Neurocomputing 73 (2010) 2783–2788

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

On the P-critical dynamics analysis of projection recurrent neural networks ☆

Chen Qiao *, Zongben Xu

Institute for Information and System Science, Faculty of Science, Xi'an Jiaotong University, Xi'an, China

ARTICLE INFO

Article history:
Received 3 September 2009
Received in revised form
22 December 2009
Accepted 4 April 2010
Communicated by W. Lu
Available online 4 May 2010

Keywords:
Dynamics analysis
Critical condition
Recurrent neural network
Nonlinear norm

ABSTRACT

Dynamics research of recurrent neural networks is very meaningful in both theoretical importance and practical significance. Recently, the study on the critical dynamics behaviors of such networks has drawn especial attention because of its application requirements. In this paper, new criteria are found to ascertain the global convergence and asymptotic stability of recurrent neural networks under the generally P-critical conditions, i.e., a discriminant matrix $M(L,\Gamma)+P$ is nonnegative definite, where $M(L,\Gamma)$ is a matrix related with the network and P is an arbitrary nonnegative definite matrix. The analysis results given in this paper improve substantially upon the existing relevant convergence and stability results in literature, including both the non-critical conclusions, i.e., the dynamics analysis under the conditions that $M(L,\Gamma)$ is positive definite, and the special critical discuss when $M(L,\Gamma)$ is nonnegative definite.

1. Introduction

Recurrent neural networks (RNNs) are mainly used to model dynamic process associated with control process, perform associative memory and solve optimization problems. The crucial foundation of the RNNs consists in their dynamical properties, such as the global convergence, asymptotic stability and exponential stability, therefore, the analysis of such dynamical behaviors is a first and necessary step for any practical design and application of RNNs. In recent years, considerable efforts have been devoted to the analysis on the stability of RNNs without and with delay (see, e.g. [2,6–9,13,17] and the references therein). For a given recurrent neural network, if we define

$$M(D,\Gamma)=D^{-1}\Gamma-\frac{\Gamma W+W^T\Gamma}{2},$$

where both D and Γ are positive definite diagonal matrices, and W is the weight matrix of the network, then by generalizing these existing stability results of RNNs, it should be noticed that most of them are on the exponential stability analysis under the conditions that for some positive definite diagonal matrix Γ, $M(\Lambda,\Gamma)$ is positive definite, where $\Lambda=\mathrm{diag}\{\lambda_1,\lambda_2,\ldots,\lambda_N\}$ with each $\lambda_i>0$ being the Lipschitz constant of g_i and $G=(g_1,g_2,\ldots,g_N)^T$ is the activation mapping of the network. On the other hand, [18,16] have proved that a RNN will be globally exponentially unstable if there is a positive definite diagonal matrix Γ, such that $M(V,\Gamma)$ is negative definite, where $V=\mathrm{diag}\{r_1,r_2,\ldots,r_N\}$ with each $r_i>0$ being the inversely Lipschitz constant of g_i, i.e., $|g_i(t)-g_i(s)|\ge r_i|t-s|$ for all $s,t\in\mathcal{R}^N$. By the definitions of Lipschitz constant and inversely Lipschitz constant, we have $r_i\le\lambda_i$ and, in the sense of nonnegative definition, the inequality relation: $M(\Lambda,\Gamma)\le M(V,\Gamma)$ holds. Since $M(\Lambda,\Gamma)>0$ is sufficient for the globally exponential stability of RNNs, and $M(V,\Gamma)\ge 0$ is necessary for RNNs to have globally stable dynamics, the questions then arise: what kinds of asymptotic behavior of RNNs will hold when $M(\Lambda,\Gamma)\le 0$ and $M(V,\Gamma)\ge 0$? The dynamics analysis of RNNs under such conditions is referred to as the *critical dynamics analysis*. Because a RNN is globally exponential stability when $M(\Lambda,\Gamma)>0$, recently, almost all of the critical dynamics investigations of RNNS are related to the special critical condition that $M(\Lambda,\Gamma)\ge 0$. While, it should be remarked that it is by no means easy to conduct a meaningful critical dynamics analysis since such analysis is much more difficult than the dynamics analysis under the non-critical condition that $M(\Lambda,\Gamma)>0$.

Up to now, there are only a few critical stability and convergence analysis of RNNs in the sense that $M(\Lambda,\Gamma)$ is nonnegative definite. For RNN with hyperbolic tangent activation function, in [14,2,3], the globally asymptotical stability and globally exponential stability of the unique equilibrium point of the network under some specific conditions of $M(\Lambda,\Gamma)\ge 0$ have been conducted. The authors of [26] have gotten the globally exponential stability of RNN with projection operator under the condition that $I-W$ is nonnegative definite (which is a special case of $M(\Lambda,\Gamma)\ge 0$). In [16], the authors have proved that a RNN with Sigmoidal activation mapping has a globally attractive equilibrium state, and when W is quasi-symmetric (i.e., there exists a positive definite diagonal matrix D, such that DW is

☆This research was supported by the National Nature Science Foundation of China (no. 70531030) and the National Basic Research Program of China (973 Program) (no. 2007CB311002).

* Corresponding author.
E-mail addresses: qiaochen@mail.xjtu.edu.cn (C. Qiao), zbxu@mail.xjtu.edu.cn (Z. Xu).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.04.008

symmetric), then RNN with nearest point projection activation mapping is global convergence on a region defined by the network. The quasi-symmetric requirement of W in [16] has been removed in [19,20]. For all that, there are still many important dynamics questions of RNNs unsettled under the critical conditions. For example, with what additional requirement will the Sigmoidal RNNs be globally exponential stability when $M(\Lambda,\Gamma)\geq 0$? For a RNN with general projection mapping, does there exist some convergence results when $M(\Lambda,\Gamma)<0$? Further, for neural network with general activation mapping, what asymptotic behaviors of it will be under the critical conditions that $M(\Lambda,\Gamma)\leq 0$ and $M(V,\Gamma)\geq 0$? All these are under our current investigation.

In the current paper, we devote to answer the question that what dynamics behavior will happen for a RNN with general projection mapping when $M(\Lambda,\Gamma)<0$. By using Lyapunov functional method and topological degree theory, it is shown that a RNN with general projection mapping is globally convergent under the condition that $M(\Lambda,\Gamma)+P\geq 0$ (here P is an arbitrary nonnegative definite matrix) if the nonlinear norm defined by the network is bounded. The dynamics analysis of RNNs under such a condition is called as the *P-critical dynamics analysis*. The *P*-critical convergence and asymptotical stability of both static and local field RNNs (which are two fundamental modeling approaches in RNNs research) are established. The results obtained here extend, to a large extent, most of the existing noncritical conclusions and the latest critical results given by [16,26,20]. Furthermore, they provide a wider application range of RNNs and can be applied directly to many concrete RNN models.

2. Model description and preliminaries

Static RNNs and local field RNNs typically represent two fundamental modeling approaches in current neural network research [25], which are, respectively, modeled by

$$\tau\frac{dx}{dt}=-x+G(Wx+q),\quad x(0)=x_0\in\mathcal{R}^N \tag{1}$$

and

$$\tau\frac{dy}{dt}=-y+WG(y)+q,\quad y(0)=y_0\in\mathcal{R}^N \tag{2}$$

where $x=(x_1,x_2,\ldots,x_N)^T$ is the neural state vector, $y=(y_1,y_2,\ldots,y_N)^T$ is the local field vector, $W=(\omega_{ij})_{N\times N}$ is the synaptic weight matrix, τ is a positive constant, q is a fixed external bias vector and $G:\mathcal{R}^N\to\mathcal{R}^N$ is the nonlinear activation mapping.

We now recall some notion and notations (taking system (1) as an example). A constant vector x^* is said to be an equilibrium state of system (1), if x^* is a zero point of the mapping $F_e(x):=-x+G(Wx+q)$, $\forall x\in\mathcal{R}^N$. x^* is said to be *stable* if any trajectory of system (1) can stay within a small neighborhood of x^* whenever the initial state x_0 is close to x^*, and it is said to be *attractive* if there is a neighborhood $\Delta(x^*)$, called the *attraction basin* of x^*, such that any trajectory of system (1) initialized from a state in $\Delta(x^*)$ will approach to x^* as time goes to infinity. x^* is said to be *globally asymptotically stable on* $\Delta(x^*)$ if it is both stable and attractive, with the attraction basin $\Delta(x^*)$. System (1) is said to be *globally convergent on* Θ if for every initial point $x_0\in\Theta$, $x(t,x_0)$ converges to an equilibrium state of system (1) (the limit of $x(t,x_0)$ may not be the same for different x_0).

For the nonlinear activation mapping $G:\mathcal{R}^N\to\mathcal{R}^N$, denote the *range* and the *fixed-point-set* of G, respectively, by $\mathbf{R}(G)$ and $\mathbf{F}(G)$. G is said to be *diagonally nonlinear* if G is defined componentwisely by $G(x)=(g_1(x_1),g_2(x_2),\ldots,g_N(x_N))^T$, where each g_i is a one-dimensional nonlinear function; G is said to be a *projection mapping* if $G\circ G=G$, or equivalently, $\mathbf{R}(G)\subseteq\mathbf{F}(G)$; G is said to be a *diagonal projection* if it is diagonally nonlinear, and each component function g_i is a one-dimensional projection; G is said to be a *nearest point projection* if there is a bounded, closed and convex subset $\Omega\subset\mathbf{R}^N$ such that $G(x)=\operatorname{argmin}_{z\in\Omega}\|x-z\|$. Obviously, nearest point projection is a special kind of projection mapping, but the inverse is not necessarily true. In the present investigation, we assume g_i is Lipschitz continuous. L_i, the *minimum Lipschitz constant* of g_i, is defined as follows:

$$L_i=\sup_{t,s\in\mathbf{R},t\neq s}\frac{|g_i(t)-g_i(s)|}{|t-s|}. \tag{3}$$

Without loss of generality, through out this paper, we assume that each $L_i>0$ and let $L=\operatorname{diag}\{L_1,L_2,\ldots,L_N\}$.

In what follows, we will give the definition of the nonlinear norm, which is similar to that of the matrix norm. Suppose that $T:\Psi\subseteq\mathcal{R}^N\to Y\subseteq\mathcal{R}^N$ is a nonlinear mapping, A is a nonsingular $N\times N$ matrix, and $\tilde{x}\in\Psi$ is a given vector. Define

$$L_{\|\cdot\|}(T,A,\tilde{x},\Psi)=\sup_{x\neq\tilde{x},x\in\Psi}\frac{\|ATx-AT\tilde{x}\|}{\|Ax-A\tilde{x}\|}. \tag{4}$$

Clearly, $L_{\|\cdot\|}(T,A,\tilde{x},\Psi)$ is a nonnegative function determined by five parameters: T, A, $\tilde{x}$, $\|\cdot\|$ and Ψ. Most important of all, $L_{\|\cdot\|}(T,A,\tilde{x},\Psi)$ can be regarded as a nonlinear generalization of the matrix norm $\|\cdot\|$ and called as the *nonlinear norm*. This is because, by defining $F=ATA^{-1}$, $y=Ax$, $\tilde{y}=A\tilde{x}$ and $\tilde{\Psi}=A\Psi$, one can get

$$L_{\|\cdot\|}(T,A,\tilde{x},\Psi)=\sup_{x\neq\tilde{x},x\in\Psi}\frac{\|ATx-AT\tilde{x}\|}{\|Ax-A\tilde{x}\|}$$
$$=\sup_{x\neq\tilde{x},x\in\Psi}\frac{\|ATA^{-1}(Ax)-ATA^{-1}(A\tilde{x})\|}{\|Ax-A\tilde{x}\|}=\sup_{y\neq\tilde{y},y\in\tilde{\Psi}}\frac{\|Fy-F\tilde{y}\|}{\|y-\tilde{y}\|}.$$

Obviously, for any given matrix B, $L_{\|\cdot\|}(B,I,\tilde{x},\Psi)=\|B\|$. Additionally, for any constant $\beta>0$, we have $L_{\|\cdot\|}(\beta T,A,\tilde{x},\Psi)=\beta L_{\|\cdot\|}(T,A,\tilde{x},\Psi)$.

3. *P*-critical global convergence results

In this section, under the *P*-critical condition that $M(L,\Gamma)+P\geq 0$, where P is an arbitrary nonnegative definite matrix, the global convergence and asymptotic stability theorem and corollaries for projection RNNs of both systems (1) and (2) will be established. We consider the networks of form (1) first.

Suppose that the nonlinear activation mapping G is continuous and $\mathbf{R}(G)$ is bounded, closed and convex. For any $v\in\mathbf{R}(G)$, define $T(v)=G(Wv+q)$, then by Brouwer's fixed point theorem, T has at least one fixed point v^*, namely, $F_e^{-1}(0)$, the equilibrium state set of (1) is not empty.

Since for any $N\times N$ positive definite matrix A, there exists an orthogonal matrix K ($K^TK=I$) such that $K^TAK=\operatorname{diag}\{\lambda_1,\lambda_2,\ldots,\lambda_N\}$ (here $\lambda_i>0$ is the eigenvalue of A), so if we define $B=K\operatorname{diag}\{\lambda_1^{1/2},\lambda_2^{1/2},\ldots,\lambda_N^{1/2}\}K^T$, then it is clear that $B=B^T$, $B^2=A$ and B is invertible. Such a matrix B is denoted by $Deco(A)$, i.e., $B=Deco(A)$.

Theorem 3.1. *Let $G:\mathcal{R}^N\to\Theta$ be a diagonal projection and each g_i be monotonically increasing and continuous, Θ be a bounded, closed and convex subset of $\mathcal{R}^N$. If there is a positive definite diagonal matrix Γ and a nonnegative definite matrix P, such that $M(L,\Gamma)+P\geq 0$, and for a $v^*\in F_e^{-1}(0)$, $L_{\|\cdot\|_2}(T,Q,v^*,\Theta)\leq 1$ (here $Q=Deco(L^{-1}\Gamma+P)$), then RNN system (1) is globally convergent on Θ when $F_e^{-1}(0)$ is disconnected. Furthermore, when v^* is the unique equilibrium point of (1), then v^* is globally asymptotically stable on Θ.*

Proof. For any trajectory $x(t)$ of (1) starting from $x_0\in\Theta$, it follows from Lemma 1 in [19] that $x(t)\in\Theta$. Let $y_0=Wx_0+q$, $y(t)=Wx(t)+q$,

$z(t)=G(y(t))$ and $u(t)=z(t)-x(t)$. Define

$V(x(t))=x^T(t)(L^{-1}\Gamma+P)x(t)-2x^T(t)(L^{-1}\Gamma+P)v^*$,

then we have

$$\tau\frac{dV(x(t))}{dt}=2\langle x(t),(L^{-1}\Gamma+P)(z(t)-x(t))\rangle-2\langle (L^{-1}\Gamma+P)v^*,z(t)-x(t)\rangle$$
$$=2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle.$$

Define

$$E_1(x(t))=\tau x^T(t)\left(\frac{(L^{-1}+I)\Gamma}{2}-(\Gamma W+W^T\Gamma)\right)x(t)-2\tau x^T(t)\Gamma q$$
$$+\tau\sum_{i=1}^N\xi_i\int_{(y_0)_i}^{y_i(t)}g_i(s)\,ds-\tau\left\{\frac{1}{2}(y(t)-x(t))^T\Gamma(y(t)-x(t))\right.$$
$$\left.-\frac{1}{2}y^T(t)\Gamma y(t)-\frac{1}{2}x^T(t)(I-L^{-1})\Gamma x(t)\right\}+\tau V(x(t)).$$

Then, a direct calculation is that

$$\frac{dE_1(x(t))}{dt}=2\left\langle\left(\frac{(L^{-1}+I)\Gamma}{2}-(\Gamma W+W^T\Gamma)\right)x(t),u(t)\right\rangle-2\langle\Gamma q,u(t)\rangle$$
$$+\langle L^{-1}\Gamma z(t),u(t)\rangle-\langle\Gamma z(t),(L^{-1}-W)u(t)\rangle$$
$$-\langle\Gamma x(t),(L^{-1}-W)u(t)\rangle+\langle\Gamma y(t),u(t)\rangle$$
$$+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle$$
$$=-2\langle\Gamma y(t),u(t)\rangle+\langle\Gamma x(t),u(t)\rangle+\langle L^{-1}\Gamma x(t),u(t)\rangle$$
$$-2\langle\Gamma x(t),Wu(t)\rangle+\langle L^{-1}\Gamma z(t),u(t)\rangle$$
$$-\langle\Gamma z(t),(L^{-1}-W)\cdot u(t)\rangle-\langle\Gamma x(t),(L^{-1}-W)u(t)\rangle$$
$$+\langle\Gamma y(t),u(t)\rangle+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle$$
$$-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle$$
$$=-\langle\Gamma(y(t)-x(t)),u(t)\rangle-\langle L^{-1}\Gamma x(t),u(t)\rangle$$
$$+\langle\Gamma(x(t)-z(t)),(L^{-1}-W)u(t)\rangle+\langle L^{-1}\Gamma z(t),u(t)\rangle$$
$$+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle$$
$$=-\langle\Gamma(y(t)-x(t)),u(t)\rangle+\langle\Gamma(z(t)-x(t)),W(z(t)-x(t))\rangle$$
$$+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle$$
$$=-\langle\Gamma(y(t)-x(t)),u(t)\rangle+\langle z(t)-x(t),(\Gamma W+W^T\Gamma)(z(t)-x(t))\rangle$$
$$+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle$$
$$=-\langle\Gamma(y(t)-x(t)),u(t)\rangle-u^T(t)\left((L^{-1}\Gamma+P)-\frac{\Gamma W+W^T\Gamma}{2}\right)u(t)$$
$$+u^T(t)(L^{-1}\Gamma+P)u(t)+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle$$
$$-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle. \quad (5)$$

Since $L^{-1}\Gamma+P-(\Gamma W+W^T\Gamma)/2\geq 0$, so we have

$$\frac{dE_1(x(t))}{dt}\leq-\langle\Gamma(y(t)-x(t)),u(t)\rangle+u^T(t)(L^{-1}\Gamma+P)u(t)$$
$$+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle. \quad (6)$$

On the other hand, from the assumption that v^* is one fixed point of the mapping $T(v):=G(Wv+q)$, we have $v^*=G(Wv^*+q)$, and then

$$u^T(t)(L^{-1}\Gamma+P)u(t)+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle$$
$$=\langle u(t),(L^{-1}\Gamma+P)u(t)\rangle+2\langle (L^{-1}\Gamma+P)x(t),u(t)\rangle$$
$$-2\langle (L^{-1}\Gamma+P)v^*,u(t)\rangle$$
$$=\langle z(t)+x(t),(L^{-1}\Gamma+P)u(t)\rangle-2\langle v^*,(L^{-1}\Gamma+P)u(t)\rangle$$
$$=\langle (z(t)-v^*)+(x(t)-v^*),(L^{-1}\Gamma+P)u(t)\rangle$$
$$=\langle (z(t)-v^*)+(x(t)-v^*),(L^{-1}\Gamma+P)((z(t)-v^*)-(x(t)-v^*))\rangle$$
$$=\langle (z(t)-G(Wv^*+q))+(x(t)-v^*),(L^{-1}\Gamma+P)((z(t)$$
$$-G(Wv^*+q))-(x(t)-v^*))\rangle$$
$$=\langle z(t)-G(Wv^*+q),(L^{-1}\Gamma+P)(z(t)-G(Wv^*+q))\rangle$$
$$-\langle x(t)-v^*,(L^{-1}\Gamma+P)(x(t)-v^*)\rangle$$
$$=\langle G(Wx(t)+q)-G(Wv^*+q),(L^{-1}\Gamma+P)(G(Wx(t)+q)$$
$$-G(Wv^*+q))\rangle-\langle x(t)-v^*,(L^{-1}\Gamma+P)(x(t)-v^*)\rangle$$
$$=\langle Q(G(Wx(t)+q)-G(Wv^*+q)),Q(G(Wx(t)+q)-G(Wv^*+q))\rangle$$
$$-\langle Q(x(t)-v^*),Q(x(t)-v^*)\rangle$$
$$=\|Q(G(Wx(t)+q)-G(Wv^*+q))\|_2^2-\|Q(x(t)-v^*)\|_2^2. \quad (7)$$

Furthermore, on noting that $L_{\|\cdot\|_2}(T,Q,v^*,\Theta)\leq 1$, one can get that

$$L_{\|\cdot\|_2}(T,Q,v^*,\Theta)=\sup_{x\neq v^*,x\in\Theta}\frac{\|QTx-QTv^*\|_2}{\|Qx-Qv^*\|_2}$$
$$=\sup_{x\neq v^*,x\in\Theta}\frac{\|QG(Wx+q)-QG(Wv^*+q)\|_2}{\|Qx-Qv^*\|_2}\leq 1,$$

this, combined with (6) and (7), implies that

$$\frac{dE_1(x(t))}{dt}\leq-\langle\Gamma(y(t)-x(t)),u(t)\rangle$$
$$=-\sum_{i=1}^N\xi_i[y_i(t)-x_i(t)]\times[g_i(y_i(t))-x_i(t)]. \quad (8)$$

In addition, since $x(t)\in\Theta=\mathbf{R}(G)$ and G is a projection, then $x(t)\in\mathbf{F}(G)$ and (8) implies

$$\frac{dE_1(x(t))}{dt}\leq-\sum_{i=1}^N\xi_i[y_i(t)-x_i(t)]\times[g_i(y_i(t))-g_i(x_i(t))]. \quad (9)$$

Since g_i is monotonically increasing with Lipschitz constant L_i, it is easy to verify that whether $y_i(t)\geq x_i(t)$ or $y_i(t)<x_i(t)$, the following inequality always holds:

$$[y_i(t)-x_i(t)]\times[g_i(y_i(t))-g_i(x_i(t))]\geq L_i^{-1}(g_i(y_i(t))-g_i(x_i(t)))^2$$
$$=L_i^{-1}(g_i((Wx(t)+q)_i)-g_i(x_i(t)))^2. \quad (10)$$

Denote by $\lambda_{min}(L^{-1}\Gamma)$ the smallest eigenvalue of the diagonal matrix $L^{-1}\Gamma$, then by (9) and (10), we get that

$$\frac{dE_1(x(t))}{dt}\leq-\sum_{i=1}^N\xi_iL_i^{-1}(g_i((Wx(t)+q)_i)-g_i(x_i(t)))^2$$
$$\leq-\lambda_{min}(L^{-1}\Gamma)\|G(Wx(t)+q)-x(t)\|_2^2, \quad (11)$$

from the fact that $L_i>0$ and Γ is positive definite, it can be deduced that $dE_1(x(t))/dt\leq 0$, and the equal sign holds if and only if $G(Wx(t)+q)=x(t)$, i.e., $x(t)\in F_e^{-1}(0)$. Furthermore, since $x(t)\in\Theta$ is bounded and $F_e^{-1}(0)$ is disconnected, then by LaSalle invariance principle [11], we know that RNN model (1) is globally convergent on Θ. Furthermore, when $F_e^{-1}(0)=\{v^*\}$, then it is easy to deduce that v^* is both attractive and stable on Θ since Θ is bounded, i.e., v^* is globally asymptotically stable on Θ. Thus, Theorem 3.1 is proved. □

Corollary 3.1. *Assume that $\Theta\subseteq\mathcal{R}^N$ is a bounded, closed and convex set, $G:\mathcal{R}^N\to\Theta$ is a diagonal projection with each g_i being monotonically increasing and continuous. If there exists a positive definite diagonal matrix Γ and a nonnegative definite diagonal matrix P, such that $M(L,\Gamma)+P\geq 0$, and $\|QLWQ^{-1}\|_2\leq 1$ (here $Q=Deco(L^{-1}\Gamma+P)$), then RNN model (1) is globally convergent on Θ when $F_e^{-1}(0)$ is disconnected. Moreover, when v^* is the unique equilibrium point of (1), then v^* is globally asymptotically stable on Θ.*

Proof. By the definition of L,Γ,P and Q, we know that Q is a positive definite diagonal matrix. Let $Q=\text{diag}\{q_1,q_2,\ldots,q_N\}$ with each $q_i>0$. For any $x,v^*\in\Theta$, since G is a diagonal projection and each L_i is the minimum Lipschitz constant of g_i, it can be deduced that

$$\|QTx-QTv^*\|_2^2=\|QG(Wx+q)-QG(Wv^*+q)\|_2^2$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} (q_i(g_i((Wx+q)_i) - g_i((Wv^*+q)_i)))^2 \\
&\le \sum_{i=1}^{N} (q_i L_i (W(x-x^*))_i)^2 \\
&= \sum_{i=1}^{N} (q_i L_i)^2 \left(\sum_{j=1}^{N} W_{ij} q_j^{-1} (q_j(x_j - v_j^*)) \right)^2 \\
&= \sum_{i=1}^{N} \left(\sum_{j=1}^{N} q_i L_i W_{ij} q_j^{-1} (q_j(x_j - v_j^*)) \right)^2 \\
&= \|QLWQ^{-1}(Q(x-v^*))\|_2^2. \qquad (12)
\end{aligned}
$$

Clearly, when $\|QLWQ^{-1}\|_2 \le 1$, then by (12), it can be deduced that

$$L_{\|\cdot\|_2}(T,Q,v^*,\Theta) = \sup_{x \ne v^*, x \in \Theta} \frac{\|QTx - QTv^*\|_2}{\|Qx - Qv^*\|_2} \le 1.$$

Corollary 3.1 is then proved from Theorem 3.1. □

Correspondingly, we have the critical global convergence and asymptotical stability conclusions for RNN system (2).

Corollary 3.2. *Suppose that $G: \mathcal{R}^N \to \Theta$ is a diagonal projection with each g_i being monotonically increasing and continuous, Θ is a bounded, closed and convex subset of $\mathcal{R}^N$, and the equilibrium state set of (2) is disconnected. Then, under the following condition (S), RNN system (2) is globally convergent on $W(\Theta)+q$. Moreover, when y^* is the unique equilibrium state of (2), then y^* is globally asymptotically stable on $W(\Theta)+q$.*

(S) There is a positive definite diagonal matrix Γ and a symmetric nonnegative definite matrix P, such that $M(L,\Gamma)+P \ge 0$ and either $L_{\|\cdot\|_2}(T,Q,v^*,\Theta) \le 1$ for a $v^* \in \mathbf{F}(T)$ (here $Q = Deco(L^{-1}\Gamma + P)$), or $\|QLWQ^{-1}\|_2 \le 1$ when P is a diagonal matrix.

Proof. For any trajectory $y(t)$ of (2) starting from $y_0 \in W(\Theta)+q$, let $y_0 = Wx_0+q$ with $x_0 \in \Theta$ and $x(t)$ be the solution of (1) with initial value $x(0)=x_0$, then by the uniqueness of solution of differential equations, it is easy to verify that $y(t) \equiv Wx(t)+q$ and there exists an equilibrium state of (1), denoted by v^*, such that $y^* = Wv^*+q$. Thus, by [25], the convergence of $y(t)$ to an equilibrium state of (2) can be shown by studying the asymptotic behavior of $x(t)$. Then, the conclusion of Corollary 3.2 readily follows from Theorem 3.1 and Corollary 3.1. □

Remark 3.1. The results achieved above exploit new global convergence and asymptotical stability of both static RNNs and local field RNNs, and can generalize or extend most of the known non-critical as well as critical dynamics results (e.g., the non-critical results summarized in [18], the critical analysis of [16,26,19,20]). In addition, on noting that the nearest point projection is a special case of the general projection mapping, the convergence and stability conclusions established in Theorem 3.1 and Corollaries 3.1 and 3.2 can be directly applied to RNNs with nearest point projection.

Applying the obtained generic dynamics analysis results of RNNs directly to the linear variational inequality problem (LVIP) [24,13], and several typical recurrent neural network models, such as the brain-state-in-a-box/domain recurrent neural networks (BSB RNNs) [12,23], the BCOp-type RNNs [1,8] and the cellular neural networks (CNNs) [4,15], we can get some useful solutions for solving LVIP and some new criteria for convergence as well as stability of the BSB RNNs, BCOp-type RNNs and the CNNs. The new results achieved for solving LVIP can extend the existing results in [24,10,13], and the conclusions for the three RNN models mention above can unify and improve further the results related in [4,5,12,18,21,22,23,24,26].

4. Illustrative examples

In this section, we provide an illustrative example to demonstrate the validity of the critical convergence and stability results formulated in the previous section.

Example 4.1. Consider the following RNN:

$$
\begin{cases}
\dfrac{dx_1(t)}{dt} = -x_1(t) + g_1(2x_1(t) + x_2(t) + 4(\sqrt{2}-1)), \\
\dfrac{dx_2(t)}{dt} = -x_2(t) + g_2(-2x_1(t) + 2x_2(t) + x_3(t) + 6), \\
\dfrac{dx_3(t)}{dt} = -x_3(t) + g_3(-2x_2(t) + 2x_3(t) + x_4(t) - 6), \\
\dfrac{dx_4(t)}{dt} = -x_4(t) + g_4(-2x_3(t) + 2x_4(t) + x_5(t) + 6), \\
\dfrac{dx_5(t)}{dt} = -x_5(t) + g_5(-2x_4(t) + 2x_5(t) + x_6(t) + 6), \\
\dfrac{dx_6(t)}{dt} = -x_6(t) + g_6(-2x_5(t) + 2x_6(t) - 5),
\end{cases}
\qquad (13)
$$

where each $g_i(s) = \frac{1}{2}(|s+1| - |s-1|)$ $(i=1,2,\ldots,6)$.

In this example, $L_i = 1$, the unique equilibrium state is $x^* = (1,1,-1,1,1,-1)^T$.

For any positive definite diagonal matrix Γ, it is easy to verify that $L^{-1}\Gamma - (\Gamma W + W^T\Gamma)/2$ is not positive definite, and further, not nonnegative definite. That is, all of the non-critical and critical conclusions in literature (see, e.g., [18,16,20]) cannot be used here. But we will show Theorem 3.1 established in Section 3 can be applied to this example. By choosing $\Gamma = \mathrm{diag}\{2,1,1/2,1/4,1/8,1/16\}$ and $P = \mathrm{diag}\{2,1,1/2,1/4,1/8,1/16\}$, we have $L^{-1}\Gamma - (\Gamma W + W^T\Gamma)/2 + P \ge 0$ and $Q = Deco(L^{-1}\Gamma + P) = \mathrm{diag}\{2, \sqrt{2}, 1, \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{2\sqrt{2}}\}$. Letting $T(v) = G(Wv+q)$ for any $v \in \Theta = [-1,1]^6$, one can get that $\mathbf{F}(T) = \{(1,1,-1,1,1,-1)^T\}$. For any $v^* \in \mathbf{F}(T)$, we want to show that

$$L_{\|\cdot\|_2}(T,Q,v^*,\Theta) = \sup_{v \ne v^*, v \in \Theta} \frac{\|QG(Wv+q) - Qv^*\|_2}{\|Qv - Qv^*\|_2} \le 1. \qquad (14)$$

On noting that for any $v \in \Theta$, $\sum_{k=1}^6 w_{2k}v_k + q_2 \ge 1$, so by the definition of g_2, we have $g_2(\sum_{k=1}^6 w_{2k}v_k + q_2) = 1$. Similarly, one can easily get that $g_3(\sum_{k=1}^6 w_{3k}v_k + q_3) = -1$, $g_4(\sum_{k=1}^6 w_{4k}v_k$

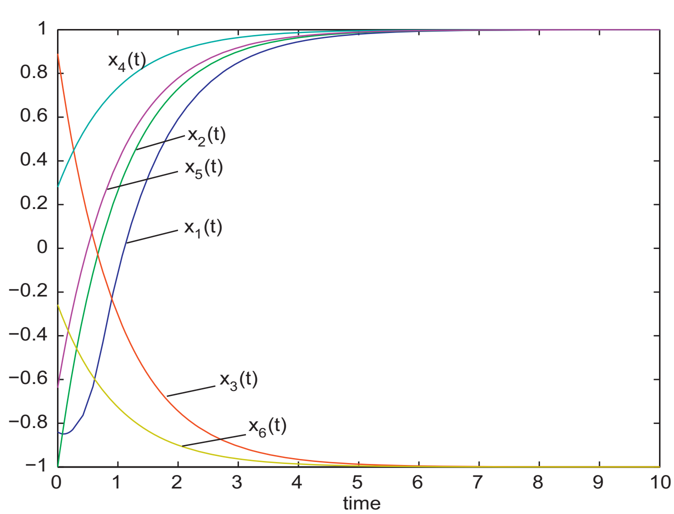

Fig. 1. Transient behaviors of RNN in system (13) with a random initial point $x_0 \in \Theta$.

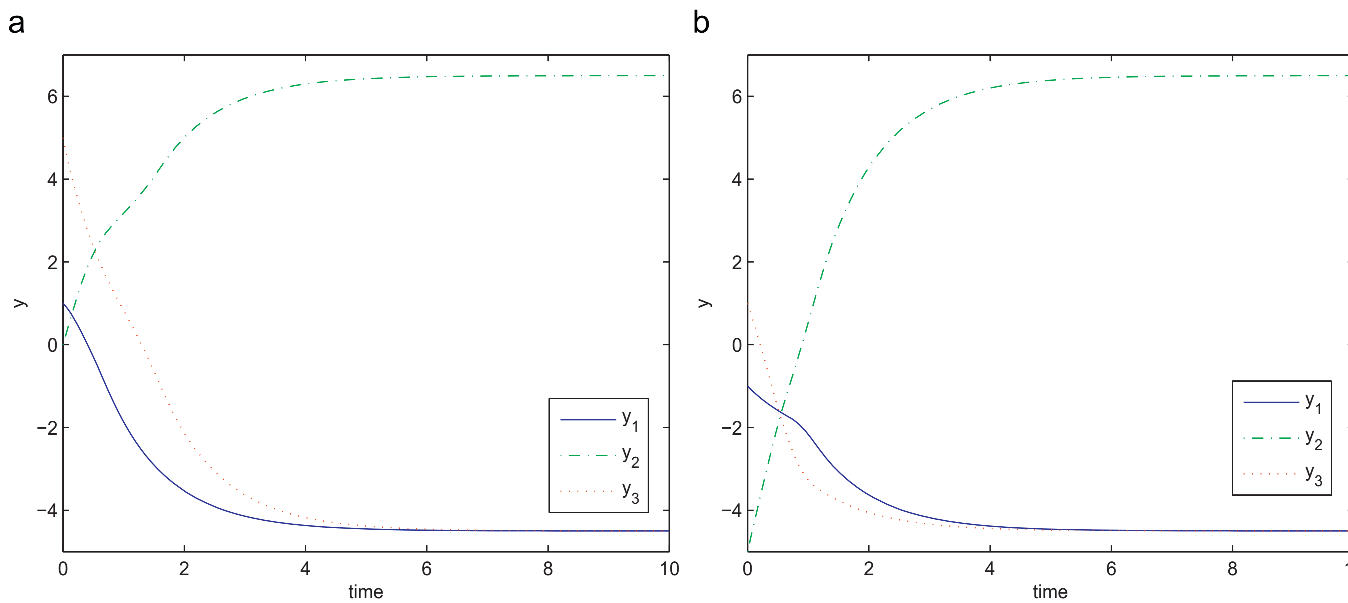

Fig. 2. Transient behaviors of RNN in system (15) with random initial points $y_0 \in W(\Theta)+q$.

$+q_4)=1$, $g_5(\sum_{k=1}^{6} w_{5k}v_k+q_5)=1$ and $g_6(\sum_{k=1}^{6} w_{6k}v_k+q_6)=-1$. Thus,

$$\|QG(Wv+q)-Qv^*\|_2^2 = 4\left(g_1\left(\sum_{k=1}^{6} w_{1k}v_k+q_1\right)-v_1^*\right)^2 = 4(1-g_1(2v_1+v_2+4(\sqrt{2}-1)))^2.$$

On the other hand,

$$\|Qv-Qv^*\|_2^2 \ge 4(1-v_1)^2+2(1-v_2)^2 \ge (2(1-v_1)+\sqrt{2}(1-v_2))^2/2.$$

If for any $v\in\Theta=[-1,1]^6$, $0\le 2(1-g_1(2v_1+v_2+4(\sqrt{2}-1)))\le(2(1-v_1)+\sqrt{2}(1-v_2))/\sqrt{2}$ always holds, then $L_{\|\cdot\|_2}(T,Q,v^*,\Theta)\le 1$. Which equals to prove $g_1(2t+w+4\sqrt{2}-3)\ge \frac{1}{\sqrt{2}}t+\frac{1}{2}w$, where $-2\le t\le 0$ and $0\le w\le 2$. For the three cases that $2t+w+4\sqrt{2}-3\ge 1$, $-1\le 2t+w+4\sqrt{2}-3\le 1$ and $2t+w+4\sqrt{2}-3\le -1$, it is easy to verify that $g_1(2t+w+4\sqrt{2}-3)\ge\frac{1}{\sqrt{2}}t+\frac{1}{2}w$ is always true. According to Theorem 3.1, system (13) is globally asymptotically stable on Θ. Fig. 1 depicts the time responses of state variables of the system with random initial point starting from Θ, which confirm that the proposed condition in Theorem 3.1 ensures the globally asymptotical stability of the RNNs.

Example 4.2. Consider a RNN defined by

$$\begin{cases} \dfrac{dy_1(t)}{dt} = -y_1(t)+1.5g_1(y_1(t))-g_2(y_2(t))-2, \\ \dfrac{dy_2(t)}{dt} = -y_2(t)+g_1(y_1(t))+1.5g_2(y_2(t))-g_3(y_3(t))+5, \\ \dfrac{dy_3(t)}{dt} = -y_3(t)+g_2(y_2(t))+1.5g_3(y_3(t))-4, \end{cases} \quad (15)$$

where $g_i(s)$ $(i=1,2,3)$ is defined as

$$g_i(s)=\begin{cases} 1, & s>1, \\ s, & -1\le s\le 1, \\ -1, & s<-1. \end{cases} \quad (16)$$

Obviously, this example takes the form of local field model, with

$$W=\begin{pmatrix} 1.5 & -1 & 0 \\ 1 & 1.5 & -1 \\ 0 & 1 & 1.5 \end{pmatrix}$$

and $q=(-2,5,-4)^T$. In this case, $L_i=1$ $(i=1,2,3)$ and the equilibrium state set of the system is $\{(-4.5,6.5,-4.5)^T\}$. For any positive definite diagonal matrix Γ, $M(L,\Gamma)=L^{-1}\Gamma-(\Gamma W+W^T\Gamma)/2$ is not nonnegative definite. By choosing $\Gamma=I$ and $P=0.5I$, one has $L^{-1}\Gamma-(\Gamma W+W^T\Gamma)/2+P\ge 0$, and, similarly as the proof in Example 4.1, we have $L_{\|\cdot\|_2}(T,Q,v^*,\Theta)\le 1$ for $v^*\in \mathbf{F}(T)=\{(-1,1,-1)^T\}$ (here $Q=Deco(L^{-1}\Gamma+P)=\sqrt{1.5}I$ and $\Theta=[-1,1]^3$). Thus, by Corollary 3.2, it follows that the unique equilibrium state of RNN (15), say, $y^*=(-4.5,6.5,-4.5)^T$, is asymptotically stable on $W(\Theta)+q$. All simulation results show that the trajectories starting from $\Psi := W(\Theta)+q$ are within the region Ψ, and convergent to y^*. The transient behaviors of $y(t)$ with random initial points starting from $W(\Theta)+q$ are depicted in Fig. 2.

5. Conclusion

Two basic dynamics behaviors: global convergence and asymptotical stability of both static and local field RNNs have been studied under the P-critical condition. It has been proved that when the nonlinear norm determined by the network is bounded, then RNN with general projection mapping possesses convergent and stable properties in the sense that, a discriminant matrix $M(L,\Gamma)+P$ is nonnegative definite, where $M(L,\Gamma)$ is a matrix related with the network, P is an arbitrary nonnegative definite matrix. Compared with the known dynamics analysis, our results extend the latest critical analysis results and generalize most of the non-critical conclusions developed in literature. The theory obtained here can be directly applied to some typical RNNs and problems, such as the brain-state-in-a-box/domain recurrent neural networks, the cellular neural networks, the BCOp-type RNNs as well as solving linear variational inequality problem. The significance of the results obtained here not only lies in providing some farther cognizance on the essentially dynamical behavior of RNNs, but also in enlarging the application field of them.

References

[1] A. Bouzerdoum, T.R. Pattison, Neural network for quadratic optimization with bound constraints, IEEE Trans. Neural Networks 4 (1993) 293–303.
[2] T.P. Chen, S.I. Amari, New theorems on global convergence of some dynamical systems, Neural Networks 14 (2001) 251–255.

[3] T.P. Chen, Global convergence of delayed dynamical systems, IEEE Trans. Neural Networks 12 (2001) 1532–1536.
[4] L.O. Chua, L. Yang, Cellular neural networks: theory and applications, IEEE Trans. Circuits Syst. 35 (1988) 1257–1290.
[5] L.O. Chua, T. Roska, Cellular Neural Networks and Visual Computing: Foundations and Applications, Cambridge University Press, UK, 2002.
[6] P.V.D. Driessche, X.F. Zou, Global attractivity in delayed Hopfield neural networks, SIAM J. Appl. Math. 58 (1998) 1878–1890.
[7] Y. Fang, T.G. Kincaid, Stability analysis of dynamical neural networks, IEEE Trans. Neural Networks 7 (1996) 996–1006.
[8] M. Forti, A. Tesi, New conditions for global stability of neural networks with applications to linear and quadratic programming problems, IEEE Trans. Circuits Syst. 42 (1995) 354–366.
[9] Z.H. Guan, G. Chen, Y. Qin, On equilibria, stability and instability of Hopfield neural networks, IEEE Trans. Neural Networks 11 (2000) 534–540.
[10] X.L. Hu, J. Wang, Design of general projection neural networks for solving monotone linear variational inequalities and linear and quadratic optimization problems, IEEE Trans. Syst. Man Cybern. 37 (2007) 1414–1421.
[11] J.P. LaSalle, The Stability of Dynamical Systems, SIAM, Philadelphia, PA, 1976.
[12] J. Li, A.N. Michel, W. Porod, Analysis and synthesis of a class of neural networks: linear systems operating on a closed hypercube, IEEE Trans. Circuits Syst. 36 (1989) 1406–1422.
[13] X.B. Liang, J. Si, Global exponential stability of neural networks with globally Lipschitz continuous activation and its application to linear variational inequality problem, IEEE Trans. Neural Networks 12 (2001) 349–359.
[14] X.W. Liu, T.P. Chen, A new result on the global convergence of Hopfield neural networks, IEEE Trans. Circuits Syst. 49 (2002) 1514–1516.
[15] J. Park, H.Y. Kim, Y. Park, S.W. Lee, A synthesis procedure for associative memories based on space-varying cellular neural networks, Neural Networks 14 (2001) 107–113.
[16] J. Peng, Z.B. Xu, H. Qiao, B. Zhang, A critical analysis on global convergence of Hopfield-type neural networks, IEEE Trans. Circuits Syst. 52 (2005) 804–814.
[17] H. Qiao, J. Peng, Z.B. Xu, Nonlinear measures: a new approach to exponential stability analysis for Hopfield-type neural networks, IEEE Trans. Neural Networks 12 (2001) 360–370.
[18] H. Qiao, J. Peng, Z.B. Xu, B. Zhang, A reference model approach to stability analysis of neural networks, IEEE Trans. Syst. Man Cybern. 33 (2003) 925–936.
[19] C. Qiao, Z.B. Xu, New critical analysis on global convergence of recurrent neural networks with projection mappings, ISNN 07, Lecture Notes in Computer Science, vol. 4493, Springer, Berlin, 2007, pp. 131–139.
[20] C. Qiao, Z.B. Xu, A critical global convergence analysis of recurrent neural networks with general projection mappings, Neurocomputing 72 (2009) 1878–1886.
[21] T. Roska, J. Vandewalle, Cellular Neural Networks, Wiley, Chichester, UK, 1995.
[22] A. Slavova, Cellular Neural Networks: Dynamics and Modelling, Kluwer Academic Publishers Pub. Dordrecht, 2003.
[23] I. Varga, G. Elek, H. Zak, On the brain-state-in-a-convex-domain neural models, Neural Networks 9 (1996) 1173–1184.
[24] Y.S. Xia, J. Wang, On the stability of globally projected dynamical systems, J. Optim. Theory Appl. 106 (2000) 129–150.
[25] Z.B. Xu, H. Qiao, J. Peng, B. Zhang, A comparative study of two modeling approaches in neural networks, Neural Networks 17 (2004) 73–85.
[26] Y.Q. Yang, J. Cao, Solving quadratic programming problems by delayed projection neural network, IEEE Trans. Neural Networks 17 (2006) 1630–1634.

Chen Qiao was born in Shaanxi Province, China. She received the B.S. degree in computational mathematics in 1999 and the M.S. degree in applied mathematics in 2002 from Xi'an Jiaotong University, Xi'an, China, where she is currently working towards the Ph.D. degree in applied mathematics. From 2005 until now, she is a lecturer in the Institute for Information and System Sciences, Xi'an Jiaotong University. Her current research interests include nonlinear systems, stability theory and neural networks.

Zong-Ben Xu received the M.S. degree in mathematics in 1981 and the Ph.D. degree in applied mathematics in 1987 from Xi'an Jiaotong University, China. In 1998, he was a postdoctoral researcher in the Department of Mathematics, The University of Strathclyde, United Kingdom. He worked as a research fellow in the Information Engineering Department from February 1992 to March 1994, the Center for Environmental Studies from April 1995 to August 1995, and the Mechanical Engineering and Automation Department from September 1996 to October 1996, at The Chinese University of Hong Kong. From January 1995 to April 1995, he was a research fellow in the Department of Computing in The Hong Kong Polytechnic University. He has been with the Faculty of Science and Institute for Information and System Sciences at Xi'an Jiaotong University since 1982, where he was promoted to associate professor in 1987 and full professor in 1991, and now serves as an authorized Ph.D. supervisor in mathematics and computer science, director of the Institute for Information and System Sciences, and Vice President of Xi'an Jiaotong University. Currently, he serves as Chief Scientist of one National Basic Research Project of China (973 Project). He has published four monographs and more than 134 academic papers on nonlinear functional analysis, optimization techniques, neural networks, evolutionary computation, and data mining algorithms, most of which are in international journals. His current research interests include nonlinear functional analysis, mathematical foundation of information technology, and computational intelligence. Dr. Xu holds the title "Owner of Chinese PhD Degree Having Outstanding Achievements" awarded by the Chinese State Education Commission (CSEC) and the Academic Degree Commission of the Chinese Council in 1991. He was awarded the first prize of the Science and Technology Award by the Ministry of Education, China (2006). He is a member of the New York Academy of Sciences.