Statistical Properties and Memory Capacities of
Chaos Associative Memory
Masahiro NAKAGAWA
Nagaoka University of Technology,
Kamitomioika 1603-1, Nagaoka, Niigata 940-2188
E-mail: masanaka@vos.nagaokaut.ac.jp

Abstract-In this report we shall propose a chaos dynamic memory model applied to the chaotic autoassociation memory. The present artificial neuron model is properly characterized in terms of a time-dependent sinusoidal activation function to involve a transient chaotic dynamics as well as the energy steepest descent strategy. It is elucidated that the present neural network has a remarkable retrieval ability beyond the conventional models with such a monotonous activation function as sigmoidal one. This advantage is found to result from the property of the analogue periodic mapping accompanied with a chaotic behaviour of the neurons as well as the symmetry of the dynamic equation.

Keywords: chaos neuron, associative memory, sinusoidal mapping

1. INTRODUCTION

Tsuura reported some results concerned with dynamics retrieval model, dynamic linking of associative memories and explored the significance of chaos in those dynamics[1,2]. Nara et al. also argued the memory search model with a chaos control[3]. So far some applications of the chaotic neural networks have been investigated by Aihara et al.[4], Nakamura and Nakagawa[5]. In practice, however, as was confirmed by Kasahara and Nakagawa,[6] the chaotic dynamic association has been found to encounter the problem such that the complete association of the embedded patterns becomes inevitably troublesome if the loading rate, i.e. \( \alpha \) is increased beyond \(-0.2\) even though the orthogonal learning model with the generalised inverse matrix is concerned. Also the chaotic dynamics with a monotonous activation function was applied to such a combinatorial optimization problem as the Travelling Salesman Problem (TSP)[6]. They clarified that a parameter controlled chaos dynamics, which is called the transient chaos through the chaos simulated annealing (CSA), has a capability to realise a more efficient search of an optimal solution in TSP beyond the earlier work with fixed parameters. Thus chaotic behaviour in neural networks has been considered to play some important roles so as to find optimal solutions in the combinatorial optimization problems.

In contrast to the above-mentioned models with monotonous activation functions, the neurodynamics with a nonmonotonous mapping was recently proposed by Morita[7], Yanai and Amari[8]. They reported that the nonmonotonous mapping in a neuron dynamics possesses a certain advantage of the storage capacity, \( \alpha \approx 0.27 \), superior than the conventional association models with such a monotonous mapping as the signum function. This finding was explained as a result of an orthogonalisation process of the apparent synaptic weight matrix as a first approximation through the nonmonotonous dynamics[8]. That is, they insist that a nonmonotonous neuron dynamics involves in itself an approximate one-step orthogonalising process. Later Shiino and Fukai analysed the memory capacity for a somewhat simplified step-like nonmonotonous activation function with the continuous time, and concluded that the complete association could be realised up to the critical loading rate \( \alpha \approx 0.42[9] \). As a related nonmonotonous model, the present author proposed a novel neuron model with a periodic activation function to construct an association model with chaotic dynamics as the discrete time and orthogonal learning model[10-13]. Therein the storage capacity was found to be promoted up to \( \alpha \approx 0.35 \) even in the search mode without any key information beyond the previously proposed monotonous dynamic model with the discrete time as was above mentioned. Recently such a chaotic dynamics was involved in the synergetic neural network[14] to construct a chaos synergetic neural network model which involves a competition dynamics between overlaps[15,16]. In the single layer structure association model, however, the memory capacity has not been dramatically improved even by the chaotic neural networks[10-13,17].

In the present paper, let us investigate a chaos associative memory (CAM) model with the autocorrelation dynamics and the statistical property of the present chaos neuron model. [18-23]

2. THEORY

First of all let us define some dynamic rules to construct a chaotic neural network below. For this purpose we shall define first the internal state and the corresponding output of the i-th neuron as \( \sigma_i \) and \( s_i \), respectively, which have to be related to each other in terms of the following sinusoidal mapping,

\[
s_i = \sin \left( \frac{\pi \sigma_i}{2} \right).
\]

The energy function of the dynamical system with N neurons may be defined by

\[
E = E_w + E_c,
\]

where an objective function, \( E_w \), and a coupling energy function, \( E_c \), are defined by

\[
E_w = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij} s_i s_j,
\]

and

\[
E_c = \sum_{i=1}^{N} w_{ij} s_i s_j.
\]
\[ E_c = \sum_{i=1}^{N} \lambda_i \sigma_i, \]  

(4)

respectively; here \( w_{ij} \) (\( 1 \leq i, j \leq N \)) are the autocorrelation memory matrix components corresponding to the interconnection strengths between the \( i \)th and \( j \)th neurons and assumed to be symmetric, i.e. \( w_{ij} = w_{ji} \). \( \lambda_i \) (\( 1 \leq i \leq N \)) are the coupling constants between \( \sigma_i \) (internal state) and \( s_i \) (output), and \( \gamma_i \) (\( 1 \leq i \leq N \)) are the Lagrange multiplier of the \( i \)th neuron corresponding to the constraints (\( 1 \leq i \leq N \)) at a retrieval point. As in the conventional autocorrelation learning model with off-diagonal components, \( w_{ij} \) can be simply defined by

\[ w_{ij} = \frac{1}{N} \sum_{r=1}^{L} \left( e^{(r)}_i e^{(r)}_j - \delta_{ij} \right), \]

(5)

where \( L \) is the number of the embedded pattern vectors and \( e^{(r)}_i (r=\{1,2,\ldots,L\}) \) is the \( r \)th component for the \( i \)th embedded pattern, and all the embedded vectors are assumed to be linearly independent each other.

Then the overlaps \( o^{(r)}_i (r=\{1,2,\ldots,L\}) \), which are regarded as the pattern matching rates, are to be defined as follows,

\[ o^{(r)}_i = \frac{1}{N} \sum_{i=1}^{N} e^{(r)}_i \text{sign}(s_i) (1 \leq r \leq L). \]

(6)

Now let us define the dynamics of the present system below. The time-dependent Ginzburg-Landau (TDGL) equation of the internal state \( \sigma_i \) may be given by

\[ \frac{D\sigma_i}{Dt} = -\frac{\partial E}{\partial s_i}, \quad (1 \leq i \leq N) \]

(7)

where the operator, \( D/Dt \), may be regarded as the forward time difference operator (for the discrete-time model) or the time differential operator (for the continuous-time model). Substituting eq.(2) into eq.(8), one readily has

\[ \frac{D\sigma_i}{Dt} = \lambda_i \sigma_i + \sum_{j=1}^{N} w_{ij} s_j(t). \]

(8)

In the above dynamic equations, the coupling constant \( \lambda_i \) is concerned with the relaxation time of the \( i \)th neuron. If one resorts on the discrete-time model under consideration, the difference operator can be replaced as follows,

\[ \frac{D\sigma_i}{Dt} = \sigma_i(t+h) - \sigma_i(t). \]

(9)

where \( h \) is the time division interval for the \( t \)-axis.

Making use of eq.(9), our dynamic equation (8) reads

\[ \sigma_i(t+h) = (1-\lambda_i \cdot h) \sigma_i(t) + h \sum_{j=1}^{N} w_{ij} s_j(t) \]

(10)

where \( \sigma_i^*(t) \) is defined by

\[ \sigma_i^*(t) = \sum_{j=1}^{N} w_{ij} s_j(t). \]

(11)

It should be borne in mind here that \( s_i(t) \) and \( \sigma_i(t) \) have to be related each other in terms of eq.(1), and that \( \tau \) is assumed to be controlled towards \( 1 \) at a retrieval point, or a fixed point corresponding to a basin in the \( N \)-dimensional phase space spanned by \( s_i \) (\( 1 \leq i \leq N \)). The symmetry of eqs.(10) and (11) under \( s_i(t) \rightarrow -s_i(t) \) and \( \sigma_i(t) \rightarrow -\sigma_i(t) \) may be confirmed in Fig.1. This property may play a crucial role for the memory retrieval since an unknown component \( s_i(t) \) to be retrieved has equal probability for \( s_i(t) > 0 \) or \( s_i(t) < 0 \). In Figs.1(a) and (b), the bifurcation diagram and the Lyapunov exponent are given for \( N = 1, \lambda = 1, \gamma = 0, \) and \( \tau = 0 \), in which the updating rule can be derived as

\[ \sigma(t+1) = -\frac{\pi \sigma_i(t)}{2}. \]

(12)

For the nonlinear mapping as eq.(12), the Frobenius-Perron equation to determine the invariant measure \( p(\sigma) \) may be solved as

\[ p(\sigma) = c(\theta \delta(\sigma - \sin(\frac{\pi \theta}{2 \tau}) \right) \rho(\theta) \]

(13)

where the summation over \( n \) has to be restricted to \( \left| \sin^{-1}(\sigma - n \pi) \right| \leq 1/\tau \) \((-1 \leq \sigma \leq 1\)). An example of the solution of eq.(13) is depicted in Fig.2. According to the solution given by eq.(13), the Lyapunov exponent \( \lambda \) for \( \tau \ll 1 \) can be approximately derived as

\[ \lambda \sim \log\left( \frac{\pi}{2\tau \cdot \log 2} \right). \]

(14)
mode with eqs.(10) and (11). Hereafter \( h, \kappa, \lambda, i, \tau, 0 \) and \( \varepsilon \) are set to 0.5, 0.8, 1, \( 10^{-4} \), and \( 10^{-10} \), respectively, if not mentioned below[10-13]. The embedded pattern vectors were randomly selected from \( 2^{N} \) binary patterns. In the present associative model, let us investigate the dynamic memory retrieval characteristics with eqs.(10) and (11) in the autoassociation mode with a key input vector \( \theta \), i.e. the initial input vector for the autoassociation with the Hamming distance \( H_d \), which corresponds to the number of the components to be set into 0, from a target pattern \( e^{(s)}_i (1 \leq s \leq L) \). The directional cosine, \( \zeta_{(s)} \), of the initial input vector \( \theta = e^{(s)}_i (0) \) with respect to a target pattern \( e^{(s)}_i \) can be evaluated in terms of

\[
\zeta_{(s)} = \frac{1}{N} \sum_{i=1}^{N} e^{(s)}_i \theta_i = \frac{N - H_d}{N} = 1 - \frac{H_d}{N}
\]

Now choosing the parameters \( N \) as 100, the memory capacities are derived as in Figs.3 for (a) \( H_d/N = 0.01 \) and (b) \( H_d/N = 0.1 \). Thus we may confirm the memory retrieval up to \( \alpha \sim 1 \) for \( H_d/N = 0.01 \) which is remarkably larger than the corresponding value derived in the conventional models with the monotonous activation function as well as the nonmonotonous model proposed by Morita[7].

![Bifurcation diagram](image1.png)

(a) Bifurcation diagram

![Lyapunov exponent](image2.png)

(b) Lyapunov exponent

Fig.1 Bifurcation diagram and the Lyapunov exponent.

![Invariant measure](image3.png)

Fig.2 An example of the invariant measure (solid line: Analytical (eq.(13)), dots: Numerical). Here \( \tau = 10^{-2} \).

According to the above-noted idea, we simply assume the following linear dynamics for \( \tau (t) \).

\[
\frac{d \tau}{dt} = \kappa (1-\tau)
\]

(15)

where \( \kappa \) is a positive relaxation constant to drive the system from chaotic \( \tau (t) = 0 \) to nonchaotic \( \tau (t) = 1 \) state. Then \( \tau (t) \) has to be initialized by setting it to \( \tau = 0 (0 < \tau < 1) \), or a sufficiently small positive number so as to induce a chaotic dynamics[22] as follows,

if \( t = 0 \lor \sqrt{\| \sigma (t+h) - \sigma (t) \|} < \varepsilon \)

\( \tau (t) = \tau (t) \lor \theta (0 < \tau < 1) \),

(16)

where \( \lor \) is for the logical OR operation. At the same time, \( s \) and \( \sigma \) are to be set to \( \theta \) to start the association process.

3. RESULTS

We shall show a few examples of the dynamic behaviour of the present model in a chaotic associative mode with eqs.(10) and (11). Hereafter \( h, \kappa, \lambda, i, \tau, 0 \) and \( \varepsilon \) are set to 0.5, 0.8, 1, \( 10^{-4} \), and \( 10^{-10} \), respectively, if not mentioned below[10-13]. The embedded pattern vectors were randomly selected from \( 2^{N} \) binary patterns. In the present associative model, let us investigate the dynamic memory retrieval characteristics with eqs.(10) and (11) in the autoassociation mode with a key input vector \( \theta \), i.e. the initial input vector for the autoassociation with the Hamming distance \( H_d \), which corresponds to the number of the components to be set into 0, from a target pattern \( e^{(s)}_i (1 \leq s \leq L) \). The directional cosine, \( \zeta_{(s)} \), of the initial input vector \( \theta = e^{(s)}_i (0) \) with respect to a target pattern \( e^{(s)}_i \) can be evaluated in terms of

\[
\zeta_{(s)} = \frac{1}{N} \sum_{i=1}^{N} e^{(s)}_i \theta_i = \frac{N - H_d}{N} = 1 - \frac{H_d}{N}
\]

Now choosing the parameters \( N \) as 100, the memory capacities are derived as in Figs.3 for (a) \( H_d/N = 0.01 \) and (b) \( H_d/N = 0.1 \). Thus we may confirm the memory retrieval up to \( \alpha \sim 1 \) for \( H_d/N = 0.01 \) which is remarkably larger than the corresponding value derived in the conventional models with the monotonous activation function as well as the nonmonotonous model proposed by Morita[7].

![Autoassociation characteristics](image4.png)

(a) \( H_d/N = 0.01 \)

![Autoassociation characteristics](image5.png)

(b) \( H_d/N = 0.1 \)

Fig.3 The autoassociation characteristics with \( s = \sin \left( \frac{\pi \sigma_i}{2 \tau} \right) \).

To see the statistical property of the present model, we shall evaluate the output distributions of the output and the error of the internal state defined as

\[
p (s) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \delta(s - s^{(t)})
\]

(18)
\[ p(N_e) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \delta(N_e - (\sigma_i(t) - e^{(r)}_i)) \]  
\[ (19) \]

where \( e^{(r)}_i \) is a target pattern such that \( \sum e^{(r)}_i \neq 0 \). The output and the error distributions are given in Figs. 4 and 5, respectively. Figure 4 may resemble of Fig. 2 derived from eq. (13).

Then, from Fig. 5, one may confirm that the error has a tendency to be concentrated around \( N_e = 0 \) even for a relatively large loading rate \( L/N \).

![Fig. 4 The output distribution defined by eq. (18) with \( s_i(t) = \sin \left( \frac{\pi}{2} \sigma_i(t) \right) \). Here \( H_{d/N} = 0.01 \).](image)

![Fig. 5 The cross-noise distribution defined by eq. (19) with \( s_i(t) = \sin \left( \frac{\pi}{2} \sigma_i(t) \right) \). Here \( H_{d/N} = 0.01 \).](image)

The probability of the success of the output of \( s_i \) such that \( \text{sign}(s_i) = e^{(r)}_i \), where \( e^{(r)}_i \) is a component of the target pattern to be retrieved, can be evaluated as shown in Fig. 6. Therein one may see that does not drastically decrease with the increase of the loading rate \( L/N \). This may be regarded as the reason why the success rate does not critically depend on the loading rate \( L/N \) as shown in Fig. 3.

![Fig. 6 The dependence of the probability of the success defined by eq. (21) with \( H_{d/N} = 0.01 \). Here \( H_{d/N} = 0.01 \).](image)

4. CONCLUDING REMARKS

In this paper the dynamic memory retrieval characteristics of such a periodic chaos neural network with periodicity control has been found to be considerably improved in comparison with the conventional chaos neural networks with such a monotonous mapping as a sigmoidal function [4,5]. Figure 3(a) shows that the association can be realised up to the loading rate \( \sim 1 \) with beyond the previous finding derived from the partial reverse dynamics with the discrete time proposed by Morita et al [7], in which \( \sim 0.27 \) at most even for the \( H_{d/N} \to 0 \). It may be also concluded that the present advantage of our model results from the compatibility between chaotic dynamics and the symmetry of the bifurcation characteristics, which can not be realized in the monotonous chaos neuron model proposed by Aihara et al [4]. From the error distribution of the present model, we may again confirm the advantage of our model beyond the conventional association models.

References

[6] M. Morita et al [7], in which \( \sim 0.27 \) at most even for the \( H_{d/N} \to 0 \).