Available online at www.sciencedirect.com
SCIENCE DIRECT
Neural Networks 18 (2005) 171–178
Neural Networks
www.elsevier.com/locate/neunet

On the global output convergence of a class of recurrent neural networks with time-varying inputs☆

Sanqing Hu, Derong Liu*

Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL 60607, USA

Received 1 April 2003; revised 29 October 2004; accepted 29 October 2004

Abstract

This paper studies the global output convergence of a class of recurrent neural networks with *globally* Lipschitz continuous and monotone nondecreasing activation functions and *locally* Lipschitz continuous time-varying inputs. We establish two sufficient conditions for global output convergence of this class of neural networks. Symmetry in the connection weight matrix is not required in the present results which extend the existing ones.

Keywords: Recurrent neural networks; Global output convergence; Time-varying input; Lyapunov diagonal stability; Lipschitz continuity; Optimization

1. Introduction

In this paper, we consider a class of continuous-time recurrent neural networks (RNNs) defined by the following model

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = \sum_{j=1}^{n} w_{ij} g_j(x_j(t)) + u_i(t), x_i(0) = x_{i0}, \quad i = 1, 2, \ldots, n,$$

or equivalently in matrix format given by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Wg(x(t)) + u(t), \ x(0) = x_0, \tag{1}$$

where $x = [x_1, x_2, \ldots, x_n]^T \in R^n$ is the state vector, $W = [w_{ij}] \in R^{n \times n}$ is a constant connection weight matrix, $u(t) = [u_1(t), u_2(t), \ldots, u_n(t)]^T \in R^n$ is a nonconstant locally Lipschitz continuous input vector function defined on $[0, +\infty)$ which is called the time-varying input, $g(x) = [g_1(x_1), g_2(x_2), \ldots, g_n(x_n)]^T$ is a nonlinear vector-valued activation function from R^n to R^n and $g(x)$ is called the output of the network (1). The RNN model (1) with constant inputs has important applications to content-addressable memory (CAM) problem (Grossberg, 1988). Recently, the RNN model (1) has been applied to solving various optimization problems such as linear programming problem (Wang, 1993, 1994), shortest path problem (Wang, 1996), sorting problem (Wang, 1995), and assignment problem (Wang, 1992, 1997).

In this paper we focus on the RNN model (1) with time-varying inputs. In the application of neural networks to CAM and optimization problems, state and/or output convergence of the network is a basic requirement. There are several reasons to consider state and output convergence of the RNNs with time-varying inputs. First, as mentioned in Hirsch (1989), time-varying inputs $u(t)$ can drive quickly $x(t)$ to some desired region of the activation space. Second, in some RNNs for optimization, it is required for their inputs to vary over time to ensure the feasibility and optimality of solutions, as elaborated in (Wang, 1992, 1993, 1994, 1995, 1996, 1997). The convergence of these RNNs needs an in-depth investigation. Third, as the inputs are also adaptive parameters, the convergence issue arises for online learning of RNNs. Fourth, the inputs can be considered as external inputs. The presence of time-varying external inputs also entails the study of convergence of RNNs. Fifth, in a cascade neural network which was studied in Hirsch (1989),

☆ This work was supported by the National Science Foundation under Grant ECS-9996428.

* Corresponding author. Tel.: +1 312 355 4475; fax: +1 312 996 6465.
E-mail addresses: shu@cil.ece.uic.edu (S. Hu), dliu@ece.uic.edu (D. Liu).

0893-6080/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2004.10.005

its inputs are the outputs of the previous layers. In summary, it is necessary and desirable to further investigate the convergence of the RNN model (1).

In general, it is very difficult to analyze the state convergence or stability of the RNN model (1) with time-varying inputs. Note that converged output vectors of the networks provide solutions to optimization problems (Wang, 1992, 1993, 1994, 1995, 1996, 1997). Therefore, output convergence analysis of the RNN model (1) is an important task. Due to the time-varying input $u(t)$ and the lack of a linear term, the RNN model (1) is different from the well-known Hopfield neural networks which have been used in some optimization problems (Forti & Tesi, 1995; Hopfield, 1982, 1984). Recently, global asymptotic stability and global exponential stability of the Hopfield neural networks have received attention (Forti & Tesi, 1995; Liang & Si, 2001; Liang & Wu, 1999; Zhang, Heng, & Fu, 1999). For the Hopfield neural networks with sigmoidal activation functions, the quasi-diagonally row-sum and column-sum dominant conditions for absolute stability were obtained (Liang & Wu, 1998), and for these networks with globally Lipschitz continuous and monotone nondecreasing activation functions, Lyapunov diagonal stability results (Forti & Tesi, 1995; Liang & Si, 2001; Zhang et al., 1999) extends many existing results shown in Liang and Si (2001) and Liang and Wu (1999).

The remainder of this paper is organized as follows. In Section 2, some preliminaries on recurrent neural networks are presented. In Section 3, global output convergence results are established. In Section 4, illustrative examples are given. Finally, in Section 5, concluding remarks are presented.

2. Assumptions and preliminaries

We assume that the function $g(\cdot)$ in (1) belongs to the class of globally Lipschitz continuous and monotone nondecreasing activation functions; that is, for $g_i(\cdot)$, there exist $\bar{\ell}_i > 0$ and $\underline{\ell}_i \geq 0$ such that

$$0 \leq \underline{\ell}_i \leq \frac{g_i(\theta) - g_i(\rho)}{\theta - \rho} \leq \bar{\ell}_i, \ \forall\, \theta,\ \rho \in R \text{ and } \theta \neq \rho, \quad i = 1, 2, \ldots, n. \tag{2}$$

It should be noted that such activation functions may not be bounded. There are many frequently used activation functions that satisfy this condition, for example, $1/(1+e^{-\theta})$, $(2/\pi)\arctan(\theta)$, max $(0,\theta)$ and $(|\theta+1|-|\theta-1|)/2$, where $\theta \in R$. We assume that the time-varying input $u_i(t)$ satisfy the following conditions

$$\lim_{t\to+\infty} u_i(t) = u_i, \tag{3}$$

where u_i are some constants, $i=1,2,\ldots,n$, i.e. we assume that $\lim_{t\to+\infty} u(t) = u$.

We will use D^+ in this paper to denote the upper right Dini derivative. For any continuous function f: $R \to R$, the upper right Dini derivative of $f(t)$ is defined as

$$D^+ f(t) = \lim_{h\to 0^+} \sup \frac{f(t+h) - f(t)}{h}.$$

Lemma 1. *(Lemma 1 in Zhang et al., 1999). Let $g(\cdot)$ be a globally Lipschitz continuous and monotone nondecreasing activation function. Then*

$$\int_v^u [g_i(s) - g_i(v)]\mathrm{d}s \geq \frac{1}{2\bar{\ell}_i}[g_i(u) - g_i(v)]^2, \ \ \forall\, v, u \in R$$

and $i = 1, 2, \ldots, n$.

The RNN model (1) is said to be globally output convergent if, given any $x_0 \in R^n$, there exists a constant vector $x^* = (x_1^*, x_2^*, \ldots, x_n^*)^T$ such that

$$\lim_{t\to+\infty} g_i(x_i(t, x_0)) = g_i(x_i^*), \quad i = 1, 2, \ldots, n.$$

An $n \times n$ matrix A is said to be Lyapunov diagonally stable if there exists a diagonal matrix $P = \text{diag}(p_1, p_2, \ldots, p_n)$ with $p_i > 0$ such that $PA + A^T P < 0$.

In the sequel, $\lambda_{\min}(A)$ $[\lambda_{\max}(A)]$ denotes the minimum [maximum] eigenvalue of A. $\|\cdot\|$ is the norm of a vector or a matrix. $\|x\| = \sqrt{x^T x}$ for a vector x and $\|A\| = \sqrt{\lambda_{\max}(A^T A)}$ for a matrix A. Let $P = \text{diag}(p_1, p_2, \ldots, p_n)$ and $p_{\max} = \max\{p_1, p_2, \ldots, p_n\}$.

3. Output convergence analysis

We require Lemma 2 in order to establish our convergence analysis results

Lemma 2. *(Lemma 3 in Liu et al., 2004). If (3) is satisfied and there exists a constant vector $x^* \in R^n$ such that $Wg(x^*) + u = 0$, then, given any $x_0 \in R^n$, the RNN model (1) has a unique solution $x\ (t;\ x_0)$ defined on $[0, +\infty)$.*

Remark 1. According to Lemma 2, given any $x_0 \in R^n$, the RNN model (1) has a unique state solution $x\ (t;\ x_0)$ defined on $[0, +\infty)$ if (3) is satisfied and if there exists a constant vector $x^* \in R^n$ such that $Wg(x^*) + u = 0$. Let $\tilde{u}(t) = u(t) - u$ and $z = (z_1, z_2, \ldots, z_n)^T = x - x^*$. Then the RNN model (1) can be transformed to the following equivalent system

$$\frac{\mathrm{d}z}{\mathrm{d}t} = WF(z(t)) + \tilde{u}(t),\ z(0) = z_0 = x_0 - x^* \tag{4}$$

where $F(z) = (f_1(z_1), \ldots, f_n(z_n))^T = g(z + x^*) - g(x^*)$ is a globally Lipschitz continuous and monotone nondecreasing activation function and $F(0) = 0$.

Theorem 1. *Suppose that $\lim_{t\to+\infty} u_i(t) = u_i$, $i = 1, 2, \ldots, n$, and there exists a vector $x^* = (x_1^*, x_2^*, \ldots, x_n^*)^T \in R^n$ such that $Wg(x^*) + u = 0$. If there exist $p_i > 0$ $(i = 1, 2, \ldots, n)$ such* *that one of the following three conditions holds,*

$$p_i w_{ii} + \sum_{j=1, j\neq i}^{n} |w_{ij}| p_j < 0, \quad i = 1, 2, \ldots, n; \tag{5}$$

$$\underline{\ell}_j > 0 \text{ and } p_j w_{jj} + \sum_{i=1, i\neq j}^{n} p_i |w_{ij}| < 0, \quad j = 1, 2, \ldots, n; \tag{6}$$

$$\underline{\ell}_i > 0 \text{ and } p_i w_{ii} \underline{\ell}_i + \sum_{j=1, j\neq i}^{n} \frac{1}{2}(p_i \bar{\ell}_j |w_{ij}| + p_j \bar{\ell}_i |w_{ji}|) < 0,$$
$$i = 1, 2, \ldots, n; \tag{7}$$

then, the output vector of the RNN model (1) is globally convergent; that is, given any $x_0 \in R^n$,

$$\lim_{t\to+\infty} g_i(x_i(t, x_0)) = g_i(x_i^*), \quad i = 1, 2, \ldots, n.$$

Proof. See Appendix A.

□

Remark 2. The matrix W in (5) and (6) is said to satisfy the so-called row dominance condition and column dominance condition, respectively (Michel & Miller, 1977).

Remark 3. Global output convergence indicates only a unique output vector after convergence. It does not necessarily imply a unique state vector after state convergence. As shown in Case 2 and Case 3 of the proof of Theorem 1, the global output convergence results given in conditions (6) and (7) of Theorem 1 can also guarantee global *state* convergence because $\underline{\ell}_i > 0$ implies that $g(\cdot)$ is a strictly monotone increasing activation function. Thus, under the conditions of (6) or (7), the state vector of the RNN (1) will converge to a unique vector starting from any initial state.

Theorem 2. *Suppose that W is Lyapunov diagonally stable (i.e., there exists a diagonal matrix* $P = diag(p_1, \ldots, p_n)$ *with* $p_i > 0$ *such that* $W^T P + PW < 0$) and

$$\int_0^{+\infty} |u_i(\tau) - u_i| d\tau < +\infty, \quad i = 1, 2, \ldots, n.$$

If there exists a vector $x^* = (x_1^*, x_2^*, \ldots, x_n^*)^T \in R^n$ *such that* $Wg(x^*) + u = 0$*, then, the output vector of the RNN model (1) is globally convergent; that is, given any* $x_0 \in R^n$,

$$\lim_{t\to+\infty} g_i(x_i(t, x_0)) = g_i(x_i^*), \quad i = 1, 2, \ldots, n.$$

Proof. Based on the conditions in Theorem 2 and Remark 1, we only need to consider system (4). We first prove that $\|F(z(t))\|$ is bounded for $[0, +\infty)$. Since W is Lyapunov diagonally stable, there exists a diagonal matrix $P = \text{diag}(p_1, p_2, \ldots, p_n)$ with $p_i > 0$ such that $W^T P + PW < 0$. Consider a function

$$E(t) = \sum_{i=1}^{n} p_i \int_0^{z_i(t)} f_i(s) ds >= 0. \tag{8}$$

The time derivative of $E(t)$ along the trajectory $z(t)$ of system (4) gives

$$\begin{aligned}\frac{dE(t)}{dt} &= F(z)^T P(WF(z) + \tilde{u}(t)) \\ &= F(z)^T [PW]^S F(z) + F(z)^T P\tilde{u}(t) \\ &\le \lambda_{\max}([PW]^S)\|F(z)\|^2 \\ &\quad + p_{\max}\|F(z)\| \, \|\tilde{u}(t)\|, \ \forall t \in [0, +\infty),\end{aligned} \tag{9}$$

where $[PA]^S = (PA + A^T P)/2$. On the other hand, in terms of (8) and Lemma 1, $\forall t = [0, +\infty)$, we have

$$\begin{aligned}E(t) &= \sum_{i=1}^{n} p_i \int_0^{z_i(t)} f_i(s) ds \ge \sum_{i=1}^{n} \frac{p_i}{2\bar{\ell}_i} f_i^2(z_i(t)) \ge \xi \sum_{i=1}^{n} f_i^2(z_i(t)) \\ &= \xi \|F(z(t))\|^2;\end{aligned} \tag{10}$$

that is, $\|F(z(t))\| \le \sqrt{E(t)/\xi}$ where $\xi = \min_{1\le i\le n}\{p_i/(2\bar{\ell}_i)\}$. Then, from (9) and $[PW]^s < 0$ we get

$$\frac{dE(t)}{dt} \le p_{\max}\|F(z)\| \, \|\tilde{u}(t)\| \le p_{\max}\sqrt{E(t)/\xi}\|\tilde{u}(t)\|;$$

that is,

$$\frac{dE(t)}{\sqrt{E(t)}} \le \frac{p_{\max}}{\sqrt{\xi}} \|\tilde{u}(t)\| dt.$$

Taking integral on both sides yields

$$\begin{aligned}2(\sqrt{E(t)} - \sqrt{E(0)}) &\le \frac{p_{\max}}{\sqrt{\xi}} \int_0^t ||\tilde{u}(s)|| ds \\ &\le \frac{p_{\max}}{\sqrt{\xi}} \int_0^{+\infty} ||\tilde{u}(s)|| ds;\end{aligned}$$

that is

$$E(t) \le \left[\sqrt{E(0)} + \frac{p_{\max}}{2\sqrt{\xi}} \int_0^{+\infty} \|\tilde{u}(s)\| ds\right]^2$$

which is bounded. As a result, $\|F(z(t))\|$ is bounded for $t \in [0, +\infty)$ based on (10). Let

$$\|F(z(t))\| \le M, \ \forall t \in [0, +\infty), \tag{11}$$

where M is a positive number. Then from (9), we have $\forall t \in [0, +\infty)$,

$$\begin{aligned}\frac{dE(t)}{dt} &\le \lambda_{\max}([PW]^S)\|F(z)\|^2 + p_{\max}\|F(z)\| \, \|\tilde{u}(t)\| \\ &\le \lambda_{\max}([PW]^S)\|F(z)\|^2 + p_{\max}M\|\tilde{u}(t)\|.\end{aligned} \tag{12}$$

Now define another differentiable function

$$\phi(t) = [E(t) + 1]\exp\left(-p_{\max}M \int_0^t \|\tilde{u}(\tau)\| d\tau\right)$$

for all $t \geq 0$. One can see that

$$\phi(t) \geq \exp\left(-p_{\max} M \int_0^t \|\tilde{u}(\tau)\| \mathrm{d}\tau\right) > 0$$

for all $t \geq 0$. Then, using (9), we derive $\forall t \in [0, +\infty)$

$$\begin{aligned}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} &= -p_{\max} M \|\tilde{u}(t)\| \phi(t) \\ &\quad + \frac{\mathrm{d}E(t)}{\mathrm{d}t} \exp\left(-p_{\max} M \int_0^t \|\tilde{u}(\tau)\| \mathrm{d}\tau\right) \\ &\leq \left[-p_{\max} M \|\tilde{u}(t)\| + \frac{\mathrm{d}E(t)}{\mathrm{d}t}\right] \\ &\quad \times \exp\left(-p_{\max} M \int_0^t \|\tilde{u}(\tau)\| \mathrm{d}\tau\right) \\ &\leq [-p_{\max} M \|\tilde{u}(t)\| + \lambda_{\max}([PW]^S) \|F(z)\|^2 \\ &\quad + p_{\max} M \|\tilde{u}(t)\|] \exp\left(-p_{\max} M \int_0^t \|\tilde{u}(\tau)\| \mathrm{d}\tau\right) \\ &= \lambda_{\max}([PW]^S) \|F(z)\|^2 \exp\left(-p_{\max} M \int_0^t \|\tilde{u}(\tau)\| \mathrm{d}\tau\right) \\ &\leq \lambda_{\max}([PW]^S) \|F(z)\|^2 \exp\left(-p_{\max} M \int_0^{+\infty} \|\tilde{u}(\tau)\| \mathrm{d}\tau\right) \\ &\triangleq -M_1 \|F(z)\|^2 \end{aligned} \tag{13}$$

where

$$M_1 = -\lambda_{\max}([PW]^S) \exp\left(-p_{\max} M \int_0^{+\infty} \|\tilde{u}(\tau)\| \mathrm{d}\tau\right) > 0.$$

Next, we will show that $\|F(z(t))\|^2 \to 0$ as $t \to +\infty$. Suppose this is not true, then for some $\varepsilon > 0$ there exists a divergent sequence $\{t_k\}$ for which $\|F(z(t_k))\|^2 \geq 2\varepsilon$. Note that

$$D^+ \|F(z(t))\|^2 \leq 2 \sum_{i=1}^{n} f_i(z_i) \cdot D^+ f_i(z_i) \cdot \dot{z}_i(t). \tag{14}$$

On one hand, $f_i(z_i)$ is bounded by (11) and consequently $\dot{z}_i(t)$ is bounded by Eq. (4). On the other hand, $F(z)$ is a globally Lipschitz continuous and monotone nondecreasing activation function and consequently $D^+ f_i(\cdot)$ is bounded by (2). Thus, the right hand side of (14) is bounded and there exists a constant $\alpha > 0$ such that

$$|D^+ \|F(z(t))\|^2| \leq \alpha.$$

Thus, on the interval

$$t_k \leq t \leq t_k + \frac{\varepsilon}{\alpha}, \ (k = 1, 2, \ldots)$$

we have

$$\|F(z(t))\|^2 \geq \|F(z(t_k))\|^2 - \alpha |t - t_k| \geq 2\varepsilon - \varepsilon = \varepsilon.$$

The above intervals can be assumed to be disjoint by taking, if necessary, a subsequence of t_k. Then, it follows from (13) that

$$\phi\left(t_k + \frac{\varepsilon}{\alpha}\right) - \phi(t_k) \leq -M_1 \varepsilon^2 / \alpha, \quad k = 1, 2, \ldots. \tag{15}$$

Since $\mathrm{d}\phi(t)/\mathrm{d}t \leq 0$ for all $t \geq 0$ and the intervals $[t_k, t_k + \varepsilon/\alpha]$ $(k = 1, 2, \ldots)$ are disjoint with each other, then

$$\phi(t_k) \leq \phi\left(t_{k-1} + \frac{\varepsilon}{\alpha}\right), \quad k = 1, 2, \ldots. \tag{16}$$

From (15) and (16) it follows that

$$\phi\left(t_k + \frac{\varepsilon}{\alpha}\right) - \phi(t_1) \leq -M_1 \frac{\varepsilon^2}{\alpha} \cdot k \to -\infty$$

as $k \to +\infty$, which contradicts to $\phi(t) > 0$. This proves that $F(z(t)) \to 0$ as $t \to +\infty$. That is, $\lim_{t \to +\infty} g_i(x_i(t, x_0)) = g_i(x_i^*)$, $i = 1, 2, \ldots, n$. □

Remark 4. The RNN model (1) with constant inputs is a special case of the general CAM network (Grossberg, 1988). In this case, the state stability follows easily by noting (9), (A1), (A4) and (A6) in the proofs of Theorems 1 and 2 and the time-varying input $\tilde{u}(t) \equiv 0$. The symmetry condition (Grossberg, 1988) is required to guarantee that every state trajectory approaches one of the equilibrium points where the activation function of the network satisfies $g'_i(\cdot) \geq 0$. Theorem 7 in Hirsch (1989) is a state stability result by assuming that W is in lower (or upper) block triangular form and its diagonal blocks are symmetric where the activation function of the network satisfies $g'_i(\cdot) \geq 0$. However, W in our Theorems 1 and 2 may not be symmetric and the activation function may not be differentiable.

Remark 5. On one hand, the time-varying input in the form of $u_i(t) = \alpha_i \, \mathrm{e}^{-\beta_i t}$ used in Wang (1992, 1993, 1994, 1995, 1996, 1997) is a special case of $u_i(t)$ in our Theorems 1 and 2 where α_i, β_i are two design parameters and $\beta_i > 0$. On the other hand, one can easily check that the connection weight matrix W of the networks in Wang (1992, 1993, 1994, 1995, 1996, 1997) satisfy $W + W^T < 0$ if W are nonsingular. Hence, the results of Wang (1992, 1993, 1994, 1995, 1996, 1997) are special cases of the present Theorem 2 if the connection weight matrices are nonsingular.

4. Illustrative examples

Example 1. Consider the RNN model (1) where

$$W = \begin{bmatrix} -10 & 90 \\ 1 & -20 \end{bmatrix}, \quad u(t) = \begin{bmatrix} 1 + \dfrac{0.001}{t+1} \\ 1 + \dfrac{0.001}{t+1} \end{bmatrix},$$

and $g_i(x_i) = (|x_i + 1| - |x_i - 1|)/2$, $i = 1, 2$. Letting $p_1 = 1$ and $p_2 = 0.1$, we can check that W satisfies condition (5). On the other hand, one can see that there exists a vector $x^* = (x_1^*, x_2^*)^T \in R^2$ such that $Wg(x^*) + u = 0$ where $u = (1, 1)^T$. Based on Theorem 1 this network is globally output

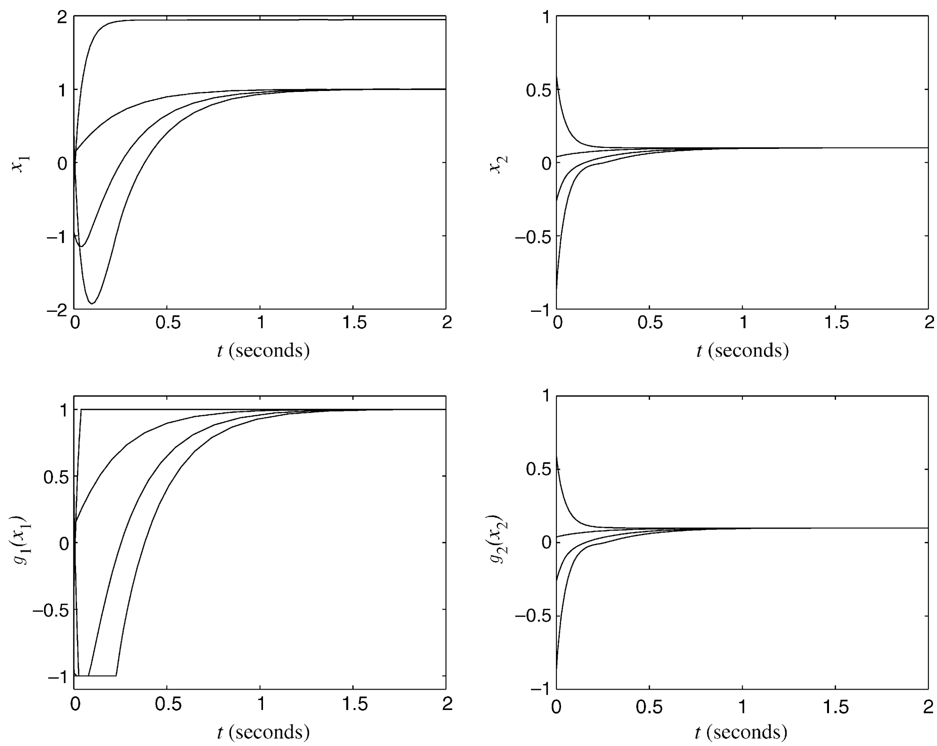

Fig. 1. Global output convergence of the network in Example 1.

convergent. Fig. 1 shows that outputs $(g_1(x_1),g_2(x_2))^T$ of this network are globally convergent to the unique point $(g_1(x_1^*), g_2(x_2^*))^T=(1,0.1)^T$ from four random initial points $x_0=(x_{10},x_{20})^T$ in the state space. Fig. 1 also shows that the steady states $(x_1, x_2)^T$ of the RNN are not unique but the steady output vector $(g_1(x_1), g_2(x_2))^T$ is unique. Since

$$\int_0^{+\infty} |u_i(\tau)-u_i|\mathrm{d}\tau = \int_0^{+\infty} \frac{0.001}{\tau+1}\mathrm{d}\tau = +\infty, \quad i=1,2,$$

Theorem 2 cannot be employed to guarantee the global output convergence of this network.

Example 2. Consider the RNN model (1) where

$$W=\begin{bmatrix} -1 & 3 \\ -1 & -1 \end{bmatrix}, \quad u(t)=\begin{bmatrix} 1+\dfrac{0.001}{t^2+1} \\ 2+\dfrac{0.001}{t^2+1} \end{bmatrix},$$

and $g_i(x_i)=\max\{0,x_i\}$, $i=1,2$. Letting $P=\mathrm{diag}(1,3)$, we can check that W is Lyapunov diagonally stable and

$$\int_0^{+\infty} |u_i(\tau)-u_i|\mathrm{d}\tau = \int_0^{+\infty} \frac{0.001}{\tau^2+1}\mathrm{d}\tau < +\infty, \quad i=1,2.$$

On the other hand, one can see that there exists a vector $x^*=(x_1^*,x_2^*)^T\in R^2$ such that $Wg(x^*)+u=0$ where $u=(1,2)^T$. Based on Theorem 2 this network is globally output convergent. However, noting that W does not satisfy condition (5) and $\underline{l}_i=0$, $i=1,2$, one can see that Theorem 1 cannot be employed to guarantee the global output convergence of this network.

5. Conclusions

In this paper, we have established global output convergence for a class of continuous-time recurrent neural networks with *globally* Lipschitz continuous and monotone nondecreasing activation functions and *locally* Lipschitz continuous time-varying inputs. Two sufficient conditions to guarantee global output convergence of this class of neural networks have been derived. Our results do not require symmetric connection weight matrix and our results extend existing results.

Appendix A. Proof of Theorem 1

Based on the conditions in Theorem 1 and Remark 1, we only need to consider system (4)

Case 1. Condition (5) is satisfied. Let $k=k(t)$ be the index such that $|p_k^{-1}f_k(z_k(t))|=\max_{1\le i\le n}|p_i^{-1}f_i(z_i(t))|$. Then, from (4) it follows that $\forall t\ge 0$,

$$\begin{aligned} D^+|z_k(t)| &\le w_{kk}|f_k(z_k(t))| + \sum_{j=1,j\ne k}^{n}|w_{kj}|\,|f_j(z_j(t))| + |\tilde{u}_k(t)| \\ &= p_k w_{kk}|p_k^{-1}f_k(z_k(t))| + \sum_{j=1,j\ne k}^{n} p_j|w_{kj}|\,|p_j^{-1}f_j(z_j(t))| \\ &\quad + |\tilde{u}_k(t)| \le \left(p_k w_{kk} + \sum_{j=1,j\ne k}^{n}|w_{kj}|p_j\right)|p_k^{-1}f_k(z_k(t))| \\ &\quad + |\tilde{u}_k(t)|, \end{aligned} \tag{A1}$$

$$D^+|f_k(t)| \le D^+ g_k(x_k)\left[\left(p_k w_{kk} + \sum_{j=1,j\ne k}^{n} |w_{kj}| p_j\right) \times |p_k^{-1} f_k(z_k(t))| + |\tilde{u}_k(t)|\right]. \tag{A2}$$

$$\text{Let } a_i = -\left[p_i w_{ii} + \sum_{j=1,j\ne i}^{n} |w_{ij}| p_j\right],\ i = 1,2,\dots,n. \tag{A3}$$

Then, from (5) one can see that $a_i > 0$, $i = 1,2,\dots,n$. Since $|\tilde{u}_i(t)| \to 0$ as $t \to +\infty$, $i = 1,2,\dots,n$, there exists $t_1 > 0$ such that $|\tilde{u}_i(t)| < \min_{1\le j\le n}(a_j)$, $\forall t \ge t_1$, $i = 1,2,\dots,n$.

We first prove that $|f_k(z_k(t))|$ is bounded on $[0,+\infty)$. Otherwise, assume that $|f_k(z_k(t))|$ is unbounded, then there must exist $t_3 > t_2 > t_1 > 0$ such that $|f_k(z_k(t_3))| > |f_k(z_k(t_2))| > p_{\max}$. Based on $|f_k(z_k(t_3))| > |f_k(z_k(t_2))|$, there must exist $t_4 \in [t_2, t_3]$ such that $|f_k(z_k(t_4))| \ge |f_k(z_k(t_2))|$ and $D^+|f_k(t_4)| > 0$. Then, from (A2) and $D^+ g_k(x_k) \ge 0$ we have

$$\begin{aligned}
D^+&|f_k(t_4)| \\
&\le D^+ g_k(x_k)\left[\left(p_k w_{kk} + \sum_{j=1,j\ne k}^{n} |w_{kj}| p_j\right)|p_k^{-1} f_k(z_k(t_4))| + |\tilde{u}_k(t_4)|\right] \\
&\le D^+ g_k(x_k)\left[\left(p_k w_{kk} + \sum_{j=1,j\ne k}^{n} |w_{kj}| p_j\right)|p_k^{-1} f_k(z_k(t_2))| + |\tilde{u}_k(t_4)|\right] \\
&\le D^+ g_k(x_k)\left[\left(p_k w_{kk} + \sum_{j=1,j\ne k}^{n} |w_{kj}| p_j\right)p_k^{-1} p_{\max} + |\tilde{u}_k(t_4)|\right] \\
&\le D^+ g_k(x_k)\left[\left(p_k w_{kk} + \sum_{j=1,j\ne k}^{n} |w_{kj}| p_j\right) + |\tilde{u}_k(t_4)|\right] \\
&\le D^+ g_k(x_k)\left[\left(p_k w_{kk} + \sum_{j=1,j\ne k}^{n} |w_{kj}| p_j\right) + \min_{1\le j\le n}(a_j)\right] \le 0
\end{aligned}$$

which contradicts $D^+|f_k(t_4)| > 0$. Therefore, $|f_k(z_k(t))|$ is bounded on $[0,+\infty)$.

Suppose that $\lim_{t\to+\infty} \sup|f_k(t)| = \delta_k$. Obviously, $0 \le \delta_k < +\infty$. Next, we will show that $\delta_k = 0$.

Assume that $\delta_k > 0$. Due to $|\tilde{u}_i(t)| \to 0$ as $t \to +\infty$, $i = 1,2,\dots,n$, there must exist a $\hat{t}_1 > 0$ such that $\forall t \ge \hat{t}_1$

$$|\tilde{u}_i(t)| \le \frac{a_i \delta_k}{4 p_{\max}}, \quad i = 1,2,\dots,n,$$

where a_i is defined in (A3). We choose a constant ε such that $0 < \varepsilon \le \delta_k/2$. We will prove that there exists a $\hat{t}_2 > \hat{t}_1$ such that $\forall t \ge \hat{t}_2$,

$$D^+|f_k(t)| \le 0.$$

If this is not true, there must exist a $\hat{t}_3 > \hat{t}_1$ such that

$$|f_k(\hat{t}_3)| \ge \delta_k - \varepsilon \text{ and } D^+|f_k(\hat{t}_3)| > 0.$$

In this case, from (A2) and $D^+ g_k(x_k) \ge 0$, it follows that

$$\begin{aligned}
D^+|f_k(\hat{t}_3)| &\le D^+ g_k(x_k)[-a_k |p_k^{-1} f_k(z_k(\hat{t}_3))| + |\tilde{u}_k(\hat{t}_3)|] \\
&\le D^+ g_k(x_k)\left[-a_k p_k^{-1}(\delta_k - \varepsilon) + \frac{a_k \delta_k p_{\max}^{-1}}{4}\right] \\
&\le D^+ g_k(x_k)\left[-a_k p_{\max}^{-1}\left(\delta_k - \frac{\delta_k}{2}\right) + \frac{a_k \delta_k p_{\max}^{-1}}{4}\right] \\
&= -\frac{a_k \delta_k p_{\max}^{-1}}{4} \cdot D^+ g_k(x_k) \le 0
\end{aligned}$$

which contradicts $D^+|f_k(\hat{t}_3)| > 0$. This shows that there exists a $\hat{t}_2 > \hat{t}_1$ such that $D^+|f_k(t)| \le 0$, $\forall t \ge \hat{t}_2$. Hence,

$$\lim_{t\to+\infty} |f_k(z_k(t))| = \lim_{t\to+\infty} \sup|f_k(z_k(t))| = \delta_k.$$

Then, one can see that there exists a $\hat{t}_4 > \hat{t}_2$ such that $\forall t \ge \hat{t}_4$,

$$|f_k(z_k(t))| > \delta_k - \varepsilon.$$

From (A1), we obtain $\forall t \ge \hat{t}_4$,

$$D^+|z_k(t)| \le -a_k |p_k^{-1} f_k(z_k(t))| + |\tilde{u}_k(t)| \le -a_k p_k^{-1}(\delta_k - \varepsilon) + \frac{a_k \delta_k p_{\max}^{-1}}{4} \le -\frac{a_k \delta_k p_{\max}^{-1}}{4} < 0.$$

As a result,

$$|z_k(t)| \le |z_k(\hat{t}_4)| - \frac{a_k \delta_k p_{\max}^{-1}}{4}(t - \hat{t}_4) \to -\infty$$

as $t \to +\infty$. This contradicts $|z_k(t)| \ge 0$. Thus, based on the above analysis, we have shown that $\delta_k = 0$ and

$$\lim_{t\to+\infty} |f_k(z_k(t))| = \lim_{t\to+\infty} \sup|f_k(z_k(t))| = 0.$$

Note that $|p_i^{-1} f_i(z_i(t))| \le |p_k^{-1} f_k(z_k(t))|$, $i = 1,2,\dots,n$, one can see that

$$\lim_{t\to+\infty} |f_i(z_i(t))| \le \lim_{t\to+\infty} p_i |p_k^{-1} f_k(z_k(t))| = 0, \quad i = 1,2,\dots,n.$$

This shows $\lim_{t\to+\infty} |f_i(z_i(t))| = 0$, or $\lim_{t\to+\infty} g_i(x_i(t, x_0)) = g_i(x_i^*)$, $i = 1,2,\dots,n$, which indicates that the output vector of the RNN model (1) is globally convergent.

Case 2. Condition (6) is satisfied. Define a function $V(z) = \sum_{i=1}^{n} p_i |z_i|$ and let

$$b_j \triangleq -\left(w_{jj} + \sum_{i=1,i\ne j}^{n} |w_{ij}| p_i / p_j\right) \underline{\ell}_j,\ j = 1,2,\dots,n.$$

Then, from (6) we have $b_j > 0$, $j = 1,2,\dots,n$. Computing the time derivative of $V(z)$ along the positive half trajectory of (4) yields

$$D^+V(z(t))$$
$$=\sum_{i=1}^{n}p_iD^+|z_i(t)|\le\sum_{i=1}^{n}p_iw_{ii}|f_i(z_i(t))|$$
$$+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}p_iw_{ij}|f_j(z_j(t))|+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$=\sum_{j=1}^{n}p_jw_{jj}|f_j(z_j(t))|+\sum_{j=1}^{n}\sum_{\substack{i=1\\ i\neq j}}^{n}p_i|w_{ij}|\,|f_j(z_j(t))|+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$=\sum_{j=1}^{n}\left(p_jw_{jj}+\sum_{i=1,i\neq j}^{n}p_i|w_{ij}|\right)|f_j(z_j(t))|+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$\le\sum_{j=1}^{n}\left(p_jw_{jj}+\sum_{i=1,i\neq j}^{n}p_i|w_{ij}|\right)\underline{\ell}_j|z_j(t)|+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$=\sum_{j=1}^{n}\left(w_{jj}+\sum_{i=1,i\neq j}^{n}|w_{ij}|p_i/p_j\right)\underline{\ell}_jp_j|z_j(t)|+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$=-\sum_{j=1}^{n}b_jp_j|z_j(t)|+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$\le-\min_{1\le j\le n}(b_j)\sum_{j=1}^{n}p_j|z_j(t)|+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$=-\min_{1\le j\le n}(b_j)V(z(t))+\sum_{i=1}^{n}p_i|\tilde{u}(t)|. \quad \text{(A4)}$$

Since $|\tilde{u}_i(t)|\to 0$ as $t\to+\infty$, $i=1,2,\ldots,n$, there exists $t_1>0$ such that

$$\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|<\min_{1\le j\le n}(b_j),\ \forall\, t\ge t_1.$$

We first prove that $V(z(t))$ is bounded on $[0,+\infty)$. Otherwise, assume that $V(z(t))$ is unbounded on $[0,+\infty)$. Then, there must exist $t_3>t_2>t_1>0$ such that $V(z(t_3))>V(z(t_2))>1$. Based on $V(z(t_3))>V(z(t_2))$, there must exist $t_4\in[t_2,t_3]$ such that $V(z(t_4))\ge V(z(t_2))>1$ and $D^+V(z(t_4))>0$. In this case, from (A4) it follows that

$$D^+V(z(t_4))\le-\min_{1\le j\le n}(b_j)V(z(t_4))+\sum_{i=1}^{n}p_i|\tilde{u}_i(t_4)|$$
$$<-\min_{1\le j\le n}(b_j)+\sum_{i=1}^{n}p_i|\tilde{u}_i(t_4)|<0$$

which contradicts $D^+V(z(t_4))>0$. Hence, $V(z(t))$ is bounded on $[0,+\infty)$.

Suppose that

$$\lim_{t\to+\infty}\sup V(z(t))=\delta.$$

Obviously, $0\le\delta<+\infty$. Next, we will show that $\delta=0$. Otherwise, assume that $\delta>0$. Due to $|\tilde{u}_i(t)|\to 0$ as $t\to+\infty$, $i=1,2,\ldots,n$, there must exist a $\hat{t}_1>0$ such that $\forall t\ge\hat{t}_1$,

$$\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|\le\frac{\delta\min_{1\le j\le n}(b_j)}{4}.$$

We choose a constant ε such that $0<\varepsilon\le\delta/2$. We will prove that there exists a $\hat{t}_2>\hat{t}_1$ such that $\forall t\ge\hat{t}_2$,

$$D^+V(z(t))\le 0.$$

If this is not true, there must exist a $\hat{t}_3>\hat{t}_1$ such that

$$V(z(\hat{t}_3))\ge\delta-\varepsilon \text{ and } D^+V(z(\hat{t}_3))>0.$$

In this case, from (A4), it follows that

$$D^+V(z(\hat{t}_3))\le-\min_{1\le j\le n}(b_j)V(z(\hat{t}_3))+\sum_{i=1}^{n}p_i|\tilde{u}_i(\hat{t}_3)|$$
$$\le-\min_{1\le j\le n}(b_j)(\delta-\varepsilon)+\sum_{i=1}^{n}p_i|\tilde{u}_i(\hat{t}_3)|$$
$$\le-\min_{1\le j\le n}(b_j)(\delta-\delta/2)+\sum_{i=1}^{n}p_i|\tilde{u}_i(\hat{t}_3)|$$
$$=-\frac{\delta\min_{1\le j\le n}(b_j)}{2}+\sum_{i=1}^{n}p_i|\tilde{u}_i(\hat{t}_3)|$$
$$\le-\frac{\delta\min_{1\le j\le n}(b_j)}{4}<0,$$

which contradicts $D^+V(z(\hat{t}_3))>0$. This shows that there exists a $\hat{t}_2>\hat{t}_1$ such that $\forall t\ge\hat{t}_2$, $D^+V(z(t))\le 0$. Hence,

$$\lim_{t\to+\infty}V(z(t))=\lim_{t\to+\infty}\sup V(z(t))=\delta.$$

Then, one can see that there exists a $\hat{t}_4>\hat{t}_2$ such that $\forall t\ge\hat{t}_4$,

$$V(z(t))\ge\delta-\varepsilon.$$

From (A4) we obtain $\forall t\ge\hat{t}_4$,

$$D^+V(z(t))\le-\min_{1\le j\le n}(b_j)V(z(t))+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$\le-\min_{1\le j\le n}(b_j)(\delta-\varepsilon)+\sum_{i=1}^{n}p_i|\tilde{u}_i(t)|$$
$$\le-\min_{1\le j\le n}(b_j)(\delta-\delta/2)+\delta\min_{1\le j\le n}(b_j)/4$$
$$=-\delta\min_{1\le j\le n}(b_j)/4<0.$$

As a result,

$$V(z(t))\le V(z(\hat{t}_4))-\frac{\delta\min_{1\le j\le n}(b_j)}{4}(t-\hat{t}_4)\to-\infty$$

as $t\to+\infty$. This contradicts $V(z(t))\ge 0$. Thus, based on the above analysis, we have shown that $\delta=0$ and $\lim_{t\to+\infty}V(z(t))=\lim_{t\to+\infty}\sup V(z(t))=0$. From $\lim_{t\to+\infty}V(z(t))=\lim_{t\to+\infty}\sum_{i=1}^{n}p_i|z_i(t)|=0$ it follows that $\lim_{t\to+\infty}z(t)=0$. Therefore, $\lim_{t\to+\infty}F(z(t))=0$; that is, $\lim_{t\to+\infty}g_i(x_i(t,x_0))=g_i(x_i^*)$, $i=1,2,\ldots,n$.

Case 3. Condition (7) is satisfied. Let the function $V(z)=(1/2)z^TPz$, where $P=\mathrm{diag}(p_1,\ldots,p_n)$. Computing the time derivative of $V(z)$ along the positive half trajectory of (4) yields

$$\begin{aligned}\frac{\mathrm{d}V(z)}{\mathrm{d}t} &= \sum_{i=1}^{n} p_i z_i \left(\sum_{j=1}^{n} w_{ij} f_j(z_j) + \tilde{u}_i(t)\right) \\ &= \sum_{i=1}^{n} p_i w_{ii} z_i f_i(z_i) + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} p_i w_{ij} z_i f_j(z_j) + \sum_{i=1}^{n} p_i z_i \tilde{u}_i(t) \\ &\le \sum_{i=1}^{n} p_i \underline{\ell}_i w_{ii} z_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} p_i \bar{\ell}_j |w_{ij}|\,|z_i|\,|z_j| \\ &\quad + \sum_{i=1}^{n} p_i |z_i|\,|\tilde{u}_i(t)| \\ &= \frac{1}{2}|z^T|(PA + A^TP)|z| + \sum_{i=1}^{n} p_i |z_i|\,|\tilde{u}_i(t)| \end{aligned} \tag{A5}$$

where $A=(a_{ij})_{n\times n}$, $a_{ii}=\underline{\ell}_i w_{ii}$, $i=1,2,\ldots,n$, $a_{ij}=\bar{\ell}_j|w_{ij}|$, $i\neq j$, and $|z|=(|z_1|,|z_2|,\ldots,|z_n|)^T$. By condition (7), we see that $PA+A^TP$ is negative definite. We choose a constant ξ such that

$$0<\xi\le\frac{\lambda_{\min}(-PA-A^TP)}{4p_{\max}}.$$

Then, from $V(z)=(1/2)z^TPz$ and (A5), it follows that

$$\begin{aligned}\frac{\mathrm{d}V(z(t))}{\mathrm{d}t} &\le -\lambda_{\min}(-PA-A^TP)\frac{1}{2}z^Tz + p_{\max}\sum_{i=1}^{n}|z_i|\,|\tilde{u}_i(t)| \\ &\le -\lambda_{\min}(-PA-A^TP)\frac{1}{2}z^Tz \\ &\quad + p_{\max}\sum_{i=1}^{n}\left(\xi z_i^2 + \frac{\tilde{u}_i^2(t)}{4\xi}\right) \\ &= [-\lambda_{\min}(-PA-A^TP) + 2p_{\max}\xi]\frac{1}{2}z^Tz \\ &\quad + \frac{p_{\max}}{4\xi}\|\tilde{u}(t)\|^2 \\ &\le -\frac{\lambda_{\min}(-PA-A^TP)}{2}\,\frac{1}{2}z^Tz + \frac{p_{\max}}{4\xi}\|\tilde{u}(t)\|^2 \\ &\le -\frac{\lambda_{\min}(-PA-A^TP)}{2p_{\max}}V(z(t)) + \frac{p_{\max}}{4\xi}\|\tilde{u}(t)\|^2. \end{aligned} \tag{A6}$$

Comparing (A6) with (A4), similar to Case 2, we can prove that $V(z(t))$ is bounded on $[0,+\infty)$ and

$$\lim_{t\to+\infty} V(z(t)) = \lim_{t\to+\infty}\sup V(z(t)) = 0.$$

From $\lim_{t\to+\infty} V(z(t)) = \lim_{t\to+\infty} z^T(t)Pz(t)/2 = 0$ it follows that $\lim_{t\to+\infty} z(t)=0$. Therefore, $\lim_{t\to+\infty} F(z(t))=0$; that is, $\lim_{t\to+\infty} g_i(x_i(t,x_0)) = g_i(x_i^*)$, $i=1,2,\ldots,n$.

References

Forti, M., & Tesi, A. (1995). New condition for global stability of neural networks with application to linear and quadratic programming problems. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, *42*, 354–366.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanism, and architectures. *Neural Networks*, *1*, 17–61.

Hirsch, M. W. (1989). Convergent activation dynamics in continuous time networks. *Neural Networks*, *2*, 331–349.

Hopfield, J. J. (1982). Neuronal networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA*, *79*, 2554–2558.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences USA*, *81*, 3088–3092.

Liang, X. B., & Si, J. (2001). Global exponential stability of neural networks with globally Lipschitz continuous activations and its application to linear variational inequality problem. *IEEE Transactions on Neural Networks*, *12*, 349–359.

Liang, X. B., & Wu, L. D. (1998). New sufficient conditions for absolute stability of neural networks. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, *45*, 584–586.

Liang, X. B., & Wu, L. D. (1999). Global exponential stability of a class of neural circuits. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, *46*, 748–751.

Liu, D., Hu, S., & Wang, J. (2004), Global output convergence of a class of continuous-time recurrent neural networks with time-varying thresholds. *IEEE Transactions on Circuits and Systems-II:* Express Briefs, 51, 161–167.

Michel, A. N., & Miller, R. K. (1977). *Qualitative analysis of large scale dynamical systems in the mathematical science and engineering*. New York: Academic Press.

Wang, J. (1992). Analog neural networks for solving the assignment problem. *Electronics Letters*, *28*, 1047–1050.

Wang, J. (1993). Analysis and design of a recurrent neural network for linear programming. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, *40*, 613–618.

Wang, J. (1994). A deterministic annealing neural network for convex programming. *Neural Networks*, *7*, 629–641.

Wang, J. (1995). Analysis and design of an analog sorting network. *IEEE Transactions on Neural Networks*, *6*, 962–971.

Wang, J. (1996). A recurrent neural network for solving the shortest path problem. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, *43*, 482–486.

Wang, J. (1997). Primal and dual assignment networks. *IEEE Transactions on Neural Networks*, *8*, 784–790.

Zhang, Y., Heng, P. A., & Fu, A. W. C. (1999). Estimate of exponential convergence rate and exponential stability for neural networks. *IEEE Transactions on Neural Networks*, *10*, 1487–1493.