Neurocomputing 106 (2013) 158–166

Contents lists available at SciVerse ScienceDirect

Neurocomputing

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

The UPPAM continuous-time RNN model and its critical dynamics study ☆

Chen Qiao*, Wenfeng Jing, Zongben Xu [1]

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

ARTICLE INFO

Article history:
Received 26 April 2012
Received in revised form
11 October 2012
Accepted 25 October 2012
Communicated by Long Cheng
Available online 30 November 2012

Keywords:
Recurrent neural network
Uniformly pseudo-projection-anti-monotone RNNs model
Critical dynamics analysis
Continuous-time

ABSTRACT

There have been numerous recurrent neural network models diversely developed for modeling or simulating the associative memory behavior of human beings in the past decades, and the existing results for each model individual are in certain sense redundant with similarity. By utilizing the innate character of general activation operators, i.e., the uniformly pseudo-projection-anti-monotone property, a unified continuous-time recurrent neural network model is introduced, which can jointly cover almost all of the known continuous-time recurrent neural network individuals. Under the critical condition which is the intrinsic bounded line of stability and instability, we develop some convergence and stability theory for the unified recurrent neural network model when the time is continuous. The study shows that the approach adopted in the present paper is powerful, particularly in the sense of unifying, simplifying and extending the currently existing various models and dynamics results of continuous-time RNNs.

1. Introduction

Recurrent neural networks (RNNs) are generally formalized as dynamic systems which can be implemented by physical means. They are assumed to capture the associative memory performance of human beings and then widely used to model dynamic process associated with control process, perform pattern recognition, solve optimization problems, and so on.

A RNN is associated with two fundamental operators: one is the synaptic connections among the neurons, and the other is the nonlinear activation functions deduced from the input-output properties of the involved neurons. The synaptic connections among the neurons are hopefully used to encode the memories we expect to have; and the activation functions of a RNN are assumed to capture the complex, nonlinear response of neurons of the brain, which are preassigned before use in general, depending on the simulation purpose or application. Once the synaptic connections are given, the characteristics of the activation functions determine the performance of the RNN. As a rule, the activation functions are monotonically nondecreasing and saturated, as suggested from neurobiology [12,24]. However, the monotonicity and boundedness of the activation functions do not sufficiently result in the expected properties of a RNN in general, so usually, various other assumptions and specifications need to be assumed additionally in the study and application of RNNs. To discover some more essential characteristics else than the nondecreasing and bounded properties of the commonly used activation functions, Xu and Qiao put forward two novel concepts: uniformly anti-monotone as well as the pseudo-projection properties by carefully examining the properties of various activation functions used in general [23]. When the number of neurons is fixed as N and each activation function is denoted by g_i $(i=1,2,\ldots,N)$, the corresponding activation operator $G=(g_1,g_2,\ldots,g_N)^T$ is called as the uniformly pseudo-projection-anti-monotone (UPPAM) operator [23]. As a framework of formalizing the activation operators, the UPPAM operator embodies most of those meaningful activation operators, e.g., signum operator, symmetric multi-valued step operator, multi-threshold operator, linear saturating operator, nearest-point projection, winner-take-all operator, etc. Results in a natural, a unified RNN model which is called as the UPPAM neural networks is presented in that paper. Consequently, it is proved that most of the typical concrete RNN models are UPPAM RNNs. For instance, the Hopfield-type neural networks [24,25], the bidirectional associative memory models [26], the Little-Hopfield neural networks [12], the recurrent back-propagation model [27], the multi-valued associative memories [28], the multi-threshold neural networks [29], the brain-state-in-a-box model [7], the bound-constraints optimization solvers, the convex optimization solvers [10], the winner-take-all automata [30], the mean-field model [31], the recurrent correlation associative memories [32], and so on. Further, a unified stability theory for the discrete-time UPPAM RNNs were developed in [23]. But, it should be noticed that in that paper,

* Corresponding author. Tel.: +86 29 82660949.
E-mail addresses: qiaochen@mail.xjtu.edu.cn (C. Qiao), wfjing@mail.xjtu.edu.cn (W. Jing), zbxu@mail.xjtu.edu.cn (Z. Xu).
[1] Tel.: +86 29 82668005.
☆ This research was supported by the National Nature Science Foundation of China (No. 11101327 and No. 11171270) and the Fundamental Research Funds for the Central Universities (No. xjj20100087 and No. 2011jdhz30).

0925-2312/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.10.026

the stability for the UPPAM RNNs was given only when the time is discrete (i.e., the RNNs were described by difference equation forms) and, the synaptic connections are symmetric. As we know, due to the easy VLSI implementation of the continuous-time RNNs, most of the existing study and applications for recurrent neural networks were on the continuous-time case, and the synaptic connections of RNNs in practice are asymmetric and very complex.

In this presented paper, we will consider the continuous-time UPPAM RNNs model, i.e., the activation operators of the continuous-time RNNs own the uniformly pseudo-projection-anti-monotone property. Such RNNs model can be generally described as

$$\tau\frac{dx(t)}{dt} = -x(t)+AG(Wx(t)+q)+b, \quad x_0 \in \mathbb{R}^N \tag{1}$$

where $x(t)=(x_1(t),x_2(t),\ldots,x_N(t))^T$ is the neural network state, $G=(g_1,g_2,\ldots,g_N)^T$ is the nonlinear activation operator deduced from all the activation functions g_i. Here, we consider G has the UPPAM property (see the next section for the precise definition). A and W both are the connective weight matrices, b,q are two fixed external bias vectors and τ is the state feedback coefficient. Since the UPPAM operator embodies most of those meaningful activation operators as mentioned above, in fact, model (1) summarizes almost all of the existing continuous-time RNNs specials, e.g., Hopfield-type neural networks, brain-state-in-a-box neural networks, bound-constraints optimization solvers, recurrent back-propagation neural networks, mean-field neural networks, convex optimization solvers, recurrent correlation associative memories neural networks, cellular neural networks, etc. Thus, analysis of the dynamics behaviors for most of the known concrete continuous-time RNNs individuals can be replaced by studying the dynamics behaviors for model (1).

For the continuous-time RNNs, we should notice that there exist two basic problems needed to be solved at present. On the one hand, we know that the analysis of the dynamical behaviors, such as the global convergence, asymptotic stability and exponential stability, is a crucial foundation for any practical design and application of them. While, for all those but not limited to those commonly used continuous-time RNN individuals, the fact is that there exist huge numbers of researches on the dynamics behaviors for each model (see, e.g. [8,14,15,17–21,36,45] and the references therein). However, those known studies have been conducted in a very separative way, and there exist lots of similarity and redundant among those results. All these prompt us to look for a more universal methodology and formalize a unified approach to jointly cover all those known diverse dynamics results.

On the other hand, by summarizing the existing convergence as well as stability for most of the continuous-time recurrent neural network individuals, Peng, Xu, etc., pointed out that the continuous-time RNN models are exponential stable under the conditions that one discriminant matrix defined by the networks is positive definite [1]. For the UPPAM RNNs model, if we define

$$S(\Gamma,P)=\Gamma P-\frac{\Gamma AW+(\Gamma AW)^T}{2}$$

where both Γ and P are diagonal matrices, and A and W are the connective weight matrices, then similar to the proof of Theorem 1 in [1], it can be generalized that most of the exponential stability analysis for UPPAM RNNs individuals are under the conditions that there exists a positive definite diagonal matrix Γ, such that $S(\Gamma,(2\Lambda-B))$ is positive definite (where Λ and B are two diagonal matrices given by the UPPAM RNNs model). On the other hand, from [9,15], we know that a RNN will be globally exponentially unstable if there is a positive definite diagonal matrix Γ such that $S(\Gamma,V^{-1})$ is negative definite, where $V=diag\{r_1,r_2,\ldots,r_N\}$ with each $r_i>0$ being the inversely Lipschitz constant of activation function f_i. The questions then arise: since $S(\Gamma,(2\Lambda-B))>0$ (i.e., $S(\Gamma,(2\Lambda-B))$ is positive definite) is sufficient for the globally exponential stability of UPPAM RNNs, and $S(\Gamma,V^{-1})\geq 0$ (i.e., $S(\Gamma,V^{-1})$ is nonnegative definite) is necessary for UPPAM RNNs to have globally stable dynamics, then what kinds of asymptotic behavior of UPPAM RNNs will hold when $S(\Gamma,(2\Lambda-B))\leq 0$ (i.e., $S(\Gamma,(2\Lambda-B))$ is negative semi-definite) and $S(\Gamma,V^{-1})\geq 0$ (i.e., $S(\Gamma,V^{-1})$ is nonnegative definite)? If there exists a diagonal matrix Q, such that

$$S(\Gamma,(2\Lambda-B))\leq S(\Gamma,Q)\leq S(\Gamma,V^{-1})$$

(here we use $X\leq Y$ to denote the condition that matrix $Y-X$ is nonnegative definite), then in particular, we want to know what will happen when $S(\Gamma,Q)=0$ (i.e., for any $x\in\mathcal{R}^N$, $x^TS(\Gamma,Q)x=0$). The dynamics analysis of RNNs under such conditions is referred to as the *critical dynamics analysis*. It should be remarked that it is by no means easy to conduct a meaningful critical dynamics study for RNNs since such exploration has essential difficult in analysis, and there exist hardly any results about this topic.

In comparison to the general critical condition that $S(\Gamma,Q)=0$, $S(\Gamma,(2\Lambda-B))=0$ is the primary case of it, and since a UPPAM RNN is globally exponential stability when $S(\Gamma,(2\Lambda-B))>0$, so in this paper, we focus our attentions on the dynamics investigations of UPPAM RNNs under the particular critical condition that $S(\Gamma,(2\Lambda-B))\geq 0$. Even so, it is still much more difficult than the dynamics analysis under the noncritical condition that $S(\Gamma,(2\Lambda-B))>0$.

Since the critical condition is the essential bounded line which can distinguish between stability and instability of continuous-time RNNs [5], the critical dynamics study has drawn special attention in recent years. For continuous-time RNNs with hyperbolic tangent activation functions, in [6,8,21], the globally asymptotical stability and globally exponential stability of the network under some certain specific critical conditions have been conducted. Yang and Cao [22] have gotten the globally exponential stability of a static neural network with projection operator under the condition that $I-W$ is nonnegative. In [1], the authors have proved that RNN with sigmoidal activation operator has a globally attractive equilibrium state, and when the synaptic connection matrix is quasi-symmetric, RNN with nearest point projection activation operator is global convergence on a region defined by the network. Some improvements on [1] have been made in [2], but it is with one requirement on the trajectories of the network, which is hard to be verified in the application. In [3,4], the critical global convergence and asymptotical stability for continuous-time RNNs with projection activation operators have been achieved. Some critical globally exponential stability of the continuous-time RNNs when activation operator owns the decreasing anti-monotone property has been studied in [5]. For all that, there are still many important dynamics questions of RNNs unsettled under the critical conditions. For example, for a continuous-time RNN with general activation operator, what asymptotic behaviors of it will be under the critical conditions? If the study on the asymptotic behaviors of such RNN could be achieved, then would the dynamics analysis for it unify and generalize the corresponding theory validated for the individual models, or further, make some new discoveries? Answering such questions are of great importance in both theory and applications.

Based on the essential characteristics of the activation operators, i.e., uniformly pseudo-projection anti-monotone properties, we devote to establish a unified continuous-time RNNs model and its critical dynamics theory in this paper. This work can answer these two problems which are put forward above, i.e., it can integrate and generalize the known stability results for almost all of the known individual models under the noncritical conditions, and, most important, it can give the general determinate method of critical stability for continuous-time RNNs, and which is only determined by the UPPAM constants and the connection matrix of the networks. The investigation lunches a visible step towards establishing a unified method for the continuous-time RNNs, and the obtained results cover and generalize almost all of the known corresponding results for continuous-time RNN individuals, which shows the powerfulness of the suggested unified model and the approach.

2. Preliminaries

In this paper, associated with the activation operator G, the domain, range and fixed-point set of G are, respectively, defined by $\mathbf{D}(G)$, $\mathbf{R}(G)$ and $\mathbf{F}(G)$. For studying RNN purpose, we only consider $\mathbf{D}(G)=\mathbf{R}(G)\subseteq\mathbb{R}^N$ case below. We further assume that $\mathbb{R}^N$ is embedded with Euclidean norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$.

For any $x=(x_1,x_2,\ldots,x_N)^T\in\mathbf{D}(G)$, write

$$G(x)=(g_1(x),g_2(x),\ldots,g_N(x))^T,\quad \forall x\in\mathbf{D}(G)$$

G is said to be diagonal if $g_i(x)=g_i(x_i)$ holds for each $i=1,2,\ldots,N$. We further introduce some useful definitions of G as follows.

Definition 2.1 (*Xu and Qiao [23]*). (i) A operator G is said to be a pseudo-projection if there exists a positive definite diagonal matrix $B=diag\{\beta_1,\beta_2,\ldots,\beta_N\}$, such that $B\mathbf{R}(G)\subseteq\mathbf{D}(G)$ and $G=GBG$ (i.e., $G(x)=G(BG(x))$, $\forall x\in\mathbf{D}(G)$). In this case, we say that G is a B-projection.

(ii) A operator G is said to be λ-uniformly anti-monotone (λ-UAM) if there is a positive constant λ such that for any $x\in\mathbf{D}(G)$ and $y\in B\mathbf{R}(G)$,

$$\langle G(x)-G(y),x-y\rangle\geq\lambda\|G(x)-G(y)\|_2^2 \tag{2}$$

(iii) A operator G is uniformly pseudo-projection-anti-monotone (UPPAM) if it is pseudo-projection and uniformly anti-monotone; specially, we say it is (B,λ)-UPPAM whenever it is B-projection and λ-UAM.

In [23], it is shown that most of the activation operators deduced from the concrete activation operators diversely appeared in RNNs literature all are the special cases of UPPAM operators, and then, the UPPAM operator provides a very appropriate, unified framework within which most of the known RNN models can be embedded and uniformly studied.

Definition 2.2. Let $\Lambda=diag\{\lambda_1,\lambda_2,\ldots,\lambda_N\}$ and $B=diag\{\beta_1,\beta_2,\ldots,\beta_N\}$. G is said to be diagonally (B,Λ)-UPPAM if each component g_i of G is a β_i-projection and λ_i-UAM.

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

Define

$$L_{\|\cdot\|}(T,A,\tilde{x},\Psi)=\sup_{x\neq\tilde{x},x\in\Psi}\frac{\|ATx-AT\tilde{x}\|}{\|Ax-A\tilde{x}\|} \tag{3}$$

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 it is called as the *nonlinear norm*. That 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, when Ψ contains $\tilde{x}$ as an interior point (i.e., Ψ contains a neighborhood of $\tilde{x}$), then for any given matrix B, $L_{\|\cdot\|}(B,I,\tilde{x},\Psi)=\|B\|$. Additionally, for any $\beta>0$, we have $L_{\|\cdot\|}(\beta T,A,\tilde{x},\Psi)=\beta L_{\|\cdot\|}(T,A,\tilde{x},\Psi)$.

Throughout the paper, the identity matrix is denoted by I. For a positive semi-definite diagonal matrix $\Delta=diag\{\delta_1,\delta_2,\ldots,\delta_N\}$, let $\Delta^{1/2}=diag\{\delta_1^{1/2},\delta_2^{1/2},\ldots,\delta_N^{1/2}\}$.

3. Critical global convergence and stability results

In this section, we establish several global convergence and asymptotic stability results for generic continuous-time UPPAM RNNs, which are under the critical conditions that the discriminant matrix defined by the network is positive semi-definite. To be simple, we denote the range of the nonlinear activation operator, i.e., $\mathbf{R}(G)$, by Θ.

Lemma 3.1. *For any $x_0\in A(\Theta)+b$, $x(t,x_0)$, the solution of (1), satisfies $x(t,x_0)\in A(\Theta)+b(t\geq 0)$.*

Proof. By the differential equation theory, we have

$$x(t,x_0)=e^{-t/\tau}x_0+\frac{1}{\tau}e^{-t/\tau}\int_0^t e^{s/\tau}(AG(Wx(s)+q)+b)\,ds$$
$$=e^{-t/\tau}x_0+(1-e^{-t/\tau})\frac{\int_0^{t/\tau}e^{-r}(AG(Wx(t-\tau r)+q)+b)\,dr}{1-e^{-t/\tau}} \tag{4}$$

where $r=(t-s)/\tau$. Since $1-e^{-t/\tau}=\int_0^{t/\tau}e^{-r}\,dr=\lim_{n\to+\infty}\sum_{i=1}^n(t/\tau n)e^{-it/\tau n}$, and $A(\Theta)+b$ is a bounded, closed and convex subset, then $P(t):=\int_0^{t/\tau}e^{-r}(AG(Wx(t-\tau r)+q)+b)\,dr$, the limit of the sum $\sum_{i=1}^n(t/\tau n)e^{-it/\tau n}(AG(Wx(t-it/n)+q)+b)$, should satisfy $P(t)/(1-e^{-t/\tau})\in A(\Theta)+b(\forall t\geq 0)$. Further, by (4), we know $x(t,x_0)\in A(\Theta)+b$ when $x_0\in A(\Theta)+b$. □

Suppose that Θ is bounded, closed and convex. For any $v\in\Theta$, define $T(v)=AG(Wv+q)+b$, 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.

Theorem 3.1. *Assume that G is diagonally (B,Λ)-UPPAM with Θ being a bounded, closed and convex subset of $\mathbb{R}^N$, and A is a nonzero diagonal matrix. If there exists a positive definite diagonal matrix Γ such that*

(i) *$M(\Gamma)=(2\Lambda-B)\Gamma-(\Gamma AW+(\Gamma AW)^T)/2$ is positive semi-definite, and*
(ii) *for one $v^*\in F_e^{-1}(0)$, $L_{\|\cdot\|_2}(T,D,v^*,\Theta)\leq 1$ (here $D=((2\Lambda-B)\Gamma)^{1/2}$),*

is the unique equilibrium point of (1), then x^* is globally asymptotically stable on Θ.

Proof. Denote $A=diag\{a_1,a_2,\ldots,a_N\}$, $\Gamma=diag\{\xi_1,\xi_2,\ldots,\xi_N\}$ and $D=diag\{d_1,d_2,\ldots,d_N\}$. For any trajectory $x(t)$ of (1) starting from $x_0\in A(\Theta)+b$, it follows from Lemma 3.1 that $x(t)\in A(\Theta)+b$. Let $y_0=Wx_0+q$, $y(t)=Wx(t)+q$, $z(t)=AG(y(t))+b$ and $u(t)=z(t)-x(t)$. In the following proof, we will use some diagonal Lyapunov functions, and the concept of diagonal Lyapunov function was first introduced in [41–43].

Define

$$P(x(t))=\tfrac{1}{2}\{(y(t)-x(t))^TA\Gamma(y(t)-x(t))-y^T(t)A\Gamma y(t)-x^T(t)(I-(2\Lambda-B))A\Gamma x(t)\}$$

and

$$Q(x(t))=x^T(t)(2\Lambda-B)\Gamma x(t)-2x^T(t)(2\Lambda-B)\Gamma v^*$$

Let

$$E_1(x(t)) = \tau x^T(t)\left(\frac{((2\Lambda-B)A+B)\Gamma}{2} - (\Gamma AW + (\Gamma AW)^T)\right)x(t) - 2\tau x^T(t)\Lambda\Gamma q$$
$$-\tau x^T(t)\Gamma Bb + \tau\sum_{i=1}^{N}\xi_i a_i\int_{(Wx_0+q)_i}^{(Wx(t)+q)_i}(a_i g_i(s)+b_i)\,ds - \tau P(x(t))$$
$$+\tau Q(x(t))$$

We will complete the proof in the following four steps.

*Step*1. We show that $\lim_{t\to+\infty} dE_1(x(t))/dt = 0$.

Note first that

$$\frac{dP(x(t))}{dt} = \langle \Lambda\Gamma(y(t)-x(t)),(W-I)u(t)\rangle - \langle \Lambda\Gamma y(t),Wu(t)\rangle$$
$$-\langle (I-(2\Lambda-B))\Lambda\Gamma x(t),u(t)\rangle$$
$$= \langle \Lambda\Gamma x(t),((2\Lambda-B)-W)u(t)\rangle - \langle \Lambda\Gamma y(t),u(t)\rangle$$

and

$$\frac{dQ(x(t))}{dt} = 2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle - 2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle$$

Since $x(t)\in A(\Theta)+b$, there exists $p(t)\in\Theta$, such that $x(t)=Ap(t)+b$, and then, $Ap(t)=x(t)-b$. Meanwhile, on noting that $((2\Lambda-B)A+B)\Gamma/2-(\Gamma AW+(\Gamma AW)^T)$ is symmetric, then a direct calculation shows

$$\frac{dE_1(x(t))}{dt} = \langle ((2\Lambda-B)A\Gamma - 2(\Gamma AW+(\Gamma AW)^T)x(t),u(t)\rangle$$
$$+\langle \Gamma Bx(t),u(t)\rangle - 2\langle \Lambda\Gamma q,u(t)\rangle - \langle \Gamma Bb,u(t)\rangle$$
$$+\langle \Lambda\Gamma z(t),Wu(t)\rangle$$
$$-\langle \Lambda\Gamma x(t),((2\Lambda-B)-W)u(t)\rangle + \langle \Lambda\Gamma y(t),u(t)\rangle$$
$$+2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle - 2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle$$
$$= \langle \Lambda(2\Lambda-B)\Gamma x(t),u(t)\rangle - 2\langle \Lambda\Gamma y(t),u(t)\rangle + \langle \Gamma BAp(t),u(t)\rangle$$
$$-2\langle \Lambda\Gamma x(t),Wu(t)\rangle + \langle (2\Lambda-B)\Lambda\Gamma z(t),u(t)\rangle$$
$$-\langle \Lambda\Gamma z(t),((2\Lambda-B)-W)u(t)\rangle$$
$$-\langle \Lambda\Gamma x(t),((2\Lambda-B)-W)u(t)\rangle + \langle \Lambda\Gamma y(t),u(t)\rangle$$
$$+2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle - 2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle$$
$$= -\langle \Lambda\Gamma(y(t)-Bp(t)),u(t)\rangle + \langle \Lambda\Gamma u(t),Wu(t)\rangle$$
$$+2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle - 2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle$$
$$= -\langle \Lambda\Gamma(y(t)-Bp(t)),u(t)\rangle + \frac{1}{2}\langle u(t),(\Gamma AW+(\Gamma AW)^T)u(t)\rangle$$
$$+2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle - 2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle$$
$$= -\langle \Lambda\Gamma(y(t)-Bp(t)),u(t)\rangle - u^T(t)((2\Lambda-B)\Gamma - \frac{\Gamma AW+(\Gamma AW)^T}{2})u(t)$$
$$+u^T(t)(2\Lambda-B)\Gamma u(t) + 2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle$$
$$-2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle \qquad (5)$$

From the assumption that v^* is a fixed point of $AG(Wv+q)+b$, we have $v^* = AG(Wv^*+q)+b$, and then

$$u^T(t)(2\Lambda-B)\Gamma u(t) + 2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle - 2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle$$
$$= \langle z(t)-x(t),(2\Lambda-B)\Gamma u(t)\rangle + 2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle$$
$$-2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle$$
$$= \langle z(t)+x(t),(2\Lambda-B)\Gamma u(t)\rangle - 2\langle v^*,(2\Lambda-B)\Gamma u(t)\rangle$$
$$= \langle (z(t)-v^*)+(x(t)-v^*),(2\Lambda-B)\Gamma u(t)\rangle$$
$$= \langle (z(t)-v^*)+(x(t)-v^*),(2\Lambda-B)\Gamma((z(t)-v^*)-(x(t)-v^*))\rangle$$
$$= \langle z(t)-v^*,(2\Lambda-B)\Gamma(z(t)-v^*)\rangle - \langle x(t)-v^*,(2\Lambda-B)\Gamma(x(t)-v^*)\rangle$$
$$= \langle A(G(Wx(t)+q)-G(Wv^*+q)),\Gamma(2\Lambda-B)A(G(Wx(t)+q)-G(Wv^*+q))\rangle$$
$$-\langle x(t)-v^*,(2\Lambda-B)\Gamma(x(t)-v^*)\rangle$$
$$= \langle DA(G(Wx(t)+q)-G(Wv^*+q)),DA(G(Wx(t)+q)-G(Wv^*+q))\rangle$$
$$-\langle D(x(t)-v^*),D(x(t)-v^*)\rangle$$
$$= \|DA(G(Wx(t)+q)-G(Wv^*+q))\|_2^2 - \|D(x(t)-v^*)\|_2^2 \qquad (6)$$

Noting that $L_{\|\cdot\|_2}(T,D,v^*,\mathbf{R}(G))\le 1$, it follows

$$L_{\|\cdot\|_2}(T,D,v^*,\mathbf{R}(G)) = \sup_{x\ne v^*,x\in\mathbf{R}(G)}\frac{\|DTx-DTv^*\|_2}{\|Dx-Dv^*\|_2}$$
$$= \sup_{x\ne v^*,x\in\mathbf{R}(G)}\frac{\|D(AG(Wx+q)+b)-D(AG(Wv^*+q)+b)\|_2}{\|Dx-Dv^*\|_2}$$
$$\le 1.$$

This, combined with (6), implies that

$$\langle u^T(t)(2\Lambda-B)\Gamma u(t) + 2\langle (2\Lambda-B)\Gamma x(t),u(t)\rangle - 2\langle (2\Lambda-B)\Gamma v^*,u(t)\rangle \le 0.$$

Since G is a B-projection and $p(t)\in\Theta$, it is clear that $G(Bp(t)) = p(t)$. Then by the conditions that $(2\Lambda-B)\Gamma-(\Gamma AW+(\Gamma AW)^T)/2$ is positive semi-definite and the diagonal nonlinear property of G, one can get from (5) that

$$\frac{dE_1(x(t))}{dt} \le -\langle \Lambda\Gamma(y(t)-Bp(t)),u(t)\rangle$$
$$= -\sum_{i=1}^{N} a_i\xi_i((Wx(t)+q)_i-\beta_i p_i(t))\cdot((a_i g_i((Wx(t)+q)_i)+b_i)-(a_i p_i(t)+b_i))$$
$$= -\sum_{i=1}^{N} a_i^2\xi_i((Wx(t)+q)_i-\beta_i p_i(t))\cdot(g_i((Wx(t)+q)_i)-g_i(\beta_i p_i(t))) \qquad (7)$$

For each component g_i of G is a β_i-projection and λ_i-UAM, we then have

$$((Wx(t)+q)_i-\beta_i p_i(t))\cdot(g_i((Wx(t)+q)_i)-g_i(\beta_i p_i(t)))$$
$$\ge \lambda_i(g_i((Wx(t)+q)_i)-g_i(\beta_i p_i(t)))^2 \qquad (8)$$

Let $\lambda_{min}(\Lambda\Gamma)$ be the smallest eigenvalue of $\Lambda\Gamma$. Then by (7) and (8), we get

$$\frac{dE_1(x(t))}{dt} \le -\sum_{i=1}^{N}\xi_i a_i^2\lambda_i(g_i((Wx(t)+q)_i)-g_i(\beta_i p_i(t)))^2$$
$$= -\sum_{i=1}^{N}\xi_i\lambda_i(a_i(g_i((Wx(t)+q)_i)-p_i(t)))^2$$
$$= -(A(G(Wx+q)-p(t)))^T\Lambda\Gamma(A(G(Wx+q)-p(t)))$$
$$\le -\lambda_{min}(\Lambda\Gamma)\|A(G(Wx+q)-p(t))\|_2^2$$
$$= -\lambda_{min}(\Lambda\Gamma)\|z(t)-x(t)\|_2^2 \qquad (9)$$

Since $dE_1(x(t))/dt$ is continuous, $x(t)\in A(\Theta)+b$ and Θ is a bounded and closed set, it follows that $dE_1(x(t))/dt$ is a uniformly continuous function of t in $[0,+\infty)$. Furthermore, by (9), we have $dE_1(x(t))/dt\le 0$ since Λ and Γ all are positive definite, which, combined with the fact that $E_1(x(t))$ is bounded, implies that $\lim_{t\to+\infty}E_1(x(t))$ exists. Thus, applying the well-known Barbalat Lemma, we obtain that $\lim_{t\to+\infty}dE_1(x(t))/dt = 0$.

*Step*2. We show that any limit point of $x(t)$ is an equilibrium state of (1).

Let x^* be any limit point of $x(t)$, i.e., $\lim_{n\to+\infty}x(t_n)=x^*$ for some positive sequence $\{t_n\}$ with $t_n\to+\infty$ as $n\to+\infty$ (notice that x^* exists since $x(t)$ is bounded).

From (9) and the fact that $\lambda_{min}(\Lambda\Gamma)>0$, we have

$$0 = \lim_{t\to+\infty}\frac{dE_1(x(t))}{dt}$$
$$\le \liminf_{t\to+\infty}(-\lambda_{min}(\Lambda\Gamma))\|z(t)-x(t)\|_2^2$$
$$\le 0$$

consequently,

$$\lim_{t\to+\infty}\|AG(Wx(t)+q)+b-x(t)\|_2 = 0 \qquad (10)$$

Thus, it can be deduced that

$$\|AG(Wx^*+q)+b-x^*\|_2=0$$

That is, x^* is an equilibrium state of (1).

*Step*3. In what follows, we still use the denotation that $y(t)=Wx(t)+q$, and specially, denote $y(t_n)=Wx(t_n)+q$. We will show $\lim_{t\to+\infty}(AG(y(t))+b)-x^*=0$.

Clearly, since x^* is the limit point of $x(t_n)$, $x(t_n)\in A(\Theta)+b$ and Θ is closed, then $x^*\in A(\Theta)+b$, i.e., there exists $p^*\in\Theta$, such that $x^*=Ap^*+b$. Further, since G is a B-projection and x^* is an equilibrium state of (1) (that is to say, $x^*=AG(Wx^*+q)+b$ holds), we can deduce that $G(Bp^*)=p^*$ and $Ap^*=AG(Wx^*+q)$.

Let

$$E_2(x(t))=\tau\sum_{i=1}^{N}\xi_i a_i^2\int_{(Wx^*+q)_i}^{y_i(t)}(g_i(r)-g_i((Wx^*+q)_i))\,dr$$

It can be deduced that

$$\begin{aligned}\frac{dE_2(x(t))}{dt}&=\sum_{i=1}^{N}\xi_i a_i^2(g_i(y_i(t))-g_i((Wx^*+q)_i))\cdot\left(\tau\frac{dy_i(t)}{dt}\right)\\&=(AG(y(t))-Ap^*)^T\Gamma AW(-x(t)+AG(y(t))+b)\\&=(AG(y(t))-Ap^*)^T\Gamma A[(-((Wx(t)+q)-Bp^*)+((Wx^*+q)-BG(y(t)))\\&\quad+W(AG(y(t))+b-x^*)+B(G(y(t))-p^*)]\\&=-(AG(y(t))-Ap^*)^T\Gamma A((Wx(t)+q)-Bp^*)\\&\quad-(Ap^*-AG(y(t)))^T\Gamma A((Wx^*+q)-BG(y(t)))\\&\quad+(AG(y(t))-Ap^*)^T\Gamma AW(AG(y(t))+b-x^*)\\&\quad+(AG(y(t))-Ap^*)^T\Gamma AB(G(y(t))-p^*)\end{aligned}\tag{11}$$

Observing that G is diagonally (B,Λ)-UPPAM and $G(Bp^*)=p^*$, we obtain

$$\begin{aligned}&(AG(y(t))-Ap^*)^T\Gamma A((Wx(t)+q)-Bp^*)\\&=(AG(y(t))-AG(Bp^*))^T\Gamma A(y(t)-Bp^*)\\&=\sum_{i=1}^{N}\xi_i a_i^2(g_i(y_i(t))-g_i(\beta_i p_i^*))(y_i(t)-\beta_i p_i^*)\\&\geq\sum_{i=1}^{N}\xi_i a_i^2\lambda_i(g_i(y_i(t))-g_i(\beta_i p_i^*))^2\\&=\sum_{i=1}^{N}\xi_i a_i^2\lambda_i(g_i(y_i(t))-p_i^*)^2\\&=(AG(y(t))-Ap^*)^T\Gamma\Lambda(AG(y(t))-Ap^*)\end{aligned}\tag{12}$$

and similarly, by the fact that $Ap^*=AG(Wx^*+q)$ and $G(y(t))=G(BG(y(t)))$, we have

$$\begin{aligned}&(Ap^*-AG(y(t)))^T\Gamma A((Wx^*+q)-BG(y(t)))\\&=\langle G(Wx^*+q)-G(BG(y(t))),A\Gamma A((Wx^*+q)-BG(y(t)))\rangle\\&\geq\langle G(Wx^*+q)-G(BG(y(t))),\Lambda A^2\Gamma(G(Wx^*+q)-G(BG(y(t))))\rangle\\&=\langle AG(Wx^*+q)-AG(BG(y(t))),\Lambda\Gamma(AG(Wx^*+q)-AG(BG(y(t))))\rangle\\&=(Ap^*-AG(y(t)))^T\Gamma\Lambda(Ap^*-AG(y(t)))\\&=(AG(y(t))-Ap^*)^T\Gamma\Lambda(AG(y(t))-Ap^*)\end{aligned}\tag{13}$$

$$\begin{aligned}&(AG(y(t))-Ap^*)^T\Gamma AW(AG(y(t))+b-x^*)\\&=(AG(y(t))-Ap^*)^T\Gamma AW(AG(y(t))+b-(AG(Wx^*+q)+b))\\&=(AG(y(t))-Ap^*)^T\Gamma AW(AG(y(t))-AG(Wx^*+q))\\&=(AG(y(t))-Ap^*)^T\frac{\Gamma AW+(\Gamma AW)^T}{2}(AG(y(t))-Ap^*)\end{aligned}\tag{14}$$

then by (11)–(14), it can be deduced that

$$\begin{aligned}\frac{dE_2(x(t))}{dt}&\leq-2(AG(y(t))-Ap^*)^T\Gamma\Lambda(AG(y(t))-Ap^*)\\&\quad+(AG(y(t))-Ap^*)^T\frac{\Gamma AW+(\Gamma AW)^T}{2}(AG(y(t))-Ap^*)\\&\quad+(AG(y(t))-Ap^*)^T\Gamma B(AG(y(t))-Ap^*)\\&=-(AG(y(t))-Ap^*)^T((2\Lambda-B)\Gamma-\frac{\Gamma AW+(\Gamma AW)^T}{2})(AG(y(t))-Ap^*)\end{aligned}$$

Thus, the positive semi-definiteness of $(2\Lambda-B)\Gamma-(\Gamma AW+(\Gamma AW)^T)/2$ implies $dE_2(x(t))/dt\leq 0$ for all $t\geq 0$, and furthermore, $\lim_{t\to+\infty}E_2(x(t))$ exists since $E_2(x(t))\geq 0$ by Lemma 2 in [2]. This, together with the fact that $\lim_{n\to+\infty}y(t_n)=Wx^*+q$, implies that $\lim_{t\to+\infty}E_2(x(t))=0$. As a result, we obtain by applying Lemma 2 in [2] to each component g_i of G that

$$\lim_{t\to+\infty}(G(y(t))-G(Wx^*+q))=0$$

that is

$$\begin{aligned}&\lim_{t\to+\infty}(AG(y(t))+b-(AG(Wx^*+q)+b))\\&=\lim_{t\to+\infty}(AG(y(t))+b-(AG(p^*)+b))\\&=\lim_{t\to+\infty}(AG(y(t))+b-(Ap^*+b))\\&=\lim_{t\to+\infty}(AG(y(t))+b-x^*)\\&=0\end{aligned}$$

and then,

$$\lim_{t\to+\infty}(AG(y(t))+b)=x^*\tag{15}$$

*Step*4. We finally prove that $\lim_{t\to+\infty}x(t)=x^*$.

By differential equation theory, $x(t)$ solves the following integral equation:

$$x(t)-x^*=e^{-(1/\tau)(t-t_0)I}(x_0-x^*)+\int_{t_0}^{t}e^{-(1/\tau)(t-s)I}\cdot\frac{1}{\tau}\cdot(AG(y(s))+b-x^*)\,ds$$

Obviously, it holds that

$$\|x(t)-x^*\|\leq e^{-(1/\tau)(t-t_0)}\|x_0-x^*\|+\int_{t_0}^{t}e^{-(1/\tau)(t-s)}\cdot\frac{1}{\tau}\cdot\|AG(y(s))+b-x^*\|\,ds\tag{16}$$

By (15), for any $\varepsilon>0$, there is a $T_\varepsilon>0$ such that, whenever $t>t_0\geq T_\varepsilon$,

$$\|AG(y(t))+b-x^*\|\leq\varepsilon$$

Therefore, we conclude from (16) that, when $t>t_0\geq T_\varepsilon$,

$$\begin{aligned}\|x(t)-x^*\|&\leq e^{-(1/\tau)(t-t_0)}\|x_0-x^*\|+\frac{\varepsilon}{\tau}\int_{t_0}^{t}e^{-(1/\tau)(t-s)}ds\\&<e^{-(1/\tau)(t-t_0)}\|x_0-x^*\|+\varepsilon\end{aligned}$$

Letting $t\to+\infty$ in the above inequality yields $\lim_{t\to+\infty}\|x(t)-x^*\|\leq\varepsilon$, which then implies $\lim_{t\to+\infty}x(t)=x^*$ since ε is arbitrary, i.e., for any trajectory $x(t)$ of (1) starting from $A(\Theta)+b$, there corresponds an equilibrium state x^* of (1) such that $\lim_{t\to+\infty}x(t)=x^*$. Furthermore, when $F_e^{-1}(0)=\{x^*\}$, then x^* is both attractive and stable on Θ since Θ is bounded, i.e., x^* is globally asymptotically stable on Θ. This completes the proof of the theorem. □

Corollary 3.1. *Assume that G is diagonally (B,Λ)-UPPAM and $\mathbf{R}(G)\subseteq\mathbb{R}^N$ is a bounded, closed and convex set, and A is a nonzero diagonal matrix. If there exists a positive definite diagonal matrix Γ, such that $(2\Lambda-B)\Gamma-(\Gamma AW+(\Gamma AW)^T)/2$ is positive semi-definite and $\|D\Lambda^{-1}AWD^{-1}\|_2\leq 1$ (here $D=((2\Lambda-B)\Gamma)^{1/2}$), then RNN system (1) is globally convergent on $\mathbf{R}(G)$. Moreover, when x^* is the unique* *equilibrium point of (1), then x^* is globally asymptotically stable on* $\mathbf{R}(G)$.

Proof. For any $v^* \in F_e^{-1}(0)$, we have

$$\|DTx-DTv^*\|_2^2 = \|D(AG(Wx+q)+b)-D(AG(Wv^*+q)+b)\|_2^2 = \sum_{i=1}^{N}(d_i a_i (g_i((Wx+q)_i)-g_i((Wv^*+q)_i)))^2. \quad (17)$$

Since g_i is λ_i-uniformly anti-monotone, then

$$(g_i((Wx+q)_i)-g_i((Wv^*+q)_i))((Wx+q)_i-(Wv^*+q)_i) \geq \lambda_i (g_i((Wx+q)_i)-g_i((Wv^*+q)_i))^2 \geq 0$$

and further

$$|(Wx+q)_i-(Wv^*+q)_i|^2 \geq \lambda_i^2 |g_i((Wx+q)_i)-g_i((Wv^*+q)_i)|^2$$

then from (17), we get

$$\begin{aligned}\|DTx-DTv^*\|_2^2 &= \sum_{i=1}^{N}(d_i a_i (g_i((Wx+q)_i)-g_i((Wv^*+q)_i)))^2 \\ &\leq \sum_{i=1}^{N}(d_i a_i \lambda_i^{-1}(W(x-v^*))_i)^2 \\ &= \sum_{i=1}^{N}(d_i a_i \lambda_i^{-1})^2 \left(\sum_{j=1}^{N} W_{ij} d_j^{-1}(d_j(x_j-v_j^*))\right)^2 \\ &= \sum_{i=1}^{N}\left(\sum_{j=1}^{N} d_i a_i \lambda_i^{-1} W_{ij} d_j^{-1}(d_j(x_j-v_j^*))\right)^2 \\ &= \|D\Lambda^{-1}AWD^{-1}(D(x-v^*))\|_2^2 \end{aligned} \quad (18)$$

Clearly, when $\|D\Lambda^{-1}AWD^{-1}\|_2 \leq 1$, then by (18), it can be deduced that

$$L_{\|\cdot\|_2}(T,D,v^*,\mathbf{R}(G)) = \sup_{x \neq v^*, x \in \mathbf{R}(G)} \frac{\|DTx-DTv^*\|_2}{\|Dx-Dv^*\|_2} \leq 1.$$

Corollary 3.1 is then proved from Theorem 3.1. □

Remark 3.1. The continuous-time recurrent neural networks have been attracting great interest either as associative memories, optimization solvers, or system simulators in science and engineering [33–35]. For different purpose, there exist varies continuous-time RNN individuals and huge numbers of dynamics behaviors analysis for each individual. Since the uniformly pseudo-projection-anti-monotone operator provides a very appropriate, unified framework within which most of the known activation operators can be embedded, the continuous-time UPPAM RNNs model, i.e., model (1), offers a uniform approach to study continuous-time RNNs. The global convergence as well as asymptotical stability results achieved in Theorem 3.1 and Corollary 3.1 exploit new dynamics analysis for continuous-time UPPAM RNNs. The obtained results remove mostly the diversity and redundancy of the dynamical conclusions existing in various RNNs models, and further, discuss that under the critical condition (which is the essential line of demarcation between stability and instability), what asymptotic behavior for generic continuous-time RNNs will be.

Yang and Cao [22] have gotten the globally exponential stability of a static neural network with projection operator under the condition that $I-W$ is nonnegative, while, this model is a kind of the UPPAM RNNs model and $(I-W) \geq 0$ is only a special case of $S(\Gamma,(2\Lambda-B)) \geq 0$. Peng et al. [1] have proved that when W is quasi-symmetric (i.e., there exists a positive definite diagonal matrix D, such that DW is symmetric), then a static neural network model with nearest point projection activation operator is critical global convergence on a region defined by the network, and this conclusion can also be summarized in our results. For a RNN with projection activation operator, in [2–4], we have proved the critical convergence of the static neural network model, obviously, it is a special case for this critical analysis of UPPAM RNNs. For Hopfield type neural network, in [44], Zeng et al. have got the globally exponential stability of recurrent neural network with time-varying delays when matrix $C-A^{(1)}-B^{(1)}$ is a nonsingular M-matrix. While, we know that $C-A^{(1)}-B^{(1)}$ being a nonsingular M -matrix is only a special case of the noncritical condition, let alone the critical condition. In [17], Guan et al. also present some noncritical results for Hopfield type neural network. For BSB-type RNNs, under the conditions that $-(\Gamma V+V^T\Gamma)$ is positive definite, or $\Gamma(I+\alpha V)$ is negative definite for a positive diagonal matrix Γ (where α is a positive parameter and V is the weight matrix of this RNN), the globally convergent analysis of such type RNNs have been revealed in [9,15]. But, it is easy to verify that such conditions are still noncritical. For cellular neural networks, when W is symmetric with $I-W$ being positive definite, or $\Gamma-(\Gamma W+W^T\Gamma)/2$ is positive definite, or ΓW is symmetric and either 1 is not an eigenvalue of any principal sub-matrix of W, or $\Gamma(I-W)$ is nonnegative definite, the stability analysis haven been proved in [13–16]. Similarly, these stability results are given under the noncritical conditions or under a special case of critical conditions. In all, the conclusions presented in this section can unify and improve the latest critical analysis for continuous-time RNN models, let alone they can extend deeply those noncritical analysis for continuous-time RNNs (see, e.g. [14,15,17–20,22] and the references therein). And specially, the achieved results generalize and extend further the existing dynamics conclusions for the brain-state-in-a-box/domain recurrent neural networks, Hopfield-type neural networks, cellular neural networks, etc.

In addition, the obtained analysis results can be applied to solve the following linear variational inequality (LVI): determining a vector x^* in a nonempty closed convex subset $\Omega \subseteq \mathcal{R}^n$ such that

$$(Qx^*+q)^T(x-x^*) \geq 0, \quad \forall x \in \Omega \quad (19)$$

where $Q=(Q_{ij})$ is an $n \times n$ real matrix and $q \in \mathcal{R}^n$ is a vector. The corresponding RNN for solving LVI is described by

$$\tau \frac{dx}{dt} = -x + P_\Omega(x-\Lambda Qx+q) \quad (20)$$

where Λ is a positive diagonal matrix.

LVI has many applications in bound constrained quadratic programming, linear complementarity problem, economic equilibrium modeling, traffic network equilibrium modeling, analysis of piecewise-linear resistive circuits, and so on [11,18,38,39]. Hu, Wang, Xia, et al. have proved that the solutions for an LVI on its constrained set can be found when matrix $(I-\Lambda Q)$ is positive semi-definite [19,37,40]. We can see that it is a special case of $S(\Gamma,(2\Lambda-B)) \geq 0$.

Applying Theorem 3.1 and Corollary 3.1 directory to the specific RNN model (20), we can have some new criteria for global convergence of such RNN and what is more important, new solution for solving LVI problems. Because quadratic and linear programming problems are special cases of LVI in terms of solutions, then our critical results can solve them efficiently as well.

4. Illustrative examples

In this section, we provide two illustrative examples to demonstrate the validity of the critical convergence and stability results formulated in the previous section. It should be noticed that the known stability and convergence results developed in literature can not be applied here.

Example 4.1. Consider the following UPPAM RNN with two neurons:

$$\begin{cases} \dfrac{dx_1(t)}{dt} = -x_1(t) + g_1(\sqrt{2}x_1(t) + \sqrt{2}x_2(t)) \\ \dfrac{dx_2(t)}{dt} = -x_2(t) + g_1(-\sqrt{2}x_1(t) + \sqrt{2}x_2(t) + 3\sqrt{2}) \end{cases} \quad (21)$$

where g_i $(i=1,\ 2)$ is defined as follows:

$$g_i(s) = \begin{cases} 1, & s > \sqrt{2} \\ s/\sqrt{2}, & s \in [-\sqrt{2}, \sqrt{2}] \\ -1, & s < -\sqrt{2} \end{cases} \quad (22)$$

In this example, $\Lambda = B = \sqrt{2}I$, $A = I$, the unique equilibrium state is $(1,1)^T$. For this UPPAM network, almost all of the existing stability conclusions cannot be used here, because in this example, the activation operator is not a nearest projection, or a general projection, i.e., the latest critical results presented in [1,3,4] cannot be applied.

In what follows, we will show the results established in this paper can be successfully applied here. By setting $\Gamma = I$, we have $(2\Lambda - B)\Gamma - (\Gamma A W + W^T A \Gamma)/2 \geq 0$. For any $v \in \mathbf{R}(G) = [-1,1]^2$, define $T(v) = G(Wv+q)$. It is easy to see that $\mathbf{F}(T) = \{(1,1)^T\}$. Let $D = ((2\Lambda - B)\Gamma)^{1/2}$. For any $v^* \in \mathbf{F}(T)$, we can verify that $L_{\|\cdot\|_2}(T, D, v^*, \mathbf{R}(G)) \leq 1$. In fact,

$$L_{\|\cdot\|_2}(T, D, v^*, \mathbf{R}(G)) = \sup_{v \neq v^*, v \in \mathbf{R}(G)} \frac{\|DTv - DTv^*\|_2}{\|Dv - Dv^*\|_2}$$

$$= \sup_{v \neq v^*, v \in \mathbf{R}(G)} \frac{\|G(Wv+q) - v^*\|_2}{\|v - v^*\|_2} \quad (23)$$

On noting that for any $v \in \mathbf{R}(G)$, $w_{21}v_1 + w_{22}v_2 + q_2 \geq \sqrt{2}$, then, by the definition of g_2, we have $g_2(w_{21}v_1 + w_{22}v_2 + q_2) = 1$ and $\|G(Wv+q) - v^*\|_2^2 = (g_1(w_{11}v_1 + w_{12}v_2 + q_1) - v_1^*)^2 = (1 - g_1(\sqrt{2}(v_1 + v_2)))^2$. Meanwhile, since $\|v - v^*\|_2^2 = (1-v_1)^2 + (1-v_2)^2 \geq (2-(v_1+v_2))^2/2$, so, if for any $v \in \mathbf{R}(G)$, $(1 - g_1(\sqrt{2}(v_1+v_2)))^2 \leq (2-(v_1+v_2))^2/2$ always holds, then we can get that $L_{\|\cdot\|_2}(T, D, v^*, \mathbf{R}(G)) \leq 1$. Since $g_1(v_1+v_2) \leq 1$ and $v_1 + v_2 \leq 2$, which equals to prove

$$1 - g_1(\sqrt{2}(v_1+v_2)) \leq (2-(v_1+v_2))/\sqrt{2} \quad (24)$$

We prove it in the following three cases.

(a) When $\sqrt{2} \leq \sqrt{2}(v_1+v_2) \leq 2\sqrt{2}$, then $g_1(\sqrt{2}(v_1+v_2)) = 1$, and (24) holds obviously;

(b) When $-\sqrt{2} \leq \sqrt{2}(v_1+v_2) \leq \sqrt{2}$, then $g_1(\sqrt{2}(v_1+v_2)) = v_1 + v_2$ and $1-(v_1+v_2) \leq \sqrt{2} - (1/\sqrt{2})(v_1+v_2)$, i.e. (24) is true;

(c) When $-2\sqrt{2} \leq \sqrt{2}(v_1+v_2) \leq -\sqrt{2}$, then it is clear that $g_1(\sqrt{2}(v_1+v_2)) = -1$ and $2 \leq \sqrt{2} - (1/\sqrt{2})(v_1+v_2)$, and (24) holds, too.

From the discussion above, it can be deduced that $L_{\|\cdot\|_2}(T, D, v^*, \mathbf{R}(G)) \leq 1$. Then by Corollary 3.1, we get that $x^* = (1,1)^T \in \Omega_e$ is globally asymptotically stable on $\mathbf{R}(G)$. Simulation results with several random initial points starting from $\mathbf{R}(G)$ are depicted in Fig. 1.

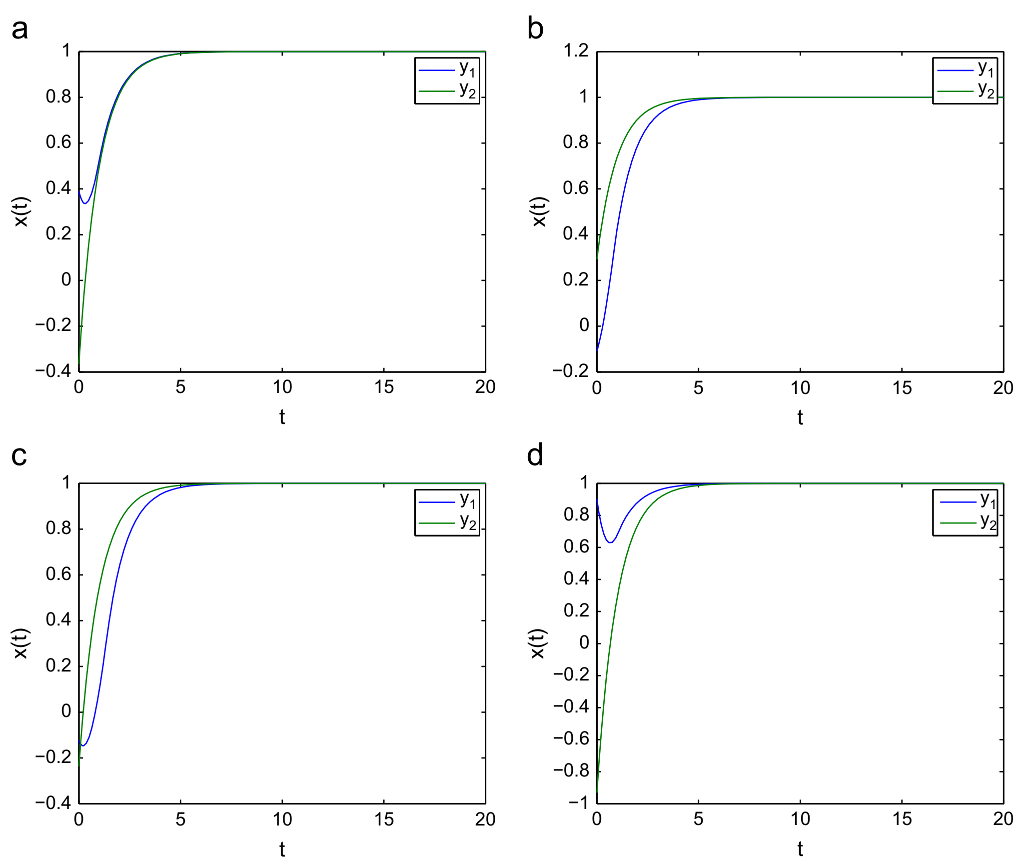


Fig. 1. Transient behaviors of RNN in system (21) with random initial points $x_0 \in \mathbf{R}(G)$.

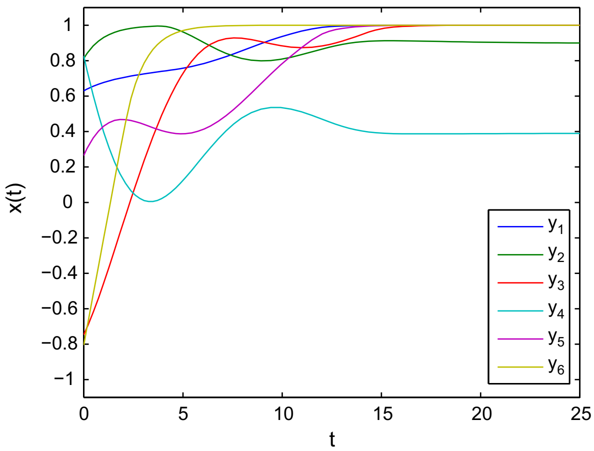


Fig. 2. Transient behaviors of RNN in system (25) with random initial point $x_0 \in [-1,1]^6$.

Example 4.2. Consider a UPPAM RNN in a 6-dimensional array of neurons

$$\begin{cases} \frac{dx_1(t)}{dt} = -x_1(t)+g_1(0.9x_1(t)-0.15\sqrt{2}x_2(t)+0.3) \\ \frac{dx_2(t)}{dt} = -x_2(t)+g_2(0.15\sqrt{2}x_1(t)+0.4x_2(t)-0.05\sqrt{6}x_3(t)) \\ \frac{dx_3(t)}{dt} = -x_3(t)+g_3(0.05\sqrt{6}x_2(t)+8/30x_3(t)-0.05\sqrt{3}x_4(t)) \\ \frac{dx_4(t)}{dt} = -x_4(t)+g_4(0.05\sqrt{3}x_3(t)+0.2x_4(t)-0.03\sqrt{5}x_5(t)) \\ \frac{dx_5(t)}{dt} = -x_5(t)+g_5(0.03\sqrt{5}x_4(t)+0.18x_5(t)) \\ \frac{dx_6(t)}{dt} = -x_6(t)+g_6(1/6x_6(t)+0.1) \end{cases} \tag{25}$$

where each g_i $(i=1,2,\ldots,6)$ is defined as

$$g_i(s) = \begin{cases} 1, & s > 1/i \\ i*s, & s \in [-1/i, 1/i] \\ -1, & s < -1/i \end{cases} \tag{26}$$

In this case, $\Lambda = B = diag\{1,1/2,1/3,1/4,1/5,1/6\}$, $A=I$. For any positive diagonal matrix Γ, $M(\Gamma) = (2\Lambda - B)\Gamma - (\Gamma AW + W^T A\Gamma)/2$ is not positive, but by taking $\Gamma = I$, it is clear that $M(\Gamma)$ is positive semi-definite. That is, the dynamics behaviors of system (25) should be considered under the critical conditions. For this example, there does not exist a result to ensure the convergence of it. On noting that $\|D\Lambda^{-1}AWD^{-1}\|_2 = 1$, where $D = ((2\Lambda - B)\Gamma)^{1/2}$, then by Corollary 3.1, it is quite easy to achieve the global convergence of system (25) on $\mathbf{R}(G) = [-1,1]^6$. The following Fig. 2 depicts the time responses of neural state variables of the system starting randomly from $[-1,1]^6$.

5. Conclusion

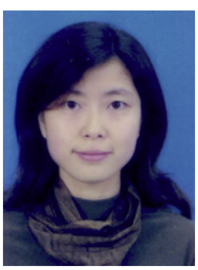
RNNs have been attracting great interest in many fields, such as optimization solvers, associative memories, system simulators in science and engineering, etc. There exist lots of RNN individuals and corresponding mathematical foundations, but the results are very often redundant with similarity.

In the present paper, based on the uniformly pseudo-projection anti-monotone property of common activation operators, the generic continuous-time UPPAM RNNs model has been introduced to unify the existing continuous-time RNNs individuals. Most important, the critical global convergence as well as asymptotic stability of the UPPAM net in general setting has also been proven. The established model and dynamics results not only unify but also jointly generalize and extend the most known conclusions of RNNs, and the approach has lunched a visible step towards establishment a general mathematical method of studying recurrent neural networks.

References

[1] J. Peng, Z.B. Xu, H. Qiao, A critical analysis on global convergence of Hopfield-type neural networks, IEEE Trans. Circuits Syst. 52 (2005) 804–814.
[2] C. Qiao, Z.B. Xu, New critical analysis on global convergence of recurrent neural networks with projection mappings, in: Lecture Notes in Computer Science, vol. 4493, Springer-Verlag, Berlin, Germany, 2007, pp. 131–139.
[3] C. Qiao, Z.B. Xu, A critical global convergence analysis of recurrent neural networks with general projection mappings, Neurocomputing 72 (2009) 1878–1886.
[4] C. Qiao, Z.B. Xu, On the P-critical dynamics analysis of projection recurrent neural networks, Neurocomputing 73 (2010) 2783–2788.
[5] C. Qiao, Z.B. Xu, Critical dynamics study on recurrent neural networks: globally exponential stability, Neurocomputing 77 (2012) 205–211.
[6] T.P. Chen, Global convergence of delayed dynamical systems, IEEE Trans. Neural Networks 12 (2001) 1532–1536.
[7] J.A. Anderson, J.W. Silverstein, S.A. Ritz, Distinctive features, categorical perception, and probability learning: some applications of a neural model, Psychol. Rev. 84 (1977) 413–451.
[8] 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.
[9] H. Qiao, J. Peng, Z.B. Xu, A reference model approach to stability analysis of neural networks, IEEE Trans. Syst. Man Cybern. 33 (2003) 925–936.
[10] 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.
[11] Y.S. Xia, J. Wang, On the stability of globally projected dynamical systems, J. Optim. Theory Appl. 106 (2000) 129–150.
[12] W.A. Little, The existence of persistent states in the brain, Math. Biosci. 19 (1974) 101–120.
[13] L.O. Chua, L. Yang, Cellular neural networks: theory and applications, IEEE Trans. Circuits Syst. 35 (1988) 1257–1290.
[14] L.O. Chua, T. Roska, Cellular Neural Networks and Visual Computing: Foundations and Applications, Cambridge University Press, UK, 2002.
[15] Z.B. Xu, H. Qiao, J. Peng, A comparative study of two modeling approaches in neural networks, Neural Networks 17 (2004) 73–85.
[16] A. Slavova, Cellular Neural Networks: Dynamics and Modelling, Kluwer Academic Publishers, 2003.
[17] Z.H. Guan, G. Chen, Y. Qin, On equilibria, stability and instability of Hopfield neural networks, IEEE Trans. Neural Networks 11 (2000) 534–540.
[18] 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.
[19] 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.
[20] 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.
[21] T.P. Chen, S.I. Amari, New theorems on global convengence of some dynamical systems, Neural Networks 14 (2001) 251–255.
[22] Y.Q. Yang, J. Cao, Solving quadratic programming problems by delayed projection neural network, IEEE Trans. Neural Networks 17 (2006) 1630–1634.
[23] Z.B. Xu, C. Qiao, Towards a unified feedback neural network theory: the uniformly pseudo-projection-anti-monotone net, Acta Math. Sin. Engl. Ser. 27 (2011) 377–396.
[24] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. U.S.A. 79 (1982) 2554–2558.
[25] J.J. Hopfield, D.W. Tank, Computing with neural circuits: a model, Science 233 (1986) 625–633.
[26] B. Kosko, Bidirectional associative memories, IEEE Trans. Syst. Man Cybern. 18 (1988) 49–60.
[27] L.B. Almedia, Backpropagation in perceptrons with feedback, in: Neural Computers, Springer-Verlag, Berlin, 1988, pp. 199–208.
[28] Z.Y. Chen, C.P. Kwong, Z.B. Xu, Multiple-valued feedback and recurrent correlation neural networks, Neural Comput. Appl. 3 (1995) 242–250.
[29] W. Banzhaf, Towards continuous models of memory, in: Proceedings of the IEEE 1st International Conference on Neural Nets, San Diego, 1987, pp. 223–230.
[30] G. Seiler, J. Nossek, Winner-take-all cellular neural networks, IEEE Trans. Circuits Syst. 40 (1993) 184–190.
[31] C. Peterson, J.R. Anderson, A mean field theory learning algorithm for neural networks, Complex Syst. 1 (1987) 995–1019.
[32] T.D. Chiueh, R.M. Goodman, Recurrent correlation associative memories, IEEE Trans. Neural Networks 2 (1991) 275–284.

[33] K.J. Hunt, D. Sbarbaro, R. Źbikowski, P.J. Gawthrop, Neural networks for control systems: a survey, Automatica 28 (1992) 1083–1112.
[34] H.G. Zhang, Y.B. Quan, Modeling, identification, and control of a class of nonlinear systems, IEEE Trans. Fuzzy Syst. 9 (2001) 349–354.
[35] M.W. Mak, K.W. Ku, Y.L. Lu, On the improvement of the real time recurrent learning algorithm for recurrent neural networks, Neurocomputing 24 (1999) 13–36.
[36] H.G. Zhang, Z.W. Liu, G.B. Huang, Z.S. Wang, Novel weighting-delay-based stability criteria for recurrent neural networks with time-varying delay, IEEE Trans. Neural Networks 21 (2010) 91–106.
[37] Y.S. Xia, G. Feng, J. Wang, A novel neural network for solving nonlinear optimization problems with inequality constraints, IEEE Trans. Neural Networks 19 (2008) 1340–1353.
[38] S.B. Liu, J. Wang, A simplified dual neural network for quadratic programming with its KWTA application, IEEE Trans. Neural Networks 17 (2006) 1500–1510.
[39] L. Cheng, Z.G. Hou, M. Tan, A neutral-type delayed projection neural network for solving nonlinear variational inequalities, IEEE Trans. Circuits Syst. II 55 (2008) 806–810.
[40] L. Cheng, Z.G. Hou, M. Tan, A delayed projection neural network for solving linear variational inequalities, IEEE Trans. Neural Networks 20 (2009) 915–925.
[41] M. Forti, On global asymptotic stability of a class of nonlinear systems arising in neural network theory, J. Differ. Equations 113 (1994) 246–264.
[42] E. Kaszkurewicz, A. Bhaya, On a class of globally stable neural circuits, IEEE Trans. Circuits Syst. I 41 (1994) 171–174.
[43] M. Forti, P. Nistri, Global convergence of neural networks with discontinuous neuron activations, IEEE Trans. Circuits Syst. I 50 (2003) 1421–1435.
[44] Z.G. Zeng, J. Wang, Improved conditions for global exponential stability of recurrent neural networks with time-varying delays, IEEE Trans. Neural Netw. 17 (2006) 623–635.
[45] Q.S. Liu, J. Wang, Finite-time convergent recurrent neural network with a hard-limiting activation function for constrained optimization with piecewise-linear objective functions, IEEE Trans. Neural Networks 22 (2011) 601–613.



Chen Qiao was born in Shaanxi Province, China. She received the B.S. degree in computational mathematics in 1999, the M.S. degree in applied mathematics in 2002, and the Ph.D. degree in applied mathematics in 2009, all from Xi'an Jiaotong University, China. Her current research interests include nonlinear dynamics systems, stability theory, neural networks and machine learning.

Wenfeng Jing received the M.S. and Ph.D. degrees in mathematics from Xi'an Jiaotong University, Xi'an, China, in 2002 and 2011, respectively. He is now with the School of Mathematics and Statistics, Xi'an Jiaotong University, where he was promoted to an associate professor in 2007. His current research interests include data mining and machine learning.

Zongben 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. He holds the title "Owner of Chinese Ph.D. 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 an academician of Chinese Academy of Sciences and a member of the New York Academy of Sciences.