Neurocomputing 175 (2016) 40–46

Contents lists available at ScienceDirect

Neurocomputing

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

The general critical analysis for continuous-time UPPAM recurrent neural networks ☆

Chen Qiao [a,b], Wen-Feng Jing [a], Jian Fang [a,b], Yu-Ping Wang [b,c,*]

[a] School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, PR China
[b] Department of Biomedical Engineering, Tulane University, New Orleans, LA 70118, USA
[c] Center of Genomics and Bioinformatics, Tulane University, New Orleans, LA 70112, USA

ARTICLE INFO

Article history:
Received 10 June 2015
Received in revised form
26 July 2015
Accepted 14 September 2015
Communicated by Ma Lifeng Ma
Available online 26 October 2015

Keywords:
Continuous-time recurrent neural network
Uniformly pseudo-projection-anti-monotone network
General critical condition
Dynamical analysis

ABSTRACT

The uniformly pseudo-projection-anti-monotone (UPPAM) neural network model, which can be considered as the unified continuous-time neural networks (CNNs), includes almost all of the known CNNs individuals. Recently, studies on the critical dynamic behaviors of CNNs have drawn special attentions due to its importance in both theory and applications. In this paper, we will present the analysis of the UPPAM network under the general critical conditions. It is shown that the UPPAM network possesses the global convergence and asymptotical stability under the general critical conditions if the network satisfies one quasi-symmetric requirement on the connective matrices, which is easy to be verified and applied. The general critical dynamics have rarely been studied before, and this work is an attempt to gain a meaningful assurance of general critical convergence and stability of CNNs. Since UPPAM network is the unified model for CNNs, the results obtained here can generalize and extend the existing critical conclusions for CNNs individuals, let alone those non-critical cases. Moreover, the easily verified conditions for general critical convergence and stability can further promote the applications of CNNs.

1. Introduction

The two basic elements of a recurrent neural network (RNN) are the synaptic connections among the neurons and the nonlinear activation functions deduced from the input–output properties of the involved neurons. For applications such as associative memory, synaptic connections among the neurons are designed to encode the memories we hope to recover. The activation functions are assumed to capture the complex, nonlinear response of neurons of the brain. For different purpose of simulations and applications, both of them are preassigned before use. So understanding their properties is very important, and especially exploring the characteristics of the activation functions is quite crucial to determine the performance of the RNNs. For the commonly used RNN individuals, the activation functions are monotonically nondecreasing and saturated. To study and apply RNNs only based on such two features are far from enough. To overcome the non-thorough descriptions of activation functions, many special cases of activation functions have been brought forward, resulting in many different RNNs individuals. Furthermore, in order to obtain more useful results of RNNs, e.g., the convergence and stability of those individuals, additional strict requirements are unavoidable to impose on the networks for the lack of in-depth descriptions on the activation functions. Obviously, since those individuals are studied separately, it is inevitable that there exist large numbers of redundancy of analysis for those individual models. In order to reduce the superabundance, establishing a harmonization methodology is a challenging work.

In [16], Xu and Qiao put forward two novel concepts: uniformly anti-monotone and the pseudo-projection properties of the activation functions, which discover more essential characteristics other than the nondecreasing and bounded properties of the commonly used activation functions. It is shown that the proposed uniformly pseudo-projection anti-monotone (UPPAM) operator can embody most of the activation operators (the precise definition of uniformly pseudo-projection-anti-monotone operator will be given in Section 2), e.g., nearest-point projection, linear saturating operator, signum operator, symmetric multi-valued step operator, multi-threshold operator, and winner-take-all operator. Thus, the UPPAM operator can be considered as a framework of formalizing most of the activation operators of RNNs.

☆This research was supported by NSFC Nos. 11101327, 11471006 and 11171270, and was partially supported by NIH R01 GM109068 and R01 MH104680.

* Corresponding author at: Department of Biomedical Engineering, Tulane University, New Orleans, LA 70118, USA.
E-mail addresses: qiaochen@mail.xjtu.edu.cn, cqiao@tulane.edu (C. Qiao), wfjing@mail.xjtu.edu.cn (W.-F. Jing), jfang3@tulane.edu (J. Fang), wyp@tulane.edu (Y.-P. Wang).

http://dx.doi.org/10.1016/j.neucom.2015.09.103
0925-2312/

In this paper, we use the concept of UPPAM operators to establish a unified model for continuous-time RNNs. Let us consider the following continuous-time UPPAM RNNs model:

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

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, and G owns the uniformly pseudo-projection-anti-monotone property. Both A and W are the connective weight matrices, b, q are two fixed external bias vectors and τ is the state feedback coefficient. The form of model (1) includes two basic kinds of continuous-time RNNs [17], i.e., the static RNNs and the local field RNNs. Furthermore, as proved in [16], most activation operators are special cases of the UPPAM operator. So, model (1) can be considered as a unified model of continuous-time RNNs and can include almost all of the existing continuous-time RNNs specials [4], e.g., Hopfield-type neural networks, Brain-State-in-a-Box neural networks, Recurrent Back-propagation neural networks, Mean-field neural networks, Bound-constraints Optimization Solvers, Convex Optimization Solvers, Recurrent Correlation Associative Memories neural networks, and Cellular neural networks. In addition, since model (1) owns the essential characteristics of the activation functions, i.e., the uniformly anti-monotone as well as the pseudo-projection properties, it can be expected that the analysis of model (1), especially the dynamics analysis can give more in-depth results and provide the unified and concise characterization of the continuous-time RNNs models. The main purpose of this paper will focus on discovering some essential global convergence and stability for the unified model (1), i.e., the critical convergence and stability.

For RNNs, one difficult problem of dynamics analysis lies in the critical analysis. Define a discriminant matrix

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

where Γ is a positive definite diagonal matrix, P is a diagonal matrix defined by the network, and W and A are the weight matrices. If there exist a positive definite diagonal matrix Γ, such that $S(\Gamma, 2\Lambda - B) > 0$ (i.e., $S(\Gamma, 2\Lambda - B)$ is positive definite), where Λ and B are the anti-monotone and pseudo-projection constant matrices of the network (the definitions of them are given in Section 2), then RNNs have exponential stability [4]. Many stability results have been achieved for RNNs individuals under various specifications of $S(\Gamma, 2\Lambda - B) > 0$ (typically, when $S(\Gamma, 2\Lambda - B) > 0$ is an M-matrix), and they are called as the non-critical dynamical analysis [1]. On the other hand, if there exists a positive definite diagonal matrix Γ such that $S(\Gamma, V)$ is negative definite, here $V = diag\{r_1, r_2, \ldots, r_N\}$ with each $r_i > 0$ being the maximum inversely Lipschitz constant of g_i (i.e., for all $s, t \in \mathcal{R}^N$, $|g_i(t) - g_i(s)| \geq r_i |t - s|$), then RNNs are globally exponentially unstable [7,1]. Since $S(\Gamma, 2\Lambda - B) > 0$ is the sufficient condition on the globally exponential stability of RNNs, and $S(\Gamma, V) \geq 0$ is the necessarycondition for RNNs to be globally stable, it is quite natural to explore the gap between $S(\Gamma, 2\Lambda - B) \leq 0$ (i.e., $S(\Gamma, 2\Lambda - B)$ is negative semi-definite) and $S(\Gamma, V) \geq 0$ (i.e., $S(\Gamma, V)$ is positive definite). Such a gap is called the general critical condition, and the dynamics analysis of RNNs under such condition is referred to as the general critical dynamics analysis.

For any application and practical design of RNNs, such as pattern recognition, associative memories, or as optimization solvers, the convergence and stability of RNNs are both prerequisite. For instance, when an RNN is used in associative memory or pattern recognition, any pattern we hope to store has to be an equilibrium point of the RNN. In addition, to ensure that each stored pattern can be retrieved even with noises, each equilibrium point must possess the stability. When the RNN is employed as an optimization solver, the possible optimal solutions correspond to the equilibrium of the RNN, and the convergence of the RNN is a guarantee of finding the optimal solutions. Since the general critical conditions can be considered essentially as the distinct region of stability and non-stability of RNNs, studying the general critical dynamic behaviors of an RNN can find broad applications.

Recently, due to the difficulty in the dynamical analysis of RNNs for general critical conditions, most of the studies on critical analysis have been focused on the special critical conditions, i.e., considering the asymptotic behaviors of RNNs under the condition that $S(\Gamma, 2\Lambda - B) \geq 0$ (this is because $S(\Gamma, 2\Lambda - B) > 0$ is already known to be globally exponential stable and $S(\Gamma, 2\Lambda - B) = 0$ is a special case of the general critical condition). Even for this special critical condition, there only exist a few results since the analysis is much more difficult than the dynamics analysis under the non-critical condition that $S(\Lambda, L) > 0$. In [15], the globally exponential stability of a static neural network with projection operator (a special kind of UPPAM operator) has been proven under the condition that I–W is nonnegative (which is a special case of $S(\Gamma, 2\Lambda - B) \geq 0$). The special critical convergence of a static neural network model with nearest point projection activation operator (special case of projection operator) on a region defined by the network has been achieved in [1] when W is quasi-symmetric. Some general critical stability conclusions for the static and the local field continuous-time RNNs with projection activation operators have been achieved in [2], but they require the network to satisfy one bounded matrix norm. In [4], for the presented unified continuous-time RNNs, namely, UPPAM RNNs, the special critical global convergence is obtained with some bound requirements on the defined nonlinear norm, but such requirements cannot be verified easily in applications. In [5], some improvements on dynamics analysis of the UPPAM networks have been obtained, while they are still under the special critical conditions.

In the present paper, we give some solutions on how to assure the convergence and stability under the general critical conditions. By applying the energy function method and LaSalle invariance principle to the unified continuous-time RNNs model (1), we obtain the global convergence and asymptotic stability under some general critical conditions, that is, $S(\Gamma, 2\Lambda - B) + \Psi$ is positive definite for one diagonal matrix Ψ. The results only require the network to satisfy some quasi-symmetric conditions on the connection matrices. Since the conclusions obtained here are for the unified RNNs model under the general critical conditions, they can sharpen and generalize, to a large extent, the latest critical results given by [1,2,4,5,15], and they can further be extended to those non-critical conclusions (see, e.g., [6–14,18–25] and the references quoted there). Furthermore, they can be applied directly to many individual RNN models mentioned above. They can be widely applied to solve the linear variational inequality and many other optimization problems, etc. Therefore, the study here provides an insight on the unified continuous-time RNNs model with critical analysis.

2. Preliminaries

For the activation operator G, the domain, range and fixed-point set of G are respectively defined by $\mathbf{D}(G)$, $\mathbf{R}(G)$ and $\mathbf{F}(G)$, and $\mathbf{D}(G) = \mathbf{R}(G) \subseteq \mathbb{R}^N$. 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

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

G is said to be diagonal if $g_i(x) = g_i(x_i)$ holds for each $i = 1, 2, \ldots, N$.

Definition 2.1 (*Xu and Qiao* [16]). (i) An operator G is said to be a pseudo-projection if there exists a positive definite diagonal matrix $B=diag\{\beta_1,\beta_2,\dots,\beta_N\}$, 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) An operator G is said to be λ-uniformly anti-monotone ($\lambda-UAM$) 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) An 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*.

Definition 2.2. Let $\Lambda=diag\{\lambda_1,\lambda_2,\dots,\lambda_N\}$ and $B=diag\{\beta_1,\beta_2,\dots,\beta_N\}$. 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 [16], it is shown that most of the activation operators of RNNs in the literature are special cases of UPPAM operators. Thus, the RNNs with UPPAM operators, i.e., the uniformly pseudo-projection-anti-monotone neural networks provide an appropriate and unified framework, within which most of the known RNN models can be embedded and uniformly studied.

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

3. General critical dynamic results

In this section, under the general critical conditions, results on global convergence and asymptotic stability for the unified continuous-time RNN model are established, which are quite easy to be verified in applications. In the following, we denote the equilibrium state set of (1) by $F_e^{-1}(0)$, and the range of nonlinear activation operator, i.e., $\mathbf{R}(G)$, by Θ. Throughout this paper, we suppose that Θ is bounded, closed and convex.

Lemma 3.1. *For any $x_0\in A(\Theta)+b$, $x(t,x_0)$, the solution of (1), satisfies $x(t,x_0)\in A(\Theta)+b$ $(t\geq0)$.*

Proof. With the differential equation theory, we have

$$\begin{aligned}x(t,x_0)&=e^{-t/\tau}x_0+\frac{1}{\tau}e^{-t/\tau}\int_0^t e^{s/\tau}(AG(Wx(s)+q)+b)\,ds\\&=e^{-t/\tau}x_0+(1-e^{-t/\tau})\frac{\int_0^{t/\tau}e^{-r}(AG(Wx(t-\tau r)+q)+b)\,dr}{1-e^{-t/\tau}}\end{aligned}\tag{3}$$

where $r=\frac{t-s}{\tau}$. Since $1-e^{-t/\tau}=\int_0^{t/\tau}e^{-r}\,dr=\lim_{n\to+\infty}\sum_{i=1}^n\frac{t}{\tau n}e^{-it/\tau n}$, 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 $\sum_{i=1}^n\frac{t}{\tau n}e^{-it/\tau n}(AG(Wx(t-\frac{it}{n})+q)+b)$, should satisfy $\frac{P(t)}{1-e^{-t/\tau}}\in A(\Theta)+b$ $(\forall t\geq0)$. Further, by (3), we know $x(t,x_0)\in A(\Theta)+b$ when $x_0\in A(\Theta)+b$.□

For any $v\in\Theta$, we define $T(v)=AG(Wv+q)+b$. Since that Θ is bounded, closed and convex, then by Brouwer's fixed point theorem, T has at least one fixed point v^*, so namely, $F_e^{-1}(0)$ 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 and a diagonal matrix Ψ, such that $(2\Lambda-B)\Gamma-\Gamma AW+\Psi$ is positive definite, QAW is symmetric (here $Q=((2\Lambda-B)\Gamma+\Psi)\Lambda^{-1}$) and Q is a positive definite diagonal matrix, then RNN model (1) is globally convergent on $A(\Theta)+b$ when $F_e^{-1}(0)$ is disconnected. Moreover, when x^* is the unique equilibrium point of (1), then x^* is globally asymptotically stable on $A(\Theta)+b$.*

Proof. Denote $A=diag\{a_1,a_2,\dots,a_N\}$, $\Gamma=diag\{\xi_1,\xi_2,\dots,\xi_N\}$, $D=diag\{d_1,d_2,\dots,d_N\}$ and $\Psi=diag\{\varphi_1,\varphi_2,\dots,\varphi_N\}$. 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 $u(t)=z(t)-x(t)$.

Let

$$\begin{aligned}E(x(t))&=\frac{\tau}{2}x^T(t)(QB-(Q+\Gamma)AW)x(t)-\tau(QAq)^Tx(t)-\tau(QBb)^Tx(t)\\&\quad+\tau\sum_{i=1}^N\xi_ia_i\int_{(Wx_0+q)_i}^{(Wx(t)+q)_i}(a_ig_i(s)+b_i)\,ds\end{aligned}$$

Since $x(t)\in A(\Theta)+b$, there exists $p(t)\in\Theta$, such that $x(t)=Ap(t)+b$. Then, $Ap(t)=x(t)-b$. Meanwhile, noting that QB, QAW and ΓAW are all symmetric, and thus $W^T(\Gamma A)=(\Gamma AW)^T=\Gamma AW$ for the case that both A and Γ are diagonal. Then, a direct calculation shows

$$\begin{aligned}\frac{dE(x(t))}{dt}&=\langle(QB-(Q+\Gamma)AW)x(t),u(t)\rangle-\langle QAq,u(t)\rangle-\langle QBb,u(t)\rangle+\langle\Gamma Az(t),Wu(t)\rangle\\&=-\langle QAWx(t),u(t)\rangle-\langle QAq,u(t)\rangle\\&\quad+\langle QBx(t),u(t)\rangle-\langle QBb,u(t)\rangle+\langle\Gamma AWz(t),u(t)\rangle-\langle\Gamma AWx(t),u(t)\rangle\\&=-\langle QA(Wx(t)+q),u(t)\rangle+\langle QB(x(t)-b),u(t)\rangle\\&\quad+\langle\Gamma AWz(t),u(t)\rangle-\langle\Gamma AWx(t),u(t)\rangle\\&=-\langle QAy(t),u(t)\rangle+\langle QBAp(t),u(t)\rangle+\langle\Gamma AWz(t),u(t)\rangle-\langle\Gamma AWx(t),u(t)\rangle\\&=-\langle QAy(t),u(t)\rangle+\langle QABp(t),u(t)\rangle+\langle\Gamma AW(z(t)-x(t)),u(t)\rangle\\&=-\langle QA(y(t)-Bp(t)),u(t)\rangle+\langle\Gamma AWu(t),u(t)\rangle\\&=-\langle QA(y(t)-Bp(t)),u(t)\rangle-\langle((2\Lambda-B)\Gamma+\Psi-\Gamma AW)u(t),u(t)\rangle\\&\quad+\langle((2\Lambda-B)\Gamma+\Psi)u(t),u(t)\rangle\end{aligned}\tag{4}$$

Since G is a B-projection and $p(t)\in\Theta$, it is clear that $G(Bp(t))=p(t)$. Denote the diagonal matrix $Q=diag\{q_1,q_2,\dots,q_N\}$. By the diagonal nonlinear property of G, one can get that

$$\begin{aligned}&-\langle QA(y(t)-Bp(t)),u(t)\rangle\\&=-\sum_{i=1}^Na_iq_i((Wx(t)+q)_i-\beta_ip_i(t))\cdot((a_ig_i((Wx(t)+q)_i)+b_i)-(a_ip_i(t)+b_i))\\&=-\sum_{i=1}^Na_i^2q_i((Wx(t)+q)_i-\beta_ip_i(t))\cdot(g_i((Wx(t)+q)_i)-g_i(\beta_ip_i(t)))\end{aligned}$$

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

$$((Wx(t)+q)_i-\beta_ip_i(t))\cdot(g_i((Wx(t)+q)_i)-g_i(\beta_ip_i(t)))\geq\lambda_i(g_i((Wx(t)+q)_i)-g_i(\beta_ip_i(t)))^2$$

Further, since both Q and Λ are positive definite diagonal matrices, we have

$$\begin{aligned}-\langle QA(y(t)-Bp(t)),u(t)\rangle&\leq-\sum_{i=1}^Na_i^2q_i\lambda_i\big(g_i\big((Wx(t)+q)_i\big)-g_i(\beta_ip_i(t))\big)^2\\&=-\sum_{i=1}^Nq_i\lambda_i\big(a_i(g_i((Wx(t)+q)_i)-p_i(t))\big)^2\\&=-(A(G(Wx+q)-p(t)))^T\Lambda Q(A(G(Wx+q)-p(t)))\\&=-u^T(t)Q\Lambda u(t)\\&=-\langle((2\Lambda-B)\Gamma+\Psi)u(t),u(t)\rangle\end{aligned}\tag{5}$$

Then from (4), we directly have

$$\frac{dE(x(t))}{dt}\leq-\langle((2\Lambda-B)\Gamma+\Psi-\Gamma AW)u(t),u(t)\rangle\tag{6}$$

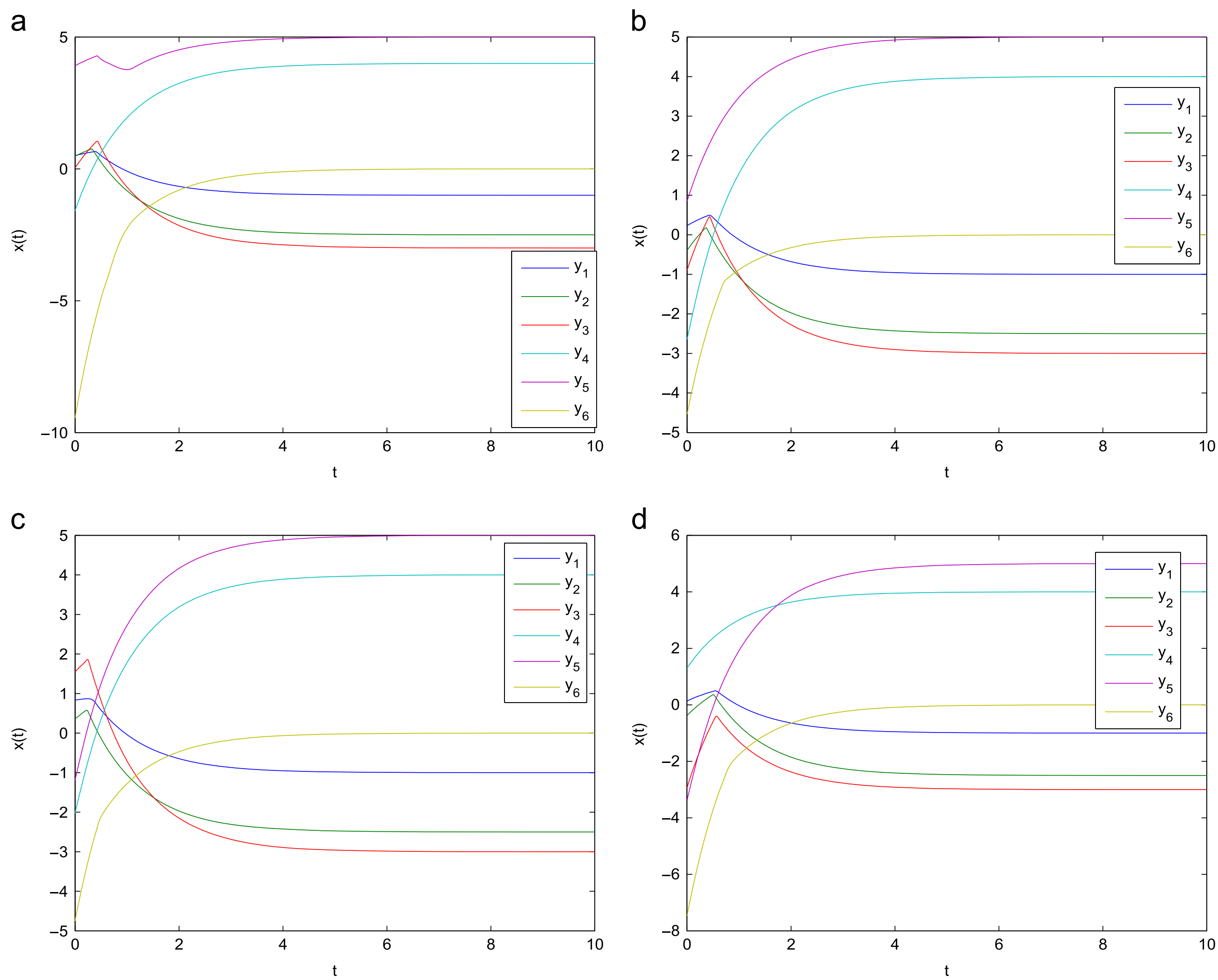


Fig. 1. Transient behaviors of RNN in system (8) with random initial points $x_0 \in \mathbf{R}(G)$.

By $S=(2\Lambda-B)\Gamma-\Gamma AW+\Psi$ being positive definite, we get that $\lambda_{min}(S)>0$. From (6), we know

$$\frac{dE(x(t))}{dt} \le -\lambda_{min}(S)\|z(t)-x(t)\|_2^2 \tag{7}$$

Obviously, it can be deduced that $\frac{dE(x(t))}{dt} \le 0$, and the equal sign holds if and only if $z(t)=x(t)$, i.e., $x(t)\in F_e^{-1}(0)$. Moreover, since $x(t)\in A(\Theta)+b$ is bounded and $F_e^{-1}(0)$ is disconnected, then by LaSalle invariance principle [26], we know that RNN model (1) is globally convergent on Θ. Furthermore, when $F_e^{-1}(0)=\{x^*\}$, it is easy to deduce that x^* is both attractive and stable on $A(\Theta)+b$ since $A(\Theta)+b$ is bounded, i.e., x^* is globally asymptotically stable on $A(\Theta)+b$. Thus, Theorem 3.1 is proved.□

Corollary 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 $\Gamma(2\Lambda-AW)$ is positive definite, then RNN model (1) is globally convergent on $A(\Theta)+b$ when $F_e^{-1}(0)$ is disconnected. Moreover, when x^* is the unique equilibrium point of (1), then x^* is globally asymptotically stable on $A(\Theta)+b$.*

Proof. Let $\Psi=B\Gamma$ in Theorem 3.1, then we have $(2\Lambda-B)\Gamma-\Gamma AW+\Psi=\Gamma(2\Lambda-AW)$. Since both Λ and Γ are diagonal and positive definite matrices, we have $Q=2\Lambda\Gamma\Lambda^{-1}=2\Gamma$ is positive definite diagonal matrix, and $QAW=2\Gamma AW$ is obviously symmetric by the positive definite requirement of $\Gamma(2\Lambda-AW)$. So the corollary follows directly from Theorem 3.1.□

Remark 3.1. Studying the dynamic behaviors of unified model (1) can provide uniform results for RNNs and thus can deduce the numerous redundancy existing in the RNNs individuals. Further, since model (1) owns the pseudo-projection as well as the anti-monotone property, then by utilizing these two properties, we can obtain some meaningful conclusions.

Recently, the dynamic studies of RNNs have attracted great interest in the critical analysis. It should be pointed out that due to the difficulty in analysis, most of them are based on the special critical conditions, i.e., discriminant matrix $S(\Gamma, 2\Lambda-B)$ is positive semi-definite. In addition, in order to assure the stability, some other restrictions are required on the networks [1,2,4,5,15]. Obviously, just studying the special critical dynamics is far from enough in both theory and applications, and additional requirements on the networks are quite hard for applications.

Theorem 3.1 and Corollary 3.1 exploit new methods to assure the global asymptotical stability and global convergence for the unified continuous-time RNN model (1). The results obtained here are under the general critical conditions, and do not need difficultly verified requirements. For Theorem 3.1, in addition to the general critical conditions, it only requires the UPPAM network to meet quasi-symmetric conditions. That is because, in the sense of positive definite, one can easily choose a diagonal matrix Ψ in Theorem 3.1

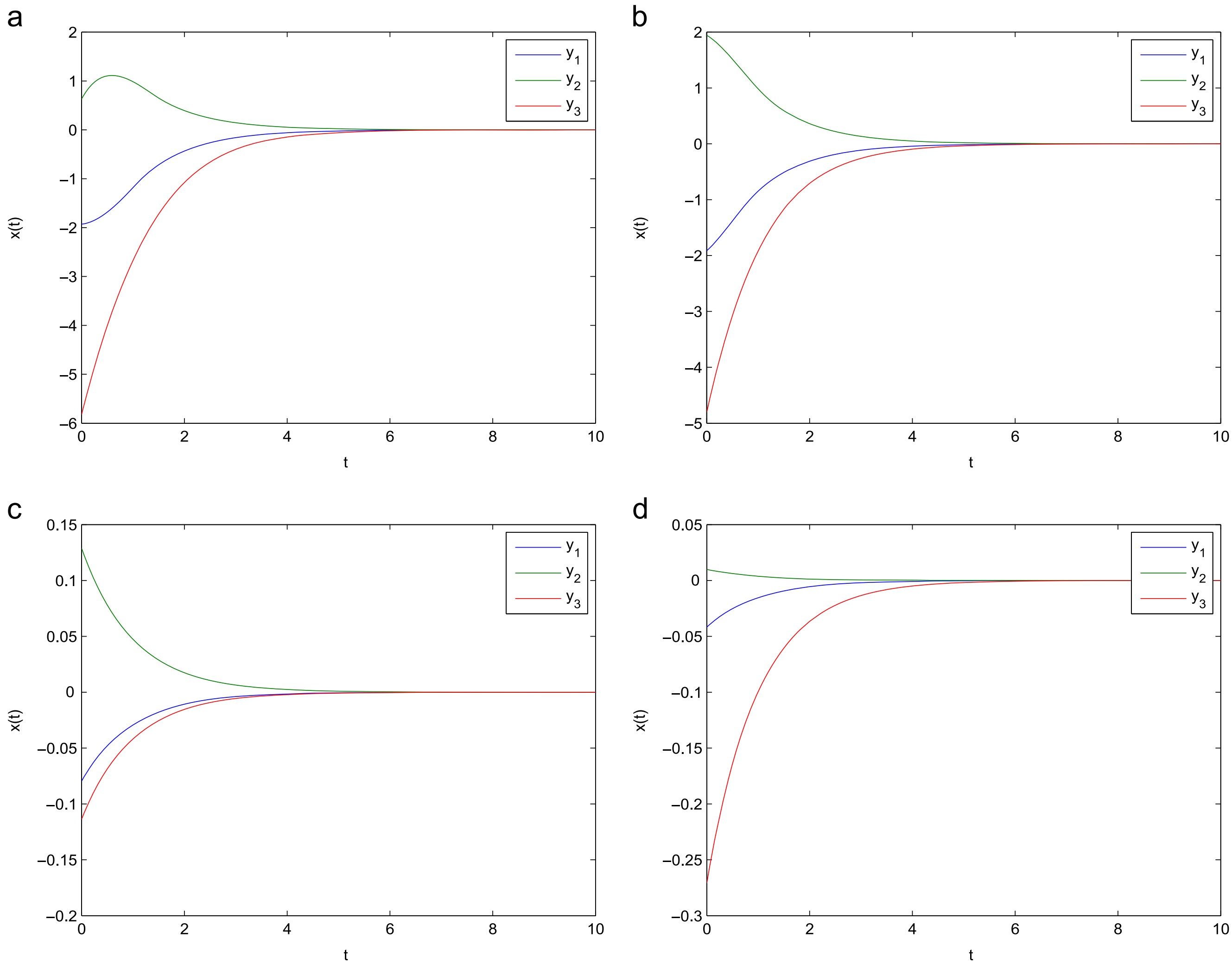


Fig. 2. Transient behaviors of RNN in system (9) with random initial points $x_0 \in \mathbf{R}(G)$.

satisfying $\Psi > (B-2\Lambda)\Gamma$, where B, Λ and Γ all are positive definite matrices. Thus, Q is positive definite. Then by Theorem 3.1, in order to assure the global stability and convergence, we only need to verify that QAW is symmetric, where both A and W are connection matrices of the network. Corollary 3.1 shows that existing critical results are the special cases of the results obtained in this paper. The critical condition that $\Gamma((2\Lambda-B)-AW) \geq 0$ in [1,4,5,15] is a special case of $(2\Lambda-B)\Gamma-\Gamma AW+\Psi \geq 0$. Further, since both Γ and B are positive definite diagonal matrices, thus $\Gamma((2\Lambda-B)-AW) \geq 0$ is a particular case of $\Gamma(2\Lambda-AW) > 0$. The latter is just the only requirement in Corollary 3.1 to guarantee the global stability and convergence for model (1). The critical dynamic conclusions of Theorem 3.1 and Corollary 3.1 not only summarize, but also deepen to a large extent most of the existing results for the RNNs individuals.

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 cannot be applied here.

Example 4.1. Consider the following UPPAM RNN:

$$\tau\frac{dx(t)}{dt} = -x(t)+AG(Wx(t)+q)+b, \quad x_0 \in \mathbb{R}^6 \tag{8}$$

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

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

In this example,

$$W=\begin{pmatrix} 3.8640 & 2.5440 & 4.7200 & 4.7480 & 2.4920 & 3.3960 \\ -1.7989 & -1.9307 & -1.9516 & -2.0704 & -1.3237 & -1.7225 \\ 2.7251 & 1.5935 & 3.5839 & 3.0946 & 1.4318 & 2.4942 \\ -2.3740 & -1.4640 & -2.6800 & -2.7505 & -1.0970 & -1.9410 \\ 1.1145 & 0.8372 & 1.1091 & 0.9812 & 0.7587 & 0.8130 \\ -1.3864 & -0.9945 & -1.7636 & -1.5848 & -0.7422 & -1.2680 \end{pmatrix}$$

and $A=diag\{-1,2,-3,4,-5,6\}$, $q=[-1.2290,-0.9133,0.1524,2.8258,-1.5383,0.9961]^T$, $b=[-1.782,1.4427,-0.1067,1.9619,3.0046,-2.7749]^T$.

In addition, $\Lambda=B=diag\{1,1/2,1/3,1/4,1/5,1/6\}$, and the equilibrium state set, $F_e^{-1}(0)$, is a single state set, which contains only one state $(-1,-2.5,-3,4,5,0)^T$.

For this UPPAM network, almost all of the existing stability conclusions cannot be used here. That is because for any positive definite diagonal matrix Γ, $(2\Lambda - B)\Gamma - \Gamma AW$ is neither positive definite nor positive semi-definite, so all the non-critical and the special critical dynamical results (see, the reference mentioned in the Introduction section) are not suitable here. In addition, since this example is for the unified RNNs model, it is totally hard to calculate the nonlinear norm of the network, and the conclusions in [2] cannot be used here.

In what follows, we will show that Theorem 3.1 established in this paper can be successfully applied here. By setting $\Gamma = diag\{1, 1/\sqrt{2}, 1/\sqrt{3}, 1/\sqrt{4}, 1/\sqrt{5}, 1/\sqrt{6}\}$ and $\Psi = \Lambda\Gamma$, we have $(2\Lambda - \mathrm{B})\Gamma - \Gamma\mathrm{AW} + \Psi \geq 0, Q = ((2\Lambda - \mathrm{B})\Gamma + \Psi)\Lambda^{-1})$ is a positive definite diagonal matrix and QAW is symmetric. Then by Theorem 3.1, it is quite easy to achieve the global convergence of network (8) on $A(\Theta) + b$ with $\Theta = [-1, 1]^6$. Fig. 1 depicts the time responses of neural state variables of the system starting randomly from $A(\Theta) + b$.

Example 4.2. Consider another UPPAM RNN:

$$\tau \frac{dx(t)}{dt} = -x(t) + AG(Wx(t) + q) + b, \quad x_0 \in \mathbb{R}^N \tag{9}$$

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

$g_i(s) = \min(\max(-1, s), 1)$

$W = (W_{ij})_{N \times N}$ with each W_{ij} is

$$W_{ij} = \begin{cases} (-1)^{i+1} \cdot \left(\dfrac{1}{N(N-1)} + \dfrac{1}{iN(N-1)} \right), & i \neq j \\ (-1)^{i+1} \cdot \left(\dfrac{1}{i} - \dfrac{1}{N} + \dfrac{1}{iN(N-1)} \right), & i = j \end{cases} \tag{10}$$

and $A = diag\{a_1, a_2, \cdots, a_N\}$ with each $a_i = (-1)^{i-1} \cdot i$. q and b are two N-dim vectors, $q_i = (-1)^{i+1}[1 + (\frac{1}{i} - \frac{1}{N} + \frac{a}{i}) + (N-i) \cdot (a + \frac{a}{i})]$ and $b_i = (-1)^i i$.

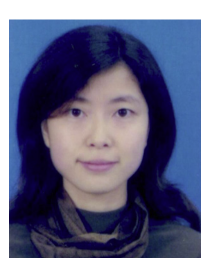
In this case, $\Lambda = B = I$. For any positive diagonal matrix $\Gamma, \Gamma(2\Lambda - B - AW)$ is not positive semi-definite, i.e., the latest critical results in [5] and other recent results for the special critical analysis, e.g., [2–4] all cannot be used for this example. When $\Gamma = diag\{1, 1/2, \ldots, 1/N\}$, it is easy to verify that $\Gamma(2\Lambda - AW)$ is positive definite. Then by Corollary 3.1, network (9) is global convergent to the unique equilibrium, i.e., the origin. Fig. 2 depicts the time responses of neural state variables of network (9) with $N=3$ starting randomly from $A(\Theta) + b$, where $\Theta = [-1, 1]^3$.

5. Conclusion

In the present paper, based on the unified RNN model, i.e., the uniformly pseudo-projection-anti-monotone RNNs model, the corresponding global convergence and the global asymptotic stability under the general critical conditions are given. In addition to the general critical conditions, our conclusions only require that the synaptic connective matrices defined by the network are quasi-symmetric. Compared the existing dynamical analysis results for RNNs, the conclusions obtained here demonstrate several advantages. Firstly, they are given under the general critical conditions, which have scarcely been studied before. Secondly, they are for the unified RNN model, so they can be applied to almost all of the existing individuals of RNNs and can be used directly in applications. Finally, with our results there is no need to verify additional intricate requirements on the network, so they can be easily used. In summary, the results achieved here are a significant step towards establishing a unified theory for the dynamics of 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, On the P-critical dynamics analysis of projection recurrent neural networks, Neurocomputing 73 (2010) 2783–2788.
[3] C. Qiao, Z.B. Xu, Critical dynamics study on recurrent neural networks: globally exponential stability, Neurocomputing 77 (2012) 205–211.
[4] C. Qiao, W.F. Jing, Z.B. Xu, The UPPAM continuous-time RNN model and its critical dynamics, Neurocomputing 77 (2013) 158–166.
[5] C. Qiao, H.B. Chen, W.F. Jing, K.F. Sun, Towards establishing a meaningful and practical dynamics results for the unified RNN model, Neurocomputing 157 (2015) 315–322.
[6] 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.
[7] 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.
[8] Y.S. Xia, J. Wang, On the stability of globally projected dynamical systems, J. Optim. Theory Appl. 106 (2000) 129–150.
[9] Z.B. Xu, H. Qiao, J. Peng, A comparative study of two modeling approaches in neural networks, Neural Netw. 17 (2004) 73–85.
[10] A. Slavova, Cellular Neural Networks: Dynamics and Modelling, Kluwer Academic Publishers, Dordrecht, NL, 2003.
[11] Z.H. Guan, G. Chen, Y. Qin, On equilibria, stability and instability of Hopfield neural networks, IEEE Trans. Neural Netw. 11 (2000) 534–540.
[12] 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 Netw. 12 (2001) 349–359.
[13] 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.
[14] H. Qiao, J. Peng, Z.B. Xu, Nonlinear measures: a new approach to exponential stability analysis for Hopfield-type neural networks, IEEE Trans. Neural Netw. 12 (2001) 360–370.
[15] Y.Q. Yang, J. Cao, Solving quadratic programming problems by delayed projection neural network, IEEE Trans. Neural Netw. 17 (2006) 1630–1634.
[16] 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.
[17] Z.B. Xu, H. Qiao, J. Peng, A comparative study of two modeling approaches in neural networks, Neural Netw. 17 (2004) 73–85.
[18] 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 Netw. 21 (2010) 91–106.
[19] Y.S. Xia, G. Feng, J. Wang, A novel neural network for solving nonlinear optimization problems with inequality constraints, IEEE Trans. Neural Netw. 19 (2008) 1340–1353.
[20] S.B. Liu, J. Wang, A simplified dual neural network for quadratic programming with its KWTA application, IEEE Trans. Neural Netw. 17 (2006) 1500–1510.
[21] 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.
[22] L. Cheng, Z.G. Hou, M. Tan, A delayed projection neural network for solving linear variational inequalities, IEEE Trans. Neural Netw. 20 (2009) 915–925.
[23] M. Forti, P. Nistri, Global convergence of neural networks with discontinuous neuron activations, IEEE Trans. Circuits Syst. I 50 (2003) 1421–1435.
[24] 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.
[25] 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 Netw. 22 (2011) 601–613.
[26] J.P. LaSalle, The Stability of Dynamical Systems, SIAM, Philadelphia, PA, 1987.



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. She is now an associate professor of Xi'an Jiaotong University. Her current research interests include nonlinear dynamics systems, neural networks and machine learning. At present, she is a Postdoc fellow of Tulane University.

Wen-Feng 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.

Jian Fang is a Ph.D. candidate of mathematics from Xi'an Jiaotong University, Xi'an, China. His current research interests include sparse SAR imaging, compressed MRI imaging, image classification, and sparse coding.

Yu-Ping Wang received the B.S. degree in applied mathematics from Tianjin University, China, in 1990, and the M.S. degree in computational mathematics and the Ph.D. degree in communications and electronic systems from Xi'an Jiaotong University, China, in 1993 and 1996, respectively. After his graduation, he had visiting positions at the Center for Wavelets, Approximation and Information Processing of the National University of Singapore and Washington University Medical School in St. Louis. From 2000 to 2003, he worked as a senior research engineer at Perceptive Scientific Instruments, Inc., and then Advanced Digital Imaging Research, LLC, Houston, Texas. In the fall of 2003, he returned to academia as an assistant professor of computer science and electrical engineering at the University of Missouri-Kansas City. He is currently an associate professor of Biomedical Engineering and Biostatistics & Bioinformatics at Tulane University School of Science and Engineering & School of Public Health and Tropical Medicine. He is also a member of Tulane Center of Bioinformatics and Genomics, Tulane Cancer Center and Tulane Neuroscience Program. His research interests have been computer vision, signal processing and machine learning with applications to biomedical imaging and bioinformatics, where he has about 150 peer reviewed publications. He has served on numerous program committees and NSF/NIH review panels, and served as editors for several journals such as Neuroscience Methods.