Absolute Exponential Stability of Recurrent Neural Networks With Generalized Activation Function

Jun Xu, Yong-Yan Cao, *Senior Member, IEEE*, Youxian Sun, and Jinshan Tang, *Senior Member, IEEE*

***Abstract*—In this paper, the recurrent neural networks (RNNs) with a generalized activation function class is proposed. In this proposed model, every component of the neuron's activation function belongs to a convex hull which is bounded by two odd symmetric piecewise linear functions that are convex or concave over the real space. All of the convex hulls are composed of generalized activation function classes. The novel activation function class is not only with a more flexible and more specific description of the activation functions than other function classes but it also generalizes some traditional activation function classes. The absolute exponential stability (AEST) of the RNN with a generalized activation function class is studied through three steps. The first step is to demonstrate the global exponential stability (GES) of the equilibrium point of original RNN with a generalized activation function being equivalent to that of RNN under all vertex functions of convex hull. The second step transforms the RNN under every vertex activation function into neural networks under an array of saturated linear activation functions. Because the GES of the equilibrium point of three systems are equivalent, the next stability analysis focuses on the GES of the equilibrium point of RNN system under an array of saturated linear activation functions. The last step is to study both the existence of equilibrium point and the GES of the RNN under saturated linear activation functions using the theory of M-matrix. In the end, a two-neuron RNN with a generalized activation function is constructed to show the effectiveness of our results.**

***Index Terms*—Absolute exponential stability (AEST), convex hull, generalized activation function class, piecewise linear function, recurrent neural networks (RNNs).**

I. Introduction

IN RECENT YEARS, absolute stability (ABST) and absolute exponential stability (AEST) analysis of neural networks have received a great deal of attention (see, for example, [2]–[5], [10], [11], [17]–[23], [33], [36], and [37]). A neural network is called ABST (AEST) if there is a unique and globally asymptotically stable (globally exponentially stable) equilibrium point for every neuron activation function belonging to a given activation function class and for every constant input vector to the neural networks. The interest in ABST and AEST analysis comes from practical problems, such as in optimization, signal processing, pattern recognition, and associative memories, when the activation functions belong to some known function class but the exact shape is not known or its model is difficult to describe mathematically. In designing a neural network, one is concerned not only with the stability of the system but also with the convergence rate, that is to say, a fast response in the neural network is usually preferred. On the other hand, in most of the applications of neural networks, the networks' convergence speed is expected to increase in order to cut down on the neural computing time. Thus, it is also important to determine the exponential stability. AEST is an interesting and significant topic in the stability analysis of neural networks. The AEST result extends many types of stability such as global asymptotical stability (GAS) [1], [3], [10], [11], [38], [43], global exponential stability (GES) [31], [33], [44], [45], and [49], local stability [46], [50], and ABST [2], [10]–[13], [37], etc., thus AEST is the strongest one.

A suitable and more generalized activation function can greatly improve the neural networks' performance. For example, the property of the activation function is important to the capacity of neural network. Morita *et al.* showed in [24] and [32] that the absolute capacity of an associative memory model can be remarkably improved by replacing the usual sigmoid activation function with a nonmonotonic activation function. Therefore, it is very significant to design a new artificial neural network with a more generalized activation function class. In recent years, many researchers have devoted themselves to attain this goal by proposing new generalized activation function classes. So far, there are five kinds of activation function classes being proposed for neural networks.

1) Sigmoid function class $\mathcal{S}$ [1], [6], [10]: a function $f(x) : \mathcal{R}^n \rightarrow \mathcal{R}^n$ is said to be in class $\mathcal{S}$ if for $i = 1, \cdots, n$, $f_i(x_i) : \mathcal{R} \rightarrow \mathcal{R}$ is a $\mathcal{C}_1$ function with $0 < f_i'(x_i)$ and $f_i(x_i)$ is bounded for all $x_i \in \mathcal{R}$.
2) The class of globally Lipschitz continuous (g.l.c.) and monotone nondecreasing activation functions $\mathcal{GL}$ [11]: for $\forall \theta, \rho \in \mathcal{R}$ and $\theta \neq \rho$, there exist constants $l_i > 0, (i = 1, 2, \cdots, n)$ such that

$$0 \leq \frac{g_i(\theta) - g_i(\rho)}{\theta - \rho} \leq l_i.$$

3) The class of partially Lipschitz continuous (p.l.c.) and monotone nondecreasing activation functions $\mathcal{PL}$ [19]:

Manuscript received November 19, 2006; revised August 7, 2007; accepted October 31, 2007. This work was supported by the National Science Foundation of China under Grants 60574019 and 60474045.

J. Xu is with the System Research Institute & Department of Advanced Technologies, Alcorn State University, Alcorn State, MS 39096 USA and also with the School of Information Technology, Jiangxi University of Finance & Economics, Nanchang 330013, China (e-mail:xujung@gmail.com).

Y.-Y. Cao is with the Department of Automation, Wuhan University of Science and Technology, Wuhan 430081, China and also with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China (e-mail: yongyancao@gmail.com).

Y. X. Sun is with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China (e-mail: yxsun@iipc.zju.edu.cn).

J. Tang is with the Image Processing and Bio-Imaging Research Laboratory, System Research Institute, Department of Advanced Technologies, Alcorn State University, Alcorn State, MS 39096 USA (e-mail:jtang@alcorn.edu).

Digital Object Identifier 10.1109/TNN.2007.2000060

for $\forall\rho \in \mathcal{R}$, there exist constants $l_i(\rho)(i = 1, 2, \cdots, n)$ such that $\forall\theta \in \mathcal{R}$ and $\theta \neq \rho$

$$0 \leq \frac{g_i(\theta) - g_i(\rho)}{\theta - \rho} \leq l_i(\rho).$$

4) The class of locally Lipschitz continuous (l.l.c.) and monotone nondecreasing activation functions $\mathcal{LL}$ [17]: for $\forall x_{i_0} \in \mathcal{R}$, there exist a ε_{i_0} and a constant $l_{i_0} > 0$, $(i = 1, 2, \cdots, n)$ such that $\forall\theta, \rho \in [x_{i_0} - \varepsilon_{i_0}, x_{i_0} + \varepsilon_{i_0}]$ and $\theta \neq \rho$

$$0 \leq \frac{g_i(\theta) - g_i(\rho)}{\theta - \rho} \leq l_{i0}.$$

5) The class of g.l.c. and monotone increasing activation functions whose derivative has upper and lower bounds $\mathcal{G}$ [18]: for $\forall\theta, \rho \in \mathcal{R}$ and $\theta \neq \rho$, there exist constants $\bar{l}_i > \breve{l}_i > 0, (i = 1, 2, \cdots, n)$ such that

$$0 < \breve{l}_i \leq \frac{g_i(\theta) - g_i(\rho)}{\theta - \rho} \leq \bar{l}_i.$$

As Hu *et al.* pointed out in [17], the aforementioned activation function classes satisfy $\mathcal{GL} \subset \mathcal{PL}$ and $\mathcal{GL} \subset \mathcal{LL}$.

In [5], [17]–[23], [33], [36], and [37], researchers studied the AEST of neural networks with the different kinds of activation function classes mentioned previously. Zhao *et al.* [33] studied the exponential stability of delayed neural networks by assuming the activation functions belonging to $\mathcal{GL}$, where $l_i \equiv 1$. The AEST of delayed bidirectional associative memory neural networks with the activation function belonging to $\mathcal{GL}$ is investigated via the Lyapunov stability theory in [23]. In [18], a necessary and sufficient condition is established for ascertaining the AEST for a class of finite delayed neural networks with the activation function belonging to $\mathcal{G}$. In [22], the AEST of a class of delayed neural networks with the activation function belonging to $\mathcal{GL}$ is studied. In [17], AEST of a class of continuous-time recurrent neural network (RNN) with the activation function belonging to $\mathcal{LL}$ is studied.

In earlier papers (Chua *et al.* [7]–[9], Roska *et al.* [25]–[28], Civalleri *et al.* [29], and Arik *et al.* [1], [3]), the activation function for cellular neural networks is assumed to be the following saturation function:

$$f(x) = 0.5\left(|x+1| - |x-1|\right). \tag{1}$$

It is considered a widely employed piecewise linear neural network, where infinite intervals with zero slope are presented in activations.

As stated earlier, the selection of activation functions is very important in designing a neural network, especially for a neural network which is absolutely exponentially stable. Recently, the stability analysis of the systems with saturation nonlinearity was studied in [15], [16], and [46]–[48]. In [15] and [16], a generalized sector bounded by piecewise linear functions was introduced for the purpose of reducing conservatism in the absolute stability analysis of systems with nonlinearity and/or uncertainty. Inspired by the idea of a generalized sector condition, and following the previously mentioned activation function classes,

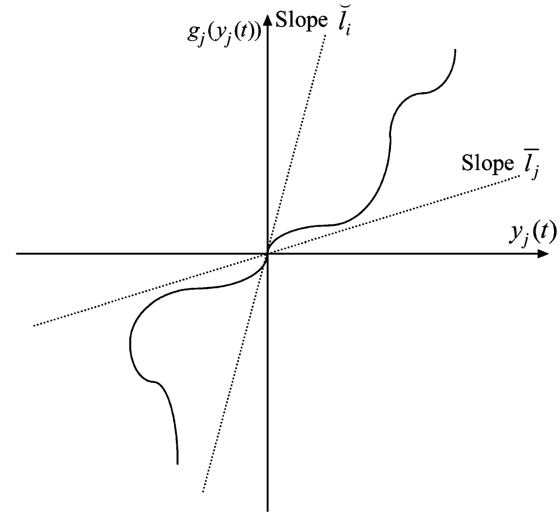

Fig. 1. Linear sector.

in this paper, we propose a new generalized activation function class for RNNs.

First, we will show the difference between our proposed generalized activation function class and the traditional activation class. For example, if we use the activation function class $\mathcal{G}$, consider the following recurrent neural networks:

$$\frac{dx(t)}{dt} = -Cx(t) + Ag\left(x(t)\right) + I \tag{2}$$

where $x(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T$ and $g(x(t)) = [g_1(x_1(t)), g_2(x_2(t)), \ldots, g_n(x_n(t))]^T$. Every component of the neuron's activation function $g(x(t))$ belongs to function class $\mathcal{G}$, respectively. Assume there is an equilibrium point x^e for the system. By coordinate transformation, we can obtain the following new system whose equilibrium point is at origin:

$$\frac{dy(t)}{dt} = -Cy(t) + Af\left(y(t)\right) \tag{3}$$

where $f(y(t)) = g(y(t)+x^e) - g(x^e)$ and $y(t) = x(t) - x^e$. For system (3), as the ith component $f_i(y_i(t))$ of $f(y(t))$ belongs to $\mathcal{G}$, we can obtain the inequalities

$$\begin{cases} 0 \leq y_j(t)\bar{l}_i \leq f_j\left(y_j(t)\right) \leq y_j(t)\breve{l}_i, & \text{as } y_j(t) \geq 0 \\ 0 \geq y_j(t)\bar{l}_i \geq f_j\left(y_j(t)\right) \geq y_j(t)\breve{l}_i, & \text{as } y_j(t) \leq 0. \end{cases} \tag{4}$$

As Fig. 1 shows, the ith component $f_i(y_i(t))$ of $f(y(t))$ belongs to a linear sector formed by two lines (the dot line) $y_j(t)\bar{l}_i$ and $y_j(t)\breve{l}_i$. Therefore, every component of the neuron's activation function of the neural network (3) belongs to a linear sector, respectively. All of these linear sectors are composed of the activation function class $\mathcal{G}$. The other activation function classes we mentioned earlier, such as $\mathcal{S}$, $\mathcal{GL}$, $\mathcal{LL}$, and $\mathcal{PL}$, have similar linear sector structures.

In the generalized activation function class proposed in this paper, we assume every component of the neuron's activation function in the neural network belongs to the convex hull, respectively. All of the convex hulls are composed of the generalized activation function class. In contrast to the slope of activation function bounded by two straight lines, such as $\mathcal{S}$, $\mathcal{G}$, $\mathcal{GL}$,

$\mathcal{LL}$, and $\mathcal{PL}$, we use two odd symmetric piecewise linear functions that are convex or concave over $\mathcal{R}$ to bound an (uncertain) activation function. The proposed generalized activation function class can describe the activation functions more flexibly and more specifically. Moreover, our new activation function class should include more activation functions, which the other paper could not have considered, such as piecewise linear activation function.

However, the flexible description of the activation function by the new function class makes the neural network model complex. This brings some difficulties in the next stability analysis. To overcome these difficulties, we will show that the equilibrium point of RNN with a generalized activation function class is GES if and only if the equilibrium point of RNN under every vertex activation function of the convex hull is GES. As the GES of the equilibrium point of two systems is equivalent, we will focus on the stability analysis of the RNN system under every vertex function of the convex hull. In what follows, the neural network under every vertex activation function is transformed into neural networks under an array of saturated linear activation functions. The existence of an equilibrium point of RNN with a saturated linear activation function is studied by Brouwer's fixed-point theorem. In the end, based on the M-matrix theory, the GES of the equilibrium point is studied. By the definition of absolute exponential stability, we can conclude AEST of RNN with the generalized activation function class.

This paper is organized as follows. In Section II, some preliminaries are given on convex combination, convex hull, convex (concave) function, exponential stability theorem and its converse, and the lemmas on Brouwer's fixed-point theorem and the M-matrix theory. In the model description, we describe in detail the new proposed generalized activation function class $\mathcal{K}$. As every component of the generalized activation function class belongs to the convex hull of two piecewise linear functions, we prove that the original RNN with generalized activation function class $\mathcal{K}$ is equivalent in GES to RNN under every vertex activation functions of the convex hull. The neural networks under all vertex activation functions are then transformed into neural networks under an array of saturated linear activation functions. In Section III, the main results of this paper are discussed. By Brouwer's fixed-point theorem, we first verify the existence of an equilibrium point. Then, using the theory of M-matrix, the GES of the RNN with a saturated linear activation function is studied. In the end, the AEST of the original RNN with the generalized activation function is obtained. In Section IV, a two-neuron RNN with a generalized activation function is constructed to show the effectiveness of our results. To demonstrate the advantage of our results, a comparison with the previous ones was drawn. In Section V, conclusions are drawn.

Notation: The following notation will be used throughout this paper: $\mathcal{R}$ denotes the set of real numbers; $\mathcal{R}^+$ denotes the set of nonnegative real numbers; and $\mathcal{R}^n$ denotes the n-dimensional Euclidean space. The notation $X \geq Y$ (respectively, $X > Y$), where X and Y are symmetric matrix, means that $X-Y$ is positive semidefinite (respectively, positive definite); $\|\cdot\|$ denotes the vector norm (or matrix norm) in finite-dimensional space; $\|\cdot\|_1$ and $\|\cdot\|_2$ denote L_1- and L_2-norm in finite-dimensional Euclidean space, respectively; and $\mathrm{co}\{x_1, x_2, \ldots, x_l\}$ denotes the

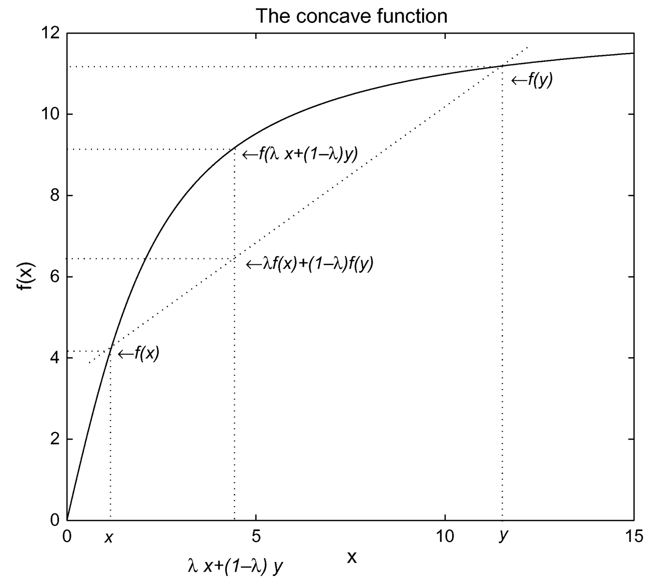

Fig. 2. Concave function.

convex hull of vectors $x_1, x_2, \ldots, x_l$, i.e., $\mathrm{co}\{x_1, x_2, \ldots, x_l\} := \{\sum_{i=1}^{l} \gamma_i x_i : \sum_{i=1}^{l} \gamma_i = 1, \gamma_i \geq 0\}$. For two integers k_1 and $k_2, k_1 < k_2$, we denote $I[k_1, k_2] = \{k_1, k_1+1, \ldots, k_2\}$. $\mathrm{sat}(x)$ denotes the saturation function, where $\mathrm{sat}(x) = 0.5(|x+1| - |x-1|)$.

II. Preliminaries and Model Description

A. Preliminaries

1) Convex Combination and Convex Hull: Let V be a vector space over $\mathcal{R}^n$. Let X be a set of elements of V. Then, a convex combination of elements from X is a linear combination of the form

$$\gamma_1 x_1 + \gamma_2 x_2 + \cdots + \gamma_n x_n$$

for some $n > 0$, where $x_i \in X$, $\gamma_i \geq 0$, $\Sigma_i \gamma_i = 1$, and $i = 1, 2, \cdots, n$.

Let $\mathrm{co}(X)$ be the set of all convex combinations from X. We call $\mathrm{co}(X)$ the convex hull of X.

2) Convex and Concave Functions: Suppose Ω is a convex set in a vector space over $\mathcal{R}$ (or $\mathcal{C}$), and suppose f is a function $f : \Omega \to \mathcal{R}$. Assume the following:

a) $f(x)$ is continuous, piecewise differentiable, $f(0) = 0$, and $(df(x)/dx)|_{x=0} > 0$;

b) $f(x)$ is odd symmetric, i.e., $f(-x) = -f(x)$.

If for any $x, y \in \Omega, x \neq y$ and any $\lambda \in (0,1)$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

we say that f is a convex function. If for any $x, y \in \Omega$ and any $\lambda \in (0,1)$

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda) f(y)$$

we say that f is a concave function (see Fig. 2). If either of the inequalities is strict, then we say that f is a strictly convex function, or a strictly concave function, respectively. Fig. 2 shows the concave function in the first quadrant.

The concave function has the following properties.

a) A function f is a concave function if and only if $-f$ is a convex function. For this reason, majority of the following discussion only focuses on the concave function. Similar results hold for the convex function.

b) On $\mathcal{R}$, a differentiable function is concave if and only if f' is monotone decreasing.

3) Exponential Stability Theorem and Its Converse: Consider the autonomous nonlinear system

$$\frac{dx(t)}{dt} = f(x) \tag{5}$$

where $f : \mathcal{R}^n \to \mathcal{R}^n$ is continuously differentiable and the Jacobian matrix $[\partial f/\partial x]$ is bounded on $\mathcal{R}^n$. Then, we have the following theorem.

Lemma 1 (Exponential Stability Theorem and Its Converse)[14], [30]: Let $x = 0$ be an equilibrium point for system (5) and let M and λ be positive constants. Then, the following two statements are equivalent.

a) The origin of system (5) is globally exponentially stable if for any $x_0 \in \mathcal{R}^n$ the trajectory of the system satisfies

$$\|x(t)\| \le M\|x_0\|e^{-\lambda t} \qquad \forall t \ge 0.$$

b) There exists a function $V : \mathcal{R}^n \to R$ satisfying the inequalities

$$\begin{aligned} c_1\|x\|^2 \le & V(x) \le c_2\|x\|^2 \\ \frac{\partial V}{\partial x} & f(x) \le -c_3\|x\|^2 \\ \left\|\frac{\partial V}{\partial x}\right\| & \le c_4\|x\| \end{aligned}$$

for some positive constants c_1, c_2, c_3, and c_4.

Proof: The proof can be seen in [14, pp. 163–165].

4) Brouwer Fixed-Point Theorem and M-Matrix Theory:

Lemma 2 (Brouwer Fixed-Point Theorem): Suppose that the continuous operator T maps closed bounded convex set $K \subset R^n$ onto itself, then the operator T has at least one fixed point in set K.

Lemma 3 [34], [38]: Let matrix $A = (a_{ij})_{n\times n}$ have nonpositive off-diagonal elements. Then, A is a nonsingular M-matrix if and only if one of the following conditions holds:

a) all principal minors of A are positive;

b) A is nonsingular and the elements of A^{-1} are all nonnegative;

c) A has all positive diagonal elements and there is a vector $x = (\xi_1, \xi_2, \ldots, \xi_n)^T$ (or y), whose elements are all positive, such that all the elements of Ax (or $A^T y$) are positive, namely

$$a_{ii}\xi_i > \sum_{j\ne i} |a_{ij}|\xi_j \left(a_{ii}\xi_i > \sum_{j\ne i} |a_{ji}|\xi_j \right), \qquad i = 1, 2, \ldots, n$$

which can be rewritten as

$$\sum_{j=1}^{n} a_{ij}\xi_j > 0 \left(\sum_{j=1}^{n} a_{ji}\xi_j > 0 \right), \qquad i = 1, 2, \ldots, n;$$

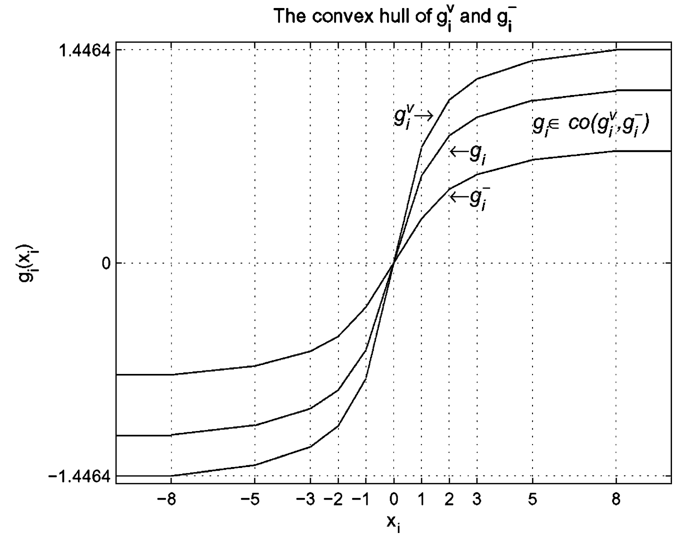

Fig. 3. Convex hull of g_i^v and g_i^-.

d) the real parts of all the eigenvalues of A are positive;

e) there is a diagonal matrix, $P = \mathrm{diag}(p_1, p_2, \ldots, p_n)$, with $p_i > 0$, $i = 1, 2, \ldots, n$, such that $PA + A^T P$ is a positive–definite matrix.

B. Model Description

The RNN with a generalized activation function is described by the following differential equation:

$$\frac{dx(t)}{dt} = -Cx(t) + Ag\left(x(t)\right) + I \tag{6}$$

where $x(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T$ and $g(x(t)) = [g_1(x_1(t)), g_2(x_2(t)), \ldots, g_n(x_n(t))]^T$. C, A, and I are constant matrices. $A = \{a_{ij}\}_{n\times n}$ denotes the feedback matrix. $C = \mathrm{diag}(c_1, c_2, \ldots, c_n)$ denotes the self-feedback matrix, which is positive–definite diagonal matrix. $I = [I_1, I_2, \ldots, I_n]^T$ denotes the external bias. For simplicity, we denote $g_i = g_i(x_i(t))$ and $g = g(x(t))$. In system (6), we assume the activation function g_i, the ith component of g, belongs to the convex hull of $\bar{g}_i$ and $\breve{g}_i$, where $\bar{g}_i$ and $\breve{g}_i$ are the convex or concave piecewise continuous differentiable linear functions, respectively. That is $g_i \in \mathrm{co}\{\bar{g}_i, \breve{g}_i\}$. Then, the activation function class for g is the convex hull of 2^n decoupled functions

$$g^l = \left[g_1^l, g_2^l, \ldots, g_n^l\right]^T, \qquad l \in I[1, 2^n] \tag{7}$$

where $g_i^l = \bar{g}_i$ or $\breve{g}_i$. Here, we denote this activation function class as $\mathrm{co}\{g^l : l \in I[1, 2^n]\}$, which is composed of all the convex hulls $\mathrm{co}\{\bar{g}_i, \breve{g}_i\}$, where $i = 1, 2, \ldots, n$. Fig. 3 shows that an activation function g_i belongs to the convex hull $\mathrm{co}\{g_i^v, g_i^-\}$ which is bounded by a pair of continuously piecewise linear concave functions g_i^v and g_i^-, where $g_i = 0.6g_i^v + 0.4g_i^-$. The convex hull lies between functions g_i^v and g_i^-.

To sum up, we denote the following function class $\mathcal{K}$.

Definition 1 (Function Class $\mathcal{K}$): We say that $g \in \mathcal{K}$ if and only if $g \in \mathrm{co}\{g^l : l \in I[1, 2^n]\}$, where $g^l = [g_1^l, g_2^l, \ldots, g_n^l]^T$ and $g_i^l = \bar{g}_i$ or $\breve{g}_i$, and g_i, the ith component of g, belongs to the

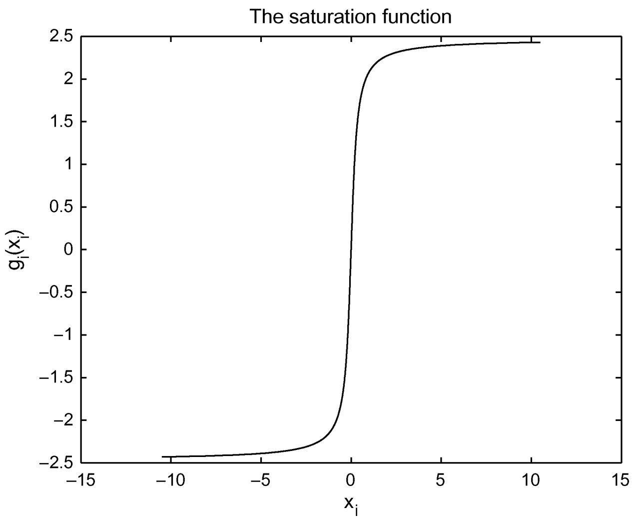

Fig. 4. Saturation function $g_i(x_i)$.

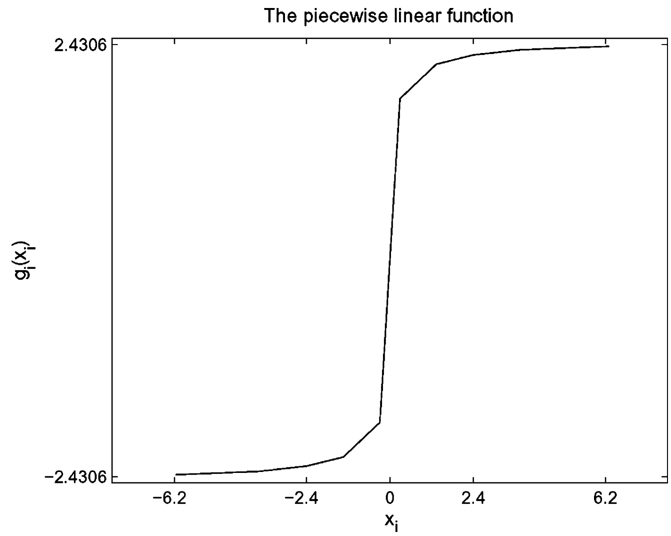

Fig. 5. Piecewise linear function $g_i'(x_i)$.

convex hull of $\bar{g}_i$ and $\breve{g}_i$, where $\bar{g}_i$ and $\breve{g}_i$ are a pair of continuously piecewise differentiable convex/concave linear functions. The convex hull is bounded by them, i.e.,

$$g_i \in \mathrm{co}\{\bar{g}_i, \breve{g}_i\}, \qquad i = 1, 2, \ldots, n. \tag{8}$$

Remark 1: As discussed earlier, the activation function class $\mathcal{K}$ is a more flexible and more specific description of the activation function than some function classes. As $\bar{g}_i = \breve{g}_i$, the activation function of neural networks (6) becomes a deterministic piecewise linear function. That is, $g_i = \breve{g}_i$ or $\bar{g}_i$. Because a smooth nonlinear function can be approximated with a piecewise linear function, we can use a piecewise linear function to approximate well-known activation functions such as the sigmoidal activation function.

Figs. 4–6 show how a saturation function $g_i(x_i) = (\pi/2)\arctan(4x_i)$ is approximated by a piecewise linear function. Fig. 4 is a saturation function $g_i(x_i) = (\pi/2)\arctan(4x_i)$. Fig. 5 is the piecewise linear function $g_i'(x_i)$, where $g_i'(x_i)$ is shown in (9) at the bottom of the page.

In Fig. 6, we see that the saturation function $g_i(x_i)$ is well approximated by the piecewise linear function $g_i'(x_i)$.

With the definition of the function class $\mathcal{K}$, we can give the definition of AEST.

$$g_i'(x_i) = \begin{cases} -\left\{\frac{\pi}{2}\arctan(32) - \frac{16\pi}{1025}\right\} + \frac{2\pi}{1025}x_i, & \text{if } x_i \in (-\infty, -6.1546] \\ -\left\{\frac{\pi}{2}\arctan(20) - \frac{10\pi}{401}\right\} + \frac{2\pi}{401}x_i, & \text{if } x_i \in [-6.1546, -3.7514] \\ -\left\{\frac{\pi}{2}\arctan(12) - \frac{6\pi}{145}\right\} + \frac{2\pi}{145}x_i, & \text{if } x_i \in [-3.7514, -2.4014] \\ -\left\{\frac{\pi}{2}\arctan(8) - \frac{4\pi}{65}\right\} + \frac{2\pi}{65}x_i, & \text{if } x_i \in [-2.4014, -1.3400] \\ -\left\{\frac{\pi}{2}\arctan(4) - \frac{2\pi}{17}\right\} + \frac{2\pi}{17}x_i, & \text{if } x_i \in [-1.3400, -0.2897] \\ 2\pi x_i, & \text{if} x_i \in [-0.2897, 0.2897] \\ \left\{\frac{\pi}{2}\arctan(4) - \frac{2\pi}{17}\right\} + \frac{2\pi}{17}x_i, & \text{if } x_i \in [0.2897, 1.3400] \\ \left\{\frac{\pi}{2}\arctan(8) - \frac{4\pi}{65}\right\} + \frac{2\pi}{65}x_i, & \text{if } x_i \in [1.3400, 2.4014] \\ \left\{\frac{\pi}{2}\arctan(12) - \frac{6\pi}{145}\right\} + \frac{2\pi}{145}x_i, & \text{if } x_i \in [2.4014, 3.7514] \\ \left\{\frac{\pi}{2}\arctan(20) - \frac{10\pi}{401}\right\} + \frac{2\pi}{401}x_i, & \text{if } x_i \in [3.7514, 6.1546] \\ \left\{\frac{\pi}{2}\arctan(32) - \frac{16\pi}{1025}\right\} + \frac{2\pi}{1025}x_i, & \text{if } x_i \in [6.1546, +\infty). \end{cases} \tag{9}$$

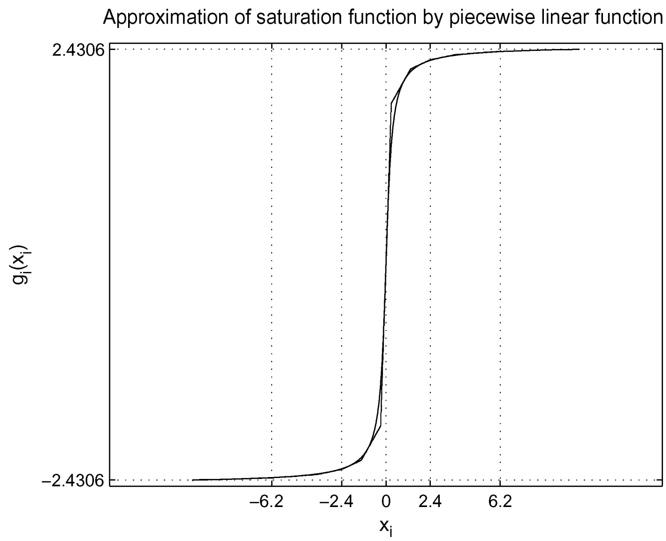

Fig. 6. Approximation of saturation nonlinear function $g_i(x_i)$ by piecewise linear function $g_i'(x_i)$.

Definition 2: A neural network (6) is said to be AEST with respect to function class $\mathcal{K}$ if it possesses a GES equilibrium point for every activation function $g \in \mathcal{K}$ and for every input vector $I \in \mathcal{R}^n$.

Next, we will prove an important lemma of this paper, which establishes the equivalence of the original RNN with the generalized activation function (6) and RNN under every vertex function of convex hull (10) regarding GES.

Lemma 4: Assume Lyapunov function $V(x)$ is differentiable and radially unbounded. Given a convex set $\mathrm{co}\{g^l : l = 1, 2, \ldots, 2^n\}$, the equilibrium point of the neural network system (6) is GES if and only if the equilibrium point of the following system is GES

$$\frac{dx(t)}{dt} = -Cx(t) + Ag^l\left(x(t)\right) + I \tag{10}$$

for every $l \in I[1, 2^n]$, where $g^l(x(t)) = [g_1^l(x_1(t)), g_2^l(x_2(t)), \ldots, g_n^l(x_n(t))]^T$ and $g_i^l(x_i(t)) = \bar{g}_i(x_i(t))$ or $\breve{g}_i(x_i(t))$, $i = 1, 2, \ldots, n$. Here, we define system (10) as the RNN under every vertex activation function of the convex hull.

Proof: The proof of the theorem is given in the Appendix.

As the GES of the equilibrium point of the previous two systems are equivalent by Lemma 4, in the following study, we will focus our attention on stability analysis of every individual system (10).

In this paper, as a class of piecewise linear functions, we define the nonlinear function $\bar{g}_i$ ($\breve{g}_i$ are similar) as

$$\bar{g}_i(x_i) = \begin{cases} \bar{k}_{i\bar{N}_i}x_i - \bar{e}_{i\bar{N}_i}, & \text{if } x_i \in \left(-\infty, -\bar{b}_{i\bar{N}_i}\right] \\ \vdots & \\ \bar{k}_{i1}x_i - \bar{e}_{i1}, & \text{if } x_i \in \left[-\bar{b}_{i2}, -\bar{b}_{i1}\right] \\ \bar{k}_{i0}x_i, & \text{if } x_i \in \left[-\bar{b}_{i1}, \bar{b}_{i1}\right] \\ \bar{k}_{i1}x_i + \bar{e}_{i1}, & \text{if } x_i \in \left[\bar{b}_{i1}, \bar{b}_{i2}\right] \\ \vdots & \\ \bar{k}_{i\bar{N}_i}x_i + \bar{e}_{i\bar{N}_i}, & \text{if } x_i \in \left[\bar{b}_{i\bar{N}_i}, +\infty\right) \end{cases} \tag{11}$$

where $0 < \bar{b}_{i_1} < \bar{b}_{i_2} < \cdots < \bar{b}_{i_{\bar{N}_i}}$. By the properties of the concave function in Section II-A2, we see that if $\bar{g}_i$ is concave, then $\bar{k}_{i_0} > \bar{k}_{i_1} > \cdots > \bar{k}_{i_{\bar{N}_i}} > -\infty$ and $0 < \bar{e}_{i_1} < \bar{e}_{i_2} < \cdots < \bar{e}_{i_{\bar{N}_i}} < \infty$. If $\bar{g}_i$ is convex, then $\bar{k}_{i_0} < \bar{k}_{i_1} < \cdots < \bar{k}_{i_{\bar{N}_i}} < \infty$ and $0 > \bar{e}_{i_1} > \bar{e}_{i_2} > \cdots > \bar{e}_{i_{\bar{N}_i}}$.

This means that

$$\bar{v}_{i\bar{h}_i} = \frac{\bar{k}_{i0} - \bar{k}_{i\bar{h}_i}}{\bar{e}_{i\bar{h}_i}} > 0, \qquad \bar{h}_i \in I\left[1, \bar{N}_i\right], \qquad i \in I[1, n].$$

The $\bar{b}_{i_{\bar{h}_i}}$ can be determined from $\bar{e}_{i_{\bar{h}_i}}$ and $\bar{k}_{i_{\bar{h}_i}}$ by the continuity of the function

$$\bar{b}_{i\bar{h}_i} = -\frac{\bar{e}_{i\bar{h}_i} - \bar{e}_{i\bar{h}_i-1}}{\bar{k}_{i\bar{h}_i} - \bar{k}_{i\bar{h}_i-1}}, \qquad \bar{h}_i \in I[1, \bar{N}_i], \qquad i \in I[1, n].$$

As an example of piecewise linear function (11), Fig. 7 shows a piecewise linear concave function with four bends and Fig. 3 shows three piecewise linear concave functions with six bends.

By some simple transformations, equations in (11) are equivalent to the following equations with respect to i:

$$\begin{aligned} \bar{g}_i(x_i) &= \bar{k}_{i\bar{h}_i}x_i + \bar{e}_{i\bar{h}_i}\mathrm{sat}\left(\frac{\bar{k}_{i0} - \bar{k}_{i\bar{h}_i}}{\bar{e}_{i\bar{h}_i}}x_i\right) \\ &= \bar{k}_{i\bar{h}_i}x_i + \bar{e}_{i\bar{h}_i}\mathrm{sat}\left(\bar{v}_{i\bar{h}_i}x_i\right). \end{aligned} \tag{12}$$

In order to use a new set of vertex functions (12) to replace the piecewise linear vertex functions (11), we cite the following useful lemma from [16].

Lemma 5 [16]: Consider the piecewise linear concave/convex function

$$\bar{g}_i(x_i) = \begin{cases} \bar{k}_{i\bar{N}_i}x_i - \bar{e}_{i\bar{N}_i}, & \text{if } x_i \in \left(-\infty, \bar{b}_{i\bar{N}_i}\right] \\ \vdots & \\ \bar{k}_{i1}x_i - \bar{e}_{i1}, & \text{if } x_i \in \left[-\bar{b}_{i2}, -\bar{b}_{i1}\right] \\ \bar{k}_{i0}x_i, & \text{if } x_i \in \left[-\bar{b}_{i1}, \bar{b}_{i1}\right] \\ \bar{k}_{i1}x_i + \bar{e}_{i1}, & \text{if } x_i \in \left[\bar{b}_{i1}, \bar{b}_{i2}\right] \\ \vdots & \\ \bar{k}_{i\bar{N}_i}x_i + \bar{e}_{i\bar{N}_i}, & \text{if } x_i \in \left[\bar{b}_{i\bar{N}_i}, +\infty\right) \end{cases} \tag{13}$$

and

$$\breve{g}_i(x_i) = \begin{cases} \breve{k}_{i\breve{N}_i}x_i - \breve{e}_{i\breve{N}_i}, & \text{if } x_i \in \left(-\infty, \breve{b}_{i\breve{N}_i}\right] \\ \vdots & \\ \breve{k}_{i1}x_i - \breve{e}_{i1}, & \text{if } x_i \in \left[-\breve{b}_{i2}, -\breve{b}_{i1}\right] \\ \breve{k}_{i0}x_i, & \text{if } x_i \in \left[-\breve{b}_{i1}, \breve{b}_{i1}\right] \\ \breve{k}_{i1}x_i + \breve{e}_{i1}, & \text{if } x_i \in \left[\breve{b}_{i1}, \breve{b}_{i2}\right] \\ \vdots & \\ \breve{k}_{i\breve{N}_i}x_i + \breve{e}_{i\breve{N}_i}, & \text{if } x_i \in \left[\breve{b}_{i\breve{N}_i}, +\infty\right). \end{cases} \tag{14}$$

For $\bar{h}_i \in I[1, \bar{N}_i]$, $\breve{h}_i \in I[1, \breve{N}_i]$, and $i \in I[1, n]$, define

$$\bar{g}_i^{\bar{h}_i}(x_i) = \bar{k}_{i\bar{h}_i}x_i + \bar{e}_{i\bar{h}_i}\mathrm{sat}\left(\bar{v}_{i\bar{h}_i}x_i\right) \tag{15}$$

and

$$\breve{g}_i^{\breve{h}_i}(x_i) = \breve{k}_{i\breve{h}_i}x_i + \breve{e}_{i\breve{h}_i}\mathrm{sat}\left(\breve{v}_{i\breve{h}_i}x_i\right). \tag{16}$$

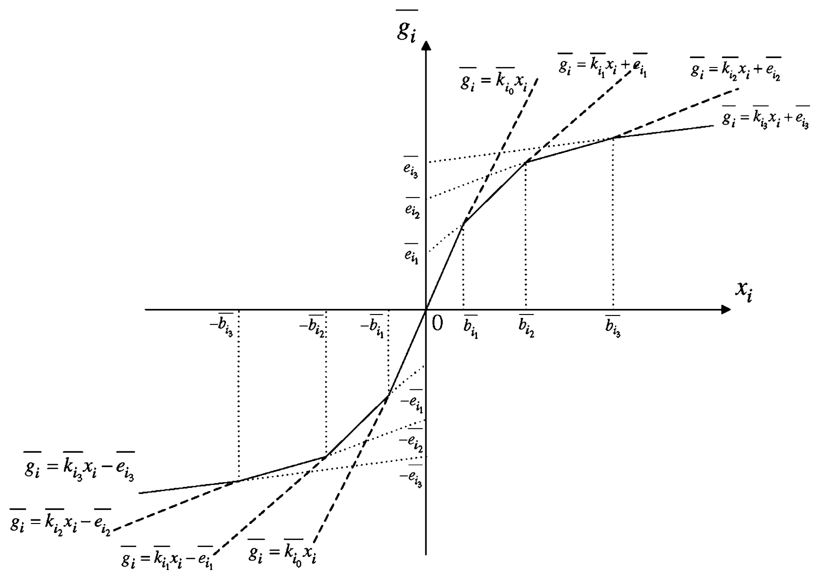

Fig. 7. Piecewise linear concave function with four bends.

Then

$$\mathrm{co}\{\bar{g}_i, \breve{g}_i\} \subset \mathrm{co}\left\{\bar{g}_i^{\bar{h}_i}, \breve{g}_i^{\breve{h}_i} : \bar{h}_i \in I[1, \bar{N}_i], \breve{h}_i \in I[1, \breve{N}_i], i \in I[1, n]\right\}.$$

Proof: The proof can refer to [16]. Thus, it is omitted here.

By Definition 1 and Lemma 5, we have

$$g_i \in \mathrm{co}\{\bar{g}_i, \breve{g}_i\} \subset \mathrm{co}\left\{\bar{g}_i^{\bar{h}_i}, \breve{g}_i^{\breve{h}_i} : \bar{h}_i \in I[1, \bar{N}_i], \breve{h}_i \in I[1, \breve{N}_i], i \in I[1, n]\right\}.$$

Thus, it is easier to see that

$$g_i \in \mathrm{co}\left\{\bar{g}_i^{\bar{h}_i}, \breve{g}_i^{\breve{h}_i} : \bar{h}_i \in I[1, \bar{N}_i], \breve{h}_i \in I[1, \breve{N}_i], i \in I[1, n]\right\}.$$

Remark 2: In [12], Forti uses piecewise linear functions with two bends, which is the special case of (13) or (14), to approximate Sigmoidal function $\mathcal{S}$. As we discussed in Remark 1, by choosing $\bar{g}_i = \breve{g}_i$, the activation function class $\mathcal{K}$ includes the activation function class used in [12] as a special case.

Remark 3: In [1], [3], [7]–[9], and [25]–[29], the activation function for cellular neural networks is assumed to be saturation function (1), which is a piecewise linear concave function with two bends. It is obvious that function (1) is the special case of (13) or (14). Again, by denoting $\bar{g}_i = \breve{g}_i$ in Definition 1, the activation function class $\mathcal{K}$ includes the function (1) as a special case.

Again, similar to Lemma 4, we first assume that the GES of the equilibrium point of every system (10) is equivalent to that of the neural networks under every vertex activations $\bar{g}_i^{\bar{h}_i}$ or $\breve{g}_i^{\breve{h}_i}$, $\bar{h}_i \in I[1, \bar{N}_i]$, $\breve{h}_i \in I[1, \breve{N}_i]$, $i \in I[1, n]$ (this assumption will be proven later in Lemma 6). Therefore, in the following study, we focus our stability analysis on system (6) under piecewise activation functions $\bar{g}_i^{\bar{h}_i}$ or $\breve{g}_i^{\breve{h}_i}$. For simplicity, we denote $g_i^{h_i} = \bar{g}_i^{\bar{h}_i}$ or $\breve{g}_i^{\breve{h}_i}$.

As

$$g_i^{h_i}(x_i) := k_{ih_i} x_i + e_{ih_i} \mathrm{sat}\left(v_{ih_i} x_i\right), \quad h_i \in I[1, N_i], \quad i \in I[1, n] \tag{17}$$

then

$$g^h = \left[g_1^{h_1}, g_2^{h_2}, \ldots, g_n^{h_n}\right], \quad h_i \in I[1, N_i], \quad i \in I[1, n]$$

has $H := N_1 \times N_2 \times \cdots \times N_n$ vertices. If we assign a number $p \in I[1, H]$ to each vertex function corresponding to $(h_1, h_2, \cdots, h_n)$, then each of the vertices has the form

$$\begin{aligned} g^h &= \left[g_1^{h_1}(x_1), g_2^{h_2}(x_2), \cdots, g_n^{h_n}(x_n)\right]^T \\ &= K_p(t)x(t) + E_p \mathrm{sat}\left(V_p x(t)\right) \end{aligned} \tag{18}$$

where $p \in I[1, H]$ and

$$\begin{aligned} K_p &= \mathrm{diag}\left(k1_{h_1}^p, k2_{h_2}^p, \cdots, kn_{h_n}^p\right) \\ E_p &= \mathrm{diag}\left(e1_{h_1}^p, e2_{h_2}^p, \cdots, en_{h_n}^p\right) \\ V_p &= \mathrm{diag}\left(v1_{h_1}^p, v2_{h_2}^p, \cdots, vn_{h_n}^p\right). \end{aligned}$$

Noting that

$$g_i \in \mathrm{co}\left\{\bar{g}_i^{\bar{h}_i}, \breve{g}_i^{\breve{h}_i} : \bar{h}_i \in I[1, \bar{N}_i], \breve{h}_i \in I[1, \breve{N}_i], i \in I[1, n]\right\}$$

thus

$$g \in \mathrm{co}\left\{\left[g_1^{h_1}, g_2^{h_2}, \ldots, g_n^{h_n}\right], : h_i \in I[1, N_i], i \in I[1, n]\right\}$$

where $g_i^{h_i} = \bar{g}_i^{\bar{h}_i}$ or $\breve{g}_i^{\breve{h}_i}$.

Then, like Lemma 4, we can prove the following Lemma.

Lemma 6: Let $H := N_1 \times N_2 \times \cdots \times N_n$ and $g_i^{h_i} = \bar{g}_i^{\bar{h}_i}$ or $\breve{g}_i^{\breve{h}_i}$. For convex sets $\mathrm{co}\{[g_1^{h_1}, g_2^{h_2}, \ldots, g_n^{h_n}] : h_i \in I[1, N_i], i \in I[1, n]\}$, assign a number $p \in I[1, H]$ to each vertex function $g_i^{h_i}$ corresponding to $(h_1, h_2, \cdots, h_n)$, then the equilibrium point of every neural network system (6) is GES if and only if each of the equilibrium points of the following systems is GES:

$$\frac{dx(t)}{dt} = -\left(C - AK_p(t)\right)x(t) + E_p\mathrm{sat}\left(V_p x(t)\right) + I \qquad \forall p \in I[1, H] \quad (19)$$

where

$$K_p = \mathrm{diag}\left(k1^p_{h_1}, k2^p_{h_2}, \cdots, kn^p_{h_n}\right)$$
$$E_p = \mathrm{diag}\left(e1^p_{h_1}, e2^p_{h_2}, \cdots, en^p_{h_n}\right)$$
$$V_p = \mathrm{diag}\left(v1^p_{h_1}, v2^p_{h_2}, \cdots, vn^p_{h_n}\right).$$

Proof: With (18), we have

$$\begin{aligned}\frac{dx(t)}{dt} &= -Cx(t) + Ag^h\left(x(t)\right) + I,\\ &= -\left(C - AK_p(t)\right)x(t) + E_p\mathrm{sat}\left(V_p x(t)\right) + I.\end{aligned}$$

We can then prove this lemma similarly to the proof of Lemma 4.

From Lemma 6, we can see the equivalence of system (6) and the neural networks under all vertex activation functions [system (19)] regarding GES. In the following, we will focus our attention on the GES of the equilibrium point of every system (19).

III. Main Results

A. Existence of Equilibrium Point

Theorem 1: For every $p \in I[1, H]$, every system (19) has an equilibrium point if the matrix $C_p = C - AK_p$ is invertible, respectively.

Proof: This theorem can be proven by the well-known Brouwer's fixed-point theorem (Lemma 2). A vector $x^e_p = (x^e_{1_p}, x^e_{2_p}, \ldots, x^e_{n_p})^T$ is an equilibrium point of system (19) if

$$-\left(C - AK_p(t)\right)x^e_p + E_p\mathrm{sat}\left(V_p \cdot x^e_p\right) + I_p = 0, \qquad p \in I[1, H]. \quad (20)$$

For any $p \in I[1, H]$, as C_p is invertible, then C_p^{-1} exists. In view of (20), we denote a mapping $T_p(x(t))$ ($T_p : \mathcal{R}^n \to \mathcal{R}^n$) for $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T \in \mathcal{R}^n$, where $T_p(x(t)) = (T_{p_1}(x_1(t)), T_{p_2}(x_2(t)), \ldots, T_{p_n}(x_n(t)))^T$ with

$$T_p\left(x(t)\right) = {C_p}^{-1}\left(E_p\mathrm{sat}\left(V_p \cdot x(t)\right) + I_p\right).$$

Then, there exists a constant matrix R_p such that

$$\begin{aligned}\|T_p\left(x(t)\right)\| &= \left\|{C_p}^{-1}\left(E_p\mathrm{sat}\left(V_p \cdot x(t)\right) + I_p\right)\right\|\\ &\le \left\|{C_p}^{-1}\right\|\left(\|E_p\|\,\|\mathrm{sat}\left(V_p \cdot x(t)\right)\| + \|I_p\|\right)\\ &\le \left\|{C_p}^{-1}\right\|\left(\|E_p\| + \|I_p\|\right)\\ &\le \|R_p\|.\end{aligned}$$

We define a sphere

$$K_p = \left\{x(t) \in \mathcal{R}^n : \|x(t)\| \le \|R_p\|\right\}.$$

One can see that

$$x(t) \in K_p \Rightarrow T_p\left(x(t)\right) \in K_p.$$

From the continuity of the nonlinear activation functions $\mathrm{sat}(\cdot)$, it follows that the mapping $T_p : K_p \to K_p$ is continuous. By Lemma 2, there exists a fixed point $x^e_p \in K_p$ of the mapping. Therefore, for every $p \in I[1, H]$, the fixed point x^e_p is the equilibrium point of every system (19), respectively. The proof is completed.

Remark 4: As Definition 1 (function class $\mathcal{K}$) shows, the activation function in this paper is allowed to take an arbitrarily large positive value and is also allowed to have infinite intervals with zero slops. The existence of an equilibrium point is guaranteed if the conditions of Theorem 1 hold.

By Theorem 1, we assume that, for any $p \in I[1, H]$, each system (19) has an equilibrium point $x^e_p = (x^e_{1p}, x^e_{2p}, \ldots, x^e_{np})^T$. In order to simplify our proofs, we will shift the equilibrium point to the origin using the transformation $y_i(t) = x_i(t) - x^e_{ip}$. Then, every system (19) can be transformed into the form

$$\begin{aligned}\frac{dy(t)}{dt} &= (-C + AK_p)y(t) + E_p\left[\mathrm{sat}\left(V_p x(t)\right) - \mathrm{sat}\left(V_p x^e_p\right)\right]\\ &= (-C + AK_p)y(t) + E_p g\left(V_p y(t)\right) \qquad (21)\end{aligned}$$

where

$$g\left(V_p y(t)\right) = \Big(g_1\left(v1^p_{h_1} y_1(t)\right), g_2\left(v2^p_{h_2} y_2(t)\right), \ldots, g_n\left(vn^p_{h_n} y_n(t)\right)\Big)^T$$

and

$$g_i\left(vi^p_{h_i} y_i(t)\right) = \mathrm{sat}\left(vi^p_{h_i} x_i(t)\right) - \mathrm{sat}\left(vi^p_{h_i} x^e_{ip}\right). \quad (22)$$

By (22), we have

$$\mathrm{sgn}\left(y_i(t)\right) g_i\left(vi^p_{h_i} y_i(t)\right) \le vi^p_{h_i}\left|y_i(t)\right|. \quad (23)$$

Equations in (21) are equivalent to the following equations:

$$\frac{dy_i(t)}{dt} = -c_i y_i(t) + \sum_{j=1}^{n} a_{ij} kj^p_{h_j} y_j(t) + ei^p_{h_i} g_i\left(vi^p_{h_i} y_i(t)\right) \quad (24)$$

where $p \in I[1, H], H := N_1 \times N_2 \times \cdots \times N_n, h_i \in I[1, N_i], i \in I[1, n]$.

The stability behavior of the trivial solution $x = x^e_p$ of each system (19) is equivalent to that of the origin of the systems (21), respectively. Thus, in the following study, we only focus attention on the stability behavior of the origin of every system (21).

B. Absolute Stability of Equilibrium Point

Theorem 2: For each system (21), if the matrix $C_p = C - AK_p, p \in I[1, H]$ is invertible and the matrix $\Pi_p = C - E_p V_p - |A| K_p$ is M-matrix, then the origin of each system (21) is GES.

Proof: As $\Pi_p = C - E_pV_p - |A|K_p$ are M-matrix, by using c) in Lemma 3, there exist constant numbers $\xi_i > 0 (i = 1, 2, \ldots, n)$ such that the following inequalities hold:

$$\left(c_i - e^p_{ih_i} v^p_{ih_i}\right)\xi_i - \Sigma_{j=1}^n |a_{ji}| k^p_{ih_i}\xi_j > 0.$$

Moreover, it is easy to prove that there exist constants λ_i such that the previous inequality still holds, namely

$$\left[\left(c_i - e^p_{ih_i} v^p_{ih_i}\right) - \lambda_i\right]\xi_i - \Sigma_{j=1}^n |a_{ji}| k^p_{ih_i}\xi_j > 0. \tag{25}$$

For each $p \in I[1, H]$, construct the radially unbounded Lyapunov function

$$V(y(t)) = \sum_{i=1}^n e^{\lambda_i t}\xi_i |y_i(t)| \tag{26}$$

In view of inequality (23), the derivative of $V(y(t))$ along the trajectories of the system (24) is given by

$$\begin{aligned}
\dot{V}(y(t)) = &\sum_{i=1}^n e^{\lambda_i t}\Big\{\operatorname{sgn}(y_i(t))\,\xi_i\big[-c_i y_i(t) + \Sigma_{j=1}^n a_{ij} k^p_{jh_j} y_j(t) \\
&\qquad + e^p_{ih_i} g_i\left(v^p_{ih_i} y_i(t)\right)\big]\Big\} \\
&+ \sum_{i=1}^n \lambda_i e^{\lambda_i t}\xi_i |y_i(t)| \\
\le &\sum_{i=1}^n \xi_i e^{\lambda_i t}\Big\{-c_i |y_i(t)| + \Sigma_{j=1}^n |a_{ij}| k^p_{jh_j} |y_j(t)| \\
&\qquad + e^p_{ih_i} v^p_{ih_i} |y_i(t)|\Big\} \\
&+ \sum_{i=1}^n \xi_i \lambda_i e^{\lambda_i t} |y_i(t)| \\
= &\sum_{i=1}^n e^{\lambda_i t}\left\{-c_i\xi_i + \Sigma_{j=1}^n |a_{ji}| k^p_{ih_i}\xi_j + e^p_{ih_i} v^p_{ih_i}\xi_i\right\} \\
&\times |y_i(t)| + \sum_{i=1}^n \xi_i \lambda_i e^{\lambda_i t} |y_i(t)| \\
= &\sum_{i=1}^n -e^{\lambda_i t}\Big\{\left[\left(c_i - e^p_{ih_i} v^p_{ih_i}\right) - \lambda_i\right]\xi_i \\
&\qquad - \Sigma_{j=1}^n |a_{ji}| k^p_{ih_i}\xi_j\Big\} |y_i(t)|.
\end{aligned}$$

Hence, with inequality (25), we have

$$\dot{V}(y(t)) < 0.$$

By integrating the previous inequality from 0 to $t (> 0)$, we have

$$\sum_{i=1}^n e^{\lambda_i t} |y_i(t)| < \sum_{i=1}^n |y_i(0)|$$

or

$$\|y(t)\|_1 = \sum_{i=1}^n |y_i(t)| < \sum_{i=1}^n |y_i(0)|\, e^{-\lambda_i t} \le \|y(0)\|_1\, e^{-\lambda' t}$$

where $\lambda' = min_i(\lambda_i)$, $i = 1, 2, \ldots, n$. By a) in Lemma 1, the origin of every system (21) is GES. The proof is completed.

Theorem 3: If the conditions in Theorem 2 hold, then the trivial solution $x = x^e$ of system (6) is AEST.

Proof: If the origin of every system (21) is GES, then the trivial solution $x = x^e_p$ of every system (19) is GES. Thus, the trivial solution $x = x^e$ of system (6) is GES with Lemma 6. By Definition 2, we can conclude that the trivial solution $x = x^e$ of system (6) is AEST. The proof is completed.

IV. Comparison and Example

In this section, we will construct an example to show the effectiveness of our results. Then, we will compare our results with the previous ones.

A. Example and Simulation

Example: Consider a neural network with two neurons, where every component of the activation functions belong to the convex hull of a pair of piecewise continuous linear functions, which is described by the following differential equation:

$$\frac{dx(t)}{dt} = -Cx(t) + Ag(x(t)) + I \tag{27}$$

where

$$C = \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix}$$
$$A = \begin{pmatrix} 2 & 8 \\ 2 & -5 \end{pmatrix}$$
$$I = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Every component of $g(x)$ belongs to a convex hull which is bounded by a pair of piecewise linear functions, namely, $g_1(x_1) \in \operatorname{co}(\bar{g}_1(x_1), x_1)$ and $g_2(x_2) \in \operatorname{co}(\bar{g}_2(x_2), x_2)$ (see Figs. 8 and 9). The functions $\bar{g}_1(x_1)$ and $\bar{g}_2(x_2)$ are defined by

$$\bar{g}_1(x_1) = \begin{cases} -1.4281, & \text{if } x_1 \in [-\infty, -3] \\ 0.3218x_1 - 0.4627, & \text{if } x_1 \in [-3, -1] \\ 0.7845x_1, & \text{if } x_1 \in [-1, 1] \\ 0.3218x_1 + 0.4627, & \text{if } x_1 \in [1, 3] \\ 1.4281, & \text{if } x_1 \in [3, +\infty] \end{cases} \tag{28}$$

$$\bar{g}_2(x_2) = \begin{cases} -1.8528, & \text{if } x_1 \in [-\infty, -4] \\ 0.1419x_2 - 1.2852, & \text{if } x_2 \in [-4, -2] \\ 0.7845x_2, & \text{if } x_2 \in [-2, 2] \\ 0.1419x_1 + 1.2852, & \text{if } x_2 \in [2, 4] \\ 1.8528, & \text{if } x_1 \in [4, +\infty]. \end{cases} \tag{29}$$

The piecewise linear function (28) has three bounds with $\bar{N}_1 = 2$ and

$$(\bar{k_{10}}, \bar{k_{11}}, \bar{k_{12}}) = (0.7845, 0.3218, 0) \tag{30}$$
$$(\bar{e_{10}}, \bar{e_{11}}, \bar{e_{12}}) = (0, 0.4627, 1.4281) \tag{31}$$
$$(\bar{b_{10}}, \bar{b_{11}}, \bar{b_{12}}) = (0, 1, 3) \tag{32}$$
$$(\bar{v_{11}}, \bar{v_{12}}) = (1, 0.5493). \tag{33}$$

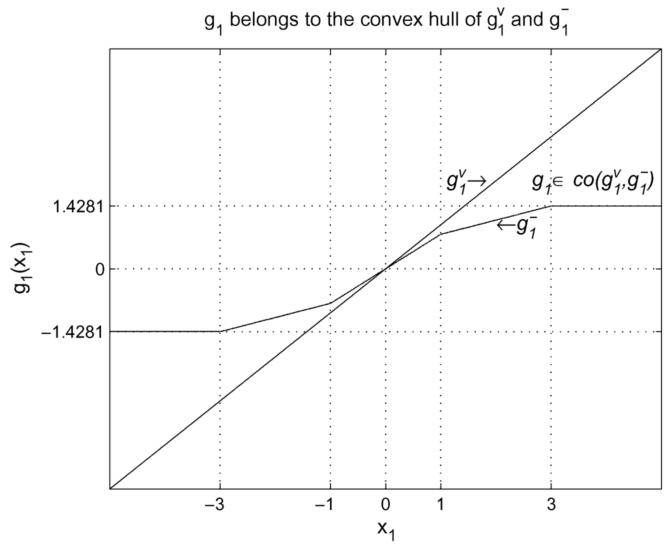

Fig. 8. The g_1 belongs to the convex hull of g_1^v and g_1^-.

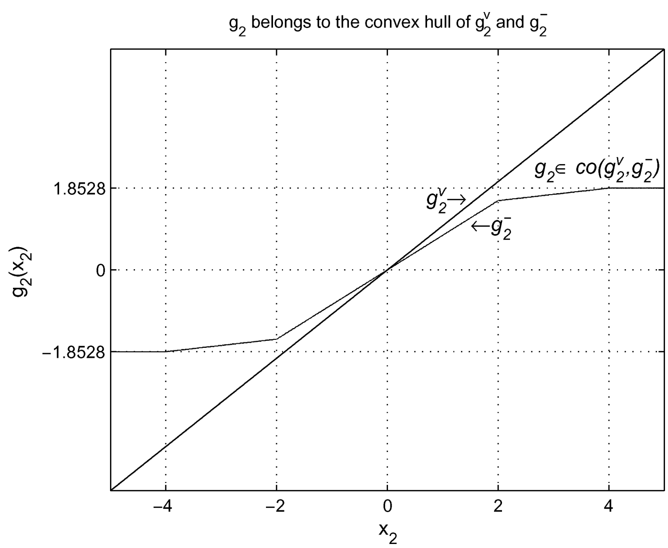

Fig. 9. The g_2 belongs to the convex hull of g_2^v and g_2^-.

The piecewise function (29) also has three bounds with $\bar{N}_2 = 2$ and

$$(\bar{k}_{20}, \bar{k}_{21}, \bar{k}_{22}) = (0.7845, 0.1419, 0) \tag{34}$$
$$(\bar{e}_{20}, \bar{e}_{21}, \bar{e}_{22}) = (0, 1.2852, 1.8528) \tag{35}$$
$$(\bar{b}_{20}, \bar{b}_{21}, \bar{b}_{22}) = (0, 2, 4) \tag{36}$$
$$(\bar{v}_{21}, \bar{v}_{22}) = (0.5, 0.4234). \tag{37}$$

Functions $\breve{g}_1(x_1) = x_1$ and $\breve{g}_2(x_2) = x_2$ have one bound.

We have the following four RNNs under the vertex activation functions:

$$\begin{cases} \dfrac{dx_1(t)}{dt} = -c_1x_1(t) + a_{11}\bar{g}_1\left(x_1(t)\right) + a_{12}\bar{g}_2\left(x_2(t)\right) + I_1 \\ \dfrac{dx_2(t)}{dt} = -c_2x_2(t) + a_{21}\bar{g}_1\left(x_1(t)\right) + a_{22}\bar{g}_2\left(x_2(t)\right) + I_2 \end{cases} \tag{38}$$

$$\begin{cases} \dfrac{dx_1(t)}{dt} = -c_1x_1(t) + a_{11}\breve{g}_1\left(x_1(t)\right) + a_{12}\breve{g}_2\left(x_2(t)\right) + I_1 \\ \dfrac{dx_2(t)}{dt} = -c_2x_2(t) + a_{21}\breve{g}_1\left(x_1(t)\right) + a_{22}\breve{g}_2\left(x_2(t)\right) + I_2 \end{cases} \tag{39}$$

$$\begin{cases} \dfrac{dx_1(t)}{dt} = -c_1x_1(t) + a_{11}\bar{g}_1\left(x_1(t)\right) + a_{12}\breve{g}_2\left(x_2(t)\right) + I_1 \\ \dfrac{dx_2(t)}{dt} = -c_2x_2(t) + a_{21}\bar{g}_1\left(x_1(t)\right) + a_{22}\breve{g}_2\left(x_2(t)\right) + I_2 \end{cases} \tag{40}$$

$$\begin{cases} \dfrac{dx_1(t)}{dt} = -c_1x_1(t) + a_{11}\breve{g}_1\left(x_1(t)\right) + a_{12}\bar{g}_2\left(x_2(t)\right) + I_1 \\ \dfrac{dx_2(t)}{dt} = -c_2x_2(t) + a_{21}\breve{g}_1\left(x_1(t)\right) + a_{22}\bar{g}_2\left(x_2(t)\right) + I_2 \end{cases} \tag{41}$$

where $c_1 = c_2 = 2$, $a_{11} = a_{22} = 1$, $a_{12} = 1.1$, and $a_{21} = 1.2$.

Next, using Theorem 2, we study the GES of the previously mentioned four RNNs with the vertex activation functions.

1) We have rewritten the neural network system (38) in form of (19). As $\bar{N}_1 = 2$ and $\bar{N}_2 = 2$, $H_{(38)} = \bar{N}_1 \times \bar{N}_2 = 2 \times 2 = 4$; so $p \in I[1,4]$. Hence

$$\begin{cases} K_1 = \mathrm{diag}(0.3218, 0.1419) \\ E_1 = \mathrm{diag}(0.4627, 1.2852) \\ V_1 = \mathrm{diag}(1, 0.5) \end{cases}$$
$$\begin{cases} K_2 = \mathrm{diag}(0.3218, 0) \\ E_2 = \mathrm{diag}(0.4627, 1.8528) \\ V_2 = \mathrm{diag}(1, 0.4234) \end{cases}$$
$$\begin{cases} K_3 = \mathrm{diag}(0, 0.1419) \\ E_3 = \mathrm{diag}(1.4281, 1.2852) \\ V_3 = \mathrm{diag}(0.5493, 0.5) \end{cases}$$
$$\begin{cases} K_4 = \mathrm{diag}(0, 0) \\ E_4 = \mathrm{diag}(1.4281, 1.8528) \\ V_4 = \mathrm{diag}(0.5493, 0.4234). \end{cases}$$

Thus, we can calculate

$$C_1 = C - |A|K_1 = \begin{pmatrix} 8.3564 & -1.1352 \\ -0.6436 & 8.2905 \end{pmatrix}$$
$$C_2 = C - |A|K_2 = \begin{pmatrix} 8.3564 & 0 \\ -0.6436 & 9.0000 \end{pmatrix}$$
$$C_3 = C - |A|K_3 = \begin{pmatrix} 9.0000 & -1.1352 \\ 0 & 8.2905 \end{pmatrix}$$
$$C_4 = C - |A|K_4 = \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix}$$
$$\Pi_1 = C - E_1V_1 - |A|K_1 = \begin{pmatrix} 7.8937 & -1.1352 \\ -0.6436 & 7.6479 \end{pmatrix}$$
$$\Pi_2 = C - E_2V_2 - |A|K_2 = \begin{pmatrix} 7.8937 & 0 \\ -0.6436 & 8.2155 \end{pmatrix}$$
$$\Pi_3 = C - E_3V_3 - |A|K_3 = \begin{pmatrix} 8.2155 & -1.1352 \\ 0 & 7.6479 \end{pmatrix}$$
$$\Pi_4 = C - E_4V_4 - |A|K_4 = \begin{pmatrix} 8.2155 & 0 \\ 0 & 8.2155 \end{pmatrix}.$$

It is easy to see that $C_1 - C_4$ are invertible. From Lemma 3, we know that $\Pi_1 - \Pi_4$ are M-matrices. Hence, by Theorem

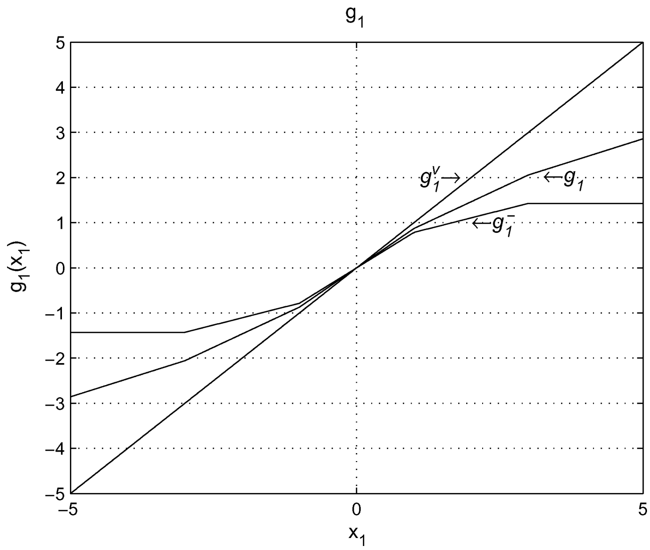

Fig. 10. The $g_1(x_1)$.

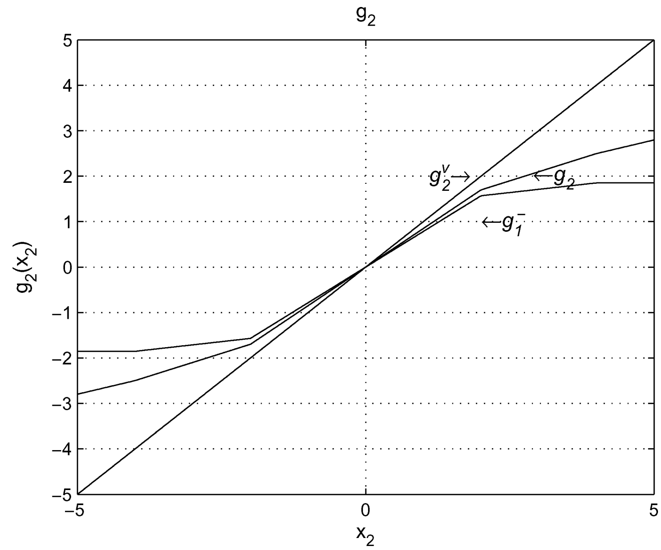

Fig. 11. The $g_2(x_2)$.

2, the trivial solution $x = x^e$ of neural networks system (38) is GES.

2) For neural networks system (39), $\breve{N}_1 = 1$ and $\breve{N}_2 = 1$, $H_{(39)} = \breve{N}_1 \times \breve{N}_2 = 1 \times 1 = 1$; so $p = 1$. Hence

$$\begin{cases} K_1 = \mathrm{diag}(1,1) \\ E_1 = \mathrm{diag}(0,0) \\ V_1 = \mathrm{diag}(0,0). \end{cases}$$

Using Theorem 2, we can calculate

$$C_1 = C - |A|K_1 = \begin{pmatrix} 7 & -8 \\ -2 & 4 \end{pmatrix}$$

$$\Pi_1 = C - E_1V_1 - |A|K_1 = \begin{pmatrix} 7 & -8 \\ -2 & 4 \end{pmatrix}.$$

It is easy to see that C_1 is invertible. From Lemma 3, we know that Π_1 is M-matrix. Hence, by Theorem 2, the trivial solution $x = x^e$ of neural networks system (39) is GES.

3) For neural networks system (40), $\bar{N}_1 = 2$ and $\breve{N}_2 = 1$, $H_{(40)} = \bar{N}_1 \times \breve{N}_2 = 2 \times 1 = 2$; so $p = 1, 2$. Hence

$$\begin{cases} K_1 = \mathrm{diag}(0.3218, 1) \\ E_1 = \mathrm{diag}(0.4627, 0) \\ V_1 = \mathrm{diag}(1, 0) \end{cases}$$

$$\begin{cases} K_2 = \mathrm{diag}(0, 1) \\ E_2 = \mathrm{diag}(1.4821, 0) \\ V_2 = \mathrm{diag}(0.5493, 0). \end{cases}$$

Using Theorem 2, we can calculate

$$C_1 = C - |A|K_1 = \begin{pmatrix} 8.3564 & -8.0000 \\ -0.6436 & 4.0000 \end{pmatrix}$$

$$C_2 = C - |A|K_2 = \begin{pmatrix} 9 & -8 \\ 0 & 4 \end{pmatrix}$$

$$\Pi_1 = C - E_1V_1 - |A|K_1 = \begin{pmatrix} 7.8937 & -8.0000 \\ -0.6436 & 4.0000 \end{pmatrix}$$

$$\Pi_2 = C - E_2V_2 - |A|K_2 = \begin{pmatrix} 8.1859 & -8.0000 \\ 0 & 4.0000 \end{pmatrix}.$$

It is easy to see that C_1 and C_2 are invertible. From Lemma 3, we know that Π_1 and Π_2 are M-matrices. Hence, by Theorem 2, the trivial solution $x = x^e$ of neural networks system (40) is GES.

4) For neural networks system (41), $\breve{N}_1 = 1$ and $\bar{N}_2 = 2$, $H_{(41)} = \breve{N}_1 \times \bar{N}_2 = 1 \times 2 = 2$; so $p = 1, 2$. Hence

$$\begin{cases} K_1 = \mathrm{diag}(1, 0.1419) \\ E_1 = \mathrm{diag}(0, 1.2852) \\ V_1 = \mathrm{diag}(0, 0.5) \end{cases}$$

$$\begin{cases} K_2 = \mathrm{diag}(1, 0) \\ E_2 = \mathrm{diag}(0, 1.8528) \\ V_2 = \mathrm{diag}(0, 0.4234). \end{cases}$$

Using Theorem 2, we can calculate

$$C_1 = C - |A|K_1 = \begin{pmatrix} 7.0000 & -1.1352 \\ -2.0000 & 8.2905 \end{pmatrix}$$

$$C_2 = C - |A|K_2 = \begin{pmatrix} 7 & 0 \\ -2 & 9 \end{pmatrix}$$

$$\Pi_1 = C - E_1V_1 - |A|K_1 = \begin{pmatrix} 7.0000 & -1.1352 \\ -2.0000 & 7.6479 \end{pmatrix}$$

$$\Pi_2 = C - E_2V_2 - |A|K_2 = \begin{pmatrix} 7.0000 & 0 \\ -2.0000 & 8.2155 \end{pmatrix}.$$

It is easy to see that C_1 and C_2 are invertible. From Lemma 3, we know that Π_1, Π_2 are M-matrices. Hence, by Theorem 2, the trivial solution $x = x^e$ of neural networks system (40) is GES.

According to 1)–4), by Theorem 3, we can conclude that the trivial solution $x = x^e$ of the neural networks (27) is AEST.

For numerical simulation, we arbitrarily select one of the activation functions in $\mathrm{co}(\bar{g}_1(x_1), x_1)$ and $\mathrm{co}(\bar{g}_2(x_2), x_2)$, respectively. For example, let $g_1(x_1) = 0.6\bar{g}_1(x_1) + 0.4x_1$ and $g_2(x_2) = 0.7\bar{g}_2(x_2) + 0.3x_2$ in system (27). The selected activation functions are shown as q_1 and q_2 in Fig. 10 and Fig. 11, respectively. The following three different initial conditions are given:

Case 1) initial state $x_{01} = (2, -3)^T$;
Case 2) initial state $x_{0_2} = (-2, 3)^T$;
Case 3) initial state $x_{0_3} = (0, 1)^T$.

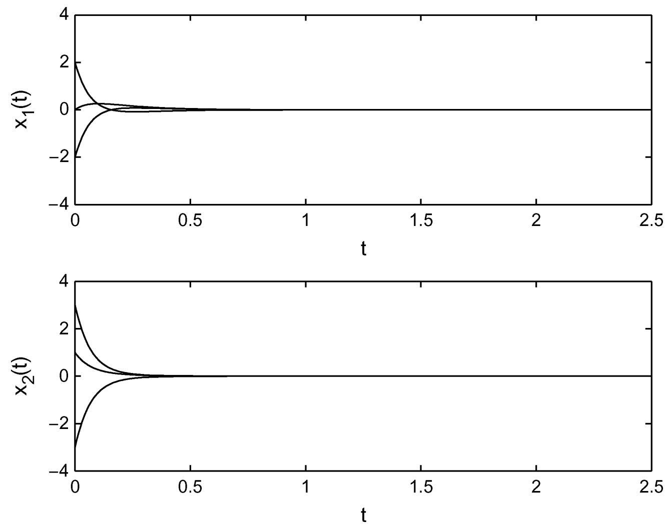

Fig. 12. Transient response of the state variables.

Fig. 12 shows the time responses of the states $x_1(t)$ and $x_2(t)$ with the three initial conditions. We see that all the states are convergent to the unique and GES equilibrium point of the system. The equilibrium point is at the origin.

B. Comparison

From neural network (27), we can compare our M-matrix criteria with some stability criteria. In [2],[10],[11],[17],[19], [20], [39], and [41], the absolute stability or absolute exponential stability were also studied. However, the stability criteria in these papers only gave some conditions on the interconnection weight matrix, such as $A \in \mathcal{M}_0, A \in \mathcal{C}_0$, and $-A \in \mathcal{P}_0$, where $\mathcal{M}_0, \mathcal{C}_0$ and $\mathcal{P}_0$ are denoted in [38] and [40]. As Cao *et al.* pointed out in [38], A may not belong to $\mathcal{M}_0, \mathcal{C}_0, -A$ nor it may belong to $\mathcal{P}_0$. Therefore, our results are less conservative than the references mentioned. In (27), we can see that matrix A does not belong to $\mathcal{C}_0$, $\mathcal{M}_0$, or $\mathcal{P}_0$; so, the stability of the equilibrium point of system (27) could not be shown by the stability criteria in the references mentioned.

V. Conclusion

In this paper, both the existence of an equilibrium point and the AEST of RNN with a generalized activation functions are addressed. Every component of the activation function of neural networks belongs to the convex hull of two piecewise linear functions. This generalized activation function allows a more flexible or more specific description for the activation functions. We demonstrate that GES of the equilibrium point of original RNN with a generalized activation function is equivalent to that of RNN under all vertex functions of convex hull. Then, the neural network under all the vertex activation functions is transformed into neural network under an array of saturated linear activation functions. Again, we demonstrate that GES of the equilibrium point of RNN system under all vertex functions of the convex hull is equivalent to that of neural network under an array of saturated linear activation functions. Because of the equivalence of three systems, the stability analysis is focused on the RNN system under an array of saturated linear activation function. In the end, a two-neuron RNN with a generalized activation function is constructed to show the effectiveness of our results. To demonstrate the advantage of our results, a comparison with the previous ones was drawn.

Appendix
Proof of Lemma 4

Proof: "Only if:" Assume (6) is GES. As $g \subset \mathrm{co}\{g^l : l \in I[1, 2^n]\}$, by set $l \in I[1, 2^n]$ in (6), we have

$$\frac{dx(t)}{dt} = -Cx(t) + Ag\left(x(t)\right) + I = -Cx(t) + Ag^l\left(x(t)\right) + I$$

which implies every system (10) is GES.

"If:" Now, we assume that every individual system (10) is GES. We need to show that system (6) is also GES.

Because every system (10) is GES, with b) in Lemma 1, there exists a function that $V : \mathcal{R}^n \to R$ satisfies the following inequalities:

$$c_1\|x\|^2 \leq V(x) \leq c_2\|x\|^2 \tag{42}$$
$$\frac{\partial V}{\partial x} f^l(x) \leq -c_3\|x\|^2 \tag{43}$$
$$\left\|\frac{\partial V}{\partial x}\right\|_{(10)} \leq c_4\|x\| \tag{44}$$

for some positive constants c_1, c_2, c_3, and c_4, where $f^l(x) = -Cx(t) + Ag^l(x(t)) + I$.

To show that system (6) is also GES, we need to prove that for $V : \mathcal{R}^n \to R$, the following two inequalities hold with b) in Lemma 1:

$$\frac{\partial V}{\partial x} f(x) \leq -c_5\|x\|^2 \tag{45}$$
$$\left\|\frac{\partial V}{\partial x}\right\|_{(6)} \leq c_6\|x\| \tag{46}$$

for positive constants c_5 and c_6, where

$$f(x) = -Cx(t) + Ag\left(x(t)\right) + I. \tag{47}$$

Then, we can use a) and b) in Lemma 1 to prove that system (6) is also GES.

Proof of Inequality (45): As assumed previously, g_i, the ith component of g, belongs to the convex hull of $\bar{g}_i$ and $\breve{g}_i$, and $\bar{g}_i$ and $\breve{g}_i$ are continuously piecewise differentiable on $\mathcal{R}$. According to the definition of the convex hull in Section II-A1, it is not difficult to prove that g_i is also continuously piecewise differentiable on $\mathcal{R}$. In view of (47), $f(x)$ is continuously piecewise differentiable on $\mathcal{R}^n$. Thus, the Jacobian matrix $[\partial f / \partial x]$ is bounded on $\mathcal{R}^n$.

In view of (43), there exists a constant c_5 such that

$$\begin{aligned}\frac{\partial V}{\partial x} f(x) &= \frac{\partial V}{\partial x} \cdot \left(-Cx(t) + Ag\left(x(t)\right) + I\right)\\ &\leq \max_{1 \leq k \leq 2^n} \left\{\nabla V(x) \cdot \left(-Cx(t) + Ag^k\left(x(t)\right) + I\right)\right\}\\ &= \max_{1 \leq k \leq 2^n} \frac{\partial V(x)}{\partial x} f^k(x)\\ &\leq -c_5\|x\|^2. \end{aligned} \tag{48}$$

Therefore, the inequality (45) holds.

Proof of Inequality (46): To prove the inequality (46), the following two inequalities have to be proven.

For system (5), the following two inequalities hold:

$$\left|\frac{d}{dt}\left[x^T(t)x(t)\right]\right| \leq 2L\,\|x(t)\|_2^2 \tag{49}$$

$$\|x(t)\|_2 \leq \|x_0\|_2 \exp(Lt) \tag{50}$$

As inequality (50) can be easily obtained by (49), we mainly prove (49). Because the Jacobian matrix $[\partial f/\partial x]$ is bounded on $x \in \mathcal{R}^n$, there is a constant positive number L such that

$$\left\|\frac{\partial f}{\partial x}\right\|_2 \leq L. \tag{51}$$

For system (6), we assume that there is an equilibrium point at the origin. Then, it is easy to know that $f(0) = 0$. As $f(x)$ is differentiable on $\mathcal{R}^n$, by mean-value theorem, there is $\xi = \theta x$, where $0 < \theta < 1$, such that

$$f(x) - f(0) = \frac{\partial f}{\partial x}|_{(x=\xi)}(x - 0).$$

In view of inequality (51), we have

$$\|f(x) - f(0)\|_2 = \|f(x)\|_2 \leq \left\|\frac{\partial f}{\partial x}\right|_{(x=\xi)} \|_2 \|x\|_2 \leq L\|x\|_2. \tag{52}$$

With system (5), we have

$$\left|\frac{d}{dt}\left[x^T(t)x(t)\right]\right| = 2\left|x^T(t)\dot{x}(t)\right| = 2\left|x^T(t)f(x)\right|. \tag{53}$$

By Hölder inequality

$$\left|x^T(t)f(x)\right| \leq \|x\|_2 \,\|f(x)\|_2 \tag{54}$$

Thus, by (52)–(54), it follows that

$$\left|\frac{d}{dt}\left[x^T(t)x(t)\right]\right| \leq 2L\,\|x(t)\|_2^2. \tag{55}$$

The last assertion (50) can be easily proven by integrating both sides of inequality (49) from t_0 to t. With simple calculation, we can obtain inequality (50).

Let $\phi(t;x)$, which begins with x, denote the solution of the system (5) at time t. Then

$$\phi(t;x) = \phi(0;x) + \int_0^t f\left(\phi\left(\tau; x(\tau)\right)\right) d\tau.$$

Taking partial derivatives with respect to x yields

$$\phi_x(t;x) = \frac{\partial \phi(t;x)}{\partial x} = \int_0^t \frac{\partial f\left(\phi\left(\tau;x(\tau)\right)\right)}{\partial \phi\left(\tau;x(\tau)\right)} \frac{\partial \phi\left(\tau;x(\tau)\right)}{\partial x} d\tau. \tag{56}$$

Differentiating with respect to t, it can be observed that ϕ_x satisfies the sensitivity equation

$$\frac{\partial \phi_x(t;x)}{\partial t} = A(\phi)\phi_x(t;x)$$

where $A(\phi) = (\partial f/\partial \phi)(\phi(t;x))$ and $\phi_x(0;x) = N$, where N is a constant matrix. Since

$$\left\|\frac{\partial f}{\partial \phi}\right\|_2 \leq L \tag{57}$$

on $\mathcal{R}$, with the inequalities (49) and (50), ϕ_x satisfies the bound

$$\|\phi_x(t;x)\|_2 \leq \|N\|_2 e^{Lt}. \tag{58}$$

Let

$$V(x) = \int_0^t \phi^T(\tau;x)\phi(\tau;x)d\tau$$

where $t > 0$. Therefore, with a) in Lemma 1 and inequality (58), we have

$$\begin{aligned}
\left\|\frac{\partial V}{\partial x}\right\|_2 &= \left\|\int_0^t 2\phi^T(\tau;x)\phi_x(\tau;x)d\tau\right\|_2 \\
&\leq \int_0^t 2\left\|\phi^T(\tau;x)\right\|_2 \|\phi_x(\tau;x)\|_2\, d\tau \|x\|_2 \\
&\leq \int_0^t 2\|M\|_2\|N\|_2\|x_0\|_2 e^{-\lambda\tau} e^{L\tau} d\tau \|x\|_2 \\
&= \frac{2\|M\|_2\|N\|_2\|x_0\|_2}{(\lambda - L)}\left[1 - e^{-(\lambda - L)t}\right] \|x\|_2.
\end{aligned}$$

Thus, by choosing $\lambda > L$, the inequality (46) is satisfied with

$$c_6 = \frac{2\|M\|_2\|N\|_2\|x_0\|_2}{(\lambda - L)}\left[1 - e^{-(\lambda - L)t}\right] > 0.$$

Using (42), (45), and (46), b) in Lemma 1 is satisfied. Thus, from Lemma 1, we can conclude that system (6) is GES. The proof is completed.

REFERENCES

[1] S. Arik and V. Tavsanoglu, "Equilibrium analysis of delayed CNN," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 45, no. 2, pp. 168–171, Feb. 1998.

[2] S. Arik and V. Tavsanoglu, "A comment on Comments on 'Necessary and sufficient condition for absolute stability of neural networks,'," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 45, no. 5, pp. 1057–7122, May 1998.

[3] S. Arik and V. Tavsanoglu, "On the global asymptotic stability of delayed cellular neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 47, no. 4, pp. 571–574, Apr. 2000.

[4] J. Cao and J. Wang, "Global asymptotic stability of a general class of recurrent neural networks with time-varying delays," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 1, pp. 34–44, Jan. 2003.

[5] J. Cao and J. Wang, "Absolute exponential stability of recurrent neural networks with Lipschitz-continuous activation functions and time delays," *Neural Netw.*, vol. 17, pp. 379–390, 2004.

[6] T. Chu, C. Zhang, and Z. Zhang, "Necessary and sufficient condition for absolute stability of normal neural networks," *Neural Netw.*, vol. 16, no. 8, pp. 1223–1227, 2003.

[7] L. O. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.*, vol. 35, no. 10, pp. 1257–1272, Oct. 1988.

[8] L. O. Chua and L. Yang, "Cellular neural networks: Applications," *IEEE Trans. Circuits Syst.*, vol. 35, no. 10, pp. 1273–1290, Oct. 1988.

[9] L. O. Chua and T. Roska, "Cellular neural networks with nonlinear and delay-type template elements," in *Proc. 1990 Int. Workshop Cellular Neural Netw. Appl.*, 2006, pp. 12–25.

[10] M. Forti, S. Manetti, and M. Marini, "Necessary and sufficient condition for absolute stability of neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 41, no. 7, pp. 491–494, Jul. 1994.

[11] M. Forti and A. Tesi, "New conditions for global stability of neural networks with application to linear and quadratic programming problems," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 42, no. 7, pp. 354–366, Jul. 1995.

[12] M. Forti, "A note on neural networks with multiple equilibrium points," *IEEE Trans. Circuits Syst. I, Fundam Theory Appl.*, vol. 43, no. 6, pp. 487–491, Jun. 1996.

[13] M. Joy, "Results concerning the absolute stability of delayed neural networks," *Neural Netw.*, vol. 13, pp. 613–616, 2000.

[14] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[15] T. Hu, B. Huang, and Z. Lin, "Absolute stability with a generalized sector condition," *IEEE Trans. Autom. Control*, vol. 49, no. 4, pp. 535–548, Apr. 2002.

[16] T. Hu and Z. Lin, "Absolute stability analysis of discrete-time systems with composite quadratic Lyapunov functions," *IEEE Trans. Autom. Control*, vol. 50, no. 6, pp. 781–797, Jun. 2005.

[17] S. Hu and J. Wang, "Absolute exponential stability of a class of continuous-time recurrent neural networks," *IEEE Trans. Neural Netw.*, vol. 14, no. 1, pp. 35–45, Jan. 2003.

[18] T. Huang, J. Cao, and C. Li, "Necessary and sufficient condition for the absolute exponential stability of a class of neural networks with finite delay," *Phys. Lett. A*, vol. 352, no. 1–2, pp. 94–98, 2006.

[19] X. B. Liang and J. Wang, "Absolute exponential stability of neural networks with a general class of activation functions," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 47, no. 8, pp. 1258–1263, Aug. 2000.

[20] X. B. Liang and J. Wang, "An additive diagonal-stability condition for absolute exponential stability of a general class of neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 11, pp. 1308–1317, 2001.

[21] H. Lu, R. Shen, and F. Chuang, "Absolute exponential stability of a class of recurrent neural networks with multiple and variable delays," *Theor. Comput. Sci.*, vol. 344, pp. 103–119, 2005.

[22] H. Lu, "Absolute exponential stability analysis of delayed neural networks," *Phys. Lett. A*, vol. 336, pp. 133–140, 2005.

[23] X. Lou and B. Cui, "Absolute exponential stability analysis of delayed bi-directional associative memory neural networks," *Chaos Solitons Fractals*, vol. 31, pp. 695–701, 2007.

[24] M. Morita, "Associative memory with nonmonotone dynamics," *Neural Netw.*, vol. 6, pp. 115–126, 1993.

[25] T. Roska, T. Boros, P. Thiran, and L. O. Chua, "Detecting simple motion using cellular neural networks," *Proc. Int. Workshop Cellular Neural Netw. Appl.*, pp. 127–138, 1990.

[26] T. Roska, C. M. Wu, M. Balsi, and L. O. Chua, "Stability and dynamics of delay-type general and cellular neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 39, no. 6, pp. 487–490, Jun. 1992.

[27] T. Roska and L. O. Chua, "Cellular neural networks with non-linear and delay-type template elements and non-uniform grids," *Int. J. Circuit Theory Appl.*, vol. 20, pp. 469–481, 1992.

[28] T. Rosks, C. M. Wu, and L. O. Chua, "Stability of cellular neural networks with dominant non-linear and delay-type templates," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 40, no. 4, pp. 270–272, Apr. 1993.

[29] P. P. Civalleri, M. Gilli, and L. Pandolfi, "On stability of cellular neural networks with delay," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 40, no. 3, pp. 157–164, Mar. 1993.

[30] S. Sastry, *Nonlinear Systems: Analysis, Stability, and Control*. New York: Springer-Verlag, 1999.

[31] S. Xu, Y. Chu, and J. Lu, "New results on global exponential stability of recurrent neural networks with time-varying delays," *Phys. Lett. A*, vol. 352, pp. 371–379, 2006.

[32] S. Yoshizawa, M. Morita, and S. I. Amari, "Capacity of associative memory using a nonmonotic neural model," *Neural Netw.*, vol. 6, pp. 167–176, 1993.

[33] W. Zhao and H. Zhang, "Globally exponential stability of neural network with constant and variable delays," *Phys. Lett. A*, vol. 352, pp. 350–357, 2006.

[34] J. Y. Zhang, "Globally exponential stability of neural networks with variable delays," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 2, pp. 288–291, Feb. 2003.

[35] J. Zhang, "Absolute stability of a class of neural networks with unbounded delay," *Int. J. Circuit Theory Appl.*, vol. 32, pp. 11–21, 2004.

[36] J. Zhang, "Absolutely exponential stability in delayed cellular neural networks," *Int. J. Circuit Theory Appl.*, vol. 30, pp. 395–409, 2002.

[37] J. Zhang, Y. Suda, and T. Iwasa, "Absolutely exponential stability of a class of neural networks with unbounded delay," *Neural Netw.*, vol. 17, pp. 391–397, 2004.

[38] J. Cao and J. Wang, "Global asymptotic stability of a general class of recurrent neural networks with time-varying delays," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 1, pp. 34–44, Feb. 2003.

[39] E. Kaszkurewicz and A. Bhaya, "Comments on 'Necessary and sufficient condition for absolute stability of neural networks'," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 42, no. 8, pp. 497–499, Aug. 1995.

[40] S. Arik and V. Tavsanoglu, "A comment on 'Comments on necessary and sufficient condition for absolute stability of neural networks'," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 45, no. 5, pp. 595–596, May 1998.

[41] M. Forti, "On global asymptotic stability of a class of nonlinear systems arising in neural networks theory," *J. Differ. Equ.*, vol. 113, pp. 246–264, 1994.

[42] X. B. Liang and J. Si, "Comments on 'New conditions for global stability of neural networks with application to linear and quadratic programming problems'," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 44, no. 11, pp. 1099–1101, Nov. 1997.

[43] Y. He, M. Wu, and J.-H. She, "An improved global asymptotic stability criterion for delayed cellular neural networks," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 250–252, Jan. 2006.

[44] S. Guo and L. Huang, "Stability analysis of Cohen-Grossberg neural networks," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 106–117, Jan. 2006.

[45] Z. Zeng and J. Wang, "Improved conditions for global exponential stability of recurrent neural networks with time-varying delays," *IEEE Trans. Neural Netw.*, vol. 17, no. 3, pp. 623–635, May 2006.

[46] Y.-Y. Cao, Z. Lin, and D. G. Ward, "An antiwindup approach to enlarging domain of attraction for linear systems subject to actuator saturation," *IEEE Trans. Autom. Control*, vol. 47, no. 1, pp. 140–145, Jan. 2002.

[47] Y.-Y. Cao, Z. Lin, and T. Hu, "Stability analysis of linear time-delay systems subject to input saturation," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 2, pp. 233–240, Feb. 2002.

[48] Y.-Y. Cao and Z. Lin, "Stability analysis of discrete-time systems with actuator saturation by a saturation-dependent Lyapunov function," *Automatica*, vol. 39, pp. 1235–1241, 2003.

[49] J. Xu, D. Y. Pi, Y.-Y. Cao, and S. M. Zhong, "On stability of neural networks by a Lyapunov functional-based approach," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 4, pp. 912–924, Apr. 2007.

[50] J. Xu, Y.-Y. Cao, D. Y. Pi, and Y. X. Sun, "An estimation of the domain of attraction for general recurrent delayed neural networks," *Neurocomputing*, to be published.

Jun Xu was born in Jiangxi, China. He received the M.S. degree in applied mathematics from the University of Electronic Science and Technology of China, Chengdu, China, in 2004, and the Ph.D. in control science and engineering from the State Key Laboratory of Industrial Control Technology, Zhejiang University, Zhejiang, China, in April 2007.

He jointed the Department of Advanced Technologies, Alcorn State University, Alcorn State, MS, as a Postdoctoral Researcher in May 2007. His research interest include neural networks, pattern recognition, machine learning, and its application to medical image processing. He authored two journal papers in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS and *Neurocomputing*.

Yong-Yan Cao (SM'02) was born in Hunan, China, in 1968. He received the B.E. and M.Sc. degrees in electrical engineering from Wuhan University of Science and Technology, Wuhan, China, in 1990 and 1993, respectively, and the Ph.D. degree in industrial automation from Zhejiang University, Zhejiang, China, in 1996.

From 1996 to 1997, he was a Postdoctoral Researcher at the Institute of Industrial Process Control, Zhejiang University. From 1997 to 1998, he was a Visiting Research Associate at the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong. From 1998 to 1999, he was an Associate Professor at the Zhejiang University. He then spent a year and a half at the Department of Measurement and Control, Duisburg University, Duisburg, Germany, as an Alexander von Humboldt Research Fellow. In May 2000, he joined the Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, as a Research Scientist. Currently, he is a Professor at the Department of Automation, Wuhan University of Science and Technology, Wuhan, China. He has published more than 50 papers in reputed journals and conferences such as the IEEE TRANSACTION ON AUTOMATIC CONTROL, *Automatica*, *Systems and Control Letters*, etc. He also has been active both in his own research and in his service as a reviewer for many journals and conferences. His current research interests include robust and nonlinear control, constrained control, time-delay systems, sampled-data systems, and fuzzy control.

Dr. Cao is an Associate Editor of *Automatica* and the Conference Editorial Board of the IEEE Control Systems Society.

Youxian Sun was born in Zhuji, Zhejiang Province, China, in 1940. He graduated from Zhejiang University in 1964. He was elected as an Academician of Chinese Academy of Engineering in 1995.

Currently, he is a Professor at the Zhejiang University, Director of Institute of Industrial Control, and Director of National Engineering Research Center of Industrial Automation. He has made significant contributions to modern control engineering. Under his leadership, the first National Engineering Research Center in China was established, and the first five-star institution in China formed, which consists of Authorization of Conferring Doctoral Degree, Postdoctoral Station, 2 National Key Specialties, National Key Laboratory, and National Engineering Research Center. He proposed two important special projects including "Industrialization of China Industrial Process Control High-Technologies," which were adopted by the government. He has taken charge of four State-Level Key Projects such as "Commercial Engineering Software for Industrial Process Modeling, Control and Optimization," "42 State-Level Special Projects," and "8 Province- or Ministry-Level Key Projects." He has authored or edited 15 books and has published more than 450 papers.

Dr. Sun serves as Vice-Chairman of IFAC Pulp and Paper Committee, Vice-President of Chinese Association of Automation, and Vice-President of China Instrument and Control Society. He has won the Second and Third Prize of National Science and Technology Progress in China once, respectively, the Prize of National Outstanding Teaching Achievements twice, and the First or Second Prize of Province- or Ministry-Level Science and Technology Progress for 16 times. He was also awarded National Model Worker in Education, People's Teacher Medal, the First National Excellent Scientist, and National Prominent Expert of the Young and the Middle-Aged in China.

Jinshan Tang (SM'03) received the Ph.D. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 1998.

He joined ATR media integration and communication research laboratories in Japan. In 2000, he went to Harvard Medical School and received postdoctoral training. In 2001, he joined the University of Virginia and worked there for about three years. After that, he was a Visiting Fellow at the National Cancer Institute, National Institutes of Health, Bethesda, MD, and as a Senior Engineer at Intel, Huston, MA. Since 2006, he has been an Assistant Professor at Alcorn State University, Alcorn, MS. His teaching interests are digital image processing, computer networks, and numerical analysis. His current research interests include image processing and pattern recognition for cancer detection and diagnosis, machine learning and its application, 3-D biometrics, MRI, and ultrasound imaging. He is an author or coauthor of more than 50 papers on image processing and pattern recognition.

Dr. Tang served as the chair of a special session on image processing and analysis for oncology at 2007 IEEE International Conference on Image Processing.