Preservation of Dynamics of Discrete Time Hopfield Neural Network: Perturbation/Quantization Analysis

Rama Murthy Garimella

January 6, 2025

PRESERVATION OF DYNAMICS OF DISCRETE-TIME HOPFIELD NEURAL NETWORK: PERTURBATION / QUANTIZATION ANALYSIS

Garimella Rama Murthy, IEEE Senior Member

Abstract— **In this research paper, $\varepsilon - perturbation$ of diagonal elements of symmetric synaptic weight matrix, $\overline{W}$ (with $\varepsilon > 0$) of Hopfield Associative Memory (HAM) (resulting in updated synaptic weight matrix $\widehat{W} = \overline{W} + \varepsilon I$) is assumed to ensure that the sufficient condition of convergence theorem is satisfied. It is proved that under such perturbation, stable states of HAM based on synaptic weight matrix $\overline{W}$ are a subset of those of HAM based on $\widehat{W}$. This result is generalized to prove that if $\widehat{W} = \overline{W} + \overline{R}$, (where $\overline{W}, \overline{R}$ have the same eigenvectors), the stable states of HAM based on $\overline{W}$ are preserved as some of those of $\widehat{W}$. It is proved that a linear system of equations with the coefficient matrix being doubly stochastic naturally arises in expressing the vector of diagonal elements of $\overline{W}$ in terms of the eigenvalues of the symmetric matrix $\overline{W}$. It is proved that (in a well defined sense), if $\overline{W}$ is positive definite, from the view point of dynamics of HAM, the threshold vector can be assumed to be a zero vector. These results are interesting from the viewpoint of preservation of 'interesting" dynamics of HAM under practical perturbation models. Based on known literature, such perturbations ensure that minimum cut in the graph associated with $\overline{W}$ is same as that associated with $\widehat{W}$. Also, preservation of interesting dynamics (e.g. stable states) under quantization of synaptic weights is also explored.**

Index Terms— Hopfield Associative Memory (HAM), Stable States, Anti-Stable States, ϵ −perturbation, Convergence Theorem

I. INTRODUCTION

In an effort to model the biological neural network, McCulloch and Pitts proposed a model of artificial neural network. Network of such neurons, called an Artificial Neural Network (ANN) was proposed to emulate the "classification" function. Artificial Neural Networks such as the Single Layer Perceptron (SLP) and Multi-Layer Perceptron (MLP) were successfully utilized in many applications. Based on such successful ANNs, Hopfield proposed a model which emulates a biological memory. Specifically, Hopfield Neural Network (HNN) based on the McCulloch-Pitts neuron acts as an associative memory.

This work was supported in part by a grant from the Ecole Centrale School of Engineering, Mahindra University, India.The author is a Professor of Computer Science at Mahindra University, Hyderabad,INDIA (email: rama.murthy@mahindrauniversity.edu.in)

Hopfield Associative Memory (HAM), a homogeneous nonlinear dynamical system was shown to exhibit convergence behavior or periodic/cyclic behavior. The convergence theorem related to HAM/HNN utilizes a quadratic energy function associated with its dynamics. For the convergence theorem to hold, a sufficient condition was imposed on the diagonal elements of the symmetric synaptic weight matrix. Researchers routinely assumed that all diagonal elements of the synaptic weight matrix are all nonnegative for the HNN to act as an associative memory. The author contemplated over approaches to ensure that such a sufficient condition holds true. In such an effort, we are naturally led to the perturbation of the symmetric synaptic weight matrix, $\overline{W}$ by a perturbation matrix resulting in a symmetric synaptic weight matrix, $\widehat{W}$ which satisfies the sufficient condition that all diagonal elements are nonnegative. There are several ways to choose the perturbation matrix. This research paper is an effort to relate the dynamics of HAM based on $\overline{W}$ with that of the symmetric matrix, $\widehat{W}$. Specifically, it is reasoned that the "interesting/desired" dynamics of HAM is preserved under "suitable" perturbation. The notions of "suitable perturbation", "interesting dynamics" are formalized and proof of preservation of "interesting dynamics" is demonstrated.

In summary, the main contributions of this research paper are

(i)Preservation of stable states of HAM under "interesting" perturbation of elements of the synaptic weight matrix $\overline{W}$.

(ii)Capitalization of freedom in choice of eigenvalues given desired corners of hypercube as eigenvectors and a specific choice of diagonal elements of $\overline{W}$. It is specifically shown that that a linear system of equations with the coefficient matrix being doubly stochastic naturally arises in expressing the vector of diagonal elements of $\overline{W}$ in terms of the eigenvalues of the symmetric matrix $\overline{W}$. (This discovery is a very general linear algebraic result not found in the literature).

(iii) Proving that when $\overline{W}$ is positive definite, there is no loss of generality in choosing the threshold vector to be zero for ensuring energy function to be a pure quadratic form in the state vector. Based on result in [21], this ensures that stable states of such a HAM are related to the cuts in the graph associated with the HAM.

This research paper is organized as follows. In Section II, relevant research literature is reviewed. In Section III, $\varepsilon - perturbation$ of synaptic weight matrix is considered and the perturbation analysis of discrete time Hopfield neural network is investigated. In Section IV, more general perturbation of synaptic weight matrix is considered and the dynamics of associated HAM is investigated. The research paper concludes in Section V.

II. Review of Related Research Literature

We now briefly explain the operation of Hopfield Neural Network (HNN) which acts as an associative memory (HAM).Consider a set of N, McCulloch-Pitts neurons which are connected to one another by edges with associated symmetric synaptic weights (i.e. edge weight from node 'i' to node 'j' is same as the weight from node 'j' to node 'i'). Thus, the synaptic weight matrix, $\overline{W}$ is a symmetric matrix (i.e. $\overline{W} = \overline{W}^T$). The connection structure of such an ANN is a weighted, undirected graph G. Each node of such a graph is associated with a threshold captured by means of a threshold vector, $\overline{T}$. Since each of the N Mc-Culloch-Pitts neurons assume only { +1, -1 } values, the state vector of such an ANN constitutes the symmetric unit hypercube, H. The network of neurons have no external input. The dynamics of the ANN is driven by the initial state vector of N neurons. Let the network state at time 'n' be denoted by

$$\overline{V}(n) = (\, v_1(n) \;\; v_2(n) \;\; v_3(n) \cdots v_{M-1}(n) \;\; v_N(n) \,) \; with$$
$$v_i(n) = +1 \; or \; -1 \; for \; all \; 1 \le i \le N \; and \; n \ge 0 \; i.e.$$
$$state \; vector \; lies \; on \; the \; unit \; hypercube \; H.$$

The Hopfield Neural Network (HNN), thus constitutes a homogeneous nonlinear dynamical system [5]. Such an ANN operates in the following modes of operation.

- Serial Mode of Operation: The state of only one neuron is updated at any given time 'n' i.e.

$$v_i(n+1) = Sign \left\{ \sum_{j=1}^{M} w_{ij} v_j(n) - t_i \right\} for \;\; n \ge 0.$$

- Fully Parallel Mode of Operation: The state of all M neurons is updated at any time 'n' i.e.

$$\overline{V}(n+1) = Sign \{ \overline{W} \; \overline{V}(n) - \overline{T} \} \; for \; n \ge 0.$$

- Other Parallel Modes of Operation: The state of more than one neuron, but strictly less than M neurons is updated at any time 'n'.

In the state space of HNN, there are distinguished states, called "stable states" and "anti-stable states". They are defined in the following manner:

Definition: A state $\overline{Z}$, lying on the symmetric unit hypercube is called a Stable State if

$$\overline{Z} = Sign \{ \overline{W} \, \overline{Z} - \overline{T} \} .$$

Definition: A state $\overline{U}$, lying on the symmetric unit hypercube is called an Anti-Stable State [4] if

$$\overline{U} = - \, Sign \{ \overline{W} \, \overline{U} - \overline{T} \} .$$

It is well known that the dynamics of Hopfield Neural Network (HNN) satisfies the following Convergence Theorem:

Theorem 1: Consider a Hopfield Neural Network operating in the serial, fully parallel modes of operation

(i) In the serial mode of operation, starting in any initial corner of hypercube, H (as the state at time '0'), the HNN always converges to a stable state if all the diagonal elements are non-negative.

Also

(ii) In the fully parallel mode of operation, starting in any initial condition (lying on, H), the HNN either converges to a stable state (if all the diagonal elements are nonnegative) or a cycle of length atmost 2 is reached.

Proof: The well known proof is documented in [ ], [ ].

Such a Theorem enables utilization of HNN as an associative memory, the so called Hopfield Associative Memory (HAM).

The proof of convergence Theorem is based on associating an energy function with the state vector of the ANN and reasoning that it is non-decreasing as a function of time. The proof of convergence Theorem requires that all the diagonal elements of the synaptic weight matrix $\overline{W}$ are non-negative. Most researchers assume that such a condition always holds true and proceed with analysis of HNN.

This research paper is based on converting a synaptic weight matrix, $\overline{W}$, some of whose diagonal elements are negative into a matrix, $\widehat{W}$ all of whose diagonal elements are non-negative. The notion of "suitable perturbation is formalized in the following sections

This research paper investigates the relationship between dynamics of HAM based on $\overline{W}$, with that of HNN based on $\widehat{W}$ (obtained by "suitable perturbation" of elements of $\overline{W}$). Detailed results are presented in the following section. The results are based on linear algebraic arguments.

III. DISCRETE TIME HOPFIELD NEURAL NETWORK: PERTURBATION ANALYSIS:

We now provide a sufficient condition which ensures that the diagonal elements of $\overline{W}$ are all nonnegative. The following lemma is documented for completeness. It is based on linear algebraic results associated with a symmetric matrix. It is included for completeness.

Lemma 1: If $\overline{W}$ is a positive definite matrix (or even a positive semi-definite matrix), then all its diagonal elements are non-negative.

Proof: The symmetric matrix $\overline{W}$ has the following spectral representation.i.e.

$$\overline{W} = \sum_{i=1}^{N} \mu_i \, \bar{f}_i \; \bar{f}_i^T \;\; , \qquad where$$

$\mu_i' s$ are the eigenvalues and $\bar{f}_i' s$ are the corresponding eigenvectors (forming an orthonormal basis).

Since, $\overline{W}$ is positive definite, $\mu_i \ge 0 \; for \; all \; 'i'$.

Hence,

$\overline{W_{JJ}} = \sum_{i=1}^{N} \mu_i \left(\bar{f}_{i_j} \right)^2$, $where \; \bar{f}_{i_j} \; is \; the \; jth \; component$ of $i^{th} \; eigenvector \; \bar{f}_i$. $Thus, \; we \; have \; that$

$$\overline{W_{JJ}} \ge 0 \; for \; all \; 'j'.$$

An alternative way to prove that the result is to use the fact that a positive definite matrix has the associated Cholesky decomposition i.e.

$\overline{W} = \overline{L} \, \overline{L}^T$, where $\overline{L}$ is a lower triangular matrix. Hence, it readily follows that

$$\overline{W}_{jj} \;\; = \;\; L_{j1}^2 + L_{j2}^2 + \cdots + L_{jj}^2 \;\; for \;\; 1 \le j \le N.$$

Thus,

$$\overline{W_{JJ}} \geq 0 \;\; for \;\; all \;\; 'j' \ldots\ldots\ldots..QED.$$

Note: It readily follows that $\overline{W}$ can never be negative definite if $\overline{W_{ii}} \geq 0 \;\; for \;\; all \;\; i.$

I)n view of the above Lemma, if $\overline{W}$ has some negative diagonal elements, it cannot be positive definite. It readily follows that the simplest way to ensure that all the diagonal elements of $\overline{W}$ are converted into non-negative values is through "$\varepsilon - perturbation$ of diagonal elements of $\overline{W}$. We have the following definition

Definition: Given a symmetric matrix, $\bar{B}$, *ε-perturbation* of such matrix is defined as (resulting in a new matrix $\hat{B}$)

$$\hat{B} = \bar{B} + \epsilon \, I$$

(i.e. all the diagonal elements of symmetric matrix B are perturbed by ϵ).

In the following discussion, we explain the details of such *ε-perturbation* idea. It constitutes one possible "sutiable" perturbation of the elements of $\overline{W}$. The concept suitable perturbation includes the following two cases:

(i)Given that the diagonal elements are negative, let the "additive perturbation matrix", $\bar{R}$ be such that $\bar{R} = \epsilon \, I$ and $\widehat{W} = \overline{W} + \bar{R}$. It readily follows that if $\bar{f}$ is an eigenvector of $\overline{W}$ corresponding to eigenvalue μ, we have $\widehat{W}\,\bar{f} = (\overline{W} + \bar{R})\,\bar{f} = \mu\,\bar{f} + \epsilon\,\bar{f} = (\mu + \epsilon)\,\bar{f}$. Thus, we have that ($\mu + \epsilon$) is an eigenvalue of $\widehat{W}$ with the corresponding eigenvector being $\bar{f}$.

(ii) The perturbation matrix, $\bar{R}$ has the same set of eigenvectors as the synaptic weight matrix, $\overline{W}$. The eigenvalues of $\overline{W}$ could be different from those of $\bar{R}$.

We prove in this research paper that under suitable perturbation,'interesting dynamics" of HAM is preserved. The notion of "interesting dynamics" is defined in the following sense:

Interesting dynamics of HAM: Based on the convergence Theorem of HAM, we consider "interesting dynamics" of it to be convergence to stable state starting in any initial state.

We now consider the case (i)

Let w_{min} be the minimum of all negative diagonal elements of $\overline{W}$ $i.e.$

$$w_{min} = \min_{1 \leq j \leq N} w_{jj} \;.$$

Let $\epsilon > |\, w_{min} \,|$. It readily follows that

$$\widehat{W} = \overline{W} + \epsilon \, I$$

is a symmetric matrix with all diagonal elements being non-negative. Thus, the HAM based on such a synaptic weight matrix converges to a stable state in the serial mode of operation. In the following lemmas, we reason that the programmed/desired stable states [1,2,3] as well as arbitrary stable states of HAM based on $\overline{W}$ are same as those of HAM based on $\widehat{W}$. We assume that the threshold vector $\bar{T} \equiv \bar{0}$. Formally, we have the following result.

Lemma 2: Consider the spectral representation of a symmetric matrix, $\overline{W}$. i.e.

$$\overline{W} = \sum_{i=1}^{N} \mu_i \; \bar{g}_i \; \bar{g}_i^T, \;\; where$$

$\bar{g}_i' s$ are the corners of hypercube (i.e. { +1, -1 } vectors) that are eigenvectors of $\overline{W}$ corresponding to eigenvalues, $\mu_i' s$. Consider $\epsilon - perturbation \; of \; \overline{W} \; i.e. \; with \; \epsilon > 0$

$$\widehat{W} = \overline{W} + \epsilon \, I \, .$$

The " programmed" stable states of HAM associated with $\overline{W}$ are all same as those associated with $\widehat{W}$ (threshold vector $\bar{T} \equiv \bar{0}$)

Proof: From [3], the eigenvector of $\overline{W}$ corresponding to positive eigenvalue that is also a corner of hypercube is the stable state of HAM with the synaptic weight matrix $\overline{W}$. Such a stable state is called the "programmed stable state". More clearly, since

$$\overline{W}\;\bar{g}_i = \mu_i \;\bar{g}_i$$

and $\bar{g}_i$ is a corner of hypercube, we have that

$$\text{Sign}\,(\overline{W}\;\bar{g}_i) = \text{Sign}\,(\mu_i\;\bar{g}_i) = \bar{g}_i \;\; if \;\; \mu > 0 \;.$$

Also,

$$\widehat{W}\,\bar{g}_i = \overline{W}\,\bar{g}_i + \epsilon\;\bar{g}_i = (\mu_i + \epsilon)\,\bar{g}_i \,.$$

Hence, $\bar{g}_i$ is also the eigenvector of $\widehat{W}$ corresponding to the eigenvalue ($\mu_i + \epsilon$). Thus,

Sign $(\widehat{W}\,\bar{g}_i) = Sign\,((\mu_i + \epsilon)\,\bar{g}_i) = \bar{g}_i$ if $\mu_i > 0, \; \epsilon > 0.$

Thus, we have the desired result Q.E.D.

Note: Suppose $Sign\;(\mu_i + \epsilon) = Sign\,(\mu_i)$. Then, from the above proof, it readily follows that the programmed stable, anti-stable states of $\overline{W}, \widehat{W}$ are same.

Note: It readily follows [3] that the proof argument generalizes to the case when $\bar{T} \neq \bar{0}$, $based \; on \; the \; result \; in$ [3]

Now, with $\widehat{\varepsilon - perturbation}$ of $\widehat{W}$, (with $\varepsilon > 0$) i.e.

$$\widehat{W} = \overline{W} + \varepsilon \, I \,,$$

we reason that any stable state (i.e. even the non-programmed ones) of HAM with $\overline{W}$ as synaptic weight matrix is same as those of HAM based on $\widehat{W}$.

Lemma 3: Every stable state of HAM based on synaptic weight matrix, $\overline{W}$ is same as that of HAM associated with synaptic weight matrix, $\widehat{W}$.

Proof: Let $\bar{g}$ be a stable state of $\overline{W}$ $i.e.$

$$Sign\,(\overline{W}\;\bar{g}) = \bar{g} \;.$$

Since $\varepsilon > 0$, we have that

$Sign(\widehat{W}\,\bar{g}) = Sign\,(\overline{W}\,\bar{g} + \varepsilon\,\bar{g}) = \bar{g}$

(*since multiplication of* $\bar{g}$ *by* ε *and addition with* $\overline{W}\,\bar{g}$, *preserves the sign structure of* $\overline{W}\,\bar{g}$).

Thus, stable states of $\overline{W}$ are preserved under the $\varepsilon - perturbation \; of \; it \; (resulting \; in \; \widehat{W})$..........QED

Note: The above lemma doesnot claim that the stable states of $\overline{W}, \widehat{W}$ are all same. In fact, we prove in Lemma 4 that if $\widehat{W}$ is positive definite, all corners of hypercube are stable states of HAM based on it. It resadily follows by choice of sufficiently large positive value of $\epsilon, \widehat{w}$ can always be made positive definite.

Note: This lemma leads to the conclusion that $\varepsilon - perturbation$ of diagonal elements of $\overline{W}$ ensures that the sufficient condition for convergence of HAM is ensured and that the stable states are preserved. This is a very interesting result.

Note: For notational convenience, stable/anti-stable states of HAM based on $\overline{W}/\widehat{W}$ are called as the stable/anti-stables states of $\overline{W}/\widehat{W}$.

- We now focus on the linear algebraic properties of the $\varepsilon - perturbation$ of $\overline{W}$.

Let the spectral representation of $\overline{W}$ be

$$\overline{W} = \sum_{i=1}^{N} \mu_i \bar{f}_i \bar{f}_i^T .$$

Hence,

$$\widehat{W}\ \bar{f}_i = (\overline{W} + \varepsilon I)\ \bar{f}_i = \overline{W}\ \bar{f}_i + \varepsilon \bar{f}_i = (\mu_i + \varepsilon)\ \bar{f}_i .$$

- Suppose,,$\overline{W}$ *is positive definite i.e.* $\mu_i > 0$ *for all 'i'.*

Since, $\varepsilon > 0$ and the eigenvalues of $\widehat{W}$ are $\mu_i + \varepsilon$ for all 'i', $\widehat{W}$ will be positive definite.

- Suppose, $\overline{W}$ *is NOT positive definite i.e. smallest eigenvalue of* $\overline{W}$ (*i.e.* μ_{min}) *is negative i.e.*

$\mu_{min} < 0$. *Suppose* $\varepsilon > -\mu_{min}$. *Then* $\mu_{min} + \varepsilon > 0$.

Consequently, $\mu_i + \varepsilon > 0$ *for all 'i'*. Thus, with such

$\varepsilon - perturbation$, $\widehat{W}$ *will be positive definite.*

- $Trace(\widehat{W}) = Trace(\overline{W} + \varepsilon I) = Trace(\overline{W}) + \varepsilon N$.
- Since, the threshold vector, $\bar{T} \equiv \bar{0}$ (*zero vector*), *the energy function* associated with the dynamics of HAM based on $\overline{W}$ is
$$\bar{V}^T(n)\ \overline{W}\ \bar{V}(n) = Trace(\overline{W}) + \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} v_i(n)\ w_{ij}\ v_j(n).$$
Hence
$$\bar{V}^T(n)\ \widehat{W}\ \bar{V}(n) = \bar{V}^T(n)\ (\overline{W} + \varepsilon\ I)\ \bar{V}(n) = \bar{V}^T(n)\ \overline{W}\ \bar{V}(n) + \varepsilon N.$$

Lemma 4: If $\widehat{W}$ is positive definite, then there are no antistable states in the dynamics of HAM associated with $\widehat{W}$. Furthermore, if $\widehat{W}$ is diagonally dominant matrix, all the corners of unit hypercube are stable states.

Proof: Suppose, $\bar{U}$ is an anti-stable state of HAM based on synaptic weight matrix, $\widehat{W}$ *i.e.*

$$\bar{U} = -Sign(\widehat{W}\ \bar{U}).$$

Hence, it readily follows that

$$\bar{U}^T\ \widehat{W}\ \bar{U} < 0.$$

It contradicts the fact that $\widehat{W}$ is a positive definite matrix.

Thus, there are no anti-stable states associated with the dynamics of HAM based on $\widehat{W}$.

Now, we consider a diagonally dominant matrix $\widehat{W}$ *i.e.*

$$\widehat{W}_{ii} > \sum_{\substack{j=1 \\ j \neq i}}^{N} |\ \widehat{W}_{ij}|\ .$$

It readily follows that by choosing sufficiently large value of ε, $\widehat{W} = \overline{W} + \varepsilon\ I$ can be ensured to be a diagonally dominant matrix.

Consider a corner of unit symmetric hypercube, $\bar{U}$. Since

$$\widehat{W}_{jj} > \sum_{\substack{k=1 \\ k \neq j}}^{N} |\ \widehat{W}_{jk}|\ , we\ have\ that$$

$$Sign(\widehat{W}\ \bar{U}) = \bar{U}\ for\ all\ corners\ of\ unit\ hypercube.$$

Thus, every corner of unit hypercube is a stable state.

QED

Note: Every diagonally dominant matrix is a positive definite matrix.

Note: By successively increasing the positive value of ε, it can be ensured that all diagonal elements of $\widehat{W}$ are positive/non-negative (satisfying the sufficient condition for the convergence theorem to hold), $\widehat{W}$ is positive definite and even $\widehat{W}$ is a diagonally dominant matrix. Such $\varepsilon - perturbation$ ensures "suitable" preservation of stable states.

The above discussion is concerned with a special perturbation of elements of $\overline{W}$. We are naturally led to more general perturbation of elements of $\overline{W}$. Thus, in the following section, we consider ROBUST HOPFIELD NEURAL NETWORK and its dynamics under special perturbation of elements of $\overline{W}$. The perturbation could be due to noise corrupting the synaptic weights.

In the above discussion, we assumed that the threshold vector is a zero vector i.e. $\bar{T} \equiv \bar{0}$. In the following lemma, we reason that (in a well defined sense), if $\overline{W}$ is a positive definite matrix, there is no loss of generality in assuming . $\bar{T} \equiv \bar{0}$.

Lemma 5: Let $\overline{W}$ be the synaptic weight matrix of HAM with the associated threshold vector being a non-zero vector i.e. $\bar{T} \neq \bar{0}$. Let $\overline{W}$ be a positive definite matrix. Consider associated HAM with synaptic weight matrix, $\widehat{W}$ i.e

$$, \widehat{W} = \begin{bmatrix} \overline{W} & -\bar{T} \\ -\bar{T}^T & \alpha \end{bmatrix}\ with\ \alpha > \sum_{i=1}^{N} |\ T_i\ |\ \ and$$

threshold vector, $\hat{T} \equiv \bar{0}$. $\bar{g}$ is a stable state of HAM with synaptic weight matrix $\overline{W}$ if and only if $\begin{bmatrix} \bar{g} \\ 1 \end{bmatrix}$ is a stable state of HAM with associated synaptic weight matrix $\widehat{W}$.

Proof: From the definition of $\widehat{W}$ *and* α, we have that

$$\widehat{W}\begin{bmatrix}\bar{g}\\1\end{bmatrix} = \begin{bmatrix}\overline{W}\ \bar{g}-\bar{T}\\1\end{bmatrix}.$$ Hence, it follows that

$$Sign\left\{\widehat{W}\begin{bmatrix}\bar{g}\\1\end{bmatrix}\right\} = \begin{bmatrix}Sign(\overline{W}\ \bar{g}-\bar{T})\\1\end{bmatrix}.$$

Thus, the claim holds. There is a one-to-one correspondence between stable states of $\overline{W}, \widehat{W}$. QED

IV. ROBUST HOPFIELD NEURAL NETWORK: DYNAMICS

In this section, we consider the dynamics of HAM based on $\overline{W}$ when its elements are subjected to interesting perturbations [6-20]. We consider some interesting cases:

CASE (A): All the elements of the symmetric matrix, $\overline{W}$ are perturbed by a common value i.e.

$$\widehat{W} = \overline{W} + \varepsilon\ \bar{e}\,\bar{e}^T, \quad where\ \bar{e}^T = (1\ \ 1\ \ \ldots.\ \ 1\ \ 1)$$
$$i.e. row\ vector\ all\ of\ whose\ elements\ are\ 1.$$

NOTE: We realized that (as discussed in [3]), the synaptic weight matrix, $\overline{W}$ can be synthesized using the corners of hypercube as the eigenvectors of $\overline{W}$. When N is odd (as reasoned in [3]), only one corner of hypercube can be the eigenvector of $\overline{W}$. But, when N is even, the orthonormal basis of eigenvectors could be the corners of hypercube. We call such an orthonormal basis of corners of hypercube as the "Hadamard Basis".

Let N be an even number. Suppose $\bar{g}$ is a stable state of $\overline{W}$ with the number of '+1" values (also, the number of -1 values) be $\frac{N}{2}$ (as in the case of Hadamard basis). Thus,

$$\bar{g}^T\bar{e} = \bar{e}^T\bar{g} = 0.$$

Hence,

$$\widehat{W}\bar{g} = \overline{W}\ \bar{g} + \varepsilon\ \bar{e}\,\bar{e}^T\ \bar{g} = \overline{W}\ \bar{g}.$$

Thus,

$$\text{Sign}\ (\widehat{W}\bar{g}) = \text{Sign}\ (\overline{W}\ \bar{g}) = \bar{g}\ .$$

Thus, $\bar{g}$ is a programmed stable state of $\overline{W}$ as well as $\widehat{W}$.

CASE (B): We now consider a generalization of $\varepsilon -$ *perturbation* of elements of symmetric matrix $\overline{W}$ and investigate whether the stable states of $\overline{W}$ are preserved under more general perturbation (of elements of $\overline{W}$).

Let $\widehat{W} = \overline{W} + \bar{R}$.

The perturbation matrix, $\bar{R}$ has the same set of eigenvectors as the synaptic weight matrix, $\overline{W}$. The eigenvalues of $\overline{W}$ could be different from those of $\bar{R}$.

For simplicity, we assume that the threshold vector is identically zero vector.

Lemma 6: Suppose $\{\mu_i\}_{i=1}^N$, $\{\delta_i\}_{i=1}^N$ are the eigenvalues of $\overline{W}, \bar{R}$ respectively.

Let $Sign\ (\mu_i + \delta_i) = Sign(\mu_i)\ for\ 1 \le i \le N$.

Under such condition, the stable states of HAM based on $\overline{W}$ are same as those of HAM based on $\widehat{W}$.

Proof: Let the spectral representation of $\overline{W}$ be

$$\overline{W} = \sum_{i=1}^{N} \mu_i\,\bar{f}_i\,\bar{f}_i^T\ , where\ \bar{f}_i's\ \ are\ the\ eigenvectors\ of$$

symmetric matrix, $\overline{W}$. Let $\bar{g}$ be a stable state of HAM based on $\overline{W}$ *i.e*

$Sign\ (\overline{W}\ \bar{g}) = Sign\left(\sum_{i=1}^N \mu_i\,\alpha_i\,\bar{f}_i\right)\ where\ \alpha_i = \bar{f}_i^T\bar{g}$.

By the hypothesis (that the eigenvectors of $\overline{W}, \bar{R}$ are same), $\mu_i + \delta_i$ is an eigenvalue of $\widehat{W}$ corresponding to the eigenvector $\bar{f}_i$ (common between $\overline{W}, \bar{R}$). Hence, we have that

$$Sign\left(\widehat{W}\bar{g}\right) = Sign\ \left(\sum_{i=1}^{N}(\mu_i+\delta_i)\,\bar{f}_i\,\bar{f}_i^T\ \bar{g}\right)$$
$$= Sign\left(\sum_{i=1}^N(\mu_i+\delta_i)\,\alpha_i\,\bar{f}_i\right).$$

By the condition that

$Sign\ (\mu_i + \delta_i) = Sign(\mu_i)$,

we have that

$$Sign\left(\widehat{W}\bar{g}\right) = \bar{g}.$$

Hence the stable states of $\overline{W}, \widehat{W}$ are all same QED.

Note: The proof readily generalizes to the case where the threshold vector is a non-zero vector.

Note: If eigenvalues of $\overline{W}, \bar{R}$ are of the same sign/polarity, the condition in the lemma is satisfied.

Note: If the set of eigenvalues of of $\overline{W}$ are permuted versions of those of $\bar{R}$, the condition in the lemma statement is satisfied.

Note: It readily follows that $\varepsilon - perturbation$ is a special case of the perturbation considered in CASE (B).

Note: From the proof, it is clear that the programmed as well as non-programmed stable states of $\overline{W}$ are same as those of $\widehat{W}$ (*under the condition on eigenvalues*).

Note: Since, the symmetric matrices $\overline{W}, \bar{R}$ have the same set of eigenvectors, they commute i.e. $\overline{W}\bar{R} = \bar{R}\overline{W}$.

From the above discussion, it is clear that a suitable perturbation of $\overline{W}$ by the symmetric matrix, $\bar{R}$ will preserve the stable states. Specifically, some conditions for perturbation are discussed in Lemma 6.

- **Vector of Eigenvalues of $\overline{W}$: Vector of Diagonal Elemens of $\overline{W}$**

We now address the general question of relationship between eigenvalue vector, $\bar{\gamma}$ of $\overline{W}$ (i.e. vector of eigenvalues of $\overline{W}$) and the diagonal element vector of $\overline{W}$ (i.e. vector whose elements are the diagonal elements of $\overline{W}$), $\breve{g}$. Let

$$\overline{W} = \sum_{i=1}^{N} \mu_i\,\bar{f}_i\,\bar{f}_i^T = \bar{P}\,\bar{D}\bar{P}^T,$$

where $\bar{f}_i's$ *are the eigenvectors of* $\overline{W}, \bar{D}$ *is the*

diagonal matrix of eigenvalues and $\bar{P}$ *is the orthogonal matrix of eigenvectors*

NOTE: The following is a general linear algebraic result associated with any symmetric matrix (It can easily be generalized to any Hermitian matrix).

Lemma 7: The eigenvalue vector, $\bar{\gamma}$ and the diagonal element vector $\breve{g}$ are related through the following system of linear equations

$$\bar{F}\,\bar{\gamma} = \breve{g},$$

where $\bar{F} = \bar{P}\ o\ \bar{P}$ with 'o' denoting the Schur product of matrices ($\bar{F}$ is a doubly stochastic matrix based on the eigenvectors of the symmetric matrix $\bar{W}$).

Proof: Since the eigenvectors of a symmetric matrix form an orthonormal basis,

$$\bar{f}_j^T \bar{f}_j = 1 \ \ for \ \ 1 \leq j \leq N \ \ and$$
$$\bar{f}_i^T \bar{f}_j = 0 \ \ for \ \ all \ \ i \neq j.$$

Also, $\bar{W}$ can be expressed as

$$\bar{W} = \sum_{i=1}^{N} \mu_i \, \bar{E}_i, where$$

$\bar{E}_i's$ *are residue matrices such that* $\sum_{i=1}^{N} \bar{E}_i = \bar{I}$, the identity matrix. Thus, we have that

$$\sum_{j=1}^{N} \ (f_{j_k})^2 = \ 1 \ \ for \ 1 \leq k \leq N, \qquad where \ f_{j_k} \ is \ the$$
$$k^{th} \ component \ of \ the \ j^{th} \ eigenvector.$$

It is clear that, by definition, eigenvectors of the symmetric matrix $\bar{W}$ are column vectors of the orthogonal matrix $\bar{P}$. Thus, letting

$$\bar{F} = \bar{P}\ o\ \bar{P} \ \ (with \ 'o' \ being \ the \ Schur \ product \ operation),$$

we have that $\bar{F}$ is a doubly stochastic matrix.

Now, it is clear that the given eigenvectors, the system of equations

$$\bar{W} = \sum_{i=1}^{N} \mu_i \, \bar{f}_i \, \bar{f}_i^T$$

leads to the following linear equations relating the vector of eigenvalues $\bar{\gamma}$, to the vector of diagonal elements of $\bar{W}\ i.e.\ \breve{g}$.

$$\bar{F}\,\bar{\gamma} = \breve{g},$$

Thus, we have a linear system of equations with the coefficient matrix being a doubly stochastic matrix QED

Corollary: If the doubly stochastic matrix, $\bar{F}$, is non-singular, then there is a unique solution for the vector of eigenvalues of $\bar{W}\ i.e.$

$$\bar{\gamma} = \bar{F}^{-1}\,\breve{g},$$

Also, if $\bar{F}$ is singular, then there are infinitely many solutions for the eigenvalue vector $\bar{\gamma}$.

Note: In view of lemmas, 6 and 7, given the eigenvalues of $\widehat{w}$ which are based on the solution of above system of linear equations (leading to a desired positive vector $\breve{g}$), for a given matrix $\bar{W}$ with known eigenvalues, we can readily determine the eigenvalues of R. It shoulde be kept in mind that in the case of Lemma 6, the eigenvectors of matrices, $\bar{W},\ \bar{R}\ and\ \widehat{W}$ are all same

Note: The Lemma 7 holds true for any symmetric matrix, $\widehat{W}$ and its diagonal element vector, $\breve{g}$, vector of eigenvalues $\bar{\gamma}$.

Note: The above Lemma can be generalized for any Hermitian matrix (whose diagonal elements are real numbers). Details are avoided for brevity.

Note: In the spirit of above Lemma, given the eigenvectors of a symmetric matrix, linear system of equations can naturally be associated with expressing any vector of N elements of W (for instance the trailing diagonal elements) in terms of the eigenvalues.

- ZERO-FORCING EIGENVALUE VECTOR:

From the convergence Theorem, it is sufficient for the diagonal elements of the synaptic weight matrix to be non-negative. Thus, for convergence Theorem to hold, $\breve{g}$ can be a zero vector. Hence, the system of linear equations reduce to

$$\bar{F}\,\bar{\gamma} = \bar{0}.$$

Thus, for zero-forcing eigenvector solution, we are interested in the vectors in the null space of $\bar{F}$. $If\ \bar{F}$ is non-singular, the zero vector is the only vector in the null space of $\bar{F}$. Also, there are infinitely many solutions for the eigenvalue vector if $\bar{F}$ is singular.

To illustrate the Lemma 7 and the above results, we provide an example:

Example: Consider the eigenvectors of $\bar{\boldsymbol{W}}$ which constitute a normalized Hadamard basis (normalized columns of a Hadamard matrix) i.e. the orthogonal matrix, $\bar{P}$ of eigenvectors (arising in the spectral representation of $\overline{W}$). Hence, it readily follows that the doubly stochastic matrix, $\bar{F}$ is given by

$$\bar{F} = \bar{P}\ o\ \bar{P}^T = \frac{1}{N}\ \bar{e}\ \bar{e}^{\ T}\ , \ where$$

$\bar{e}$ is a column vector of 'ones' and 'o' denotes the Schur product. Thus, $\bar{F}$ is a doubly stochastic singular matrix of rank one. Thus, given a non-negative diagonal element vector, $\breve{g}$, (sum of all the element of the non-negative diagonal vector is the trace of the matrix, $\bar{W}$) one possible solution for $\bar{\gamma}$ is

$$\bar{\gamma} = \frac{1}{N}\ Trace\ (\overline{W})\ \bar{e}\ ,$$

where $\bar{e}$ is a vector all ones. Infinitely many solutions for $\bar{\gamma}$ are generated using the vectors in the null space of the matrix $\bar{F}$. They are given by

$$\bar{\gamma} = \frac{1}{N}\ Trace\ (\overline{W})\ \bar{e}\ +\ \alpha\,\bar{J}, \qquad where$$

$\bar{J}$ lies in the null space of doubly stochastic matrix $\bar{F}$.

We now investigate the operations on symmetric matrices that preserve the stable/anti-stable states:

- **Algebraic Structure of Symmetric Matrices with same programmed (desired)/ Non-programmed Stable States:**

The following Lemma sheds light on operations on symmetric matrices which preserve the stable/anti-stable states.

Lemma 8: The following inferences hold true (i) If $\bar{f}$ is a stable state of symmetric matrices, $\bar{A}, \bar{B}$, then $\bar{f}$ is a stable state of $\bar{A} + \bar{B}$,
(ii) If $\bar{f}$ is an anti-stable state of symmetric matrices, $\bar{A}, \bar{B}$, then $\bar{f}$ is an anti-stable state of $\bar{A} + \bar{B}$,
(iii) If $\bar{f}_1, \bar{f}_2$ are cycle states of length 2 of symmetric matrices, $\bar{A}, \bar{B}$, then $\bar{f}_1, \bar{f}_2$ are cycle states of length 2 of of $\bar{A} + \bar{B}$,
(iv) Suppose $\bar{f}$ is a programmed/desired stable state of $\bar{A}$. It is also a programmed/desired stable or anti-stable state of $\bar{A} + \bar{B}$ if and only if , it is a programmed stable or anti-stable state of $\bar{B}$ ($\bar{f}$ can be an eigenvector of $\overline{B}$ corresponding to zero eigenvalue i.e. $\bar{f}$ lies in null space of $\overline{B}$).

Proof:

(i) $\bar{f}$ is a stable state of symmetric matrices, $\bar{A}, \bar{B}$ *i.e.*

$$Sign(\bar{A}\,\bar{f}) = \bar{f} \text{ and } Sign(\bar{B}\,\bar{f}) = \bar{f}.$$

Let $\bar{A}\,\bar{f} = \bar{h}$ and $\bar{B}\bar{f} = \bar{k}$. Hence, we have that Sign ($(\bar{A}+\bar{B})\,\bar{f}$) = Sign $(\bar{h}+\bar{k})$ = $\bar{f}$.

(ii) Similar reasoning as in (i) gives the desired result

(iii) Let $\bar{f}_1, \bar{f}_2$ are cycle states of length 2 of symmetric matrices, $\bar{A}, \bar{B}$ *i.e.*

$Sign(\bar{A}\bar{f}_1) = \bar{f}_2$ and $Sign(\bar{A}\bar{f}_2) = \bar{f}_1$

$Sign(\bar{B}\bar{f}_1) = \bar{f}_2$ and $Sign(\bar{B}\bar{f}_2) = \bar{f}_1$.

As in (i) above,

$$Sign((\bar{A}+B)\,\bar{f}_1) = Sign(\bar{A}\bar{f}_1 + \bar{B}\bar{f}_1) = \bar{f}_2 \text{ and}$$

$Sign((\bar{A}+B)\,\bar{f}_2) = \bar{f}_1$ i.e. cycles of length 2 are also preserved under the addition of two associated symmetric matrices.

(iv) Suppose $\bar{f}$ is a programmed/desired stable state of $\bar{A}$ *i.e.*

$$\bar{A}\bar{f} = \mu\bar{f} \text{ with } \mu > 0 \text{ and } Sign(\bar{A}\bar{f}) = \bar{f}.$$

Now suppose $\bar{f}$ is also a programmed/desired stable state of

$$\bar{A}+\bar{B} \text{ i.e.}$$

$$(\bar{A}+\bar{B})\,\bar{f} = \theta\bar{f} \text{ with } \theta > 0 \text{ i.e. } Sign((\bar{A}+\bar{B})\,\bar{f}) = \bar{f}.$$

Hence, it readily follows that

$$\bar{B}\bar{f} = (\theta-\mu)\bar{f}.$$

Thus, $\bar{f}$ is a programmed/ desired stable/anti-stable state of $\bar{B}$.

Similar reasoning can be applied to programmed/desired anti-stable state of $\bar{A}$ ....…………………………..QED

- Graph Theoretic Significance of Perturbation Models:

In [21], it is proved that the global optimum stable state in the Hopfield Neural Network corresponds to the minimum cut in the graph associated with HAM. In view of the results in Section III, Section IV, we infer that the suitable perturbations of elements of $\bar{W}$, preserve the minimum cut in the associated graphs.

V. DISCRETE TIME HOPFIELD NEURAL NETWORK: QUANTIZATION OF SYNAPTIC WEIGHTS:

In many practical implimentations of discrete time Hopfield neural network (hardware and software implimentations), it is necessary to quantize the elements of synaptic weight matrix so that the elements are integers (e.g. quantization to 8 bits). Under such perturbation of elements of W, it is necessary to investigation of preservation of interesting dynamics of HAM (e.g. the stable states). As discussed in earlier sections, we are led to the perturbation of only the eigenvalues:

- The results in earlier sections readily apply for quantization of elements of W based on perturbation of only the eigenvalues.

CLAIM: We now consider "sign preserving" perturbation of eigenvalues of W i.e. When an eigenvalue of W is perturbed so that its "sign"/"polarity" remains invariant (i.e. positive/negative eigenvalue remains positive/negative). Under such condition it readily follows that "programmed/desired" stable/anti-stable states are preserved after quantization. Also, when some eigenvalues are very small (in absolute value), it is possible to ZERO them out i.e. the dimension of null space of W is increased. Based on the results of the author in [3], detailed results can be derived. Details are avoided for brevity. W

We now derive an interesting sufficient condition which ensures that the stable states are preserved under perturbation. Suppose the perturbation matrix is R i.e. $\widetilde{W} = \bar{W} + \bar{R}$. Let $\bar{f}$ be a stable state of $\bar{W}$ *and let* $Sign(\bar{W}\bar{f}) = Sign(\bar{R}\bar{f})$. Then it readily follows that $\bar{f}$ is also a stable state of $\widetilde{W}$.

VI. CONCLUSIONS

In this research paper, it is shown that the simplest perturbation of diagonal elements of synaptic weight matrix of HAM ensures that the convergence theorem is satisfied and at the same time stables states of HAM dynamics are preserved. It is formally proved that even under more general perturbation model, the stable states of associated HAM are preserved. It is proved that given a symmetric matrix $\bar{W}$, its vector of eigenvalues and its vector of diagonal elements of $\bar{W}$ are related through a linear system of equations with the coefficient matrix being a doubly stochastic matrix.

REFERENCES:

[1] G. Rama Murthy and Moncef Gabbouj,"On the Design of Hpfield Neural Networks: Synthesis of Hopfield Type Associative Memories," International Joint Conference on Neural Networks (IJCNN), 2015

[2] G. Rama Murthy, Vamshi, Devaki and Divya,"Synthesis/Programming of Hopfield Associative Memory," Proceedings of International Conference on Machine Learning and Data Science (ICMLDS 2019), December 2019 (ACM Digital Library)

[3] G. Rama Murthy, "Toward Optimal synthesis of Discrete Time Hopfield Neural Network," IEEE Transactions on Neural Networks and Learning Systems,pp 1-6, March 2022

[4] G. Rama Murthy and B. Nischal," Hopfield-Amari Neural Network : Minimization of Quadratic forms," The 6th International Conference on Soft Computing and Intelligent Systems, Kobe Convention Center (Kobe Portopia Hotel) November 20-24, 2012, Kobe, Japan.

[5] J. J. Hopfield,"Neural Networks and Physical Systems with emergent computational abilities," Proceedings of National Academy of Sciences, USA, vol. 79, 1982, pp. 2554-2558

[6] Liu, Peng, Jun Wang, and Zhigang Zeng. "An overview of the stability analysis of recurrent neural networks with multiple equilibria." IEEE Transactions on Neural Networks and Learning Systems (2021).

[7] Kobayashi, Masaki. "Synthesis of complex-and hyperbolic-valued Hopfield neural networks." Neurocomputing 423 (2021): 80-88.

[8] Feng, Naiqin, Xiuqin Geng, and Bin Sun. "Study on Neural Network Integration Method Based on Morphological Associative Memory Framework." Neural Processing Letters 53, no. 6 (2021): 3915-3945.

[9] Dai, Xiaoliang, Hongxu Yin, and Niraj K. Jha. "NeST: A neural network synthesis tool based on a grow-and-prune paradigm." IEEE Transactions on Computers 68, of no. 10 (2019): 1487-1497.

[10] Zeng, Zhigang, and Jun Wang. "Design and analysis of high-capacity associative memories based on a class of discrete-time recurrent neural networks." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 38, no. 6 (2008): 1525-1536.

[11] Reis, Alcir G., José Luis Acebal, Rogério M. Gomes, and Henrique E. Borges. "Space-vector structure based synthesis for hierarchically coupled associative memories." In 2006 Ninth Brazilian Symposium on Neural Networks (SBRN'06), pp. 190-195. IEEE, 2006.

[12] Cao, Jinde, and Jun Wang. "Global asymptotic stability of a general class of recurrent neural networks with time-varying delays." IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 50, no. 1 (2003): 34-44.

[13] Cai, Guanghui, and Jihong Shi. "Estimation of attraction domain and exponential convergence rate of Hopfield-type associative memory neural network." In Proceedings of the 4th World Congress on Intelligent Control and Automation (Cat. No. 02EX527), vol. 3, pp. 2435-2439. IEEE, 2002.

[14] Liu, Derong, and Zanjun Lu. "A new synthesis approach for feedback neural networks based on the perceptron training algorithm." IEEE transactions on Neural Networks 8, no. 6 (1997): 1468-1482.

[15] Lillo, Walter E., David C. Miller, Stefen Hui, and Stanislaw H. Zak. "Synthesis of brain-state-in-a-box (BSB) based associativememories." IEEE transactions on Neural Networks 5, no. 5 (1994): 730-737

[16] Liu, Derong, and Anthony N. Michel. "Sparsely interconnected neural networks for associative memories with applications to cellular neural networks." IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing 41, no. 4 (1994): 295-307.

[17] Jang, J-SR, and C-T. Sun. "Functional equivalence between radial basis function networks and fuzzy inference systems." IEEE transactions on Neural Networks 4, no. 1 (1993): 156-159.

[18] Michel, Anthony N., and Jay A. Farrell. "Associative memories via artificial neural networks." IEEE Control Systems Magazine 10, no. 3 (1990): 6-17.

[19] Farrell, J. A., and A. N. Michel. "A synthesis procedure for Hopfield's continuous-time associative memory." IEEE Transactions on Circuits and Systems 37, no. 7 (1990): 877-884.

[20] Krishnamurthy, Ashok K., Stanley C. Ahalt, Douglas E. Melton, and Prakoon Chen. "Neural networks for vector quantization of speech and images." IEEE journal on selected areas in Communication, October 1990.

[21] J. Bruck and M. Blaum,"Neural Networks, Error Correcting Codea and Polynomials over the binary n-cube," IEEE Transactions on Information Theory, vol. 35(5), pp.976-989, 1989

Author.

Dr. Garimella Rama Murthy received Ph.D. degree in Computer Engineering from Purdue University, West Lafayette, USA in 1989. He worked as a Member of Technical Staff at Bellcore, USA. After returning to India, he worked as an Associate Professor at IIIT-Hyderabad before joining Mahindra Ecole Centrale in July 2018 as a professor in the Computer Science department. He has about 20 years of teaching experience. His current research interests are in Machine Learning, Performance Evaluation of Computing and Communication Systems, Design of autonomous driving vehicles, Internet of Things etc. He is a Senior Member of ACM and a senior member of IEEE. He received many awards including the Rastriya Gaurav Award.