Complex Valued Amari_Hopfield Neural Network: Convergence Theorem

Rama Garimella and Shaik Salma

December 17, 2023

Complex Valued Amari – Hopfield Neural Network: Convergence Theorem

Garimella Ramamurthy
Computer Science and Engineering
Mahindra University
Bahadurpally, Hyderabad,
Telangana - 500043
Email: rama.murthy@mahindrauniversity.edu.in

Shaik Salma
Computer Science and Engineering
Mahindra University
Bahadurpally, Hyderabad,
Telangana - 500043
Email: salma20pcse004@mahindrauniversity.edu.in

***Abstract*—In this research paper, a simplified expression for the energy function of a Complex Hopfield neural network is derived. Based on that expression, a simplified proof of the convergence theorem is presented. Several interesting results related to the Convergence theorem are proved.**

1. Introduction

Research on artificial neural networks (ANN) was initiated and progressed to emulate the biological neural network. Mcculloch Pitts proposed a model of artificial neuron to emulate the classification function based on the linear separability of patterns. Since the synaptic weights are fixed, it is realized that such a model of neuron has no training ability. Rosenblatt introduced the perceptron model of neuron with training ability. It was shown that based on a learning law, the perceptron weights coverage to a hyperplane which separates 2 classes provided they are linearly separable. As a natural generalization, a single-layer perceptron was innovated to classify patterns belonging to multiple classes that are linearly separable. Werbos pioneered the multi-layer perceptron (MLP) which can classify even non-linearly separable patterns. Also, researchers proposed ANN's in which inputs, synaptic weights, and thresholds are complex numbers leading to the research area of complex-valued neural networks.

In an effort to emulate biological memory, Hopfield proposed the Hopfield neural network [1]. Amari proposed an ANN which is similar to the Hopfield neural network. Several researchers proposed complex-valued Hopfield neural networks [2].

The complex-valued Hopfield neural network proposed is based on an interesting complex signum function (different from that utilized in [2]) and the associated complex unit hypercube. The associated convergence theorem (in the spirit of real-valued Hopfield Associative Memory. In the context of a real-valued Hopfield neural network, Bruck etal [ ] explored the connection between stable states and the cuts in the graph associated with the Hopfield neural network(HNN).

In this research paper, the authors are motivated to establish the relationship between the stable states of the specific complex valued HNN and the cuts in the graph associated with it. More generally, graph theoretic codes associated with complex-valued HAM are proposed for investigation.

This research paper is organized as follows. In section 2, the relevant research literature is reviewed. In section 3, the quadratic energy function associated with a complex Hopfield neural network (CHNN) is evaluated and interpreted from the point of view of the graph associated with such ANN.

2. Review of Research Literature:

Hopfield as well as Amari independently proposed an Artificial neural network (ANN) which acts as an associative memory [1]. Goles and Foglemann proved a convergence theorem which confirms that such an ANN acts as a content addressable associative memory [4]. The proof is based on associating a quadratic energy function with the dynamics.

Zurade etal proposed a Complex Valued Hopfield Neural Network and applied it to Image Denoising [2]. There are other research efforts related to complex-valued associative memories.

3. Complex Hopfield Neural Network: Cuts in Graphs:

3.1. Basics of Novel Complex Hopfield Neural Network:

In [3], the authors proposed a novel complex Hopfiled neural network and proved the convergence theorem under certain conditions. We briefly summarize the dynamics of such an artificial neural network.

Consider an ANN with 'M' neurons. Let the state of each neuron assume values in the set

$$L = \{1 + j1, 1–j1, −1 + j1, −1–j1\}$$

Thus, the state space of the ANN is the unit Complex Hypercube, H i.e;

$$H = \{(x_1, x_2, ..., x_M) : x_i \in L \quad for \quad 1 \leq i \leq M\}$$

(i.e. x_i is the state of i^{th} neuron). The ANN constitutes a nonlinear dynamical system that operates in the following modes of operation. Let $\bar{W}$ be the Hermitian synaptic weight matrix and Let $\bar{T}$ be the Threshold vector with complex-valued elements. The synaptic weight matrix provides weights of the links connecting 'M' artificial neurons, to each other.

The activation function employed by each neuron is the complex Signum function: CSIGN( ), where

CSIGN (a+jb) = Sign(a)+j Sign(b),

Where Sign( ) is the signum function.

Serial Mode: At each time; the state of only one neuron is updated. The state updation in serial mode is done in the following manner.
Let v_i(n) be the state of i^th neuron at time n.

$$v_i(n + 1) = CSign\{\sum_{j=1}^{M} w_{ij}v_j(n) − t_i\}$$

Fully Parallel Mode: The state vector of the ANN is $\bar{V}(n)$ with components lying in the set L. The state updation in this mode is performed in the following manner.

$$\bar{V}(n + 1) = CSign[\bar{W}\bar{V}(n) − \bar{T}]$$

Other Parallel Modes: The state Updation is performed at more than one node (but strictly less than M nodes).

Thus, the Complex Hopfield Neural Network is a homogeneous (no external input), non-linear dynamical system operating in the above nodes of operation starting with an initial condition lying on the complex hypercube.

In the state space of the Complex Hopfield Associative Memory (i.e. Complex Hypercube) there are distinguished states called as STABLE STATES. They are defined in the following manner.
.

Stable State: A state vector, $\bar{U}$ lying on the complex hypercube is called a stable state if and only if

$$\bar{U} = CSign(\bar{W}\bar{U} − \bar{T})$$

Similarly

Anti-Stable State: A state vector, $\bar{Z}$ lying on the complex hypercube is called an anti-stable state if and only if

$$\bar{Z} = −CSign(\bar{W}\bar{Z} − \bar{T})$$

Convergence Theorem: The following theorem summarizes the functioning of the Complex Hopfield neural network as an associative memory.

Theorem: Consider a Complex Hopfield neural network. The following dynamics hold true when the diagonal elements of 'W' are non-negative.

1) In the serial mode, the ANN always converges to a stable state starting in any initial state.
2) In the fully parallel mode, the ANN converges to a stable state or a cycle of length at most 2 is reached.

We now explore the relationship between the energy function associated with the Hermitian synaptic weight matrix

$$\bar{W} = \bar{W_R} + j\bar{W_I}$$

and the energy function based on the real-valued matrices

$$\{\bar{W_R}, \bar{W_I}.\}$$

Lemma: Let $\bar{W}$ be the synaptic weight matrix of the novel Complex Hopfield neural network. Let

$$\bar{W} = \bar{W_R} + j\ \bar{W_I}$$

Where $\bar{W_R}$ is real valued symmetric matrix and $\bar{W_I}$ is the real valued anti-symmetric matrix.

The Hermitian form associated with $\bar{W}$ is equivalent to the sum of two quadratic forms associated with $\bar{W_R}$ i.e.

$$\bar{X^*}\bar{W}\bar{X} = \bar{X_R^T}\bar{W_R}\bar{X_R} + \bar{X_I^T}\bar{W_R}\bar{X_I}$$

Proof: Let $\bar{X^*}$ be the conjugate transpose of complex valued elements of vector $\bar{X}$

$$\bar{X^*}\bar{W}\bar{X} = \bar{X_R^T}\bar{W_R}\bar{X_R} + \bar{X_I^T}\bar{W_R}\bar{X_I}$$

$$\bar{X^*}\bar{W}\bar{X} = (X_R^T − jX_I^T)(\bar{W_R} + j\bar{W_I})(\bar{X_R} + jX_I)$$
$$= (X_R^T − jX_I^T)(\bar{W_R}\bar{X_R} + j\bar{W_R}X_I + j\bar{W_I}\bar{X_R} − W_IX_I$$

$$= X_R^T\bar{W_R}\bar{X_R} + \cancel{jX_R^T\bar{W_R}X_I} + jX_R^TW_IX_R − \cancel{X_R^T\bar{W_I}X_I}$$
$$\cancel{−jX_I^T\bar{W_R}\bar{X_R}} + X_I^TW_RX_I + \cancel{X_I^T\bar{W_I}\bar{X_R}} + jX_I^TW_IX_I$$

$$\left.\begin{array}{l} W_R^T{=}W_R \text{ and } W_I^T = -\ W_I \end{array}\right\} \rightarrow X_R^TW_IX_R = X_I^TW_IX_I = 0$$

i.e. Anti-Symmetric

$$\bar{X^*}\bar{W}\bar{X} = X_R^T(W_R + jW_I)X_R + X_I^T(W_R + jW_I)X_I$$

Where W_I = W_I^T = 0

$$= X_R^T\bar{W_R}\bar{X_R} + X_I^T\bar{W_R}X_I$$

Using the Lemma ( ), we provide a simple proof of the convergence theorem below

Simple Proof Of Convergence Theorem:

From the lemma (1), the energy function (i.e. quadratic form associated with Hermitian Synaptic weight matrix) reduces to the sum of two quadratic forms associated with $\bar{W}_R$. Thus, we directly invoke the proof of the Convergence theorem associated with the real-valued Hopfield neural network [4] and arrive at the desired conclusion.

The above lemma is significant in the reuse that the energy function(quadratic form) associated with the dynamics of the Complex-Hopfield neural network reduces to that associated with a related real Hopfield neural network. Thus, the following lemma is associated with the real Hopfield neural network is interesting as it relates the global optimum stable state of a Hopfield neural network.

Lemma ( ): Let N = ($\bar{H}$, T) be a real Hopfield neural network with $\bar{H}$ being the symmetric (real-valued) synaptic weight matrix and the threshold vector, $\bar{T}$ being zero vector i.e; $\bar{T} \equiv \bar{0}$. The problem of finding state $\bar{V}$ for which the energy E (E = $\bar{V}^T H \bar{V}$) is maximum is equivalent to finding a minimum cut in the weighted graph corresponding to 'N.'

Proof: Refer Bruck and Blaum research paper.

In view of the previous 2 lemmas, the minimum cut in the graph associated with the Complex Hopfield network is related to the global optimum stable.

To prove the main result of the preservation of stable states, we need the following definition.
Definition: ϵ - perturbation of a matrix, $\bar{A}$ is defined as

$$\hat{A} = \bar{A} + \epsilon I \qquad with \quad \epsilon > 0$$

i.e; the diagonal elements of $\hat{A}$ are obtained by adding 'ϵ' to the corresponding diagonal elements of $\bar{A}$.

$$Let\ \hat{W} = W + \epsilon I$$

From the above definition, it is clear that, if some of the diagonal elements of $\bar{W}$ are negative, by adding a suitable value of 'ϵ', it can be ensured that all the diagonal elements of $\hat{W}$ are positive/non-negative. Specifically

$$\delta = \Big\{ Max_i |W_{ii}| : W_{ii} < 0 \Big\}$$

Let $\epsilon \geq \delta$

With such a choice of 'ϵ', we now prove the following lemma, where we assume that the threshold vector, $\bar{T} \equiv \bar{0}$ (zero vector).

Lemma: Consider 'ϵ' (with ϵ>0) perturbation of synaptic weight matrix, $\bar{W}$ i.e. Let

$$\hat{W} = W + \epsilon I$$

The stable states of Hopfield Associative Memory(HAM) based on $\bar{W}$ are a subset of the HAM based on $\hat{W}$.

Proof: Suppose $\bar{h}$ is a stable state of $\bar{W}$

$$Sign(\bar{W}\bar{h}) = \bar{h} \qquad Hence$$

$$Sign(\bar{W}\bar{h}) = Sign(\bar{W}\bar{h} + \epsilon\bar{h})$$

Let the i^{th} component of $\bar{W}\bar{h}$ = a+jb.
Since $\epsilon > 0$, we have that

$$Sign\Big\{(a+jb)\epsilon\Big\} = Sign(\epsilon) + jsign(b)$$

Thus, we have that

$$Sign(\hat{W}\bar{h}) = \bar{h}$$

Hence, the stable states of $\hat{W}$ are a subset of those of $\bar{W}$. Thus 'ϵ'- perturbation preserves the stable states of $\bar{W}$, in $\hat{W}$.

Note: It can be reasoned that under a more general perturbation model, the stable states are preserved. We now consider one such generalized perturbation model.

3.1.1. Generalized Perturbation Model:.

Let $\bar{W}$ be additively perturbed by a symmetric matrix, $\bar{R}$ i.e

$$\hat{W} = \bar{W} + \bar{R}$$

. We consider the case, where the perturbation matrix, $\bar{R}$ has the same set of eigenvectors as those of $\bar{W}$. But the eigenvalues of $\bar{W}$ can be different from those of $\bar{R}$. As assumed earlier, the threshold vector, $\bar{T} \equiv \bar{0}$ (i.e. Zero vector). We have the following

Lemma:

$$Let\ \Big\{\mu_i\Big\}_{i=1}^{N}, \Big\{\theta_i\Big\}_{i=1}^{N}$$

be the eigenvalues of $\bar{W}$, $\bar{R}$ respectively. Also, let

$$Sign(\mu_i + \theta_i) = Sign(\mu_i) \quad for \quad 1 \leq i \leq N$$

Under such assumptions, the stable states of Complex Hopfield Neural Network based on $\bar{W}$ are the same as those of Hopfield Neural Network(HNN) based on $\hat{W}$.

Proof:

Follows from a similar argument as those of ϵ-perturbation (since eigenvectors of $\bar{W}$, $\bar{R}$ are the same). It is avoided for brevity.

4. Conclusion

In this research paper, a simplified expression for the energy function of a Complex Hopfield neural network is presented. Based on that result a simplified proof of convergence theorem is provided. Several results on the Convergence theorem are provided.

References

[1] J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," Proceedings of the National Academy of Sciences, USA: Vol.79. 1989, PP.2554-2558, DOI:10.1073/pnas.79.8.2554.

[2] Mehmet Kerem Muezzinoglu; Cüneyt Güzeliş; Jacek M. Zurada, "A new design method for the complex-valued multistate Hopfield associative memory," IEEE Transactions on Neural Networks, Volume: 14, Issue: 4, July 2003, PP. 891-899, DOI: 10.1109/TNN.2003.813844.

[3] G. Rama Murthy and D. Praveen, "Complex-valued neural associative memory on the complex hypercube," Proceedings of 2004. IEEE Conference on Cybernetics and Intelligent Systems (CIS 2004), DOI: 10.1109/ICCIS.2004.1460492.

[4] Eric Goles-Chacc, Françoise Fogelman-Soulie, Didier Pellegrin, "Decreasing energy functions as a tool for studying threshold networks," Discrete Applied Mathematics, Vol. 12, PP. 261-277, 1985, doi.org/10.1016/0166-218X(85)90029-0.