Dynamic Analysis of a General Class of Winner-Take-All Competitive Neural Networks

Yuguang Fang, *Fellow, IEEE,* Michael A. Cohen, and Thomas G. Kincaid, *Member, IEEE*

Abstract—This paper studies a general class of dynamical neural networks with lateral inhibition, exhibiting winner-take-all (WTA) behavior. These networks are motivated by a metal–oxide–semiconductor field effect transistor (MOSFET) implementation of neural networks, in which mutual competition plays a very important role. We show that for a fairly general class of competitive neural networks, WTA behavior exists. Sufficient conditions for the network to have a WTA equilibrium are obtained, and rigorous convergence analysis is carried out. The conditions for the network to have the WTA behavior obtained in this paper provide design guidelines for the network implementation and fabrication. We also demonstrate that whenever the network gets into the WTA region, it will stay in that region and settle down exponentially fast to the WTA point. This provides a speeding procedure for the decision making: as soon as it gets into the region, the winner can be declared. Finally, we show that this WTA neural network has a self-resetting property, and a resetting principle is proposed.

Index Terms—Competition, convergence analysis, lateral inhibition, neural networks, neurodynamics, shunting and additive, very large scale integration (VLSI) neural networks, winner-take-all (WTA).

I. INTRODUCTION

NEURAL networks which pick the maximum from a collection of inputs are known as the winner-take-all (WTA) networks [1]. The operation of this network is a mode of extreme contrast enhancement where only the maximally stimulated neuron responds and all other neurons in the network are inhibited. Such networks have been used extensively in decision making, pattern recognition, and competitive-learning networks (see [2]–[5], [27], and references therein).

The current literature frequently describes WTA networks that are constructed using lateral inhibition among the neurons so that the system is a competitive neural network. Intuitively, if the competitive system is initiated from a fair start with sufficiently strong inhibition, the players in this competition will go to two extremes: win or lose, so that WTA behavior can be expected [6]. However, if the lateral inhibition is weak or lateral excitation is involved in the competition, the dynamics of the system can be very complex as illustrated by

Ermentrout [7] and Lemmon and Kumar [8]. Thus, conditions for WTA behavior are desirable for the practical design of such neural networks. There are many strategies for WTA network design. MAXNET [9] is an architecture of mutual inhibition to select a maximum, though its efficiency and implementation are a problem. Motivated by self-organizing algorithms, some iterative neural networks are designed to pick the largest number in a data set (see [10], [11], and references therein). Although these networks are convenient for computer computation, they may be very hard to implement in analog hardware. Lazzaro *et al.* [12] designed, fabricated, and tested a series of compact complementary metal–oxide–semiconductor (CMOS) integrated circuits that realize the WTA function. Wawryn and Strzeszewski [29] designed a programmable WTA neural network using current mode circuit and demonstrated its effectiveness. However, both of these works only analyzed the circuits at steady state and did not provide the transient dynamics, which is obviously a very important issue for WTA network design. Majani and Abu-Mostafa [13] generalized the WTA analog networks to the K-WTA networks, but only gave conditions for local stability of the network. Calvert and Marinov [30] studied an analog Hopfield-type neural network [20], which can identify the K largest components of a list of N numbers with rigorous stability analysis, which follows our prior work [28].

All the WTA neural networks except [12] discussed above assumed that the external inputs are the initial conditions, i.e., the network in operation has to be initialized by the input vectors. Following Lazzaro *et al.* [12], Kane and Kincaid [14] proposed an analog implementation of WTA neural networks in which the external inputs are kept on during the operation. The WTA neural circuit is implemented with the metal–oxide–semiconductor field effect transistors (MOSFETs), some preliminary steady-state analysis, and experimental results have been presented, where the triode region of the MOSFETs have been used. A four-neuron neural network implemented by MOSFETs is shown in Fig. 1. Each neuron consists of one capacitor with capacitance C, one resistor with conductance G, and three MOSFETs. Let us consider the first neuron. I_1 is the external current, the current flowing into the MOSFET with inhibition from jth neuron $(j = 2, 3, 4)$ is $h(v_1, v_j)$, which is given by what we call the MOSFET function in series with the voltage source V_T [23]

$$h(x,y) = \begin{cases} K\left[2(x+V_T)y - (x+V_T)^2\right], & y \geq 0, 0 \leq x + V_T \leq y \\ Ky^2, & y \geq 0, x + V_T > y \\ 0, & \text{otherwise} \end{cases}$$
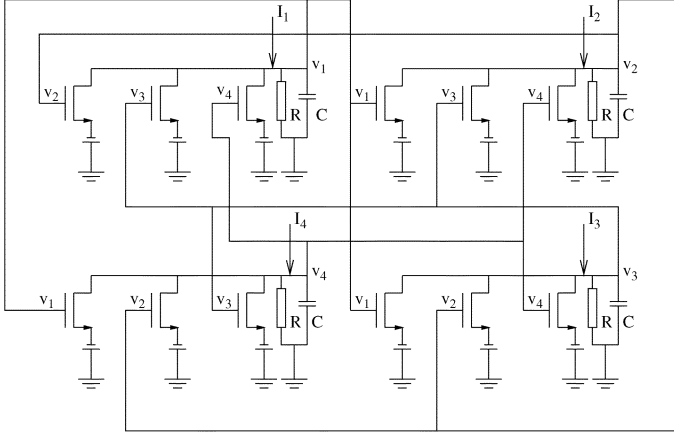
Fig. 1. Four-neuron MOSFET implemented neural network.

where V_T is the voltage threshold and K is the constant determined by the physical characteristics of the MOSFET. From circuit theory, we easily obtain

$$I_1 = C\dot{v}_1 + Gv_1 + h(v_1, v_2) + h(v_1, v_3) + h(v_1, v_4)$$

i.e.,

$$C\dot{v}_1 = -Gv_1 + I_1 - \sum_{j \neq 1} h(v_1, v_j)$$

where the overdot denotes the time derivative. Similarly, for the ith neuron, we have

$$C\dot{v}_i = -Gv_i + I_i - \sum_{j \neq i} h(v_i, v_j), \qquad i = 1, 2, 3, 4.$$

Motivated by this model, but allowing $h(\cdot, \cdot)$ to be a general function, we arrive at the following general class of N-neuron dynamical neural networks:

$$C\dot{v}_i = -Gv_i + I_i - \sum_{j \neq i} h_j(v_i, v_j), \qquad i = 1, 2, \dots, N \quad (1)$$

where v_i is the activation of the ith neuron, $G > 0$ is the decaying factor, I_i is the external input of the ith neuron, and $h_j(v_i, v_j)$ is the inhibition on the ith neuron from the jth neuron, which shows the dependency on the ith neuron.

This class of neural networks turns out to be very general and includes two well-known intensively studied models—the additive and shunting neural networks for the purpose of WTA behavior [13], [16], [17]. If we choose $C = 1$ and $h_i(x, y) = h(x, y) = Kd(y)$ where $d(y)$ is a nondecreasing nonnegative function, we obtain the additive model [13]

$$\dot{v}_i = -Gv_i + I_i - \sum_{j \neq i} Kd(v_j), \qquad i = 1, 2, \dots, N. \quad (2)$$

If we choose $C = 1$ and $h_i(x, y) = K(x + V_T)d(y)$, then we obtain the shunting neural network [17]

$$\dot{v}_i = -Gv_i + I_i - K(v_i + V_T) \sum_{j \neq i} d(y_j), \qquad i = 1, 2, \dots, N \quad (3)$$

where V_T is a threshold constant. When I_i is allowed to be time varying, model (1) also contains the model studied by Tonkin *et al.* [27].

By specifying the function $h_i(x, y)$, we can obtain a specific neural network. This is very important for practical design from a very rich class of models which share similar properties, yet possess different implementations. Moreover, this model has two advantages. The first is the choice of circuit elements. Additive models and shunting models may be hard to implement, while our general models can be easily implemented via the well-known MOSFET circuits. The second is the generality of our models, which offers us more choices to select the appropriate models for better implementation.

This paper is devoted to the study of the WTA behavior of neural networks which belong to the general type (1). It is organized as follows. In Section II, we present notation and definitions. A general set of sufficient conditions for the WTA point to exist for the neural network (1) is given in Section III. In Section IV, we show that neural networks with a few specified functions $h_i(x, y)$ will settle down to the WTA point and the winner can be declared by observing the outputs of the network. We show that if the neural network gets into the WTA region, it will stay there, never get out of that region, and it will settle down at the WTA point. This suggests a quick decision making procedure. We do not need to wait for the system to settle down; as long as it gets into the WTA region, a decision can be made immediately. We also present a resetting procedure for new inputs. If we do not need the system, we simply switch off the inputs, then the neural network will return to the fair starting point, and wait for new inputs to be tested. In consecutive input testing, we should reset the system to the fair starting point. This can be easily done by switching off the inputs for a while between these tests. For a few special cases, we also give the conditions under which the resetting of the neural networks to the fair starting conditions is not necessary. Some illustrative numerical results are presented in Section V and conclusions are drawn in Section VI.

II. NOTATION, DEFINITIONS, AND PRELIMINARIES

Before we study the properties of the neural network (1), we first introduce some notation and definitions. Let $I_1, I_2, \dots, I_N$ denote the external inputs, and $I_{\max}$ and I_{submax} be the largest and the second largest external inputs. Without loss of generality for analysis, we assume throughout this paper that $I_i > 0$ and $I_i \neq I_j$ ($i \neq j, i, j = 1, 2, \dots, N$). (If some values I_i are equal, we can randomly add small numbers to the input, with high probability to make all inputs distinct, i.e., a randomization technique.) We use $\dot{x}$ to denote the time derivative of x. A point $v = (v_1, v_2, \dots, v_N)^T$ (where the superscript T denotes the transpose of a vector or a matrix) is called a WTA point if it is an equilibrium point of (1) and only the component corresponding to the largest input is positive while other components are nonpositive. Let $v_{\max}$ denote the state variable corresponding to the largest input $I_{\max}$. Then, v is the WTA point

if it is such an equilibrium point of (1) satisfying $v_{\max} > 0$ and $v_j \leq 0$ $(j \neq \max)$. This definition is really motivated by our intuition: when the external inputs stimulate the neural network, at the very end, the strongest will win the competition while the weakest will lose everything. For convenience, let the set $C^+ = \{v = (v_1, \ldots, v_N)^T | v_{\max} > 0, v_j \leq 0, j \neq \max\}$ be called the WTA region. Let $x^+ = \max\{0, x\}$.

A function $h(x, y)$ is called a Lipschitz function (or satisfies the Lipschitz condition) in the domain of interest if there exists a constant $L > 0$ such that $|h(x_1, y) - h(x_2, y)| \leq L|x_1 - x_2|$ and $|h(x, y_1) - h(x, y_2)| \leq L|y_1 - y_2|$ in the domain of interest. Define the function class $\mathcal{F}$ to be the following:

$$\mathcal{F} = \{h(x, y) \geq 0 | h(x, y) \text{ is continuous, nondecreasing}$$
$$\text{in } x \text{ and } y, h(x, y) = 0 \text{ for } y \leq 0\}.$$

The function in this class has the following interpretation: when the activation in one neuron is negative, i.e., inactive, then it does not inhibit other neurons. However, when it is active, it will inhibit other neurons. The stronger the activation, the stronger the inhibition. This can be easily observed from (1). Therefore, this function class is a reasonable characterization for the inhibition process in the WTA networks.

Definition: We say that the neural network (1) is *convergent* if the trajectory of (1) converges to an equilibrium point as time goes to infinity. We say (1) is exponentially convergent to x_0 if there exist an $M > 0$ and a $\sigma > 0$ such that $\|x(t) - x_0\| \leq Me^{-\sigma t}$ for any $t \geq 0$ where $\| \cdot \|$ is the Euclidean norm [19]; we call σ the *convergence rate*. We say that neural network (1) is *(exponentially, asymptotically) stable* if system (1) is (exponentially, asymptotically) stable in the sense of Lyapunov [22].

It is known [22] that asymptotic (exponential) stability is equivalent to stability in the Lyapunov sense and convergence (exponential convergence) as time goes to infinity. For most cases we consider, stability is easy to verify, and we will only concentrate on the convergence property in this paper.

The following lemma will be needed in the subsequent development.

Lemma 2.1:
 a) (Comparison Principle) Let $g(t, x)$ be a scalar continuous function, let $m(t)$ be the solution of the scalar differential equation: $\dot{m}(t) = g(t, m(t))$ with $m(t_0) = m_0$. Then, for any continuous function $x(t)$ satisfying the differential inequality: $\dot{x}(t) \leq g(t, x(t))$ with $x(t_0) = m_0$, we have $x(t) \leq m(t)$ for any $t \geq t_0$. Similarly, for any continuous function $x(t)$ satisfying the differential inequality: $\dot{x}(t) \geq g(t, x(t))$ with $x(t_0) = m_0$, we also have $x(t) \geq m(t)$ for any $t \geq t_0$.
 b) If the continuous function $y(t) \geq 0$ satisfying $\dot{y}(t) \leq -\alpha y(t) + g(t)$ where $\alpha > 0$ and $g(t)$ exponentially converges to zero, i.e., there exist positive numbers $M > 0$ and $\beta > 0$ such that $|g(t)| \leq Me^{-\beta t}$, then $y(t)$ also exponentially converges to zero. In fact, there exists positive number M_1 such that $|y(t)| \leq M_1 e^{-\gamma t}$, where $\gamma < \min\{\alpha, \beta\}$.

Proof: The rigorous version of a) and its proof can be found in [21]. We only need to prove b). From $\gamma < \min\{\alpha, \beta\}$, there exists a small positive number ϵ such that $\gamma + \epsilon < \min\{\alpha, \beta\}$. Then, from the comparison principle, we have

$$0 \leq y(t) \leq e^{-\alpha t} y(0) + M \int_0^t e^{-\alpha(t-\tau)} g(\tau) d\tau$$

$$\leq e^{-\alpha t} y(0) + M \int_0^t e^{-\alpha(t-\tau)} e^{-\beta \tau} d\tau$$

$$\leq e^{-\gamma t} y(0) + \int_0^t e^{-(\gamma+\epsilon)(t-\tau)} e^{-(\gamma+\epsilon)\tau} d\tau$$

$$= e^{-\gamma t} \left[y(0) + Mt e^{-\epsilon t} \right]. \tag{4}$$

However, $y(0) + Mt e^{-\epsilon t}$ is bounded. Therefore, there exists a positive number $M_1 > 0$ such that $y(0) + Mt e^{-\epsilon t} \leq M_1$. Taking this into the above inequality, we can complete the proof of b). □

III. SUFFICIENT CONDITIONS FOR WTA POINT TO EXIST

In order to make the system (1) be a WTA network, we must guarantee that the WTA point exists for system (1) so that a decision can be made when the system settles down. In this section, we present some sufficient conditions for the existence of the WTA point for system (1).

Theorem 3.1: Suppose that the function $h_i(x, y)$ belongs to the class $\mathcal{F}$. Then, system (1) has a WTA point if and only if

$$I_j \leq h_{\max}(0, I_{\max}/G), \qquad j \neq \max, \qquad j = 1, 2, \ldots, N \tag{5}$$

where the subscript max denotes the index such that $I_{\max}$ is the largest input.

Proof: Without loss of generality, we assume that $I_1 > I_2 > \cdots > I_N > 0$. In this case, $I_{\max} = I_1$ and $h_{\max}(x, y) = h_1(x, y)$. *Sufficiency*: Suppose that (5) is true, we want to show that system (1) has an equilibrium point v such that $v_1 > 0$ and $v_j \leq 0$ $(j \neq 1)$. We only need to show that there is a WTA point in the WTA region $C^+ = \{v | v_1 > 0, v_j \leq 0, j \neq 1\}$. In fact, in this region, because $h_i(x, y)$ belongs to the class $\mathcal{F}$, $h_i(x, y) = 0$ for $y \leq 0$, we have $-Gv_1 + I_1 = 0$, i.e., $v_1 = I_1/G > 0$ and $-Gv_j + I_j - h_1(v_j, v_1) = 0$; we only need to show that for any $j \neq 1$, the equation $-Gv_j + I_j - h_1(v_j, I_1/G)$ has nonpositive solution. Let $F(x) = -Gx + I_j - h_1(x, I_1/G)$. Then $F(x)$ is a continuous function, $F(0) = I_j - h_1(0, I_1/G) \leq 0$, and for $T > 0$, we have $F(-T) = GT + I_j - h_1(-T, I_1/G) \geq GT - h_1(0, I_1/G) > 0$ for sufficiently large T. From the intermediate theorem for continuous functions, we conclude that there exists a point $v_j \in [-T, 0]$ such that $F(v_j) = 0$. This proves the existence of a WTA point.

Necessity: We need to show that if the system (1) has a WTA point, then (5) must be true. In fact, suppose that (5) is not true, then there is $i \neq 1$ such that $I_i > h_1(0, I_1/G)$. Let v be the WTA point, i.e., $v_1 > 0$ and $v_j \leq 0$ $(j \neq 1)$. We know that $v_1 = I_1/G > 0$ and $-Gv_j + I_j - h_1(v_j, I_1/G) = 0$. In particular, $-Gv_i + I_i - h_1(v_i, I_1/G) = 0$. From this and the monotonic property for $h_1(x, y)$ in x, we have $Gv_i = I_i - h_1(v_i, I_1/G) \geq I_i - h_1(0, I_1/G) > 0$, which contradicts the fact that $v_i \leq 0$. Therefore, (5) must be true. □
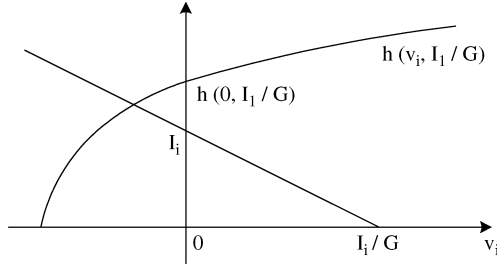
Fig. 2. Graphic proof of Corollary 3.2.

Theorem 3.1 is for the fixed set of inputs. If the maximal input can be actuated at any neuron (this is usually true in practice because we do not know which input will be the largest), the following result should be used for fabrication.

Corollary 3.1: Suppose that the function $h_i(x,y)$ belongs to the class $\mathcal{F}$. Then, system (1) has a WTA point if

$$I_j \leq \min_{1 \leq i \leq N} h_i(0, I_{\max}/G), \qquad j \neq \max, \qquad j = 1,2,\ldots,N.$$

Proof: This is straightforward from Theorem 3.1.

If all neurons are identical, then the inhibition function $h_i(x,y)$ will be the same for any $i \in \{1,2,\ldots,N\}$. Let $h_i(x,y) = h(x,y)$, then we have the following.

Corollary 3.2: If the function $h(x,y)$ belongs to class $\mathcal{F}$, then the neural network

$$C\dot{v}_i = -Gv_i + I_i - \sum_{j \neq i} h(v_i, v_j), \qquad i = 1,2,\ldots,N \quad (6)$$

has a WTA point if and only if

$$I_j \leq h(0, I_{\max}/G), \qquad j \neq \max, \qquad j = 1,2,\ldots,N.$$

Proof: The proof in Theorem 3.1 for this case is best illustrated graphically in Fig. 2.

For the additive neural networks, we have the following.

Corollary 3.3: Suppose that $d(y)$ is a nondecreasing function with $d(y) = 0$ for $y \leq 0$, then the system (2) has a WTA point if and only if $I_j \leq K d(I_{\max}/G)$ for $j \neq \max$. In particular, if $d(y) = y^+$, and $I_j \leq (K/G)I_1$ or $K \geq G$, system (2) has a WTA point.

For the shunting neural networks, if we use function $h_i(x,y) = K(x + V_T)d(y)$, we have the following.

Corollary 3.4: Suppose that $d(y)$ is a nondecreasing function with $d(y) = 0$ for $y \leq 0$, then system (3) has a WTA point if and only if $I_j \leq K V_T d(I_{\max}/G)$ for $j \neq \max$. In particular, let $d(y) = y^+$, if $I_j \leq (K V_T/G)I_{\max}$ ($j \neq \max$), or $K V_T \geq G$, then system (3) has a WTA point.

For the MOSFET implemented neural networks, we have the following.

Corollary 3.5: The MOSFET implemented neural network (1) with MOSFET function $h_i(x,y)$ has a WTA point if and only if either

$$\frac{1}{G} \geq \max\left\{\frac{V_T}{I_{\max}}, \frac{1}{2KV_T}\frac{I_j}{I_{\max}} + \frac{V_T}{2I_{\max}}\right\}$$

or

$$\frac{1}{\sqrt{K}}\frac{\sqrt{I_j}}{I_{\max}} \leq \frac{1}{G} < \frac{V_T}{I_{\max}}$$

for all $j \neq \max$.

IV. WTA Behavior and Convergence Analysis

Our purpose in this paper is to study the WTA behavior of the neural network (1). In the last section, we provided a set of conditions for system (1) to have a WTA equilibrium point. However, the existence of a WTA point alone is not enough to say that network (1) is a WTA network. We still need to show that network (1) will eventually settle down to the WTA point so that the largest input can be picked. For this reason, we have to perform the convergence analysis for network (1). It seems impossible to design a network (1) that will always converge to the WTA point for any external inputs no matter where it starts because of the nonlinearity in the function $h_i(x,y)$. In order to get the natural winner from a closed competitive system, a fair start should be expected. A natural fair starting point is the origin, i.e., each competitor has no activity at the very beginning. In this paper, we will always start system (1) from the origin, and a resetting procedure will be needed whenever necessary. For a limited input space, global convergence is possible without any resetting. This issue will be discussed at the end of this section.

There has been intensive research on stability (including convergence) of dynamical neural networks (see [15], [16], [18]–[20], [24]–[28], and references therein). The Cohen–Grossberg theorem [15] provided the most general convergence theorem in neural networks, and a general Lyapunov function construction guideline was proposed. It has been shown that the additive and shunting neural networks with symmetric interconnections between neurons own global Lyapunov functions, so the convergence of the networks can be concluded. Hirsch [18] has written a series of papers on competitive or cooperative systems, and has established a set of remarkable results on the characterization of the limit set structures for such systems. Recently, Michel *et al.* [24]–[26], Tonkin *et al.* [27], and Fang and Kincaid [28] have used various techniques to study the stability and convergence property of Hopfield-type dynamical neural networks.

System (1) may not be converted into the Cohen–Grossberg model. Furthermore, although system (1) is competitive, the limiting theorems by Hirsch [18] cannot be applied for the WTA purpose. It seems that there is no direct available technique to study the convergence or stability for system (1) in its general form. In this section, we present some convergence results for system (1) for a few special cases.

We first want to show the boundedness of the trajectory of neural network (1).

Theorem 4.1: If the functions $h_i(x,y)$ ($i = 1,2,\ldots,N$) belong to the class $\mathcal{F}$, the trajectory of a neural network (1) is bounded.

Proof: We only need to show that the trajectory of system (1) will eventually stay in a bounded set. If $v_i > I_i/G$, then, from the nonnegativity of $h_i(x,y)$, we have

$$C\dot{v}_i = -Gv_i + I_i - \sum_{j \neq i} h_j(v_i, v_j) \leq -Gv_i + I_i < 0$$

hence $v_i(t)$ will decrease and finally $v_i(t) \leq I_i/G$, so there is a positive constant $U > 0$ such that $v_i(t) \leq U$ ($i = 1,2,\ldots,N$). If $v_i < 0$, then we have

$$C\dot{v}_i \geq -Gv_i + I_i - \sum_{j \neq i} h_j(0, U)$$

where we use the fact that $h_i(x, y)$ is nondecreasing in both x and y; therefore, if $v_i(t) < (I_i - \sum_{j \neq i} h_j(0, U))/G$, then $\dot{v}_i(t) > 0$, and $v_i(t)$ will increase, so there exists a constant B such that $v_i(t) \geq B$. This completes the proof. $\square$

The trajectory of (1) has the so-called order-preserving property which we formalize as follows.

Theorem 4.2: Suppose that functions $h_i(x, y)$ $(i = 1, 2, \ldots, N)$ are continuous functions satisfying $h_1(x, x) = h_2(x, x) = \cdots = h_N(x, x)$ for any x, then neural network (1) is order preserving: if $I_i > I_j$ and $v_i(t_0) > v_j(t_0)$, then for any $t \geq t_0$, $v_i(t) > v_j(t)$.

Proof: Let $\Delta(t) = v_i(t) - v_j(t)$, then $\Delta(t)$ is a continuous function and $\Delta(t_0) > 0$. Suppose that the claim in the theorem is not true, then there exists a $t^* > t_0$ such that $\Delta(t^*) = 0$ and $\Delta(t) > 0$ for $t_0 < t < t^*$. At t^*, we have $v_i(t^*) = v_j(t^*)$. Thus, subtracting the jth equation from the ith equation in (1), we obtain

$$
\begin{aligned}
C\dot{\Delta}(t^*) =& -G\Delta(t^*) + (I_i - I_j) \\
& - [h_j(v_i(t^*), v_j(t^*)) - h_i(v_j(t^*), v_i(t^*))] \\
& - \sum_{k \neq i, j} [h_k(v_i(t^*), v_k(t^*)) - h_k(v_j(t^*), v_k(t^*))] \\
=& I_i - I_j > 0.
\end{aligned}
$$

Since $\dot{\Delta}(t)$ is also a continuous function, from the above inequality, there exists a small $\delta > 0$ $(\delta < t^* - t_0)$ such that $\dot{\Delta}(t) > 0$ in the interval $[t^* - \delta, t^* + \delta]$, i.e., $\Delta(t)$ is strictly increasing in this interval, hence $\Delta(t^*) > \Delta(t^* - \delta) > 0$, this contradicts the choice of t^*. $\square$

Remark: Condition $h_1(x, x) = h_2(x, x) = \cdots = h_N(x, x)$ is reasonable because $h_i(x, y)$ is the inhibition between neurons. If both neurons have the same activity, they should have the same inhibition on each other. We will call this the equal mutual inhibition condition. It is easy to verify that this condition holds for the additive neural network (2), the shunting neural network (3), and the MOSFET implemented neural network. We also notice from the proof that when the above equalities only approximately hold, Theorem 4.2 is still valid.

From system (1), it is easy to see that when the system starts from the resting state, i.e., the fair starting condition, all state variables will increase and get into the positive orthants. As we expect, the state variable corresponding to the largest input will stay positive while the other state variables will gradually decrease. From the order-preserving property and the system equations, at least one of the state variables will always be positive. If network (1) has a fair start, then it is the winning variable that will always be positive. Therefore, it suffices to study the conditions under which the system will enter the WTA region. The following result guarantees that whenever the network enters the WTA region, the network will stay in that region and converge to the WTA point. Hence, the system at least locally has WTA behavior. In fact, if the inhibition functions $h_i(x, y)$ $(i = 1, 2, \ldots, N)$ are smooth, we can show that the WTA point is (locally) exponentially stable (i.e., asymptotically stable in Lyapunov sense with convergence exponentially fast).

Theorem 4.3: Let C^+ denote the WTA region. Suppose that functions $h_i(x, y)$ $(i = 1, 2, \ldots, N)$ belong to the class $\mathcal{F}$ and satisfy the equal mutual inhibition (EMI) condition: $h_1(x, x) = h_2(x, x) = \cdots = h_N(x, x)$ for any x. Then, whenever the trajectory of (1) starting from the origin enters

the WTA region C^+, it will stay there forever. Moreover, if functions $h_i(x, y)$ $(i = 1, 2, \ldots, N)$ satisfy the Lipschitz condition, it will converge exponentially to the WTA point with convergence rate at least γ for any $\gamma < G$. Furthermore, if $h_i(x, y)$ $(i = 1, 2, \ldots, N)$ are differentiable, then the WTA point is (locally) exponentially stable.

Proof: Proof is given in the Appendix. $\square$

This theorem also suggests a speeding procedure for decision making. If network (1) has a fair start, i.e., initiates near the origin, then whenever the network enters the WTA region, we can stop. The winner can be declared because the corresponding WTA region has the same property as the WTA point for the purpose of maximum selection. This will reduce the processing time significantly.

Remark: As we observe from the definition of the exponential convergence, the time taken for network (1) to converge to the neighborhood of the WTA point depends not only on the convergence rate, but also on the constant M. Exponential convergence does not mean that the transient time is short. As we can observe from the proof, the exponential convergence takes significant effect only when the network gets into the WTA region. Hence, we have to design the neural network in such a way that the transient time is as short as possible. It will be very difficult to characterize the transient response as we observe in control literature.

Now, in order to prove the global convergence, we only need to show that the trajectory of (1) starting from the origin cannot stay in the region where at least two components are positive. However, for the general functions $h_i(x, y)$ $(i = 1, 2, \ldots, N)$, the convergence of network (1) is unknown. The Cohen–Grossberg theorem is difficult to apply to general model (1) because of the inseparability of the function $h_i(x, y)$. It is obvious that system (1) is a competitive system. However, Hirsch's results [18] on the characterization of limit sets cannot be used to prove the WTA convergence. Though the general convergence analysis is still under investigation, we will in this paper show that for a few subclasses of system (1) the trajectory starting from the origin will not forever stay in the region where at least two components are positive. It will enter the WTA region and converges to the WTA point. In this way, we have completed the convergence analysis for these few subclasses of WTA neural networks. More importantly, the convergence analysis presented may provide a possible approach for the analysis of general neural system (1).

Consider the following three classes of neural networks:

$$
\dot{v}_i = -Gv_i + I_i - K(D + v_i)\sum_{j \neq i} v_j^+, \qquad i = 1, 2, \ldots, n \quad (7)
$$

$$
\dot{v}_i = -Gx_i + I_i - K\sum_{j \neq i} \left[v_j^+\right]^2, \qquad i = 1, 2, \ldots, n \quad (8)
$$

$$
\dot{v}_i = -Gv_i + I_i - K\sum_{j \neq i} v_j^+, \qquad i = 1, 2, \ldots, n. \quad (9)
$$

Obviously, the last two neural networks belong to the additive models and the first neural network belongs to the shunting models, which are the special cases of neural network (1). The interconnections among neurons are the same, i.e., the synaptic connections have the same strength and the decay factors are the same. Therefore, these three classes of neural networks are very easy to implement. We will show that these neural networks,

though simple, possess WTA behavior and can be used as a maximum decoder.

Assume that G, K, and D are all positive numbers.

Theorem 4.4: Under the following conditions, respectively,

1) $I_{\text{submax}} < D(KD - G)$ for system (7);
2) $I_{\text{submax}} \leq K I_{\max}^2/G^2$ and $I_{\max} - I_{\text{submax}} > G^2/(4K)$ for system (8);
3) $K/G > 1$ for system (9);

the trajectories of neural networks (7)–(9) starting from the fair condition, i.e., the origin, will always converge to the WTA point, and hence they are the WTA networks.

Proof: The proof is given in the Appendix.   □

Remark: The condition for boundedness of the inputs will be called the boundedness condition while the lower boundedness of the difference between the largest input and second largest input will be called the resolution condition.

From the proof, we observe that in order to complete the convergence analysis for general system (1), it suffices to show either there are no limiting equilibrium points in the region where at least two components are positive or system (1) does not stay forever in the region. The readers may want to convince themselves by studying the simple 2-D additive neural network (9). In this case, although there are other stable equilibrium points in non-WTA regions, under certain conditions, the fair starting point (or its small neighborhood) is not in their domains of attraction, and if the network starts from the domain of attraction of a non-WTA equilibrium, the network will converge to the non-WTA equilibrium point, which leads to a wrong decision.

As we see from the proof of 2) and 3) above, it is enough to show that a certain error function is unbounded. In the proof of the next theorem, we provide another approach for convergence analysis of the neural network of type (1).

Theorem 4.5: Consider the neural network

$$C\dot{v}_i(t) = -Gv_i(t) + I_i - \sum_{j \neq i} h(v_i(t), v_j(t)), \qquad i = 1, 2, \ldots, N$$

(10)

where function $h(x, y)$ belongs to the class $\mathcal{F}$ and satisfies the following conditions: there exist positive numbers V_T, Γ_1, Γ_2, and $\gamma > 1$ such that $h(x, y) - h(x, x) \geq \Gamma_1(y - x)^\gamma$ if $0 < x \leq y \leq x + V_T$ and $h(x, y) - h(x, x) \geq \Gamma_2$ if $x > 0$ and $y > x + V_T$, and there exists a continuous function $g(y)$ such that $h(x, y) = g(y)$ for $x \geq y > 0$, i.e., the inhibition function $h(x, y)$ has a saturation property. If the following resolution condition and the boundedness condition hold:

$$I_{\max} - I_{\text{submax}} > \Gamma_1(\gamma - 1) \left[\frac{G}{\Gamma_1 \gamma}\right]^{\gamma/(\gamma-1)}$$

$$I_{\text{submax}} < \min\{h(0, I_{\max}/G), h(0, a)\}$$

where

$$a = \max\{V_T, (I_{\max} - I_{\text{submax}} + \Gamma_2)/G\}$$

then neural network (10) starting from the origin will always converge to the WTA equilibrium point, so system (10) is a WTA network.

Proof: Without loss of generality, we assume that $I_1 > I_2 > \cdots > I_N > 0$. If system (10) starts from the origin, then from the order-preserving property, we must have $v_1 > v_2 > \cdots > v_N$. From the assumption, we have $I_2 < h(0, I_1/G)$.

From Corollary 3.2, a WTA point of (10) exists. From Theorem 4.3, it suffices to show that the trajectory of (10) starting from the origin does not stay forever in the region where at least two components are positive. Suppose that this is not true, then we must have $v_1(t) > v_2(t) > 0$ for all $t > 0$. The first two equations of (10) become

$$C\dot{v}_1 = -Gv_1 + I_1 - h(v_1, v_2) - \sum_{j \neq 1, 2} h(v_1, v_j) \quad (11)$$

$$C\dot{v}_2 = -Gv_2 + I_1 - h(v_2, v_1) - \sum_{j \neq 1, 2} h(v_2, v_j). \quad (12)$$

If $v_j \leq 0$, then $h(v_1, v_j) = h(v_2, v_j) = 0$. If $v_j > 0$, then from $v_1 > v_2 > v_j$, we have $h(v_1, v_j) = h(v_2, v_j) = g(v_j)$. So for $j \neq 1, 2$, we always have $h(v_1, v_j) - h(v_2, v_j) = 0$. Letting $e(t) = v_1(t) - v_2(t)$, subtracting (12) from (11) and noticing that $h(v_1, v_2) = g(v_2) = h(v_2, v_2)$, we obtain

$$C\dot{e} = -Ge + (I_1 - I_2) + h(v_2, v_1) - h(v_1, v_2)$$
$$= -Ge + (I_1 - I_2) + h(v_2, v_1) - h(v_2, v_2). \quad (13)$$

System (10) starts from the origin, and hence initially $e(t) < V_T$. From the property of the function $h(x, y)$, we have for $e(t) \leq V_T$, we have $h(v_2(t), v_1(t)) - h(v_2(t), v_2(t)) \geq \Gamma_1 e(t)^\gamma$. From (13), we obtain

$$C\dot{e}(t) \geq (I_1 - I_2) - Ge(t) + \Gamma_1 e(t)^\gamma$$
$$\geq (I_1 - I_2) - G\left(\frac{G}{\Gamma_1 \gamma}\right)^{1/(\gamma-1)} + \Gamma_1\left(\frac{G}{\Gamma_1 \gamma}\right)^{\gamma/(\gamma-1)}$$
$$= (I_1 - I_2) - \Gamma_1(\gamma - 1)\left(\frac{G}{\Gamma_1 \gamma}\right)^{\gamma/(\gamma-1)}$$
$$> 0$$

where we have used the resolution condition and the fact that the function $ax^\gamma - bx + c$ is minimized at $[b/(2a)]^{1/(\gamma-1)}$ on $[0, +\infty)$ for $a > 0$, $b > 0$, $c > 0$, and $\gamma > 1$. Thus, $e(t)$ will eventually exceed V_T and never go below V_T. Let T_0 denote the instant such that $e(t) \geq V_T$ for $t \geq T_0$. Then, from (13), we have

$$C\dot{e} \geq -Ge + (I_1 - I_2) + \Gamma_2, \text{ or } \dot{e} \geq -\frac{G}{C}e + \frac{(I_1 - I_2) + \Gamma_2}{C}.$$

From the comparison principle, we obtain for $t \geq T_0$

$$e(t) \geq e^{-(G/C)(t-T_0)}e(T_0) + \frac{(I_1 - I_2) + \Gamma_2}{G}\left(1 - e^{-(G/C)(t-T_0)}\right)$$

i.e.,

$$v_1(t) \geq v_2(t) + e^{-(G/C)(t-T_0)}e(T_0)$$
$$+ \frac{(I_1 - I_2) + \Gamma_2}{G}\left(1 - e^{-(G/C)(t-T_0)}\right). \quad (14)$$

Let $b(\delta) = \max\{V_T, [(I_1 - I_2) + \Gamma_2](1 - \delta)/G\}$. From the assumption

$$I_2 < \min\{h(0, I_1/G), h(0, a)\}$$

there exists a small positive number $\delta > 0$ such that $I_2 < h(0, b(\delta))$. Also, there exists a $T > T_0 > 0$ such that for $t > T$, from (14) and $e(T_0) \geq V_T$, we have $v_1(t) \geq v_2(t) + b(\delta)$. From (12), we obtain for $t \geq T_0$

$$C\dot{v}_2(t) \leq -Gv_2(t) + I_2 - h(v_2(t), v_1(t))$$
$$\leq -Gv_2(t) + I_2 - h(v_2(t), v_2(t) + b(\delta)).$$

Let $m(t)$ be the solution of

$$C\dot{m}(t) = -Gm(t) + I_2 - h\left(m(t), m(t) + b(\delta)\right)$$
$$m(T) = v_2(T).$$

Because $I_2 < h(0, b(\delta))$, this system has a negative stable equilibrium point and for sufficiently large t, we will have $m(t) < 0$. However, from the comparison principle, we have $v_2(t) \leq m(t)$ for $t \geq T$, and so for sufficiently large t, we have $v_2(t) < 0$. This contradicts the assumption that $v_2(t) > 0$. This shows that the trajectory must get out of that region and enter the WTA region in the future time. From Theorem 4.3, the trajectory will exponentially converge to the WTA point. This completes the proof. $\square$

It can be easily verified that the MOSFET function satisfies all the conditions of Theorem 4.5 with $\gamma = 2$, $\Gamma_1 = K$, and $\Gamma_2 = KV_T^2$, and so for the MOSFET implemented network, we have the following.

Corollary 4.6: Consider neural network (1) with $h_i(x, y)$ defined as the MOSFET function. Under the resolution condition $I_{\max} - I_{\text{submax}} \geq G^2/4K$ and the upper boundedness condition

$$I_{\text{submax}} < KV_T \left[\max\left\{V_T, 2(I_{\max} - I_{\text{submax}})/G \right.\right.$$
$$\left.\left. + 2KV_T^2/G - V_T\right\}\right]$$

the MOSFET implemented network starting from the origin will exponentially converge to the WTA point. $\square$

The proof may provide an approach to the convergence analysis of (1). It may be possible to obtain a good lower bound for the trajectory of the winner and get upper bounds for the losers so that imposed conditions on the inputs may lead to a contradiction. This issue will be studied in the future.

As we have mentioned, in order to effectively use this kind of WTA network (1), we have to start from the origin, i.e., the fair starting point. One way is to use switching circuits to reset the network to the origin whenever a new input vector is about to be tested. However, this needs some additional circuits to implement. Fortunately, network (1) has an intrinsic resetting procedure, which will be discussed next. The result shows that we only need to switch off the inputs for some time and network (1) will accomplish the resetting task.

Theorem 4.7 (Self-Resetting Theorem): If functions $h_i(x, y)$ $(i = 1, 2, \ldots, N)$ belong to class $\mathcal{F}$ and satisfy Lipschitz condition, then when the external inputs are switched off, i.e., $I_i = 0$ $(i = 1, 2, \ldots, N)$, neural network (1) will globally exponentially converge to the origin, the fair starting point, with the convergence rate γ where $\gamma < G/C$, hence the network has an intrinsic resetting property.

Proof: When the external inputs are switched off, the network reduces to the system

$$C\dot{v}_i = -Gv_i - \sum_{j \neq i} h_j(v_i, v_j), \qquad i = 1, 2, \ldots, N. \quad (15)$$

First, we want to show that (15) has a unique equilibrium point. Obviously, the origin is one equilibrium point of (15). Let $v = (v_1, \ldots, v_N)^T$ be any equilibrium point. Then, $-Gv_i - \sum_{j \neq i} h_j(v_i, v_j) = 0$, and so $v_i \leq 0$ for any $i = 1, 2, \ldots, N$. Since $h_i(x, y)$ belongs to class $\mathcal{F}$, we have $h(x, y) = 0$ for $y \leq 0$. Therefore, we have

$0 = -Gv_i - \sum_{j \neq i} h(v_i, v_j) = -Gv_i$, so $v_i = 0$, i.e., the origin is the unique equilibrium point of (15).

Next, we want to show that if there exists a t_1 such that $v_i(t) \leq 0$, then for all $t \geq t_1$, $v_i(t) \leq 0$. This is almost obvious, because if there exists another t_2 such that $v_2(t_2) \geq 0$, we have

$$C\dot{v}_i(t_2) = -Gv_i(t_2) - \sum_{j \neq i} h_j\left(v_i(t_2), v_j(t_2)\right) \leq -Gv_i(t_2) \leq 0$$

so $v_i(t)$ cannot cross the boundary.

From this argument, we observe that there exists a $T > 0$ and m such that for all $t \geq T$, we have (without loss of generality, we can still use such indexing)

$$v_i(t) \geq 0, \qquad i = 1, 2, \ldots, m$$
$$v_j(t) \leq 0, \qquad j = m+1, \ldots, N.$$

Then, system (15) is reduced to the following system:

$$C\dot{v}_i = -Gv_i - \sum_{k \neq i}^{m} h_k(v_i, v_k), \qquad i = 1, 2, \ldots, m \quad (16)$$

$$C\dot{v}_j = -Gv_j - \sum_{k=1}^{m} h_k(v_j, v_k), \qquad j = m+1, \ldots, N. \quad (17)$$

From (16), we have $C\dot{v}_i \leq Gv_i$, hence we have $0 \leq v_i(t) \leq e^{-(G/C)(t-T)}v_i(T)$. Thus, there exists a $M > 0$ such that $|v_i(t)| \leq Me^{-(G/C)(t-t_0)}$, i.e., (16) globally exponentially converges to its origin.

Because $h_k(x, y)$ satisfies the Lipschitz condition, there exists a constant $L > 0$ such that $h_k(0, y) \leq L|y|$. Also, since $v_j(t) \leq 0$ for $t \geq T$ $(j = m+1, \ldots, N)$, we have $h_k(v_j(t), y) \leq L|y|$. Thus, from (17), for $t \geq T$ and $j \in \{m+1, \ldots, N\}$, we have

$$C\dot{v}_j(t) \geq -Gv_j(t) - \sum_{k=1}^{m} h_k\left(0, v_k(t)\right) \geq -Gv_j(t) - L\sum_{j=1}^{m} v_k(t).$$

Applying the same procedure as in the proof of Lemma 2.1(b) and noticing that $0 \leq v_k(t) \leq Me^{-(G/C)(t-t_0)}$ $(k = 1, 2, \ldots, m)$, we can conclude that $v_j(t)$ will globally exponentially converge to zero with convergence rate $\gamma < G/C$. This completes the proof. $\square$

For a very large scale integration (VLSI) circuit designer interested in the MOSFET implemented neural networks, we summarize our results for this particular case for convenience.

Theorem 4.8: For the MOSFET implemented neural network (1) with MOSFET function $h_i(x, y)$, under

1) the gain condition

$$\frac{KV_T}{G} \geq 1;$$

2) the resolution condition and lower bound condition

$$I_{\max} - I_{\text{submax}} \geq \frac{G^2}{4K}, \qquad I_{\max} > \frac{G^2}{K};$$

the trajectory of the MOSFET implemented neural network starting from the fair condition (i.e., the origin) will exponentially converge to the WTA point, hence the network has WTA behavior. Furthermore, the network has intrinsic self-resetting property: the network will exponentially converge to the fair

condition when the external inputs are switched off. The time constant for the network is approximately equal to the time constant of the RC circuit, i.e., C/G, and hence the convergence rate of the network is approximately equal to G/C.

Proof: The gain condition and the lower bound condition guarantee the existence of a WTA point from Corollary 3.6. We will use the procedure in the proof of Theorem 4.5 to show the convergence of the equilibrium point. It suffices to show that under the gain condition $e(t)$ in (13) diverges to infinity, which leads to the contradiction. Following the similar argument, we obtain that there exists T_0 such that $e(t) \geq V_T > 0$ for all $t \geq T_0$. Taking the MOSFET function $h_i(x, y)$ in (13), we obtain

$$\dot{e}(t) = -Ge + I_1 - I_2 + K\left[2(v_2 + V_T)v_1 - (v_2 + V_T)^2\right] - Kv_2^2$$
$$= (2KV_T - G)e + I_1 - I_2 - KV_T^2 + 2Kv_2e \geq ae + b$$

where $a = 2KV_T - G$ and $b = I_1 - I_2 - KV_T^2$. From the gain condition, we know that $a > 0$. From the comparison principle, we obtain for $t \geq T_0$

$$e(t) \geq e^{a(t-T_0)}e(T_0) + b\int_{T_0}^{t} e^{a(t-\tau)}d\tau$$
$$= -\frac{b}{a} + \left(e(T_0) + \frac{b}{a}\right)e^{a(t-T_0)}.$$

From $e(T_0) \geq V_T$ and the resolution condition, we obtain

$$e(T_0) + \frac{b}{a} \geq V_T + \frac{(I_1 - I_2) - KV_T^2}{2KV_T - G} \geq \frac{(KV_T/G - 1/2)^2 G}{4K(2KV_T/G - 1)} > 0.$$

Hence, $\lim_{t\to\infty} e(t) = +\infty$, and so $v_1(t)$ is unbounded; this contradicts the fact that the trajectory is bounded. Therefore, the network starting from the fair condition will exponentially converge to the WTA point. The resetting property can be obtained from Theorem 4.7 (noticing that we absorb the capacitance C in G). This completes the proof. □

In the above, we assume that the network starts from the origin, the fair starting condition. Intuitively, this is reasonable, because for a fair competition you have to allow each player a fair start. However, in certain cases, if some inputs are strong enough, they can overcome the initial bias and still win the competition. In terms of the dynamical neural networks, it means that for certain input range, network (1) may always converge to the WTA point no matter where it starts. Therefore, the reset is not necessary. This will require that: 1) the network should have only one equilibrium point, i.e., the WTA point; and 2) the network should globally exponentially converge. We only give a result for systems (7) and (9) in this paper. Further research on this issue will be investigated in the future.

Theorem 4.9:
1) Neural network (7) has only one asymptotically stable equilibrium point, which is the WTA equilibrium point if the following three conditions hold: $G < KD$, $I_{\text{submax}} < D(KD - G)$ and the resolution condition $I_{\text{submax}}/I_{\text{max}} \leq G/KD$. Therefore, network (7) will always converge to the WTA point for any inputs satisfying the given conditions no matter where it starts.
2) If $K > G$ and the resolution condition $I_{\text{max}}/I_{\text{submax}} > K/G$, then network (9) has only one asymptotically stable equilibrium point, which is the WTA point. This implies that network (9) will always converge to the WTA point for

any external inputs I_i satisfying the resolution conditions no matter where the network starts.

Proof:
1) The first condition guarantees that the WTA point exists. The second condition implies that there are no equilibrium points having at least two positive components. Without loss of generality, we assume that $I_1 > I_2 > \cdots > I_N > 0$. We want to show that there is no other equilibrium point other than the WTA point. Suppose that this is not true, then there is an equilibrium point v such that $v_i > 0$ $(i \neq 1)$ and $v_j \leq 0$ $(j \neq i)$. So from (7), we have $v_i = I_i/G$ and

$$0 = -Gv_1 + I_1 - K(v_1 + D)v_i$$
$$= -Gv_1 + I_1 - K(v_1 + D)I_i/G.$$

Since $v_1 \leq 0$, we must have $I_1 \geq K(0 + D)I_i/G$, i.e., $I_i/I_1 \leq G/(KD)$. This contradicts the resolution condition. So network (7) has only one equilibrium point. Using the same Lyapunov function and applying Cohen–Grossberg theorem, we know that system (7) will converge to the equilibrium point, therefore, it will converge to the WTA point no matter where it starts. This completes the proof of 1).

2) We still assume that $I_1 > I_2 > \cdots > I_N > 0$. Similar to the proof of Theorem 4.4 (3), condition $K > G$ guarantees that the trajectory of (9) starting from any initial condition cannot stay forever in the region which has at least two components are positive. (Otherwise, $v_i(t) - v_j(t)$ will become unbounded.) Thus, there is no equilibrium point in the region where at least two components are positive. The resolution condition implies that there are no equilibrium points in the other regions except the WTA region, and hence the WTA point is the only equilibrium point. Next, we want to show the global convergence. We know that there exists an instant T_0 such that the trajectory enters the region $C_i = \{v | v_i > 0, v_j \leq 0 (j \neq i)\}$ after T_0. Next, we want to show that the trajectory either cannot get out of this region or gets into the region $C_j = \{v | v_j > 0, v_k \leq 0 (k \neq j)\}$ where $I_j > I_i$. In fact, in order for the trajectory to get out of this region into C_j, $v_i(t)$ must decrease while $v_j(t)$ must increase, so $v_i(t) - v_j(t)$ must decrease. There must be an instant $T_1 > T_0$ such that $v_i(T_1) = v_j(T_1) > 0$. Suppose that $I_i > I_j$, then since

$$\dot{v}_1(t) - \dot{v}_2(t) = (K - G)(v_i(t) - v_j(t)) + (I_i - I_j)$$
$$> 0$$

around T_1, i.e., $v_i(t) - v_j(t)$ is increasing around the instant T_1, which contradicts the fact that $v_i(t) - v_j(t)$ is decreasing. Hence, we must have $I_j > I_i$. Thus, there exists an i such that the trajectory will stay in the region C_i forever. Next, we want to show that $i = 1$. Suppose not, then there exists a $T_2 > T_1$ such that for any $t \geq T_2$, we have $v(t) \in C_i$. In this case, we have $\lim_{t\to\infty} v_i(t) = I_i/G$, and the first equation of (9) becomes for $t \geq T_2$

$$\dot{v}_1(t) = -Gv_1(t) + I_1 - Kv_i(t).$$

Therefore, we have

$$\lim_{t\to\infty} v_1(t) = \frac{1}{G}\left(I_1 - K\frac{I_i}{G}\right) = \frac{1}{G}\left(I_1 - K\frac{I_2}{G}\right) > 0.$$

For sufficiently large t, we have $v_1(t) > 0$, which contradicts the fact that $v \in C_i$. So the trajectory will stay in the region C_1, hence exponentially converge to the WTA point. This completes the proof. $\square$

Remark: Theorem 4.9(2) can also be proved by the Cohen–Grossberg theorem. The proof presented here may provide better understanding of the system behavior.

As a final thought, we remark that in this paper we assume that the resistances and capacitances in all neurons are identical. In practice, inaccuracies for capacitors and resistors are unavoidable. This issue can be addressed by the robust stability analysis: the exponential convergence (stability) has intrinsic robustness property [22]. Fang and Kincaid [28] used the matrix measure to obtain some sufficient conditions for exponential stability. Wang and Michel [24] and Ye *et al.* [26] have studied the robust stability of the Hopfield-type neural networks. Their techniques may be helpful for the robust convergence analysis for network (1). The stability of differential inclusions [25] may also shed light on the robustness issue. The rigorous analysis of the robustness problem for (1) has not been carried out, which forms a very challenging future research direction.

V. ILLUSTRATIVE EXAMPLES

In this section, we present illustrative examples to demonstrate our results obtained in this paper and further confirm the consistency between the theoretical results and the numerical results.

Example 5.1: We first give an example for the MOSFET neural network, which has been presented by Kane and Kincaid [14]. For a three-neuron network, we use capacitors of 100 picofarads, with resistors of 113 kΩ and with MOSFETs with physical parameters $K = 30 \ \mu$A/V^2 and threshold $V_T = 1$ V. The current inputs are $I_1 = 20 \ \mu$A, $I_2 = 17 \ \mu$A, and $I_3 = 3.2 \ \mu$A. Since $I_1 - I_2 = 3 \ \mu$A, and $1/4KR^2 = 0.66 \ \mu$A, hence the resolution condition in Corollary 4.6 is satisfied. Also, $KV_T^2(2KRV_T - 1) = 173.4 \ \mu$A, and the boundedness condition in Corollary 4.6 is also valid. Notice that the first condition in Theorem 3.1 is also easily established. Therefore, the network with these inputs will converge to the WTA point. The first graph in Fig. 3 is the trajectories of the state variables, and it can be easily seen that this simulation is consistent with the experimental results in [14]. The second graph in Fig. 3 shows the WTA behavior with self-resetting. We first test the WTA behavior with above input currents, then switch off the input for some time, then swap the input values between I_1 and I_2. The simulation shows that at beginning, the network picks up the winner I_1. After we switch off the currents, the network settles down automatically to the fair starting condition. When the values of I_1 and I_2 are swapped, the network then picks up the new winner I_2. This simulation confirms our theoretical results for the MOSFET WTA networks.

Example 5.2: We continue to study the system in Example 5.1. However, we increase the number of neurons in the system to demonstrate how input currents affect the transient response of neural network (1). All parameters K, V_T C, and R remain the same as in Example 5.1. From Figs. 4–6, we observe that the difference between the largest input and the second largest input affects the transient response significantly; the smaller the
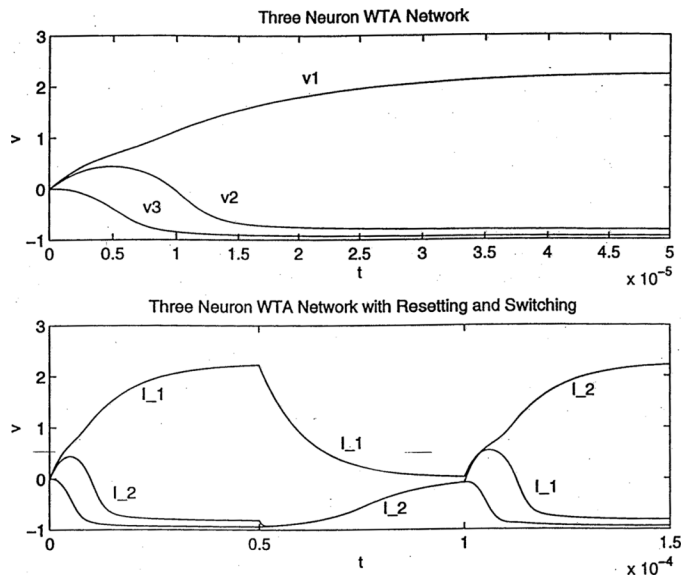

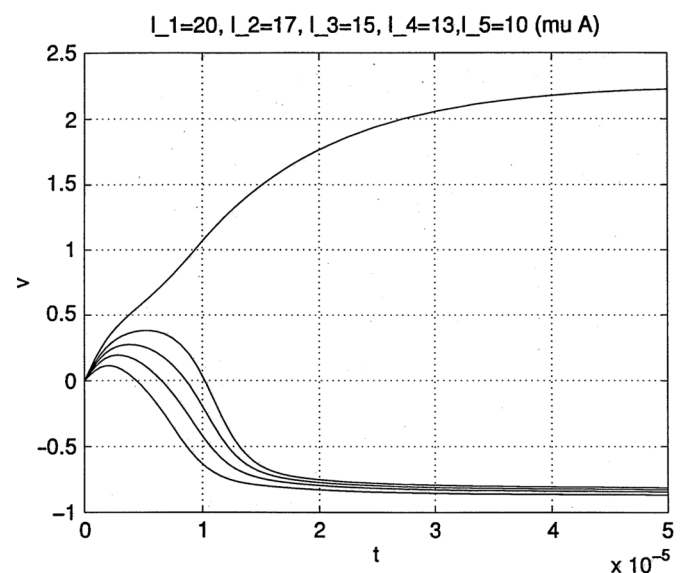Fig. 3. WTA behavior for the MOSFET neural network.

Fig. 4. Five-neuron network with WTA behavior.

difference, the slower the convergence, which is consistent with our intuition.

VI. CONCLUSION

In this paper, we have studied a new class of dynamical WTA neural networks. We have obtained a set of sufficient conditions for the existence of a WTA point, then carried out the convergence analysis of the networks for a few subclasses. Some sufficient conditions are obtained to guarantee that the systems belonging to a few subclasses of type (1) starting from the fair start will converge to WTA points. A self-resetting procedure is also proposed to effectively apply such WTA networks. We also obtain a few conditions for networks to have WTA behavior which do not need resetting. The convergence analysis of *general* system (1) is still under investigation. We conjecture that as long as system (1) has a WTA point, it will converge to the WTA equilibrium point if the system starts from the origin.
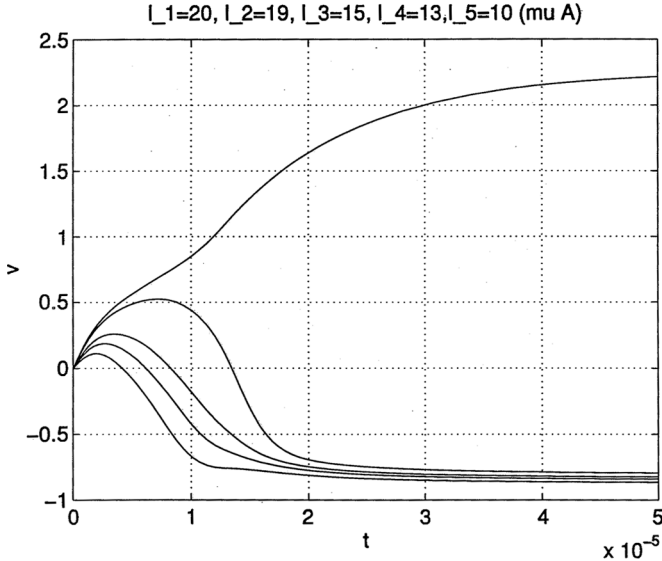
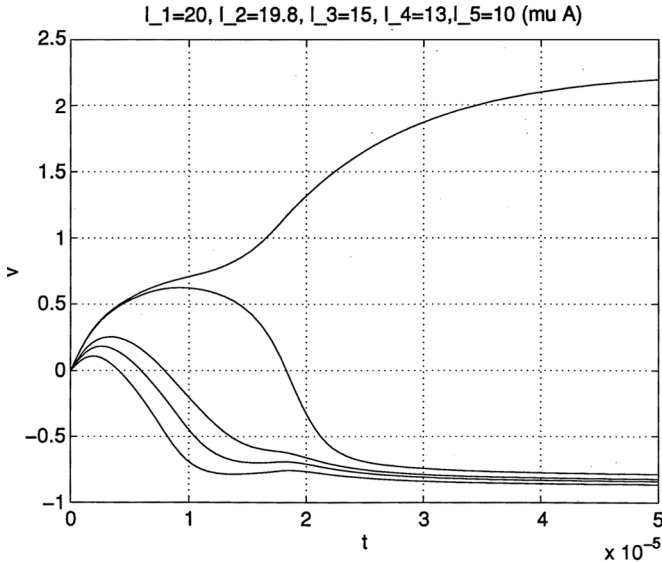Fig. 5.   Five-neuron network with WTA behavior.

Fig. 6.   Five-neuron network with WTA behavior.

APPENDIX

Proof of Theorem 4.3: Since $h_i(0,0) = 0$ and $I_i > 0$, initially, system (1) will get into the positive orthant. Without loss of generality in the proof, we assume that $I_1 > I_2 > \cdots > I_N > 0$. Then, $C^+ = \{v = (v_1, v_2, \ldots, v_N)^T | v_1 > 0 \text{ and } v_j \leq 0, j \neq 1\}$. From (1) and the nonnegativeness of $h_i(x,y)$, we have $C\dot{v}_i(t) \leq -Gv_i(t) + I_i$, so

$$v_i(t) \leq e^{-(G/C)t}v_i(0) + \int_0^t e^{-(G/C)(t-\tau)}(I_i/C)d\tau$$

$$= \frac{I_i}{G}\left(1 - e^{-(G/C)t}\right) \leq I_i/G.$$

Suppose the first claim in Theorem 4.3 is not true, i.e., network (1) enters the WTA region and then gets out of the region later on. Then, there exist $t_1 > 0$ and $t_2 > 0$ such that $v_2(t_1) = v_2(t_2) = 0$ [notice that we have used the order-preserving property of the trajectory of (1)], i.e., $v_2(t)$ gets into

the region C^+ at t_1 and gets out of the region C^+ at t_2. In this case, we must have $\dot{v}_2(t_1) \leq 0$ and $\dot{v}_2(t_2) > 0$ (otherwise, if $\dot{v}_2(t_1) > 0$, then $v_2(t) > v_2(t_1) = 0$ for the time t sufficiently close to t_1 from the right, hence the network does not enter the WTA region at t_1, which contradicts the definition of t_1; a similar argument applies to the other case). On the other hand, in $[t_1, t_2]$, we have $v_i(t) \leq 0$ ($i \neq 1$), therefore, $\dot{v}_1(t) = -Gv_1(t) + I_1 \geq -G(I_1/G) + I_1 = 0$, so we have $v_1(t_1) \leq v_1(t_2)$. Moreover, from (1), we have (using the fact that $h_1(0,y)$ is nondecreasing in y)

$$C\dot{v}_2(t_2) = -Gv_2(t_2) + I_2 - \sum_{j \neq 2} h_j(v_2(t_2), v_j(t_2))$$

$$= I_2 - h_1(0, v_1(t_2))$$

$$\leq I_2 - h(0, v_1(t_1)) = \dot{v}_2(t_1) \leq 0.$$

This contradicts the fact that $\dot{v}_2(t_2) > 0$, and this completes the proof of the first part.

Next, we want to show the exponential convergence. Let t_0 be the instant the trajectory enters the WTA region C^+. Then, from the first part of the theorem, for any $t \geq t_0$, we have

$$C\dot{v}_1(t) = -Gv_1(t) + I_1$$
$$C\dot{v}_j(t) = -Gv_j(t) + I_j - h_1(v_j(t), v_1(t)), \qquad j \neq 1. \quad (18)$$

Solving the first equation, we obtain

$$v_1(t) = I_1/G + Me^{-(G/C)(t-t_0)}, \text{where } M = v_1(t_0) - I_1/G.$$

Since function $h_1(x,y)$ is a Lipschitz function, there is a positive constant $L > 0$ such that $|h_1(x_1,y) - h_1(x_2,y)| \leq L|x_1 - x_2|$ and $|h_1(x,y_1) - h_1(x,y_2)| \leq L|y_1 - y_2|$. For any $j \neq 1$, we have $C\dot{v}_j = -Gv_j + I_j - h_1(v_j, v_1)$. Let v^* be the WTA point, then $0 = -Gv_j^* + I_j - h(v_j^*, I_1/G)$. Let $e(t) = v_j(t) - v_j^*$, then from the above two equations, we obtain

$$C\dot{e}(t) = -Ge(t) - (h_1(v_j(t), v_1(t)) - h_1(v_j^*, I_1/G)). \quad (19)$$

For simplicity, we let

$$g_1(t) = \frac{h_1(v_j(t), v_1(t)) - h_1(v_j^*, v_1(t))}{v_j(t) - v_j^*}$$

$$g_2(t) = \frac{h_1(v_j^*, v_1(t)) - h_1(v_j^*, I_1/G)}{v_1(t) - I_1/G}.$$

From (19), we obtain (noticing that $v_1(t) \leq I_1/G$)

$$C\dot{e}(t) \geq -Ge(t) - [h_1(v_j(t), v_1(t)) - h_1(v_j^*, v_1(t))]$$
$$\geq -[G + g_1(t)]e(t)$$

from the comparison principle and noticing that $g_1(t) \geq 0$, we can easily obtain that

$$e(t) \geq \exp\left(-\int_{t_0}^t (G + g_1(\tau))/C d\tau\right) \geq -|e(0)|e^{-(G/C)(t-t_0)}. \quad (20)$$

On the other hand, from (19), we have

$$C\dot{e}(t) = -Ge(t) - g_1(t)e(t) - g_2(t)(v_1(t) - I_1/G)$$
$$\leq -(G + g_1(t))e(t) + |M|Le^{-(G/C)(t-t_0)}.$$

Following the same procedure as in the proof of Lemma 2.1(b) and the fact that $g_1(t) \geq 0$, we can easily obtain that there exists a positive number $M_1 > 0$ such that $e(t) \leq M_1 e^{-\gamma(t-t_0)}$ for any $\gamma < G/C$. Combining this and (20), we can obtain that there exists a positive number $M_2 > 0$ such that $|e(t)| \leq M_2 e^{-\gamma(t-t_0)}$ for any $\gamma < G/C$. This completes the proof of the second claim in the theorem.

If functions $h_i(x, y)$ $(i = 1, 2, \ldots, N)$ are differentiable, around the WTA point v^*, the linearized system is a triangular system with negative diagonal entries, hence a stable matrix, thus the WTA point is exponentially stable. This completes the proof. $\qquad\square$

Proof of Theorem 4.4: Because these three systems are special systems of type (1) whose inhibition functions belong to class $\mathcal{F}$ and satisfy the EMI condition, as we discussed before, we only need to show that the WTA points exist and the trajectories do not stay forever in the regions where at least two components are positive.

1) If (1) is true, from the positivity, then we have $KD - G > 0$. From Corollary 3.5, we know that system (7) has a WTA point. We first want to apply the Cohen–Grossberg theorem to show that trajectory (7) converges to an invariant set. Making the following variable substitution $\omega_i = v_i + D$, then system (7) becomes

$$\dot{\omega}_i = \omega_i \left\{ \left[(GD + I_i)\omega_i^{-1} - G \right] - \sum_{j \neq i} (\omega_j - D)^+ \right\}.$$

This has the form of the Cohen–Grossberg theorem equation

$$\dot{\omega}_i = a_i(\omega_i) \left\{ b_i(\omega_i) - \sum_{j=1}^{N} c_{ij} d_j(\omega_j) \right\}$$

where

$$a_i(\omega_i) = \omega_i$$
$$b_i(\omega_i) = (GD + I_i)\omega_i^{-1} - G$$
$$c_{ij} = 0 \text{ if } j = i, \text{ and } c_{ij} = K \text{ otherwise}$$
$$d_i(\omega_i) = [\omega_i - D]^+.$$

These functions meet the requirements of the Cohen–Grossberg theorem. Therefore, the trajectory approaches the largest bounded invariant set contained in the region where the time derivative of the corresponding Lyapunov function vanishes. According to Cohen–Grossberg [15], the Lyapunov function for system (7) is given by

$$V(\omega) = -\sum_{i=1}^{N} \int_0^{\omega_i} b_i(\xi_i)\dot{d}_i(\xi_i)d\xi_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} c_{ij}d_i(\omega_i)d_j(\omega_j)$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_N)^T$. From this, the Lyapunov function can be easily obtained in the original variable v_i in an explicit form

$$V(v) = \sum_{i=1}^{N} \left\{ Gv_i - (GD + I_i)\left[\ln(v_i + D) - \ln(D)\right] \right\} u(v_i)$$

$$+ \frac{1}{2}\sum_{i=1}^{N}\sum_{j \neq i} K^2 v_i^+ v_j^+$$

where $u(\cdot)$ is the unit step function.

The time derivative of the Lyapunov function is given by

$$\dot{V}(v) = -\sum_{i=1}^{N} (v_i + D)^{-1} u(v_i)$$

$$\times \left\{ -Gv_i + I_i - K(v_i + D)\sum_{j \neq i}^{N} v_j^+ \right\}^2$$

$$\leq 0.$$

According to the Cohen–Grossberg theorem [15], the trajectory of system (7) will converge to the invariant set contained in the set $\{v | \dot{V}(v) = 0\}$, i.e., the set

$$\mathcal{A} = \left\{ v \Big| u(v_i) \left\{ -Gv_i + I_i - K(v_i + D)\sum_{j \neq i}^{N} v_j^+ \right\}^2 = 0, \right.$$

$$\left. i = 1, 2, \ldots, N \right\}. \quad (21)$$

Next, we want to show that system (7) will enter the WTA region. Suppose that this is not true, then there are two positive components in the trajectory $v(t)$ for all t. Without loss of generality, assume that $I_1 > I_2 > \cdots > I_N > 0$. Then, when system (7) starts from the origin, the trajectory will enter the positive orthant initially. So, from the order-preserving property and the above assumption, we must have $v_1(t) > v_2(t) > 0$ for all $t > 0$. The trajectory will converge to the invariant set $\mathcal{A}$, so there exists a point $\bar{v} = (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_N)^T \in \mathcal{A}$ such that the trajectory $v(t)$ will converge to this point $\bar{v}$. Moreover, $\bar{v}_1 \geq \bar{v}_2 \geq 0$. It is easy to show that $\bar{v}_1 > 0$. We first want to show that $\bar{v}_2 > 0$. In fact, if $\bar{v}_2 = 0$, i.e., $\lim_{t \to \infty} v_2(t) = 0$, therefore, we can easily show that $\lim_{t \to \infty} \sum_{j \neq 1} v_j^+(t) = 0$, so $\lim_{t \to \infty} v_1(t) = I_1/G$. However, it is also easy to show that

$$\lim_{t \to \infty} \left(-Gv_2(t) + I_2 - K\left(v_2(t) + D\right)\sum_{j \neq 2} v_j^+(t) \right)$$

$$= I_2 - (KD/G)I_1 < 0.$$

This implies that $\lim_{t \to \infty} \dot{v}_2(t) < 0$, and $v_2(t)$ must become negative, which is contradicting the assumption that $v_2(t) > 0$ for all $t > 0$. Thus, we must have $\bar{v}_2 > 0$. Since $I_1 > I_2$, we can also easy to deduce that $\bar{v}_1 > \bar{v}_2$. Since $\bar{v} \in \mathcal{A}$, we have

$$-G\bar{v}_1 + I_1 - K(D + \bar{v}_1)\sum_{j \neq 1}^{N} \bar{v}_j^+ = 0$$

$$-G\bar{v}_2 + I_2 - K(D + \bar{v}_2)\sum_{j \neq 2}^{N} \bar{v}_j^+ = 0.$$

Let $Q = \sum_{j=1}^{N} \bar{v}_j^+$, then we have

$$-G\bar{v}_1 + I_1 - K(D + \bar{v}_1)(Q - v_1) = 0$$
$$-G\bar{v}_2 + I_2 - K(D + \bar{v}_2)(Q - \bar{v}_2) = 0. \quad (22)$$

Subtracting the second equation from the first equation, we obtain

$$(\bar{v}_1 - \bar{v}_2)\left[K^{-1}(KQ - KD + G) - (\bar{v}_1 + \bar{v}_2) \right] = K^{-1}(I_1 - I_2) > 0.$$

From $\bar{v}_1 > \bar{v}_2$, we must have $KQ - KD + G > 0$, i.e.,

$$Q > K^{-1}(KD - G). \tag{23}$$

Solving the two equations in (22) for $\bar{v}_1, \bar{v}_2$ in an explicit form, we have

$$\bar{v}_1 = \frac{1}{2}K^{-1}(KQ - KD + G)$$
$$\pm \frac{1}{2}K^{-1}\sqrt{(KQ - KD + G)^2 - 4K(I_1 - KDQ)}$$
$$\bar{v}_2 = \frac{1}{2}K^{-1}(KQ - KD + G)$$
$$\pm \frac{1}{2}K^{-1}\sqrt{(KQ - KD + G)^2 - 4K(I_2 - KDQ)}$$

We want to show that the $\pm$ in the second equation must be a minus sign. In fact, suppose that it is a plus sign. If $\pm$ in the first equation is minus, obviously, $\bar{v}_1 < \bar{v}_2$, which is a contradiction. If $\pm$ in the first equation is a plus sign, then since $I_1 > I_2$, we also have $\bar{v}_1 < \bar{v}_2$, which contradicts the assumption. So $\pm$ in the second equation must be a minus. However, in this case, from the positivity of $\bar{v}_2$, we must have $4K(I_2 - KDQ) > 0$, or $I_2/(KD) > Q$. Combining this with (23), we have $I_2/(KD) > (KD - G)/K$, i.e., $I_2 > D(KD - G)$, which contradicts the condition in (1).

Therefore, we have proved that the trajectory of (7) starting from the origin cannot stay forever in the region where at least two components are positive, thus the trajectory will enter the WTA region at a certain instant. From Theorem 4.3, the trajectory will exponentially converge to the WTA point. This proves the first part of the theorem.

2) The first condition in (2) guarantees that system (8) has a WTA point from Corollary 3.4. We only need to show that the trajectory of (8) cannot stay forever in the region which has at least two positive components. Suppose that this is not true, then we must have for all time $t > 0$ $v_1(t) > v_2(t) > 0$ (without loss of generality, we assume that $I_1 > I_2 > \cdots > I_N > 0$). The first two equations of (8) become

$$\dot{v}_1 = -Gv_1 + I_1 - Kv_2^2 - \sum_{j \neq 1,2} K\left[v_j^+\right]^2$$
$$\dot{v}_2 = -Gv_2 + I_2 - Kv_1^2 - \sum_{j \neq 1,2} K\left[v_j^+\right]^2.$$

Let $e(t) = v_1(t) - v_2(t)$. Subtracting the second equation from the first equation, we obtain

$$\dot{e}(t) = -Ge(t) + (I_1 - I_2) + K\left(v_1^2 - v_2^2\right)$$
$$= -Ge(t) + (I_1 - I_2) + K(v_1 + v_2)e(t)$$
$$\geq -Ge(t) + (I_1 - I_2) + Ke(t)^2$$
$$= K\left(e(t) - \frac{G}{2K}\right)^2 + (I_1 - I_2) - \frac{G^2}{4K}$$
$$\geq (I_1 - I_2) - \frac{G^2}{4K} > 0.$$

From the comparison principle, $\lim_{t \to \infty} e(t) = \infty$. This contradicts the fact that the trajectory of (8) is bounded, so are $v_1(t)$ and $v_2(t)$. This completes the proof of the second part.

3) There is a simple proof for this part without using the Cohen–Grossberg theorem. We only need to show that the tra-

jectory of (9) starting from the origin does not stay forever in the region where at least two components are positive. Suppose that this is not true, then noticing fact $I_1 > I_2 > 0$ and the order-preserving property, we must have $v_1(t) > v_2(t) > 0$ for all $t > 0$. Let $e(t) = v_1(t) - v_2(t)$, then we have

$$\dot{e}(t) = -Ge(t) + (I_1 - I_2) - \left[\sum_{j \neq 1} Kv_j^+ - \sum_{j \neq 2} Kv_j^+\right]$$
$$= (K - G)e(t) + (I_1 - I_2).$$

Since $K > G$, this 1-D system diverges to infinity. On the other hand, the trajectory of system (9) is bounded (Theorem 4.1), so is $e(t) = v_1(t) - v_2(t)$. This is a contradiction. Therefore, the trajectory of (9) starting from the origin will enter the WTA region. From Theorem 4.3, we conclude that system (9) starting from the origin will exponentially converge to the WTA point. This completes the proof. $\square$

REFERENCES

[1] J. A. Feldman and D. H. Ballard, "Connectionist models and their properties," *Cogn. Sci.*, vol. 6, pp. 205–254, 1982.

[2] S. Haykin, *Neural Networks: A Comprehensive Foundation*. New York: Macmillan, 1994.

[3] S. Kaski and T. Kohonen, "Winner-take-all networks for physiological models of competitive learning," *Neural Netw.*, vol. 7, pp. 973–984, 1994.

[4] T. Kohonen, "Physiological interpretation of the self-organizing map algorithm," *Neural Netw.*, vol. 6, pp. 895–905, 1993.

[5] B. Nabet and R. B. Pinter, *Sensory Neural Networks: Lateral Inhibition*. Boston, MA: CRC Press, 1991.

[6] S. Grossberg, "Contour enhancement, short term memory, and constancies in reverberating neural networks," *Studies Appl. Math.*, vol. 52, pp. 213–257, 1973.

[7] B. Ermentrout, "Complex dynamics in winner-take-all neural nets with slow inhibition," *Neural Netw.*, vol. 5, pp. 415–431, 1992.

[8] M. Lemmon and B. V. K. V. Kumar, "Emulating the dynamics for a class of laterally inhibited neural networks," *Neural Netw.*, vol. 2, pp. 193–214, 1989.

[9] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE Acoust. Speech Signal Process. Mag.*, vol. 4, no. 4, pp. 4–22, Apr. 1987.

[10] J. F. Yang, C. M. Chen, W. C. Wang, and J. Y. Lee, "A general mean-based iterative winner-take-all neural network," *IEEE Trans. Neural Netw.*, vol. 6, no. 1, pp. 14–24, Jan. 1995.

[11] J. C. Yen, F. J. Chang, and S. Chang, "A new winners-take-all architecture in artificial neural networks," *IEEE Trans. Neural Netw.*, vol. 5, no. 5, pp. 838–843, Sep. 1994.

[12] J. Lazzaro, S. Ryckebusch, M. A. Mahowald, and C. A. Mead, "Winner-take-all networks of O(N) complexity," in *Advances in Neural Information Processing Systems I*, D. S. Touretzky, Ed. San Mateo, CA: Morgan Kaufmann, 1989, pp. 703–711.

[13] E. Majani, R. Erlanson, and Y. Abu-Mostafa, "On the K-winners-take-all network," in *Advances in Neural Information Processing Systems I*, D. S. Touretzky, Ed. San Mateo, CA: Morgan Kaufmann, 1989, pp. 634–642.

[14] J. S. Kane and T. G. Kincaid, "Optoelectric winner-take-all VLSI shunting neural networks," *IEEE Trans. Neural Netw.*, vol. 6, no. 5, pp. 1275–1279, Sep. 1995.

[15] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel pattern formation by competitive neural networks," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-13, no. 5, pp. 815–826, Sep.-Oct. 1983.

[16] S. Grossberg, "Nonlinear neural networks: Principles, mechanisms, and architectures," *Neural Netw.*, vol. 1, pp. 17–61, 1988.

[17] S. Grossberg, "Competition, decision, and consensus," *J. Math. Anal. Appl.*, vol. 66, pp. 470–493, 1978.

[18] M. W. Hirsch, "Systems of differential equations which are competitive or cooperative, I: Limit sets," *SIAM J. Math. Anal.*, vol. 13, pp. 167–179, 1982.

[19] M. W. Hirsch, "Convergent activation dynamics in continuous time networks," *Neural Netw.*, vol. 2, pp. 331–349, 1989.

[20] J. J. Hopfield, "Neurons with graded responses have collective computational properties like those of two-state neurons," *Proc. Nat. Acad. Sci. USA*, vol. 81, pp. 3088–3092, 1984.

[21] V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities: Theory and Applications*. New York: Academic, 1969, vol. I, Ordinary Differential Equations.

[22] M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd ed. : Prentice-Hall, 1993, pp. 248–249.

[23] A. S. Sedra and K. C. Smith, *Micro-Electronic Circuits*, 3rd ed. New York: HRW, 1991.

[24] K. Wang and A. N. Michel, "Robustness and perturbation analysis of a class of artificial neural networks," *Neural Netw.*, vol. 7, no. 2, pp. 251–260, 1994.

[25] K. N. Wang and A. N. Michel, "Stability analysis of differential inclusions in Banach space with applications to nonlinear systems with time delays," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 43, no. 8, pp. 617–626, Aug. 1996.

[26] H. Ye, A. N. Michel, and K. N. Wang, "Robust stability of nonlinear time-delay systems with applications to neural networks," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 43, no. 7, pp. 532–543, Jul. 1996.

[27] S. P. Tonkin, R. B. Pinter, and B. Nabet, "Stability of a class of directionally sensitive asymmetric nonlinear neural networks," *Neural Netw.*, vol. 9, no. 4, pp. 555–565, 1996.

[28] Y. Fang and T. G. Kincaid, "Stability analysis of dynamical neural networks," *IEEE Trans. Neural Netw.*, vol. 7, no. 4, pp. 996–1006, Jul. 1996.

[29] K. Wawryn and B. Strzeszewski, "Current mode circuits for programmable WTA neural network," *Analog Integrated Circuits Signal Process.*, vol. 27, pp. 49–69, 2001.

[30] B. D. Calvert and C. A. Marinov, "Another K-winner-take-all analog neural network," *IEEE Trans. Neural Netw.*, vol. 11, no. 4, pp. 829–838, Jul. 2000.
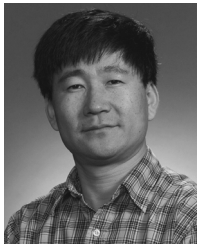
Yuguang Fang (F'08) received the Ph.D. degree in systems engineering from Case Western Reserve University, Cleveland, OH, in January 1994 and the Ph.D. degree in electrical engineering from Boston University, Boston, MA, in May 1997.

He was an Assistant Professor at the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Hoboken, from July 1998 to May 2000. He then joined the Department of Electrical and Computer Engineering, University of Florida, Gainesville, in May 2000, as an Assistant Professor, got an early promotion to an Associate Professor with tenure in August 2003 and to a Full Professor in August 2005. He holds a University of Florida Research Foundation (UFRF) Professorship from 2006 to 2009, a Changjiang Scholar Chair Professorship with Xidian University, Xi'an, China, from 2008 to 2011, and a Guest Chair Professorship with Tsinghua University, China, from 2009 to 2012. He has published over 250 papers in refereed professional journals and conferences. He is also active in professional activities.

Dr. Fang received the National Science Foundation Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002, and is the recipient of the Best Paper Award at the IEEE International Conference on Network Protocols (ICNP) in 2006 and the recipient of the IEEE TCGN Best Paper Award at the IEEE High-Speed Networks Symposium, IEEE Globecom in 2002. He is a Fellow of IEEE and a member of ACM. He is currently serving as the Editor-in-Chief for the IEEE WIRELESS COMMUNICATIONS and serves/served on several editorial boards of technical journals including the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE WIRELESS COMMUNICATIONS MAGAZINE and *ACM Wireless Networks*. He was an editor for the IEEE TRANSACTIONS ON MOBILE COMPUTING and currently serves on its Steering Committee. He has been actively participating in professional conference organizations such as serving as the Steering Committee Co-Chair for QShine from 2004 to 2008, the Technical Program Committee Vice-Chair for IEEE INFOCOM'2005, the Area Technical Program Committee Chair for IEEE INFOCOM (2009–2010), the Technical Program Symposium Co-Chair for IEEE Globecom'2004, and a member of Technical Program Committee for IEEE INFOCOM (1998, 2000, 2003–2008) and a member of Technical Program Committee for ACM Mobihoc (2008–2009).

Michael A. Cohen was born and raised in Whitestone, NY. He received the undergraduate degree in pure mathematics from Massachusetts Institute of Technology (MIT), Cambridge, in 1969 and the Ph.D. degree in experimental psychology from Harvard University, Cambridge, MA, in 1979.

During 1972–1974, he was a student at the MIT Linguistics Department and a Staff Programmer at Research Lab of Electronics. He left and took a job from 1975 to 1976 as a Professional Programmer. From 1979 to 1980, he was a Postdoctoral Fellow at New York University. In 1980, he became a Research Assistant Professor at Boston University, Boston, MA, and later an Associate Professor in 1989. He has expertise in many areas and has published papers in dynamical systems, child language development, mathematical measurement theory, visual psychophysics and audition, and currently works in cardiovascular physiology.
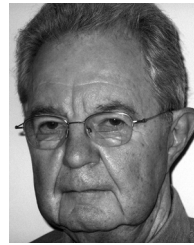
Thomas G. Kincaid (S'58–M'65) was born in Canada in 1937. He received the B.Sc. degree in engineering physics from Queen's University, Kingston, U.K., in 1959, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1962 and 1965, respectively.

In 1965, he joined the General Electric Company at the Corporate Research and Development Center, Schenectady, NY, where he worked on sonar system design and power line carrier communications. In 1977, he became Manager of the program to develop nondestructive evaluation systems and guided the development of industrial tomography, ultrasonic microscopy, and eddy current sensor array systems. In 1983, he became a Professor of Electrical, Computer and Systems Engineering at Boston University, Boston, MA, chairing the Department until 1994. He retired in 2005. His research interests include signals and systems, and dynamical neural networks.