



ANTI-SPURIOUS-STATE NEURAL NETWORK USING NONLINEAR OUTER PRODUCT AND DYNAMIC SYNAPSES

M. Xia*, M. Shen†, J. Wang‡, L. Weng*, C. Yan‡

Abstract: Associative memory (AM) is a very important part of the theory of neural networks. Although the Hebbian learning rule is always used to model the associative memory, it easily leads to spurious state because of the linear outer product method. In this work, nonlinear function constitution and dynamic synapses, against a spurious state for associative memory neural network are proposed. The model of the dynamic connection weight and the updating scheme of the states of neurons are presented. Nonlinear function constitution improves the conventional Hebbian learning rule to be a nonlinear outer product method. The simulation results show that both nonlinear function constitution and dynamic synapses can effectively enlarge the attractive basin. Comparing to the existing memory models, associative memory of neural network with nonlinear function constitution can both enlarge the attractive basin and increase the storage capacity. Owing to dynamic synapses, the attractive basin of the stored patterns is further enlarged, at the same time the attractive basin of the spurious state is diminished. But the storage capacity is decreased by using the dynamic synapses.

Key words: *associative memory, spurious state, nonlinear function constitution, dynamic synapse*

Received: April 24, 2012

DOI: 10.14311/NNW.2016.26.022

Revised and accepted: April 27, 2016

1. Introduction

The associative information processing, for instance the associative memory, plays an important role in many functions of brain [4, 18, 26]. And it can be used for some application problems, such as image processing [4, 22], data clustering [11], regression analysis [5], and control systems [27, 35]. The Hopfield networks are often used to carry out the AM. The Hopfield networks have stable steady states

*Min Xia – Corresponding author, Liguo Weng, Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing University of Information Science and Technology, Ningliu Road 219, Nanjing China, 210044, E-mail: xiamin@nuist.edu.cn

†MaoYang Shen, Jianfeng Wang, Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing University of Information Science and Technology

‡Chen Yan, College of Information Science and Technology, Donghua University

viz. equilibrium points [3, 6, 9, 19, 37]. The patterns can be stored in the network by setting of synaptic weight, but a big number of spurious states exist in the traditional Hopfield networks [20]. According to the Hebbian learning rule [36], the synaptic weight matrix is static, and therefore the neuron's local field is linear with the state of the neuron. There is no depression for the neuron's local field. Thus, the neuron's local field is linearly increased, which leads to the "diffused spurious memory".

In most of traditional neural network models [7, 15, 25, 31, 38–40], the synaptic weights are assumed to be 'static', or changing their weight values only on a slow time pace. Accordingly, once the weights of the synapses are for example determined by a covariance rule, they are fixed during the network evolution process. However, synaptic plasticity across many time scales from the order of days to the order of milliseconds [1, 2, 13, 16, 17, 23, 41], has been widely discovered. Short-time synaptic plasticity such as facilitation and depression may play an important role in neuronal response. In this sense, the synapses of neurons can be dynamic rather than static [21]. These findings affect the transmission properties of single neurons, as well as network function and behavior. The dynamics of the network with dynamic synapses has attracted much attention recently [10, 12, 24, 32, 42]. Most of the models of dynamic synapses are constructed within the context of spiking neurons and under the principle of synaptic transmission resources [1, 13, 29, 30]. In traditional works, synaptic weights cannot be changed after learning. If the stored patterns have large overlap, some synaptic weights will be very big, which leads to the spurious state. In order to reduce this influence, synaptic depression is introduced in this paper. Due to the synaptic depression, the synaptic weights will be depressed, especially for a big value of synaptic weights, which gives the network a driving force to get out of the spurious state.

In a contrast to many studies on binary-valued discrete time associative models, in this paper, the nonlinear function constitution, a new method against spurious state is proposed. This method improves the conventional Hebbian learning rule to become a nonlinear outer product method. Then we propose an anti-spurious-state neural network with static synapses (ASS-SSNN). The simulation results show that on one hand, our methods can effectively enlarge the attractive basin; on the other hand, associative memory neural network with nonlinear function constitution method can increase storage capacity. We further propose a novel anti-spurious-state neural network with dynamic synapses for auto-associative memory (ASS-DSNN) based on ASS-SSNN model. The experiments show that the ASS-DSNN model not only keeps the advantage of ASS-SSNN model, but also has more ample dynamic features than associative neural network without depressing synapses. The ASS-DSNN can further enlarge the attractive basin for memory state, reduce the attractive basin for spurious state, but the storage capacity is reduced.

The paper is organized as follows. In Section 2, the associative memory model with nonlinear function constitution and dynamic synapses is proposed. Section 3 presents the simulation results of the new model, and discusses the attractive basin and storage capacity of the proposed model. Finally, the Section 4 concludes this paper.

2. The network with nonlinear function constitution and dynamic synapses

The structure of the discrete Hopfield neural network is shown in Fig. 1, where N is the number of neurons in the network, I_i is the external input to neuron i , that is constant, and it is assumed to be zero. S_i is the state of neuron i . w_{ij} is the connection weight from neuron j to neuron i .

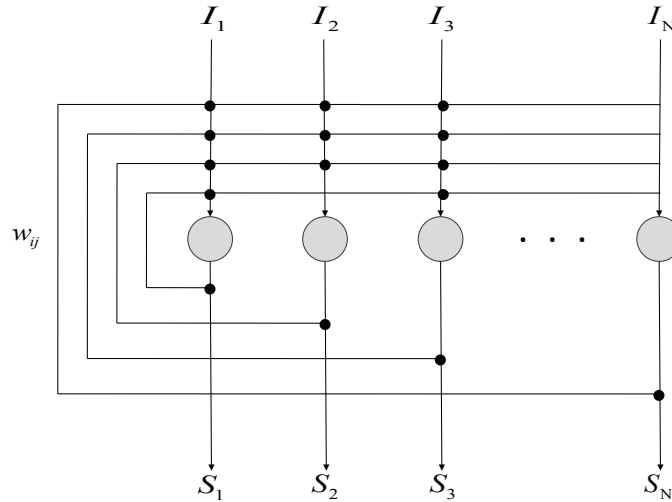


Fig. 1 The neural network construction.

2.1 The traditional associative memory model

The neuron in the network is characterized by binary state $\{s_i\}_{i=1}^N = \{0, 1\}$. The patterns $\{X^u = (x_1^u, x_2^u, \dots, x_N^u)\}$ are stored in the network, where $u = 1, 2, \dots, p$ denotes the label of the patterns. Each element $\{X^u\}$ of the u -th memory pattern is generated independently with $\text{prob}[x_i^u = 1] = 1 - \text{prob}[x_i^u = 0] = r$, where r is a coding level of the network. For the task of the recognition of p stored patterns, the synaptic coupling strength w_{ij} is usually given by the Hebbian learning rule. In such rule, the weights are described as

$$\begin{cases} w_{ij} = \frac{1}{p} \sum_{u=1}^p x_i^u x_j^u & \text{for } i \neq j, \\ w_{ii} = 0. \end{cases} \quad (1)$$

In the traditional associative memory model, the state of each neuron in the network is updated as follows

$$s_i(t+1) = f(h_i(t) - \theta_i) = f\left(\sum_{j=1, j \neq i}^N w_{ij} s_j(t) - \theta_i\right), \quad (2)$$

$$f(u_i(t)) = \begin{cases} 1 & u_i(t) \geq 0, \\ 0 & u_i(t) < 0, \end{cases} \quad (3)$$

where $s_i(t)$ is the output of neuron i at time t . The update is done synchronously, and the network dynamics continue until the convergence. θ_i is the threshold of the local field. $h_i(t)$ is the local field of neuron i ,

$$h_i(t) = \sum_{j=1, j \neq i}^N w_{ij} s_j(t). \quad (4)$$

The Eqs. (1)–(4) describe the traditional model of the associative memory (THNN). There are two conclusions about the Hebbian learning rule.

Conclusion 1: if the stored patterns X^k ($k = 1, 2, 3, \dots, p$) are orthogonal, thus $X^k \cdot X^l = 0$, then,

$$\begin{aligned} WX^k &= \frac{1}{p} \left[\sum_{l=1}^p X^l (X^l)^T \right] X^k = \frac{1}{p} \left[\sum_{l=1, l \neq k}^p X^l (X^l)^T X^k + X^k (X^k)^T X^k \right] \\ &= \frac{1}{p} X^k [(X^k)^T X^k] \approx \frac{r}{p} X^k, \end{aligned}$$

where r is a coding level of the network. The above equation indicates that each stored pattern X^k is the attractor of the network.

Conclusion 2: if the stored patterns X^k ($k = 1, 2, 3, \dots, p$) are not orthogonal, thus,

$$\begin{aligned} WX^k &= \frac{1}{p} \left[\sum_{l=1}^p X^l (X^l)^T \right] X^k = \frac{1}{p} \left[\sum_{l=1, l \neq k}^p X^l (X^l)^T X^k + X^k (X^k)^T X^k \right] \\ &\approx \frac{r}{p} X^k + \frac{1}{p} \sum_{l=1, l \neq k}^p X^l (X^l)^T X^k, \end{aligned}$$

where $\frac{r}{p} X^k$ is the useful signal for pattern recall, and $\frac{1}{p} \sum_{l=1, l \neq k}^p X^l (X^l)^T X^k$ is regarded as the noise for the memory. According to the Hebbian learning rule, the synaptic weight matrix is static, weights are constrained as they are normalized by the number of stored patterns, p , so that they are limited to the range 0 to 1. But $\frac{1}{p}$ is just a normalized term, which cannot change the memory result, synaptic weight $w_{ij} = \frac{1}{p} \sum_{l=1}^p x_i^l x_j^l$ is still linear increased by $x_i^l x_j^l$. According to the Hebbian learning rule, the neuron's field is linear with the output of the neuron. The proportional coefficient is determined by the static matrix of the synaptic weight, and the synaptic weight increases proportionally. If the element of noise vector is too large, and the corresponding vector element of X^k is 0, then the network cannot evolve to the attractor X^k , which results in a diffused spurious state.

Here, an example of the spurious state is presented. In Fig. 2, there are three memory patterns, which are denoted as P_1 , P_2 , and P_3 . These three patterns lap

over each other, and the overlap parts are marked as $P_a = P_1 \cap P_2$, $P_b = P_1 \cap P_3$, $P_c = P_3 \cap P_2$, and $P_d = P_1 \cap P_2 \cap P_3 = P_a \cap P_b \cap P_c$. We use the Pattern P_2 to investigate the emergence of a spurious state. Fig. 2 shows that the P_d is the overlap of three patterns. For the neurons of P_d , the connection weight is the highest. For the same reason, the connection weight of P_a , P_b , and P_c are higher than that of P_1 , P_2 , and P_3 . If we use the Hebb rule for learning, the local field is linearly increased to the output of other neurons, and the proportional coefficient is determined by static matrix of the synaptic weight, the synaptic weight increases proportionally without depression, which can easily lead to a “diffused spurious memory”. Thus, in the memory process for pattern P_2 , not only $P_a - P_d$, and $P_c - P_d$ can be retrieved, but the $P_b - P_d$ can frequently be retrieved as well. In this way, memory result of the network may be $P = P_2 + P_b - P_d$. Here, $P_b - P_d$ is the spurious state. The state P_2 is desirable, and we expect the spurious state is to be avoided.

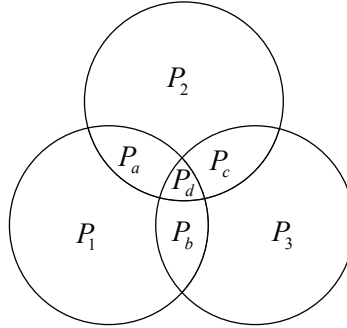


Fig. 2 Diffused spurious state.

2.2 Associative Memory with the nonlinear function constitution (ASS-SSNN)

For reduction of the diffused spurious state of the THNN, ASS-SSNN is proposed, which is implemented by a nonlinear function constitution. This method is a nonlinear outer product method. When using the nonlinear function constitution, the associative local field is described as

$$\hat{h}_i(t) = \sum_{j=1, j \neq i}^N \frac{1}{a} \left(\frac{1}{1 + e^{-b \cdot w_{ij} s_j(t)}} - 0.5 \right), \quad (5)$$

where a and b are scale parameters. If we assume that $s_j(t) = 1$ for $j = 1, 2, 3, \dots, N$, and w_{ij} is the variable, then the output result of nonlinear integration can be described in Fig. 3. The figure shows that the range of the output is determined by value a , the gradient is determined by value b . w_{ij} is the static synapse matrix, unchangeable within time. The local field of the neuron is nonlinear with the output of the neuron, avoiding the unrestrained increase of $\hat{h}_i(t)$.

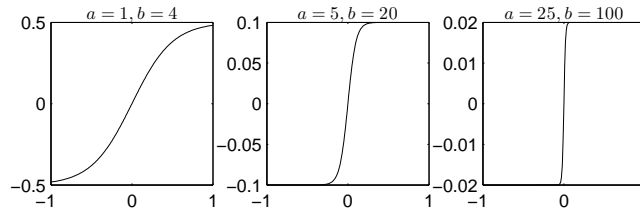


Fig. 3 Different values of a and b for nonlinear integration.

Due to the nonlinear integration, the state of neuron i at time t is updated as

$$s_i(t + 1) = f(\hat{h}_i(t) - \hat{\theta}_i) = f\left(\sum_{j=1, j \neq i}^N \frac{1}{a} \left(\frac{1}{1 + e^{-b \cdot w_{ij} s_j(t)}} - 0.5\right) - \hat{\theta}_i\right). \quad (6)$$

Here, $\hat{\theta}$ is the threshold of the local field. ASS-SSNN can be described by Eqs. (1), (3), (5) and (6). In ASS-SSNN the local field of a neuron is nonlinear with the output of the neuron, avoiding the unrestrained increase of $\hat{h}_i(t)$. Due to the nonlinear function constitution the number of spurious states will be reduced.

2.3 Associative Memory with nonlinear function constitution and dynamic synapses (ASS-DSNN)

Synapses in real neurons tend to exhaust their resources, i.e., their strength decreases when used. If we consider the dynamic synapses [13,36], the synaptic weight discussed above, w_{ij} evolves within time t as follows [30]

$$w_{ij}(t) = w_{ij} \cdot r_j(t), \quad (7)$$

where w_{ij} (see Eq. (1)) are static terms that are determined by the stored patterns. $r_j(t)$ is the term produced by the synaptic depression, and it is updated with following the dynamics [1, 21]

$$r_j(t + 1) = r_j(t) + \frac{1 - r_j(t)}{\tau} - \delta r_j(t) s_j(t), \quad (8)$$

where $r_j(t)$ denotes the fraction of synaptic resources available for a postsynaptic current, $s_j(t)$ is the output of presynaptic j -th neuron, τ and δ ($0 < \delta < 1$) are the parameters for synaptic depression. At every time step, a fraction of unavailable resources $\frac{1 - r_j(t)}{\tau}$ is recovered, while a fraction of available resources $\delta s_j(t) r_j(t)$ becomes unavailable due to the activity of the presynaptic neuron. The authors have proved that the fixed point of this model is the same as that of a network without synaptic depression [33]. Thus, the local field of the neuron i can be described as

$$\bar{h}_i(t) = \sum_{j=1, j \neq i}^N \frac{1}{a} \left(\frac{1}{1 + e^{-b \cdot w_{ij}(t) s_j(t)}} - 0.5\right). \quad (9)$$

Due to the dynamic synapses, the state of neuron i at time t is updated as

$$s_i(t+1) = f(\bar{h}_i(t) - \bar{\theta}_i) = f\left(\sum_{j=1, j \neq i}^N \frac{1}{a} \left(\frac{1}{1 + e^{-b \cdot w_{ij}(t) s_j(t)}} - 0.5\right) - \bar{\theta}_i\right). \quad (10)$$

The associative memory in a network with nonlinear function constitution and dynamic synapses (ASS-DSNN) can be given by Eqs. (1), (3), and (7–10).

The closeness between the state of the network $s(t)$ and the stored pattern x^u at time t is characterized by an overlap defined as

$$z^u(t) = \frac{1}{Nk(1-k)} \sum_{i=1}^N (x_i^u - k) s_i(t). \quad (11)$$

If the overlap $z^u(t)$ is close to 1, the stored pattern x^u is assumed to be retrieved at time step t .

3. Simulation results of the model

In order to evaluate the validity of the proposed model for anti-spurious-state, we must investigate the attractive basin¹ of different models with the same number of stored patterns. In this paper, the number of neurons in the network is assumed to be 2000, and the coding level r of the network is 0.1, which is the sparse coding. In order to get a high accuracy of the simulation results, each experiment is repeated for 100 times, then the average overlap of these results is computed. In this work, all stored patterns are used for simulations. Each stored pattern is simulated 100 times with added noise. For a single pattern, the added noise is randomly produced with same probability for each simulation. In the simulations for scale parameters and threshold, each of 100 simulations has a different starting point with 15% added noise (15% added noise means that the overlap between the initial state of the neural network and stored pattern is 85%).

First of all, the values of scale parameters a and b should be confirmed. In practice, when $a = 25$, $b = 100$, the network could provide a good performance, the simulations for scale parameters are presented in Tab. I. Indeed, if the factor a

Threshold θ	$a = 1, b = 4$	$a = 4, b = 20,$	$a = 25, b = 100$
	Overlap z^u		
0.25	0.9922	0.9939	1.0000
0.2	0.6017	0.6467	0.9994
0.15	0.5333	0.5539	0.8411
0.1	0.3972	0.4187	0.6050

Tab. I The scale parameters of ASS – DSNN ($p = 300$).

¹Attractive basin $[a,1]$ means that the neural network can recall the pattern when the overlap between the initial state of the neural network and stored pattern is from a to 1.

is just a scale parameter, it cannot affect the nonlinearity. But it can influence the range of a local field. The factor a can control the threshold less than 1. Fig. 4 shows how memory performance changes with variation in parameter b . The simulation shows that the network can get the best performance when $b = 100$. Thus, in this work, the values of a and b are set as 25 and 100 respectively. Then, the threshold θ of THNN and the threshold $\hat{\theta}$ of ASS-SSNN should be confirmed. Tab. II and Tab. III present the numerical results for THNN and ASS-SSNN respectively. From the simulation, the θ of THNN can be set as 0.42, the $\hat{\theta}$ of ASS-SSNN can be set as 0.26. When thresholds of 0.42 and 0.26 are reached for THNN and ASS-SSNN, the best performance can be reached.

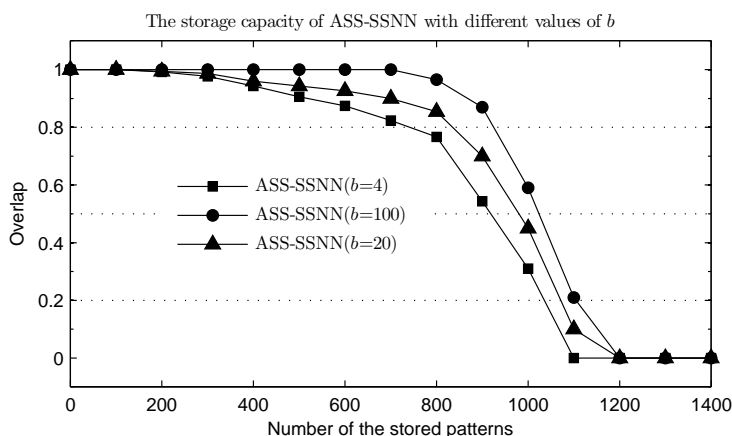


Fig. 4 Performance of the storage capacity for ASS-SSNN with $b = 4$, $b = 20$, and $b = 100$ (all simulations are with 15% initial noise).

The number of stored patterns p	300	400	500	600	700
Loading rate α	0.15	0.2	0.25	0.3	0.35
Threshold θ	Overlap z^u				
0.6	0.9850	0.9800	0.9350	0.8350	0.8050
0.5	0.9900	1.0000	0.9850	0.9900	0.9494
0.4	1.0000	1.0000	1.0000	0.9950	0.9828
0.3	1.0000	0.9978	0.9950	0.9283	0.8522
0.2	0.9111	0.6872	0.7178	0.5950	0.6389
0.1	0.5422	0.5483	0.6106	0.5500	0.5461

Tab. II The threshold of THNN.

In order to find how sensitive the results to parameters τ and δ are, the associative memory results of different values τ and δ are presented. In the simulation the initial state of the network are with noise of a stored pattern. The added noise is 40%, 50%, and 60% respectively. Fig. 5 shows that the network can provide good performance for different added noise when $\delta = 0.01$. Fig. 6 indicated that

The number of stored patterns p	300	500	700	800
Loading rate α	0.15	0.25	0.35	0.4
Threshold $\hat{\theta}$	Overlap z^u			
0.6	0	0	0	0
0.5	0.9950	0	0	0
0.4	1.0000	0.9600	0.8094	0.7200
0.3	1.0000	1.0000	0.9544	0.9583
0.2	1.0000	0.9983	0.9906	0.8761
0.1	0.7467	0.7033	0.6617	0.6289

Tab. III The threshold of ASS-SSNN.

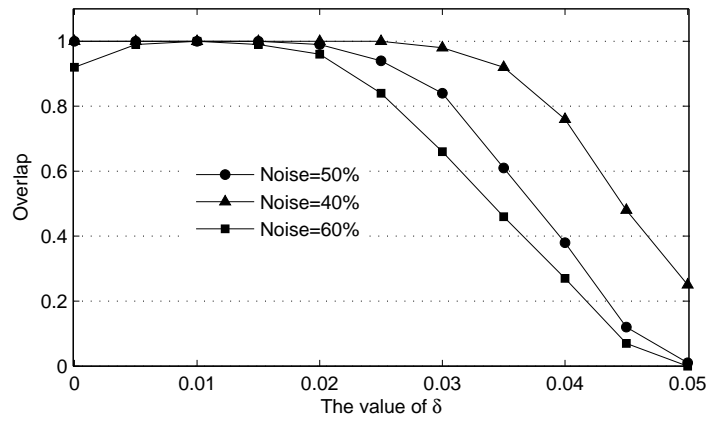


Fig. 5 Associative memory results of different values δ using different initial noise for ASS-DSNN ($p = 300$).

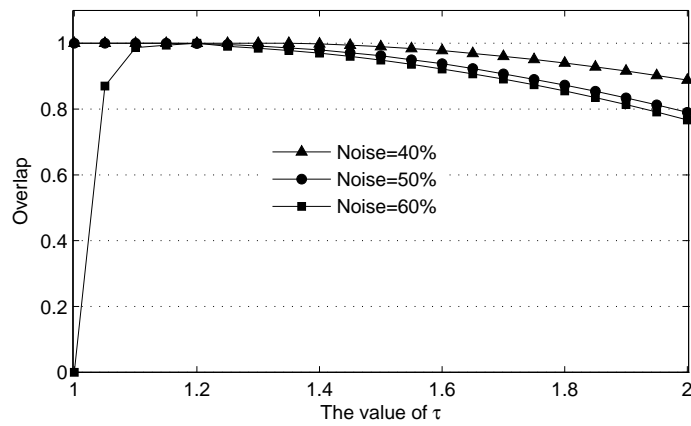


Fig. 6 Associative memory results of different values τ using different initial noise for ASS-DSNN ($p = 300$).

if only $\tau = 1.2$, the pattern can be recalled when the added noise is 60%. Thus, in this work, the values of parameters τ and δ are set as 1.2 and 0.01 respectively, and $r_j(0) = 1$. Then the threshold $\bar{\theta}$ of ASS-DSNN should be confirmed. Tab. IV presents the numerical results for ASS-DSNN. From the simulation, the $\bar{\theta}$ of ASS-DSNN can be set as 0.26, that the network can provide the best performance.

Stored patterns p	300	400	500	600	700	800
Loading rate α	0.15	0.2	0.25	0.3	0.35	0.4
Threshold $\bar{\theta}$	Overlap z^u					
0.6	0	0	0	0	0	0
0.5	0.9450	0.6000	0	0	0	0
0.4	1.0000	0.9850	0.9600	0.9600	0.6594	0
0.3	1.0000	0.9950	0.9900	1.0000	0.9944	0.9678
0.2	1.0000	0.9989	0.9722	0.9883	0.8611	0.9494
0.1	0.6978	0.7283	0.7067	0.6933	0.6622	0.6450

Tab. IV *The threshold of ASS-DSNN.*

Then we investigate how the nonlinear function constitution and the dynamic synapses influence the storage capacity. The loading rate α [20] ($\alpha = p/N$, where p is the number of stored patterns, N is the number of network nodes) is required to be less than the storage capacity α_c [20] ($\alpha_c = p_{\max}/N$, where p_{\max} is the maximum number of patterns the network can store, N is the number of network nodes) to ensure the successful retrieval of the pattern. In this work, if the memory recall error is less than 0.01, viz. the overlap is more than 0.99, the pattern is assumed to be memorized. From the simulation results in Tab. II, Tab. III and Tab. IV, the resolution of threshold are $\theta = 0.42$, $\bar{\theta} = \hat{\theta} = 0.26$ to arrive at the storage capacity in this work. Fig. 7 gives the performance of storage capacity for the THNN, ASS-SSNN and ASS-DSNN. Fig. 7 shows that the network with nonlinear function constitution has a bigger storage capacity than that of the THNN. But with the dynamic synapses the storage capacity in the model of ASS-DSNN is less than that of ASS-SSNN. This result indicates that the dynamic synapses decrease the storage capacity of network.

Tab. V presents the numerical results of storage capacity for THNN, ASS-SSNN and ASS-DSNN. The Storage capacity of the THNN is about 0.3375, while the storage capacity of ASS-SSNN is 0.377 and the storage capacity of ASS-DSNN

Different models	Maximal patterns stored	Threshold	Storage capacity
THNN	675	0.42	0.3375
ASS-SSNN	754	0.26	0.3770
ASS-DSNN	717	0.26	0.3585

Tab. V *The storage capacity of THNN, ASS-SSNN and ASS-DSNN ($N=2000$).*

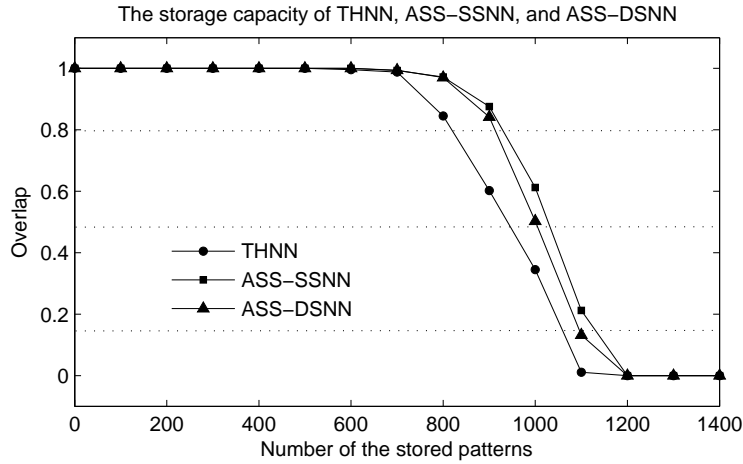


Fig. 7 Performance of the storage capacity for THNN, ASS-SSNN and ASS-DSNN (all simulations are with 15% initial noise).

is 0.3585. Due to the nonlinear function constitution, the storage capacity of the network is enlarged, but the storage capacity can be decreased by using dynamic synapses. Due to the characteristics of dynamic synapses, the network is easily to jump from one pattern to another, or stay at an intermediate state between different stored patterns, which is the reason why the storage capacity of ASS-DSNN is less than ASS-SSNN.

Fig. 8 presents the simulation results of attractive basin for THNN, ASS-SSNN and ASS-DSNN. In this simulation, the number of stored pattern should be set less than $\min(717, 754, 675)$ to ensure that the number of stored patterns is lower than the stored capacity. The result shows that the attractive basin of the THNN can

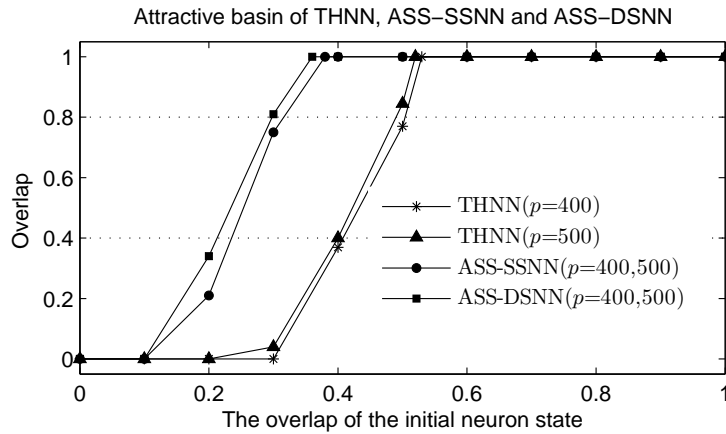


Fig. 8 Performance of attractive basin for THNN, ASS-SSNN and ASS-DSNN.

be influenced by p , and the attractive basin decreases when p increases. But the simulation implies that the number of stored patterns has no effect on the attractive basin of ASS-SSNN and ASS-DSNN. The simulation results in Fig. 8 also show that the nonlinear function constitution can effectively increase the ability of error tolerance i.e. to enlarge the attractive basin. The ASS-DSNN can further enlarge the attractive basin for memory state, reduce the attractive basin for spurious state, and has better ability of error tolerance.

Tab. VI gives the numerical results of attractive basin for the THNN, ASS-SSNN, ASS-DSNN. When the number of stored patterns is 400, the attractive basin of the THNN is $[0.53, 1]$, but the attractive basin of the THNN is $[0.52, 1]$ as $p = 500$. The attractive basin of the ASS-SSNN is $[0.38, 1]$, and that of ASS-DSNN is $[0.36, 1]$. This result indicates that the nonlinear function constitution and dynamic synapses can efficiently reduce the spurious state.

Model	Threshold	Critical noise	Attractive basin
THNN (p=400)	0.42	47%	$[0.53, 1]$
THNN (p=500)	0.42	48%	$[0.52, 1]$
ASS-SSNN (p=400,500)	0.26	62%	$[0.38, 1]$
ASS-DSNN (p=400,500)	0.26	64%	$[0.36, 1]$

Tab. VI *Attractive basin of THNN, ASS-SSNN and ASS-DSNN ($N = 2000$).*

4. Summary and discussion

In this work, we investigated the influence of nonlinear function constitution and dynamic synapses on the performance of the associative memory in the terms of attractive basin and storage capacity. In the traditional associative model, the local field of neuron is defined as

$$w_{ij} = \frac{1}{p} \sum_{u=1}^p x_i^u x_j^u.$$

Due to the Hebbian learning rule, the neuron's state is linear with the output of a neuron, there is no upper bound for the local field, which results in diffused spurious states. Due to the nonlinear function constitution for local field, the conventional Hebbian learning rule with linear outer product method can be improved, the neural network avoids the unrestrained increase in local field, which can prevent the network from evolving to the spurious states. That is why the nonlinear function constitution can effectively enlarge the attractive basin. Furthermore, the associative memory of neural network with nonlinear function constitution, the storage capacity can be increased remarkably comparing to the existing sequence memory models.

The electrophysiological experiments with animals have proven that synaptic weights are not "static" in real neurons. In natural neurons synapses tend to exhaust their resources, i.e. their strength decreases being used. In this work, the

dynamic synapse is introduced to reflect better the operation of a real neural system. In traditional work, synaptic weights cannot be changed after learning. If the stored patterns have a large overlap, some synaptic weights will be too big, which leads to a spurious state. In order to reduce this influence, synaptic depression is introduced in this paper. Thanks to the synaptic depression, the synaptic weights will be depressed, especially for a big value of synaptic weights, which provides the network with a driving force to get out of the spurious state. Thus, the attractive basin of the stored patterns is enlarged, but the storage capacity is decreased by using dynamic synapses.

In the existing literature, there are many methods how to improve the associative memory, such as Willshaw's clipped Hebbian learning [8], Koh's "coherent and collective firing" [14]. Clipped Hebbian learning is a good work for pattern recall, the strength of the synaptic connection from i -th neuron to j -th neuron can be written as

$$w_{ij} = \min \left(1, \sum_{u=1}^p x_i^u \cdot x_j^u \right) \in \{0, 1\},$$

which is different from the learning method in this work. In this work, the highest value of w_{ij} is very small because of the sparse coding. Thus, even if $b = 100$, $e^{-bw_{ij}}$ is still big enough to influence the last results. So, even if $b = 100$, or the value of b is larger than 100, our model is still different from the clipped Hebbian network of Willshaw. In Willshaw's work, the storage capacity is improved, and the performance is based on the measures of a unit usage. The model proposed in this paper is not of the unit use, but it is a good point for our future works. In a coherent and collective firing model, neurons fire collectively in a phase coherent manner, which may increase the storage capacity. In previous works, the neural networks were with pairwise neuron interactions. The coherent and collective firing gives a good suggestion for future works. In addition, the stability of the neural network is also an important research point of associative memory. It determines whether the associative memory model can be competent for some practical application problems. Some works have been done for this point so far [28, 34], which can improve the stability of memory.

Acknowledgement

This work is supported in part by the National Natural Science Foundation of PR China (grant No. 61105115, No. 61503192), Six Talent Peaks Program of Jiangsu Province (2014-XXRJ-007), Laboratory open project of NUIST, PRF-PSRA of Jiangsu Province (BY2015007-01) and Natural Science Foundation of Jiangsu Province (BK20161533, BK20131002).

References

- [1] ABBOTT L.F., VARELA J.A., SEN K. Synaptic Depression and Cortical Gain Control. *Science*. 1997, 275(5297), pp. 221–224, doi: 10.1126/science.275.5297.221.
- [2] ABBOTT L.F., REGEHR W.G. Synaptic computation. *Nature*. 2004, 431(7010), pp. 796–803, doi: 10.1038/nature03010.

- [3] AGHAJARI Z.H., TESHNEHLAB M., JAHED MOTLAGH M.R. A novel chaotic hetero-associative memory. *Neurocomputing*. 2015, 167(1), pp. 352–358, doi: 10.1016/j.neucom.2015.04.060.
- [4] AHMADIAN K., GAVRILOVA M. Transiently chaotic associative network for fingerprint image analysis. *Neural Network World*. 2010, 20(3), pp. 389–403.
- [5] AMIRI M., DAVANDE A., CHARTIER S. Feedback associative memory based on a new hybrid model of generalized regression and self-feedback neural networks. *Neural Networks*. 2010, 23(7), pp. 892–904, doi: 10.1016/j.neunet.2010.05.005.
- [6] DU K.L., SWAMY M.N.S. *Neural Networks and Statistical Learning*. New York: Springer, 2014, doi: 10.1007/978-1-4471-5571-3-7.
- [7] FENG W., YANG S.X., WU H. Improved robust stability criteria for bidirectional associative memory neural networks under parameter uncertainties. *Neural Computing and Applications*. 2014, 25(5), pp. 1205–1214, doi: 10.1007/s00521-014-1600-6.
- [8] GRAHAM B., WILLSHAW D. Capacity and information efficiency of the associative net. *Network Computation in Neural Systems*. 1997, 8(1), pp. 35–54, doi: 10.1088/0954-898X_8_1_005.
- [9] HEUSEL J., LOWE M., VERMET F. On the capacity of an associative memory model based on neural cliques. *Statistics and Probability Letters*. 2015, 106, pp. 256–261, doi: 10.1016/j.spl.2015.07.026.
- [10] JIAO X., WANG R. Nonlinear dynamic model and neural coding of neuronal network with the variable coupling strength in the presence of external stimuli. *Applied Physics Letters*. 2005, 87(8), pp. 083901, doi: 10.1063/1.1957120.
- [11] JOSE A., GOMEZ W. Automatic clustering using nature-inspired metaheuristics: A survey. *Applied Soft Computing*. 2016, 41, pp. 192–213, doi: 10.1016/j.asoc.2015.12.001.
- [12] KANAZAWA Y., ASAI T., IKEBE M. A novel CMOS circuit for depressing synapse and its application to contrast-invariant pattern classification and synchrony detection. *International Journal of Robotics and Automation*. 2004, 19(4), pp. 206–212, doi: 10.2316/Journal.206.2004.4.206-2716.
- [13] KISTLER W., HEMMEN J. Short-term Synaptic Plasticity and Network Behavior. *Neural Computation*. 1999, 11(7), pp. 1579–1594, doi: 10.1162/089976699300016151.
- [14] KOH Y., TAKATSUKA K. Increasing memory capacity and reducing spurious states in neural networks by introducing coherent and collective firing. *Neural Computation*. 2009, 21(5), pp. 1321–1334, doi: 10.1162/neco.2008.07-08-827.
- [15] LI L., YANG J., WU W. Intuitionistic fuzzy Hopfield neural network and its stability. *Neural Network World*. 2011, 21(5), pp. 461–472, doi: 10.14311/NNW.2011.21.027.
- [16] MAASS W., BISSHOP C.M. *Pulsed Neural Networks*. MIT Press. 2001.
- [17] MARKRAM H., TSODYKS M. Redistribution of Synaptic Efficacy between Neocortical Pyramidal Neurons. *Nature*. 1996, 382(6594), pp. 807–810, doi: 10.1038/382807a0.
- [18] NIE X., CAO J. Stability analysis for the generalized Cohen-Grossberg neural networks with inverse Lipschitz neuron activations. *Computers and Mathematics with Applications*. 2009, 57(9), pp. 1522–1536, doi: 10.1016/j.camwa.2009.01.003.
- [19] OHONEN T. Correlation matrix memories. *IEEE Transactions on Computers*. 1972, 100(4), pp. 353–359, doi: 10.1109/TC.1972.5008975.
- [20] OKADA M. Notions of associative memory and sparse coding. *Neural Networks*. 1996, 9(8), pp. 1429–1458, doi: 10.1016/S0893-6080(96)00044-5.
- [21] PANTIC L., TORRES J.J., KAPPEN H.J., GIELEN S.C.A.M. Associative memory with dynamic synapses. *Neural Computation*. 2002, 14(12), pp. 2903–2923, doi: 10.1162/089976602760805331.
- [22] QUIAN R. Neuronal codes for visual perception and memory. *Neuropsychologia*. 2016, 83, pp. 227–241, doi: 10.1016/j.neuropsychologia.2015.12.016.
- [23] SARGSYAN A.R., MELKONYAN A.A., PAPTAEODOROPOULOS C. A model synapse that incorporates properties of short- and long-term synaptic plasticity. *Neural Networks*. 2003, 16(8), pp. 1161–1177, doi: 10.1016/S0893-6080(03)00135-7.

- [24] SMITH L.S., DAGMAR S.F. Robust sound onset detection using integrate-and-fire neurons with depression synapses. *IEEE Transactions on Neural Networks*. 2004, 15(5), pp. 1125–1134, doi: 10.1109/TNN.2004.832831.
- [25] SMOESH K., MANU P.S. Pattern recall analysis of the Hopfield neural network with a genetic algorithm. *Computers and Mathematics with Applications*. 2010, 60(4), pp. 1049–1057, doi: 10.1016/j.camwa.2010.03.061.
- [26] TANG Y., QIU R., FANG J., MIAO Q., XIA M. Adaptive lag synchronization in unknown stochastic chaotic neural networks with discrete and distributed time-varying delays. *Physica Letters A*. 2008, 372(24), pp. 4425–4433, doi: 10.1016/j.physleta.2008.04.032.
- [27] TANG Y., GAO H., ZHANG W., KURTHS J. Leader-following consensus of a class of stochastic delayed multi-agent systems with partial mixed impulses. *Automatica*. 2015, 53(1), pp. 346–354, doi: 10.1016/j.automatica.2015.01.008.
- [28] TANG Y., XING X., KARIMI H., KOCATEV L., KURTHS J. Tracking control of networked multi-agent systems under new characterizations of impulses and its applications in robotic systems. *IEEE Transactions on Industrial Electronics*. 2016, 63(2), pp. 1299–1307, doi: 10.1109/TIE.2015.2453412.
- [29] TSODYKS M.V., PAWELZIK K., MARKRAM H. Neural networks with dynamic synapses. *Neural Computation*. 1998, 10(4), pp. 821–835, doi: 10.1162/089976698300017502.
- [30] TSODYKS M.V., MARKRAM H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences of the United States of America*. 1997, 94(2), pp. 719–723.
- [31] WANG X., LI C., HUANG T., DUAN S. A Weakly Connected Memristive Neural Network for Associative Memory. *Neural Processing Letters*. 2014, 40(3), pp. 275–288, doi: 10.1007/s11063-013-9328-3.
- [32] WANG R., JIAO X. Stochastic model and neural coding of large-scale neuronal population with variable coupling strength. *Neurocomputing*. 2006, 69(7), pp. 778–785, doi: 10.1016/j.neucom.2005.05.010.
- [33] WANG Z., FAN H. Dynamics of a continuous-valued discrete-time Hopfield neural network with synaptic depression. *Neurocomputing*. 2007, 71(1), pp. 181–190, doi: 10.1016/j.neucom.2007.01.004.
- [34] WANG F., LIU M. Global exponential stability of high-order bidirectional associative memory (BAM) neural networks with time delays in leakage terms. *Neurocomputing*. 2016, 177(12), pp. 515–528, doi: 10.1016/j.neucom.2015.11.052.
- [35] WENG L., XIA M., WANG W., LIU Q. Crew exploration vehicle (CEV) attitude control using a neural-immunology/memory network. *International Journal of Systems Science*. 2015, 46(1), pp. 152–158, doi: 10.1080/00207721.2013.775389.
- [36] WICKRAMASINGHE L.K., ALAHAKOON L.D., SMITH-MILES K.A. novel Episodic Associative Memory model for enhanced classification accuracy. *Pattern Recognition Letters*. 2007, 28(10), pp. 1193–1202, doi: 10.1016/j.patrec.2007.02.012.
- [37] XIA M., WENG L., WANG Z., FANG J. Sequence memory based on an oscillatory neural network. *Science China Information Science*. 2014, pp. 1–12, doi: 10.1007/s11432-013-4998-z.
- [38] XIA M., FANG J., WANG Y., WANG Z. Dynamic depression control of chaotic neural networks for associative memory. *Neurocomputing*. 2010, 73(4), pp. 776–783, doi: 10.1016/j.neucom.2009.10.015.
- [39] XIA M., FANG J., PAN F., BAI E. Robust sequence memory in sparsely-connected networks with controllable steady-state period. *Neurocomputing*. 2009, 72(13), pp. 3123–3130, doi: 10.1016/j.neucom.2009.03.004.
- [40] XIA M., WENG L., WANG Z. Sequence Memory based on Coherent Spin-interaction Neural Networks. *Neural Computation*. 2014, 26, pp. 2944–2961, doi: 10.1162/NECO_a_00663.
- [41] XIA M., ZHANG Y.C., YE X.L. Efficient Associative Memory Based on a Nonlinear Function Constitution and Dynamic Synapses. *Advanced Materials Research*. 2011, 225–226, pp. 479–482, doi: 10.4028/www.scientific.net/AMR.225-226.479.

- [42] XIA M., FANG J., TANG Y. Efficient multi-sequence memory with controllable steady-state period and high sequence storage capacity. *Neural Computing and Applications*. 2011, 20(1), pp. 17–24, doi: 10.1007/s00521-010-0453-x.