A Critical Analysis on Global Convergence of Hopfield-Type Neural Networks

Jigen Peng, Zong-Ben Xu, Hong Qiao, *Member, IEEE*, and Bo Zhang

***Abstract*—This paper deals with the global convergence and stability of the Hopfield-type neural networks under the critical condition that $M_1(\Gamma) = L^{-1}D\Gamma - (\Gamma W + W^T\Gamma)/(2)$ is nonnegative for any diagonal matrix Γ, where W is the weight matrix of the network, $L = \mathrm{diag}\{L_1, L_2, \ldots, L_N\}$ with L_i being the Lipschitz constant of g_i and $G(u) = (g_1(u_1), g_2(u_2), \ldots, g_N(u_N))^T$ is the activation mapping of the network. Many stability results have been obtained for the Hopfield-type neural networks in the noncritical case that $M_1(\Gamma)$ is positive definite for some positive definite diagonal matrix Γ. However, very few results are available on the global convergence and stability of the networks in the critical case. In this paper, by exploring two intrinsic features of the activation mapping, two generic global convergence results are established in the critical case for the Hopfield-type neural networks, which extend most of the previously known globally asymptotic stability criteria to the critical case. The results obtained discriminate the critical dynamics of the networks, and can be applied directly to a group of Hopfield-type neural network models. An example has also been presented to demonstrate both theoretical importance and practical significance of the critical results obtained.**

***Index Terms*—Attractive, global convergence, Hopfield-type neural networks, stability analysis.**

I. Introduction

HOPFIELD-TYPE neural networks have been extensively studied in the past years due to their applicability in solving associative memory, pattern recognition, and optimization problems as well as their easy VLSI implementation (see, e.g., [4], [14], [15], [17], [18], [22], [38] and the references quoted there). The applicability and efficiency of such networks hinge upon their dynamics, and therefore the analysis of dynamic behaviors of the networks is a first and necessary step for any practical design and application of the networks.

For Hopfield-type neural networks two kinds of dynamic behaviors may be expected: asymptotically stable dynamics and globally convergent dynamics. The asymptotically stable dynamics is vital for such a network to have the capability of error correction when used as an associative memory, whereas the globally convergent dynamics is necessary for preventing the network from oscillatory or chaotic behavior. In particular, the globally convergent dynamics implies that every trajectory of the network can converge to some equilibrium state, so that, when used as an associative memory, every state in the underlying space can serve as a key to recovering certain stored memory and therefore the state space is totally covered by distinct basins of the stored memories. Moreover, when applied as an optimization solver, the globally convergent dynamics implies the guaranteed convergence of the neural network algorithm to an optimal solution from every initial guess.

It is thus highly desirable to establish criteria for both global convergence and asymptotic stability of the Hopfield-type neural networks. In the past decades, considerable effort has been devoted to the asymptotic stability analysis and, in particular, the global exponential stability analysis of the networks (see, e.g., [2], [6], [8], [9], [11], [12], [16], [19], [20], [25], [32], [36], [39], [40] and the references therein). As a result, a set of very generic, in-depth, criteria for global exponential stability has been obtained for the Hopfield-type networks (see Theorem 1 below). In contrast, however, very few results are available on a generic, in-depth, global convergence analysis except those deduced from the global exponential stability analysis. In [6], a global convergence analysis was conducted for a specific class of neural networks. It should be remarked that it is by no means easy to conduct a meaningful global convergence analysis. This is because such analysis is essentially equivalent to a stability analysis of the system in the critical case in the sense that $M_1(\Gamma) = L^{-1}D\Gamma - (\Gamma W + W^T\Gamma)/(2)$ is nonnegative for any diagonal matrix Γ, where W is the weight matrix of the network, $L = \mathrm{diag}\{L_1, L_2, \ldots, L_N\}$ with L_i being the Lipschitz constant of g_i and $G(u) = (g_1(u_1), g_2(u_2), \ldots, g_N(u_N))^T$ is the activation mapping of the network (see Section II for details). (Note that many stability results have been obtained for the noncritical case that $M_1(\Gamma)$ is positive definite for some positive definite diagonal matrix Γ.)

The purpose of this paper is to present such a critical analysis on global convergence of the Hopfield-type neural networks. Our approach is based on two intrinsic features of the nonlinear activation mapping in conjunction with the energy function approach combined with *a priori* decay estimates on solutions of the system. With such an approach, we establish two generic critical global convergence theorems. These results generalize some of the existing results on global exponential stability of the Hopfield-type neural networks to the critical case in the sense that under the corresponding critical conditions, the global convergence of the Hopfield-type neural networks can be justified.

Manuscript received November 19, 2002; revised August 31, 2004. The work of J. Peng and Z.-B. Xu was supported by the National Natural Science Foundation of China under Grant 10101019 and the Hi-Tech R & D (863) Program under Grant 2001AA113182. This paper was recommended by Associate Editor W.-C. Fang.

J. Peng and Z.-B. Xu are with Institute for Information and System Sciences, Faculty of Science, Xi'an Jiaotong University, Xi'an, China (e-mail: jgpeng@xjtu.edu.cn; zbxu@xjtu.edu.cn).

H. Qiao is with Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China, and also with the School of Informatics, University of Manchester, Manchester M60 1QD, U.K. (e-mail: hong.qiao@mail.ia.ac.cn).

B. Zhang is with Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing 100080, China, and also with the School of MIS, Coventry University, Coventry CV1 5FB, U.K. (e-mail: b.zhang@amt.ac.cn).

Digital Object Identifier 10.1109/TCSI.2005.844366

Our results not only apply directly to many concrete examples of the Hopfield-type neural networks, but also extend, to a large extent, the results of Chen and Amari [6].

The remaining part of the paper is organized as follows. In Section II, we summarize as well as give a unified proof and account for the existing global exponential stability results of the Hopfield-type neural networks. Section III explores two intrinsic features of the activation mappings. The main results are presented in Section IV, which also contains applications of the general results in several special cases and a comparison with the known global exponential stability results. To illustrate the general results obtained an example is provided in Section V, where the existing global exponential stability results are not applicable. Some concluding remarks are presented in Section VI.

We conclude this section by introducing some notations. Denote by $\mathbb{R}^N$ the N-dimensional real vector space with the vector norm $\|\cdot\|$. For a given $N \times N$ matrix A, denote by $\|A\|$ and $\mu(A)$ its matrix norm induced by the given vector norm $\|\cdot\|$ and its corresponding matrix measure, respectively, defined by

$$\|A\| = \sup_{\|x\|=1} \|Ax\| \quad \mu(A) = \lim_{\lambda \to 0^+} \frac{\|I + \lambda A\| - 1}{\lambda}$$

where I denotes the identity matrix. For any matrix A, A^T stands for the transpose of A. Given an operator $F : V \to \mathbb{R}^N$, we denote by $\Re(F)$ the range of $F, F^{-1}(0)$ the null set, i.e., $F^{-1}(0) = \{x \in V \mid F(x) = 0\}$, and F^{-1} the inverse operator from $\Re(F)$ to V (if it exists).

II. Global Exponential Stability: A Unified Account

In this section, we summarize some of the existing results on global exponential stability/instability of the Hopfield neural networks. We shall also demonstrate how those diversely developed results in (e.g., [2], [10], [12], [16], [19], [25], [40]) can be uniformly and concisely stated and justified. These results will serve as a prototype of comparison and reference with our main theorems in Section IV. We first introduce some definitions and notations needed throughout this paper.

Throughout this paper, we consider Hopfield-type neural networks of the form

$$C_i \frac{du_i}{dt} = -\frac{u_i}{R_i} + \sum_{j=1}^{N} w_{ij} g_j(u_j) + I_i, \qquad i = 1, 2, \ldots, N \tag{2.1}$$

or the general form

$$\frac{du}{dt} = -Du + WG(u) + I \tag{2.2}$$

where $u = (u_1, u_2, \ldots, u_N)^T$ with u_i being the local field, $g_i(u_i)$ is the state of neuron i, C_i, R_i, and I_i are fixed physical parameters, $D = \operatorname{diag}\{d_1, d_2, \ldots, d_N\}$ is a positive diagonal matrix associated with C_i and $R_i, G = (g_1, g_2, \ldots, g_N)^T$ is the activation mapping of the network with g_i being the activation function acted on neuron $i, I = (I_1, I_2, \ldots, I_N)^T$ is the external input vector imposed on the network, $W = (w_{ij})_{N \times N}$ is the weight matrix with w_{ij} being the synaptic connection value between neuron i and neuron j, and N is the number of neurons in the network.

We assume throughout that there exists a unique solution $u(t, u_0)$ to (2.2) for any given initial data u_0 in $\mathbb{R}^N$. (This is the case, e.g., when G is locally Lipschitz continuous.) As usual, the solution $u(t, u_0)$ is also called a *trajectory* of (2.2) through u_0, denoted henceforth by $\Gamma(u_0)$. Recall that a constant vector u^* is said to be an equilibrium state of the system (2.2) if u^* is a zero point of the operator F, defined by

$$Fx = -Dx + WG(x) + I \qquad \forall x \in \mathbb{R}^N \tag{2.3}$$

that is, $F(u^*) = 0$. All the equilibrium states of (2.2) is denoted by $F^{-1}(0)$. The equilibrium state u^* is said to be *stable* if any trajectory of (2.2) can stay within a small neighborhood of u^* whenever the initial data u_0 is close to u^*, and is said to be *attractive* if there is a neighborhood $\Xi(u^*)$, called the *attraction basin* of u^*, such that any trajectory of (2.2) initialized from a state in $\Xi(u^*)$ will approach u^* as time goes to infinity. An equilibrium state is said to be *asymptotically stable* if it is both stable and attractive, whilst the equilibrium state u^* is said to be *exponentially stable* if there exist two positive constants M and α such that

$$\|u(t, u_0) - u^*\| \leq Me^{-\alpha t} \cdot \|u_0 - u^*\| \qquad \forall t \geq 0. \tag{2.4}$$

Further, u^* is said to be *globally asymptotic stable* if it is asymptotically stable and $\Xi(u^*) = \mathbb{R}^N$. A system (say, (2.2)) is said to be *globally convergent* if $\lim_{t \to +\infty} u(t, u_0) = u^*$ for every initial point $u_0 \in \mathbb{R}^N$ (the limit of $u(t, u_0)$ may not be the same for different u_0), whilst it is said to be *exponentially convergent* if it is globally convergent with $u(t, u_0)$ and its limit u^* satisfying (2.4).

We now recall some basic facts from dynamical system theory [23]. Let V be a subset of $\mathbb{R}^N$ and let $E : V \to \mathbb{R}$ be a continuously differentiable function. Then, the following are true.

1) E is said to be an energy function of system (2.2) if it decreases along the trajectory of (2.2), that is, $(dE(u(t, u_0))/dt) \leq 0$ for all $u_0 \in V$.
2) The energy function E is said to be *strict* if its derivative along the trajectories vanishes only at the equilibria of (2.2), i.e.,

$$\frac{dE(u(t, u))}{dt} = 0 \quad \text{iff } u \in F^{-1}(0).$$

With the above notion, we can summarize some existing results on global exponential stability of the Hopfield-type neural networks.

Theorem 1: Suppose

$$G(u) = (g_1(u_1), g_2(u_2), \ldots, g_N(u_N))^T$$

with each g_i being monotonically increasing and Lipschitz continuous (that is, $|g_i(s) - g_i(t)| \leq L_i|s - t|$ for all $s, t \in \mathbb{R}$ with L_i being a positive constant). For any diagonal matrix $\Gamma = \operatorname{diag}\{\xi_1, \xi_2, \ldots, \xi_N\}$, write

$$M_1(\Gamma) = L^{-1}D\Gamma - \frac{\Gamma W + W^T\Gamma}{2} \tag{2.5}$$

where $L = \operatorname{diag}\{L_1, L_2, \ldots, L_N\}$. Then (2.2) has a unique equilibrium state u^*, and moreover, u^* is globally exponentially stable if there is a positive definite diagonal matrix Γ such that, for any $i \in \{1,2,\ldots,N\}$, one of the following conditions $(\mathbf{D})$ and $(\mathbf{D1})$–$(\mathbf{D8})$ is satisfied.

$(\mathbf{D})$ $M_1(\Gamma)$ is positive definite.
$(\mathbf{D1})$ $L_i^{-1}d_i\xi_i - \xi_i w_{ii} > \sum_{j\neq i}^N \xi_j|w_{ij}|$.
$(\mathbf{D2})$ $L_i^{-1}d_i\xi_i - \xi_i w_{ii} > \sum_{j\neq i}^N \xi_j|w_{ji}|$.
$(\mathbf{D3})$ $L_i^{-1}d_i\xi_i - \xi_i w_{ii} > (1/2)\sum_{j\neq i}^N |\xi_j w_{ji} + \xi_i w_{ij}|$.
$(\mathbf{D4})$ the matrix measure $\mu(\Gamma W - L^{-1}D\Gamma) < 0$.
$(\mathbf{D5})$ $\Gamma W + W^T\Gamma$ is nonnegative.
$(\mathbf{D6})$ $L_i^{-1}d_i\xi_i > \{\xi_i w_{ii} + \sum_{j\neq i}^N \xi_j|w_{ij}|\}^+$.
$(\mathbf{D7})$ $d_i\xi_i > L_i\xi_i(w_{ii})^+ + \sum_{j\neq i}^N L_j\xi_j|w_{ij}|$.
$(\mathbf{D8})$ $d_i\xi_i > L_i\xi_i(w_{ii})^+ + (1/2)\sum_{j\neq i}^N\{L_j\xi_j|w_{ij}| + \xi_i L_i|w_{ji}|\}$ where $(a)^+ = \max\{0,a\}$.

Proof: Under the assumption of positive definiteness of the matrix $M_1(\Gamma)$ (namely, condition $(\mathbf{D})$), it is easy to see that there is a unique equilibrium state u^* to system (2.2). Let

$$E(u) = \sum_{i=1}^N \xi_i \int_{u_i^*}^{u_i} [g_i(s) - g_i(u_i^*)]\, ds. \tag{2.6}$$

It is easy to verify, similarly as in [40], that E is a strict energy function of system (2.2). Thus, the global convergence of the trajectory $\Gamma(u_0)$ follows immediately from the LaSalle invariance principle [23]. Further, an exponential estimate as (2.4) on the decay of solution $u(t,u_0)$ can be established similarly as in [25], [40].

It is clear that each of conditions (D4) and (D5) implies (D). Note that any of conditions (D1)–(D3) and (D6)–(D8) can imply that $L^{-1}D - W$ is a nonsingular M-matrix [3] so, by ([26], Lemma 1), there is a positive definite diagonal matrix Γ such that $M_1(\Gamma)$ is positive definite. Thus, any of conditions (D1)–(D3) and (D6)–(D8) implies (D). Theorem 1 is thus proved. □

Remark 1: Theorem 1 was proved previously by many authors under certain specific conditions as listed above. In particular, the same or similar results have been established in [11], [25], [40] under condition (D), in [13], [10], [28], [32] under (D1), in [9], [13] under (D2), in [39] under (D3), in [20] under (D5), and in [5], [10] under (D6)–(D8).

The following theorem summarizes some results on instability of the Hopfield-type neural networks.

Theorem 2: Assume that

$$G(u) = (g_1(u_1), g_2(u_2), \ldots, g_N(u_N))^T$$

with each g_i being monotonically increasing and Lipschitz continuous with Lipschitz constant L_i. Let us assume further that g_i is inversely Lipschitz continuous, i.e., $|g_i(s) - g_i(t)| \geq l_i|s-t|$ for all $s,t \in \mathbb{R}$, where $l_i > 0$ is a constant, and let, for any diagonal matrix $\Gamma = \mathrm{diag}\{\xi_1,\xi_2,\ldots,\xi_N\}$

$$M_2(\Gamma) = l^{-1}D\Gamma - \frac{\Gamma W + W^T\Gamma}{2} \tag{2.7}$$

where $l = \mathrm{diag}\{l_1,l_2,\ldots,l_N\}$. Suppose there is a positive definite diagonal matrix Γ such that, for any $i \in \{1,2,\ldots,N\}$, one of the following conditions (I) and (I1)–(I4) is satisfied.

(I) $M_2(\Gamma)$ is negative definite.
(I1) $-l_i^{-1}d_i\xi_i + \xi_i w_{ii} > \sum_{j\neq i}^N \xi_j|w_{ij}|$.
(I2) $-l_i^{-1}d_i\xi_i + \xi_i w_{ii} > \sum_{j\neq i}^N \xi_j|w_{ji}|$.
(I3) $-l_i^{-1}d_i\xi_i + \xi_i w_{ii} > (1/2)\sum_{j\neq i}^N |\xi_j w_{ji} + \xi_i w_{ij}|$.
(I4) The matrix measure $\mu(l^{-1}D\Gamma - \Gamma W) < 0$.

Then, (2.2) has a unique equilibrium state u^*, and further, u^* is globally exponentially unstable in the sense that, for some positive constants M and α, there holds

$$\|u(t) - u^*\| \geq Me^{\alpha t}\|u_0 - u^*\| \qquad \forall t \geq 0 \tag{2.8}$$

where $u(t)$ is any trajectory of (2.2) starting from u_0.

Proof: It is easy to verify that each of conditions (I1)–(I4) implies condition (I). So we assume without loss of generality that condition (I) is satisfied.

We first prove that (2.2) has a unique equilibrium state u^*. By the homeomorphism theorem (cf. [29]), it suffices to verify that F (defined in (2.3)) is both invertible and proper. (Note that F is proper if $F(u) \to \infty$ whenever $u \to \infty$ [29].) Let $\lambda_{\max}$ be the maximum singular value of $M_2(\Gamma)$. Then, $\lambda_{\max} < 0$ since $M_2(\Gamma)$ is negative definite. For any $u, v \in \mathbb{R}^N$, we have by the Cauchy-Schwarz inequality and the assumptions on g_i that

$$\begin{aligned}
&\max_{1\geq i\geq N}\{L_i\xi_i\}\cdot\|F(u)-F(v)\|\cdot\|u-v\| \\
&\quad\geq \sum_{i=1}^N \xi_i|F_i(u)-F_i(v)|\cdot|g_i(u_i)-g_i(v_i)| \\
&\quad\geq \langle -D\Gamma(u-v) + \Gamma W(G(u)-G(v)), G(u)-G(v)\rangle \\
&\quad\geq -(G(u)-G(v))^T M_2(\Gamma)(G(u)-G(v)) \\
&\quad\geq -\lambda_{\max}\cdot\|G(u)-G(v)\|^2 \\
&\quad\geq -\lambda_{\max}\cdot\min_{1\leq i\leq N}\{l_i^2\}\cdot\|u-v\|^2
\end{aligned}$$

which implies that F is both invertible and proper. Thus, F is a homeomorphism on $\mathbb{R}^N$ so that F has a unique zero point u^*.

To prove the exponential instability of u^*, let us define, for any trajectory $u(t) = (u_1(t), u_2(t), \ldots, u_N(t))^T$ of (2.2) starting from u_0

$$E(t) = \sum_{i=1}^N \gamma_i \int_{u_i^*}^{u_i(t)} [g_i(s) - g_i(u_i^*)]\, ds, \qquad t \geq 0.$$

From the fact that g_i is inversely Lipschitz continuous and monotonically increasing, it follows that

$$\begin{aligned}
E(t) &\geq \sum_{i=1}^N \gamma_i \int_{u_i^*}^{u_i(t)} l_i(s-u_i^*)\, ds \\
&\geq \frac{1}{2}\sum_{i=1}^N \gamma_i l_i (u_i(t)-u_i^*)^2 \\
&\geq \frac{1}{2}\min_{1\leq i\leq N}\{\gamma_i l_i\}\cdot\|u(t)-u^*\|^2
\end{aligned} \tag{2.9}$$

and

$$\begin{aligned}
E(t) &\leq \sum_{i=1}^N \gamma_i[g_i(u_i(t)) - g_i(u_i^*)]\cdot[u_i(t)-u_i^*] \\
&\leq \sum_{i=1}^N \gamma_i l_i^{-1}[g_i(u_i(t)) - g_i(u_i^*)]^2 \\
&\leq \max_{1\leq i\leq N}\{\gamma_i l_i^{-1}\}\cdot\|G(u(t)) - G(u^*)\|^2.
\end{aligned} \tag{2.10}$$

The derivative of $E(t)$ can be estimated similarly, and we have

$$\begin{aligned}
&\frac{dE(t)}{dt}\\
&=\sum_{i=1}^{N}\gamma_i(g_i(u_i(t))-g_i(u_i^*))\cdot\frac{du_i(t)}{dt}\\
&=-\sum_{i=1}^{N}\Bigg\{\gamma_i d_i(g_i(u_i(t))-g_i(u_i^*))(u_i(t)-u_i^*)\\
&\quad-\gamma_i(g_i(u_i(t))-g_i(u_i^*))\sum_{j=1}^{N}w_{ij}(g_j(u_j(t))-g_j(u_j^*))\Bigg\}\\
&\geq-\sum_{i=1}^{N}l_i^{-1}d_i\gamma_i[g_i(u_i(t))-g_i(u_i^*)]^2\\
&\quad+\sum_{i=1}^{N}\gamma_i(g_i(u_i(t))-g_i(u_i^*))\sum_{j=1}^{N}w_{ij}(g_j(u_j(t))-g_j(u_j^*))\\
&=-(G(u(t))-G(u^*))^TM_2(\Gamma)(G(u(t))-G(u^*))\\
&\geq-\lambda_{\max}\cdot\|G(u(t))-G(u^*)\|^2\\
&\geq-\frac{\lambda_{\max}\cdot E(t)}{\max_{1\leq i\leq N}\left\{\gamma_i l_i^{-1}\right\}}=:\alpha\cdot E(t).
\end{aligned}$$

Integrating this inequality yields $E(t)\geq e^{\alpha t}\cdot E(0)$ for all $t\geq 0$. This together with (2.9) and (2.10) gives the inequality (2.8). The proof is thus completed. □

Theorem 1 means that the positive definiteness of $M_1(\Gamma)$ is sufficient for the global exponential stability of (2.2), whilst Theorem 2 implies that the nonnegative definiteness of $M_2(\Gamma)$ is necessary for (2.2) to have globally stable dynamics. It is thus clear that there is a gap between the positive definiteness of $M_1(\Gamma)$ and the nonnegative definiteness of $M_2(\Gamma)$. The analysis on stability of the Hopfield-type neural network (2.2) in this gap will be referred to as the *critical stability analysis*. In the remainning part of this paper we will focus on exploiting this gap and present some critical global convergence theorems on the network (2.2). More precisely, we will answer such a question: *what happens to the asymptotic behavior of* (2.2) *when the positive definiteness of* $M_1(\Gamma)$ *in Theorem 1 is replaced by the nonnegative definiteness of* $M_1(\Gamma)$? To this end, we first explore some intrinsic features of the activation mapping G in the next section.

III. Intrinsic Characteristics of Activation Mappings

In this section, we explore two intrinsic features of the activation mapping which are essential in establishing a critical stability analysis of the Hopfield network (2.2). To begin with, we introduce the following definition.

Definition 1: Let $G:\mathbb{R}^N\to\mathbb{R}^N$ be a nonlinear mapping.

1) G is said to be diagonally nonlinear (or, is of diagonal nonlinearity) if G is defined componentwise by

$$G(x)=(g_1(x_1),g_2(x_2),\ldots,g_N(x_N))^t$$

where each g_i is a one-dimensional (1-D) nonlinear function.

2) G is said to be a nearest point projection if there is a bounded, closed, convex subset $\Omega\subset\mathbb{R}^N$ such that

$$G(x)=\arg\min_{z\in\Omega}\|x-z\|$$

Such a mapping is denoted by P_Ω, i.e., $G(x)=P_\Omega(x)$.

3) G is said to be a diagonal projection if it is diagonally nonlinear, and each component function g_i is a 1-D nearest point projection.

4) G is said to be uniformly anti-monotonic if there is a constant $\alpha_0>0$ such that

$$\langle G(x)-G(y),x-y\rangle\geq\alpha_0\|G(x)-G(y)\|^2,\ x,y\in\mathbb{R}^N \tag{3.11}$$

α_0 is called the anti-monotonic constant of G.

5) G is said to be diagonally uniform anti-monotonic if it is diagonally nonlinear, and each g_i, as a 1-D function, is uniformly anti-monotonic with the anti-monotonic constant α_i.

Examples 1–3 show that the activation mappings appeared in most of the currently-known neural networks naturally possess the uniformly anti-monotonic property.

Example 1: The nearest point projection $G=P_\Omega$ is uniformly anti-monotonic with $\alpha_0=1$.

It is known (see, e.g., [21]) that P_Ω is the nearest point projection of $\mathbb{R}^N$ onto Ω if and only if it satisfies

$$\langle x-P_\Omega(x),P_\Omega(x)-z\rangle\geq 0\qquad\forall z\in\Omega,\qquad x\in\mathbb{R}^N.$$

Taking $z=P_\Omega(y)$ with any $y\in\mathbb{R}^N$ then gives

$$\langle x-P_\Omega(x),P_\Omega(x)-P_\Omega(y)\rangle\geq 0.$$

Similarly, $\langle y-P_\Omega(y),P_\Omega(y)-P_\Omega(x)\rangle\geq 0$. Adding these two inequalities leads to the result

$$\langle P_\Omega(x)-P_\Omega(y),x-y\rangle\geq\|P_\Omega(x)-P_\Omega(y)\|^2\qquad\forall x,y\in\mathbb{R}^N \tag{3.12}$$

which means that G is uniformly anti-monotonic with $\alpha_0=1$.

Such a projection nonlinearity is commonly used in neural networks of the brain-state-in-a-box (BSB) type (see, e.g., [24], [35]) and of the optimization type (see, e.g., [11], [12], [37], [36]).

Example 2: The diagonal projection $G(x)=(g_1(x_1),g_2(x_2),\ldots,g_N(x_N))^T$ is diagonally uniformly anti-monotonic with $\alpha_0=1$.

In the BSP-type neural networks (see, e.g., [24], [35]), cellular neural networks (CNNs) (see, e.g., [7], [30], [33]) and the bound constrain optimization neural networks (BCOp-type NNs) (see, e.g., [4], [11], [27]), either the following specific 1-D nearest point projection $g_i:\mathbb{R}\to[a_i,b_i]$ is used:

$$g_i(s)=\begin{cases}a_i, & s\leq a_i\\ s, & s\in[a_i,b_i]\\ b_i, & s\geq b_i\end{cases} \tag{3.13}$$

or

$$g_i(s)=\frac{1}{2}(|s+1|-|s-1|). \tag{3.14}$$

Example 3: Let $G(x)=(g_1(x_1),g_2(x_2),\ldots,g_N(x_N))^T$, where each g_i is monotonically increasing and Lipschitz continuous. Then, G is diagonally uniformly anti-monotonic with anti-monotonic constant $\alpha_0 = \min_{1\leq i\leq N}\{L_i^{-1}\}$, where L_i is the minimum Lipschitz constant of g_i defined by

$$L_i = \sup_{t\neq s}\frac{|g_i(t)-g_i(s)|}{|t-s|} < +\infty. \tag{3.15}$$

In fact, this follows directly from the inequality

$$(g_i(t)-g_i(s))\cdot(t-s) \geq \frac{1}{L_i}|g_i(t)-g_i(s)|^2. \tag{3.16}$$

In many neural networks such as the Hopfield-type neural networks ([17], [18]), neural networks of the bidirectional associative memory (BAM) type ([22] and the recurrent back-propagation (ReBP) neural networks ([1], [31]), the nonlinear activation mapping G is usually defined via a set of so-called sigmoidal functions g_i, which may be defined, for example, by

$$g_i(s) = \tanh\left(\frac{\delta_i s}{2}\right) = \frac{1-e^{-\delta_i s}}{1+e^{-\delta_i s}} \tag{3.17}$$

where $\delta_i > 0$ is a parameter controlling the slope of the sigmoidal curve. It is easy to verify that $L_i = (\delta_i)/(2)$ in this case so that G is diagonally uniformly anti-monotonic with $\alpha_0 = 2\min_{1\leq i\leq N}\{\delta_i^{-1}\}$.

Definition 2: A function g is said to be a generalized sigmoidal function if it possesses the following three properties.

1) The range $\Re(g)$ is bounded, that is, there are two real constants a, b such that $a \leq g(x) \leq b$ for all $x \in \mathbb{R}$.
2) It is strictly monotonically increasing and continuously differentiable.
3) The derivative g' attains its maximum at a unique point.

The set of generalized sigmoidal functions is denoted by $\Im$.

By Definition 2, for any $g \in \Im$, there is a unique point, say, x_0 such that

$$|g'(x_0)| = \sup_{x\in\mathbb{R}}|g'(x)| := L(g).$$

Clearly, $L(g)$ gives the minimum Lipschitz constant of function g. Note that any generalized sigmoidal function g is uniformly anti-monotonic with the anti-monotonic constant $\alpha = L(g)$.

It is easy to verify that the function g_i, defined in (3.17), is a generalized sigmoidal function. However, the 1-D nearest point projections, defined in (3.13) and (3.14), are not such functions.

We now explore some useful properties of generalized sigmoidal functions.

Lemma 1: Let $g \in \Im$. Then, for any $x \in \mathbb{R}$, there exists a strictly monotonically decreasing function $\varphi_x : [0,+\infty) \to [0,1]$ such that $\varphi_x(0) = 1, \lim_{s\to+\infty}\varphi_x(s) = 0$ and

$$|g(x+s)-g(x)| \leq \varphi_x(|s|)\cdot L(g)\cdot|s| \qquad \forall s\in\mathbb{R}. \tag{3.18}$$

Proof: For arbitrarily fixed $x \in \mathbb{R}$, define the function $\phi_x : [0,+\infty) \to \mathbb{R}$ by

$$\phi_x(s) = \sup_{|t|>s}\frac{g(x+t)-g(x)}{t}, \qquad s\in[0,+\infty).$$

Since g is monotonically increasing and continuously differentiable, then $0 \leq \phi_x \leq L(g)$. Clearly, ϕ_x is monotonically decreasing and satisfies

$$|g(x+s)-g(x)| \leq \phi_x(|s|)|s| \qquad \forall s\in\mathbb{R}. \tag{3.19}$$

If ϕ_x has the property that $\phi_x(s) < L(g)$ for all $s > 0$, then it is easy to verify that the function φ_x, defined by

$$\varphi_x(s) = \frac{s\phi_x(s)}{(1+s)\cdot L(g)} + \frac{1}{1+s}, \qquad s\geq 0 \tag{3.20}$$

is strictly monotonically decreasing in $[0,+\infty)$ and satisfies the required inequality (3.18), which proves the lemma. Thus, to prove the lemma, it is enough to show that $\phi_x(s) < L(g)$ for all $s > 0$.

First, it follows, on noting that g is bounded, that $\lim_{t\to\infty}(g(x+t)-g(x))/t = 0$ so, for any $s > 0$, there exists a positive number $\delta(s)$ such that

$$\phi_x(s) = \sup_{\delta(s)\geq|t|>s}\frac{g(t+x)-g(x)}{t}.$$

We assume without loss of generality that

$$\phi_x(s) = \sup_{\delta(s)\geq t>s}\frac{g(t+x)-g(x)}{t}.$$

Note next that there is a unique x_0 such that $L(g) = |g'(x_0)|$ since $g \in \Im$. The required result will be shown by distinguishing between two cases.

Case 1. $x_0 \notin (x, x+s)$: In this case, by the well-known mean-value theorem we can choose an $x_1 \in (x, x+s)$ such that $s\cdot g'(x_1) = g(x+s)-g(x)$. Since $g'(x)$ attains its maximum value $L(g)$ only at x_0, then $g'(x_1) < L(g)$. It thus follows that

$$\begin{aligned}\phi_x(s) &= \sup_{\delta(s)\geq t>s}\frac{g(t+x)-g(x+s)+g(x+s)-g(x)}{t}\\ &\leq \sup_{\delta(s)\geq t>s}\frac{L(g)(t-s)+s\cdot g'(x_1)}{t}\\ &= \sup_{\delta(s)\geq t>s}\left[L(g)-\frac{s}{t}(L(g)-g'(x_1))\right]\\ &\leq L(g)-\frac{s}{\delta(s)}(L(g)-g'(x_1)) < L(g).\end{aligned}$$

Case 2. $x_0 \in (x, x+s)$: By the mean-value theorem again there exists an $x_2 \in (x, (x_0+x)/2)$ such that

$$\frac{x_0-x}{2}\cdot g'(x_2) = g\left(\frac{x+x_0}{2}\right)-g(x).$$

Similarly, as in Case 1 it can be deduced that $g'(x_2) < L(g)$ and

$$\begin{aligned}\phi_x(s) &= \sup_{\delta(s)\geq t>s}\frac{g(t+x)-g\left(\frac{x+x_0}{2}\right)+g\left(\frac{x+x_0}{2}\right)-g(x)}{t}\\ &\leq \sup_{\delta(s)\geq t>s}\frac{L(g)\left(t-\frac{x_0-x}{2}\right)+\frac{x_0-x}{2}\cdot g'(x_2)}{t}\\ &= \sup_{\delta(s)\geq t>s}\left[L(g)-\frac{x_0-x}{2t}(L(g)-g'(x_2))\right]\\ &\leq L(g)-\frac{x_0-x}{2\delta(s)}(L(g)-g'(x_2)) < L(g).\end{aligned}$$

Thus, in either case, we have concluded that $\phi(s) < L(g)$, as expected. This completes the proof of Lemma 1. □

Lemma 2: For any real-valued uniformly anti-monotonic function g defined on $\mathbb{R}$, we have

$$\int_s^t (g(\theta)-g(s))d\theta \geq \frac{\alpha}{2}(g(t)-g(s))^2 \qquad \forall t,s\in\mathbb{R} \tag{3.21}$$

where α is the anti-monotonic constant of g.

Proof: Let

$$f(t)=\int_s^t (g(\theta)-g(s))d\theta-\frac{\alpha}{2}(g(t)-g(s))^2.$$

If g is continuously differentiable, then it can be shown that

$$f'(t)=(1-\alpha g'(t))(g(t)-g(s))$$

and $|\alpha g'(t)|\le 1$. This implies that $f'(t)\ge 0$ when $t\ge s$, and $f'(t)\le 0$ when $t\le s$. Thus, $f(t)$ attains its unique minimum at $t=s$. In view of the fact that $f(s)=0$, the required estimate (3.21) follows in this case.

Since any continuous function can be arbitrarily approximated by continuously differentiable functions, then (3.21) also holds when g is continuous. The proof is thus completed. □

IV. Critical Global Convergence Results

In this section, we shall establish two generic critical global convergence theorems for the Hopfield-type neural networks (2.2) in the case when G is defined either from a set of generalized sigmoidal functions or from a set of nearest point projections. Our approach is based on the energy function method together with the two intrinsic features of the nonlinear activation mapping established in Section III and *a priori* decay estimates for solutions of the system. These generic theorems are then specified and compared with the known global asymptotic stability results previously developed in, e.g., [2], [4], [6], [9]–[11], [19], [20], [25], [40].

We first consider the case when G is defined from a set of generalized sigmoidal functions.

Theorem 3: Assume that $G=(g_1,g_2,\dots,g_N)^T$ is diagonally nonlinear with each $g_i\in\Im$. Let $L=\mathrm{diag}\{L(g_1),L(g_2),\dots,L(g_N)\}$, where $L(g_i)$ is the minimum Lipschitz constant of g_i. If there is a positive diagonal matrix $\Gamma=\mathrm{diag}\{\xi_1,\xi_2,\dots,\xi_N\}$ such that the matrix

$$M_1(\Gamma)=L^{-1}D\Gamma-\frac{\Gamma W+W^T\Gamma}{2} \qquad (4.22)$$

is nonnegative definite, then system (2.2) has a unique equilibrium state u^*, and u^* is globally attractive.

Proof: First, by the well-known Brouwer fixed point theorem, system (2.2) has at least one equilibrium state (i.e., $F^{-1}(0)\ne\emptyset$). In fact, it is easy to see that $u\in F^{-1}(0)$ if and only if u is a fixed point of the operator T defined by $T(u)=D^{-1}WG(u)$. Since each component function g_i of G is a generalized sigmoidal function, then the range of G is bounded and hence T is compact. So, by Brouwer's fixed point theorem, T has at least one fixed point.

Next, choose $u^*=(u_1^*,u_2^*,\dots,u_N^*)^T\in F^{-1}(0)$ arbitrarily. We want to show that any trajectory $u(t)$ of (2.2) will converge to u^* as $t\to\infty$, which implies that u^* is both the unique equilibrium state of (2.2) and globally attractive and hence proves the theorem. To this end, for any given trajectory $u(t)=(u_1(t),u_2(t),\dots,u_N(t))^T$ of (2.2), define

$$E(u(t))=\sum_{i=1}^N \xi_i\int_{u_i^*}^{u_i(t)}(g_i(s)-g_i(u_i^*))\,ds. \qquad (4.23)$$

The proof will be broken down into the following several steps.

Step 1) To show that the limit $\lim_{t\to+\infty}E(u(t))$ exists.

Note first that $E(u(t))\ge 0$ for all $t\ge 0$ since g_i is monotonically increasing. Now, a direct calculation using (2.2) gives

$$\frac{dE(u(t))}{dt}=\sum_{i=1}^N \xi_i(g_i(u_i(t))-g_i(u_i^*))\frac{d(u_i(t)-u_i^*)}{dt}=\sum_{i=1}^N \xi_i(g_i(u_i)-g_i(u_i^*))\left\{-d_i(u_i-u_i^*)+\sum_{j=1}^N w_{ij}(g_j(u_j)-g_j(u_j^*))\right\}. \qquad (4.24)$$

Since, by Lemma 1 and the monotonically increasing property of g_i

$$(g_i(u_i(t))-g_i(u_i^*))\cdot(u_i(t)-u_i^*)\ge\frac{1}{\varphi_{u_i^*}(|u_i(t)-u_i^*|)\cdot L(g_i)}|g_i(u_i(t))-g_i(u_i^*)|^2$$

then it follows from (4.24) that

$$\begin{aligned}\frac{dE(u(t))}{dt}&\le-\sum_{i=1}^N\frac{\xi_i d_i}{\varphi_{u_i^*}(|u_i(t)-u_i^*|)\cdot L(g_i)}|g_i(u_i(t))-g_i(u_i^*)|^2\\&\quad+\sum_{i=1}^N\sum_{j=1}^N w_{ij}\xi_i[g_i(u_i(t))-g_i(u_i^*)]\cdot[g_j(u_j(t))-g_j(u_j^*)]\\&=-\sum_{i=1}^N\frac{\xi_i d_i[1-\varphi_{u_i^*}(|u_i(t)-u_i^*|)]}{\varphi_{u_i^*}(|u_i(t)-u_i^*|)\cdot L(g_i)}|g_i(u_i(t))-g_i(u_i^*)|^2\\&\quad-(G(u(t))-G(u^*))^TM_1(\Gamma)(G(u(t))-G(u^*)).\end{aligned}$$

Noting that $M_1(\Gamma)$ is nonnegative definite and that $0\le\varphi_{u_i^*}(|u_i(t)-u_i^*|)\le 1$ for all $i=1,2,\dots,N$, we have

$$\begin{aligned}\frac{dE(u(t))}{dt}&\le-\sum_{i=1}^N\frac{\xi_i d_i(1-\varphi_{u_i^*}(|u_i(t)-u_i^*|))}{L(g_i)\cdot\varphi_{u_i^*}(|u_i(t)-u_i^*|)}|g_i(u_i(t))-g_i(u_i^*)|^2\\&\le-\beta\cdot\sum_{i=1}^N\frac{1-\varphi_{u_i^*}(|u_i(t)-u_i^*|)}{\varphi_{u_i^*}(|u_i(t)-u_i^*|)}|g_i(u_i(t))-g_i(u_i^*)|^2\end{aligned} \qquad (4.25)$$

where

$$\beta=\min\left\{\frac{\xi_i d_i}{L(g_i)}:1\le i\le N\right\}>0.$$

Thus, $E(u(t))$ is monotonically decreasing and therefore the limit $\lim_{t\to+\infty}E(u(t))$ exists.

Step 2) We want to show that

$$\liminf_{t\to+\infty}|g_i(u_i(t))-g_i(u_i^*)|=0 \qquad (4.26)$$

for all $i\in\{1,2,\dots,N\}$.

If (4.26) were not true for some $i_0\in\{1,2,\dots,N\}$, then there would be two positive constants ε and δ such that, whenever $t\ge\delta$

$$|g_{i_0}(u_{i_0}(t))-g_{i_0}(u_{i_0}^*)|\ge\varepsilon. \qquad (4.27)$$

Thus, by (4.25), we deduce that

$$\begin{aligned} 0 < \varepsilon^2 \cdot \int_{\delta}^{+\infty} \frac{1-\varphi_{u_{i_0}^*}(|u_{i_0}(t)-u_{i_0}^*|)}{\varphi_{u_{i_0}^*}(|u_{i_0}(t)-u_{i_0}^*|)}\, dt \\ \le \int_{\delta}^{+\infty} \frac{1-\varphi_{u_{i_0}^*}(|u_{i_0}(t)-u_{i_0}^*|)}{\varphi_{u_{i_0}^*}(|u_{i_0}(t)-u_{i_0}^*|)} \\ \times |g_{i_0}(u_{i_0}(t)) - g_{i_0}(u_{i_0}^*)|^2\, ds \\ \le -\int_{\delta}^{+\infty} \frac{dE(u(t))}{dt}\, dt \\ = E(u(\delta)) - \lim_{t\to+\infty} E(u(t)) < +\infty. \qquad (4.28) \end{aligned}$$

However, by (4.27) and the Lipschitz continuity of g_{i_0}, we have

$$|u_{i_0}(t) - u_{i_0}^*| \ge \varepsilon / L(g_{i_0}) \qquad \forall t > \delta.$$

This, combined with the strictly decreasing property of $\varphi_{u_{i_0}}^*$, leads to the fact that, whenever $t \ge \delta$

$$\frac{1-\varphi_{u_{i_0}^*}(|u_{i_0}(t)-u_{i_0}|)}{\varphi_{u_{i_0}^*}(|u_{i_0}(t)-u_{i_0}^*|)} \ge \frac{1}{\varphi_{u_{i_0}^*}(\varepsilon/L(g_{i_0}))} - 1 > 0.$$

This implies that the first integral in (4.28) is divergent, which contradicts to the estimate (4.28). The result (4.26) is thus proved.

Step 3) To show that $\lim_{t\to+\infty} \|G(u(t)) - G(u^*)\| = 0$.

By Step 2 there is a subsequence $\{t_n\}$ such that $t_n \to +\infty$ as $n \to \infty$ and

$$\lim_{n\to+\infty} \|G(u(t_n)) - G(u^*)\| = 0.$$

Without loss of generality we may assume that $u(t_n) \to v^*$ as $n \to \infty$ (since $u(t)$ is bounded). Then, by the continuity of G, we have $G(v^*) = G(u^*)$ so, and in view of the strictly increasing property of each $g_i, v^* = u^*$. Thus it follows from (4.23) that $\lim_{n\to\infty} E(u(t_n)) = E(u^*) = 0$, which together with Step 1 implies that $\lim_{t\to+\infty} E(u(t)) = 0$.

On the other hand, by applying Lemma 2 to (4.23), it is seen that

$$E(u(t)) \ge \sum_{i=1}^{N} \frac{\xi_i}{2L(g_i)} [g_i(u_i(t)) - g(u_i^*)]^2.$$

Thus,

$$0 = \lim_{t\to+\infty} E(u(t)) \ge c \cdot \limsup_{t\to+\infty} \{\|G(u(t)) - G(u^*)\|^2\} \ge 0$$

with

$$c = \min_{1\le i\le N} \left\{ \frac{\xi_i}{2L(g_i)} \right\} > 0.$$

Consequently, $\lim_{t\to+\infty} \|G(u(t)) - G(u^*)\| = 0$.

Step 4) To show that $\lim_{t\to+\infty} u(t) = u^*$.

By the differential equation theory, $u(t)$, as a trajectory of (2.2), also solves the following integral equation:

$$\begin{aligned} u(t) - u^* = e^{-(t-s)D}(u(s) - u^*) \\ + \int_s^t e^{-(t-\tau)D} W[G(u(\tau)) - G(u^*)]\, d\tau, \quad t > s. \end{aligned}$$

Let $d_{\min} = \min_{1\le i\le N} d_i$. Then, for all $t > s$, there holds

$$\begin{aligned} \|u(t) - u^*\| \le e^{-(t-s)d_{\min}} \|u(s) - u^*\| \\ + \int_s^t e^{-(t-\tau)d_{\min}} \|W\| \cdot \|G(u(\tau)) - G(u^*)\|\, d\tau. \quad (4.29) \end{aligned}$$

Since, by Step 3, $\lim_{t\to+\infty} \|G(u(t)) - G(u^*)\| = 0$, then, for any $\varepsilon > 0$, there is a $T_\varepsilon > 0$ such that, whenever $t \ge T_\varepsilon$,

$$\|G(u(t)) - G(u^*)\| \le \frac{d_{\min}}{\|W\|} \cdot \varepsilon.$$

Therefore, we conclude from (4.29) that, when $t > s \ge T_\varepsilon$,

$$\begin{aligned} \|u(t) - u^*\| \le e^{-(t-s)d_{\min}} \|u(s) - u^*\| \\ + \varepsilon \cdot d_{\min} \int_s^t e^{-(t-\tau)d_{\min}}\, d\tau \\ < e^{-(t-s)d_{\min}} \|u(s) - u^*\| + \varepsilon. \end{aligned}$$

Letting $t \to +\infty$ in the above inequality yields $\lim_{t\to\infty} \|u(t) - u^*\| \le \varepsilon$, which implies $\lim_{t\to+\infty} u(t) = u^*$ since ε is arbitrary. This, on noting the remark at the beginning of the second paragraph, completes the proof of the theorem. □

As mentioned earlier (cf. the paragraph after Definition 2), a nearest point projection does not satisfy the generalized sigmoid function condition. So Theorem 3 can not be applied directly to the case when G is a nonlinear nearest point projection. The conclusion of Theorem 3, however, is still partly true in this case.

Theorem 4: Let $G = P_\Omega = (g_1, g_2, \ldots, g_N)^T$ be a diagonal projection with a bounded, closed, and convex subset $\Omega \subset \mathbb{R}^N$ and let $D = I_{N\times N}$ (the identity matrix). Assume that there is a positive diagonal matrix $\Gamma = \operatorname{diag}\{\xi_1, \xi_2, \ldots, \xi_N\}$ such that (i) ΓW is symmetric and (ii) $\Gamma(I - W)$ is nonnegative definite. Then system (2.2) is globally convergent on $W(\Omega) + I$, that is, for any trajectory $u(t)$ of (2.2) starting from $W(\Omega) + I$, there corresponds an equilibrium state u^* of (2.2) such that $\lim_{t\to+\infty} u(t) = u^*$.

Proof: Given any trajectory $u(t)$ of (2.2) starting from $u_0 \in W(\Omega) + I$, let $u_0 = Wv_0 + I$ with $v_0 \in \Omega$ and let $v(t)$ be the solution of the following system with initial value $v(0) = v_0$

$$\frac{dv(t)}{dt} = -v(t) + P_\Omega(Wv(t) + I), \qquad v(0) = v_0 \in \Omega. \quad (4.30)$$

Then, by the uniqueness result of differential equations, it is easy to verify that $u(t) \equiv Wv(t) + I$. Further, it is easy to see that $u^* = Wv^* + I$ is an equilibrium state of (2.2) if v^* is an equilibrium state of (4.30). Thus, the convergence of $u(t)$ to an equilibrium state of (2.2) can be shown by studying the asymptotic behavior of $v(t)$.

Let $E_1(v(t)) = (1/2)v^T(t)(\Gamma - \Gamma W)v(t) - v^T(t)\Gamma I$. Then, since ΓW is symmetric, we have

$$\frac{dE_1(v(t))}{dt} = \langle (\Gamma - \Gamma W)v(t) - \Gamma I, -v(t) + P_\Omega(Wv(t) + I) \rangle.$$

Noting that $v(t)$ is bounded with bounded derivatives, it follows that the derivative $dE_1(v(t))/dt$ is a uniformly continuous function of t in $[0,+\infty)$. Further, if we assume that $v(t)\in\Omega$ for all $t\geq 0$, then by using the diagonal projection property of P_Ω we obtain that

$$\begin{aligned}\frac{dE_1(v(t))}{dt}&=-\langle\Gamma(Wv(t)+I-v(t)),P_\Omega(Wv(t)+I)-P_\Omega(v(t))\rangle\\&=-\sum_{i=1}^{N}\xi_i[(Wv(t)+I)_i-v_i(t)]\\&\quad\times[g_i((Wv(t)+I)_i)-g_i(v_i(t))]\leq 0\end{aligned}\tag{4.31}$$

which implies that the limit $\lim_{t\to+\infty}E_1(v(t))$ exists. Thus, applying the well-known Barbalat Lemma [34, p. 123] to $E_1(v(t))$, we obtain that $dE_1(v(t))/dt\to 0$ as $t\to+\infty$.

On the other hand, let v^* be any limit point of $v(t)$, i.e., $\lim_{n\to\infty}v(t_n)=v^*$ for some positive sequence $\{t_n\}$ with $t_n\to\infty$ as $n\to\infty$. Then, from (4.31), it can be deduced that

$$\sum_{i=1}^{N}\xi_i[(Wv^*+I)_i-v_i^*][g_i((Wv^*+I)_i)-g_i(v_i^*)]=0.$$

This, combined with the projection property of each g_i, implies that $(Wv^*+I)_i=v_i^*$ for all $i=1,2,\ldots,N$, that is, v^* is an equilibrium state of (4.30). Let $u^*=Wv^*+q$. Then, it is clear that u^* is an equilibrium state of (2.2) and $\lim_{n\to\infty}u(t_n)=u^*$. Define

$$E_2(t)=\sum_{i=1}^{N}\xi_i\int_{u_i^*}^{u_i(t)}(g_i(s)-g_i(u_i^*))\,ds.$$

Then, by the anti-monotonic property of P_Ω (see Example 2), it can be seen that, for all $t\geq 0$

$$\begin{aligned}\frac{dE_2(t)}{dt}&=\sum_{i=1}^{N}\xi_i(g_i(u_i(t))-g_i(u_i^*))\frac{du_i(t)}{dt}\\&=(P_\Omega(u(t))-P_\Omega(u^*))^T\Gamma W(-v(t)+P_\Omega(u(t)))\\&=(P_\Omega(u(t))-P_\Omega(u^*))^T\Gamma W\\&\qquad\cdot(-v(t)+v^*+P_\Omega(u(t))-P_\Omega(u^*))\\&=-(P_\Omega(u(t))-P_\Omega(u^*))^T\Gamma(u(t)-u^*)\\&\quad+(P_\Omega(u(t))-P_\Omega(u^*))^T\Gamma W(P_\Omega(u(t))-P_\Omega(u^*))\\&\leq-(P_\Omega(u(t))-P_\Omega(u^*))^T(\Gamma-\Gamma W)\\&\quad\cdot(P_\Omega(u(t))-P_\Omega(u^*)).\end{aligned}$$

Thus, the nonnegative definiteness of $\Gamma-\Gamma W$ implies $dE_2(t)/dt\leq 0$ for all $t\geq 0$, that is, $E_2(t)$ is monotonically decreasing with t increasing, so the limit $\lim_{t\to+\infty}E_2(t)$ exists. This, together with the fact that $\lim_{n\to\infty}u(t_n)=u^*$, implies that $\lim_{t\to+\infty}E_2(t)=0$. As a result, we obtain by applying Lemma 2 to each component g_i of P_Ω that

$$\lim_{t\to+\infty}(P_\Omega(u(t))-P_\Omega(u^*))=0.$$

Arguing similarly as in the proof of Theorem 3 (cf. Step 4), it can be shown that $\lim_{t\to+\infty}(u(t)-u^*)=0$.

Now to complete the proof, we have to verify that $v(t)\in\Omega$ for all $t\geq 0$. First, it is easy to see that, as the solution of (4.30), $v(t)$ also solves the integral equation

$$v(t)=e^{-t}v_0+\int_0^t e^{-r}P_\Omega(Wv(t-r)+I)\,dr,\qquad t\geq 0.\tag{4.32}$$

Since P_Ω is a projection into the closed convex set Ω, we obtain that, as the limit of the sum

$$\sum_{i=1}^{n}\frac{t}{n}e^{-\frac{it}{n}}P_\Omega(Wv(t-itn^{-1})+I)$$

the integral term in (4.32), denoted by $Q(t)$, satisfies that $(1-e^{-t})^{-1}Q(t)\in\Omega$ for all $t\geq 0$. Thus, it follows from (4.32) that $v(t)\in\Omega$ when $v_0\in\Omega$. This completes the proof. □

Theorems 3 and 4 can be specified in many particular cases with the aid of Examples 1–3, and we have the following corollaries.

Corollary 1: Assume that $G=(g_1,g_2,\ldots,g_N)^\mathsf{T}$ is diagonally nonlinear with each g_i being generalized sigmoidal function with the minimum Lipschitz constants L_i. Then, system (2.2) has a unique equilibrium state u^* that is globally attractive if there is a set of positive parameters $\{\xi_1,\xi_2,\ldots,\xi_N\}$ such that one of the following conditions (d1)–(d8) is satisfied: for $i=1,2,\ldots,N$.

(d1) $L_i^{-1}\xi_i-\xi_i w_{ii}\geq\sum_{j\neq i}^{N}\xi_j|w_{ij}|$.
(d2) $L_i^{-1}\xi_i-\xi_i w_{ii}\geq\sum_{j\neq i}^{N}\xi_j|w_{ji}|$.
(d3) $L_i^{-1}\xi_i-\xi_i w_{ii}\geq(1/2)\sum_{j\neq i}^{N}|\xi_j w_{ji}+\xi_i w_{ij}|$.
(d4) The matrix measure $\mu(\Gamma(W-I))\leq 0$.
(d5) $\Gamma W+W^T\Gamma$ is nonnegative definite.
(d6) $\xi_i\geq L_i\{\xi_i w_{ii}+\sum_{j\neq i}^{N}\xi_j|w_{ij}|\}^+$.
(d7) $\xi_i\geq L_i\xi_i(w_{ii})^++\sum_{j\neq i}^{N}L_j\xi_j|w_{ij}|$.
(d8) $\xi_i\geq L_i\xi_i(w_{ii})^++(1/2)\sum_{j\neq i}^{N}\{L_j\xi_i|w_{ij}|+\xi_j L_i|w_{ji}|\}$ where $\Gamma=\mathrm{diag}\{\xi_1,\xi_2,\ldots,\xi_N\}$, and $(a)^+=\max\{0,a\}$ for any number a.

Proof: By the M-matrix theorems (cf. [3]), it is easy to verify that each of conditions (d1)–(d8) sufficiently guarantees the nonnegative definiteness of $M_1(\Gamma)$, defined in Theorem 3, for the positive definite diagonal matrix Γ. The corollary thus follows immediately from Theorem 3. □

Corollary 2: Assume that $G=P_\Omega=(g_1,g_2,\ldots,g_N)^T$ is a diagonal projection and that W is symmetric. If either (i) the real part of each eigenvalue of W is not larger than 1 or (ii) the matrix measure $\mu(W)\leq 1$, then (2.2) with $D=I_{N\times N}$ is globally convergent on $W(\Omega)+I$ in the sense of Theorem 4.

Proof: It is easy to verify that each of (i) and (ii) can imply the nonnegative definiteness of $I_{N\times N}-W$. Then, Corollary 2 follows immediate from Theorem 4. □

Corollary 3: Assume that $P_\Omega=(p_1,p_2,\ldots,p_N)^T$ is diagonally nonlinear with each p_i being 1-D projection, defined by

(3.13), and that $Q = (Q_{ij})_{N\times N}$ is symmetric. If $Q \geq 0$ (that is, Q is nonnegative definite), then the neural network model

$$\frac{dv_i}{dt} = -v_i + p_i\left(v_i - \sum_{j=1}^{N} Q_{ij}v_j + q_i\right), \quad i = 1,2,\ldots,N \tag{4.33}$$

is globally convergent on $\Omega = \Pi_{i=1}^{N}[a_i, b_i]$, that is, for any trajectory $v(t)$ starting from Ω, there corresponds an equilibrium state v^* such that $\lim_{t\to+\infty} v(t) = v^*$.

Proof: Let $W = I_{N\times N} - Q$ and let $I = (q_1, q_2, \ldots, q_N)^T$. Then (4.33) can be written in the compact form as (4.30). Corollary 3 thus follows from Theorem 4 on noting the relationship between (4.30) and (2.2) with $D = I_{N\times N}$ as remarked in the first paragraph in the proof of Theorem 4. □

Remark 2: Note that conditions (d1)–(d8) in Corollary 1 are exactly the critical version of conditions (D1)–(D8) in Theorem 1 in the sense that all the strict inequality signs, $>$, are replaced here with nonstrict inequality signs, $\geq$. So Theorem 3 and Corollary 1 are generalizations into the critical case of Theorem 1 (and therefore all the known global asymptotic stability results in, e.g., [2], [4], [6], [9]–[11], [19], [20], [25], [40]) in the sense that, under the corresponding critical conditions, global attractivity is valid instead of the global asymptotic stability. Note, in particular, that Corollary 1 was proved recently in [5] for the specific case when $g_i(x) = \tanh(x)$ under the conditions (d6)–(d8). Corollary 1 directly generalizes this result in the sense that the same result is true for any generalized sigmoid functions rather than the specific sigmoidal function $\tanh(x)$. Corollary 2 is new.

Remark 3: Model (4.33) has recently been employed as an optimization solver for the linear quadratic function

$$J(x) = x^TQx + x^Tq, x = (x_1, x_2, \ldots, x_N)^T \tag{4.34}$$

with the constraints: $a_i \leq x_i \leq b_i, i = 1,2,\ldots,N$ (see, e.g., [12]). For the case when Q is positive definite, it was shown in [25], [36] that (4.33) is globally and exponentially convergent to the unique optimization solution under some additional assumptions (say, I-Q is nonsingular). For the general case when Q is nonnegative, the dynamic behavior of (4.33) is rather complicated since it may have many trajectories going to infinite. However, Corollary 3 implies that, as long as the initial value is taken to be in the constrained region, network (4.33) will eventually converge to an equilibrium state that may be an optimization solution (4.34).

V. Example

In this section, we give one example to illustrate the results obtained in previous sections.

Example 4: Consider the neural networks

$$\frac{du_1(t)}{dt} = -u_1(t) + g_1(u_1(t)) + 2g_2(u_2(t)) + 2 \tag{5.35}$$

$$\frac{du_2(t)}{dt} = -u_2(t) - 3g_1(u_1(t)) + 0.5g_2(u_2(t)) - 1 \tag{5.36}$$

for $t \geq 0$, where the activation functions g_1 and g_2 are defined as follows:

$$g_1(u) = \frac{4}{1+e^{-u}} \quad g_2(u) = \frac{2}{1+e^{-2u}} - 1, \qquad u \in \mathbb{R}.$$

It is easy to calculate that the minimum Lipschitz constants L_i of $g_i (i = 1,2)$ are both equal to 1. Thus, by setting $\Gamma = \mathrm{diag}(3,2)$, we get that

$$M_1(\Gamma) = L^{-1}D\Gamma - \frac{\Gamma W + W^T\Gamma}{2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is clearly seen that the matrix $M_1(\Gamma)$ is nonnegative definite. Thus, by Theorem 3 the system (5.35)–(5.36) has a unique equilibrium point u^*, and u^* is globally attractive. In fact, $u^* = (4, -13.5)^T$.

On the other hand, it is easy to verify that, for any positive diagonal matrix Γ, the matrix $M_1(\Gamma)$ is not positive definite. Further, it can be shown that the conditions (d6), (d7), and (d8) of Corollary 1, which were used by Chen and Amiri [6], are not satisfied.

VI. Conclusion

The global convergence of the Hopfield-type neural networks has been studied under the critical condition in the sense that $M_1(\Gamma)$, defined in Theorems 1 and 3, is nonnegative definite for any positive definite diagonal matrix Γ. Two intrinsic properties of the nonlinear activation mapping were first explored and then used in conjunction with the energy function approach and *a priori* decay estimates for solutions of the system to establish two generic critical global convergence theorems for the Hopfield-type neural networks. The results obtained generalized some of the existing results on global exponential stability of the Hopfield-type neural networks to the critical case. The general results developed have been specified in several particular cases which have been compared with some of the known global exponential stability results. Furthermore, our results can be applied directly to many concrete examples of the Hopfield-type neural networks. Finally, an example has been presented to demonstrate both theoretical importance and practical significance of the critical results obtained. It is expected that the critical global convergence results obtained in this paper can be extended to the time-delay case and to other neural network systems.

Acknowledgment

The authors thank the referees for their suggestions and comments on the original version of this paper.

References

[1] L. B. Almeida, "Backpropagation in perceptrons with feedback," in *Neural Computers*, R. Eckmiller and C. Von der Malsburg, Eds. New York: Springer-Verlag, 1988, pp. 199–208.

[2] S. Arik, "Global asymptotic stability of a class of dynamical neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl., Fundam. Theory Appl.*, vol. 47, no. 4, pp. 568–571, Apr. 2000.

[3] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, 2nd ed. New York: Academic, 1994.

[4] A. Bouzerdoum and T. R. Pattison, "Neural network for quadratic optimization with bound constraints," *IEEE Trans. Neural Netw.*, vol. 4, no. 2, pp. 293–303, Mar. 1993.

[5] T. P. Chen and S. I. Amari, "Stability of asymmetric Hopfield networks," *IEEE Trans. Neural Netw.*, vol. 12, no. 1, pp. 159–163, Jan. 2001.
[6] ——, "New theorems on global convergence of some dynamical systems," *Neural Netw.*, vol. 14, pp. 251–255, 2001.
[7] L. O. Chua and L. Yang, "Cellular neural networks: Theory and applications," *IEEE Trans. Circuits Syst.*, vol. 35, no. 12, pp. 1257–1290, Dec. 1988.
[8] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks," *IEEE Trans. Syst., Man Cybern.*, vol. SMC-13, no. 10, pp. 815–826, Oct. 1983.
[9] P. V. D. Driessche and X. F. Zou, "Global attractivity in delayed Hopfield neural networks," *SIAM J. Appl. Math.*, vol. 58, no. 6, pp. 1878–1890, 1998.
[10] Y. Fang and T. G. Kincaid, "Stability analysis of dynamical neural networks," *IEEE Trans. Neural Netw.*, vol. 7, no. 4, pp. 996–1006, Jul. 1996.
[11] M. Forti and A. Tesi, "New conditions for global stability of neural networks with application to linear and quadratic programming problems," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 42, no. 7, pp. 354–366, Jul. 1995.
[12] T. L. Friesz, D. H. Bernstein, N. J. Mehta, R. L. Tobin, and S. Ganjlizadeh, "Day-to-day dynamic network disequilibrium and idealized traveler information systems," *Operations Res.*, vol. 42, pp. 1120–1136, 1994.
[13] Z. H. Guan, G. Chen, and Y. Qin, "On equilibria, stability, and instability of Hopfield neural networks," *IEEE Trans. Neural Netw.*, vol. 11, no. 2, pp. 534–540, Mar. 2000.
[14] S. Haykin, *Neural Networks: A Comprehensive Foundation*. New York: Macmillan, 1994.
[15] J. A. Hertz, A. Krogh, and R. G. Palmer, *Introduction to Theory of Neural Computation*. Reading, MA: Addison-Wesley, 1994.
[16] M. W. Hirsch, "Convergent activation dynamics in continuous time networks," *Neural Netw.*, vol. 2, pp. 331–349, 1989.
[17] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," in *Proc. Nat. Acad. Sci.*, vol. 79, 1982, pp. 2554–2558.
[18] J. J. Hopfield and D. W. Tank, "Computing with neural circuits: A model," *Science*, vol. 233, pp. 625–633, 1986.
[19] J. C. Juang, "Stability analysis of Hopfield-type neural networks," *IEEE Trans. Neural Netw.*, vol. 10, no. 6, pp. 1366–1374, Nov. 1999.
[20] E. Kaszkurewicz and A. Bhays, "On a class of globally stable neural circuits," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 41, no. 1, pp. 171–174, Jan. 1994.
[21] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*. New York: Academic Press, 1980.
[22] B. Kosko, "Bidirectional associative memories," *IEEE Trans. Syst., Man Cybern.*, vol. 18, no. 1, pp. 49–60, Jan./Feb. 1988.
[23] J. P. LaSalle, *The Stability of Dynamical Systems*. Philadelphia, PA: SIAM, 1976.
[24] J. Li, A. N. Michel, and W. Porod, "Analysis and synthesis of a class of neural networks: Linear systems operating on a closed hypercube," *IEEE Trans. Circuits Syst.*, vol. 36, no. 11, pp. 1406–1422, Nov. 1989.
[25] X. B. Liang and 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.*, vol. 12, no. 2, pp. 349–359, Mar. 2001.
[26] X. B. Liang and J. Wang, "Absolute exponential stability of neural networks with a general class of activation functions," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 47, no. 8, pp. 1258–1263, Aug. 2000.
[27] ——, "A recurrent neural network for nonlinear optimization with a continuously differentiable objective function and bound constrains," *IEEE Trans. Neural Netw.*, vol. 11, no. 6, pp. 1251–1262, Nov. 2000.
[28] K. Matsuoka, "On absolute stability of neural networks," *Trans. Inst. Electron. Inform. Commum. Eng. Japan*, pp. 536–542, 1991.
[29] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equation in Several Variables*. New York: Academic Press, 1970.
[30] J. Park, H. Y. Kim, Y. Park, and S. W. Lee, "A synthesis procedure for associative memories based on space-varying cellular neural networks," *Neural Netw.*, vol. 14, pp. 107–113, 2001.
[31] F. J. Pineda, "Generalization of back-propagation to recurrent neural networks," *Phys. Rev. Lett.*, vol. 59, pp. 2229–2232, Nov. 1987.
[32] H. Qiao, J. Peng, and Z. B. Xu, "Nonlinear measures: A new approach to exponential stability analysis for Hopfield-type neural networks," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 360–370, Mar. 2001.
[33] T. Roska and J. Vandewalle, *Cellular Neural Networks*. Chichester, U.K.: Wiley, 1995.
[34] J. J. E. Slotine and W. Li, *Applied Nonlinear Control*. Englewood Cliffs, NJ: Prentice-Hall, 1991.
[35] I. Varga, G. Elek, and H. Zak, "On the brain-state-in-a-convex-domain neural models," *Neural Netw.*, vol. 9, no. 7, pp. 1173–1184, 1996.
[36] Y. S. Xia and J. Wang, "On the stability of globally projected dynamical systems," *J. Optim. Theory Appl.*, vol. 106, pp. 129–150, Jul. 2000.
[37] ——, "A general methodology for designing globally convergent optimization neural networks," in *IEEE Trans. Neural Netw.*, vol. 9, Nov. 1998, pp. 1331–1343.
[38] Z. B. Xu and C. P. Kwong, "Associative memories," in *Neural Networks Systems: Techniques and Applications*, C. T. Leondes, Ed. New York: Academic Press, 1998, vol. 3, pp. 183–258.
[39] H. Yang and T. S. Dillon, "Exponential stability and oscillation of Hopfield graded response neural networks," *IEEE Trans. Neural Netw.*, vol. 5, no. 5, pp. 719–729, Sep. 1994.
[40] Y. P. Zhang, A. Heng, and A. W. C. Fu, "Estimate of exponential convergence rate and exponential stability for neural networks," *IEEE Trans. Neural Netw.*, vol. 10, no. 6, pp. 1487–1493, Nov. 1999.

Jigen Peng received the M.S. degree in mathematics and the Ph.D. degree in applied mathematics from Xi'an Jiaotong University, Xi'an, China, in 1992 and 1998, respectively.

From January to July 1999, and from May to October 2001, he was a Senior Associate in the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong. He is currently a Professor in the Institute for Information and System Sciences, Research Center for Applied Mathematics, Xi'an Jiaotong University. He has published more than 30 academic papers on functional analysis, transport theory, stability of dynamical systems, and neural networks. His current research interests include nonlinear operator theory with applications in engineering, control of dynamical systems, and neural networks.

Zong-Ben Xu received the M.S. degree in mathematics and the Ph.D. degree in applied mathematics from Xi'an Jiaotong University, Xi' an, China, in 1981 and 1987, respectively.

In 1988, he was a Postdoctoral Researcher in the Department of Mathematics, University of Strathclyde, Strathclyde, U.K. 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, Hong Kong. From January 1995 to April 1995, he was a Research Fellow in the Department of Computing, The Hong Kong Polytechnic University, Hong Kong. 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 was Dean of the Faculty of Science, and now serves as an authorized Ph.D. Supervisor in mathematics and computer science, Vice-President of the University, and Director of the Institute for Information and System Sciences. He has published four monographs and more than 110 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 "Chinese Ph.D. degree holder with outstanding achievements," awarded by the Chinese State Education Commission (CSEC) and the Academic Degree Commission of the Chinese Council in 1991, and get scientific awards from the university, and CSEC several times. He is a member of the New York Academy of Sciences.

Hong Qiao (M'03) received the B.Eng. degree in hydraulics and control and the M.Eng. degree in robotics from Xi'an Jiaotong University, Xi'an, China, the M.Phil. degree in robotics control from the Industrial Control Centre, University of Strathclyde, Strathclyde, U.K., and the Ph.D. degree in robotics and AI from De Montfort University, Leicester,U.K., in 1995.

She was a University Research Fellow at De Montfort University from 1995 to 1997. She was a Research Assistant Professor from 1997 to 2000, and an Assistant Professor from 2000 to 2002, both in the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong. Since January 2002, she has been a Lecturer in the School of Informatics, University of Manchester, Manchester U.K. Currently, she is also a Professor in the Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing, China (on leave from the University of Manchester). She first proposed the concept "attractive regions in the environment" which has been applied successfully by herself in robotic assembly, robotic grasping and part recognition. The work has been reported by the Magazine "*Advanced Manufacturing Alert*" (New York: Wiley, 1999). Her current research interests include information-based strategy investigation, robotics, and intelligent agents, animation, machine learning (neural networks, SVM) and pattern recognition.

Dr. Qiao is a member of the Program Committee of the IEEE International Conference on Robotics and Automation in 2001, 2002, 2003, and 2004. She is currently an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, PART B.

Bo Zhang received the B.Sc. degree in mathematics from Shandong University, Jinan, China, the M.Sc. degree in mathematics from Xi'an Jiaotong University, Xi'an, China, and the Ph.D. degree in applied mathematics from University of Strathclyde, Strathclyde, U.K. in 1983, 1985, and 1992, respectively.

He was Post-Doctoral Research Fellow in the Department of Mathematics, Keele University, Keele, U.K., from January 1992 to December 1994, and in the Department of Mathematical Sciences, Brunel University, Uxbridge, U.K., from January 1995 to October 1997. In November 1997, he joined the School of Mathematical and Informational Sciences at Coventry University, Coventry, U.K., as a Senior Lecturer, where he was promoted to Reader in Applied Mathematics in September 2000 and Professor of Applied Mathematics in September 2003. Currently, he is also a Professor in the Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China (on leave from Coventry University). From November 1999 to September 2000 he was on leave working at the Department of Electronic Engineering,t City University of Hong Kong, Hong Kong as a Visiting Research Fellow and from 1985 to 1988 he was a Lecturer in the Department of Mathematics, Xi'an Jiaotong University. His current research interests include direct and inverse scattering problems, computational electromagnetics, numerical analysis, partial differential equations, and machine learning.

Dr. Zhang is a member of the American and London Mathematical Societies and the Society for Industrial and Applied Mathematics (SIAM).