Critical dynamics study on recurrent neural networks: Globally exponential stability☆

Chen Qiao *, Zongben Xu 1

Institute for Information and System Science, Faculty of Science, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

1. Introduction

Recurrent neural networks (RNNs) are dynamic systems that can be implemented by physical means, and they are mainly used to model dynamic process associated with control process, perform associative memory and solve optimization problems. In the present paper, we consider the generic continuous-time RNNs modeled by the following nonlinear differential equation:

\[
\frac{d y(t)}{d t} = -D y(t) + W F(y(t) + b) + q, \quad y_0 \in \mathbb{R}^N, \tag{1}
\]

where \( y = (y_1, y_2, \ldots, y_N)^T \) is the neural network state, \( D = \text{diag}(d_1, d_2, \ldots, d_N) \) is a positive matrix with each \( d_i \) being the state feedback coefficients, \( W = (w_{ij})_{N \times N} \) is the connective weight matrix, \( A \) is an \( N \times N \) diagonal matrix, \( b, q \) are two fixed external bias vector and \( F : \mathbb{R}^N \to \mathbb{R}^N \) is the nonlinear activation mapping.

As we know, there are two fundamental models which can summarize most of the existing RNNs specials [23], i.e., depending upon whether neural states or local fields states are taken as basic variables, a RNN can frequently be classified either as a local field neural network model or as a static neural network model. These two basic RNNs models are extensively applied in learning, pattern recognition, associative memory, solving optimization problems, etc. It should be noted that by taking \( A = I \) and \( b = 0 \), model (1) corresponds the local field neural network model, and by choosing \( W = I \) and \( q = 0 \), model (1) refers to the static neural network model. Actually, model (1) describes uniformly various continuous-time RNNs models studied in literature, for example, Hopfield-type neural networks, recurrent back-propagation neural networks, mean-field neural networks, recurrent correlation associative memories neural networks, bound-constraints optimization solvers, convex optimization solvers, brain-state-in-a-box neural networks, cellular neural networks, and so on.

The analysis of the dynamical behaviors, such as the global convergence, asymptotic stability and exponential stability is a first and necessary step for any practical design and application of RNNs. In recent years, considerable efforts have been devoted to the analysis on the stability and convergence of RNNs without and with delay (see, e.g., the contributions of [2–4,6–8,11,13,17,18] and the references therein). If we define

\[
S(A, P) = AP^{-1} - \frac{AAW + (AAW)^T}{2},
\]

where both \( A \) and \( P \) are positive definite diagonal matrices, and \( W \) is the weight matrix of the network, then by generalizing these existing stability results of RNNs, it should be noticed that most of them are on the exponential stability analysis under the...
conditions that for one positive definite diagonal matrix \( A, S(A, L) \) is positive definite, where \( L = \text{diag}(l_1, l_2, \ldots, l_N) \) with each \( l_i > 0 \) being the Lipschitz constant of \( f_i \) and \( F = (f_1, f_2, \ldots, f_N) \) is the activation mapping of the network. On the other hand, [18, 14] have proved that a RNN will be globally exponentially unstable if there is a positive definite diagonal matrix \( A \) such that \( S(A, V) \) is negative definite, where \( V = \text{diag}(r_1, r_2, \ldots, r_N) \) with each \( r_i > 0 \) being the inversely Lipschitz constant of \( f_i \) (i.e., for all \( s.t \in \mathbb{R}^N, |f_i(t) - f_i(s)| \geq r_i |t - s| \)). From the definitions of Lipschitz constant and inversely Lipschitz constant, we have \( r_i \leq l_i \). For any positive definite diagonal matrix \( Q = \text{diag}(q_1, q_2, \ldots, q_N) \) with each \( q_i \) satisfying

\[
    r_i \leq q_i \leq l_i, \quad i = 1, 2, \ldots, N,
\]

then by using \( A \leq B \) to denote the condition that matrix \( B - A \) is nonnegative definite, it is clear that in the sense of nonnegative definition, there holds the following inequality relation:

\[
    S(A, L) \leq S(A, Q) \leq S(A, V),
\]

where \( A \) is a given positive definite diagonal matrix. Since \( S(A, L) > 0 \) (i.e., \( S(A, L) \) is positive definite) is sufficient for the globally exponential stability of RNNs, and \( S(A, V) \geq 0 \) (i.e., \( S(A, V) \) is nonnegative definite) is necessary for RNNs to have globally stable dynamics, the questions then arise: what kinds of asymptotic behavior of RNNs will be under our current investigation.

For all that, there are still many important dynamics questions of RNNs unsettled under the critical conditions. For example, with what additional requirement will the Sigmoidal RNNs be globally exponential stability when \( S(A, L) \geq 0 \), and what kinds of asymptotical behavior will be when \( S(A, Q) = 0 \)? For a RNN with general projection mapping, does there exist some convergence results when \( S(A, L) < 0 \), or, under the general critical condition that \( S(A, Q) = 0 \)? Further, for neural network with a generic activation mapping, what asymptotic behaviors of it will be under the critical conditions that \( S(A, L) \geq 0 \), or further, \( S(A, Q) = 0 \)? All these are under our current investigation.

In the current paper, we devote to answer the question that what dynamics behavior will happen for RNNs when \( S(A, L) \geq 0 \). By exploring some intrinsic features processed naturally by nonlinear activation mappings, e.g., the decreasing anti-monotone as well as the uniformly anti-monotone properties, and by applying the energy function method, it is shown that a RNN has a unique equilibrium state and which is globally asymptotically stable if there exists a positive definite diagonal matrix \( A \), such that \( S(A, L) \) is nonnegative definite. What is more important, based on matrix measure theory, we get that a RNN is globally exponentially stable under the conditions that one matrix (which is defined by the network and is very similar to \( S(A, L) \)) is positive definite at the unique equilibrium state. The results obtained here sharpen and generalize, to a large extent, most of the existing non-critical conclusions (see, e.g., [1–3, 6–8, 10, 11, 13, 17, 18, 22, 23] and the references quoted there) as well as the latest critical results given by [4, 5, 12, 14, 24]. Furthermore, they can provide a wider application range for RNNs and can be applied directly to many concrete RNN models, e.g., the well-known recurrent back-propagation neural networks, Hopfield-type neural networks, recurrent correlation associative memories with exponential activation mapping, etc.

2. Preliminaries

In the present investigation, we denote by \( \mathbb{R}^N \) the \( N \)-dimensional real vector space with norm \( \| \cdot \| \). Given any \( N \times N \) matrix \( A^T \) be its transpose, \( \sigma(A) \) be its spectral set (i.e., all eigenvalues of \( A \)), \( |A| \) and \( \mu(A) \) be its matrix norm and matrix measure that are, respectively, defined by

\[
    |A| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}, \quad \mu(A) = \lim_{\lambda \rightarrow 0^+} \frac{\|I + \lambda A\| - 1}{\lambda},
\]

where \( I \) is the identity matrix. Use \( \sigma(A) = \max(\Re \lambda : \lambda \in \sigma(A)) \) to denote the maximum abscissa of \( \sigma(A) \). Associated with a mapping \( F \), denoted by \( D(F) \) and \( R(F) \) its domain and range, respectively, \( F \) is said to be diagonally nonlinear if \( F \) is defined componentwise by

\[
    F(x) = (f_1(x_1), f_2(x_2), \ldots, f_N(x_N))',
\]

where each \( f_i \) is a one-dimensional nonlinear function. Assume that \( f_i \) is Lipschitz continuous. \( L_i \), the minimum Lipschitz constant of \( f_i \), is defined as follows

\[
    L_i = \sup_{t_j, t_s, t_x \neq 0} \frac{|f_i(t) - f_i(s)|}{|t - s|}.
\]

Without loss of generality, through out this paper, we let \( L = \text{diag}(L_1, L_2, \ldots, L_N) \) and assume that each \( L_i > 0 \). The equilibrium state set of system (1) is denoted by \( \Omega \).

**Definition 1** (Qiao et al. [18]). \( F \) is said to be \( \beta \)-uniformly anti-monotone (\( \beta \)-UAM) if there is a positive constant \( \beta > 0 \) such that for any \( x, y \in D(F) \), it satisfies

\[
    \|F(x) - F(y)\|^2 \leq \beta \cdot \langle F(x) - F(y), x - y \rangle.
\]

Obviously, a \( \beta \)-UAM mapping is a monotone mapping.

**Definition 2.** For a given point \( x^* \in D(F) \), \( F \) is called \( C_F \)-decreasing anti-monotone (\( C_F \)-DAM) at \( x^* \), if there exists a constant \( C_F > 0 \) and a monotonically decreasing function \( \beta_{F}(\cdot) : [0, +\infty) \rightarrow (0, C_F] \) which attains its maximum only at zero (i.e., \( \beta_{F}(t) = C_F \) if and only if \( t = 0 \)), such that

\[
    \|F(x) - F(x^*)\|^2 \leq \beta_{F}(\|x - x^*\|) \cdot \langle F(x) - F(x^*), x - x^* \rangle, \quad \forall x \in D(F).
\]

**Definition 3.** For a given point \( x^* \in D(F) \), \( F \) is called diagonally \( C_F \)-decreasing anti-monotone (diagonally \( C_F \)-DAM) at \( x^* \), if it is
diagonally nonlinear, and there exists a positive matrix $C_f = \text{diag}(C_{f_1}, C_{f_2}, \ldots, C_{f_n})$ such that each $f_i$ is $C_f$-DAM at $x^e$.

**Remark 1.** Many typical activation mappings naturally process the decreasing anti-monotone property, e.g., Sigmoids mappings [14, Lemma 1], the nearest point projection mappings, etc. The following example shows the DAM property possessed by 1-D nearest point projection. Consider:

$$
f(x) = \begin{cases} 
2(1 - \frac{x}{2}), & 0 \leq t \leq 1, \\
\frac{1}{b}, & t > 1.
\end{cases}
$$

Define:

$$
\beta_{\delta}(t) = \frac{2(1 - \frac{1}{2b})}{b}, \quad 0 \leq t \leq 1.
$$

Let $x^e = 0$ and $C_f = 2$. Obviously, $\beta_{\delta} : [0, +\infty) \to (0, C_f]$ is monotonically decreasing and satisfies $\beta_{\delta}(0) = C_f$. Further, for any $t > 0$, $\beta_{\delta}(t) \in (0, C_f)$ holds always. This is, $f$ is $C_f$-DAM.

In fact, the RNNs whose activation mappings owning the decreasing anti-monotone property are very usual, and the typical models include: Hopfield-type neural networks, recurrent back-propagation neural networks, bidirectional associative memory neural networks, bound-constraints optimization neural networks, brain-state-in-a-box /domain type neural networks, cellular neural networks and so on. All these RNN models have been widely applied in various fields of science and engineering.

### 3. Main results

The following lemma gives the globally asymptotically stable result of network (1). To begin with, it should be noted that network (1) has at least one equilibrium state. In fact, when $R(F)$ is bounded, the operator $G$ defined by $G(y) = D^{-1}(y - W(F(y) + b) + q)$ is compact. Hence, by the well-known Brouwer fixed point theorem, $G$ has at least one fixed point; that is, network (1) has at least one equilibrium state $y^e$. If we can prove that any trajectory of (1) will converge to $y^e$ as $t \to +\infty$, then, by the arbitrariness of $y^e$ and the boundedness of $R(F)$, it follows that $y^e$ is the unique equilibrium state of (1) and it is globally asymptotically stable.

**Lemma 1.** Suppose that $A = \text{diag}(a_1, a_2, \ldots, a_N)$, $f_i$ is continuous and strictly monotonically increasing, $R(F)$ is bounded. Let $L = \text{diag}(L_1, L_2, \ldots, L_N)$, where each $L_i$ is the minimum Lipschitz constant of $f_i$, and $y^e \in \Omega$. If $F$ is diagonal L-DAM at $x^e = Ay^e + b$, then $y^e$ is the unique equilibrium state of network (1), and it is globally asymptotically stable whenever there is a positive definite diagonal matrix $A = \text{diag}(a_1, a_2, \ldots, a_N)$ such that

$$
S(AL) = ADL^{-1} - \frac{AAW + (A + W)^T}{2} < 0
$$

is nonnegative definite.

**Proof.** The main object of the proof is to show that any trajectory of (1) converges to $y^e$ as $t \to +\infty$.

First, by the assumption that $F$ is diagonal L-DAM at $x^e$, we have $F$ is diagonally nonlinear and there exist monotonically decreasing functions $\beta_{\delta} : [0, +\infty) \to (0, L_i]$ ($i = 1, 2, \ldots, N$) satisfying $\beta_{\delta}(0) = L_i$, and

$$
|f_i(s) - f_i(x_i^e)|^2 \leq \beta_{\delta}(s - x_i^e) - \beta_{\delta}(s - x_i^e), \quad \forall s \in D(f_i).
$$

Define

$$
E(y(t)) = \sum_{i=1}^{N} \int_{x_i^e}^{y_i(t)+b_i} \left( f_i(s) - f_i(y_i^e + b_i) \right) ds,
$$

then

$$
\frac{dE(y(t))}{dt} = \sum_{i=1}^{N} \int_{x_i^e}^{y_i(t)+b_i} \left( f_i(s) - f_i(y_i^e + b_i) \right) ds,
$$

and

$$
\frac{dE(y(t))}{dt} = \sum_{i=1}^{N} \int_{x_i^e}^{y_i(t)+b_i} \left( f_i(s) - f_i(y_i^e + b_i) \right) ds.
$$

By the definition of $B_\delta$, it follows that $E(y(t))$ is monotonically decreasing and

$$
\lim_{t \to +\infty} E(y(t)) = 0
$$

exists since $E(y(t))$ is nonnegative. If we can prove that there is a subsequence $\{t_n\}$ such that

$$
\lim_{t \to +\infty} E(y(t)) = 0
$$

exists, implies that $\lim_{t \to +\infty} x(t) = \lim_{t \to +\infty} y(t) = 0$. Therefore, we conclude from (9) that $\lim_{t \to +\infty} y(t) = 0$.
since ε is arbitrary. That is, when t → +∞, any trajectory of (1) converges to y∗.

Now to complete the proof. We need to verify that there exists a subsequence (tn) such that \( \lim_{n \to +\infty} |f(x(t_n)) - f(x^*)| = 0 \). Obviously, if for all l ∈ \{1, 2, \ldots, N\},

\[
\lim_{t \to +\infty} \inf |f(x(t_n)) - f(x^*)| = 0,
\]

(10)

then the above proposition holds clearly. In what follows, we will show the correctness of (10).

If (10) were not true for some \( t_0 \in \{1, 2, \ldots, N\} \), then there would be two positive constants \( t_0 \) and \( \varepsilon \), such that

\[
|f(x(t_n)) - f(x^*)| \geq \varepsilon,
\]

whenever \( t \geq t_0 \). This, combined with the definition of minimum Lipschitz constant of \( f_k \) and the monotonically decreasing property of \( \beta_{x,x} \), leads to the fact that for any \( t \geq t_0 \), \( |x(t_0) - x^*| \geq \varepsilon \cdot L^{-1} \), and further, \( \beta_{x,x} ((x(t_0) - x^*)) \leq \beta_{x,x} (\varepsilon \cdot L^{-1}) \). Thus, by (7), we have

\[
\frac{dE(x(t))}{dt} \leq -\dot{\lambda}_0 \dot{\lambda}_0 \left( |f(x(t_0)) - f(x^*)|^2 \right) \leq -\dot{\lambda}_0 \dot{\lambda}_0 \left( |f(x(t_0)) - f(x^*)|^2 \right).
\]

Since \( \lim_{t \to +\infty} E(x(t)) \) exists and \( \beta_{x,x} (\varepsilon \cdot L^{-1}) > L^{-1} \), when integrate both side of the above inequality from \( t_0 \) to \( +\infty \), it can be deduced that

\[
\lim_{t \to +\infty} E(x(t)) - E(x(t_0)) \leq -\dot{\lambda}_0 \dot{\lambda}_0 \left( |f(x(t_0)) - f(x^*)|^2 \right) \leq -\dot{\lambda}_0 \dot{\lambda}_0 \left( |f(x(t_0)) - f(x^*)|^2 \right).
\]

(12)

The left side of (12) is a constant, while the right side of it approaches to \( -\infty \), which is a contradiction. The result of (10) is thus proved. This completes the proof of the lemma. □

Based on Lemma 1, we further give the globally exponential stability conclusion of network (1). In what follows, when the nonlinear activation mapping \( F(x) \) is locally continuously differentiable at \( x^* \), then \( F(x)^{1/2} \) is the derivative of \( F(x) \) at \( x^* \), is denoted by \( F'(x^*) \).

**Theorem 1.** Assume that \( A = \text{diag}[a_1, a_2, \ldots, a_N] \), \( R(F) \) is bounded and \( y^* \in \Omega \). Let \( L_k = \text{diag}[L_{k,1}, L_{k,2}, \ldots, L_{k,N}] \) with each \( L_k \) being the minimum Lipschitz constant of \( f_k \). If \( F \) is diagonal \( L \)-DAM at \( x^* \Rightarrow (Ay^* + b) \) and locally continuously differentiable at \( x^* \), each \( f_k \) is continuous and strictly monotonically increasing, then \( y^* \) is the unique equilibrium state of network (1) and it is globally exponentially stable if there exist two positive definite diagonal matrices \( \Gamma \) and \( \Lambda \), such that \( S(\Gamma, \Lambda) = \Gamma DL^{-1} - \Gamma AWF + (\Gamma AWF)^{1/2} \) is nonnegative definite and \( K(A,F(x^*)) = AD(F(x^*)) - AW AWF + (AWF)^{1/2} \) is positive definite.

**Proof.** Note first that \( F(x^*) \) is a positive definite diagonal matrix since \( F \) is diagonal nonlinear and each \( f_k \) is strictly monotonically increasing.

By Lemma 1, we already known that network (1) has a unique equilibrium state \( y^* \), and it is globally asymptotically stable on \( \mathbb{R}^N \) when \( S(\Gamma, \Lambda) \geq 0 \), \( F \) is diagonal \( L \)-DAM at \( x^* \) with \( R(F) \) being bounded, and each \( f_k \) is continuous and strictly monotonically increasing. Thus, we only need to show \( y^* \) is ultimately globally exponentially stable under the assumption that \( K(A,F(x^*)) \) is positive definite and \( F \) is locally continuously differentiable at \( x^* \). This is equivalent to justify the unique equilibrium state of the following network:

\[
A\frac{dy(t)}{dt} = -ADy(t) + AWF(Ay(t) + b) + Aq, \quad y_0 \in \mathbb{R}^N.
\]

(13)

i.e., \( y^* \), is globally exponentially stable, where \( A \) is the chosen diagonal matrix with which matrix \( K(A,F(x^*)) \) is positive definite. Define \( T: \mathbb{R}^N \to \mathbb{R}^N \) by

\[
T(y) = -ADy + AWF(Ay + b) + Aq.
\]

Obviously, \( T \) is locally continuously differentiable at \( y^* \). Let \( T'(y^*) \) be the derivative of \( T(y) \) at \( y^* \). According to the theory developed recently in [21], we only need to justify \( \lambda(T'(y^*)) < 0 \), and then, the globally exponential stability of network (1) can be verified. Since \( A \) and \( F'(x^*) \) all are diagonal matrices, a direct calculation gives that

\[
\lambda(T'(y^*)) = \lambda((-AD + AWF(Ay + b)) + Aq).
\]

Next, we will show that \( \lambda((-AD + AWF(Ay + b)) + Aq) < 0 \) when \( K(A,F(x^*)) > 0 \). Let \( \mu(A,F(x^*)) = -K(A,F(x^*)) \). By the positive definition assumption of \( K(A,F(x^*)) \), it follows that \( \mu(A,F(x^*)) \), the eigenvalues of \( K(A,F(x^*)) \), all are negative, namely,

\[
\lambda(T'(y^*)) = \lambda((-AD + AWF(Ay + b)) + Aq).
\]

(14)

For any real matrix \( M_{N \times N} \), \( \mu_2(M) \), the matrix measure induced by the 2-norm \( \|x\| = \sqrt{\sum_{i=1}^{N} x_i^2} \), is defined as

\[
\mu_2(M) = \lim_{\rho \to 0^+} \frac{\|I + \rho M\|_{2}}{\rho},
\]

and it has the property \( \mu_2(M) = \max_{l} \lambda_l \left(M + M^T\right) \). When all the eigenvalues of \( M + M^T \) are negative, we know \( \mu_2(M) < 0 \). Thus, by (14), we have \( \mu_2((-AD + AWF(Ay + b)) + Aq) < 0 \), i.e., \( \mu_2((-AD + AWF(Ay + b)) + Aq) < 0 \).

Let \( H = (h_{ij})_{N \times N} := (-AD + AWF(Ay + b) + Aq) \) and \( \gamma = \text{diag}[\gamma_1, \gamma_2, \ldots, \gamma_N] \) be \( 2 \), then \( \mu_2(H) < 0 \). Define \( \gamma = \text{diag}[\gamma_1^{1/2}, \gamma_2^{1/2}, \ldots, \gamma_N^{1/2}] \). Denote by \( \mu_2 \), the matrix measure induced by the vector norm \( \|x\|_2 = \sqrt{\sum_{i=1}^{N} x_i^2} \), one can get that

\[
\mu_2(H) = \frac{\|I + \rho \gamma\|_2}{\rho},
\]

\[
\mu_2(H) = \frac{\gamma_1^{1/2} \gamma_2^{1/2} \cdots \gamma_N^{1/2}}{\rho},
\]

(15)

In accordance with the fact that for any \( N \times N \) matrix \( M \) and any invertible matrix \( P_{N \times N} \), the eigenvalues of \( P^{-1}M \) are equal to that of \( A \), it can be deduced from (15) that

\[
\mu_2(H) = \frac{1}{2} \mu_{\text{max}}(Y^{-1/2}H^T + HY^{-1/2}) = \frac{1}{2} \mu_{\text{max}}(Y^{-1/2}H + H^T Y^{-1/2}).
\]

(16)
We now prove that when $M_{N,N}$ is a Hermite matrix and $B_{N,N}$ is a positive definite diagonal matrix, then

$$\lambda_{\text{max}}(B^{-1}M) = \max_{x \neq 0} \frac{x^TMx}{x^TBx}.$$

Since for any invertible matrix $Q$, $P(Q)p^{-1} = PQ$, thus $\lambda(Q) = \lambda(PQ)$. By the definition of Rayleigh quotient, it follows that

$$\lambda_{\text{max}}(B^{-1}M) = \lambda_{\text{max}}(B^{-1/2}B^{-1/2}M) = \lambda_{\text{max}}(B^{-1/2}(B^{-1/2}M))$$

$$= \max_{y \neq 0} \frac{y^T(B^{-1/2}MB^{-1/2})y}{y^Ty}$$

(17)

By (16) and (17) and the fact that $\max_{x \neq 0} x^T(H^2 + (H)^2)x < 0$ (which is deduced by $\mu_2(H) < 0$), we have

$$\mu_1(H) = \frac{1}{2} \max_{x \neq 0} \frac{x^T(H^2 + (H)^2)x}{x^TMx}$$

$$\leq \frac{1}{2} \max_{x \neq 0} x^T(H^2 + (H)^2)x \cdot \min_{x \neq 0} x^TMx = \lambda_1 < 0.$$

(18)

On noting that for any real matrix $B$,

$$\alpha(B) = \inf_{\|B\| \leq 1} \mu(B),$$

where $\Psi$ denotes the set of all equivalent norm of $\| \cdot \|$ [19,20], then

$$\alpha(H) = \alpha(-AD + AWA^T(x^*)) < 0.$$

This completes the proof of Theorem 1. □

The following corollary replaces verifying the positive definiteness of matrix $K(A,F(x^*))$ with certifying the relationship of two matrices defined by the network in the sense of componentwise comparison, and it is more available for applications.

**Corollary 1.** Consider network (1) with $A=I$. Assume that $R(F)$ is bounded and $y^* \in \Omega$. Let $L = \text{diag}(L_1, L_2, \ldots, L_n)$ with each $L_i$ being the minimum Lipschitz constant of $f_i$. If $F$ is diagonal L-DAM at $x^* (=Ay^*+b)$ and locally continuously differentiable at $x^*$, then $y^*$ is the unique equilibrium state of network (1), and it is globally exponentially stable if $L > F(x^*)$ in the sense of componentwise comparison and there exists a positive definite diagonal matrix $G$, such that $S(G,L) = GD^{-1} - (GA^T)(GA^T)^T/2$ is nonnegative definite.

**Proof.** It is obvious that $S(G,L) = K(G,L)$ when $A=I$, and if $L > F(x^*)$ in the sense of componentwise comparison, then the nonnegative definiteness of $S(G,L)$ implies the positive definiteness of $K(G,F(x^*))$. Thus, Corollary 1 follows from Theorem 1 directly. □

**Remark 2.** Lemma 1, Theorem 1 and Corollary 1 give some definite answers to the problem that what kinds of dynamics behavior will be for network (1) under the critical conditions. In addition, as the special cases of network (1), the two basic RNN models: local field neural network model and static neural network model, are extensively applied in learning, pattern recognition, associative memory, solving optimization problems, etc. The applicability and efficiency of such applications crucially hinge upon their dynamics, and therefore the analysis of dynamical behaviors of such two networks becomes imperative (actually it is a first step) for any practical design and application of the networks. Due to their intrinsic difficulty, these two RNN models have been fallen short of a generic, in-depth theoretical analysis under the critical conditions. While, by Lemma 1, Theorem 1 and Corollary 1, we can easily provide the corresponding dynamic results for them and the results are original. Either Lemma 1, Theorem 1 and Corollary 1, or the results achieved for local field neural network model and static neural network model can not only sharpen and extend most of the existing non-critical conclusions, but also generalize, the latest critical results for RNNs, see, e.g., [1–8,10–14,17,18,22–24].

---

**Fig. 1.** Transient behaviors of RNN in network (19) with random initial points.
4. Illustrative example

In what follows, we provide an illustrative example to demonstrate the validity of the critical dynamics results formulated in Section 3. Consider the following Hopfield-type RNN:

\[
\begin{align*}
\frac{dy_1}{dt} &= -y_1(t) + f_1(y_1(t)) + 3f_2(y_2(t)) - (4 - \ln 3), \\
\frac{dy_2}{dt} &= -y_2(t) - f_1(y_1(t)) + 0.2f_2(y_2(t)) + 0.8 - \ln 3,
\end{align*}
\]

where \( f_2(y) = (4/(1 + e^{-y})) \) \((i = 1, 2)\).

In this example, \( D = A = I \), \( W = \begin{pmatrix} 1 & 3 \\ 3 & 0 \end{pmatrix} \), \( b = \overrightarrow{0} \), \( q = \begin{pmatrix} \ln 3 \\ -\ln 3 \end{pmatrix} \), each \( L_i \) is 1 and the equilibrium state set is \( \Omega_e = \{ -\ln 3, -\ln 3 \} \).

For any positive definite diagonal matrix \( I \), it is easy to verify that \( \Gamma L^{-1}((\Gamma W + W^T I)/2) \) is not positive definite, so almost all of the exponential stability conditions of such kind of RNN are not suitable here, see, e.g., [9, 18, 23]. Meanwhile, by the results established in [14], one can only get the globally attractive conditions of network (19), and, it is easy to verify that the critically exponentially stable conditions used in [3] do not satisfy for this example. We will show Theorem 1 established in Section 3 can be applied here to achieve the global exponentially stable for network (19).

By choosing \( \Gamma = \text{diag}(1, 3) \), we have

\[ \Gamma L^{-1} - \Gamma W^T W \Gamma 2 \geq 0. \]

On the other hand, for \( y_e = (\ln 3, -\ln 3)^T \in \Omega_e \), we know \( x_0 = A y_e + b = (\ln 3, -\ln 3)^T \) and further, \( F(x_0) = \begin{pmatrix} \ln 3 \\ -\ln 3 \end{pmatrix} \). It follows that

\[ \Gamma F(x_0)^T - \Gamma W^T W \Gamma 2 > 0. \]

Meanwhile, by the definition of \( f_2 \), it is clear that each \( f_2 \) is continuous and strictly monotonically increasing, \( f_2 \) is locally continuously differentiable at \( y_e \). Since \( f_2 \) is sigmoidal, then by Remark 1, it is \( L \)-DAM. That is, the conditions in Theorem 1 all hold. According to Corollary 1, network (19) is globally exponentially stable. The following Fig. 1 depicts the time responses of state variables of the network with random initial points, which can confirm that the proposed conditions in Theorem 1 ensure the globally exponential stability of RNNs.

5. Conclusion

In this work, we have developed the critical stability theory of RNNs with general forms. Based on exploring some intrinsic properties of the networks, and by using the energy function method and matrix measure theory, it is shown that a RNN with decreasing anti-monotone activation mapping has a unique equilibrium state and which is globally asymptotically stable under the critical conditions. Further, the RNN is globally exponentially stable under the conditions that a discriminant matrix determined by the network is positive definite at the unique equilibrium state. The obtained critical dynamics results extend directly the existing non-critical conclusions without adding any further requirements, and at the same time, they generalize almost all of the critical conclusions. The achieved conclusions can not only provide a wider application range for RNNs, but also can be applied directly to many concrete RNN models.

References


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.

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