SIAM J. APPL. MATH.
Vol. 58, No. 6, pp. 1878–1890, December 1998

011

GLOBAL ATTRACTIVITY IN DELAYED HOPFIELD NEURAL NETWORK MODELS*

P. VAN DEN DRIESSCHE† AND XINGFU ZOU†

Abstract. Two different approaches are employed to investigate the global attractivity of delayed Hopfield neural network models. Without assuming the monotonicity and differentiability of the activation functions, Liapunov functionals and functions (combined with the Razumikhin technique) are constructed and employed to establish sufficient conditions for global asymptotic stability independent of the delays. In the case of monotone and smooth activation functions, the theory of monotone dynamical systems is applied to obtain criteria for global attractivity of the delayed model. Such criteria depend on the magnitude of delays and show that self-inhibitory connections can contribute to the global convergence.

Key words. neural network, delay, global attractivity, monotone dynamical system

AMS subject classifications. 34C35, 34D20, 34D45, 34K20, 92B20

PII. S0036139997321219

1. Introduction.

There has recently been increasing interest in the potential applications of the dynamics of artificial neural networks in signal and image processing. Among the most popular models in the literature of artificial neural networks is the following continuous time model described by a system of ordinary differential equations for $u_i(t)$, the voltage on the input of neuron i at time t:

$$C_i \frac{du_i(t)}{dt} = -\frac{u_i(t)}{R_i} + \sum_{j=1}^{n} T_{ij}\, g_j(u_j(t)) + I_i, \quad i = 1, 2, \dots, n. \tag{1.1}$$

Here $n \geq 2$ is the number of neurons in the network. For neuron i, $C_i > 0$ and $R_i > 0$ are the neuron amplifier input capacitance and resistance, respectively, and I_i is the constant input from outside the system. The $n \times n$ matrix $T = (T_{ij})$ represents the connection strengths between neurons, and if the output from neuron j excites (resp., inhibits) neuron i, then $T_{ij} > 0$ (resp., < 0). The matrix T is assumed to be irreducible, i.e., the network is strongly connected. The functions g_j are neuron activation functions. This model for n neurons was proposed by Hopfield [13] with an electrical circuit implementation and is thereafter referred to in the literature as a Hopfield-type neural network. In Hopfield's analysis [13], T is assumed symmetric, and functions g_j are assumed to be C^∞ sigmoid functions.

Hopfield [13] realized that in hardware implementation, time delays occur due to finite switching speeds of the amplifiers. A single time delay $\tau > 0$ was first introduced into (1.1) by Marcus and Westervelt [19]. They considered the following system of differential equations with delay:

$$C_i \frac{du_i(t)}{dt} = -\frac{u_i(t)}{R_i} + \sum_{j=1}^{n} T_{ij}\, g_j(u_j(t-\tau)) + I_i, \quad i = 1, 2, \dots, n. \tag{1.2}$$

*Received by the editors May 9, 1997; accepted for publication (in revised form) October 17, 1997; published electronically August 18, 1998. This research was partially supported by the Natural Sciences and Engineering Research Council of Canada.

http://www.siam.org/journals/siap/58-6/32121.html

†Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3P4 (pvdd@math.uvic.ca, xzou@math.uvic.ca).

System (1.2) has much more complicated dynamics than (1.1) due to the incorporation of delay. For results on this system, see, for example, Bélair, Campbell, and van den Driessche [3], Marcus and Westervelt [19], Wu [26], and Wu and Zou [27].

Systems (1.1) and (1.2) can be rewritten for $u = (u_1, u_2, \ldots, u_n)^T$ as

$$\frac{du(t)}{dt} = -Bu(t) + A\,g(u(t)) + J \tag{1.3}$$

and

$$\frac{du(t)}{dt} = -Bu(t) + A\,g(u(t-\tau)) + J, \tag{1.4}$$

respectively, where $B = \mathrm{diag}(b_1, b_2, \ldots, b_n)$ with $b_i = 1/(R_iC_i)$, the $n \times n$ irreducible connection matrix $A = (a_{ij})$ with $a_{ij} = T_{ij}/C_i$, $g(u) = (g_1(u_1), g_2(u_2), \ldots, g_n(u_n))^T$ and $J = (J_1, J_2, \ldots, J_n)^T$ with $J_i = I_i/C_i$ for $i, j = 1, 2, \ldots, n$.

Gopalsamy and He [10] recently considered a modification of (1.4) by incorporating different delays $\tau_{ij} \geq 0$ in different communication channels (from neuron j to neuron i), namely,

$$\frac{du_i(t)}{dt} = -b_iu_i(t) + \sum_{j=1}^{n} a_{ij}g_j(u_j((t-\tau_{ij})) + J_i, \quad i = 1, 2, \ldots, n. \tag{1.5}$$

Clearly (1.3) and (1.4) are special cases of (1.5). The initial conditions associated with (1.5) are of the form

$$u_i(s) = \phi_i(s) \quad \text{for } s \in [-\tau, 0], \text{ where } \tau = \max_{1 \leq i,\, j \leq n} \tau_{ij},$$

where it is usually assumed that $\phi_i \in C([-\tau, 0], R)$, $i = 1, 2, \ldots, n$.

Hopfield-type neural networks (1.3) and (1.4) and their various generalizations have attracted the attention of the scientific community (e.g., mathematicians, physicists, and computer scientists), due to their promising potential for the tasks of classification, associative memory, and parallel computation and their ability to solve difficult optimization problems. When a neural circuit is employed as an associative memory, the existence of many equilibrium points is a necessary feature. However, in applications to parallel computation and signal processing involving the solution of optimization problems, it is required that there be a well-defined computable solution for all possible initial states. From a mathematical viewpoint, this means that the network should have a unique equilibrium point that is globally attractive. Indeed, earlier applications to optimization problems have suffered from the existence of a complicated set of equilibria (see Tank and Hopfield [25]). Thus, the global attractivity of systems is of great importance for both practical and theoretical purposes and has been the major concern of most authors dealing with (1.3) and (1.4) and their generalizations (e.g., (1.5)). See, for example, Bélair [2], Cao and Wu [5], Cohen and Grossberg [6], Forti [8], Forti, Manetti, and Marini [9], Gopalsamy and He [10], Hirsch [12], Kelly [15], Li, Michel, and Porod [17], Michel, Farrell, and Porod [18], and Matsuoka [20].

To the best of the authors' knowledge, existing results on (1.3) and (1.4) and their generalizations have been obtained under the assumption that the nonlinear neuron activation functions are sigmoid, that is,

(H_1) For each $j \in \{1, 2, \ldots, n\}$, $g_j \in C^1(R)$, $g_j'(x) > 0$ for $x \in R$, and $g_j'(0) = \sup_{x \in R} g_j'(x) > 0$.

(H_2) For each $j \in \{1, 2, \ldots, n\}$, $g_j(0) = 0$ and $g_j(x)$ saturates at ± 1, i.e., $\lim_{x\to\pm\infty} g_j(x) = \pm 1$.

Hence g_j's have always been assumed to be continuously differentiable and monotonically increasing with $|g_j(x)| \leq 1$.

Recently, Morita [21], and Yoshizawa, Morita, and Amari [28] have shown that the absolute capacity of an associative memory model can be remarkably improved by replacing the usual sigmoid activation functions with nonmonotonic activation functions. Therefore, it seems that for some purposes, nonmonotonic (and not necessarily smooth) functions might be better candidates for neuron activation in designing and implementing an artificial neural network. In many electronic circuits, amplifiers that have neither monotonically increasing nor continuously differentiable input-output functions are frequently adopted. For example, in designing an optimization network for matrix inversion, Jang, Lee, and Shin [14] used cubic-like input-output functions (nonmonotone) in implementation, while Tank and Hopfield [25] designed a linear programming network that is also of the form (1.1) with all g_j equal and piecewise linear (nonsmooth). This is practical motivation for relaxing conditions (H_1), (H_2) to (A_1), (A_2) below.

In their proofs of the global stability results, previous authors almost all made use of the monotonicity and/or smoothness of the activation functions to construct Liapunov functions or functionals and to estimate their derivatives along the solutions. Therefore, some of their arguments may not be valid without the monotonicity and/or smoothness condition(s). As an example of this aspect, let us consider (1.3) with B equal to the identity matrix and A skew symmetric (i.e., $A^T = -A$). With these assumptions, it has been shown (see Forti [8, Corollary 3], or Matsuoka [20, p. 497]) that if g_j, $j = 1, 2, \ldots, n$, are continuously differentiable, bounded, and strictly increasing, then for every $J \in R^n$, system (1.3) has a unique equilibrium that is globally asymptotically stable. However, consider the following specific system

$$\frac{du(t)}{dt} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} u(t) + \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix} \begin{pmatrix} \sin u_1(t) \\ \sin u_2(t) \end{pmatrix}, \tag{1.6}$$

where $k > 0$. It is easily verified that all the assumptions as stated above are satisfied except for the monotonicity condition on the activation functions $g_j(u_j) = \sin u_j$, $j = 1, 2$. When $k > 0$ is sufficiently large (e.g., $k > \frac{3\pi}{2}$), system (1.6) has a nonzero equilibrium in addition to the zero equilibrium, and hence the above result about the global convergence fails to be valid for system (1.6) for such k.

The purpose of this paper is to establish some criteria for the global attractivity of (1.4) and (1.5). In section 2, such global conditions are obtained by constructing Liapunov functions and functionals without assuming monotonicity or smoothness of the activation functions g_j, $j = 1, \ldots, n$. Instead of (H_1) and (H_2), we assume the following:

(A_1) For each $j \in \{1, 2, \ldots, n\}$, $g_j : R \to R$ is globally Lipschitz with Lipschitz constant L_j, i.e., $|g_j(u_j) - g_j(v_j)| \leq L_j |u_j - v_j|$ for all u_j, $v_j \in R$.

(A_2) For each $j \in \{1, 2, \ldots, n\}$, $|g_j(x)| \leq M_j$, $x \in R$ for some constant $M_j > 0$.

Assumptions (A_1) and (A_2) are clearly weaker than (H_1) and (H_2), and the criteria found are independent of the magnitude of the delay(s). In section 3, we employ the theory of monotone dynamical systems to derive sufficient conditions under the hypotheses of monotonicity and differentiability of the activation functions. We prove that for excitatory connections ($a_{ij} \geq 0$, for $j \neq i$), if system (1.3) is globally attractive, then system (1.5) is also globally attractive provided that the diagonal delays τ_{ii}

corresponding to negative a_{ii} are sufficiently small. This conclusion is also true for some networks with inhibitory connections.

2. Criteria independent of the magnitude of delays. Gopalsamy and He [10] also assumed (H_1) and (H_2) in the proof of their principal result, but a slight modification shows that their stability criterion remains valid for (1.5) with only assumptions (A_1) and (A_2).

THEOREM 2.1. *Suppose* (A_1) *and* (A_2) *hold and* $\tau_{ij} \geq 0$, $i, j = 1, \ldots, n$. *If*

$$\alpha \triangleq \max_{1\leq j\leq n} \left\{ \frac{L_j}{b_j} \sum_{i=1}^{n} |a_{ij}| \right\} < 1, \tag{2.1}$$

then, for every input J, system (1.5) *has a unique equilibrium u^* that is globally asymptotically stable, independent of the delays.*

The proof follows by using a fixed point argument and the same Liapunov functional as used by Gopalsamy and He [10], but with the neuron gains replaced by the Lipschitz constants L_j, namely,

$$V(u)(t) = \sum_{i=1}^{n} \left(|u_i(t) - u^*| + \sum_{j=1}^{n} |a_{ij}| L_j \int_{t-\tau_{ij}}^{t} |u_j(s) - u_j^*| ds \right).$$

By constructing a different Liapunov functional, we can obtain a new criterion that is, in general, independent of Theorem 2.1. Recall that the spectral norm of matrix Z is defined as

$$\|Z\|_2 = \left(\max \left\{ \lambda : \lambda \quad \text{is an eigenvalue of} \quad Z^T Z \right\}\right)^{1/2}.$$

THEOREM 2.2. *Suppose* (A_1) *and* (A_2) *hold and* $\tau_{ij} \geq 0$, $i, j = 1, 2, \ldots, n$. *If*

$$\beta \triangleq \max_{1\leq i\leq n} \left\{ \frac{1}{b_i} \sum_{j=1}^{n} |a_{ij}| \right\} + \max_{1\leq j\leq n} \left\{ L_j^2 \sum_{i=1}^{n} \frac{|a_{ij}|}{b_i} \right\} < 2, \tag{2.2}$$

then, for every input J, system (1.5) *has a unique equilibrium that is globally asymptotically stable, independent of the delays.*

Proof. For every fixed input J, let $\Psi : \; R^n \to R^n$ be defined by

$$\Psi(u) = B^{-1} A g(u) + B^{-1} J, \qquad u \in R^n.$$

Then u^* is an equilibrium of (1.5) if and only if u^* is a fixed point of Ψ. From (A_1) and (A_2), we know that Ψ is continuous and

$$\begin{aligned}
\left\|\Psi(u) - B^{-1}J\right\|_2 &= \left\|B^{-1}Ag(u)\right\|_2 \\
&\leq \|B^{-1}A\|_2 \, \|g(u)\|_2 \\
&\leq \|B^{-1}A\|_2 \left(\sum_{i=1}^{n} M_i^2 \right)^{1/2} \triangleq M\|B^{-1}A\|_2, \quad u \in R^n.
\end{aligned}$$

Thus, Ψ maps the closed ball

$$\Omega = \left\{ u \in R^n : \; \|u - B^{-1}J\|_2 \leq M\|B^{-1}A\|_2 \right\}$$

into itself. By Brouwer's fixed point theorem, Ψ has a fixed point u^* in Ω that is an equilibrium of (1.5).

We next prove the global asymptotic stability of u^*. Let $x(t) = u(t) - u^*$. Then (1.5) becomes

$$\frac{dx_i(t)}{dt} = -b_i x_i(t) + \sum_{j=1}^{n} a_{ij} f_j(x_j(t-\tau_{ij})), \quad i = 1, 2, \ldots, n, \tag{2.3}$$

where

$$f_j(x_j) = g_j(x_j + u_j^*) - g_j(u_j^*), \;\; j = 1, 2, \ldots, n.$$

Clearly, u^* is globally asymptotically stable for (1.5) if and only if the trivial solution of (2.3) is globally asymptotically stable.

Let

$$V(x)(t) = \sum_{i=1}^{n} \frac{1}{b_i} x_i^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{|a_{ij}|}{b_i} \int_{t-\tau_{ij}}^{t} f_j^2(x_j(s))\, ds. \tag{2.4}$$

We calculate and estimate the derivative of $V(x)(t)$ along (2.3) as follows:

$$\begin{aligned}
&\frac{dV(x)(t)}{dt}\\
&= -2\sum_{i=1}^{n} x_i^2(t) + 2\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{a_{ij}}{b_i} x_i(t) f_j(x_j(t-\tau_{ij}))\\
&+ \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{|a_{ij}|}{b_i} f_j^2(x_j(t)) - \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{|a_{ij}|}{b_i} f_j^2(x_j(t-\tau_{ij}))\\
&\le -2\sum_{i=1}^{n} x_i^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{|a_{ij}|}{b_i} \left\{x_i^2(t) + f_j^2(x_j(t-\tau_{ij}))\right\}\\
&+ \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{|a_{ij}|}{b_i} f_j^2(x_j(t)) - \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{|a_{ij}|}{b_i} f_j^2(x_j(t-\tau_{ij}))\\
&\le -2\sum_{i=1}^{n} x_i^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{|a_{ij}|}{b_i} x_i^2(t)\\
&+ \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{|a_{ij}|}{b_i} L_j^2 x_j^2(t)\\
&\le \left(-2 + \max_{1\le i\le n} \left\{ \frac{1}{b_i} \sum_{j=1}^{n} |a_{ij}| \right\} + \max_{1\le j\le n} \left\{ L_j^2 \sum_{i=1}^{n} \frac{|a_{ij}|}{b_i} \right\} \right) \sum_{i=1}^{n} x_i^2(t)\\
&\triangleq -(2-\beta) \sum_{i=1}^{n} x_i^2(t).
\end{aligned}$$

Now, by a standard Liapunov-type theorem in functional differential equations (see, e.g., Kuang [16, Chapter 2, Corollary 5.2]), for $\beta < 2$ the trivial solution of (2.3) is globally asymptotically stable, and therefore, u^* is global asymptotically stable for (1.5). □

Conditions (2.1) and (2.2) are explicit from (1.5) and hence are convenient to verify in practice. But both of them have the disadvantage of neglecting the signs of entries in the connection matrix A, and thus, differences between excitatory and inhibitory effects might be ignored. In general, this is overly restrictive. In the case of no delay (i.e., $\tau_{ij} = 0,\ \ i, j = 1, 2, \dots, n$) and $b_i = 1,\ \ i = 1, 2, \dots, n$, Atiya [1] and Hirsch [12] obtained conditions with this same disadvantage. Forti [8] and Matsuoka [20] recently attempted to overcome this disadvantage for networks without delay. The Liapunov functions in Forti [8] and Matsuoka [20] depend strongly on the monotonicity of the activation functions and the hypothesis $\tau_{ij} = 0,\ \ i, j = 1, 2, \dots, n$, and hence fail to work for (1.5) assuming only (A_1) and (A_2).

When A is symmetric and functions $g_j(u_j)$ are all equal and sigmoid, Bélair [2] established a criterion for global attractivity of (1.4) with all $b_i = 1$ and input $J = 0$. We next extend this result (Bélair [2, Theorem 3.2]) to the general case of (1.4) with delay τ with only assumptions (A_1) and (A_2). To this end, we introduce the Liapunov–Razumikhin technique developed by Haddock and Terjéki [11]. See Bélair [2] for a simplified version.

Consider the functional differential equation

$$\frac{dx(t)}{dt} = f(x_t), \tag{2.5}$$

where $f: C \to R^n$, with $C \triangleq C([-\tau, 0], R^n)$, and $x_t \in C$ is defined by $x_t(s) = x(t+s)$ for $s \in [-\tau, 0]$. Let $V: R^n \to R$ be a Liapunov function, which we define as a C^1 function. The (upper right-hand) derivative of V along a solution $x(t, \phi)$ of (2.6) is defined by

$$\dot{V}(\phi) = \overline{\lim}_{h\to 0^+} \frac{[V(\phi(0) + hf(\phi)) - V(\phi(0))]}{h} = \sum_{i=1}^{n} \frac{\partial V(\phi(0))}{\partial x_i} f_i(\phi), \tag{2.6}$$

where f_i denotes the ith component of f.

THEOREM 2.3 (see Bélair [2, Theorem 3.1]). *Suppose that f is continuous and maps bounded sets in C into bounded sets in R^n, and $f(0) = 0$. Assume that there exists a Liapunov function $V(x)$ and a constant N such that*

$$V(0) = 0, \quad V(x) > 0 \quad \textit{for all} \quad 0 \neq |x| < N, \quad \dot{V}(0) = 0 \tag{2.7}$$

and

$$\dot{V}(\phi) < 0 \quad \textit{for all} \quad 0 \neq \|\phi\| < N \quad \textit{such that} \quad \max_{-\tau\le s\le 0} V(\phi(s)) = V(\phi(0)).$$

Then, the solution $x = 0$ of (2.5) is asymptotically stable. In addition, for each solution $x(\phi)(\cdot)$ such that $\|x_t(\phi)\| < N$ for all $t \geq 0$, $x_t(\phi) \to 0$ in C as $t \to \infty$.

Now we use the above to establish the following result in terms of the spectral norm.

THEOREM 2.4. *Suppose (A_1) and (A_2) hold and $\tau \geq 0$. If*

$$\gamma \triangleq \frac{L}{b}\|A\|_2 < 1, \tag{2.9}$$

where $L = \max_{1\le i\le n} L_i$ and $b = \min_{1\le i\le n} b_i$, then, for any input J, system (1.4) has a unique equilibrium that is globally asymptotically stable, independent of the delay.

Proof. For every input J, the same argument as in the proof of Theorem 2.2 shows that (1.4) has an equilibrium u^*. Set $x(t) = u(t) - u^*$. Then, $u(t)$ is a solution of (1.4) if and only if $x(t)$ solves

$$\frac{dx(t)}{dt} = -Bx(t) + AF(x_t), \tag{2.10}$$

where $F(x_t) = (f_1(x_t), \cdots, f_n(x_t))^T$ is defined by

$$f_j(x_t) = g_j(x_j(t-\tau) + u_j^*) - g_j(u_j^*). \tag{2.11}$$

As in the proof of Theorem 2.2, we only need to prove that the trivial solution of (2.10) is globally asymptotically stable. Now, consider the Liapunov function $V(x) = \frac{1}{2}\|x\|_2^2$. Then along (2.10)

$$\begin{aligned}\dot{V}(x) = & -x^T(t)Bx(t) + x^T(t)AF(x_t) \\ & \leq -x^T(t)Bx(t) + \|x(t)\|_2 \cdot \|AF(x_t)\|_2 \\ & \leq -x^T(t)Bx(t) + \|x(t)\|_2 \cdot \|A\|_2 \cdot \|F(x_t)\|_2.\end{aligned} \tag{2.12}$$

From (2.11) and (A_1), we obtain

$$\|F(x_t)\|_2^2 = \sum_{j=1}^{n} f_j^2(x_t) \leq \sum_{j=1}^{n} L_j^2 x_j^2(t-\tau) \leq L^2\|x(t-\tau)\|_2^2. \tag{2.13}$$

Thus, if (2.9) holds, then for those t satisfying

$$x(t) \neq 0 \quad \text{and} \max_{s\in[-\tau,0]} \|x(t+s)\|_2 = \|x(t)\|_2,$$

we have

$$\begin{aligned}\dot{V}(x) \leq & -x^T(t)Bx(t) + L\|A\|_2 \cdot \|x(t)\|_2^2 \\ & = \sum_{i=1}^{n} (L\|A\|_2 - b_i)\ x_i^2(t) \\ & \leq (L\|A\|_2 - b) \sum_{i=1}^{n} x_i^2(t) < 0.\end{aligned} \tag{2.14}$$

By Theorem 3.1, it follows that the trivial solution of (2.10) is globally asymptotically stable, and therefore, u^* is globally asymptotically stable for (1.4). □

Remark 2.1. *When A is normal (i.e., $AA^T = A^TA$), then*

$$\|A\|_2 = \max\{|\lambda| : \lambda \ \textit{ is an eigenvalue of } \ A\}.$$

So, Theorem 3.2 *in Bélair* [2] *is a special case of the above theorem with A symmetric (hence normal), $b_i = 1$ and $g_j(\cdot) = g(\cdot)$, where $g(\cdot)$ is sigmoid and $L_j = g'(0)$. Cao and Wu* [5, p. 1534] *and Kelly* [15, p. 239] *also obtained similar conditions to* (2.9) *for Hopfield-type models with delays and without delay, respectively, but they also assumed differentiability and monotonicity of the activation functions.*

REMARK 2.2. *If all* $b_i = 1$, *and* A *is doubly stochastic (i.e.,* A *is nonnegative with all row sums and column sums equal to* 1*), then* $\|A\|_2 = 1$. *This follows from Perron–Frobenius theory since* $A^T A(1,1,\ldots,1)^T = (1,1,\ldots,1)^T$ *(see, e.g., Berman and Plemmons* [4, p. 27]*). Then,* (2.1), (2.2), *and* (2.9) *all reduce to* $\max_{1\le j\le n} L_j \le 1$.

EXAMPLE 2.1. *Consider the all-excitatory doubly stochastic connection matrix studied in Marcus and Westervelt* [19] *and Wu* [26], *i.e.,* A *with* $a_{ii} = 0$, $a_{ij} = \frac{1}{n-1}$ *for* $i \ne j$. *Then,* (2.1) *becomes*

$$\alpha \triangleq \max_{1\le j\le n} \frac{L_j}{b_j} < 1. \tag{2.15}$$

If we assume that for $j = 1,2,\ldots,n$, $b_j = 1$ *and* $g_j(u) = g(u)$ *is increasingly sigmoid with neuron gain* $g'(0) = \sup_{x\in R} g'(x) > 0$, *then,* (2.15) *further reduces to*

$$g'(0) < 1. \tag{2.16}$$

It has been proved by Wu [26] *that* (2.16) *is a sufficient and necessary condition for such a network* (1.4) *to have a unique equilibrium that is a global attractor. Note that in this case,* (2.2) *and* (2.9) *also reduces to* (2.16) *as pointed out in Remark* 2.2.

We next give examples to show that (2.1), (2.2), and (2.9) are independent, in the sense that for each of them there exists a system (1.4) for which one of Theorems 2.1, 2.2, or 2.3 applies but the other two fail.

EXAMPLE 2.2. *Consider a doubly stochastic connection matrix* A, *and let* $b_i \le 1/2$ *for* $i = 1,2,\ldots,n$. *If we choose* L_i *such that* $0 < L_i < b_i$ *for* $i = 1,2,\ldots,n$, *and* $\max_{1\le i\le n} L_i \ge \min_{1\le i\le n} b_i$, *then,* (2.1) *is satisfied but both* (2.2) *and* (2.9) *fail.*

EXAMPLE 2.3. *Consider the connection matrix*

$$A = \begin{pmatrix} 1 & \frac{1}{4} \\ 1 & \frac{1}{4} \end{pmatrix},$$

and let $L_1 = 1$, $L_2 = 2$ *and* $\frac{13}{8} < b_1 \le b_2 \le 2$. *Then,* (2.2) *is satisfied but* (2.1) *and* (2.9) *fail.*

EXAMPLE 2.4. *Consider an example of network* (1.4) *with* $\tau \ge 0$ *described by*

$$\frac{du(t)}{dt} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} u(t) + \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \sin(\frac{2}{\sqrt{3}} u_1(t-\tau)) \\ \sin(\frac{2}{\sqrt{3}} u_2(t-\tau)) \end{pmatrix} + \begin{pmatrix} J_1 \\ J_2 \end{pmatrix}. \tag{2.17}$$

Functions $g_j(v) = \sin(\frac{2v}{\sqrt{3}})$ *satisfy* (A_1) *and* (A_2) *with* $L_j = \frac{2}{\sqrt{3}}$ *for* $j = 1,2$. *Then,* $\alpha = \frac{2}{\sqrt{3}} > 1$ *and* $\beta = 1 + \frac{4}{3} > 2$. *Thus, neither* (2.1) *nor* (2.2) *is satisfied. But* $\|A\|_2 = \frac{1}{\sqrt{2}}$, *and thus* (2.9) *gives* $\gamma = \sqrt{\frac{2}{3}} < 1$. *Therefore, by Theorem* 2.4, *system* (2.17) *has a globally asymptotically stable equilibrium.*

3. Criteria depending on delays. The conditions obtained in the previous section are all independent of delay. In this section, we establish some criteria for the global attractivity that depend on the magnitude of the delay(s).

As mentioned previously, there has been a lot of work on the global convergence of Hopfield-type models without delay (see, for example, Atiya [1], Forti [8], Hirsch [12], Kelly [15], Michel, Farrell, and Porod [18], and Matsuoka [20]). Since global attractivity implies the existence of a unique equilibrium, and the incorporation of the delay into the model does not change the structure of equilibria, sufficient conditions for the global attractivity of (1.3) established in the above-mentioned works guarantee the existence of a unique equilibrium for the delayed model (1.4) as well as (1.5). In the remainder of this section, we will employ the theory of monotone dynamical systems to prove that under some additional conditions, all solutions of (1.5) converge to this equilibrium provided that some delays are sufficiently small.

Since the above mentioned work on the model without delay have all assumed differentiability and monotonicity of the activation functions, the discussion of this section will be under (A_2) and the following hypothesis

(H_1^*) For each $j \in \{1, 2, \dots, n\}$, $g_j \in C^1(\mathbb{R})$ and $0 < g_j'(x) \le sup_{x\in R} g_j'(x) < +\infty$ for $x \in R$.

Note that (H_1^*) and (A_2) are similar to (H_1) and (H_2), respectively, but (H_1^*) does not require $g_j'(x)$ to attain its supremum at 0, and (A_2) means that $g_j(x)$ saturates, but not necessarily at ± 1.

For convenience of statement, we let $\beta_j = \sup_{x\in R} g_j'(x)$, $W = (w_{ij})$, and $Q = (q_{ij})$, where

$$w_{ij} = a_{ij} \frac{\beta_i}{\sqrt{b_i b_j}} \quad \text{and} \quad q_{ij} = a_{ij} \frac{\beta_i}{b_i}$$

for $i, j = 1, 2, \dots, n$. Then, we have the following results (see Matsuoka [20, (17), (20), (8)]), in terms of matrix measures (see, e.g., Coppel [7, p. 41]).

THEOREM 3.1. *Assume* (H_1^*) *and* (A_2) *hold. If one of the following conditions is satisfied, then for every J, system* (1.3) *has a unique equilibrium that is globally asymptotically stable:*

$$(i) \quad \mu_2(W) \triangleq \lambda_{max} \left\{ \frac{W + W^T}{2} \right\} < 1,$$

$$(ii) \quad \mu_1\left(\frac{W + W^T}{2}\right) \triangleq \max_{1\le i\le n} \left(w_{ii} + \sum_{j\ne i}^{n} \frac{|w_{ij} + w_{ji}|}{2} \right) < 1,$$

$$(iii) \quad \mu_1(Q) \triangleq \max_{1\le i\le n} \left(q_{ii} + \sum_{j\ne i}^{n} |q_{ji}| \right) < 1.$$

The outline of the proof of the above theorem is as follows. First, by the same argument as in the proof of Theorem 2.2, we know that for every input J, (1.3) has an equilibrium u^*. Then, by using a Liapunov function similar to that used in Matsuoka [20, (16)], namely,

$$V(x) = \sum_{i=1}^{n} \beta_i \int_0^{x_i} [g_i(s + u_i^*) - g_i(u_i^*)]\, ds,$$

where $x = (x_1, x_2, \dots, x_n)^T$ and $x = u - u^*$, we can arrive at the conclusion of Theorem 3.1 (i). Theorem 3.1 (ii) is just a corollary of Theorem 3.1 (i) because every eigenvalue of a symmetric matrix is not greater than any measure of the matrix. Theorem 3.1 (iii) can be proved by employing another Liapunov function, namely,

$$V(x) = \sum_{i=1}^{n} \frac{\beta_i}{b_i} |x_i|,$$

where x is as above.

The above conditions in Theorem 3.1 (ii) and (iii) are easily computable, and they have the merit that negative diagonal entries of the connection matrix A contribute to the global convergence of the system (1.3) through the w_{ii} and q_{ii} terms.

From now on, we will consider (1.5) with $\tau = \max_{1\le i,j\le n} \tau_{ij}$. In order to apply the theory of monotone dynamical systems, we need some definitions. Let the partial order $\le$ on R^n be the usual componentwise ordering. The partial order $\le$ on $C = C([-\tau,0],R^n)$ will mean that $\phi(\theta) \le \psi(\theta)$ for $\theta \in [-\tau,0]$. The inequality $x < y$ $(x \ll y)$ between two vectors in R^n will mean that $x \le y$ and $x_i < y_i$ for some (all) i. The inequality $\phi < \psi$ for functions in C will mean that $\phi \le \psi$ and $\phi \ne \psi$ while $\phi \ll \psi$ will mean that $\phi(\theta) \ll \psi(\theta)$ for all $\theta \in [-\tau,0]$. The ordering assigned to C as above is called the standard ordering for C.

Now, when we come to the delayed system (1.5), the negativity of diagonal entries and $\tau_{ii} > 0$ would make (1.5) fail to be monotone in the sense of Smith [23] under the standard ordering in C. Motivated by Smith and Thieme [24], we equip C with another nonstandard ordering in C and find conditions under which the semiflows generated by the solutions of (1.5) are strongly order preserving in terms of this nonstandard ordering in C.

Let D be an $n \times n$ essentially nonnegative matrix, that is, $D + \lambda I$ is entrywise nonnegative for all sufficiently large λ. Define

$$K_D = \left\{\psi \in C : \psi \ge 0 \text{ and } e^{-tD}\psi(t) \ge e^{-sD}\psi(s), \quad -\tau \le s \le t \le 0\right\}.$$

It is easy to see that K_D is a cone in the space C, that is, K_D is closed in C, closed under addition and under scalar multiplication by nonnegative scalars and $K_D \bigcap (-K_D) = \emptyset$. Moreover, K_D is a normal cone, which means that every order interval is a bounded set in C. Now as a cone in C, K_D induces a partial order on C, denoted by $\le_D$, in the usual way, namely, $\phi \le_D \psi$ if and only if $\psi - \phi \in K_D$. We write $\phi <_D \psi$ to indicate that $\phi \le_D \psi$ and $\phi \ne \psi$.

The following theorem is from Smith and Thieme [24, Theorem 3.5 and Lemma 3.7].

THEOREM 3.2. *Consider the functional differential equation*

$$\frac{dx(t)}{dt} = f(x_t), \tag{3.1}$$

where $f \in C^1(C,R^n)$. *Assume that the following conditions hold:*

(SM'_D) *For every* $\phi \in C$ *and every* $\psi \in K_D$ *with* $\psi \gg 0$,

$$df(\phi)\psi - D\psi(0) \gg 0.$$

(I'_D) *If* $\phi \in C$, $\psi \in K_D$ *and* P *is a (nonempty) proper subset of* $\{1,2,\dots,n\}$ *such that* $\psi_p \gg 0$ *for* $p \in P$ *and* $\psi_j(0) = 0$ *for* $j \notin P$, *then for some* $k \notin P$

$$(df(\phi)\psi)_k > 0.$$

Then, the semiflow Φ *generated by the solutions of* (3.1) *is strongly order preserving on* C *in terms of the ordering* $\le_D$.

Now, we take $D = \mathrm{diag}(d_1,\dots,d_n)$ with

$$d_i = -b_i - r_i, \quad i = 1,2,\dots,n,$$

where $r_i > 0$, $i = 1, 2, \dots, n$ are constants to be determined later. Then, D is essentially nonnegative. Recall that for (1.5), we have

$$f_i(\phi) = -b_i\phi_i(0) + \sum_{j=1}^{n} a_{ij} g_j(\phi_j(-\tau_{ij})) + J_i, \quad i = 1, 2, \dots, n. \tag{3.2}$$

Let $\phi \in C$ and $\psi \in K_D$. Then

$$\begin{aligned}
&(df(\phi)\psi)_i - (D\psi(0))_i \\
&= -b_i\psi_i(0) + \sum_{j=1}^{n} a_{ij} g_j'(\phi_j(-\tau_{ij}))\psi_j(-\tau_{ij}) + (b_i + r_i)\psi_i(0) \\
&= r_i\psi_i(0) + a_{ii} g_i'(\phi_i(-\tau_{ii}))\psi_i(-\tau_{ii}) + \sum_{j \neq i}^{n} a_{ij} g_j'(\phi_j(-\tau_{ij}))\psi_j(-\tau_{ij}).
\end{aligned}$$

But $\psi \in K_D$ implies that $\psi(0) \geq e^{-sD}\psi(s)$ for all $s \in [-\tau, 0]$, and hence

$$\psi_i(0) \geq e^{-\tau_{ii}(-d_i)}\psi_i(-\tau_{ii}) = e^{-\tau_{ii}(b_i + r_i)}\psi_i(-\tau_{ii}).$$

Therefore, for $\phi \in C$ and $\psi \in K_D$ we have

$$\begin{aligned}
&(df(\phi)\psi)_i - (D\psi(0))_i \\
&\geq \left(r_i e^{-\tau_{ii}(b_i + r_i)} + a_{ii} g_i'(\phi_i(-\tau_{ii}))\right)\psi_i(-\tau_{ii}) + \sum_{j \neq i}^{n} a_{ij} g_j'(\phi_j(-\tau_{ij}))\psi_j(-\tau_{ij}).
\end{aligned}$$

Since A is irreducible, it follows that (SM_D') holds for f given by (3.2) provided $a_{ij} \geq 0$ for $i \neq j$ and

$$r_i e^{-\tau_{ii}(b_i + r_i)} \geq |a_{ii}|\beta_i \quad \text{for all} \quad i \quad \text{with} \quad a_{ii} < 0. \tag{3.3}$$

But (3.3) is satisfied if and only if

$$\tau_{ii} \leq \frac{\ln \frac{r_i}{|a_{ii}|\beta_i}}{b_i + r_i} \quad \text{for all} \quad i \quad \text{with} \quad a_{ii} < 0. \tag{3.4}$$

Now, if we take $r_i = e|a_{ii}|\beta_i$, then (3.4) becomes

$$\tau_{ii} \leq \frac{1}{b_i + e|a_{ii}|\beta_i} \quad \text{for all} \quad i \quad \text{with} \quad a_{ii} < 0. \tag{3.5}$$

Hence, (SM_D') holds provided $a_{ij} \geq 0$ for $i \neq j$, and those τ_{ii} corresponding to negative a_{ii} are sufficiently small (e.g., as estimated by (3.5)).

We next verify that (I_D') also holds for f given by (3.2). To this end, let $\phi \in C$, $\psi \in K_D$, and P be a proper subset of $\{1, 2, \dots, n\}$ such that $\psi_p \gg 0$ for $p \in P$ and $\psi_i(0) = 0$ for $i \notin P$. Then, for each $i \notin P$, $\psi_i(-\tau_{ii}) = 0$, since $\psi_i(-\tau_{ii}) \leq e^{d_i\tau_{ii}}\psi_i(0)$. Since A is irreducible, there is some $i \notin P$ such that

$$\begin{aligned}
(df(\phi)\psi)_i &= -b_i\psi_i(0) + \sum_{j=1}^{n} a_{ij} g_j'(\phi_j(-\tau_{ij}))\psi_j(-\tau_{ij}) \\
&= \sum_{j \neq i}^{n} a_{ij} g_j'(\phi_j(-\tau_{ij}))\psi_j(-\tau_{ij}) \\
&= \sum_{j \in P}^{n} a_{ij} g_j'(\phi_j(-\tau_{ij}))\psi_j(-\tau_{ij})
\end{aligned}$$

is positive provided $a_{ij} \geq 0$ for $j \neq i$. Thus (I'_D) is also satisfied if $a_{ij} \geq 0$ for $i \neq j$.

Combining the above arguments with 3.2 gives the following.

PROPOSITION 3.3. *Assume that $a_{ij} \geq 0$ for $i \neq j$ and that* (3.5) *is satisfied. Then, the semiflow Φ generated by the solutions of* (1.5) *is strongly order preserving on C in terms of the ordering $\leq_D$.*

Now we are in the position to state and prove our theorem for the global convergence of model (1.5) with A irreducible.

THEOREM 3.4. *Assume that $a_{ij} \geq 0$ for $i \neq j$, and the diagonal delays τ_{ii} corresponding to negative a_{ii} are sufficiently small (say, satisfy* (3.5)*). Then, under the conditions of Theorem* 3.1*, system* (1.5) *has a unique, globally attractive equilibrium for every J.*

Proof. As mentioned at the beginning of this section, under the conditions of Theorem 3.1, system (1.5) has a unique equilibrium u^*. It is easy to see that solutions of (1.5) are bounded due to (A_2). Combining the above facts with Proposition 3.3 and the global convergence theorem in Smith [23, p. 18, Theorem 3.1], we know that u^* is globally attractive. ∎

The above result shows that the off-diagonal delays τ_{ij} have no effect on the order preserving property of the semiflow generated by the solutions of (1.5), provided that the off-diagonal connection weights are *positive*. Furthermore, from the argument before (3.3), we can see that if a_{ii} is *nonnegative*, then the corresponding delay τ_{ii} of any magnitude is also "harmless" for this purpose. For negative a_{ii}, (3.5) gives an estimate for the "smallness" of the corresponding τ_{ii}.

REMARK 3.1. *If every cycle of length ≥ 2 in the digraph corresponding to matrix A is positive, then there exists a signature matrix S so that SAS has all off-diagonal entries ≥ 0 (see, e.g., Bélair, Campbell, and van den Driessche* [3]*), and thus Theorem* 3.4 *is applicable by the transformation $u \longrightarrow Su$ to* (1.5)*.*

The following example illustrates Theorem 3.3.

EXAMPLE 3.1. *Consider an example of network* (1.4) *described by*

$$\frac{du(t)}{dt} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} u(t) + \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \tanh\, 2u_1(t-\tau) \\ tanh\, 2u_2(t-\tau) \end{pmatrix}.$$

Then none of (2.1), (2.2), (2.9) *is satisfied, but all the conditions in Theorem* 3.1 *are satisfied. Thus, by Theorem* 3.4*, $u^* = 0$ is globally attractive for $0 \leq \tau < \frac{1}{1+e}$. Linear stability analysis (see, e.g., Bélair, Campbell, and van den Driessche* [3, Section 2]*) shows that $u^* = 0$ is locally stable for this range of τ. Thus, $u^* = 0$ is in fact globally asymptotically stable for $0 \leq \tau < \frac{1}{1+e}$.*

In conclusion, we note that the theory of monotone dynamical systems was also used by Olien and Bélair [22, section 4] in a two-neuron Hopfield network with specific activation functions that satisfy hypotheses (H_1) and (H_2). For this two-neuron model with $a_{ii} \geq 0$, assuming that $a_{12}a_{21} > 0$, the system is cooperative and irreducible; thus there is generic convergence to a stable equilibrium (Olien and Bélair [22, Theorem 1 and Corollary 4]). Our Theorem 3.4 is for global convergence and gives additional conditions on the matrix A for the unique equilibrium to be globally attractive (for all delays if $a_{ii} \geq 0$, and for small τ_{ii} if $a_{ii} < 0$).

Acknowledgments. We thank the referees for their suggestions that led to an improvement of some results.

REFERENCES

[1] A. F. Atiya, *Learning on a general network,* in Neural Information Processing Systems, D. Z. Anderson, ed., American Institute Physics, New York, 1988, pp. 22–30.

[2] J. Bélair, *Stability in a model of a delayed neural network,* J. Dynam. Differential Equations, 5 (1993), pp. 607–623.

[3] J. Bélair, S. A. Campbell, and P. van den Driessche, *Frustration, stability and delay-induced oscillations in a neural network model,* SIAM J. Appl. Math., 56 (1996), pp. 245–255.

[4] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences,* Classics Appl. Math., SIAM, Philadephia, PA, 1994.

[5] Y. J. Cao and Q. H. Wu, *A note on stability of analog neural networks with time delays,* IEEE Trans. Neural Networks, 7 (1996), pp. 1533–1535.

[6] M. A. Cohen and S. Grossberg, *Absolute stability and global pattern formation and parallel memory storage by competitive neural networks,* IEEE Trans. Systems Man Cybernet., SMC13 (1983), pp. 815–821.

[7] W. A. Coppel, *Stability and Asymptotic Behavior of Differential Equations,* D. C. Heath and Company, Boston, 1965.

[8] M. Forti, *On global asymptotic stability of a class of nonlinear systems arising in neural network theory,* J. Differential Equations, 113 (1994), pp. 246–264.

[9] M. Forti, S. Manetti, and M. Marini, *A condition for global convergence of a class of symmetric neural networks,* IEEE Trans. Circuits Systems, 39 (1992), pp. 480–483.

[10] K. Gopalsamy and X. He, *Stability in asymmetric Hopfield nets with transmission delays,* Phys. D, 76 (1994), pp. 344–358

[11] J. Haddock and J. Terjéki, *Liapunov-Razumikhin functions and an invariance principle for functional differential equations,* J. Differential Equations, 48 (1983), pp. 95–122.

[12] M. W. Hirsch, *Convergent activation dynamics in continuous time networks,* Neural Networks, 2 (1989), pp. 331–349.

[13] J. J. Hopfield, *Neurons with graded response have collective computational properties like those of two-stage neurons,* Proc. Nat. Acad. Sci. U.S.A., 81 (1984), pp. 3088–3092.

[14] J. Jang, S. Lee, and S. Shin, *An optimization network for matrix inversion,* in Neural Information Processing Systems, D. Z. Anderson, ed., American Institute of Physics, New York, 1988, pp. 397–401.

[15] D. G. Kelly, *Stability in contractive nonlinear neural networks,* IEEE Trans. Biomed. Engrg., 37 (1990), pp. 231–242.

[16] Y. Kuang, *Delay Differential Equations with Applications in Population Dynamics,* Academic Press, Boston, 1993.

[17] J. H. Li, A. N. Michel, and W. Porod, *Qualitative analysis and synthesis of a class of neural networks,* IEEE Trans. Circuits Systems, 35 (1998), pp. 976–986.

[18] A. N. Michel, J. A. Farrell, and W. Porod, *Qualitative analysis of neural networks,* IEEE Trans. Circuits Systems, 36 (1989), pp. 229–243.

[19] C. M. Marcus and R. M. Westervelt, *Stability of analog neural networks with delay,* Phys. Rev. A, 39 (1989), pp. 347–359.

[20] K. Matsuoka, *Stability conditions for nonlinear continuous neural networks with asymmetric connection weights,* Neural Networks, 5 (1992), pp. 495–500.

[21] M. Morita, *Associative memory with non-monotone dynamics,* Neural Networks 6 (1993), pp. 115–126.

[22] L. Olien and J. Bélair, *Bifurcations, stability, and monotonicity properties of a delayed neural network model,* Phys. D, 102 (1997), pp. 349–363.

[23] H. Smith, *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems,* Math. Surveys Monographs 41, AMS, Providence, RI, 1995.

[24] H. Smith and H. Thieme, *Strongly order preserving semiflows generated by functional differential equations,* J. Differential Equations, 93 (1991), pp. 332–363.

[25] D. Tank and J. J. Hopfield, *Simple neural optimization networks: An A/D converter, signal decision circuit and a linear programming circuit,* IEEE Trans. Circuits Systems, 33 (1986), pp. 533–541.

[26] J. Wu, *Symmetric functional-differential equations and neural networks with memory,* Trans. Amer. Math. Soc., to appear.

[27] J. Wu and X. Zou, *Patterns of sustained oscillations in neural networks with delayed interactions,* Appl. Math. Comput., 73 (1995), pp. 55–75.

[28] S. Yoshizawa, M. Morita, and S. Amari, *Capacity of associate memory using a nonmonotonic neuron model,* Neural Networks, 6 (1993), pp. 167–176.