Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Almost periodic dynamics of a class of delayed neural networks with discontinuous activations

by

Wenlian Lu, and Tianping Chen

Preprint no.: 53 2007

Almost periodic dynamics of a class of delayed neural networks with discontinuous activations

Wenlian Lu[1], Tianping, Chen[2]

Abstract

We use the concept of the Filippov solution to study the dynamics of a class of delayed dynamical systems with discontinuous right-hand side, which contains the widely-studied delayed neural network models with almost periodic self-inhibitions, interconnections weights and external inputs. We prove that diagonal dominant conditions can guarantee the existence and uniqueness of an almost periodic solution as well as its global exponential stability. As special cases, we derive a series of results on the dynamics of delayed dynamical systems with discontinuous activations and periodic coefficients or constant coefficients, respectively. Furthermore, from the proof of the existence and uniqueness of the solution, we prove that the solution of a delayed dynamical system with high-slope activations actually approximates to the Filippov solution of the dynamical system with discontinuous activations.

Key words: Delayed integro-differential system, discontinuous activation, almost periodic function, global exponential stability

[1]Max Planck Institute for Mathematics in the Sciences, 04103, Leipzig, Germany, and Lab. of Mathematics for Nonlinear Sciences, School of Mathematical Sciences, Fudan University, 200433, Shanghai, P. R. China.

[2]Lab. of Mathematics for Nonlinear Sciences, School of Mathematical Sciences, Fudan University, 200433, Shanghai, P. R. China.

Email Addresses: wenlian@mis.mpg.de, wenlian.lu@gmail.com (Wenlian Lu); tchen@fudan.edu.cn (corresponding author: Tianping Chen)

1 Introduction

The purpose of this paper is to study the following delayed integro-differential equations:

$$\begin{aligned} \frac{du_i(t)}{dt} &= -d_i(t)u_i(t) + \sum_{j=1}^{n} a_{ij}(t) g_j(u_j(t)) \\ &\quad + \sum_{j=1}^{n} \int_0^{\infty} g_j(u_j(t-s)) d_s K_{ij}(t,s) + I_i(t),\ i = 1, \cdots, n, \end{aligned} \tag{1}$$

where $d_i(t),\ a_{ij}(t),\ I_i(t),\ i,j = 1,\cdots,n$, are some functions from $\mathbb{R}^+$ to $\mathbb{R}$, $g_i(\cdot)$, $i = 1,\cdots,n$, are some nondecreasing functions from $\mathbb{R}$ to $\mathbb{R}$, and for any $t \in \mathbb{R}$, $d_sK_{ij}(t,s)$, $i,j = 1,\cdots,n$, are Lebesgue-Stieltjes measures with respect to s.

The motivation of studying the system (1) originates from the study of the well-known recurrently connected neural networks, which have been extensively studied in both theory and applications. The neural networks can be modelled by the following differential equations:

$$\frac{du_i(t)}{dt} = -d_i u_i(t) + \sum_{j=1}^{n} a_{ij} g_j(u_j(t)) + I_i,\ i = 1,\cdots,n \tag{2}$$

known as Hopfield neural networks (Hopfield & Tank 1984, 1986) and

$$\frac{du_i(t)}{dt} = a_i(u_i(t))\Big[- d_i u_i(t) + \sum_{j=1}^{n} a_{ij} g_j(u_j(t)) + I_i\Big],\ i = 1,\cdots,n \tag{3}$$

known as Cohen-Grossberg neural networks (Cohen & Grossberg 1983), where $u_i(t)$ denotes the state variable of the i-th neuron, d_i represents the self-inhibition with which the i-th neuron will reset its potential to the resting state in isolations when disconnected from the network, a_{ij} denotes the connection strength of j-th neuron on the i-th neuron, $g_i(\cdot)$ denotes the activation function of i-th neuron, I_i denotes the external input to the i-th neuron, and $a_i(\cdot)$ denotes amplification function of the i-th neuron. There are a lot of papers in literature discussing the local and global stability of these systems. For reference, see Cohen & Grossberg (1983), Hirsch (1989), Michel & Gray (1990), Forti & Tesi (1995), Lu & Chen (2003) and others.

In practice, time delays inevitably occur due to the finite switching speed of the amplifiers and communication time. Thus, the neural networks can be modeled by the following delayed differential equations:

$$\frac{dx_i(t)}{dt} = -d_i x_i(t) + \sum_{j=1}^{n} a_{ij} g_j(x_j(t)) + \sum_{j=1}^{n} b_{ij} g_j(x_j(t-\tau_{ij})) + I_i, \ i = 1, \cdots, n, \tag{4}$$

where b_{ij} denotes the delayed feedback of the j-th neuron on the i-th neuron. There are also many papers discussing the stability of delayed neural networks. See Gopalsamy & He (1991), Civalleri et.al. (1993), Blair et.al. (1996), Cao & Zhou (1998), Joy (2000), Chen (2001), Lu et.al. (2003) for references. In these papers, various conditions based on Lyapunov functionals were given guaranteeing the global stability.

Furthermore, the interconnections also contain asynchronous terms. In general, the interconnection weights a_{ij}, b_{ij}, self-inhibitions d_i and inputs I_i should vary through time. Therefore, we need to study the nonautonomous dynamical systems with time-varying self-inhibitions, connections, and inputs:

$$\begin{aligned}\frac{du_i(t)}{dt} &= -d_i(t)u_i(t) + \sum_{j=1}^{n} a_{ij}(t) g_j(u_j(t)) \\ &\quad + \sum_{j=1}^{n} b_{ij}(t) g_j(u_j(t-\tau_{ij})) + I_i(t), \ i = 1, \cdots, n\end{aligned} \tag{5}$$

Recently, a number of researchers have investigated the existence and global attraction of the periodic solution (Gopalsamy & Sariyasa 2002, Cao 2003, Lu & Chen 2004, Zhou et.al. 2004, Zheng & Chen 2004, Chen et.al. 2005) or almost periodic solution (Huang & Cao 2003, Lu & Chen 2005b) for these non-autonomous delayed differential systems, assuming that the system is periodic or almost periodic respectively. In particular, Lu & Chen (2005b) presented the generalized delayed differential system model (1) unifying discrete delays and distribution delays and studied its dynamical behaviors.

However, all these works were based on the assumption that the activation functions are continuous even globally Lipshitz. As mentioned by Forti & Nistri (2003), a brief review on some common neural network models reveals that neural networks with discontinuous activations are of importance and do frequently arise in practice. For example, consider the classical Hopfield neural networks with graded response neurons Hopfield & Tank (1984). The standard assumption is that the activations used are in high-gain limit, where they closely approach discontinuous and comparator functions. As shown in Hopfield & Tank (1984, 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, the activation function $g_i(\cdot)$ is selected as the sign function $sign(s)$.

Also, a conceptually analogous model based on hard comparators is also used to describe the discrete-time neural networks in Harrer et.al. (1992). Another important example is the neural networks introduced in Kennedy & Chua (1988) to solve linear and nonlinear programming problems. Those networks exploit constrained neurons with a diode-like input-output activations. Again, in order to satisfy the constraints, the diodes are required to possess a very high slope in the conducting region, i.e., they should approximate the discontinuous characteristic of an ideal diode Chua et.al. (1987). When dealing with dynamical systems possessing high-slope nonlinear elements, it is often of advantage to model them with a system of differential equations with discontinuous right-hand side, rather than studying the case where the slope is high but of finite value (Utkin 1977).

In the last few years, there are several papers studying neural networks with discontinuous activations. Forti & Nistri (2003) discussed the absolute stability of Hopfield neural networks (2) with bounded and discontinuous activations. Lu & Chen (2005a) proved the global convergence for Cohen-Grossberg neural networks (3) with unbounded and discontinuous activations. Also, Lu & Chen (2006), Forti et.al. (2005) studied the dynamics of delayed neural networks (4). Papini & Taddei (2005) discussed the periodic solution of the periodic delayed neural networks (5) with discontinuous activations and periodic parameters. In all these papers, the authors use the solution in the Filippov sense (Filippov 1967) to handle differential equations with discontinuous right-hand side. The concept of the solution in the sense of Filippov is useful in engineering applications. Since a Filippov solution is a limit of the solutions of a sequence of ordinary differential equations with continuous right-hand side. Hence, we can model a system which is near a discontinuous system and expect that the Filippov trajectory of the discontinuous system will be close to the real trajectories. This approach is of significance in many applications. For instance, variable structure control, non-smooth analysis, etc, (see Utkin 1977, Aubin & Cellina 1984, Paden & Sastry 1987). In fact, the solution in the Filippov sense satisfies the corresponding differential inclusions induced by the convex extension of discontinuity (see definitions 1 and 2 in the latter section for details).

The generalized viability of the differential inclusions was investigated in the textbooks Aubin & Cellina (1984) Aubin (1991). Periodicity and almost periodicity for differential inclusions or Filippov systems have been studied in the recent decades. Methodologically, the existence of a periodic solution of differential inclusion or differential system with discontinuous right-hand side (despite that some researchers did not study the Filippov solution) can be proved by the fixed-point theory, i.e., the periodic boundary condition can be regarded as a fixed point of certain evolving operator (Hu & Papageorgiou 1995, Li & Xue 2002, Bader & Kryszewski 2003, Li & Kloeden 2006, Dhage 2006, Zuev 2006). And, some authors constructed a sequence of differential systems with continuous right-hand sides having periodic solutions and proved that the solution sequence converges to a periodic solution of the original differential inclusion (Haddad 1981, Filippakis & Papageorgiou 2006). As for the stability, the first approximation was used to deal with the local asymptotical stability for periodic differential inclusions (Smirnov 1995) and Lyapunov method was extended to handle the global stability (Bacciotti et.al. 2000, Bacciotti 2005). Furthermore, similar method was utilized to study the almost periodic solution of almost periodic differential inclusions, especially with delays. See Andres (1999) and Ivanov (1997) for references.

Continuing with our previous work (Lu & Chen 2005a, 2006), in this paper, we consider the delayed dynamical system (1) with discontinuous activations and time-varying coefficients. We also introduce the concept of solutions in the Filippov sense for delayed dynamical system (1) and prove its existence by the idea introduced by Haddad (1981): we construct a sequence of delayed systems, of which the activations are with high-slope and converge to the discontinuous activations. Firstly, we prove that under diagonal dominant conditions, the sequence of solutions has at least a subsequence converging to a solution of the system (1) with discontinuous activations by a well-known diagonal-selection argument. Secondly, we use the Lyapunov functional method to obtain an asymptotical almost periodic solution which leads to the existence of an almost periodic solution (Yoshizawa 1975). We also use this kind of Lyapunov functional to obtain the global exponential stability of this almost periodic solution. Since a periodic function or a constant can be viewed as a special almost periodic function, the results also apply to the systems with periodic (or constant) self-inhibition, connection weights and outer inputs. Furthermore, from the proof of the existence and uniqueness of the solution, we can conclude that each solution sequence of the delayed dynamical system with high-slope activations which converges to the discontinuous activations will actually converge to the unique solution of delayed dynamical system (1) in the Filippov sense with discontinuous activations.

This paper is organized as follows. In section 2, we present some necessary definitions and assumptions. We present and prove the main result in section 3. As consequences, in section 4, we give some corollaries for some specific cases. We conclude this paper in section 5.

2 Preliminaries

In this section, we present some definitions and assumptions which will be used throughout the paper. First of all, we present the definition of a solution for the delayed differential equations (1) with discontinuous right-hand side.

Consider the following system:

$$\frac{dx}{dt} = f(x) \tag{6}$$

where $f(\cdot)$ is not continuous. Filippov (1967) proposed the following definition of the solution for the system (6).

Definition 1 *A set-value map defined as*

$$\phi(x) = \bigcap_{\delta>0} \bigcap_{\mu(N)=0} \overline{co}\Big[f(\overline{\mathcal{O}}(x,\delta) - N)\Big] \tag{7}$$

where $\overline{co}(E)$ is the closure of the convex hull of some set E, $\overline{\mathcal{O}}(x,\delta) = \{y \in \mathbb{R}^n : \|y-x\| \leq \delta\}$, and $\mu(N)$ is the Lebesgue measure of the set N. A solution of the Cauchy problem for (6) 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:

$$\frac{dx}{dt} \in \phi(x), \qquad a.e.\ t \in [0,T) \tag{8}$$

Furthermore, Aubin & Cellina (1984), Aubin (1991), Haddad (1981) have proposed following functional differential inclusion with memory:

$$\frac{dx}{dt}(t) \in F(t, A(t)x) \tag{9}$$

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

$$[A(t)x](\theta) = x_t(\theta) = x(t+\theta) \tag{10}$$

Inspired by these works, in this paper, we use the definition of a solution for the delayed differential systems introduced by Lu & Chen (2006), Forti et.al. (2005), Papini & Taddei (2005), which generalize the previous concepts. We denote $\overline{co}[g_i(s)] = [g_i^-(s), g_i^+(s)]$ and $\overline{co}[g(x)] = \overline{co}[g_1(x_1)] \times \overline{co}[g_2(x_2)] \times \cdots \times \overline{co}[g_n(x_n)]$, where $\times$ denotes the Cartesian product.

Definition 2 *For a continuous function $\phi(\theta) = [\phi_1(\theta), \cdots, \phi_n(\theta)]^\top$ and a measurable function $\lambda(\theta) = [\lambda_1(\theta), \cdots, \lambda_n(\theta)]^\top \in \overline{co}[g(\phi(\theta))]$ for almost all $\theta \in (-\infty, 0]$, an absolute continuous function $u(t) = u(t,\phi,\lambda) = [u_1(t), \cdots, u_n(t)]^\top$ associated with a measurable* *function* $\gamma(t) = [\gamma_1(t), \cdots, \gamma_n(t)]^\top$ *is said to be a solution of the Cauchy problem for the system (1) on* $[0, T)$ *(T might be* ∞*) with initial value* $(\phi(\theta), \lambda(\theta))$*,* $\theta \in (-\infty, 0]$*, if*

$$\begin{cases} \frac{du_i(t)}{dt} = -d_i(t)u_i(t) + \sum\limits_{j=1}^{n} a_{ij}(t)\gamma_j(t) & \\ + \int_0^\infty \gamma_j(t-s)d_sK_{ij}(t,s) + I_i(t) & a, e.\ t \in [0, T), \\ \gamma_i(t) \in K[g_i(u_i(t))] & a.e.\ t \in [0, T), \\ u_i(\theta) = \phi_i(\theta) & \theta \in (-\infty, 0], \\ \gamma_i(\theta) = \lambda_i(\theta) & a.e.\ \theta \in (-\infty, 0], \end{cases} \tag{11}$$

holds for all $i = 1, \cdots, n$*.*

As for the almost periodicity, we use the following concept introduced by Levitan & Zhikov (1982), Yoshizawa (1975).

Definition 3 *A continuous function* $x(t) : \mathbb{R} \to \mathbb{R}^n$ *is said to be almost periodic on* $\mathbb{R}$ *if for any* $\epsilon > 0$*, it is possible to find a real number* $l = l(\epsilon) > 0$*, for any interval with length* $l(\epsilon)$*, there exists a number* $\omega = \omega(\epsilon)$ *in this interval such that* $||x(t + \omega) - x(t)|| < \epsilon$*, for all* $t \in \mathbb{R}$*.*

In the sequel, we also need some assumptions for the delayed system (1):

Assumption $\mathbf{A_1}$: Every $g_i(\cdot)$ is nondecreasing and local Lipschizian, except on a set of isolated points $\{\rho_k^i\}$. More precisely, for each $i = 1, \cdots, n$, $g_i(\cdot)$ is monotonically nondecreasing and continuous except on a set of isolated points $\{\rho_k^i\}$, where the right and left limits $g_i^+(\rho_k^i)$ and $g_i^-(\rho_k^i)$ satisfy $g_i^+(\rho_k^i) > g_i^-(\rho_k^i)$; in each compact set of $\mathbb{R}$, $g_i(\cdot)$ has only finite number of discontinuities; moreover, denote the set of discontinuities by order $\{\rho_k^i :\ \rho_{k+1}^i > \rho_k^i,\ k \in Z\}$ and there exist positive constants $G_{i,k} > 0$, $i = 1, \cdots, n$, $k \in Z$ such that $|g_i(\xi) - g_i(\zeta)| \le G_{i,k}|\xi - \zeta|$ holds for all $\xi, \zeta \in (\rho_k^i, \rho_{k+1}^i)$.

Assumption $\mathbf{A_2}$: $d_i(t)$ and $a_{ij}(t)$ are all continuous functions, $i, j = 1, \cdots, n$ and $d_i(t) \ge \delta > 0$, $a_{ii}(t) < 0$ hold for all $i = 1, \cdots, n$ and $t \in \mathbb{R}$; for any $s \in \mathbb{R}$, the Lebesgue-Stieltjes measures $d_sK_{ij}(t, s) : t \mapsto d_sK_{ij}(t, s)$ are continuous, i.e., $\lim\limits_{h \to 0} \int_0^\infty |d_sK_{ij}(t +$

$h, s) - d_s K_{ij}(t, s)| = 0$ holds for all $i, j = 1, \cdots, n$ and $t \in \mathbb{R}$, and $d_s K_{ij}(t, s)$ is dominated by some Lebesgue-Stieltjes $d\bar{K}_{ij}(s)$ independent of t satisfying $\int_0^\infty e^{\delta s}|d\bar{K}_{ij}(s)| < +\infty$ for all $i, j = 1, \cdots, n$ and some $\delta > 0$. Here the domination means $|d_s K_{ij}(t, s)| \leq |d\bar{K}_{ij}(s)|$ i.e., $\int_0^\infty f(s)|d_s K_{ij}(t, s)| \leq \int_0^\infty f(s)|d\bar{K}_{ij}(s)|$ holds for all $t \geq 0$ and any nonnegative measurable function $f(\cdot)$; moreover, $d_i(t)$, $a_{ij}(t)$, $I_i(t)$, and $d_s K_{ij}(t, s)$ all possess almost periodic property, i.e., for any $\epsilon > 0$, there exists $l = l(\epsilon)$ such that for any interval $[\alpha, \alpha + l]$, there exists $\omega \in [\alpha, \alpha + l]$ such that

$$\begin{array}{ll} |d_i(t+\omega) - d_i(t)| < \epsilon & |a_{ij}(t+\omega) - a_{ij}(t)| < \epsilon \\ |I_i(t+\omega) - I_i(t)| < \epsilon & \displaystyle\int_0^\infty |d_s K_{ij}(t+\omega, s) - d_s K_{ij}(t, s)| < \epsilon \end{array}$$

hold for all $i, j = 1, \cdots, n$ and $t \in \mathbb{R}$;

Assumption $\mathbf{A_3}$: The initial condition satisfies that $\phi(\theta) \in C((-\infty, 0], \mathbb{R}^n)$ is bounded and $\lambda(\theta)$ is measurable and essentially bounded.

Throughout this paper, we use the following notations. We denote the solution of system (11) by $u(t, \phi, \lambda)$ ($u(t)$ for simplicity). We denote by $\|u\|_\xi$ a norm of vector $u = [u_1, \cdots, u_n]^\top \in \mathbb{R}^n$: $\|u\|_\xi = \sum_{i=1}^n \xi_i |u_i|$ where $\xi_i > 0$, $i = 1, \cdots, n$. $\mathcal{O}(V, \epsilon)$ denotes the open ϵ-neighborhood of a set $V \subset \mathbb{R}^n$: $\mathcal{O}(V, \epsilon) = \{y \in \mathbb{R}^n : \inf_{x \in V} \|y - x\| < \epsilon\}$ for some norm $\|\cdot\|$. $\mathbb{Z}$ denotes the integer set and $\mathbb{N}$ denotes the natural number set.

3 Main Results

In this section, we give the main results of the paper.

Main Theorem Suppose that the assumptions $A_1 - A_3$ are satisfied. If there exist constants $\xi_i > 0$, $i = 1, \cdots, n$, and $\delta > 0$ such that $d_i(t) \geq \delta$ and

$$\xi_i a_{ii}(t) + \sum_{j=1, j \neq i}^n \xi_j |a_{ji}(t)| + \sum_{j=1}^n \xi_j \int_0^\infty e^{\delta s}|d\bar{K}_{ji}(s)| < 0 \tag{12}$$

hold for all $t \geq 0$ and $i = 1, \cdots, n$, then,

1. for every initial value (ϕ, λ), the system (1) has a unique solution in the sense (11);

2. there exists a unique almost periodic solution $u^*(t)$ for the system (1), which is globally exponentially stable, i.e., for any other solution $u(t)$ with the initial condition (ϕ, λ), there exists a constant $M = M(\phi, \lambda) > 0$ such that

$$\|u(t) - u^*(t)\|_\xi \leq Me^{-\delta t}$$

holds for all $t \geq 0$.

We divide the proof of Main Theorem into several steps. First, we construct a sequence of delayed dynamical systems with high-slope continuous activations and prove the solutions are uniformly bounded. Second, based on the uniform boundedness and the compactness, we prove that for each initial value, the dynamical system (1) has at least a solution, which is a clustering point of the solution sequence of the delayed dynamical system sequence with high-slope activations converging to the discontinuous activations. Third, we prove that all solutions of the system (1) is globally exponentially asymptotically stable, which also implies that for any initial value, the solution is unique. Finally, we prove that the system (1) has a unique almost periodic solution, which is surely globally exponentially attractive.

Similar to the idea proposed in Haddad (1981), the solution of the system (1) in the sense (11) can be regarded as an approximation of the solutions of delayed neural networks with high-slope activations. This is the main idea in proving the existence of the solution and an almost periodic solution. More precisely, define a family of functions Ξ containing $f(x) = [f_1(x_1), f_2(x_2), \cdots, f_n(x_n)]^T \in C(\mathbb{R}^n, \mathbb{R}^n)$ satisfying

1. every $f_i(\cdot)$ is monotonically nondecreasing, for all $i = 1, 2, \cdots, n$;
2. every $f_i(\cdot)$ is uniformly locally bounded, i.e., for any compact set $Z \subset \mathbb{R}^n$, there exists a constant $M > 0$ independent of f such that $|f_i(x)| \leq M$ holds for all $x \in Z$ and $i = 1, \cdots, n$;

3. every $f_i(\cdot)$ is locally Lipschitzean continuous, i.e., for any compact set $Z \subset \mathbb{R}^n$, there exists $\lambda > 0$ such that $|f_i(\xi) - f_i(\zeta)| \le \lambda|\xi - \zeta|$ holds for all ξ, $\zeta \in Z$, and $i = 1, 2, \cdots, n$.

For any $f \in \Xi$, by the theory given in Hale (1977), the following system:

$$\begin{cases} \frac{du_i^f}{dt}(t) = -d_i(t)u_i^f(t) + \sum\limits_{j=1}^n a_{ij}(t)\sigma_j^f(t) \\ + \sum\limits_{j=1}^n \int_0^\infty \sigma_j^f(t-s)d_sK_{ij}(t,s) + I_i(t) \\ u_i^f(\theta) = \phi_i(\theta),\ \theta \in (-\infty, 0] \\ \sigma_i^f(\theta) = \begin{cases} \lambda_i(\theta), & \theta \le 0 \\ f_i(u_i^f(\theta)), & \theta \ge 0 \end{cases} \quad i = 1, \cdots, n, \end{cases}$$

admits a unique solution $u_f(t) = [u_1(t), u_2(t), \cdots, u_n(t)]^\top$ on $[0, T)$, where T might be ∞.

Step 1. we prove that the solutions $u^f(t)$ are uniformly bounded with respect to $f \in \Xi$.

Lemma 1 *Under the condition (12), the solutions $u^f(t)$ are uniformly bounded with respect to $f \in \Xi$, i.e. there exists $M = M(\phi, \lambda) > 0$, which is independent of $f \in \Xi$, such that $\|u^f(t)\|_\xi \le M$ holds for all $f \in \Xi$ and $t \ge 0$. Consequently, the existence interval of $u^f(t)$ can be extended to $[0, \infty)$.*

Proof: Define

$$V^f(t) = \sum_{i=1}^n \xi_i|u_i^f(t)|e^{\delta t} + \sum_{i,j=1}^n \xi_i \int_0^\infty \int_{t-s}^t |\sigma_j^f(\theta)|e^{\delta(s+\theta)}d\theta|d\bar{K}_{ij}(s)|$$

Differentiating it, we have

$$\begin{aligned}
\frac{d}{dt}V^f(t) &= \sum_{i=1}^{n}\delta e^{\delta t}\xi_i|u_i^f(t)| + \sum_{i=1}^{n}\xi_i e^{\delta t} sign(u_i^f(t))\Big\{ -d_i(t)u_i^f(t) \\
&+ a_{ii}(t)f_i(u_i^f(t)) + \sum_{j=1,j\neq 1}^{n} a_{ij}(t)f_j(u_j^f(t)) + \sum_{j=1}^{n}\int_0^\infty \sigma_j^f(t-s)d_sK_{ij}(t,s)\Big\} \\
&+ \sum_{i=1}^{n}\xi_i e^{\delta t} sign(u_i^f(t))I_i(t) + \sum_{i,j=1}^{n}\xi_i|f_j(u_j^f(t))|e^{\delta t}\int_0^\infty e^{\delta s}|d\bar{K}_{ij}(s)| \\
&- \sum_{i,j=1}^{n}\xi_j e^{\delta t}\int_0^\infty |\sigma_j^f(t-s)||\bar{K}_{ij}(s)| \\
&\le \sum_{i=1}^{n}\xi_i|u_i^f(t)|e^{\delta t}(-d_i(t)+\delta) + \sum_{i=1}^{n} e^{\delta t}|f_i(u_i^f(t))|\Big\{ a_{ii}(t)\xi_i \\
&+ \sum_{j=1,j\neq i}^{n}|a_{ji}(t)|\xi_j + \sum_{j=1}^{n}\xi_j\int_0^\infty e^{\delta s}|d\bar{K}_{ji}(s)|\Big\} + e^{\delta t}\hat{I} \\
&\le e^{\delta t}\hat{I}
\end{aligned}$$

where $\hat{I} = \sup_{t\ge 0}\|I(t)\|_\xi < +\infty$. It follows that

$$\begin{aligned}
\|u^f(t)\|_\xi &\le e^{-\delta t}V^f(t) = e^{-\delta t}[\int_0^t \dot{V}^f(s)ds + V^f(0)] \le e^{-\delta t}\int_0^t e^{\delta s}\hat{I}ds + e^{-\delta t}V^f(0) \\
&\le \frac{\hat{I}}{\delta}(1-e^{-\delta t}) + e^{-\delta t}V^f(0) < \frac{\hat{I}}{\delta} + V^f(0)
\end{aligned}$$

Noting that $V^f(0)$ is independent of $f \in \Xi$, we obtain the uniform boundedness of the solutions $u^f(t)$ by letting $M = \frac{\hat{I}}{\delta} + V^f(0)$. Moreover, $f(\cdot)$ is locally Lipschtiz continuous, we conclude that the existence interval of the solution $u^f(t)$ can be extended to the infinite interval $[0,+\infty)$ with the results given in Hale (1977). Lemma 1 is proved.

Now, for any sequence $\Big\{g^m(x) = [g_1^m(x_1),\cdots,g_n^m(x_n)]^\top\Big\}_{m\in\mathbb{N}} \in \Xi$ satisfying

$$\lim_{m\to\infty} d_H(Graph(g^m(K)), \overline{co}[g(K)]) = 0, \quad for\ all\ K\subset\mathbb{R}^n, \tag{13}$$

where $d_H(\cdot,\cdot)$ denotes the Hausdorff metric of $\mathbb{R}^n$, we construct a sequence of delayed systems with high-slope continuous activations as follows:

$$\begin{aligned}\frac{du_i^m(t)}{dt} &= -d_i(t)u_i^m(t) + \sum_{j=1}^n a_{ij}(t)\sigma_j^m(t) \\ &\quad + \sum_{j=1}^n \int_0^\infty \sigma_j^m(t-s)d_sK_{ij}(t,s) + I_i(t),\ i = 1, \cdots, n, \end{aligned} \tag{14}$$

where $u_i^m(\theta) = \phi_i(\theta)$, $\theta \in (-\infty, 0]$, and $\sigma_j^m(\theta) = \begin{cases} \lambda_j(\theta) & \theta \le 0 \\ g_j^m(u_j(\theta)) & \theta > 0 \end{cases}$. For instance, let $\{\rho_{k,i}\}$ be the set of discontinuous points of $g_i(\cdot)$. Pick a strictly decreasing sequence $\{\delta_{k,i,m}\}$ with $\lim\limits_{m\to\infty} \delta_{k,i,m} = 0$ and define $I_{k,i,m} = [\rho_{k,i} - \delta_{k.i.m}, \rho_{k,i} + \delta_{k,i,m}]$ such that for every $k_1 \ne k_2$, $I_{k_1,i,m} \bigcap I_{k_2,i,m} = \emptyset$ hold. Then, we define functions $g_i^m(\cdot)$ as follows:

$$g_i^m(s) = \begin{cases} g_i(s) & s \notin \bigcup\limits_{k\in\mathbb{Z}} I_{k,i,m}, \\ \frac{g_i(\rho_{k,i}+\delta_{k,i,m}) - g_i(\rho_{k,i}-\delta_{k,i,m})}{2\delta_{k,i,m}}[s - \rho_{k,i} - \delta_{k,i,m}] + g_i(\rho_{k,i} + \delta_{k,i,m}) & s \in I_{k,i,m}. \end{cases}$$

It can be seen that the sequence $\{g^m(\cdot)\}_{m\in\mathbb{N}} \subset \Xi$ satisfies condition (13).

Step 2. We will point out that the solution sequence of the system sequence (14) converges to a solution of the system (1) in the sense (11).

Lemma 2 *Under the assumptions of the Main Theorem, for each initial value pair* (ϕ, λ)*, the system (1) has a solution in the sense (11) on the whole time interval* $[0, \infty)$.

Proof: By lemma 1, we know that all the solutions $\{u^m(t)\}_{m\in\mathbb{N}}$ are uniformly bounded, which implies that $\{\dot{u}^m(t)\}_{m\in\mathbb{N}}$ is uniformly essentially bounded. By the Arzela-Ascoli lemma and diagonal selection principle, we can select a subsequence of $\{u^m(t)\}_{m\in\mathbb{N}}$ (still denoted by $u^m(t)$) such that $u^m(t)$ converges uniformly to a continuous function $u(t)$ on any compact interval of $\mathbb{R}$. Since $\{\dot{u}^m(t)\}_{m\in\mathbb{N}}$ is uniformly essentially bounded, for any $T > 0$, $u(t)$ is Lipschitz continuous on $[0, T]$. This implies that $\dot{u}(t)$ exists for almost all $t \in [0, T]$ and is bounded almost everywhere in $[0, T]$.

We claim that $\{\dot{u}^m(t)\}_{m\in\mathbb{N}}$ weakly converges to $\dot{u}(t)$ on the space $L^\infty([0, T], \mathbb{R}^n)$.

Because $C_0^\infty([0,T],\mathbb{R}^n$ is dense in the Banach space $L^1([0,T],\mathbb{R}^n)$ and is the conjugate space $L^\infty([0,T],\mathbb{R}^n)$). Therefore, for each $p(t)\in C_0^\infty([0,T],\mathbb{R}^n)$, we have

$$\int_0^T \langle \dot{u}^m(t)-\dot{u}(t),p(t)\rangle dt = -\int_0^T \langle \dot{p}(t),u^m(t)-u(t)\rangle dt.$$

By the uniform essential boundedness of $\{\dot{u}^m(t)\}_{m\in\mathbb{N}}$ and the Lebesgue dominant convergence theorem, we conclude that $\{\dot{u}^m(t)\}_{m\in\mathbb{N}}$ weakly converges to $\dot{u}(t)$ on the space $L^\infty([0,T],\mathbb{R}^n)$.

By the Mazur's convexity theorem (see page 120-123, Yoshida 1978), for any m, we can find finite number of constants $\alpha_l^m\geq 0$ satisfying $\sum\limits_{l=m}^{\infty}\alpha_l^m=1$, such that

$$\lim_{m\to\infty} y^m(t)=u(t),\ unifomly\ on\ [0,T],\quad \lim_{m\to\infty}\dot{y}^m(t)=\dot{u}(t),\ a.e.\ t\in[0,T],$$

where $y^m(t)=\sum\limits_{l=m}^{\infty}\alpha_l^m u^l(t)$.

Let $\eta_j^m(t)=\sum\limits_{l=m}^{\infty}\alpha_l^m\sigma_j^l(u_j(t))$. Then,

$$\begin{aligned}\dot{y}_i^m(t) &= -d_i(t)y_i^m(t)+\sum_{j=1}^n a_{ij}(t)\eta_j^m(t)\\ &\quad+\sum_{j=1}^n\int_0^\infty \eta_j^m(t-s)d_sK_{ij}(t,s)+I_i(t),\ i=1,\cdots,n.\end{aligned}$$

Let $\varphi^m(t)=\int_0^t\eta^m(s)ds$, which is absolutely continuous and has uniformly essentially bounded derivative. By the same arguments, we can find $\gamma^m(t)=\sum\limits_{l=m}^{\infty}\beta_l^m\eta^l(t)$ such that

$$\lim_{m\to\infty}\gamma^m(t)=\gamma(t),\ almost\ everywhere\ in\ t\in(-\infty,T],\ \gamma(t)\ is\ measurable.$$

Now, denoting $z^m(t)=\sum\limits_{l=m}^{\infty}\beta_l^m y^m(t)$, we have

$$\begin{aligned}\dot{z}_i^m(t) &= -d_i(t)z_i^m(t)+\sum_{j=1}^n a_{ij}(t)\gamma_j^m(t)\\ &\quad+\sum_{j=1}^n\int_0^\infty \gamma_j^m(t-s)d_sK_{ij}(t,s)+I_i(t),\ i=1,\cdots,n.\end{aligned}\tag{15}$$

Let $m \to \infty$, by Lebesgue dominant convergence theorem, we obtain

$$\begin{aligned}\dot{u}_i(t) &= -d_i(t)u_i(t) + \sum_{j=1}^{n} a_{ij}(t)\gamma_j(t) \\ &\quad + \sum_{j=1}^{n} \int_0^{\infty} \gamma_j(t-s) d_s K_{ij}(t,s) + I_i(t),\ i = 1, \cdots, n, a.e.\ t \in [0,T].\end{aligned}$$

The remaining is to prove $\gamma(t) \in \overline{co}[g(u(t))]$ on $t \in [0,T]$. Since $u^m(t)$ converges to $u(t)$ uniformly with respect to $t \in [0,T]$ and $\overline{co}[g(\cdot)]$ is an upper-semi-continuous set-valued map, for any $\epsilon > 0$, there exists $N > 0$ such that $g^m(u^m(t)) \in \mathcal{O}(\overline{co}[g(u(t))], \epsilon)$ holds for all $m > N$ and $t \in [0,T]$. Noting that $\overline{co}[g(u(t))]$ is convex and compact, we conclude $\gamma^m(t) \in \mathcal{O}(\overline{co}[g(u(t))], \epsilon)$, which implies $\gamma(t) \in \mathcal{O}(\overline{co}[g(u(t))], \epsilon)$ holds for any $t \in [0,T]$. Because of the arbitrariness of ϵ, we conclude that $\gamma(t) \in \overline{co}[g(u(t))],\ t \in [0,T]$. Since T is also arbitrary, the solution can be extended to $[0, \infty)$. This completes the proof.

Remark 1 *In the proof of Lemma 2, by the Arzela-Ascoli lemma and diagonal selection principle, we select a subsequence $u^{m_k}(t)$ of the sequence $u^m(t)$, which is uniformly convergent and $\dot{u}^{m_k}(t)$ is weakly convergent. But it is not enough to guarantee the convergence of $\dot{u}^{m_k}(t)$ almost everywhere. That is the reason why we need to cite the Mazur convexity theorem. By Mazur convexity theorem, we select a new subsequence which is a convex combination of the original sequence and prove that it converges almost everywhere. This technique is repeatedly used in the paper. For example, we use it to obtain the output $\gamma(t)$ of the activations by selecting a sequence $\gamma^m(t)$ and proving $\lim_{m\to\infty} \gamma^m(t) = \gamma(t)$.*

Step 3. We will point out that any solution of the system (1) in the sense (11) is asymptotically stable by lemma 3.

Lemma 3 *Suppose that the assumptions of the Main Theorem are satisfied. For any two solutions $u(t) = u(t, \phi, \lambda)$ and $v(t) = v(t, \psi, \chi)$ of the system (1) in the sense (11) associated with the outputs $\gamma(t)$ and $\mu(t)$ and initial value pairs (ϕ, λ) and (ψ, χ) respectively,* *there exists a constant $M = M(\phi,\psi,\lambda,\chi)$ satisfying that $M(\phi,\phi,\lambda,\lambda) = 0$ holds for all (ϕ,λ) such that*

$$\|u(t) - v(t)\|_{\xi} \leq Me^{-\delta t}$$

holds for all $t \geq 0$. Moreover, the solution of the system (1) in the sense (11) is unique.

Proof: Let $u(t) = [u_1(t), \cdots, u_n(t)]^\top$ be a solution of

$$\frac{d}{dt}u_i(t) \;=\; -d_i(t)u_i(t) + \sum_{j=1}^{n} a_{ij}(t)\gamma_j(t) + \sum_{j=1}^{n}\int_0^\infty \gamma_j(t-s)d_sK_{ij}(t,s) + I_i(t),$$

and $v(t) = [v_1(t), \cdots, v_n(t)]^\top$ be a solution of

$$\frac{d}{dt}v_i(t) \;=\; -d_i(t)v_i(t) + \sum_{j=1}^{n} a_{ij}(t)\mu_j(t) + \sum_{j=1}^{n}\int_0^\infty \mu_j(t-s)d_sK_{ij}(t,s) + I_i(t),$$

Then,

$$\begin{aligned}\frac{d}{dt}\Big[u_i(t) - v_i(t)\Big] \;&=\; -d_i(t)\Big[u_i(t) - v_i(t)\Big] + \sum_{j=1}^{n} a_{ij}(t)\Big[\gamma_j(t) - \mu_j(t)\Big] \\ &\quad + \sum_{j=1}^{n}\int_0^\infty \Big[\gamma_j(t-s) - \mu_j(t-s)\Big]d_sK_{ij}(t,s),\; i = 1,\cdots,n.\end{aligned}$$

Let

$$L_1(t) = \sum_{i=1}^{n}\xi_i|u_i(t) - v_i(t)|e^{\delta t} + \sum_{i,j=1}^{n}\xi_j\int_0^\infty\int_{t-s}^{t}|\gamma_j(\theta) - \mu_j(\theta)|e^{\delta(s+\theta)}d\theta|d\bar{K}_{ij}(s)|$$

and $M = M(\phi,\psi,\lambda,\chi) = L_1(0)$. By the chain rule (see Clarke 1983 or Lu & Chen 2005a for details), differentiating it gives

$$\begin{aligned}
\frac{d}{dt}L_1(t) &= \sum_{i=1}^{n}\delta e^{\delta t}\xi_i|u_i(t)-v_i(t)| + \sum_{i=1}^{n}\xi_i e^{\delta t} sign\Big(u_i(t)-v_i(t)\Big) \\
&\quad \Big\{-d_i(t)[u_i(t)-v_i(t)] + a_{ii}(t)[\gamma_i(t)-\mu_i(t)] + \sum_{j=1,j\neq 1}^{n} a_{ij}(t)[\gamma_j(t)-\mu_j(t)] \\
&+ \sum_{j=1}^{n}\int_0^\infty [\gamma_j(t-s))-\mu_j(t-s)]d_sK_{ij}(t,s)\Big\} + \sum_{i,j=1}^{n}\xi_i|\gamma_j(t)-\mu_j(t)| \\
&\quad e^{\delta t}\int_0^\infty e^{\delta s}|d\bar{K}_{ij}(s)| - \sum_{i,j=1}^{n}\xi_j e^{\delta t}\int_0^\infty |\gamma_j(t-s)-\mu_j(t-s)||\bar{K}_{ij}(s)| \\
&\le \sum_{i=1}^{n}\xi_i|u_j(t)-v_j(t)|e^{\delta t}(-d_i(t)+\delta) + \sum_{i=1}^{n} e^{\delta t}|\gamma_j(t)-\mu_j(t)|\Big\{a_{ii}(t)\xi_i \\
&+ \sum_{j=1,j\neq i}^{n}|a_{ji}(t)|\xi_j + \sum_{j=1}^{n}\xi_j\int_0^\infty e^{\delta s}|d\bar{K}_{ji}(s)|\Big\} \\
&\le 0,
\end{aligned}$$

which implies $\|u(t)-v(t)\|_\xi \le L_1(0)e^{-\delta t} = M(\phi,\psi,\lambda,\chi)e^{-\delta t}$. It is clear that $M(\phi,\phi,\lambda,\lambda) = 0$. Therefore, the solution in unique. Lemma 3 is proved.

In lemma 2, we have proved that some subsequence of $u^m(t)$ converges to the solution $u(t)$. In fact, we can prove that $u^m(t)$ itself converges to the solution $u(t)$.

Proposition 1 *Suppose that the assumptions of the Main Theorem are satisfied. For any function sequence* $\{\tilde{g}^m(x) = (\tilde{g}_1^m(x_1),\cdots,\tilde{g}_n^m(x_n))^T : m = 1,2,\cdots\} \subset \Xi$ *satisfying the condition (13) on any compact set in* $\mathbb{R}^n$*, let* $\tilde{u}^m(t) = [\tilde{u}_1^m(t),\cdots,\tilde{u}_n^m(t)]^\top$ *be the solution of the following system:*

$$\begin{aligned}
\frac{d\tilde{u}_i^m}{dt} &= -d_i(t)\tilde{u}_i^m(t) + \sum_{j=1}^{n} a_{ij}(t)\tilde{g}_j(\tilde{u}_j^m(t)) + \sum_{j=1}^{n}\int_0^\infty \tilde{\sigma}_j^m(t-s)d_sK_{ij}(t,s) + I_i(t), \\
\tilde{u}_i^m(\theta) &= \phi_i(\theta),\ \theta\in[-\infty,0],\ \tilde{\sigma}_i^m(\theta) = \begin{cases} \lambda_i(\theta) & \theta \le 0 \\ \tilde{g}_i^m(\tilde{u}_i^m(\theta)) & \theta \ge 0 \end{cases},\ i = 1,\cdots,n, \qquad (16)
\end{aligned}$$

and $u(t)=u(t,\phi,\lambda)$ be the solution of the delayed system (1) in the sense (11) with initial value (ϕ,λ). Then, $\tilde{u}^m(t)$ uniformly converges to $u(t)$ on any finite time interval $[0,T]$.

Proof: First, we prove that $u^m(t)$ converges to the solution of the delayed system (1) in the sense (11) by reduction to absurdity. Assume that there exist $T>0$, $\epsilon_0\geq 0$, and an subsequence of integers $\{m_k\}_{k\in\mathbb{N}}$ such that

$$\max_{t\in[0,T]}|u^{m_k}(t)-u(t)|\geq\epsilon_0. \tag{17}$$

By the same arguments used in the proof of lemma 2, we can select a subsequence $\{u^{m_{k_l}}\}_{l\geq 0}$ of $\{u^{m_k}\}_{k\geq 0}$, which converges to a solution $v(t)=v(t,\phi,\lambda)$ of the delayed system (1) in the sense (11) uniformly in any finite interval $[0,T]$ with the initial value (ϕ,λ). By lemma 3, $u(t)=v(t)$, which leads a contradiction with (17). This completes the proof.

Remark 2 *Proposition 1 indicates that the solution $v(t)=v(t,\phi,\lambda)$ of the delayed system (1) in the sense (11) does not depend on the choice of the sequence $\{g^m(x)\}_{m\in\mathbb{N}}\subset\Xi$ satisfying the condition (13).*

The following lemma points out that any solution is asymptotically almost periodic (Yoshizawa 1975).

Lemma 4 *Suppose that the assumptions of the Main Theorem are satisfied. Let $u(t,\phi,\lambda)$ be a solution of the system (1) in the sense (11). For any $\epsilon>0$, there exist $T>0$ and $l=l(\epsilon)$ such that any interval $[\alpha,\alpha+l]$ contains an ω such that*

$$\|u(t+\omega)-u(t)\|_\xi\leq\epsilon$$

holds for all $t\geq T$.

Proof: We introduce the following auxiliary functions

$$\begin{aligned}\epsilon_i(t,\omega) &= u_i(t+\omega)[d_i(t+\omega)-d_i(t)]+\sum_{j=1}^n\gamma_j(t+\omega)[a_{ij}(t+\omega)-a_{ij}(t)]\\ &+ \int_0^\infty\sum_{j=1}^n\gamma_j(t+\omega-s)d[K_{ij}(t+\omega,s)-K_{ij}(t,s)]+I_i(t+\omega)-I_i(t)\end{aligned} \tag{18}$$

From the assumption A_2 and the boundedness of $u(t)$ and $\gamma(t)$, one can see that for any $\epsilon > 0$, there exists $l = l(\epsilon) > 0$ such that every interval $[\alpha, \alpha + l]$ contains at least one number ω with $\sum_{i=1}^{n} \xi_i |\epsilon_i(t, \omega)| < \frac{\delta}{2}\epsilon$ for all $t \geq 0$.

Denote $z(t) = u(t + \omega) - u(t)$. Then,

$$
\begin{aligned}
\frac{dz_i(t)}{dt} &= -d_i(t)z_i(t) + \sum_{j=1}^{n} a_{ij}(t)[\gamma_j(t+\omega) - \gamma_j(t)] \\
&\quad + \sum_{j=1}^{n} \int_0^\infty [\gamma_j(t+\omega-s) - \gamma_j(t-s)] d_s K_{ij}(t,s) + \epsilon_i(t,\omega).
\end{aligned}
$$

Let

$$
L_2(t) = \sum_{i=1}^{n} \xi_i |z_i(t)| e^{\delta t} + \sum_{i,j=1}^{n} \xi_i \int_0^\infty \int_{t-s}^{t} |\gamma_j(\theta+\omega) - \gamma_j(\theta)| e^{\delta(\theta+s)} d\theta |d\bar{K}_{ij}(s)|.
$$

Pick a sufficiently large T such that $e^{-\delta t} L_2(0) < \frac{\epsilon}{2}$ holds for all $t \geq T$. Differentiating $L_2(t)$ gives

$$
\begin{aligned}
\frac{dL_2(t)}{dt} &= \sum_{i=1}^{n} \xi_i \delta e^{\delta t} |z_i(t)| + \sum_{i=1}^{n} \xi_i e^{\delta t} sign(z_i(t)) \Big\{ - d_i(t) z_i(t) \\
&+ a_{ii}(t)[\gamma_i(t+\omega) - \gamma_i(t)] + \sum_{j=1, j\neq i} a_{ij}(t)[\gamma_j(t+\omega) - \gamma_j(t)] \\
&+ \sum_{j=1}^{n} \int_0^\infty [\gamma_j(t+\omega-s) - \gamma_j(t-s)] d_s K_{ij}(t,s) \\
&+ \epsilon_i(t,\omega) \Big\} + \sum_{i,j=1}^{n} \xi_i e^{\delta t} |\gamma_j(t+\omega) - \gamma_j(t)| \int_0^\infty e^{\delta s} |d\bar{K}_{ij}(s)| \\
&- \sum_{i,j=1}^{n} \xi_i e^{\delta t} \int_0^\infty |\gamma_j(t+\omega-s) - \gamma_j(t-s)| |d\bar{K}_{ij}(s)| \\
&\leq \sum_{i} \xi_i e^{\delta t} |z_i(t)| (-d_i(t) + \delta) + \sum_{i=1}^{n} |\gamma_j(t+\omega) - \gamma_j(t)| e^{\delta t} \Big\{ \xi_i a_{ii}(t) \\
&+ \sum_{j=1, j\neq i} \xi_j |a_{ji}(t)| + \sum_{j=1}^{n} \xi_j \int_0^\infty e^{\delta s} |d\bar{K}_{ji}(s)| \Big\} + \sum_{i=1}^{n} \xi_i e^{\delta t} |\epsilon_i(t,\omega)| \\
&\leq e^{\delta t} \frac{\delta}{2} \epsilon, \ t \geq T.
\end{aligned}
$$

Therefore,

$$\begin{aligned}\sum_{i=1}^{n}\xi_i|z_i(t)| &\leq e^{-\delta}L_2(t) = e^{-\delta}\left[L_2(0) + \int_0^t \dot{L}_2(s)ds\right] \\ &\leq e^{-\delta t}L_2(0) + e^{-\delta t}\int_0^t e^{\delta s}ds\frac{\delta}{2}\epsilon < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\end{aligned}$$

holds for all $t \geq T$, which competes the proof.

Step 4. Now, we are to prove that the system (1) has at least an almost periodic solution in the sense (11).

Lemma 5 *Under the assumptions of the Main Theorem, the system (1) has at least one almost periodic solution in the sense (11).*

Proof: Let $u(t) = u(t, \phi, \lambda)$ be a solution of system (11). Pick a sequence $\{t_k\}_{k\in\mathbb{N}}$ satisfying $\lim_{k\to\infty} t_k = \infty$ and $\sup_{t\geq 0}\sum_{i=1}^{n}\xi_i|\epsilon_i(t, t_k)| \leq \frac{1}{k}$, where $\epsilon_i(t, t_k)$, $i = 1, \cdots, n$, are the auxiliary functions (18) defined in the proof of lemma 4.

Let $u^k(t) = u(t + t_k)$ and $\gamma^k(t) = \gamma(t + t_k)$. It is clear that the sequence $\{u(t + t_k)\}_{k\in\mathbb{N}}$ is uniformly continuous and bounded. By the Arzela-Ascoli lemma and diagonal selection principle, we can select a subsequence of $u(t + t_k)$ (still denoted by $u(t + t_k)$), which converges to some absolutely continuous function $u^*(t)$ uniformly on any compact interval $[0, T]$.

In the following, we will prove that $u^*(t)$ is an almost periodic solution of system (1) in the sense (11). First, we prove that $u^*(t)$ is a solution of the system (1) in the sense (11). With the notations above, we have

$$\begin{aligned}&\frac{du_i(t + t_k)}{dt} = -d_i(t)u_i(t + t_k) + \sum_{j=1}^{n} a_{ij}(t)\gamma_j(t + t_k) \\ &+ \sum_{j=1}^{n}\int_0^{\infty}\gamma_j(t + t_k - s)d_sK_{ij}(t, s) + I_i(t) + \epsilon_i(t, t_k),\ i = 1, \cdots, n.\end{aligned}$$

With the similar method used in the proof of lemma 2, we can select a subsequence from $u(t+t_k)$ (still denoted by $u(t+t_k)$) and constants $\nu_l^k \geq 0$ with finite $\nu_l^k > 0$ satisfying $\sum\limits_{l=k}^{\infty} \nu_l^k = 1$ such that

1. $v^k(t) = \sum\limits_{l=k}^{\infty} \nu_l^k u(t+t_l)$ converges to a Lipschitz continuous function $u^*(t)$ uniformly on $[0,T]$; $\{\dot{v}^k(t)\}$ converges to $\dot{v}^*(t)$ for almost all $t \in [0,T]$.

2. $\zeta^k(t) = \sum\limits_{l=k}^{\infty} \nu_l^k \gamma(t+t_l)$ converges to a measurable function $\zeta(t)$ for almost all $t \in [0,T]$.

Moreover, for each k, we have

$$\begin{aligned}\frac{dv_i^k(t)}{dt} &= -d_i(t)v_i^k(t) + \sum_{j=1}^{n} a_{ij}(t)\zeta_j^k(t) \\ &\quad + \sum_{j=1}^{n} \int_0^{\infty} \zeta_j^k(t-s) d_s K_{ij}(t,s) + I_i(t) + \bar{\epsilon}_i(t,k),\ i = 1, \cdots, n,\end{aligned}$$

where $\bar{\epsilon}_i(t,k) = \sum\limits_{l=k}^{\infty} \nu_l^k \epsilon_i(t,t_k)$. Letting $k \to \infty$, we have

$$\begin{aligned}\frac{du_i^*(t)}{dt} &= -d_i(t)u_i^*(t) + \sum_{j=1}^{n} a_{ij}(t)\zeta_j(t) \\ &\quad + \sum_{j=1}^{n} \int_0^{\infty} \zeta_j(t-s) d_s K_{ij}(t,s) + I_i(t),\ i = 1, \cdots, n.\end{aligned}$$

Repeating the proof of lemma 2, we can prove $\zeta(t) \in \overline{co}[g(u^*(t))]$, which means that $u^*(t)$ is a solution of the system (1) in the sense (11).

Second, we prove that $u^*(t)$ is almost periodic. By lemma 4, for any $\epsilon > 0$, there exist $K > 0$ and $l = l(\epsilon)$ such that each interval $[\alpha, \alpha + l]$ contains an ω such that

$$\|u(t+t_k+\omega) - u(t+t_k)\|_{\xi} < \epsilon$$

holds for all $k \geq K$ and $t \geq 0$. As $k \to \infty$, we conclude that $\|u^*(t+\omega) - u^*(t)\|_{\xi} < \epsilon$ holds for all $t \geq 0$. This implies that $u^*(t)$ is an almost periodic function. The proof is completed.

Proof of Main Theorem: By lemma 5, we know that there exists an almost periodic solution for the system (1) in the sense (11). By lemma 3, we have

$$\|u(t) - u^*(t)\|_\xi = O(e^{-\delta t}). \tag{19}$$

Finally, we prove that the almost periodic solution of the system (1) is unique. In fact, suppose that $u^*(t)$ and $v^*(t)$ are two almost periodic solutions of the system (1). Applying lemma 3 again gives

$$\|v^*(t) - u^*(t)\|_\xi = O(e^{-\delta t}). \tag{20}$$

From Levitan & Zhikov (1982), one can see that if $u^*(t)$ and $v^*(t)$ are two almost periodic functions satisfying (20), then $v^*(t) = u^*(t)$. Therefore, the almost periodic solution of the system (1) is unique. Theorem 1 is proved.

Remark 3 *Main Theorem in this paper has close relation and essential difference from the previous works (Lu & Chen 2005a, 2005b, Lu & Chen 2006).*

1. *In case that the activations are Lipschitz continuous, the results were established by Lu & Chen (2005b). In this paper, we generalize the results to the case when the Lipschtiz constants tend to infinite.*

2. *Different from our previous paper Lu & Chen (2006), the model discussed in this paper is universal, which unifies discrete delays and distribution delays, in particular, includes the model studied in Lu & Chen (2006) and many others. Moreover, in this paper, we discuss almost periodicity, which includes the stability of the equilibrium studied by Lu & Chen (2006) and periodicity studied in Papini et.al (2006) as special cases. We also investigate the uniqueness of the almost solution, which was not concerned in Lu & Chen (2006). In this sense, this paper is an advance of the previous works (Forti & Nistri 2003, Forti et.al. 2005, Lu & Chen 2005a, 2006).*

4 Applications of the main result

In this section, we discuss the following delayed dynamical systems as specific cases in the Main Theorem.

Case 1. $d_sK_{ij}(t,s) = b_{ij}(t)\delta(s-\tau_{ij})$, where $\delta(s)$ denotes the Dirac-δ function. Then, the delayed system (1) becomes a system with discrete delays:

$$\begin{aligned}\frac{du_i(t)}{dt} &= -d_i(t)u_i(t) + \sum_{j=1}^{n} a_{ij}(t)g_j(u_j(t)) \\ &\quad + \sum_{j=1}^{n} b_{ij}(t)g_j(u_j(t-\tau_{ij})) + I_i(t),\ i=1,\cdots,n. \end{aligned} \tag{21}$$

In this case, the assumptions with $d_sK_{ij}(t,s)$ can be simplified as that all $b_{ij}(t)$, $i,j = 1,\cdots,n$, are almost-periodic functions.

Case 2. $d_sK_{ij}(t,s) = b_{ij}(t)k_{ij}(s)ds$. In this case, the delayed system (1) becomes the following system with distributed delays:

$$\begin{aligned}\frac{du_i(t)}{dt} &= -d_i(t)u_i(t) + \sum_{j=1}^{n} a_{ij}(t)g_j(u_j(t)) \\ &\quad + \sum_{j=1}^{n} b_{ij}(t)\int_0^{\infty} k_{ij}(s)g_j(u_j(t-s))ds + I_i(t),\ i=1,\cdots,n. \end{aligned} \tag{22}$$

As Direct consequences of Main Theorem, we have

Corollary 1 *Suppose that $a_{ij}(t)$, $b_{ij}(t)$ and $I_i(t)$ are continuous almost periodic functions and the activations satisfy assumption A_1. If there exist positive constants ξ_i, $i=1,\cdots,n$, and $\delta>0$ such that $d_i(t)\geq\delta$ and*

$$\xi_i a_{ii}(t) + \sum_{j=1, j\neq i}^{n} \xi_j|a_{ji}(t)| + \sum_{j=1}^{n}\xi_j b_{ji}^{*} < 0 \tag{23}$$

hold for all $t\geq 0$ and $i=1,\cdots,n$, where $b_{ji}^{*} = \max_{t\in\mathbb{R}}|b_{ji}(t)|$, then, *(i). for any initial data satisfying assumption A_3, the dynamical system with discrete delays (21) has a unique solution in the sense (11);*

(ii). the system (21) has a unique almost periodic solution $u^*(t)$, which is globally exponentially stable.

Corollary 2 *Suppose that $a_{ij}(t)$, $b_{ij}(t)$, $I_i(t)$ are almost periodic functions and the activations satisfy assumption A_1. If there exist positive constants ξ_i, $i = 1, \cdots, n$, and $\delta > 0$ such that $d_i(t) \geq \delta$ and*

$$\xi_i a_{ii}(t) + \sum_{j=1, j\neq i}^{n} \xi_j |a_{ji}(t)| + \sum_{j=1}^{n} \xi_j b_{ji}^* \int_0^\infty e^{\delta s} |k_{ji}(s)| ds < 0 \tag{24}$$

hold for all $t \geq 0$ and $i = 1, \cdots, n$, then,

(i). for any initial data with assumption A_3, the dynamical system with distributed delays (22) has a unique solution in the sense (11);

(ii). the system (22) has a unique almost periodic solution $u^*(t)$, which is globally exponentially stable.

Since any periodic function can be regarded as an almost-periodic function, all the results apply to periodic case. Now, replacing assumption A_2, we assume the following assumption.

Assumption $\mathbf{A_4}$: $d_i(t)$, $a_{ij}(t)$ are all continuous functions, $i, j = 1, \cdots, n$ and $d_i(t) \geq d_i > 0$, $a_{ii}(t) < 0$ hold for all $i = 1, \cdots, n$ and $t \in \mathbb{R}$; for any $s \in \mathbb{R}$, $d_s K_{ij}(t, s) : t \mapsto d_s K_{ij}(t, s)$ is continuous with respect to $t \in \mathbb{R}$, i.e., $\lim_{h\to 0} \int_0^\infty |dK_{ij}(t+h, s) - d_s K_{ij}(t, s)| = 0$ hold for all $i, j = 1, \cdots, n$ and $t \in \mathbb{R}$, $d_s K_{ij}(t, s) : s \mapsto d_s K_{ij}(t, s)$ is a Lebesgue-Stieltjes measure, $i, j = 1, \cdots, n$, satisfying $|d_s K_{ij}(t, s)| \leq |d\bar{K}_{ij}(s)|$, where there exists $\delta > 0$ such that $\int_0^\infty e^{\delta s} |d\bar{K}_{ij}(s)| < +\infty$; moreover, there exists $\omega > 0$ such that

$$d_i(t + \omega) = d_i(t) \quad a_{ij}(t + \omega) = a_{ij}(t)$$
$$I_i(t + \omega) = I_i(t) \quad dK_{ij}(t + \omega, s) = d_s K_{ij}(t, s)$$

hold for all $i, j = 1, \cdots, n$ and $t \in \mathbb{R}$.

Thus, we have

Corollary 3 *Suppose that the discontinuous activations satisfy assumptions A_1 and the assumption A_4 is satisfied. If there exist positive constants ξ_i, $i = 1, \cdots, n$, and $\delta > 0$ such that $d_i(t) \geq \delta$ and*

$$\xi_i a_{ii}(t) + \sum_{j=1, j\neq i}^{n} \xi_j |a_{ji}(t)| + \sum_{j=1}^{n} \xi_j \int_0^\infty e^{\delta s} |d\bar{K}_{ji}(s)| < 0$$

hold for all $t \geq 0$ and $i = 1, \cdots, n$, then,

(i). for each initial data with assumption A_3, the system (1) has a unique solution in the sense (11);

(ii). there exists a unique periodic solution $u^*(t)$ for system (1), which is globally exponentially stable.

Furthermore, a constant can be regarded as a periodic function with any period. Therefore, for the following delayed system

$$\begin{aligned} \frac{du_i(t)}{dt} &= -d_i u_i(t) + \sum_{j=1}^{n} a_{ij} g_j(u_j(t)) \\ &\quad + \sum_{j=1}^{n} \int_0^\infty g_j(u_j(t-s)) d_s K_{ij}(s) + I_i, \quad i = 1, \cdots, n, \end{aligned} \tag{25}$$

we have

Corollary 4 *Suppose that the discontinuous activations satisfy assumptions A_1. If there exist positive constants ξ_i, $i = 1, \cdots, n$, and $\delta > 0$ such that $d_i \geq \delta$ and*

$$\xi_i a_{ii} + \sum_{j=1, j\neq i}^{n} \xi_j |a_{ji}| + \sum_{j=1}^{n} \xi_j \int_0^\infty e^{\delta s} |d\bar{K}_{ji}(s)| \leq 0$$

hold for all $t \geq 0$ and $i = 1, \cdots, n$, then, *(i). for each initial data with the assumption A_3 , the system (25) has a unique solution in sense of (11);*

(ii). the system (25) has a unique equilibrium u^*, which is globally exponentially stable.

As for the system

$$\begin{aligned} \frac{du_i(t)}{dt} &= -d_i u_i(t) + \sum_{j=1}^{n} a_{ij} g_j(u_j(t)) \\ &\quad + \sum_{j=1}^{n} b_{ij} g_j(u_j(t-\tau_{ij})) + I_i(t),\ i = 1, \cdots, n, \end{aligned} \tag{26}$$

which has been studied by Lu & Chen (2006) and Forti et.al. (2005) with $\tau_{ij} = \tau$, $i, j = 1, \cdots, n$, we have

Corollary 5 *Suppose that the discontinuous activations satisfy assumptions A_1. If there exist positive constants ξ_i, $i = 1, \cdots, n$, and $\delta > 0$ such that $d_i(t) \geq \delta$ and*

$$\xi_i a_{ii} + \sum_{j=1, j\neq i}^{n} \xi_j |a_{ji}| + \sum_{j=1}^{n} \xi_j |b_{ji}| < 0 \tag{27}$$

holds for all $t \geq 0$ and $i = 1, \cdots, n$, then,

(i). for each initial data with the assumption A_3 , the system (26) has a unique solution in sense of (11);

(ii). the system (26) has a unique equilibrium u^*, which is globally exponentially stable.

5 Conclusions

In this paper, we study the almost periodic dynamics of a class of delayed integro-differential systems with discontinuous activations. This class of delayed differential systems include delayed Hopfield and cellular neural networks with discontinuous activations and almost periodic coefficients as well as periodic or constant coefficients as special cases. We prove that under some diagonal dominant conditions, this system has a unique almost periodic solution, which is globally exponential stable. As direct consequences, we obtain several results for the systems with periodic and constant coefficients.

Acknowledgements

We are very grateful to the anonymous reviewers' comments, which are very helpful for us in revising the paper. This project is supported by NSF of China under Grant 60374018, 60574044.

References

Andres, J. (1999). Almost-periodic and bounded solutions of Carathéodory differential inclusions. Differential Integral Equations, 12:6, 887–912.

Aubin J. P. & Cellina, A. (1984). Differential Inclusions, Springer-Verlag, Berlin.

Aubin, J. P. (1991). Viability theory. Birkhäuser, Boston.

Bacciotti, A. et.al. (2000). Discontinuous Ordinary Differential Equations and Stabilization. *Tesi di Dottorato di Ricerca in Mathematica, Universita degli STUDI di Firenze*.

Bacciotti, A., (2005). Generalized solutions of differential inclusions and stability. Ital. J. Pure Appl. Math., 17, 183-192.

Bader R. & Kryszewski, W. (2003). On the solution sets of differential inclusions and the periodic problem in Banach space. Nonlinear Analysis, 54, 707-754.

Belair, J., Campbell, S. A., & Driessche, P. (1996). Stability and delay-induced oscillations in a neural network model. SIAM J. Appl. Math., 56, 245-255.

Cao J. & Zhou, D. (1996). Stability analysis of delayed neural networks, Neural Networks, 11, 1601-1605.

Cao, J. (2003). New results Concerning exponential stability and Periodic solutions of delayed cellular neural networks. Phys. Lett. A, 307, 136-147.

Chen, T. P. (2001). Global Exponential Stability of Delayed Hopfield Neural Networks. Neural Networks, 14:8, 977-980.

Chen, T. P., Lu, W., & Chen, G. (2005). Dynamical behaviors of a large class of general delayed neural networks. Neural Computation, 17, 949-968.

Chua, L. O., Desoer, C. A., & Kuh, E. S. (1987). Linear and nonlinear Circuits. New York, Macgraw-Hill.

Civalleri, P. P., Gilli, L. M., & Pandolfi, L. (1993). On stability of cellular neural networks with delay. IEEE Trans. Circuit Syst., 40,157-164.

Clarke, F. H. (1983). Optimization and nonsmooth analysis, Wiley, New York.

Cohen, M. A. & Grossberg, S. (1983). Absolute stability and global pattern formation and parallel memory storage by competitive neural networks. IEEE Transactions on Syst., Man Cybern., SMC-13, 815-821.

Dhage, B. C., (2006). Existence of extremal solutions for discontinuous functional integral equations. Appl. Math. Lett., 19, 881-886.

Filippakis, M. E. & Papageorgiou, N. S. (2006). Periodic solutions for differential inclusions in $\mathbb{R}^{\mathbb{N}}$. Arch. Math. (Brno), 42, 115-123.

Filippov, A. F. (1967). Classical solution of differential equations with multivalued right-hand side. SIAM J. Control, 5:4, 609-621.

Forti, M. & Tesi, A. (1995). New condition for global stability of neural networks with application to linear quadratic programming problems. IEEE Trans. Circuits Syst.-1, 42, 354-366.

Forti, M. & Nistri, P. (2003). Global convergence of neural networks with discontinuous neuron activations. IEEE Trans. Circuits Syst.I. 50:11, 1421-1435.

Forti, M., Nistri, P., & Papini, D. (2005). Global exponential stability and global convergence of delayed neural networks with infinite gain. IEEE Trans. Neural networks 16(6), 1449-1463.

Gopalsamy, K. & He, X. (1991). Stability in asymmetric Hopfield nets with transmission delays. Phys. D. 76, 344-358.

Gopalsamy, K. & Sariyasa, (2001). Time delays and stimulus-dependent pattern formation in periodic environments in isolated neurons. IEEE Trans. Neural Networks, 13:2, 551-563.

Ivanov, A. G. (1997). On the equivalence of differential inclusions and controlled almost periodic systems. (Russian) Differ. Uravn. 33:7 (1997), 876–884; translation in Differential Equations, 33:7, 879–887.

Haddad, G. (1981). Monotone viable trajectories for functional differential inclusions. J. Diff. Equ., 42, 1-24.

Hale, J. (1977). Theory of functional differential equations, Springer-Verlag, New York.

Harrer, H., Nossek, J. A., & Stelzl, R. (1992). An analog implementation of discrete-time neural networks. IEEE Trans. Neural Networks, 3 , 466-476.

Hirsch, M. W. (1989). Convergent activation dynamics in continuous time networks. Neural Networks, 2, 331-349.

Hopfield, J. J. & Tank, D. W. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. Proc. Nat. Acad. Sci., 79, 3088-3092.

Hopfield, J. J. & Tank, D.W. (1986). Computing with neural circuits:Amodel. Science, 233, 625-633.

Hu, S. & Papageorgiou, N. S. (1995). On the existence of perioidic solutions for nonconvex-valued differential inclusions in $\mathbb{R}^N$. Proc. Amer. Math. Soc., 123:10, 3043-3050.

Huang, X. & Cao, J. (2003). Almost periodic solution of shunting inhibitory cellular neural networks with time-varying delays. Phys. Lett. A, 314:3, 222-231.

Joy, M. P. (2000). Results concerning the absolute stability of delayed neural networks. Neural Networks, 13, 613-616.

Kennedy, M. P. & Chua, L. O. (1988). Neural networks for nonlinear programming. IEEE Trans. Circuits Syst.-1., 35, 554-562.

Levitan, B. M. & Zhikov, V. V. (1982). Almost periodic functions and differential equations, Cambridge University Press, New York.

Li D. & Kloeden, P. E. (2006). On the dynamics of nonautonomous periodic general dynamical systems and differential inclusions. J. Diff. Equ., 224, 1-38.

Li, G & Xue, X. (2002). On the existence of periodic solutions for differential inclusions. J. Math. Anal. Appl., 276, 168-183.

Lu, W. & Chen, T. P. (2003). New Conditions on Global Stability of Cohen-Grossberg Neural Networks. Neural Computation, 15:5, 1173-1189.

Lu, W., Rong, L., & Chen, T. P. (2003). Global convergence of delayed neural networks systems. Int. J. Neural Networks, 13:3, 193-204.

Lu, W. & Chen, T. P. (2004). On periodic dynamical systems. Chin. Ann. Math., 25B:4, 455-462.

Lu, W. & Chen, T. P. (2005a). Dynamical behaviors of Cohen-Grossberg neural networks with discontinuous activation functions. Neural Networks. 18, 231-242.

Lu, W. & Chen, T. P. (2005b). Global exponential stability of almost periodic solution for a large class of delayed dynamical systems. Sci. Chin. Ser. A Math., 48:8, 1015-1026.

Lu, W. & Chen, T. P. (2006) Dynamical Behaviors of delayed neural network systems with discontinuous activation functions. Neural Computation. 18:3, 683-708.

Michel, A. N. & Gray, D. L. (1990). Analysis and synthesis of neural networks with lower block triangular interconnecting structure. IEEE Trans. Circuits Syst., 37, 1267-1283.

Paden B. E. & Sastry, S. S. (1987). Calculus for computing Filippov's differential inclusion with application to the variable structure control of robot manipulator. IEEE Trans. Circuits Syst., 34, 73-82.

Papini, D. & Taddei, V. (2005). Global exponential stability of the periodic solution of the delayed neural networks with discontinuous activations. Physics Letters A. 343, 117-128.

Petrovskaya, N. B. (2006). On oscillations in discontinuous Galerkin discretization schemes for steady problems. SIAM J. Sci. Comput., 27:4, 1329-1346.

Smirnov, G. V. (1995). Weak asymptotic stability at first approximation for periodic differential inclusions. NoDEA, 2, 445-461.

Utkin, V. I. (1977) Variable structure systems with sliding modes. IEEE Trans. Automat. Contr., AC-22:2, 212-222.

Yoshida, K. (1978). Functional analysis. Grundlehren der Mathematicchen Wissenschaften, Springer-Verlag, New York.

Yoshizawa, T. (1975). Stability theory and the existence of periodic solutions and almost periodic solutions, Springer-Verlag, New York.

Zheng, Y. & Chen, T. (2004). Global exponential stability of delayed periodic dynamical systems. Phys. Lett. A, 322:5-6, 344-355.

Zhou, J., Liu, Z., & Chen, G. (2004). Dynamics of delayed periodic neural networks. Neural Networks, 17:1, 87-101.

Zuev, A. V. (2006). On periodic solutions of ordinary differential equations with discontinuous right-hand side. Math. Notes, 79:4, 518-527.