

- [18] A. J. Snell, W. E. Spear, P. G. LeComber, and K. Mackenzie, "Application of amorphous silicon thin film transistors in integrated circuits," *Appl. Phys.*, vol. A26, pp. 83–86, Jul. 1981.
- [19] Y. Nara, Y. Kudou, and M. Matsumura, "Application of amorphous-silicon thin-film transistors in three-dimensional integrated circuits," *Jap. J. Appl. Phys.*, vol. 22, no. 6, pp. L370–L372, Jun. 1983.
- [20] C. E. Stafstrom, P. C. Schwindt, and W. E. Crill, "Repetitive firing in layer V neurons from cat neocortex in vitro," *J. Neurophysiol.*, vol. 52, no. 2, pp. 264–277, Aug. 1984.
- [21] C. Mead, "Neuromorphic electronic systems," *Proc IEEE*, vol. 78, no. 10, pp. 1629–1636, Oct. 1990.
- [22] G. Indiveri, E. Chicca, and R. Douglas, "A VLSI array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 211–221, Jan. 2006.
- [23] C. Diorio, P. Hasler, B. A. Minch, and C. A. Mead, "A single transistor silicon synapse," *IEEE Trans. Electron Devices*, vol. 43, no. 11, pp. 1972–1980, Nov. 1996.
- [24] B. Linares-Barranco and T. Terrano-Gotarredona, "Exploiting memorance in adaptive asynchronous spiking neuromorphic nanotechnology systems," in *Proc. 9th IEEE Conf. Nanotechnol.*, Jul. 2009, pp. 978–981.
- [25] K. D. Cantley, A. Subramaniam, H. Stiegler, R. Chapman, and E. Vogel, "Spike timing-dependent synaptic plasticity using memristors and nanocrystalline silicon TFT memories," in *Proc. 11th IEEE Conf. Nanotechnol.*, Aug. 2011, pp. 421–425.
- [26] K. D. Cantley, A. Subramaniam, H. J. Stiegler, R. A. Chapman, and E. M. Vogel, "Hebbian Learning in Spiking Neural Networks with Nanocrystalline Silicon TFTs and Memristive Synapses," *IEEE Trans. Nanotechnol.*, vol. 10, no. 5, pp. 1066–1073, Sep. 2011.
- [27] M. J. Powell, C. van Berkel, I. D. French, and D. H. Nicholls, "Bias dependence of instability mechanisms in amorphous silicon thin-film transistors," *Appl. Phys. Lett.*, vol. 51, no. 16, pp. 1242–1244, Oct. 1987.
- [28] Y.-S. Tai, J.-W. Tsai, H.-C. Chung, and F.-C. Su, "Instability mechanisms for the hydrogenated amorphous silicon thin-film transistors with negative and positive bias stresses on the gate electrodes," *Appl. Phys. Lett.*, vol. 67, no. 1, pp. 76–78, Jul. 1995.
- [29] M. R. Esmaili-Rad, A. Sazonov, and A. Nathan, "Absence of defect state creation in nanocrystalline silicon thin film transistors deduced from constant current stress measurements," *Appl. Phys. Lett.*, vol. 91, no. 11, pp. 113511-1–113511-3, Sep. 2007.
- [30] D. R. Allee, L. T. Clark, B. D. Vogt, R. Shringarpure, S. M. Venugopal, S. G. Uppili, K. Kaftanoglu, H. Shivalingaiyah, Z. P. Li, J. J. Ravindra Fernando, E. J. Bawolek, and S. M. O'Rourke, "Circuit-level impact of a-Si:H thin-film-transistor degradation effects," *IEEE Trans. Electron Devices*, vol. 56, no. 6, pp. 1166–1176, Jun. 2009.
- [31] N. Joye, A. Schmid, Y. Leblebici, T. Asai, and Y. Amemiya, "Fault-tolerant logic gates using neuromorphic CMOS circuits," in *Proc. Res. Microelectron. Electron. Conf.*, Jul. 2007, pp. 249–252.
- [32] W. Liu, *MOSFET Models for SPICE Simulation: Including BSIM3v3 and BSIM4*, 1st ed. New York, NY, USA: Wiley, Feb. 2001.
- [33] C.-H. Lee, A. Sazonov, and A. Nathan, "High hole and electron mobilities in nanocrystalline silicon thin-film transistors," *J. Non Crystalline Solids*, vol. 352, nos. 9–20, pp. 1732–1736, Jun. 2006.

Effect of Input Noise and Output Node Stochastic on Wang's k WTA

John Sum, *Senior Member, IEEE*, Chi-Sing Leung, *Member, IEEE*, and Kevin Ho

Abstract—Recently, an analog neural network model, namely Wang's k WTA, was proposed. In this model, the output nodes are defined as the Heaviside function. Subsequently, its finite time convergence property and the exact convergence time are analyzed. However, the discovered characteristics of this model are based on the assumption that there are no physical defects during the operation. In this brief, we analyze the convergence behavior of the Wang's k WTA model when defects exist during the operation. Two defect conditions are considered. The first one is that there is input noise. The second one is that there is stochastic behavior in the output nodes. The convergence of the Wang's k WTA under these two defects is analyzed and the corresponding energy function is revealed.

Index Terms—Convergence analysis, energy function, input noise, k WTA, output node stochastic.

I. INTRODUCTION

In conventional design, a traditional k WTA network consists of n nodes and n^2 connections [2], [10], [13]. By reformulating the k winners selection problem as a linear program, Jun Wang and his coworkers succeeded a much simpler structure to realize a k WTA network based on the idea of dual neural network (DNN) [5], [17], [18]. Here, we call it the Wang's k WTA model. This model consists of n output nodes (which are defined as Heaviside activation functions), $2n$ connections, and a hidden node (which behaves as a recurrent state variable).¹ Because of the Heaviside activation function, Wang in [18] proved the finite time convergence of this k WTA and later Xiao *et al.* in [20] derived the exact convergence time.

Because of its simple structure, this model is particularly suitable for hardware implementation. However, as known in the studies of fault tolerant neural network. Hardware implementation can never be perfect [8], [19], [21], [22]. For example, the electrical noise [21] from the power supplies or thermal effect may be injected into the input signals. In addition, the offset voltage of comparators may drift randomly [22]. In this regard, it is valuable to analyze the effects of input noise and output node stochastic on the dynamic behavior of the Wang's k WTA model.

Manuscript received May 26, 2012; accepted March 24, 2013. Date of publication May 7, 2013; date of current version August 16, 2013. This work was supported in part by Research Grants from Taiwan National Science Council under Grant 100-2221-E-126-015 and Grant 101-2221-E-126-016, and a Research Grant under Grant CityU 115612 from the Research Grants Council of the Government of the Hong Kong Special Administrative Region.

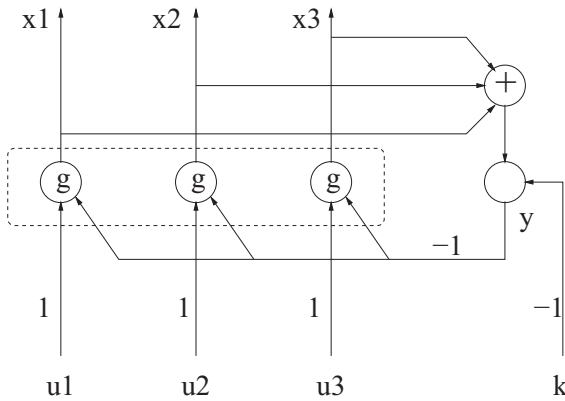
J. Sum is with the Institute of Technology Management, National Chung Hsing University, Taichung 40227, Taiwan (e-mail: psum@yahoo.com.hk).

C.-S. Leung is with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: eeleung@cityu.edu.hk).

K. Ho is with the Department of Computer Science and Communication Engineering Providence University, Taichung 43301, Taiwan (e-mail: kevinijho@yahoo.com).

Digital Object Identifier 10.1109/TNNLS.2013.2257182

Note that an identical model has been proposed independently in [9].


 Fig. 1. Structure of a Wang's k WTA.

In [16], we already analyzed the effect of output node stochastic on the Wang's k WTA model. The node stochastic is defined as a Logit function. That is to say, the probability that a node will output one follows the logistic function. This model is commonly used to model the faulty behavior of a neuron [1], [6], [8], [14] because of imperfect hardware implementation.

In this brief, we consider the Wang's k WTA model under two defect conditions. The first one is that there are input noises in the input signals. We assume that the inputs are corrupted by zero mean Gaussian noise. The second one is that the output nodes have stochastic behavior. We call the Wang's k WTA model under the two defect conditions as the stochastic Wang's k WTA model. Our key contributions are as follows.

- 1) We show that there exists a unique equilibrium point in the dynamics of the stochastic Wang's k WTA model. In addition, we investigate the position of the unique equilibrium point.
- 2) Since existing an equilibrium point does means that the state of the network will converge to the equilibrium point, this brief further shows that the dynamics of the stochastic Wang's k WTA model converges the unique equilibrium point.
- 3) We also reveal the energy function of the stochastic Wang's k WTA model. From the energy function, we briefly discuss the rate convergence of the stochastic Wang's k WTA model.

In the next section, the original Wang's k WTA model will be reviewed. The mathematical model of the stochastic Wang's k WTA model is then defined in Section III. Comprehensive analysis of this stochastic k WTA model is presented in Section IV. Finally, we conclude this brief in Section V.

II. WANG'S k WTA

Fig. 1 shows the structure of the Wang's k WTA model. It consists of three inputs. For a general n inputs k WTA network, the inputs are denoted as u_1, u_2, \dots, u_n , and the outputs are denoted as x_1, x_2, \dots, x_n . Without loss of generality, we assume that the values of u_i s are all distinct and bounded by zero and one. Following Wang's notation, we

denote $\bar{u}_n, \bar{u}_{n-1}, \dots, \bar{u}_{n-k+1}$ as the k largest numbers, and $0 \leq \bar{u}_1 < \dots < \bar{u}_n \leq 1$. The operation of the Wang's k WTA network is given by the following state-space system:¹

$$\epsilon \frac{dy(t)}{dt} = \sum_{i=1}^n x_i(t) - k \quad (1)$$

$$x_i(t) = g(u_i - y(t)) \quad (2)$$

where $g(s)$ is a comparator function

$$g(s) = \begin{cases} 1, & \text{if } s > 0 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

for $i = 1, 2, \dots, n$. The value k in (1) is specified by the user. The state-variable y corresponds to the output of the rightmost node in Fig. 1. By (1) and (2), $y(t)$ converges in finite time [18], [20].

III. STOCHASTIC WANG'S k WTA

As hardware implementation can never be perfect [21], [22], there are some stochastic behavior during the operation. For example, the electrical noise from the power supplies or thermal effect may be injected into the input signals. In addition, the output nodes may have stochastic behavior. This behavior may come from the random drift in the offset voltage of comparators [22].

A. Model

The i th input is denoted as $\tilde{u}_i(t) = u_i + \delta_i(t)$. It is composed of a constant value u_i and a random noise $\delta_i(t)$, where the probability density function of $\delta_i(t)$ is given by the following:

$$P(\tilde{u}_i(t)|u_i) = \sqrt{\frac{\alpha_I}{2\pi}} \exp\left(-\frac{\alpha_I(\tilde{u}_i(t) - u_i)^2}{2}\right) \quad (4)$$

the parameter α_I corresponds to the inverse of the noise variance.

The i th output at time t is denoted as $\tilde{x}_i(t) \in \{0, 1\}$. If the comparator function (3) has the stochastic behavior, the dynamics of the stochastic Wang's k WTA is given as follows:

$$\epsilon \frac{dy}{dt} = \sum_{i=1}^n \tilde{x}_i(t) - k \quad (5)$$

where the probability mass function of the output is given by

$$\begin{aligned} \text{Probit}(\tilde{x}_i(t)) &= 1|y(t), \tilde{u}_i(t) \\ &= \sqrt{\frac{\alpha_O}{2\pi}} \int_{y(t)}^{\infty} \exp\left(-\frac{\alpha_O(z - \tilde{u}_i(t))^2}{2}\right) dz \end{aligned} \quad (6)$$

¹[18, eqs. (29) and (30)].

for all $i = 1, 2, \dots, n$. The parameter α_O in (6) controls the shape of the Probit function. It controls the stochastic behavior of the output nodes. If this stochastic behavior comes from the random drift of the offset voltage, the parameter α_O corresponds to the inverse of the random drift's variance and the distribution of the random drift is zero mean Gaussian.

The output variable $x_i(t)$ is now a binary random variable. The definition of the dynamical change of $y(t)$ in (5) is not formal. Precisely, the dynamical change of $y(t)$ should be given in integral form as follows:

$$\epsilon(y(t + \tau) - y(t)) = \sum_{i=1}^n \int_t^{t+\tau} \tilde{x}_i(\tau) d\tau - k\tau \quad (7)$$

or

$$y(t + \tau) = y(t) + \frac{1}{\epsilon} \left\{ \sum_{i=1}^n \int_t^{t+\tau} \tilde{x}_i(\tau) d\tau - k\tau \right\}. \quad (8)$$

For τ is small, we have

$$y(t + \tau) = y(t) + \frac{\tau}{\epsilon} \left\{ \sum_{i=1}^n \text{Prob}(\tilde{x}_i(t) = 1|y(t), u_i) - k \right\} \quad (9)$$

where

$$\begin{aligned} \text{Prob}(\tilde{x}_i(t) = 1|y(t), u_i) \\ = \int_{-\infty}^{\infty} \text{Probit}(\tilde{x}_i(t) = 1|y(t), \tilde{u}_i) P(\tilde{u}_i|u_i) d\tilde{u}_i. \end{aligned} \quad (10)$$

By (4) and (6), (10) can be rewritten as follows:

$$\begin{aligned} \text{Prob}(\tilde{x}_i(t) = 1|y(t), u_i) \\ = \frac{\sqrt{\alpha_O \alpha_I}}{2\pi} \int_{y(t)}^{\infty} \left[\int_{-\infty}^{\infty} \exp\left(-\frac{\alpha_O(z - \tilde{u}_i)^2}{2}\right) \right. \\ \left. \times \exp\left(-\frac{\alpha_I(\tilde{u}_i - u_i)^2}{2}\right) d\tilde{u}_i \right] dz. \end{aligned} \quad (11)$$

It should be noted that

$$\begin{aligned} \int_{-\infty}^{\infty} \exp\left(-\frac{\alpha_O(z - \tilde{u}_i)^2}{2}\right) \exp\left(-\frac{\alpha_I(\tilde{u}_i - u_i)^2}{2}\right) d\tilde{u}_i \\ = \int_{-\infty}^{\infty} \exp\left(-\frac{\alpha_O(z - \tilde{u}_i)^2}{2} - \frac{\alpha_I(\tilde{u}_i - u_i)^2}{2}\right) d\tilde{u}_i. \end{aligned} \quad (12)$$

Let $\delta_i = \tilde{u}_i - u_i$, we can obtain

$$\begin{aligned} & -\frac{\alpha_O(z - \tilde{u}_i)^2}{2} - \frac{\alpha_I(\tilde{u}_i - u_i)^2}{2} \\ & = -\frac{\alpha_O}{2}(z - u_i - \delta_i)^2 - \frac{\alpha_I}{2}\delta_i^2 \\ & = -\frac{\alpha_O}{2}(z - u_i)^2 - \frac{\alpha_O + \alpha_I}{2}\delta_i^2 + \alpha_O(z - u_i)\delta_i \\ & = -\frac{\alpha_O \alpha_I (z - u_i)^2}{2(\alpha_O + \alpha_I)} - \frac{\alpha_O + \alpha_I}{2} \left(\delta_i - \frac{\alpha_O(z - u_i)}{\alpha_O + \alpha_I} \right)^2. \end{aligned} \quad (13)$$

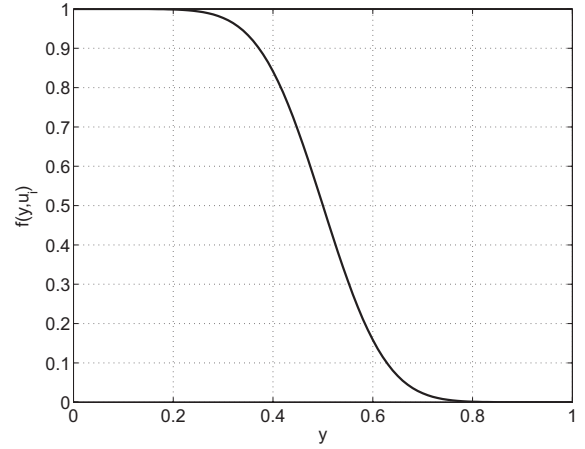


Fig. 2. $f(y, u_i)$ versus y for $u_i = 0.5$ and $\alpha = 100$.

Put (13) in (12), we can obtain

$$\begin{aligned} & \int_{-\infty}^{\infty} \exp\left(-\frac{\alpha_O(z - \tilde{u}_i)^2}{2} - \frac{\alpha_I(\tilde{u}_i - u_i)^2}{2}\right) d\tilde{u}_i \\ & = \exp\left(-\frac{\alpha_O \alpha_I (z - u_i)^2}{2(\alpha_O + \alpha_I)}\right) \\ & \quad \times \int_{-\infty}^{\infty} \exp\left[-\frac{\alpha_O + \alpha_I}{2} \left(\delta_i - \frac{\alpha_O(z - u_i)}{\alpha_O + \alpha_I} \right)^2\right] d\delta_i \\ & = \sqrt{\frac{2\pi}{\alpha_O + \alpha_I}} \exp\left(-\frac{\alpha_O \alpha_I (z - u_i)^2}{2(\alpha_O + \alpha_I)}\right). \end{aligned} \quad (14)$$

Put (14) in (11), we finally obtain

$$\begin{aligned} \text{Prob}(\tilde{x}_i(t) = 1|y(t), u_i) \\ = \sqrt{\frac{\alpha_O \alpha_I}{2\pi(\alpha_O + \alpha_I)}} \int_{y(t)}^{\infty} \exp\left(-\frac{\alpha_O \alpha_I (z - u_i)^2}{2(\alpha_O + \alpha_I)}\right) dz. \end{aligned} \quad (15)$$

For simplicity, we let²

$$f(y(t), u_i) = \text{Prob}(\tilde{x}_i(t) = 1|y(t), u_i) \quad (16)$$

$$\alpha = \frac{\alpha_O \alpha_I}{\alpha_O + \alpha_I}. \quad (17)$$

With (16) and (17), the probability mass function of the i th output node is given by the following:

$$f(y(t), u_i) = \sqrt{\frac{\alpha}{2\pi}} \int_{y(t)}^{\infty} \exp\left(-\frac{\alpha(z - u_i)^2}{2}\right) dz \quad (18)$$

where the parameter α describes the overall effect of the input noise and the output node stochastic behavior. Fig. 2 shows the shape of $f(y, u_i)$ for $u_i = 0.5$ and $\alpha = 100$.

It should be noted from (9), (10), (15), and (18) that the analysis on the dynamic behavior of the stochastic Wang's k WTA can equally be applied to other special cases as shown in Table I.

²One should note that $f(y(t), u_i)$ specifies the firing rate of the i th output node.

TABLE I
 SPECIAL CASES

α_I	α_O	α	Description
$< \infty$	$< \infty$	$< \infty$	Input noise and stochastic output nodes
$< \infty$	$= \infty$	$= \alpha_I$	Input noise only
$= \infty$	$< \infty$	$= \alpha_O$	Stochastic output nodes only
$= \infty$	$= \infty$	$= \infty$	Original Wang's kWTA

B. Illustrative Example

Fig. 3 shows four cases on the dynamics of $y(t)$. The values of the input variables are 0.5, 0.7, 0.8, 0.4, 0.1, and 0.3, where $\epsilon = 0.0001$. The parameter α is set to 20, 40, 100, and ∞ . The stochastic model with $\alpha = \infty$ corresponds to the original Wang's kWTA. It is clear that the convergence behavior of $y(t)$ and the convergence time are quite different from those obtained in the original Wang's kWTA. For the case that $y(0) = 0$ ($y(0) = 1$), the dynamics of $y(t)$ can be roughly distinguished in two phases. In the initial phase, the change of $y(t)$ is rapid. Once $y(t)$ has reached a value above 0.5 (below 0.7), it increases (decreases) in a rate proportional to $\log(t)$. If the value of α increases, the rate further reduces.

IV. ANALYSIS

With defects in the model, it is critical to investigate if the network converges. If $y(t)$ converges, where will it be. To answer these questions, we need to analyze the dynamics of $y(t)$. This section first shows that $y(t)$ converges. Then, the convergent of $y(t)$ is derived. Finally, the energy function is revealed. Without loss of generality, our analysis is based on the general form of $f(y(t), u_i)$ in (16).

A. Equilibrium Point

By (9), (16), and (18), we obtain

$$y(t + \tau) = y(t) + \frac{\tau}{\epsilon} \left\{ \sum_{i=1}^n f(y(t), u_i) - k \right\} \quad (19)$$

for τ is small.

Theorem 1 (Existence of Equilibrium): Consider stochastic Wang's kWTA with bounded inputs u_1, u_2, \dots, u_n . For any constant α , there exists a unique equilibrium point y^* such that $\sum_{i=1}^n f(y^*, u_i) - k = 0$. In addition, if α is sufficient large, $y^* = (\bar{u}_{n-k} + \bar{u}_{n-k+1})/2$, where $\alpha = (\alpha_O \alpha_I) / (\alpha_O + \alpha_I)$.

Proof: For the uniqueness property, we let

$$F(y) = \sum_{i=1}^n f(y(t), u_i) - k. \quad (20)$$

It is thus clear that $F'(y) < 0$ as $f'(y, u_i) < 0$ for all $i = 1, \dots, n$. In virtue of $\lim_{y \rightarrow -\infty} F(y) > 0$ and $\lim_{y \rightarrow \infty} F(y) < 0$, there exists a unique y^* such that $F(y^*) = 0$.

Recall that $0 < \bar{u}_1 < \bar{u}_2 < \dots < \bar{u}_n < 1$. For $\alpha \gg 1$, we can assume that

$$f(y^*, \bar{u}_1) = \dots = f(y^*, \bar{u}_{n-k-1}) = 0$$

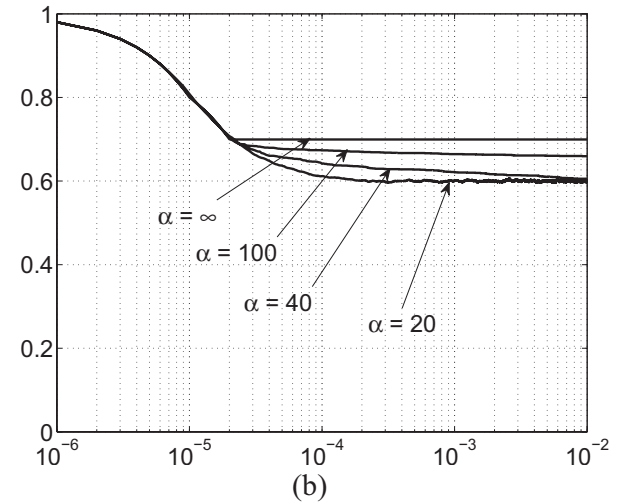
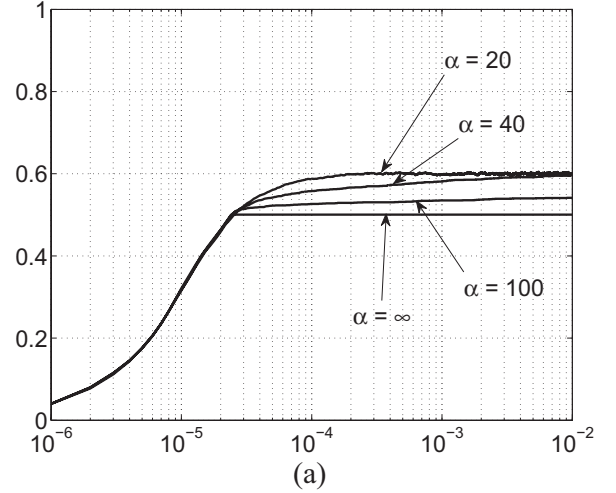


Fig. 3. $y(t)$ versus t . (a) $y(0) = 0$ and (b) $y(0) = 1$. Note that the scale on the time is in log scale. The inputs are 0.5, 0.7, 0.8, 0.4, 0.1, and 0.3 respectively. $k = 2$, $\epsilon = 0.0001$, and $\alpha = 20, 40, 100$ and ∞ .

$$\text{and } f(y^*, \bar{u}_{n-k+1}) = \dots = f(y^*, \bar{u}_n) = 1.$$

In this regard

$$f(y^*, \bar{u}_{n-k}) + f(y^*, \bar{u}_{n-k+1}) - 1 = 0$$

or

$$f(y^*, \bar{u}_{n-k}) = 1 - f(y^*, \bar{u}_{n-k+1}).$$

On the other hand, we can obtain by (18) that

$$\begin{aligned} & \sqrt{\frac{\alpha}{2\pi}} \int_{y^*}^{\infty} \exp\left(-\frac{\alpha(z - \bar{u}_{n-k})^2}{2}\right) dz \\ &= \frac{1}{2} - \sqrt{\frac{\alpha}{2\pi}} \int_0^{y^* - \bar{u}_{n-k}} \exp\left(-\frac{\alpha z^2}{2}\right) dz. \end{aligned}$$

Note that $y^* < \bar{u}_{n-k+1}$

$$\begin{aligned} & \sqrt{\frac{\alpha}{2\pi}} \int_{y^*}^{\infty} \exp\left(-\frac{\alpha(z - \bar{u}_{n-k+1})^2}{2}\right) dz \\ &= \frac{1}{2} + \sqrt{\frac{\alpha}{2\pi}} \int_0^{\bar{u}_{n-k+1} - y^*} \exp\left(-\frac{\alpha z^2}{2}\right) dz. \end{aligned}$$

Thus $f(y^*, \bar{u}_{n-k}) = 1 - f(y^*, \bar{u}_{n-k+1})$ and hence

$$\begin{aligned} & \sqrt{\frac{\alpha}{2\pi}} \int_0^{y^* - \bar{u}_{n-k}} \exp\left(-\frac{\alpha z^2}{2}\right) dz \\ &= \sqrt{\frac{\alpha}{2\pi}} \int_0^{\bar{u}_{n-k+1} - y^*} \exp\left(-\frac{\alpha z^2}{2}\right) dz. \end{aligned}$$

Therefore, $y^* - \bar{u}_{n-k} = \bar{u}_{n-k+1} - y^*$ and then $y^* = (\bar{u}_{n-k} + \bar{u}_{n-k+1})/2$. The proof is completed. ■

From the above theorem, we can find that the firing rates of the output nodes with inputs \bar{u}_{n-k} and \bar{u}_{n-k+1} are

$$\frac{1}{2} - \sqrt{\frac{\alpha}{2\pi}} \int_0^{\frac{\bar{u}_{n-k+1} - \bar{u}_{n-k}}{2}} \exp\left(-\frac{\alpha z^2}{2}\right) dz$$

and

$$\frac{1}{2} + \sqrt{\frac{\alpha}{2\pi}} \int_0^{\frac{\bar{u}_{n-k+1} - \bar{u}_{n-k}}{2}} \exp\left(-\frac{\alpha z^2}{2}\right) dz$$

respectively.

B. Convergence Proof

Now, we can state the convergence theorem regarding the convergence of $y(t)$ defined by (4)–(6).

Theorem 2 (Convergence): Consider stochastic Wang's k WTA with bounded inputs u_1, u_2, \dots, u_n . For any constant α , $\lim_{t \rightarrow \infty} y(t) = y^*$. If furthermore, α (i.e., $\alpha_O \alpha_I / \alpha_O + \alpha_I$) is sufficient large, $\lim_{t \rightarrow \infty} y(t) = (\bar{u}_{n-k} + \bar{u}_{n-k+1})/2$.

Proof: For the first part of the theorem, we let y^* be the equilibrium point. Therefore, we can have the equality that

$$\sum_{i=1}^n f(y^*, u_i) = k. \quad (21)$$

For large t , $y(t + \tau) = y^* + \Delta y(t + \tau)$ and $y(t) = y^* + \Delta y(t)$. From (19), we can obtain

$$\begin{aligned} \Delta y(t + \tau) &= \Delta y(t) + \frac{\tau}{\epsilon} \left\{ \sum_{i=1}^n f(y(t), u_i) - k \right\} \\ &= \Delta y(t) + \frac{\tau}{\epsilon} \left\{ \sum_{i=1}^n (f(y(t), u_i) - f(y^*, u_i)) \right\}. \end{aligned} \quad (22)$$

Note that $\sum_{i=1}^n f(y, u_i)$ is a continuous function of y . By Mean Value Theorem, there exists $\zeta(t) \in [y(t), y^*]$ such that

$$\sum_{i=1}^n f(y(t), u_i) - f(y^*, u_i) = \sum_{i=1}^n f'(\zeta(t), u_i) \Delta y(t). \quad (23)$$

In (23)

$$f'(\zeta(t), u_i) = \frac{\partial}{\partial y} f(y, u_i) \Big|_{y=\zeta(t)}.$$

Therefore, from (22) and (23) we can obtain

$$\Delta y(t + \tau) = \left\{ 1 + \frac{\tau}{\epsilon} \sum_{i=1}^n f'(\zeta(t), u_i) \right\} \Delta y(t). \quad (24)$$

From (18), we can obtain

$$f'(\zeta(t), u_i) = -\sqrt{\frac{\alpha}{2\pi}} \exp\left(-\frac{\alpha(\zeta(t) - u_i)^2}{2}\right). \quad (25)$$

Therefore

$$-\sqrt{\frac{\alpha}{2\pi}} \leq f'(\zeta(t), u_i) < 0$$

for all $u_i, \zeta(t) \in R$. As for any arbitrary values of α, n , and ϵ , we can define $0 < \tau < \frac{\epsilon}{n} \sqrt{\frac{2\pi}{\alpha}}$ such that

$$0 < 1 + \frac{\tau}{\epsilon} \sum_{i=1}^n f'(\zeta(t), u_i) < 1.$$

Hence from (24), we can readily show that

$$\begin{aligned} & \lim_{m \rightarrow \infty} \Delta y(t + m\tau) \\ &= \lim_{m \rightarrow \infty} \prod_{q=0}^m \left\{ 1 + \frac{\tau}{\epsilon} \sum_{i=1}^n f'(\zeta(t + q\tau), u_i) \right\} \Delta y(t) \\ &= 0. \end{aligned} \quad (26)$$

In other words, $y(t)$ converges to y^* .

If furthermore, α is sufficient large, by Theorem 1, $\lim_{t \rightarrow \infty} y(t) = (\bar{u}_{n-k} + \bar{u}_{n-k+1})/2$. Hence, the proof is completed. ■

C. Energy Function

It should also be noted that (19) can be rewritten as follows:

$$y(t + \tau) = y(t) - \frac{\tau}{\epsilon} \frac{\partial V}{\partial y} \Big|_{y=y(t)} \quad (27)$$

where $V(y)$ is the energy function.

Theorem 3 (Energy Function): The energy function governing the dynamics of the stochastic Wang's k WTA is given by³

$$\begin{aligned} V(y) &= \left(k - \frac{n}{2}\right) y + \frac{1}{2} \sum_{i=1}^n (y - u_i) \operatorname{erf}\left(\sqrt{\frac{\alpha}{2}}(y - u_i)\right) \\ &+ \frac{1}{\sqrt{2\pi\alpha}} \sum_{i=1}^n \exp\left(-\frac{\alpha}{2}(y - u_i)^2\right) + \frac{1}{2} \sum_{i=1}^n u_i. \end{aligned} \quad (28)$$

Proof: From (27), we can obtain the energy by finding the following indefinite integral:

$$V(y) = \int F(y) dy$$

where $F(y)$ is defined in (20). Denote the cumulative distribution function of the standard Gaussian distribution as $\Phi(s)$. Then, $f(y, u_i) = 1 - \Phi(\sqrt{\alpha}(y - u_i))$. As a result, We obtain

$$V(y) = (k - n)y + \sum_{i=1}^n \int \Phi(\sqrt{\alpha}(y - u_i)) dy. \quad (29)$$

$\Phi(s)$ can be expressed in term of the error function $\operatorname{erf}(s/\sqrt{2})$ as follows:

$$\Phi(s) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{s}{\sqrt{2}}\right).$$

³The $\operatorname{erf}(\cdot)$ function in (28) is the error function. It is defined as follows:

$$\operatorname{erf}(s) = \frac{2}{\sqrt{\pi}} \int_0^s \exp(-\eta^2) d\eta.$$

Let $w_i = \sqrt{\frac{\alpha}{2}}(y - u_i)$, (29) can be rewritten as follows:

$$\begin{aligned} V(y) &= \left(k - \frac{n}{2}\right)y + \frac{1}{2} \sum_{i=1}^n \int \operatorname{erf}\left(\sqrt{\frac{\alpha}{2}}(y - u_i)\right) dy \\ &= \left(k - \frac{n}{2}\right)y \\ &\quad + \sqrt{\frac{1}{2\alpha}} \sum_{i=1}^n \left\{ \sqrt{\frac{\alpha}{2}}(y - u_i) \operatorname{erf}\left(\sqrt{\frac{\alpha}{2}}(y - u_i)\right) \right. \\ &\quad \left. + \frac{1}{\sqrt{\pi}} \exp\left(-\frac{\alpha(y - u_i)^2}{2}\right) \right\}. \end{aligned} \quad (30)$$

Therefore

$$\begin{aligned} V(y) &= \left(k - \frac{n}{2}\right)y + \frac{1}{2} \sum_{i=1}^n (y - u_i) \operatorname{erf}\left(\sqrt{\frac{\alpha}{2}}(y - u_i)\right) \\ &\quad + \frac{1}{\sqrt{2\pi\alpha}} \sum_{i=1}^n \exp\left(-\frac{\alpha(y - u_i)^2}{2}\right) + \kappa \end{aligned} \quad (31)$$

where κ is a constant depended on the boundary condition of $V(y)$.

To find the constant, we need to consider the situation that $\alpha \rightarrow \infty$. In such case, the k WTA network reduces to the original Wang's k WTA. Let $V_\infty(y)$ denote the corresponding energy function

$$\begin{aligned} V_\infty(y) &= \left(k - \frac{n}{2}\right)y + \frac{1}{2} \sum_{i=1}^n (y - u_i) \operatorname{sgn}(y - u_i) + \kappa \\ &= \left(k - \frac{n}{2}\right)y + \frac{1}{2} \sum_{i=1}^n |y - u_i| + \kappa. \end{aligned} \quad (32)$$

Clearly

$$V_\infty(0) = \frac{1}{2} \sum_{i=1}^n u_i + \kappa, \quad V_\infty(1) = k - \frac{1}{2} \sum_{i=1}^n u_i + \kappa.$$

To be consistent with boundary conditions obtained in [16] for the original Wang's k WTA, i.e., 1) $V(0) = \sum_{i=1}^n u_i$, and 2) $V(1) = k$, the constant term κ is defined as follows:

$$\kappa = \frac{1}{2} \sum_{i=1}^n u_i.$$

Put it in (31), we thus can obtain (28). The proof is completed. ■

D. Discussion

In the original Wang's model, if the initial state $y(0)$ is equal to 0 (1), the state of the network converges to \bar{u}_{n-k} (\bar{u}_{n-k+1}). In the stochastic Wang's model, it is clear from Theorem 2 that the state of the network converges to $y^* = (\bar{u}_{n-k} + \bar{u}_{n-k+1})/2$.

In addition, (27) and (28) tell us that the dynamic change of the stochastic model can be treated as a gradient descent algorithm that minimizes the energy function $V(y)$ given by (28). Fig. 4 shows the shape of this energy function if the inputs are 0.5, 0.7, 0.8, 0.4, 0.1, and 0.3; $\alpha = 20, 40, 100$, and ∞ , respectively. Here $k = 2$. From the figure, it is clear that $V(y)$ has just one minimum for $\alpha < \infty$. The shape of

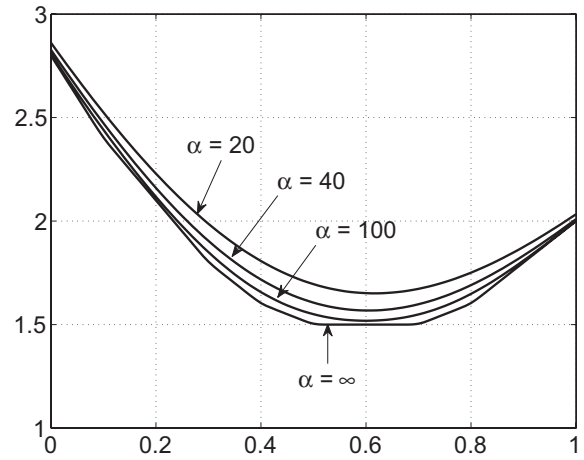


Fig. 4. Energy function $V(y)$ against y for the cases that $\alpha = 20, 40, 100$, and ∞ . Here, the number of winners is 2 (i.e., $k = 2$) and the inputs are 0.5, 0.7, 0.8, 0.4, 0.1, and 0.3, respectively.

$V(y)$ around the minimum is almost flat for $\alpha = 100$. This can also explain why the convergence rate of $y(t)$ when it is close to the convergent is very small. That means, even if the input noise is small (α_I is large) or the stochastic behavior of the output nodes is small (α_O is large), the convergence rate of $y(t)$ is very slow when $y(t)$ is around the equilibrium point. For $\alpha \rightarrow \infty$, i.e., the output nodes are deterministic, the model becomes the original Wang's model. In this case, $V(y)$ is a continuous piecewise linear function as shown in Fig. 4. The value of $V(y)$ is constant for all $y \in (0.5, 0.7)$.

V. CONCLUSION

In addition to the work done by Jun Wang on DNN-based k WTA, we analyzed in this brief the convergence behavior of the Wang's k WTA model when defects exist during the operation. Two defect conditions were considered. The first one was that there are input noise. The second one was that there was stochastic behavior in the output nodes. The existence of equilibrium was proved and stated in Theorem 1. The convergence behavior had analyzed and stated in Theorem 2. The corresponding energy function governing the dynamics of the model was revealed and stated in Theorem 3. Under very mild condition, which was $\alpha \gg 1$, we shown that the state variable $y(t)$ converged to the middle point in between the \bar{u}_{n-k} and \bar{u}_{n-k+1} .

All these results provided an in-depth understanding on the convergence rate of the Wang's k WTA model when its input nodes were corrupted by additive noise and its output nodes were not perfectly implemented as Heaviside functions. From the shape of the energy function and the simulation results, it was found that the convergence rate of $y(t)$ was small when it was close to the equilibrium, which was proportional to $\log(t)$. This convergence rate further reduced when the parameter α increased. Although the results depicted in this brief focused on the Wang's k WTA with input noise and stochastic output node, the analytical results applied to the cases that the Wang's k WTA with only input node noise (i.e., $\alpha = \alpha_I$, $\alpha_I < \infty$, and $\alpha_O \rightarrow \infty$) and the Wang's k WTA with only output node stochastic (i.e., $\alpha = \alpha_O$, $\alpha_O < \infty$, and $\alpha_I \rightarrow \infty$).

REFERENCES

- [1] G. Bolt, "Fault tolerant in multi-layer perceptrons, networks," Ph.D. dissertation, Dept. Comput. Sci., Univ. York, York, U.K., 1992.
- [2] B. A. Calvert and C. Marinov, "Another k -winners-take-all analog neural network," *IEEE Trans. Neural Netw.*, vol. 11, no. 4, pp. 829–838, Apr. 2000.
- [3] H. D. David, *Order Statistics*, 2nd ed. New York, NY, USA; Wiley, 1981.
- [4] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ, USA: Pearson Education, 2009.
- [5] X. Hu and J. Wang, "An improved dual neural network for solving a class of quadratic programming problems with its k -winners-take-all application," *IEEE Trans. Neural Netw.*, vol. 19, no. 12, pp. 2022–2031, Dec. 2008.
- [6] S. Judd and P. W. Munro, "Nets with unreliable hidden nodes learn error-correcting codes," in *Advances in Neural Information Processing Systems*, vol. 5, S. J. Hanson, J. D. Cowan, and C. L. Giles, Eds. San Mateo, CA, USA: Morgan Kaufmann, 1993, pp. 89–96.
- [7] H. K. Khalil, *Nonlinear Systems*. New York, NY, USA: MacMillan, 1992.
- [8] C. S. Leung and J. Sum, "A fault tolerant regularizer for RBF networks," *IEEE Trans. Neural Netw.*, vol. 19, no. 3, pp. 493–507, Mar. 2008.
- [9] Q. Liu, C. Dang, and J. Cao, "A novel recurrent neural network with one neuron and finite-time convergence for k -winners-take-all operation," *IEEE Trans. Neural Netw.*, vol. 21, no. 7, pp. 1140–1148, Jul. 2010.
- [10] E. Majani, R. Erlanson, and Y. Abu-Mostafa, "On the K -winner-take-all network," in *Advances in Neural Information Processing Systems*, D. Touretzky, Ed. San Mateo, CA, USA: Morgan Kaufmann, 1989, pp. 634–642.
- [11] C. A. Marinov and B. A. Calvert, "Performance analysis for a k -winners-take-all analog neural network: Basic theory," *IEEE Trans. Neural Netw.*, vol. 14, no. 4, pp. 766–780, Apr. 2003.
- [12] C. A. Marinov, B. A. Calvert, R. Costea, and V. Bucata, "Time evaluation for analog KWTA processors," in *Proc. Eur. Congr. Comput. Methods Appl. Sci. Eng.*, Jul. 2004, pp. 1–7.
- [13] C. A. Marinov and R. L. Costea, "Time-oriented synthesis for a WTA continuous-time neural network affected by capacitive cross-coupling," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 6, pp. 1358–1370, Jun. 2010.
- [14] C. H. Sequin and R. D. Clay, "Fault tolerance in feedforward artificial neural networks," *Neural Netw.*, vol. 4, pp. 111–141, Feb. 1991.
- [15] J. Sum, C. S. Leung, P. K. S. Tam, G. H. Young, W. K. Kan, and L. W. Chan, "Analysis for a class of winner-take-all model," *IEEE Trans. Neural Netw.*, vol. 10, no. 1, pp. 64–71, Jan. 1999.
- [16] J. Sum, C. S. Leung, and K. Ho, "Analysis on Wang's kWTA with stochastic output nodes," in *Proc. 16th Int. Conf. Neuro-Inf. Process.*, 2011, pp. 268–275.
- [17] J. Wang and Z. Guo, "Parametric sensitivity and scalability of k -winners-take-all networks with a single state variable and infinity-gain activation functions," in *Proc. 7th Int. Symp. Neural Netw.*, 2010, pp. 77–85.
- [18] J. Wang, "Analysis and design of a k -winners-take-all model with a single state variable and the heaviside step activation function," *IEEE Trans. Neural Netw.*, vol. 21, no. 9, pp. 1496–1506, Sep. 2010.
- [19] L. Wang, "Noise injection into inputs in sparsely connected Hopfield and winner-take-all neural networks," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 27, no. 5, pp. 868–870, Oct. 1997.
- [20] Y. Xiao, Y. Liu, C.-S. Leung, J. Sum, and K. Ho, "Analysis on the convergence time of dual neural network-based k -WTA," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 4, pp. 676–682, Apr. 2012.
- [21] M. Hu, H. Li, Q. Wu, G. S. Rose, and Y. Chen, "Memristor crossbar based hardware realization of BSB recall function," in *Proc. Int. Joint Conf. Neural Netw.*, Jun. 2012, pp. 1–7.
- [22] J. He, S. Zhan, D. Chen, and R. L. Geiger, "Analyses of static and dynamic random offset voltages in dynamic comparators," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 5, pp. 911–919, May 2009.

Controllability and Observability of Boolean Control Networks With Time-Variant Delays in States

Lijun Zhang and Kuize Zhang

Abstract—This brief investigates the controllability and observability of Boolean control networks with (not necessarily bounded) time-variant delays in states. After a brief introduction to converting a Boolean control network to an equivalent discrete-time bilinear dynamical system via the semi-tensor product of matrices, the system is split into a finite number of subsystems (constructed forest) with no time delays by using the idea of splitting time that is proposed in this brief. Then, the controllability and observability of the system are investigated by verifying any so-called controllability constructed path and any so-called observability constructed paths in the above forest, respectively, which generalize some recent relevant results. Matrix test criteria for the controllability and observability are given. The corresponding control design algorithms based on the controllability theorems are given. We also show that the computing complexity of our algorithm is much less than that of the existing algorithms.

Index Terms—Boolean control network, controllability, observability, semi-tensor product of matrices, time delay.

I. INTRODUCTION

In recent years, systems biology—the study of the behavior and relationships of all the cells, proteins, DNA, and RNA in a biological system, named cellular networks—has emerged as a notable field of study [1]. The Boolean network, introduced first by [2] and then developed by [3] and [4], and others, has become a powerful tool in describing, analyzing, and simulating the cellular networks and genetic regulatory networks. It was pointed out in [11] that "One of the major goals of systems biology is to develop a control theory for complex biological systems." Hence, it is of practical significance to investigate the control problems of Boolean networks.

Controllability and observability are two basic and challenging problems in the theory of Boolean control networks. The concept of the controllability of Boolean control networks was proposed in [11]. However, no unified test criterion for the controllability and observability was available until 2009, when Cheng and Qi [9] first gave the definitions and corresponding test criteria for the controllability and observability of Boolean control networks

Manuscript received August 27, 2012; accepted February 3, 2013. Date of publication May 13, 2013; date of current version August 16, 2013. This work was supported in part by the National Natural Science Foundation of China under Grant 61174047, Program for New Century Excellent Talents in University of Ministry of Education of China and Basic Research Foundation of Northwestern Polytechnical University (No. JC201230). The work of L. Zhang was supported by the China Scholarship Council. (Corresponding author: K. Zhang.)

L. Zhang is with the College of Automation, Harbin Engineering University, Harbin 150001, China and also with the School of Marine Technology, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: zhanglj@hrbeu.edu.cn).

K. Zhang is with the College of Automation, Harbin Engineering University, Harbin 150001, China and also with the Department of Mathematics, University of Turku, Turku FIN-20014, Finland (e-mail: zkz0017@163.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNNLS.2013.2246187