u \setminus \{0\}$ is embedded in 1-D manifold. Trajectories starting from the points of the following set of initial points:

$$S_{\text{init}} = \left\{ \begin{pmatrix} \nu \\ \nu \\ \nu \\ \nu \end{pmatrix} \middle| \, \nu \in \mathbb{R}_+ \right\} \cup \mathcal{E}_u$$

converge to $\mathcal{E}_u$. However, such trajectories are uninteresting and cannot be observed in practice since these trajectories are strictly restricted on some 1-D or nondimensional manifold, any small disturbance to the initial points can destroy the convergence.

Any trajectories starting from points in $\mathbb{R}^4_+ \setminus S_{\text{init}}$ are interesting and converge to either the first or the second attractor of S_A. Whether an interesting trajectory converges to the first or the second attractor of S_A depends on the location of the initial points. Taking $C = 500$, $w = 40$, and $h = 1$, Fig. 3 shows that an interesting trajectory converges to the first attractor of S_A and the corresponding energy evolution, while Fig. 4 shows that another interesting trajectory converges to the second attractor of S_A and its corresponding energy evolution.

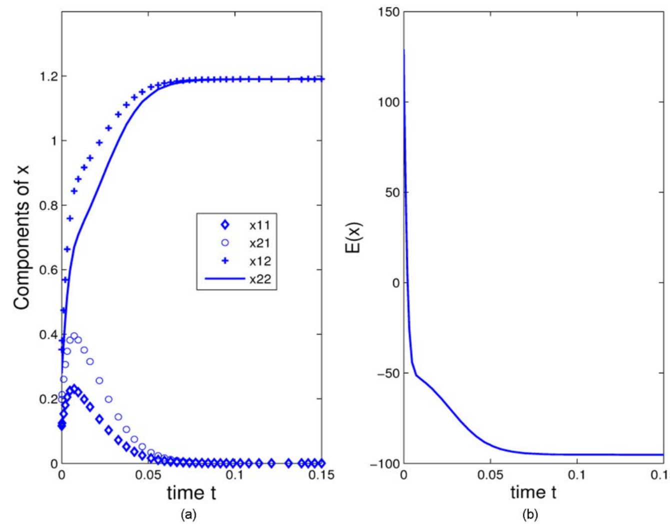


Fig. 3. (a) Trajectory starting from a randomly generated point $x(0) = (0.1159, 0.1981, 0.3525, 0.2793)^T$ converges to the first attractor of S_A. (b) Corresponding evolution of the energy function.

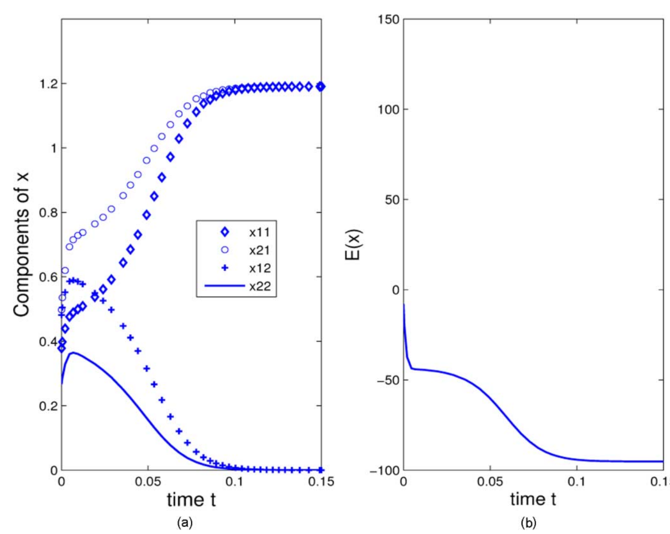
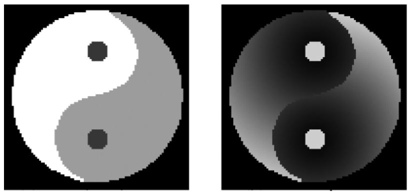


Fig. 4. (a) Trajectory starting from a randomly generated point $x(0) = (0.3783, 0.4977, 0.4812, 0.2675)^T$ converges to the second attractor of S_A. (b) Corresponding evolution of the energy function.

B. Group Binding Capability

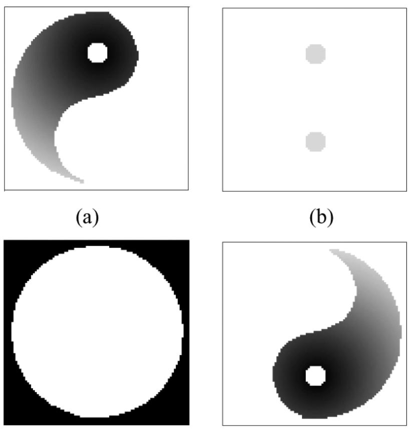
In this section, we will illustrate the binding capability for stored patterns of the proposed LV RNN (11). When the network converges to a stable attractor x^*, each neuron can be either active ($x^*_{i\alpha} > 0$) or inactive ($x^*_{i\alpha} = 0$). The set of active neurons in each layer indicates a stored pattern, and the value of $x^*_{i\alpha}$ carries analog information of the inputs. The relationship between the input $h_i, 1 \le i \le N$, and the output $x^*_{i\alpha}, 1 \le i \le N, 1 \le \alpha \le L$, can be simply calculated as

$$x^*_{i\alpha} = h_i + \frac{\sum_{j=1}^{N} w_{ij} x^*_{j\alpha}}{C}. \tag{23}$$



(a) (b)

Fig. 5. (a) Stored image which contains four groups indicated by the four different gray values, respectively, and (b) input image which contains gray information. Both images have the size 95 × 95.



(a) (b) (c) (d)

Fig. 6. Binding result of the network. The four groups are simultaneously bound into four different layers, respectively. (a) Layer 1. (b) Layer 2. (c) Layer 3. (d) Layer 4.

If C is sufficiently large, then $x^*_{i\alpha} \approx h_i$. This relationship would be very important in designing the network for practical applications.

In practical applications, patterns are stored as stable attractors in the network. Since each stable attractor has an attractive domain, the network runs from any initial vector in the domain can converge to the attractor and then extract the stored patterns. This indicates that the network is quite robust in the sense of extracting stored patterns.

Next, we use some images with size of 95×95 to further illustrate the group binding capability of the network. Fig. 5(a) shows an image which contains four groups indicated by four different gray values, respectively. In out experiment, the image is row-by-row vectorized into a vector with dimension of 95×95, and the value of each element, denoted by $p_i, 1 \leq i \leq 95 \times 95$, is the gray value of the corresponding pixel. To store the four groups of this image into network (11), the connection weight w_{ij} can be calculated as

$$w_{ij} = \begin{cases} 1, & \text{if } p_i = p_j \\ -1, & \text{otherwise.} \end{cases}$$

Next, we show that the groups can be simultaneously bounded into different layers of the network, respectively. We use an image with the same structure of patterns but each pattern carries varying gray information as the input $h_i, 1 \leq i \leq 95 \times 95$, of the network. Fig. 5(b) shows the input image. The input image is row-by-row vectorized into a vector with dimension of 95×95, and the value of each element h_i is the gray value of the corresponding pixel. By randomly choosing an initial vector $x(0)$ and running the network, the network converges to a stable attractor, and the set of active neurons in each layer indicates the extracted group. Meanwhile, the values of the active neurons indicate the gray information of the input image. Fig. 6 shows that the four groups of the input image are simultaneously bounded into the four different layers by the network, respectively.

The CLM energy function provides an energy-based approach for feature binding. The proposed network of LV RNN (11) can be used to implement the CLM for feature binding. There are some other models of energy-based approaches for feature binding, for example, the spin models [4]–[6] and RL [7], [8]. These models lack biological plausibility since they either require iterative discrete cluster update procedures or complex normalizing nonlinearities. The model of LV RNNs can be derived from the conventional membrane dynamics of neurons. It has straightforward neural circuit interpretation and fast convergence property. In [3], LT RNNs are used to implement the CLM for feature binding and sensory segmentation. However, the rigorous theories of using RNNs to implement the CLM have not been established. The proposed network of LV RNN (11) in this paper provides a clear understanding of the dynamics of the network for implementing the CLM.

VI. Conclusion

The CLM is an interesting model of neural networks. It has found important applications in spatial feature binding. Foundations of implementing the CLM by LV RNNs have been established in this paper. The theory contains three parts. First, necessary and sufficient conditions of minimum points of the CLM energy function are obtained. Second, it is proven that each interesting trajectory of the LV RNN is convergent. Third, the set of minimum points of the CLM energy function is equal to the set of stable attractors of the LV RNN. The methods of this paper could be further developed to establish rigorous foundations of implementing the CLM by other RNNs, say LT RNNs. It is believed that by establishing these important theories, more practical applications could be found.

Acknowledgment

The author would like to thank the anonymous reviewers for giving many good suggestions that improved the quality of this paper.

References

[1] H. Retter, "A spatial approach for feature linking," in *Proc. Int. Neural Netw. Conf.*, Paris, France, 1990, vol. 2, pp. 898–901.
[2] C. von der Malsburg, "Binding in models of perception and brain function," *Current Opinion Neurobiol.*, vol. 5, pp. 520–526, 1995.
[3] H. Wersing, J. J. Steil, and H. Ritter, "A competitive layer model for feature binding and sensory segmentation," *Neural Comput.*, vol. 13, no. 2, pp. 357–387, 2001.
[4] D. Geman, S. Geman, C. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 609–628, Jul. 1990.
[5] L. Herault and R. Horaud, "Figure-ground discrimination: A combinatorial optimization approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 899–914, Sep. 1993.

[6] R. Opara and F. Wörgötter, "A fast and robust cluster update algorithm for image segmentation in spin-lattice models without annealing—Visual latencies revisited," *Neural Comput.*, vol. 10, no. 6, pp. 1547–1566, 1998.
[7] A. Rosenfed, R. A. Hummel, and S. W. Zucker, "Scene labeling by relaxation operations," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-6, no. 6, pp. 420–433, Jun. 1976.
[8] R. A. Hummel and S. W. Zucker, "On the foundations of relaxation labeling processes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 5, no. 3, pp. 267–287, Mar. 1983.
[9] M. Pelillo and A. M. Fanelli, "Autoassociative learning in relaxation labeling networks," *Pattern Reecognit. Lett.*, vol. 18, no. 1, pp. 3–12, 1997.
[10] J. Kittler and J. Illingworth, "Relaxation labeling algorithms—A review," *Image Vis. Comput.*, vol. 3, no. 4, pp. 206–216, 1985.
[11] H. Wersing and H. Ritter, "Feature binding and relaxation labeling with the competitive layer model," in *Proc. Eur. Symp. Artif. Neural Netw.*, Apr. 1999, pp. 295–300.
[12] R. Ben-Yishai, R. Lev Bar-Or, and H. Sompolinsky, "Theory of orientation tuning in visual cortex," in *Proc. Nat. Acad. Sci.*, 1995, vol. 92, pp. 3844–3848.
[13] 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.
[14] R. Dougleas, C. Koch, M. Mahowald, K. Matin, and H. Suarez, "Recurrent excitation in neocortical circuits," *Science*, vol. 269, pp. 981–985, 1995.
[15] 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.
[16] 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.
[17] R. Hahnloser, "On the piecewise analysis of linear threshold neural networks," *Neural Netw.*, vol. 11, pp. 691–697, 1998.
[18] 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.
[19] Z. Yi, L. Zhang, J. Yu, and K. K. Tan, "Permitted an forbidden sets in discrete-time linear threshold recurrent neural networks," *IEEE Trans. Neural Netw.*, vol. 20, no. 6, pp. 952–963, Jun. 2009.
[20] 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.
[21] 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, Mar. 2003.
[22] Z. Yi and K. K. Tan, "Multistability analysis of discrete recurrent neural networks with unsaturating piecewise linear transfer functions," *IEEE Trans. Neural Netw.*, vol. 15, no. 2, pp. 329–336, Mar. 2004.
[23] 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.
[24] S. Weng, H. Wersing, J. J. Steil, and H. Ritter, "Learning lateral interactions for feature binding and sensory segmentation from prototypic basis interactions," *IEEE Trans. Neural Netw.*, vol. 17, no. 4, pp. 843–862, Jul. 2006.
[25] T. Fukai and S. Tanaka, "A simple neural network exhibiting selective activation of neuronal ensembles: From winner-take-all to winner-share-all," *Neural Comput.*, vol. 9, pp. 77–97, 1997.
[26] T. Asai, T. Fukai, and S. Tanaka, "A subthreshold MOS circuit for the Lotka-Volterra neural network producing the winner-take-all solutions," *Neural Netw.*, vol. 12, pp. 211–216, 1999.
[27] T. Asai, M. Ohtani, and H. Yonezu, "Analog integrated circuits for the Lotka-Volterra competitive neural networks," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 1222–1231, Sep. 1999.
[28] 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.
[29] Z. Yi and K. K. Tan, *Convergence Analysis of Recurrent Neural Networks*. Boston, MA: Kluwer, 2004.
[30] 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.

Zhang Yi (M'08) received the Ph.D. degree in mathematics from the Institute of Mathematics, Chinese Academy of Science, Beijing, China, in 1994.

Currently, he is a Professor at Sichuan University, Chengdu, China, and acts as the Dean of the College of Computer Science. He is the founder of the Machine Intelligence Laboratory, College of Computer Science, Sichuan University. He has published more than 100 refereed articles. 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). He holds two patents in China for his data mining system and seal recognition system based on neural networks methods. His research interests include machine intelligence, neural networks, data mining, and machine learning. Currently, he is especially interested in neocortex computing towards intelligent machines.