

Some Remarks on the Stability of Discrete-Time Complex-Valued Multistate Hopfield Neural Networks

Fidelis Zanetti de Castro¹

Federal Institute of Education, Science and Technology of Espírito Santo at Serra, Serra, ES, Brazil

Marcos Eduardo Valle²

Department of Applied Mathematics, University of Campinas, Campinas, SP, Brazil

In this paper, we review three discrete-time complex-valued Hopfield neural networks (CvMHNNs) proposed recently in the literature. Contrary to what has been stated, we provide examples in which the sequences produced by these CvMHNN fails to converge under the usual conditions on the synaptic weight matrix, that is, the synaptic weight matrix is hermitian with non-negative diagonal elements. Furthermore, we present one CvMHNN model that always settle down to a stationary state under the usual conditions on the synaptic weights.

1 Introduction

The Hopfield neural network (HNN) is one of the most important neural networks conceived for the storage and recall of bipolar vectors [5]. According to the neural network terminology, the discrete-time bipolar HNN is a single layer fully-connected recurrent neural network with hard-limiting neurons [3]. Despite the limited representational capability of single-layer recurrent neural network with hard-limiting neurons [2], the HNN has been successfully applied for computer vision modeling, signal reconstruction, image analysis, and optimization [4, 10]. It turns out, however, that most applications of the HNN require that the network yields a convergent sequence for any initial state. For this reason, there is a vast literature on the stability analysis of HNNs [6, 11].

Besides the widely used hard-limiting neurons, complex-valued neurons offer an advantageous representation of information [1, 6, 11]. For instance, a complex-valued neuron can appropriately treat phase and the information contained in phase. In particular, a complex-valued multistate neuron with K states generalizes the bipolar hard-limiting neuron by yielding one of the K -th power roots of one. A Hopfield neural network with such neurons is referred to as a complex-valued multistate Hopfield neural network (CvMHNN) [6].

As far as we know, CvMHNNs have been extensively studied since the early 1990s [1]. One of the most important contribution on CvMHNNs is the paper of Jankowski et al. [6]. Mistakenly, Jankowski et al. stated that a CvMHNN, operating asynchronously, always settle down at an equilibrium state if the synaptic weight matrix satisfies the usual conditions: the matrix is hermitian ($w_{ij} = \bar{w}_{ji}$) with non-negative diagonal ($w_{ii} \geq 0$). In 2014, however, Zhou and Zurada

¹fidelis@ime.unicamp.br

²valle@ime.unicamp.br

showed that the CvMHNN of Jankowski et al. may fail to yield a convergent sequence if $w_{ii} = 0$ [11]. Inspired by the theoretical findings of Zhou and Zurada, in this paper we show that three CvMHNN models introduced by Kobayashi in the last years may also fail to settle down at an equilibrium state [7, 8, 9]. Furthermore, as pointed out by Zhou and Zurada, the condition $w_{ii} = 0$ is often used in applications of the CvMHNN. For instance, some design methods to implement an associative memory using CvMHNN, including the generalized projection rule, requires $w_{ii} = 0$. In the light of this remark, in this paper we propose a slight variation of the CvMHNN which, operating in the asynchronous update mode, always settle down to at a stationary state under the usual conditions on the synaptic weight matrix.

This paper is organized as follows: Next section presents some general concepts relative to CvMHNNs. In Section 3 we review the models proposed by Kobayashi and, contrary to what has been stated, we provide examples in which they fail to settle down to a stationary state. A slight variation of the CvMHNN which yields a convergent sequence for any initial state is proposed in Section 4. We finish the paper with some concluding remarks at Section 5.

2 Complex-Valued Multistate Hopfield Neural Networks

Let $\mathcal{S} = \{s_0, s_1, \dots, s_{K-1}\} \subset \mathbb{C}$ be the set of all possible K states of a complex-valued multistate neuron. Also, let us consider discrete-time CvMHNNs composed by N complex-valued multistate neurons, whose dynamic is given as follows. Given a initial state vector $\mathbf{x}(0) \in \mathcal{S}^N$, we recursively define a sequence $\{\mathbf{x}(t)\}_{t \geq 0}$ using an asynchronous update model. Specifically, the next vector $\mathbf{x}(t + \Delta t)$ is obtained by updating the i th neuron as follows

$$x_i(t + \Delta t) = \begin{cases} f(v_i(t)), & v_i(t) \in \mathcal{D}, \\ x_i(t), & \text{otherwise,} \end{cases} \quad (1)$$

where f is a complex-valued function from $\mathcal{D} \subset \mathbb{C}$ to $\mathcal{S} \subset \mathbb{C}$ and

$$v_i(t) = \sum_{j=1}^N w_{ij}x_j(t), \quad (2)$$

is the activation potential of the i th neuron at iteration t . Here, $w_{ij} \in \mathbb{C}$ denotes the j th synaptic weight of the i th neuron.

Remark 1. Like the original Hopfield neural network [5], we assume that a CvMHNN are updated randomly and in an asynchronous manner, i.e., a single neuron is randomly selected and updated according to (1).

Usually, the *direct method of Lyapunov* is used to ensure the convergence of the sequence $\{\mathbf{x}(t)\}_{t \geq 0}$ produced by a CvMHNN. In mathematical terms, we must find a real-valued energy function $E : \mathcal{S}^N \rightarrow \mathbb{R}$ which is bounded and decreasing along any non-stationary trajectory, i.e., the inequality

$$\Delta E(\mathbf{x}(t)) = E(\mathbf{x}(t + \Delta t)) - E(\mathbf{x}(t)) < 0, \quad (3)$$

must holds true whenever $\mathbf{x}(t + \Delta t) \neq \mathbf{x}(t)$. In this case, we know that the time evolution of a CvMHNN yields a minimum value of E . Equivalently, the network comes to rest at an equilibrium state [5]. The energy function given by

$$E(\mathbf{x}) = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \bar{x}_i w_{ij} x_j, \tag{4}$$

where \bar{x} denotes the complex conjugate of x , is predominantly used for the analysis of the stability of a CvMHNN [6, 7, 8, 9].

3 The CvMHNN Models of Kobayashi

Consider an integer number $K > 1$, referred to as the *resolution factor*, and define the angle $\Delta\theta = \pi/K$, called *phase-quanta*. In [9], Kobayashi introduces a complex-valued multistate activation function, which we denote here by csgn_1 . This function has domain \mathbb{C}^* and co-domain the set

$$\mathcal{S}_1 = \{1, e^{2i\Delta\theta}, e^{4i\Delta\theta}, \dots, e^{2(K-1)i\Delta\theta}\}. \tag{5}$$

Formally, csgn_1 is defined by

$$\text{csgn}_1(z) = \begin{cases} 1, & 0 \leq \arg(z) < \Delta\theta, \\ e^{2i\Delta\theta}, & \Delta\theta \leq \arg(z) < 3\Delta\theta, \\ \vdots & \vdots \\ e^{2(K-1)i\Delta\theta}, & (2K-3)\Delta\theta \leq \arg(z) < (2K-1)\Delta\theta \\ 1, & (2K-1)\Delta\theta \leq \arg(z) < 2\pi. \end{cases} \tag{6}$$

The evolution equation of the CvMHNN model proposed in [9] is based on (1) using $f = \text{csgn}_1$. We would like to point out, however, that Kobayashi does not address the case $v_i(t) = 0$.

Possibly unaware of the stability issues of CvMHNN, Kobayashi states – without presenting a formal proof – that the energy function given by (4) always decreases assuming an asynchronous update mode and the usual conditions on the synaptic weights. The following example, however, shows that this statement is not true.

Example 1. Let $K = 2$. Note that, in this case, the network admits only the following four states:

$$\mathbf{x}_1 = [1, 1]^T, \quad \mathbf{x}_2 = [-1, 1]^T, \quad \mathbf{x}_3 = [1, -1]^T, \quad \text{and} \quad \mathbf{x}_4 = [-1, -1]^T. \tag{7}$$

Consider the synaptic weight matrix given by

$$W = \begin{bmatrix} 0 & \mathbf{i} \\ -\mathbf{i} & 0 \end{bmatrix} \in \mathbb{C}^{2 \times 2}. \tag{8}$$

Starting at any of the four states, the network yields a random walk at the diagram depicted in Figure 1a). For instance, the probability of changing from the state \mathbf{x}_1 to \mathbf{x}_2 is $1/2$, which corresponds to the probability of selecting the first neuron to be updated. In fact, if $\mathbf{x}(t) = \mathbf{x}_1$ and the first neuron is chosen to be updated, we obtain $v_1(t) = \mathbf{i}$ and $x_1(t + \Delta t) = -1$. Since there always exists a fifty percent probability to change states, the sequence given by (1) with $f \equiv \text{csgn}_1$ does not converge.

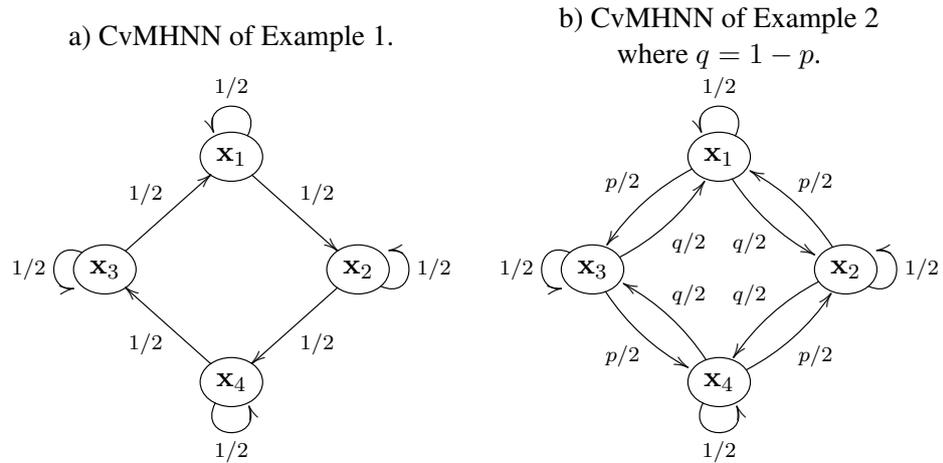


Figure 1: Diagram representing the Markov process of CvMHNN models.

In [7], Kobayashi introduces a second activation function that corresponds to a slight modification of csgn_1 . We denote this second activation function by csgn_2 . Like the function csgn_1 , the domain of csgn_2 is \mathbb{C}^* , but the co-domain is the set $\mathcal{S}_2 = \{e^{i\Delta\theta}, e^{3i\Delta\theta}, \dots, e^{(2K-1)i\Delta\theta}\}$. Formally, we have

$$\text{csgn}_2(z) = \begin{cases} e^{i\Delta\theta}, & 0 \leq \arg(z) < 2\Delta\theta, \\ e^{3i\Delta\theta}, & 2\Delta\theta \leq \arg(z) < 4\Delta\theta, \\ \vdots & \vdots \\ e^{(2K-1)i\Delta\theta}, & 2(K-1)\Delta\theta \leq \arg(z) < 2\pi. \end{cases} \quad (9)$$

Note that the activation function csgn_2 can be obtained from csgn_1 using rotations by half of the phase-quanta. Precisely, we have

$$\text{csgn}_2(z) = \phi^{-1}\left(\text{csgn}_1(\phi(z))\right), \quad \forall z \in \mathbb{C}^*, \quad (10)$$

where $\phi : \mathbb{C} \rightarrow \mathbb{C}$ is the clockwise rotation map $\phi(z) = e^{-i\Delta\theta/2}z$ and $\phi^{-1}(z) = e^{i\Delta\theta/2}z$ is its inverse. Follows from (10) that the CvMHNN models given by (1) using $f \equiv \text{csgn}_1$ and $f \equiv \text{csgn}_2$ are equivalent. In particular, contrary to what have been stated in [7], we cannot ensure that the CvMHNN with $f \equiv \text{csgn}_2$ will always settle down at an equilibrium state.

Finally, a third activation function, which we shall denote by csgn_3 , is introduced by Kobayashi in [8]. In mathematical terms, the activation function $\text{csgn}_3 : \mathbb{C}^* \rightarrow \mathcal{S}_1$ is defined as follows using

a certain probability p in a Bernoulli trial:

$$\text{csgn}_3(z) = \begin{cases} 1, & 0 \leq \arg(z) < \Delta\theta, \\ \left\{ \begin{array}{l} 1, \quad \text{with probability } p \\ e^{2i\Delta\theta}, \quad \text{with probability } 1-p \end{array} \right\}, & \arg(z) = \Delta\theta \\ e^{2i\Delta\theta}, & \Delta\theta < \arg(z) < 3\Delta\theta, \\ \vdots & \vdots \\ e^{2(K-1)i\Delta\theta}, & (2K-3)\Delta\theta < \arg(z) < (2K-1)\Delta\theta \\ \left\{ \begin{array}{l} e^{2(K-1)i\Delta\theta}, \quad \text{with probability } p \\ 1, \quad \text{with probability } 1-p \end{array} \right\}, & \arg(z) = (2K-1)\Delta\theta \\ 1, & (2K-1)\Delta\theta < \arg(z) < 2\pi. \end{cases} \quad (11)$$

Note that csgn_3 differs from csgn_1 if $\arg(z) = (2k-1)\Delta\theta$ for some $k = 1, \dots, K$. Precisely, a random draw is carried out if $\arg(z) = (2k-1)\Delta\theta$, for $k \in \{1, \dots, K\}$, and the function csgn_3 rotates either clockwise or counterclockwise to yield an element of \mathcal{S}_1 given by (5). Moreover, in the deterministic case in which $p = 0$, we have $\text{csgn}_3 \equiv \text{csgn}_1$. Despite been a stochastic process when $0 < p < 1$, a CvMHNN defined by (1) with $f \equiv \text{csgn}_3$ may fail to settle at an equilibrium state. The following example illustrates this remark.

Example 2. Consider the synaptic weight matrix $W \in \mathbb{C}^{2 \times 2}$ given by (8) and the resolution factor $K = 2$. In this case, the network admits only the following four states: $\mathbf{x}_1 = [1, 1]^T$, $\mathbf{x}_2 = [-1, 1]^T$, $\mathbf{x}_3 = [1, -1]^T$, and $\mathbf{x}_4 = [-1, -1]^T$. Starting at any of these four states, the network yields a random walk at the diagram depicted in Figure 1b). For instance, the probability of changing from the state \mathbf{x}_1 to \mathbf{x}_3 is $p/2$, which corresponds to the probability to obtain $x_2(t + \Delta t) = -1$. Now, the probability to select the second neuron is $1/2$ and the probability to obtain $\text{csgn}(-i) = 1$ is p . Thus, we have $\Pr[x_2(t + \Delta t) = -1] = p/2$. Since there always is a fifty percent probability to change states, the sequence given by (1) with $f \equiv \text{csgn}_3$ does not converge.

4 Modified CvMHNN and its stability analysis

Let us now introduce an activation function, denoted by csgn_r , for which a CvMHNN always settles down at an stationary state in an asynchronous update mode under the usual conditions on the synaptic weights. Basically, the activation function csgn_r is a restriction of both csgn_1 and csgn_3 to the set

$$R = \{z \in \mathbb{C}^* : \arg(z) \neq (2k-1)\Delta\theta, \forall k = 1, \dots, K\}. \quad (12)$$

Formally, the activation function $\text{csgn}_r : R \rightarrow \mathcal{S}_1$ is defined by

$$\text{csgn}_r(z) = \begin{cases} 1, & 0 \leq \arg(z) < \Delta\theta, \\ e^{2i\Delta\theta}, & \Delta\theta < \arg(z) < 3\Delta\theta, \\ \vdots & \vdots \\ e^{2(K-1)i\Delta\theta}, & (2K-3)\Delta\theta < \arg(z) < (2K-1)\Delta\theta \\ 1, & (2K-1)\Delta\theta < \arg(z) < 2\pi. \end{cases} \quad (13)$$

Note that csgn_r is not defined for a complex number z such that $\arg(z) = (2k - 1)\Delta\theta$, $k \in \{1, 2, \dots, K\}$. Therefore, the neuron of a CvMHNN given by (1) with $f \equiv \text{csgn}_r$ is updated at time t if and only if $\arg(v_i(t)) \neq (2k - 1)\Delta\theta$ for some $k \in \{1, 2, \dots, K\}$. The following theorem address the stability of the new CvMHNN model.

Theorem 1. The sequence defined by (1) with $f(\cdot) = \text{csgn}_r(\cdot)$ is convergent for any initial state \mathbf{x} in an asynchronous update mode if the weights satisfy $w_{ji} = \bar{w}_{ij}$ and $w_{ii} \geq 0$.

Proof. Clearly, the energy function defined by (4) is real-valued and bounded from below. Let us show that $\Delta E = E(\mathbf{x}') - E(\mathbf{x}) < 0$ if $\mathbf{x} \equiv \mathbf{x}(t)$ and $\mathbf{x}' \equiv \mathbf{x}(t + \Delta t)$ for some $t \geq 0$. Since we have assumed an asynchronous update mode, only the μ th neuron changes its state at iteration t . In other words, we have $x'_\mu \neq x_\mu$ but $x'_j = x_j$ for all $j \neq \mu$. Moreover, we know that $w_{\mu\mu}$ is a real number and $|x_\mu| = |x'_\mu| = 1$. Thus, we have $\bar{x}_\mu w_{\mu\mu} x_\mu = \bar{x}'_\mu w_{\mu\mu} x'_\mu = w_{\mu\mu}$ and the variation of the energy function is

$$\Delta E = E(\mathbf{x}') - E(\mathbf{x}) = -\frac{1}{2} \left[\sum_{j \neq \mu} (\bar{x}'_\mu - \bar{x}_\mu) w_{\mu j} x_j + \sum_{i \neq \mu} \bar{x}_i w_{i\mu} (x'_\mu - x_\mu) \right].$$

Recall that the conjugate of $\sum_{i \neq \mu} \bar{x}_i w_{i\mu} (x'_\mu - x_\mu)$ is $\sum_{j \neq \mu} (\bar{x}'_\mu - \bar{x}_\mu) w_{\mu j} x_j$ because $w_{ij} = \bar{w}_{ji}$. Also, the activation potential of the μ th neuron at iteration t is $v_\mu = \sum_{j=1}^n w_{\mu j} x_j$. Hence, we have

$$\Delta E = -\text{Re} \left\{ (\bar{x}'_\mu - \bar{x}_\mu) \sum_{j \neq \mu} w_{\mu j} x_j \right\} = -\text{Re} \left\{ (\bar{x}'_\mu - \bar{x}_\mu) (v_\mu - w_{\mu\mu} x_\mu) \right\} = -T_1 - T_2,$$

where $T_1 = \text{Re} \{ \bar{x}'_\mu v_\mu \} - \text{Re} \{ \bar{x}_\mu v_\mu \}$ and $T_2 = w_{\mu\mu} \text{Re} \{ 1 - x_\mu \bar{x}'_\mu \}$. From the Cauchy-Schwarz inequality, we derive $\text{Re} \{ x_\mu \bar{x}'_\mu \} = x_{\mu 0} x'_{\mu 0} + x_{\mu 1} x'_{\mu 1} < |x_\mu| |x'_\mu| = 1$ because x'_μ and x_μ are not parallel vectors in \mathbb{R}^2 . As a consequence, the inequality $\text{Re} \{ 1 - \bar{x}_\mu x'_\mu \} > 0$ holds true and, since $w_{\mu\mu} \geq 0$, we conclude that $T_1 = w_{\mu\mu} \text{Re} \{ 1 - x_\mu \bar{x}'_\mu \} \geq 0$. Let us now show that $T_2 \geq 0$. We know that $v_\mu = |v_\mu| e^{i\alpha} \in \mathbb{R}$ and $x_\mu = e^{i\theta} \neq x'_\mu = \text{csgn}_r(v_\mu) = e^{i\theta'}$. Thus, we have

$$\text{Re} \{ \bar{x}'_\mu v_\mu \} = |v_\mu| \text{Re} \left\{ e^{-i\theta'} e^{i\alpha} \right\} = |v_\mu| \text{Re} \left\{ e^{i(\alpha - \theta')} \right\} = |v_\mu| \cos(\alpha - \theta') = |v_\mu| \cos(|\alpha - \theta'|).$$

Similarly, we derive $\text{Re} \{ \bar{x}_\mu v_\mu \} = |v_\mu| \cos(|\alpha - \theta|)$. From the definition of csgn_r , however, we have $|\alpha - \theta'| < \Delta\theta$ while $|\alpha - \theta| > \Delta\theta$. Then, $|\alpha - \theta'| < |\alpha - \theta|$ and $\cos(|\alpha - \theta'|) > \cos(|\alpha - \theta|)$. Consequently, $|v_\mu| \cos(|\alpha - \theta'|) > |v_\mu| \cos(|\alpha - \theta|)$ which implies the inequality $T_1 = \text{Re} \{ \bar{x}'_\mu v_\mu \} - \text{Re} \{ \bar{x}_\mu v_\mu \} > 0$. Concluding, since both $T_1 > 0$ and $T_2 \geq 0$, we have $\Delta E = -T_1 - T_2 < 0$ whenever $x_\mu(t + \Delta t) \neq x_\mu(t)$ for some index $\mu \in \{1, 2, \dots, N\}$. \square

5 Conclusions and Remarks

In this paper, we reviewed Hopfield neural network models whose neuron states reside in a finite set of unit complex numbers. Specifically, we focused our attention on the complex-valued multi-state Hopfield neural networks (CvMHNNs) introduced recently by Kobayashi [7, 8, 9]. Precisely,

we pointed out that these CvMHNN models do not always settle down at an equilibrium state. In addition, we proposed a new activation function csgn_r , which is obtained by restricting the domain of the activation functions of Kobayashi to the set R defined by (12). We proved that a CvMHNN with the function csgn_r , always settle down to an equilibrium state under the usual conditions for any initial state.

References

- [1] N. N. Aizenberg and I. N. Aizenberg. CNN based on multivalued neuron as a model of associative memory for gray-scale images. In *Proceedings of the 2nd International Workshop on Cellular Neural Networks and Their Applications*, pages 36–42, 1992.
- [2] M. W. Goudreau, C. L. Giles, S. T. Chakradhar, and D. Chen. First-order versus second-order single-layer recurrent neural networks. *IEEE Transactions on Neural Networks*, 5(3):511–513, May 1994.
- [3] S. Haykin. *Neural Networks and Learning Machines*. Prentice-Hall, Upper Saddle River, NJ, 3rd edition edition, 2009.
- [4] J. Hopfield and D. Tank. Neural computation of decisions in optimization problems. *Biological Cybernetics*, 52:141–152, 1985.
- [5] J. J. Hopfield. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences*, 79:2554–2558, Apr. 1982.
- [6] S. Jankowski, A. Lozowski, and J. Zurada. Complex-Valued Multi-State Neural Associative Memory. *IEEE Transactions on Neural Networks*, 7:1491–1496, 1996.
- [7] M. Kobayashi. Fast recall for complex-valued hopfield neural networks with projection rules. 2017:1–6, 05 2017.
- [8] M. Kobayashi. Symmetric complex-valued hopfield neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 28(4):1011–1015, April 2017.
- [9] M. Kobayashi. Multistate vector product hopfield neural networks. *Neurocomputing*, 272:425 – 431, 2018.
- [10] Y. Sun. Hopfield neural network based algorithms for image restoration and reconstruction. II. Performance analysis. *IEEE Transactions on Signal Processing*, 48(7):2119–2131, Jul 2000.
- [11] W. Zhou and J. M. Zurada. Stability condition for discrete time multi-valued recurrent neural networks in asynchronous update mode. In *2014 International Joint Conference on Neural Networks (IJCNN)*, pages 3402–3405, July 2014.