Dynamical Behaviors of Delayed Neural Network Systems with Discontinuous Activation Functions

Wenlian Lu
02108041@fudan.edu, wenlian@mis.mpg.de
Tianping Chen
tchen@fudan.edu.cn
Laboratory of Nonlinear Mathematics Science, Institute of Mathematics, Fudan University, Shanghai, 200433, People’s Republic of China

In this letter, without assuming the boundedness of the activation functions, we discuss the dynamics of a class of delayed neural networks with discontinuous activation functions. A relaxed set of sufficient conditions is derived, guaranteeing the existence, uniqueness, and global stability of the equilibrium point. Convergence behaviors for both state and output are discussed. The constraints imposed on the feedback matrix are independent of the delay parameter and can be validated by the linear matrix inequality technique. We also prove that the solution of delayed neural networks with discontinuous activation functions can be regarded as a limit of the solutions of delayed neural networks with high-slope continuous activation functions.

1 Introduction

It is well known that recurrently connected neural networks (RCNNs), proposed by Cohen and Grossberg (1983) and Hopfield (1984; Hopfield & Tank, 1986), have been extensively studied in both theory and applications. They have been successfully applied in signal processing, pattern recognition, and associative memories, especially in static image treatment. Such applications heavily rely on dynamical behaviors of the neural networks. Therefore, analysis of dynamical behaviors is a necessary step for the practical design of neural networks.

In hardware implementation, time delays inevitably occur due to the finite switching speed of the amplifiers and communication time. What is more, to process moving images, one must introduce time delays in the signals transmitted among the cells (see Civalleri, Gilli, & Pabdolfi, 1993). Neural networks with time delay have much more complicated dynamics due to the incorporation of delays. The model concerning delay is described

Neural Computation 18, 683–708 (2006) © 2006 Massachusetts Institute of Technology
as follows:

\[
\frac{d x_i(t)}{dt} = -d_i x_i(t) + \sum_{j=1}^{n} a_{ij} g_j(x_j(t)) + \sum_{j=1}^{n} b_{ij} g_j(x_j(t - \tau)) + I_i,
\]

\[i = 1, 2, \ldots, n,\] (1.1)

where \(n\) is the number of units in a neural network, \(x_i(t)\) is the state of the \(i\)th unit at time \(t\), and \(g_j(x_j(t))\) denotes the output of \(j\)th unit at time \(t\). \(a_{ij}\) denotes the strength of the \(j\)th unit on the \(i\)th unit at time \(t\), and \(b_{ij}\) denotes the strength of the \(j\)th unit on the \(i\)th unit at time \(t - \tau\). \(I_i\) denotes the input to the \(i\)th unit, \(\tau\) corresponds to the transmission delay and is a nonnegative constant, and \(d_i\) represents the positive rate with which the \(i\)th unit will reset its potential to the resting state in isolation when disconnected from the network and the external inputs \(I_i\).

System (1.1) can be rewritten as

\[
\frac{d x(t)}{dt} = -D x(t) + A g(x(t)) + B g(x(t - \tau)) + I,
\]

where \(x = (x_1, x_2, \ldots, x_n)^T\), \(g(x) = (g_1(x_1), g_2(x_2), \ldots, g_n(x_n))^T\), \(I = (I_1, I_2, \ldots, I_n)^T\), \(T\) denotes transpose, \(D = \text{diag}\{d_1, d_2, \ldots, d_n\}\), \(A = \{a_{ij}\}\) is the feedback matrix, and \(B = \{b_{ij}\}\) is the delayed feedback matrix.

Some useful results on the stability analysis of delayed neural networks (DNNs) have already been obtained. Readers can refer to Chen (2001), Zeng, Weng, and Liao (2003), Lu, Rong, and Chen (2003), Joy (2000), and many others. In particular, Lu et al. (2003) and Joy (2000) provided some effective criteria based on LMI.

All discussion in these articles is based on the assumption that the activation functions are continuous and even Lipshitzean. As Forti and Nistri (2003) pointed out, a brief review of some common neural network models reveals that neural networks with discontinuous activation functions are important and frequently arise in practice. For example, consider the classical Hopfield neural networks with graded response neurons (see Hopfield 1984). The standard assumption is that the activations are used in high-gain limit where they closely approach discontinuous and comparator functions. As shown by Hopfield (1984; Hopfield & Tank, 1986), the high-gain hypothesis is crucial to make negligible the connection to the neural network energy function of the term depending on neuron self-inhibitions, and to favor binary output formation—for example, a hard comparator function sign(s).

A conceptually analogous model based on hard comparators are discrete-time neural networks discussed by Harrer, Nossek, and Stelzl (1992). Another important example concerns the class of neural networks introduced by Kennedy and Chua (1988) to solve linear and nonlinear
programming problems. Those networks exploit constrained neurons with diode-like input-output activations. Again, in order to guarantee satisfaction of the constraints, the diodes are required to possess a very high slope in the conducting region, that is, they should approximate the discontinuous characteristic of an ideal diode (see Chua, Desoer, & Kuh, 1987). And when dealing with dynamical systems possessing high-slope nonlinear elements, it is often advantageous to model them with a system of differential equations with a discontinuous right-hand side, rather than studying the case where the slope is high but of finite value (see Utkin, 1977).

Forti and Nistri (2003) applied the concepts and results of differential equations with discontinuous right-hand side introduced by Filippov (1967) to investigate the global convergence of neural networks with discontinuous neuron activations. Furthermore, they also discussed various types of convergence behaviors of global stability, such as convergence in finite time and convergence in measure. Useful sufficient conditions were obtained to ensure global convergence. But the discontinuous activations are assumed bounded. Lu and Chen (2005) studied global stability of a more general neural network model: Cohen-Grossberg neural networks. In this letter, the discontinuous activation functions were not assumed bounded. However, in both articles, the models do not involve time delays.

In this letter, we introduce a new concept of solution for delayed neural networks with discontinuous activation functions. Without assuming the boundedness and the continuity of the neuron activations, we present sufficient conditions for the global stability of neural networks with time delay based on linear matrix inequality, and we discuss their convergence. Moreover, we explore the importance of the concept of the solution presented in this letter.

2 Preliminaries

In this section, we present some definitions used in this letter.

**Definition 1.** (See Preliminaries in Forti & Nistri, 2003.) Suppose $E \subset \mathbb{R}^n$. Then $x \mapsto F(x)$ is called a set-value map from $E \leftrightarrow \mathbb{R}^n$, if to each point $x$ of a set $E \subset \mathbb{R}^n$, there corresponds a nonempty set $F(x) \subset \mathbb{R}^n$. A set-value map $F$ with nonempty values is said to be upper semicontinuous at $x_0 \in E$, if for any open set $N$ containing $F(x_0)$, there exists a neighborhood $M$ of $x_0$ such that $F(M) \subset N$. $F(x)$ is said to have a closed (convex, compact) image if for each $x \in E$ $F(x)$ is closed (convex, compact). $\text{Graph}(F(E)) = \{(x, y)|x \in E, \text{ and } y \in F(x)\}$, where $E$ is subset of $\mathbb{R}^n$.

More details about set value maps can be found in Aubin and Frankowska (1990).
**Definition 2.** Class $\bar{G}$ of functions: Let $g(x) = (g_1(x_1), g_2(x_2), \ldots, g_n(x_n))^T$. We call $g(x) \in \bar{G}$, if for all $i = 1, 2, \ldots, n$, $g_i(\cdot)$ satisfies:

1. $g_i(\cdot)$ is nondecreasing and continuous, except on a countable set of isolated points $\{\rho_i^k\}$, where the right and left limits $g_i^+(\rho_i^k)$ and $g_i^-(\rho_i^k)$, satisfy $g_i^+(\rho_i^k) > g_i^-(\rho_i^k)$. Moreover, in every compact set of $R$, $g_i(\cdot)$ has only finite discontinuous points.

2. Denote the set of points $\{\rho_i^k : i = 1, \ldots, n; k = \ldots, -2, -1, 0, 1, 2, \ldots\}$ of discontinuity in the following way: for any $i$, $\rho_i^{k+1} > \rho_i^k$, and there exist constants $G_{i,k} > 0$, $i = 1, \ldots, n; k = \ldots, -2, -1, 0, 1, 2, \ldots$, such that

$$0 \leq \frac{g_i(\xi) - g_i(\zeta)}{\xi - \zeta} \leq G_{i,k} \text{ for all } \xi \neq \zeta \text{ and } \xi, \zeta \in (\rho_i^k, \rho_i^{k+1}).$$

**Remark 1.** Forti and Nistri (2003) assumed that discontinuous activations satisfy the first condition of definition 2. We impose some local Lipschitz continuity on each interval that does not contain any points of discontinuity. Furthermore, we do not assume the activation functions are bounded, which is required by Forti and Nistri (2003).

Note that the function $g_i(\cdot)$ is undefined at the points, where $g_i(\cdot)$ is discontinuous. Such discontinuous functions $\bar{G}$ include a number of neuron activations of interest for applications—for example, the standard hard comparator function $\text{sign}(\cdot)$:

$$\text{sign}(s) = \begin{cases} 1 & s > 0 \\ -1 & s < 0 \end{cases}. \quad (2.1)$$

It is clear that if $g(\cdot) \in \bar{G}$, the right-hand side of equation 1.2 is discontinuous. Therefore, we have to explain the meaning of a solution of the Cauchy problem associated with equation 1.2 before further investigation. Filippov (1967) developed a concept of a solution for differential equations with a discontinuous right-hand side, which was used by Forti et al. (2003) and Lu and Chen (2005) to investigate the stability of neural networks with discontinuous activation functions. In the following, we apply this framework in discussing delayed neural networks with discontinuous activation functions.

Now we introduce the concept of Filippov solution. Consider the following system,

$$\frac{dx}{dt} = f(x), \quad (2.2)$$

where $f(\cdot)$ is not continuous.
Definition 3. A set-value map is defined as

\[ \phi(x) = \bigcap_{\delta > 0} \bigcap_{\mu(N) = 0} K[f(B(x, \delta) - N)] \] (2.3)

where \( K(E) \) is the closure of the convex hull of set \( E \), \( B(x, \delta) = \{ y : \| y - x \| \leq \delta \} \), and \( \mu(N) \) is Lebesgue measure of set \( N \). A solution of the Cauchy problem for equation 2.2 with initial condition \( x(0) = x_0 \) is an absolutely continuous function \( x(t), t \in [0, T] \), which satisfies \( x(0) = x_0 \), and differential inclusion:

\[ \frac{dx}{dt}(t) \in \phi(x), \quad a.e. \ t \in [0, T]. \] (2.4)

The concept of the solution in the sense of Filippov is useful in engineering applications because the Filippov solution is a limit of solutions of ordinary differential equations (ODEs) with the continuous right-hand side. Thus, we can model a system that is a near discontinuous system and expect that the Filippov trajectory of the discontinuous system will be close to the real trajectories. This approach is important in many applications, such as variable structure control and nonsmooth analysis (see Utkin, 1997; Aubin & Cellina, 1984; Paden & Sastry, 1987).

Moreover, Haddad (1981), Aubin (1991), and Aubin and Cellina (1984) gave a functional differential inclusion with memory as follows:

\[ \frac{dx}{dt}(t) \in F(t, A(t)x), \] (2.5)

where \( F : \mathbb{R} \times C([−\tau, 0], \mathbb{R}^n) \mapsto \mathbb{R}^n \) is a given set-value map and

\[ [A(t)x](\theta) = x_1(\theta) = x(t + \theta). \] (2.6)

Now we denote \( K[g(x)] = (K[g_1(x_1)], K[g_2(x_2)], \ldots, K[g_n(x_n)])^T \) where \( K[g_i(x_i)] = [g_i^−(x_i), g_i^+(x_i)] \). We extend the concept of Filippov solution to the delayed differential equations 1.2 as follows:

\[ \frac{dx}{dt}(t) \in -Dx(t) + AK[g(x(t))] + BK[g(x(t - \tau))] + I, \quad \text{for almost all } t. \]

Equivalently,

\[ \frac{dx}{dt}(t) = -Dx(t) + A\alpha(t) + B\beta(t - \tau) + I, \]

where \( \alpha(t) \in K[g(x(t))] \) and \( \beta(t) \in K[g(x(t))] \).
In the sequel, we assume $\alpha(t) = \beta(t)$. It is reasonable. The output of the system should be consistent over time. Therefore, we will consider the following delayed system,

$$\frac{dx}{dt}(t) = -Dx(t) + A\alpha(t) + B\alpha(t - \tau) + I, \quad \text{for almost all } t,$$

where output $\alpha(t)$ is measurable, and

$$\alpha(t) \in K[g(x(t))], \quad \text{for almost all } t.$$

**Definition 4.** A solution of the Cauchy problem for the delayed system 2.7 with initial condition $\phi(\theta) \in C([-\tau,0], \mathbb{R}^n)$ is an absolutely continuous function $x(t)$ on $t \in [0, T]$, such that $x(\theta) = \phi(\theta)$ for $\theta \in [-\tau, 0]$, and

$$\frac{dx}{dt}(t) = -Dx(t) + A\alpha(t) + B\alpha(t - \tau) + I \quad \text{a.e. } t \in [0, T].$$

where $\alpha(t)$ is measurable and for almost $t \in [0, T], \alpha(t) \in K[g(x(t))]$.

**Remark 2.** Concerning the solution of the ODEs or functional differential equations (FDEs) with a discontinuous right-hand side, there are various definitions, such as Euler solutions and generalized sampling solutions. Among them, Carathéodory and a weak solution set are widely studied. Liz and Pouso (2002) gave some general results for the existence of the solutions of first-order discontinuous FDEs subject to nonlinear boundary conditions. The Carathéodory solution set is defined as follows (here, we compare these solution set in the case without time delay).

Consider the following ODE,

$$\frac{dx(t)}{dt} = f(x(t)), \quad t \in [0, T].$$

with initial condition $x(0) = x_0$.

An absolutely continuous function $\xi(t)$ is said to be a Carathéodory solution if

$$\frac{dx(t)}{dt} = f(x(t)), \quad \text{a.e. } t \in [0, T]$$

$$x(0) = x_0.$$
A function $\zeta(t) \in L^1([0, T])$ is said to be a weak solution in $L^1([0, T])$ if for each $p(t) \in C^\infty_0([0, T])$, there holds
\[
\int_0^T \frac{dp(t)}{dt} \zeta(t) dt = - \int_0^T f(\zeta(t)) p(t) dt. \tag{2.11}
\]

It is clear that the Carathéodory solution set is surely a subset of the Filippov solution set if it involves a discontinuous right-hand side. On the other hand, the weak solution set might contain a discontinuous solution. But if we focus on the absolutely continuous weak solutions, this solution set is equivalent to the Carathéodory solution set. Both are subsets of the Filippov solution set. Spraker and Biles (1996) pointed out that in a one-dimensional case, the Carathéodory solution set equals the Filippov solution set if and only if $\{x, 0 \in \phi(x)\} = \{x, f(x) = 0\}$, where $\phi(\cdot)$ is defined as in definition 3. Otherwise, the two solution sets are different. For example, consider the following one-dimensional ODE,
\[
\frac{dx(t)}{dt} = -x - q(x), \tag{2.12}
\]
where
\[
q(x) = \begin{cases} 
1 & \text{if } \rho > 0 \\
-1 & \text{if } \rho < 0 \\
\frac{1}{2} & \text{if } \rho = 0 
\end{cases}. \tag{2.13}
\]

The initial condition is $x(0) = 0$. It is easy to see that $\{x, 0 \in \phi(x)\} = \{0\}$ and $\{x, f(x) = 0\} = \emptyset$. It can be seen that system (2.12) has neither a Carathéodory solution nor a weak absolutely continuous solution. On the other hand, equation 2.12 has the Filippov solution $x(t) = 0$.

**Definition 5.** (Equilibrium) $x^*$ is said to be an equilibrium of system 2.8, if there exists $\alpha^* \in K[g(x^*)]$, such that
\[
0 = -Dx^* + A\alpha^* + B\alpha^* + I.
\]

**Definition 6.** If $x^*$ is an equilibrium of the system 2.8, $x^*$ is said to be globally asymptotically stable if for any solution $x(t)$ of equation 2.8 whose existence interval is $[0, +\infty)$, we have
\[
\lim_{t \to \infty} x(t) = x^*.
\]
Moreover, $x(t)$ is said to be exponentially asymptotically stable globally if there exist constants $\epsilon > 0$ and $M > 0$, such that

$$\|x(t) - x^*\| \leq Me^{-\epsilon t}.$$ 

The letter is organized as follows. In section 3, we discuss the existence of the equilibrium point and the solution for system 2.8. The stability of the equilibrium point and the convergence of the output of the delayed neural networks are studied in section 4. Some numerical examples are presented in section 5. We conclude this letter in section 6.

3 Existence of an Equilibrium Point and Solution

In this section, we prove that under some conditions, system 2.8 has an equilibrium point and a solution in the infinite time interval $[0, \infty)$.

3.1 Existence of an Equilibrium Point. First, we investigate the existence of an equilibrium point. For this purpose, consider the following differential inclusion,

$$\frac{dy}{dt} \in -Dy(t) + TK[g(y(t))] + I.$$ \hspace{1cm} (3.1)

where $y(t) = (y_1(t), y_2(t), \ldots, y_n(t))^T$, $D$, $K[g(\cdot)]$, and $I$ are the same as those in system 2.8.

We need the following result.

**Theorem A.** (Lu and Chen, 2005, theorem 2). Suppose $g \in \tilde{G}$. If there exists a positive definite diagonal matrix $P$ such that $-PT - T^T P$ is positive definite, then there exists an equilibrium point of system 3.1; that is, there exist $y^* \in \mathbb{R}^n$ and $\alpha^* \in K[g(y^*)]$, such that

$$0 = -Dy^* + T\alpha^* + I.$$ \hspace{1cm} (3.2)

By theorem A, we can prove theorem 1.

**Theorem 1.** If there exist a positive definite diagonal matrix $P = \text{diag}\{P_1, P_2, \ldots, P_n\}$ and a positive definite symmetric matrix $Q$ such that

$$\begin{bmatrix}
-P A - A^T P - Q - P B \\
- B^T P & Q
\end{bmatrix} > 0,$$

then there exists an equilibrium point of system (see equation 2.8).
Proof. By the Schur complement theorem (see Boyd, Ghaoui, Feron, & Balakrishnan (1994), inequality 3.3 is equivalent to $-PA + \lambda^2 P > PB Q^{-1} B^T P + Q$. By the inequality $[Q^{-\frac{1}{2}} B^T P - Q^{\frac{1}{2}}]^T [Q^{-\frac{1}{2}} B^T P - Q^{\frac{1}{2}}] \geq 0$, $PBQ^{-1}B^TP + Q \geq PB + B^TP$ holds.

Then inequality 3.3 becomes

$$-PA + (A+B)^T P > 0. \tag{3.4}$$

From theorem A, there exists an equilibrium point $x^* \in \mathbb{R}^n$ and $\alpha^* \in K[g(x^*)]$ such that

$$0 = -Dx^* + (A+B)\alpha^* + I, \tag{3.5}$$

which implies that $\alpha^*$ is an equilibrium point of system 2.8. Theorem 1 is proved.

Suppose that $x^* = (x^*_1, x^*_2, \ldots, x^*_n)^T$ is an equilibrium point of system 2.8, that is, there exists $\alpha^* = (a^*_1, a^*_2, \ldots, a^*_n)^T \in K[g(x)]$ such that equation 3.5 satisfies.

Let $u(t) = x(t) - x^*$ be a translation of $x(t)$ and $\gamma(t) = \alpha(t) - \alpha^*$ be a translation of $\alpha(t)$. Then $u(t) = (u_1(t), u_2(t), \ldots, u_n(t))^T$ satisfies

$$\frac{du(t)}{dt} = -Du(t) + Ay(t) + By(t - \tau), \quad \text{for almost } t, \tag{3.6}$$

where $\gamma(t) \in K[g^*(u(t))], g_i^*(s) = g_i(s + x^*_i) - \gamma_i^*, i = 1, 2, \ldots, n$. To simplify, we still use $g_i(s)$ to denote $g_i^*(s)$.

Therefore, instead of equation 2.8, we will investigate

$$\frac{du(t)}{dt} = -Du(t) + Ay(t) + By(t - \tau) \quad \text{for almost } t \tag{3.7}$$

where $\gamma(t) \in K[g(u(t))], g(\cdot) \in \tilde{G}$, and $0 \in K[g_i(0)]$, for all $i = 1, 2, \ldots, n$.

It can be seen that the dynamical behavior of equation 2.8 is equivalent to that of equation 3.7. Namely, if there exists a solution $u(t)$ for equation 3.7, then $x(t) = u(t) + x^*$ must be a solution for equation 2.8; moreover, if all trajectories of equation 3.7 converge to the origin, then the equilibrium $x^*$ must be globally stable for system 2.8, as defined in definition 6. Therefore, instead of equation 2.8, we will investigate the dynamics of system 3.7.

3.2 Viability. In this section, we investigate the viability of system 3.7, that is, there exists at least one solution of system 3.7 on $[0, +\infty)$, which is the prerequisite to study global stability. First, we give following lemma on matrix inequalities:
Lemma 1. If $P = \text{diag}(P_1, P_2, \ldots, P_n)$ with $P_i > 0$, $Q$ is a positive definite symmetric matrix such that

$$Z = \begin{bmatrix}
-PA - A^T P - Q - PB \\
-B^T P \\
Q
\end{bmatrix} > 0, \tag{3.8}
$$

then the following two statements hold true:

1. There are a small, positive constant $\varepsilon < \min_i d_i$, a positive diagonal matrix $\hat{P} = \text{diag}(\hat{P}_1, \hat{P}_2, \ldots, \hat{P}_n)$, and a positive definite symmetric matrix $\hat{Q}$, such that

$$Z_1 = \begin{bmatrix}
-2D + \varepsilon I & A & B \\
A^T & \hat{P}A + A^T \hat{P} + \hat{Q} & \hat{P}B \\
B^T & B^T \hat{P} & -\hat{Q}
\end{bmatrix} \leq 0. \tag{3.9}
$$

2. There are a small, constant $\varepsilon > 0$, a diagonal matrix $\tilde{P} = \text{diag}(\tilde{P}_1, \tilde{P}_2, \ldots, \tilde{P}_n)$ with $\tilde{P}_i > 0$, and a positive definite symmetric matrix $\tilde{Q}$, such that

$$Z_2 = \begin{bmatrix}
-2D & A & B \\
A^T & \tilde{P}A + A^T \tilde{P} + \varepsilon I & \tilde{P}B \\
B^T & B^T \tilde{P} & -\tilde{Q}
\end{bmatrix} \leq 0. \tag{3.10}
$$

Proof. Let $\hat{P} = \alpha P$, $\hat{Q} = \alpha Q$, where $P$ and $Q$ are defined in inequality 3.8, and $\varepsilon$ and $\alpha$ are constants determined later. Then, for any $x, y$, and $z \in \mathbb{R}^n$, we have

$$[x^T, y^T, z^T]Z_1 \begin{bmatrix}
x \\
y \\
z
\end{bmatrix} = -2x^T Dx + \varepsilon x^T x + 2x^T Ay + 2x^T Bz + \alpha y^T (PA + A^T P)y + \alpha y^T Qy + 2\alpha y^T P B z - \alpha z^T Q z + \alpha (e^{\varepsilon t} - 1)y^T Q y$$

$$= -2x^T Dx + \varepsilon x^T x + 2x^T Ay + 2x^T Bz - \alpha [y^T, z^T]Z \times \begin{bmatrix}
y \\
z
\end{bmatrix} + \alpha (e^{\varepsilon t} - 1)y^T Q y$$

$$\leq -x^T Dx + 2x^T Ay - \alpha y^T [\lambda I - (e^{\varepsilon t} - 1)Q] y - x^T (D - \varepsilon I)x + 2x^T Bz - \alpha \lambda z^T z$$

$$= -[D^{1/2} x - D^{-1/2} Ay]^T [D^{1/2} x - D^{-1/2} Ay] + y^T A^T D^{-1} Ay - \alpha y^T [\lambda I - (e^{\varepsilon t} - 1)Q] y$$
\[-[(D - \varepsilon I)^{1/2}x - (D - \varepsilon I)^{-1/2}Bz]^T \left[(D - \varepsilon I)^{1/2}x - (D - \varepsilon I)^{-1/2}Bz\right] + z^T B^T (D - \varepsilon I)^{-1}Bz - \alpha \lambda z^T z \leq 0, \]  
(3.11)

where \( \lambda = \lambda_{\min}(Z) > 0 \). Pick \( \varepsilon \), satisfying \( \varepsilon < \min_i \{d_i\} \) and \( e^{\varepsilon \tau} < \frac{\lambda}{\|Q\|_2} + 1 \), and \( \alpha \) satisfying

\[
\alpha > \max \left\{ \frac{\|A^T D^{-1} A\|_2}{\lambda_{\min}(\lambda I - (e^{\varepsilon \tau} - 1)Q)}, \frac{\|B^T (D - \varepsilon I)^{-1}B\|_2}{\lambda} \right\},
\]

where \( \|X\|_2 = \sqrt{\lambda_{\max}(X^T X)} \) and \( \lambda_{\max}(Z) \) and \( \lambda_{\min}(Z) \) denote the maximum and minimum eigenvalue of the square matrix \( Z \), respectively. Then

\[
\begin{bmatrix} x^T, y^T, z^T \end{bmatrix} Z_1 \begin{bmatrix} x \\ y \\ z \end{bmatrix} \leq 0
\]

holds for any \( x, y, z \in \mathbb{R}^n \), which implies \( Z_1 \leq 0 \). In a similar way, we can prove equation 3.10.

To prove the existence of the solution for system 3.7, we will construct a sequence of functional differential systems and prove that the solutions of these systems converge to a solution of system 3.7. Specifically, consider the following Cauchy problem:

\[
\begin{cases}
\frac{dx}{dt}(t) = -Dx(t) + Ay(t) + B\gamma(t - \tau), & a.e. \ t \in [0, T] \\
a \text{ measurable function } \gamma(t) \in K[g(x(t))], \text{ for almost } t \in [0, T] \\
x(\theta) = \phi(\theta) & \theta \in [-\tau, 0].
\end{cases}
\]

(3.12)

Let \( C = C(\mathbb{R}^n, \mathbb{R}^n) \), and define a family of functions

\[
\Xi = \{ f(x) : f(x) = [f_1(x_1), f_2(x_2), \ldots, f_n(x_n)]^T \in C \}
\]

satisfying:

1. Every \( f_i(\cdot) \) is nondecreasing for all \( i = 1, 2, \ldots, n \)
2. Every \( f_i(\cdot) \) is uniformly locally bounded, that is, for any compact set \( Z \subset \mathbb{R}^n \), there exists a constant \( M > 0 \) independent of \( f \) such that
   \[
   |f_i(x)| \leq M \quad \text{for all } x \in Z \ i = 1, \ldots, n.
   \]
3. Every \( f_i(\cdot) \) is locally Lipschitzian continuous, that is, for any compact set \( Z \subset \mathbb{R}^n \), there exists \( \lambda > 0 \) such that for any \( \xi, \zeta \in Z, \) and
\[ i = 1, 2, \ldots, n, \text{we have} \]
\[ |f_i(\xi) - f_i(\zeta)| \leq \lambda|\xi - \zeta|. \]

4. \( f_i(0) = 0, \text{for all } i = 1, 2, \ldots, n. \)

As pointed out in Hale (1977), if \( f \in X_i \), then the following system,
\[
\begin{aligned}
\frac{du_f(t)}{dt} &= -Du_f(t) + Af(u_f(t)) + Bf(u_f(t - \tau)) \\
u_f(\theta) &= \phi(\theta) \quad \theta \in [-\tau, 0],
\end{aligned}
\]  
(3.13)

has a unique solution \( u_f(t) = (u_1(t), u_2(t), \ldots, u_n(t))^T \) on \([0, +\infty)\). Moreover, we can prove the following result.

**Theorem 2.** If the matrix inequality 3.3 is satisfied, then for any \( \phi \in C \), there exists \( M = M(\phi) > 0 \) such that

\[ |u_f(t)| \leq Me^{-\frac{t}{2}}, \quad \text{for all } t > 0 \text{ and } f \in \Xi, \]  
(3.14)

where \( \varepsilon > 0 \) is a constant.

**Proof.** Let

\[ V_2(t) = e^{tT}u_f^T(t) + 2\sum_{i=1}^{n}e^{tT}P_i \int_{0}^{u_f(t)} f_i(\rho)d\rho \]
\[ + \int_{t-\tau}^{t} f^T(u_f(s))Qf(u_f(s))e^{(s+\tau)}ds, \]

where \( \varepsilon, \hat{P}, \) and \( \hat{Q} \) are of the second conclusion of lemma 2. And differentiate \( V_2(t) \):

\[
\frac{dV_2(t)}{dt} = \varepsilon e^{tT}u_f(t)^T + 2e^{tT}P_i \int_{0}^{u_f(t)} f_i(\rho)d\rho - e^{tT}f^T(u_f(t-	au))Qf(u_f(t-	au))
\]
\[ + e^{tT}f^T(u_f(t))\hat{Q}f(u_f(t)). \]  
(3.15)

Pick \( \varepsilon < \min_i d_i \). Then

\[ \varepsilon \int_{0}^{u_f(t)} f_i(\rho)d\rho \leq \varepsilon u_f(t) f_i(u_f(t)) \leq d_i u_f(t) f_i(u_f(t)). \]
By matrix inequality 3.9, we have
\[
\frac{dV_2(t)}{dt} \leq e^{\epsilon t} \left[ u_f^T(t), f^T(u_f(t)), f^T(u_f(t - \tau)) \right] Z_1 \begin{bmatrix} u_f(t) \\ f(u_f(t)) \\ f(u_f(t - \tau)) \end{bmatrix}.
\]
\[
\leq 0
\]
Therefore, \( u_f(t)^T u_f(t)e^{\epsilon t} \leq V_2(t) \leq V_2(0) \leq M \), where \( M \) is a constant independent of \( f \in \mathcal{E} \). Let \( M_1 = \sqrt{M} \). We have
\[
\|u_f(t)\| \leq M_1 e^{-\frac{\epsilon t}{2}} \quad \text{for all } t > 0 \text{ and } f \in \mathcal{E}.
\]
(3.16)

Construct a sequence of systems with high-slope continuous activation functions, and prove that the sequence converges to a solution of system 3.7.

Let \( \{\rho_{k,i}\} \) be the set of discontinuous points of \( g_i(\cdot) \). Pick a strictly decreasing sequence \( \{\delta_{k,i,m}\}_m \) with \( \lim_{m \to \infty} \delta_{k,i,m} = 0 \) such that \( I_{k_1,i,m} \bigcap I_{k_2,i,n} = \emptyset \) holds for any \( k_1 \neq k_2 \) and \( m, n \in \mathbb{N} \), where \( I_{k,i,m} = [\rho_{k,i} - \delta_{k,i,m}, \rho_{k,i} + \delta_{k,i,m}] \).

Define functions \( \{g^m(x) = (g_{1}^m(x_1), \ldots, g_{n}^m(x_n))^T, m = 1, \ldots, \) as follows:

\[
g_i^m(s) = \begin{cases} 
g_i(s) & \text{if } s \in \bigcup_k I_{k,i,m} \\
g_i(\rho_{k,i} + \delta_{k,i,m}) - g_i(\rho_{k,i} - \delta_{k,i,m}) \frac{2 \delta_{k,i,m}}{\rho_{k,i} + \delta_{k,i,m} - \rho_{k,i} - \delta_{k,i,m}} [x - (\rho_{k,i} + \delta_{k,i,m})] & \text{if } 0 \neq \rho_{k,i} \text{ and } s \in I_{k,i,m} \\
g_i(\rho_{k,i} + \delta_{k,i,m}) & \text{if } 0 = \rho_{k,i} \text{ and } s \in [\rho_{k,i}, \rho_{k,i} + \delta_{k,i,m}] \\
g_i(\rho_{k,i} - \delta_{k,i,m}) & \text{if } 0 = \rho_{k,i} \text{ and } s \in [\rho_{k,i} - \delta_{k,i,m}, \rho_{k,i}] \\
g_i(\rho_{k,i} - \delta_{k,i,m}) - g_i(\rho_{k,i} + \delta_{k,i,m}) \frac{2 \delta_{k,i,m}}{\rho_{k,i} + \delta_{k,i,m} - \rho_{k,i} - \delta_{k,i,m}} [x - \rho_{k,i}] & \text{if } 0 = \rho_{k,i} \text{ and } s \in [\rho_{k,i} - \delta_{k,i,m}, \rho_{k,i}] \end{cases}
\]
(3.17)

It can be seen that every \( \{g^m(x), m = 1, \ldots, \) satisfies

- \( g^m(x) \in \mathcal{E} \).
- For any compact set \( Z \subset \mathbb{R}^n \),
\[
\lim_{m \to \infty} d^*(\text{Graph}(g^m(Z)), \text{Graph}(K[g(Z)])) = 0
\]
where
\[
d^*(A, B) = \sup_{x \in A} \inf_{y \in B} \|x - y\|.
\]
- For every continuous point \( s \) of \( g_i(\cdot) \), there exists \( m_0 \in \mathbb{N} \) such that \( g_i^m(s) = g_i(s) \), for all \( m \geq m_0, i = 1, 2, \ldots, n \).
Let $u^m(t) = (u_1^m(t), \ldots, u_n^m(t))^T$ be the solution of the following system:

$$
\frac{d u^m(t)}{dt} = -D u^m(t) + A g^m(u^m(t)) + B g^m(u^m(t - \tau))
$$

$u^m(\theta) = \phi(\theta) \quad \theta \in [-\tau, 0]. \tag{3.18}
$}

Next, we will prove that system 3.7 (or, equivalently, system 2.8) has at least one solution.

**Theorem 3 (Viability theorem).** If the matrix inequality 3.3 holds and $g \in \widetilde{G}$, then the system 3.7 has a solution $u(t) = (u_1(t), \ldots, u_n(t))^T$ for $t \in [0, \infty)$.

**Proof.** By theorem 2, we know that all the solutions $\{u^m(t)\}$ of system 3.18 are uniformly bounded, which implies that $\{\frac{d u^m(t)}{dt}\}$ are also uniformly bounded. By the Arzela-Ascoli lemma and the diagonal selection principle, we can select a sub-sequence of $\{u^m(t)\}$ (still denoted by $\{u^m(t)\}$) such that $u^m(t)$ uniformly converges to a continuous function $u(t)$ on any compact set of $R$.

Because the derivative of $\{u^m(t)\}$ is uniformly bounded, it can be seen that for any fixed $T > 0$, $u(t)$ is Lipschitz continuous on $[0, T]$. Therefore, $\frac{d u(t)}{dt}$ exists for almost all $t$ and is bounded on $[0, T]$.

For each $p(t) \in C_0^\infty([0, T], R^n)$ (noticing that $C_0^\infty([0, T], R^n)$ is dense in the Banach space $L^1([0, T], R^n)$),

$$
\int_0^T \left( \frac{d u^m(t)}{dt} - \frac{d u(t)}{dt} \right) p(t) \, dt = - \int_0^T \frac{d p(t)}{dt} (u^m(t) - u(t)) \, dt
$$

holds and $\{\frac{d u^m(t)}{dt}\}$ is uniformly bounded. Therefore, $\frac{d u^m(t)}{dt}$ weakly converge to $\frac{d u(t)}{dt}$ on $L^\infty([0, T], R^n) \subset L^1([0, T], R^n)$. By Mazur’s convexity theorem (see Yosida, 1978), we can find constants $a_i^m \geq 0$ with $\sum_{l=m}^\infty a_i^l = 1$, and for any $m$ only finite $a_i^m \neq 0$ such that $y^m(t) = \sum_{l=m}^\infty a_i^l u^l(t)$. Then $y^m(\theta) = \phi(\theta)$, if $\theta \in [-\tau, 0]$, and

$$
\lim_{m \to \infty} y^m(t) = u(t), \text{ on } [0, T] \text{ uniformly} \tag{3.19}
$$

$$
\lim_{m \to \infty} \frac{d y^m(t)}{dt} = \frac{d u(t)}{dt}, \text{ for all almost } t \in [0, T]. \tag{3.20}
$$

Let $\gamma^m(s) = \sum_{l=m}^\infty a_i^l g^l(u^l(s))$. Then

$$
\frac{d y^m(t)}{dt} = -D y^m(t) + A y^m(t) + B y^m(t - \tau). \tag{3.21}
$$
Finally, we will prove that there exists a measurable function $\gamma(t)$ that is limit of a sub-sequence of $\gamma^m(t)$ and satisfies

$$\frac{du(t)}{dt} = -Du(t) + Ay(t) + B\gamma(t - \tau), \quad \text{for almost } t \in [0, T].$$

First, we consider the time interval of $t \in [0, \tau]$. For $s \in [-\tau, 0]$, $g^i(u^i(s)) = g^i(\phi(s))$. According to the boundedness of $\phi(\cdot)$, the uniform boundedness of $g_m(\cdot)$ on the image of $\phi$, and the limitedness of the set of discontinuous points of $g(\cdot)$ on the image of $\phi$, we can find a sub-sequence of $g^m$ (still denoted by $g^m$) and a measurable function $\gamma(t)$ on $[-\tau, 0]$ such that $g^m(\phi(s))$ converges to $\gamma(s)$ for all $s \in [-\tau, 0]$. Therefore, $\lim_{m \to \infty} \gamma^m(\phi(s)) = \gamma(s)$ for all $s \in [-\tau, 0]$

For $t \in [0, \tau]$ and then $t - \tau \in [-\tau, 0]$, we can find a measurable function on $[0, \tau]$, still denoted by $\gamma(t)$, such that

$$\gamma(t) = \lim_{m \to \infty} \gamma^m(t) = A^{-1} \lim_{m \to \infty} \left[ \frac{dy^m}{dt} + Dy^m(t) + B\gamma(t - \tau) \right]$$

$$= A^{-1} \left[ \frac{du}{dt} + Du(t) + B\gamma(t - \tau) \right] \quad \text{for almost } t \in [0, \tau].$$

Similarly, we can construct a measurable function $\gamma(t)$, $t \in [0, T]$, such that

$$\frac{du}{dt} = -Du(t) + Ay(t) + B\gamma(t - \tau) \quad \text{for almost } t \in [0, T]. \quad (3.22)$$

Then we will prove $\gamma(t) \in K[g(u(t))]$.

Both $y^m(t)$ and $u^m(t)$ converge to $u(t)$ uniformly, and $K[g(\cdot)]$ is an upper-semicontinuous set-valued map. Therefore, for any $\epsilon > 0$, there exists $N > 0$ such that for all $m > N$ and $t \in [0, T]$, we have $g^m(u^m(t)) \in O(K[g(u(t))], \epsilon)$, where $O(K[g(u(t))], \epsilon) = \{x \in \mathbb{R}^n : d(x, K[g(u(t))]) < \epsilon\}$. Because $K[g(\cdot)]$ is convex and compact, $\gamma^m(t) \in O(K[g(u(t))], \epsilon)$ for $t \in [0, T]$. Letting $m \to \infty$, we have $\gamma(t) \in O(K[g(u(t))], \epsilon)$ for $t \in [0, T]$. Because of the arbitrariness of $\epsilon$, we conclude that

$$\gamma(t) \in K[g(u(t))] \quad t \in [0, T]. \quad (3.23)$$

Because $T$ is arbitrary, the solution $u(t)$ can be extended to infinite time interval $[0, +\infty)$. Theorem 3 is proved.

### 4 Global Asymptotic Stability

In this section, we study the global stability of system 3.7.
Theorem 4 (global exponential asymptotic stability). If the matrix inequality 3.3 holds and \( g(\cdot) \in \bar{G} \), then for any solution \( u(t) \) on \([0, \infty)\) of system 3.7, there exists \( M = M(\phi) > 0 \) such that

\[
\|u(t)\| \leq Me^{-\frac{\varepsilon}{2}t} \text{ for all } t > 0,
\]

where \( \varepsilon \) is given by matrix inequality 3.9. Equivalently, for any solution \( x(t) \) on \([0, \infty)\) of system 2.8, we have

\[
\|x(t) - x^*\| \leq Me^{-\frac{\varepsilon}{2}t} \text{ for all } t > 0.
\]

Proof. Let

\[
V_3(t) = e^{\varepsilon t} u^T(t) u(t) + 2 \sum_{i=1}^{n} \varepsilon e^{\varepsilon t} \hat{P}_i \int_{0}^{u_i(t)} g_i(\rho) d\rho + \int_{t-\tau}^{t} \gamma(s)^T \hat{Q} \gamma(s) e^{\varepsilon(s+\tau)} ds.
\]

Notice that for \( p_i(s) = \int_{0}^{s} g_i(\rho) d\rho \), we have \( \partial_p p_i(s) = \{ v \in R : g_i^-(s) \leq v \leq g_i^+(s) \} \).

Differentiating \( V_3(t) \) by the chain rule (for details, see Bacciotti, Conti, & Marcellini, 2000; Clarke, 1983; or Lu & Chen, 2005, for details), we have

\[
\frac{dV_3(t)}{dt} = \varepsilon e^{\varepsilon t} u^T(t) u(t) + 2 e^{\varepsilon t} u^T(t) [-Du + A\gamma(t) + B\gamma(t - \tau)]
+ 2 e^{\varepsilon t} \gamma(t) \hat{P} [-Du(t) + A\gamma(t) + B\gamma(t - \tau)]
+ \varepsilon e^{\varepsilon t} \sum_{i=1}^{n} \hat{P}_i \int_{0}^{u_i} g_i(\rho) d\rho - e^{\varepsilon t} \gamma^T(t - \tau) \hat{Q} \gamma(t - \tau)
+ e^{\varepsilon(t+\tau)} \gamma^T(t) \hat{Q} \gamma(t).
\]

Because \( \varepsilon < \min_i d_i \), we have

\[
\varepsilon \int_{0}^{u_i} g_i(\rho) d\rho \leq \varepsilon u_i(t) \gamma_i(t) \leq d_i u_i(t) \gamma_i(t)
\]

and

\[
\frac{dV_3(t)}{dt} \leq e^{\varepsilon t} [u^T(t), \gamma^T(t), \gamma^T(t - \tau)] Z_1 \begin{bmatrix} u(t) \\ \gamma(t) \\ \gamma(t - \tau) \end{bmatrix}
\leq 0.
\]
Then, $u(t)^T u(t) \leq V_3(0)e^{-\varepsilon t}$ and

$$\|u(t)\|_2 \leq \sqrt{V_3(0)e^{-\frac{1}{2} t}}$$

$$\|x(t) - x^*\|_2 \leq \sqrt{V_3(0)e^{-\frac{1}{2} t}}$$

hold.

**Remark 3.** From theorem 4, the uniqueness of the equilibrium is also obtained.

**Corollary 1.** If condition 3.3 holds and $g_i(\cdot)$ is locally Lipschitz continuous, then there exist $\varepsilon > 0$ and $x^* \in \mathbb{R}^n$ such that for any solution $x(t)$ on $[0, \infty)$ of system 1.1, there exists $M = M(\phi) > 0$ such that

$$\|x(t) - x^*\| \leq Me^{-\frac{1}{2} \varepsilon t} \quad \text{for all } t > 0$$

where $\varepsilon$ is given by the matrix inequality 3.9.

If every $x^*_i$ is a continuous point of the activation functions $g_i(\cdot)$, $i = 1, \ldots, n$. For the outputs, we have $\lim_{t \to \infty} g_i(x_i(t)) = g_i(x^*_i)$. Instead, if for some $i$, $x^*_i$ is a discontinuous point of the activation function $g_i(\cdot)$, we can prove the outputs converge in measure (also see Forti & Nistri, 2003).

**Theorem 5 (convergence of output).** If the matrix inequality 3.3 holds and $g(\cdot) \in \bar{G}$, then the output $\alpha(t)$ of system 2.7 converges to $\alpha^*$ in measure, that is, $\mu - \lim_{t \to \infty} \alpha(t) = \alpha^*$

**Proof.** Define

$$V_5(t) = u^T(t)u(t) + 2 \sum_{i=1}^{n} \hat{P}_i \int_0^{u_i} g_i(\rho) d\rho + \int_{t-\tau}^{t} \gamma(s)^T \hat{Q} \gamma(s) ds,$$

where $\hat{P}$, $\hat{Q}$, and $\epsilon$ are those in the matrix inequality 3.10 of lemma 1.

Differentiate $V_5(t)$:

$$\frac{dV_5(t)}{dt} = 2u^T(t)[-Du(t) + A\gamma(t) + B\gamma(t - \tau)] + 2\gamma^T(t)\hat{P}[-Du(t)$$

$$+ A\gamma(t) + B\gamma(t - \tau)] + \gamma^T(t)\hat{Q}\gamma(t) - \gamma^T(t - \tau)\hat{Q}\gamma(t - \tau)$$

$$+ \epsilon \gamma(t)^T \gamma(t) - \epsilon \gamma(t)^T \gamma(t)$$
\[ [u^T(t), \gamma^T(t), \gamma^T(t - \tau)] Z_2 \begin{bmatrix} u(t) \\ \gamma(t) \\ \gamma(t - \tau) \end{bmatrix} - \varepsilon \gamma^T(t) \gamma(t) \leq -\varepsilon \gamma^T(t) \gamma(t). \] (4.2)

Then
\[ V_5(t) - V_5(0) \leq -\varepsilon \int_0^t \gamma^T(s) \gamma(s) ds. \]

Since \( \lim_{t \to \infty} V_5(t) = 0 \),
\[ \int_0^\infty \gamma^T(s) \gamma(s) ds \leq -\frac{1}{\varepsilon} V_5(0). \]

For any \( \epsilon_1 > 0 \), let \( E_{\epsilon_1} = \{ t \in [0, \infty) : \| \gamma(t) \| > \epsilon_1 \} \):
\[ \frac{V_5(0)}{\epsilon} \geq \int_0^\infty \gamma^T(s) \gamma(s) ds \geq \int_{E_{\epsilon_1}} \gamma^T(s) \gamma(s) \geq \epsilon_1^2 \mu(E_{\epsilon_1}). \]

Therefore, \( \mu(E_{\epsilon_1}) < \infty \). From proposition 2 in Forti and Nistri (2003), one can see that \( \gamma(t) \) converges to zero in measure, that is, \( \mu - \lim_{t \to \infty} \gamma(t) = 0 \). Therefore, \( \mu - \lim_{t \to \infty} \alpha(t) = \alpha^* \).

**Remark 4.** From the proof of theorem 5, one can see that the equilibrium of output \( \alpha^* \) is also unique.

Similar to the concept of the Filippov solution for a system of ordinary differential equations with a discontinuous right-hand side, we propose the concept of the solution for the delayed system 2.8. Suppose \( g(\cdot) \in \tilde{G} \) is bounded. The solution of system 2.8 can be regarded as an approximation of the solutions of delayed neural networks with high-slope gain functions. From the proof of viability, one can see that any limit of the solutions of delayed neural networks with high-slope activation functions, which converge to the discontinuous activations \( \alpha(t) \), is a solution of system 2.8. More precisely, we give the following result:

**Theorem 6.** Suppose \( g(\cdot) \in \tilde{G} \) is bounded, and the function sequence \( \{g^m(x) = (g^m_1(x_1), g^m_2(x_2), \ldots, g^m_n(x_n))^T : m = 1, 2, \ldots \} \) satisfies:

1. \( \{g^m_i(\cdot)\} \) is nondecreasing for all \( i = 1, 2, \ldots, n \).
2. \( \{g^m_i(\cdot)\} \) is locally Lipschitz for all \( i = 1, 2, \ldots, n \).
3. For any compact set $Z \subset \mathbb{R}^n$,
\[
\lim_{m \to \infty} d^\star(\text{Graph}(g^m(Z)), \text{Graph}(K_g(Z))) = 0.
\]
x_m(t) is the solution of the following system:
\[
\begin{align*}
\frac{dx_m}{dt} &= -Dx_m(t) + A g_m(x_m(t)) + B g_m(x_m(t - \tau)) + I \\
x_m(\theta) &= \phi(\theta) \quad \theta \in [-\tau, 0].
\end{align*}
\] (4.3)

Then there exists a sub-sequence $m_k$ such that for any $T > 0$, $x_{m_k}(t)$ converges uniformly to an absolutely continuous Function $x(t)$ on $[0, T]$, which is a solution of system 2.8 on $[0, \infty)$. And for any sub-sequence $m_k$ that converges uniformly to an absolutely continuous function $x(t)$ on any $[0, T]$, $x(t)$ must be a solution of system 2.8 on $[0, \infty)$; moreover, if the solution of the system 2.8 is unique, then the sequence $x_m(t)$ itself uniformly converges to $x(t)$ on any $[0, T]$.

The proof is similar to that of theorem 3. Details are omitted.

**Remark 5.** There are close relationships and essential differences between this article and that of Forti and Nistri (2003). The dynamical behaviors of neural networks with discontinuous activations were first investigated in Forti and Nistri (2003). However, in that article, the authors assumed that the discontinuous neuron activations are bounded. We do not make that assumption here. Thus, the discussion of the existence of the equilibrium and solution of system 2.7 is much more difficult. Furthermore, the model in this article is a delayed neural network. Here, we introduce a new concept for the solution of delayed neural networks with discontinuous activation functions. Obviously, if a delayed feedback matrix $B = 0$, the conclusion that Forti and Nistri (2003) proposed for global stability can be regarded as a corollary.

5 Numerical Examples

In this section, we present several numerical examples to verify the theorems we have given in the previous sections.

Consider a two-dimensional neural network with time delay,
\[
\frac{dx(t)}{dt} = -Dx(t) + A g(x(t)) + B g(x(t - \tau)) + I,
\] (5.1)

where $x(t) = (x_1(t), x_2(t))^T \in \mathbb{R}^2$ denotes the state and $\alpha(t) = (\alpha_1(t), \alpha_2(t))^T$ denotes the output satisfying $\alpha_1(t) \in K[g_1(x_1(t))], \alpha_2(t) \in K[g_2(x_2(t))]$. 
In examples 1, 2, and 3, we assume

\[
D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} -\frac{1}{4} & 2 \\ -10 & -\frac{1}{4} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{5} & \frac{1}{10} \\ -\frac{1}{7} & \frac{1}{10} \end{bmatrix},
\]

and \(g_1(s) = g_2(s) = g(s) = s + \text{sign}(s), \tau = 1\). By the Matlab LMI and Control Toolbox, we obtain

\[
P = \begin{bmatrix} 5.8083 & 0 \\ 0 & 1.1796 \end{bmatrix}, \quad Q = \begin{bmatrix} 1.1515 & 0.2456 \\ 0.2456 & 0.4056 \end{bmatrix},
\]

such that

\[
Z = \begin{bmatrix} -PA - A^T P -PB \\ -B^T P \\ Q \end{bmatrix} > 0.
\]

By theorem 4, system 5.1 is globally exponentially stable for any input \(I \in \mathbb{R}^2\).

Figure 1: Dynamical behaviors of the solution of the neural networks 5.2.
5.1 Example 1. Consider the following system:

\[
\begin{align*}
\dot{x}_1 &= -x_1(t) - \frac{1}{4}[x_1(t) + \text{sign}(x_1(t))] + 2[x_2(t) + \text{sign}(x_2(t))] \\
&\quad + \frac{1}{5}[x_1(t - 1) + \text{sign}(x_1(t - 1))] + \frac{1}{10}[x_2(t - 1) + \text{sign}(x_2(t - 1))] + 6 \\
\dot{x}_2 &= -x_2(t) - 10[x_1(t) + \text{sign}(x_1(t))] - \frac{1}{4}[x_2(t) + \text{sign}(x_2(t))] \\
&\quad - \frac{1}{7}[x_1(t - 1) + \text{sign}(x_1(t - 1))] + \frac{1}{10}[x_2(t - 1) + \text{sign}(x_2(t - 1))] + 10.
\end{align*}
\]

Let \(\phi_1(\theta) = -e^{10\theta}, \phi_2(\theta) = \sin(20\theta), \theta \in [-1, 0]\), be the initial condition.

By the previous analysis, system 5.2 is globally exponentially stable. The trajectories of \(x_1(t)\) and \(x_2(t)\) are shown in Figure 1, and the trajectories of output, \(\alpha_1(t)\) and \(\alpha_2(t)\), are shown in Figure 2. The equilibrium of system 5.2 is
It is clear that 0.1974 and −1.7346 are continuous points of the function \( \rho + \text{sign}(\rho) \). Thus, the output \( \lim_{t \to \infty} \alpha(t) = (1.1974, -2.7346)^T \).

**5.2 Example 2.** In this example, we change the inputs and consider the following system,

\[
\begin{align*}
\dot{x}_1(t) &= -x_1(t) - \frac{1}{4}[x_1(t) + \text{sign}(x_1(t))] + 2[x_2(t) + \text{sign}(x_2(t))] \\
&
\quad + \frac{1}{5}[x_1(t-1) + \text{sign}(x_1(t-1))] + \frac{1}{10}[x_2(t-1)] + \frac{43}{20} \text{sign}(x_2(t-1)) \\
\dot{x}_2(t) &= -x_2(t) - 10[x_1(t) + \text{sign}(x_1(t))] - \frac{1}{4}[x_2(t) + \text{sign}(x_2(t))] \\
&
\quad - \frac{1}{7}[x_1(t-1) + \text{sign}(x_1(t-1))] + \frac{1}{10}[x_2(t-1)] + \text{sign}(x_2(t-1)) + \frac{1399}{140},
\end{align*}
\]

(5.3)
with the same initial condition as in example 1. The equilibrium of system 5.3 is $(0, 0)^T$, and the equilibrium of outputs is $(1, -1)^T$. It can be seen that 0 is a discontinuous point of the activation function $g(\rho) = \rho + \text{sign}(\rho)$ and $K[g(0)] = [1, -1]$. In this case, the solution trajectories $(x_1(t), x_2(t))^T$ converge to the equilibrium, as indicated by Figure 3. The outputs $(\alpha_1(t), \alpha_2(t))^T$ cannot converge in norm, but by theorem 5, they converge in measure, as indicated by Figure 4.

5.3 Example 3. We use this example to verify the validity of theorem 6. Consider the following system,

\[
\begin{align*}
\dot{x}_1 &= -x_1(t) + \text{sign}(x_1(t)) + \text{sign}(x_2(t)) + \text{sign}(x_1(t-1)) + \text{sign}(x_2(t-1)) \\
\dot{x}_2 &= -x_2(t) + \text{sign}(x_1(t)) + \text{sign}(x_2(t)) + \text{sign}(x_1(t-1)) + \text{sign}(x_2(t-1)).
\end{align*}
\]

and construct a sequence of systems as follows,

\[
\begin{align*}
\dot{x}_{m,1} &= -x_{m,1}(t) + \tan h(mx_{m,1}(t)) + \tan h(mx_{m,2}(t)) + \tan h(mx_{m,1}(t-1)) \\
&\quad + \tan h(mx_{m,2}(t-1))
\end{align*}
\]
Figure 5: Variation of $\text{error}_m$ with respect to $m$.

\[
\dot{x}_{m,2} = -x_{m,2}(t) + \tan h(mx_{m,1}(t)) + \tan h(mx_{m,2}(t)) + \tan h(mx_{m,1}(t-1)) \\
+ \tan h(mx_{m,2}(t-1)),
\]  

(5.5)

with the same initial condition $\phi_1(s) = -4$ and $\phi_2(s) = 10$ for $s \in [-1, 0]$. Pick 20 sample points $1 = t_1 < t_2 < \ldots < t_{20} = 3$, and define

\[
\text{error}_m = \max_k \| x(t_k) - x_m(t_k) \|.
\]

Figure 5 indicates that $\text{error}_m$ converges to zero with respect to $m$. Therefore, the solutions $\{x_m(t)\}$ of systems 5.5 with high-slope activation functions converge to the solution $x(t)$ of the delayed neural network with discontinuous activation.

6 Conclusion

In this letter, we considered the global stability of delayed neural networks with discontinuous activation functions, which might be unbounded. We extend the Filippov solution to the case of delayed neural networks. Under some conditions, the existence of equilibrium and absolutely continuous solution on infinite time interval is proved. Thus, the Lyapunov-type method
can be used to study stability. In this way, we obtain an LMI-based criterion for global convergence. If some components of the equilibrium are discontinuous points of the activation function, the output does not converge in norm. In this case, we prove that the output converges in measure. Furthermore, we point out that the solution of the delayed neural networks with discontinuous activation functions can be regarded as a limit of the solution sequence of delayed systems with high-slope activation functions.

Acknowledgments

We are very grateful to the reviewers for their comments, which were helpful to us in revising the article. This work is supported by National Science Foundation of China 60374018 and 60574044 and also supported by the Graduate Student Innovation Foundation of Fudan University.

References


Received September 24, 2004; accepted July 11, 2005.