Spherical Seperability: Associative Memories

Rama Murthy Garimella and Jagannadha Swamy Tata

February 9, 2021

Spherical Seperability : Associative Memories

1st Garimella Ramamurthy
School of Computer Science and Engineering
Mahindra University
Hyderabad, India
rama.murthy@mahindrauniversity.edu.in

2nd Tata Jagannadha Swamy
Electronics and Communication Engineering
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
tatajagan@gmail.com

***Abstract*—In this research paper, based on the concept of spherical seperabilityy, novel associative memories are proposed. The dynamics of the associative memories is shown to lead to a stable state or a cycle of length atmost 2, starting in an initial condition.**

***Index Terms*—Artificial Neural Networks, Spherical Seperability, Novel Associative memories, Hopfield Neural Networks, Hopfield Associative Memory**

I. INTRODUCTION

In an effort to understand the operation of biological neuron, McCulloch-Pitts proposed a model of biological neuron. This model of neuron was motivated by the concept of "linear separability" of patterns that need to be classified into two or more classes. Since McCulloch-Pitts neuron doesn't have the "training" ability, perceptron model was proposed. Using Perceptron learning law it was established that single "layer" perceptron can classify linearly seperable patterns. In an effort to classify non-linearly seperable patterns, Multi-Layer perceptron (MLP) was conceived and utilized successfully in many applications. Using McCulloch-Pitts neuron model, Hopfield proposed an Artificial Neural Network (ANN) which acts as an associative memory. The author, in his research efforts proposed the concept of "spherical separability of patterns and established that linear separability implies spherical separability but not the other way [1], [2]. Thus, a natural question that remained was whether it is possible to propose an ANN based on spherical separability that acts as an associative memory? This research paper is an effort to answer such a question.

In more clear terms, Artificial Neural Networks(ANNs), such as Single Layer Perceptron(SLP) were proposed based on the concept of linear seperability of patterns [3]. This model of neuron based on McCulloch-Pitts neuron was successfully utilized to arrive at Hopfield Associate Memory(HAM). The authors proposed the concept of "spherical seperabilityy" and reasoned that linear seperability implies spherical spereability (under mild conditions) but not the otherway. ANNs based on spherical seperability were proposed successfully [1],[2]. A natural question that remained was whether an associative memory can be arrived at using the concept of spherical seperability.

Vapnick, in an effort to increase the noise immunity of perceptron proposed the concept of Support Vector Machine (SVM). SVM in a well defined sense (maximization of "MARGIN") leads to the concept of "optimal linear seperability". Various researchers investigated SVM design in higher dimensions by suitably projecting the patterns (that are not linearly seperable) so that they become linearly seperable (using suitable kernel functions) in higher dimensional space. Various interesting theorems related to design of SVM's (such as Mercer's theorem) were proved.

Also, Radial Basis Function Neural Networks(RBFNN's) are proposed in which the activation function at each neuron computes the distance between an input vector and centering vector at the neuron. The centering vectors correspond to centers of clusters of patterns.

This paper is organized as follows. In section II, various models of associative memory based on spherical separability are proposed.In section III, simulation results are presented. In section IV, recurrent laered neural networks are discussed. In setion V, we briefly discuss synthesis of associative memories based on spherical seperability. Conclusions are reported in section VI.

II. SPHERICAL SEPERABILITY :ASSOCIATIVE MEMORY ARCHITECTURES :

We first summarize relevant details related to the concept of spherical seperability of patterns. It was first introduced in [1].

Definition: Patterns belonging to two classes are(in Euclidean space) said to be spherically seperable if and only if there exists a "hypersphere" which seperates the patterns belonging to the two classes.

Note:Patterns belonging to M-classes are "spherically seperable" if and only if any pair of classes are spherically seperable.

Note: The distance metric can be more general than a Euclidean distance (e.g.Hamming distance).

Note: It can easily be proved/reasoned that patterns(which are in a bounded region of Euclidean space) that are linearly seperable are necessarily spherically seperable but not the other way.

Note: In 2-Dimensions, Spherical seperability is easy to visualize. The patterns are seperable by circles (surrounding the patterns). The ANNs proposed in this research paper are motivated by RBFNNs.

• Spherical Seperability based Associative Memories:
Consider an Artificial Neural Network (ANN) represented by an un-directed graph G = (V, E). The vertices correspond to neurons and the edge weights represent synaptic weights. Let the synaptic weight matrix be labeled as W. The neurons are in state {+1 or -1}, i.e. state $V_i(n)$ of i^{th} neuron at time 'n' is {+1 or -1} i.e. $V_i(n) \in${+1 or -1}. Thus, the state vector of such ANN at time 'n', i.e. $\bar{V}(n)$ lies on the symmetric unit hypercube. Furthermore each neuron is associated with a centering vector $\bar{U}_i$ lying on the symmetric unit hypercube.
Unlike, Hopfield neural network, the state updation of the ANN is based on the concept of spherical seperability. As in the case of Hopfield Associative Memory, our novel ANN also operates in the serial mode or fully parallel mode. Based on the state updation (in serial and parallel modes of operation), we propose five different ANN architectures. Let $\bar{T}$ be the threshold vector associated with the set of neurons(say 'M' of them). i.e. $\bar{T} = [t_1, t_2, t_3....t_M]^T$. The architectures are presented in the chronological order in which they were discovered. Later we relate them.

Architecture 1: Serial Mode:

$$V_i(n+1) = Sign\{d_H(\bar{V}(n), \bar{U}_i) - t_i\},$$

where $d_H(\bar{V}(n), \bar{U}_i)$ is the Hamming distance between {+1 or -1} vectors $\bar{V}(n), \bar{U}_i$. Thus at any time 'n+1', only state of one neuron is updated.
Fully parallel mode

$$\bar{V}(n+1) = \begin{bmatrix} Sign\{d_H(\bar{V}(n), \bar{U}_1) - t_1\} \\ Sign\{d_H(\bar{V}(n), \bar{U}_2) - t_2\} \\ \vdots \\ Sign\{d_H(\bar{V}(n), \bar{U}_M) - t_M\} \end{bmatrix}$$

i.e. at any time 'n+1', state of all the neuron is updated parallelly in the fully parallel mode of operation.

Architecture 2: Serial Mode:

$$V_i(n+1) = Sign\{d_E(\bar{V}(n), \bar{U}_i) - t_i\},$$

where $d_E(\bar{V}(n), \bar{U}_i)$is the Euclidean distance between {+1 , -1} vectors $\bar{V}(n), \bar{U}_i$ i.e. at any time 'n+1', only the sate of one of the neurons is updated.
Fully Parallel Mode i.e. all the components of Mx1 state vector are updated at the same time 'n+1'.

$$\bar{V}(n+1) = \begin{bmatrix} Sign\{d_E(\bar{V}(n), \bar{U}_1) - t_1\} \\ Sign\{d_E(\bar{V}(2), \bar{U}_2) - t_2\} \\ \vdots \\ Sign\{d_E(\bar{V}(n), \bar{U}_M) - t_M\} \end{bmatrix}$$

Note: In the above two architectures, the synaptic weight matrix is not utilized in state updation.

Architecture 3: Serial Mode:

$$V_i(n+1) = Sign\{d_H(\bar{W}\bar{V}(n), \bar{U}_i) - t_i\},$$

where $\bar{W}$is the synaptic weight matrix.

Note: Architecture 1 is a special case of this architecture with $\bar{W} = \bar{I}$.
Fully parallel mode:

$$\bar{V}(n+1) = \begin{bmatrix} Sign\{d_H(\bar{W}\bar{V}(n), \bar{U}_1) - t_1\} \\ Sign\{d_H(\bar{W}\bar{V}(n), \bar{U}_2) - t_2\} \\ \vdots \\ Sign\{d_H(\bar{W}\bar{V}(n), \bar{U}_M) - t_M\} \end{bmatrix}$$

Architecture 4: Serial Mode:

$$V_i(n+1) = Sign\{d_E(\bar{W}\bar{V}(n), \bar{U}_i) - t_i\},$$

Fully parallel mode:

$$\bar{V}(n+1) = \begin{bmatrix} Sign\{d_E(\bar{W}\bar{V}(n), \bar{U}_1) - t_1\} \\ Sign\{d_E(\bar{W}\bar{V}(n), \bar{U}_2) - t_2\} \\ \vdots \\ Sign\{d_E(\bar{W}\bar{V}(n), \bar{U}_M) - t_M\} \end{bmatrix}$$

Note:In all the above architectures of associative memory state updation can take place at some nodes but not all nodes. This updation corresponds to other parallel modes of operation.
We now propose another interesting architecture.

Architecture 5: Serial Mode:

$$V_i(n+1) = Sign\{d_M(\bar{V}(n), \bar{U}_i) - t_i\},$$

Where $d_M(\bar{(}V(n), \bar{U}_i)$ is the Mahalonibis distance between $\{+1, -1\}$ vectors $\bar{V}(n), \bar{U}$ i.e.

$$d_M(\bar{V}(n), \bar{U}_i) = (\bar{V}(n) - \bar{U}_i)^T S(\bar{V}(n) - \bar{U}_i),$$

with S being any positive definite matrix. Thus, with such updation in serial mode of operation, state of only one neuron is updated. As in the case of above architectures, state updation can be done in the fully parallel mode.

Note: Architecture 4 is a case related to architecture 5.
In all the above architectures, the initial condition i.e $\bar{V}(0)$ can be chosen in one of the following ways

i) $\bar{V}(0)$ is same as one of the centering vectors i.e.$\bar{U_i}'s$.
ii)$\bar{V}(0)$ is not equal to any of the centering vectors.

Also, unlike Hopfield neural network, since the distances are non negative, in the spherical speperability based associative memories, the components of threshold vector are chosen to be positive real numbers. Specifically in architectures based on Hamming distance, the thresholds are chosen to be positive integers.
• In all the five architectures, one can choose the centering vectors to be non-orthogonal corner's of hypercube and investigate the dynamics of associated artificial neural networks.
•Also, by choosing the initial conditions i.e. $\bar{V}(0)'s$ to be orthogonal corners of hypercube, the dynamics of proposed ANN's can be investigated.

The above associative memory architectures were naturally motivated by the definition of spherical seperability. We now prove that some of the architectures are equivalent to one another, under some conditions.

• Lemma 1: Architecture 2 is related to Architecture 1 from the point of view of dynamics:

Proof:By definition of Euclidean distance

$$d_E(\bar{V}(n), \bar{U}_i) = (\bar{V}(n) - \bar{U}_i)^T(\bar{V}(n) - \bar{U}_i)$$

where $\bar{V}(n), \bar{U}_i$ lie on the symmetric unit hypercube.

$$\begin{aligned} d_E(\bar{V}(n), \bar{U}_i) &= \bar{V}^T(n)V(n) - 2\bar{V}^T(n)\bar{U}_i + \bar{U}_i^T\bar{U}_i \\ &= N - 2 < \bar{V}(n), \bar{U}_i > + \bar{U}_i^T\bar{U}_i \end{aligned}$$

Now, since $\bar{V}(n), \bar{U}_i$ lie on the unit hypercube, their inner product i.e.
$< \bar{V}(n), \bar{U}_i >$ = (Number of places where $\bar{V}(n)$, $\bar{U}_i$ agree) - (Number of places $\bar{V}(n)$, $\bar{U}_i$ disagree).

Using the definition of Hamming distance

$$< \bar{V}(n), \bar{U}_i >= N - 2d_H(\bar{V}(n), \bar{U}_i).$$

Hence

$$\begin{aligned} d_E(\bar{V}(n), \bar{U}_i) &= N - 2[N - 2d_H(\bar{V}(n), \bar{U}_i)] + N \\ &= 4d_H(\bar{V}(n), \bar{U}_i) \end{aligned}$$

Hence the dynamics of architecture 1, architecture 2 are equivalent.

Note: If the threshold vector in architecture 2 is scaled by 4 of that in architecture 1, then their dynamics is equivalent .

Q.E.D.

• Lemma 2: Architecture 3, 4 are related from the point of view of dynamics of nonlinear system:

Proof: Now, we relate architecture 3, 4.

$$\begin{aligned} d_E(\bar{W}\bar{V}(n), \bar{U}_i) &= (\bar{W}\bar{V}(n) - \bar{U}_i)^T(\bar{W}\bar{V}(n) - \bar{U}_i) \\ &= \bar{V}^T(n)\bar{W}^TW\bar{V}(n) - 2(\bar{V}^T(n)W^T\bar{U}_i) + \bar{U}_i^T\bar{U}_i. \end{aligned}$$

It should be noted that for any non-singular symmetric matrix, W, W^TW is always positive definite. Suppose $\bar{W}^TW = I$ (i.e. W is an orthogonal matrix or more generally $W^TW = D$ a diagonal matrix). Then, we have that

$$\begin{aligned} d_E(\bar{W}\bar{V}(n), \bar{U}_i) &= (\bar{V}^T(n)\bar{V}(n) - 2 < \bar{W}\bar{V}(n), \bar{U}_i > + \bar{U}_i^T\bar{U}_i \\ &= N - 2 < \bar{W}\bar{V}(n), \bar{U}_i > + N \\ &2N - 2[N - 2d_H(\bar{W}\bar{V}(n), \bar{U}_i)] \\ &= 4d_H(\bar{W}\bar{V}(n), \bar{U}_i) \end{aligned}$$

Thus, the dynamics of architectures 3, 4 are related when W is an orthogonal symmetric matrix (or more generally $W^TW = D$, a diagonal matrix).

Q.E.D.

Note: If the threshold vector in architecture 4 is scaled by 4 of that in architecture 3 then their dynamics is same.

It readily follows that if W = I (an identity matrix), then the dynamics of all 5 architectures are related.

III. Simulation Results

Let the components of threshold vector always be positive numbers. In the case of Hamming distance they must be positive integers, but in the case of Euclidean distance, they can be any positive real number. W is a symmetric matrix.

Example1:
Architecture 1(H2 Matrix):

$$H_2 = i.e.Hadamard\ matrix\ \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{1}$$

and

$$\begin{aligned} V_0^T &= [-1, 1] = \bar{V}(0) \\ U_1^T &= [1, 1] \\ U_2^T &= [1, -1] \\ [t_1, t_2]^T &= [1, 1] \end{aligned}$$

$$V = [V(1), V(2), V(3), V(4)] = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{2}$$

Thus the dynamics leads to cycle of length 2.
Architecture 2(H2 Matrix):

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{3}$$

and

$$\begin{aligned} V_0^T &= [-1, 1] = \bar{V}(0) \\ U_1^T &= [1, 1] \\ U_2^T &= [1, -1] \\ [t_1, t_2]^T &= [1, 1] \end{aligned}$$

$$V = [V(1), V(2), V(3), V(4)] = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{4}$$

Thus, the dynamics leads to stable state (cycle of length 1).
Architecture 3(H2 Matrix):

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{5}$$

and

$$\begin{aligned} V_0^T &= [-1, 1] = \bar{V}(0) \\ U_1^T &= [1, 1] \\ U_2^T &= [1, -1] \\ [t_1, t_2]^T &= [1, 1] \end{aligned}$$

$$W = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \tag{6}$$

$$V = [V(1), V(2), V(3), V(4)] = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{7}$$

Thus dynamics leads to cycle of length 2.
Architecture 4(H2 Matrix):

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{8}$$

and

$$V_0^T = [-1, 1] = \bar{V}(0)$$

$$U_1^T = [1, 1]$$

$$U_2^T = [1, -1]$$

$$[t_1, t_2]^T = [1, 1]$$

$$W = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad (9)$$

$$V = [V(1), V(2), V(3), V(4)] = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (10)$$

Thus dynamics leads to stable state.

Example2:

Architecture1(H4 Matrix):

$$V_0^T = [1; -1; 1; -1] = \bar{V}(0)$$

$$U_1^T = [1; 1; 1; 1]$$

$$U_2^T = [1; -1; 1; -1]$$

$$U_3^T = [1; 1; -1 - 1]$$

$$U_4^T = [1; -1; -1; 1]$$

$$[t_1, t_2, t_3, t_4]^T = [1, 1, 1, 1]$$

$$V = [V(1), V(2), V(3), V(4)] = \begin{bmatrix} 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (11)$$

Thus dynamics leads to cycle of length 2.

Architecture 2(H4 Matrix):

$$V_0^T = [1; -1; 1; -1] = \bar{V}(0)$$

$$U_1^T = [1; 1; 1; 1]$$

$$U_2^T = [1; -1; 1; -1]$$

$$U_3^T = [1; 1; -1 - 1]$$

$$U_4^T = [1; -1; -1; 1]$$

$$[t_1, t_2, t_3, t_4]^T = [1, 1, 1, 1]$$

$$V = [V(1), V(2), V(3), V(4)] = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (12)$$

Thus dynamics leads to cycle of length 2.

Architecture3(H4 Matrix):

$$V_0^T = [1; -1; 1; -1] = \bar{V}(0)$$

$$U_1^T = [1; 1; 1; 1]$$

$$U_2^T = [1; -1; 1; -1]$$

$$U_3^T = [1; 1; -1 - 1]$$

$$U_4^T = [1; -1; -1; 1]$$

$$[t_1, t_2, t_3, t_4]^T = [1, 1, 1, 1]$$

$$W = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{bmatrix} \quad (13)$$

$$V = [V(1), V(2), V(3), V(4)] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (14)$$

Thus dynamics leads to stable state.

Architecture4(H4 Matrix):

$$V_0^T = [1; -1; 1; -1] = \bar{V}(0)$$

$$U_1^T = [1; 1; 1; 1]$$

$$U_2^T = [1; -1; 1; -1]$$

$$U_3^T = [1; 1; -1 - 1]$$

$$U_4^T = [1; -1; -1; 1]$$

$$[t_1, t_2, t_3, t_4]^T = [1, 1, 1, 1]$$

$$W = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{bmatrix} \quad (15)$$

$$V = [V(1), V(2), V(3), V(4)] = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (16)$$

Thus dynamics leads to cycle of length 2.

Example3:

Architecture 1 (H8 Matrix):

$$V_0^T = [1; -1; 1; -1; 1; -1; 1; -1]; = \bar{V}(0)$$

$$U_1^T = [1; 1; 1; 1; 1; 1; 1; 1];$$

$$U_2^T = [1; -1; 1; -1; 1; -1; 1; -1];$$

$$U_3^T = [1; 1; -1; -1; 1; 1; -1; -1];$$

$$U_4^T = [1; -1; -1; 1; 1; -1; -1; 1];$$

$$U_5^T = [1; 1; 1; 1; -1; -1; -1; -1];$$

$$U_6^T = [1; -1; 1; -1; 1; -1; 1; -1];$$

$$U_7^T = [1; 1; -1; -1; -1; -1; 1; 1];$$

$$U_8^T = [1; -1; -1; 1; -1; 1; 1; -1];$$

$$[t_1, t_2, t_3, t_4, t_5, t_6, t_8, t_8]^T = [1, 1, 1, 1, 1, 1, 1, 1];$$

$$V = \begin{bmatrix} 1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (17)$$

Thus, dynamics leads to cycle of length 2.
It is clear that all five architectures utilize some type of distance measures. In view of the simulation results, we expect the following observations to be true in the case of Hopfield Associative Memory(HAM/HNN).

• In serial mode of operation of Hopfield Neural Network (HNN/HAM), the Hamming distance between the successive corners of hypercube reached (starting in an initial condition) and the associated stable state is non-increasing or non-decreasing.

• If the initial conditions are orthogonal, the domains of attraction of associated stable states are mostly disjoint, if 'W' is synthesized using orthogonal corners of hypercube (that are eigenvaectors of 'W').

In the above examples, it is empirically shown that in the serial mode as well as fully parallel mode, starting with any initial condition, the dynamical systems (corresponding to all architectures of ANNs) converge to a stable state or a cycle of length 2, when the centering vectors at all the nodes constitute orthogonal corners of unit hypercube. Formally, we state the following convergence theorem.

Theorem: With orthogonal corners of hypercube as centering vectors at all neurons, the associative memories in all the above architectures converge to a stable state in serial mode. In the fully parallel mode atmost a cycle of length 2, is reached.

Proof: The dynamical systems in above ANN architectures converge to a fixed point starting with an initial condition by the Brauer's fixed pint in serial mode theorem. A direct proof follows from theorem 6 in [6].

IV. Recurrent Layered Neural Networks

Traditionally, in Artificial Neural Networks (ANN) literature, feedforward neural networks received lot of attention (due to many applications as in the case of Multi-Layer Perception). Also, researchers(Like Jordan, Elman) have proposed recurrent ANNs(like finite impulse, infinite impulse networks) and applied them to the inputs with temporal dependence. Also, recurrent networks such as Long Short Term Memory (LSTM) found many applications.

Hopfield proposed an associative memory where the associated graph architecture (capturing the connectivity structure) is fully connected(clique). The authors conceived the idea of placing the artificial neurons in multiple layers and impose constrains based on the interconnections between neurons placed in multiple layers (i.e. the synaptic weight matrix is highly structured). Some interesting recurrent layered neural networks are explicitly identified below.

I. Tail Biting Recurrent Layered Network: In this ANN architecture, the neurons are placed in layers and the network contains feedforward connections from input layer to the output layer. But the outputs of neurons in the output layer are fedback to the neurons in the input layer. Thus, the connectivity graph constitutes a tail biting trellis. With 2 neurons in each layer and 3 layers, the architecture of such ANN is illustrated in the following figure 1.

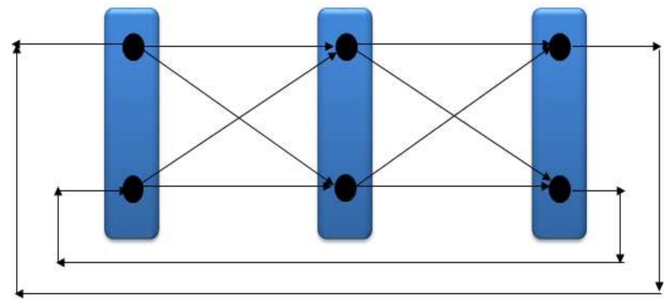

Fig. 1. Tail Biting Recurrent Layered Network

II. Nearest Neighbor Recurrent Layered Network: In this architecture, the neurons in a layer are connected to the neurons in immediate/nearest neighbor layers (i.e. connections are symmetric). As in architecture (I), the neurons outputs in the output layer are fedback to the inputs of neurons in the input layer. Effectively there are feedback connections from next layer to the neurons in current layer.

III. The architecture of recurrent layered neural network can be based on an architecture that corresponds to one of the many possible Structured Directed Graphs. Such recurrent ANNs proposed in this section potentially lead to interesting associative memories with good properties.

V. Synthesis of Spherical Seperability Based Associative Memories

We now formulate an interesting problem that deals with synthesizing an associative memory (based on spherical seperability) with desired stable states as in the case of Hopfield Associative memory. In [4],[5] the authors provided a satisfactory solution to the problem, in the case of Hopfield Associative memory.

Synthesis Problem: Given a set of desired corners of unit hypercube as memories, synthesize the matrix, W such that those memories are the fixed points/stable states. More generally, synthesize a dynamical system whose fixed points are desired stable states.

Unlike Hopfield Associative Memory(HAM), the corners of hypercube that are eigenvectors of W donot necessarily lead to fixed points in general. In that sense programming problem seems to be difficult for synthesis of associative memory based on spherical seperability. Apparently we expect the following observation to be true in the case of spherical seperability based associative memories.

• The number of spurious states is small.

• The time to converge to a fixed point starting in an initial condition is very small.

In the spirit of HAM, complex valued Hopfield Associative Memories (HAMs) are proposed by the authors [6], [7]. Such Complex Valued Neural Networks [CVNNs] based on spherical seperability are natural generalizations of architectures proposed in this research paper.

VI. Conclusion

In this research paper, 5 novel architectures for associative memories based on spherical seperability are discussed. The relationship between them is proved. It is reasoned that starting in an initial condition, the nonlinear dynamical systems converge to a stable state or a cycle of length atmost 2, starting in an intial condition. We expect the proposed associative memories to have better performance than those studied earlier by researchers.

References

[1] Garimella, Rama Murthy, Ganesh Yaparla, and Rhishi Pratap Singh. "Optimal Spherical Separability: Artificial Neural Networks." In International Work-Conference on Artificial Neural Networks, pp. 327-338. Springer, Cham, 2017.

[2] Ganesh, Yaparla, Rhishi Pratap Singh, and Garimella Rama Murthy. "Pattern classification using quadratic neuron: An experimental study." In 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-6. IEEE, 2017.

[3] Haykin, Simon S. "Neural networks and learning machines/Simon Haykin." (2009).

[4] Murthy, Garimella Rama, and Moncef Gabbouj. "On the design of Hopfield Neural Networks: Synthesis of hopfield type associative memories." In 2015 International Joint Conference on Neural Networks (IJCNN), pp. 1-8. IEEE, 2015.

[5] RamaMurthy, Garimella, Vamshi Devaki and Divya, "Synthesis or Programming of Hopfield Associative Memory," Proceedings of International Conference on Machine Learning and Data Science(ICMLDIS2019), December 2019, (ACM Digital Library).

[6] G.Ramamurthy and D.Praveen, Complex-valued Neural Associative Memories on the Complex Hypercube, proceedings of the IEEE conference on Cybernetics and Information Systems (CIS 2004), Singapore, Dec 2004.

[7] Munugoti, Sai Dileep, Anil Rayala, and Rama Murthy Garimella. "Novel Ceiling Neuronal Model Artificial Neural Networks." (2015).