Statistical Properties and Memory Capacities of Chaos Associative Memory

Masahiro NAKAGAWA
Nagaoka University of Technology,
Kamitomioka 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

Tsuda 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. α 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, α_c~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 α_c~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 α_c~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 σ_i and s_i, respectively, which have to be related to each other in terms of the following sinusoidal mapping,

$$s_i = f_i(\sigma_i) = \sin\left(\frac{\pi}{2}\frac{\sigma_i}{\tau}\right). \qquad (1)$$

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

$$E = E_w + E_c, \qquad (2)$$

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

$$E_w = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} s_i s_j, \qquad (3)$$

and

$$E_c = \sum_{i=1}^{N} \lambda_i \int ds_i \, \sigma_i \quad , \tag{4}$$

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

$$w_{ij} = \frac{1}{N} \sum_{r=1}^{L} \left(e^{(r)}_i e^{(r)}_j - \delta_{ij} \right) \quad , \tag{5}$$

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

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

$$o^{(r)} = \frac{1}{N} \sum_{i=1}^{N} e^{(r)}_i \, \mathrm{sign}\left(s_i\right) \quad (1 \le r \le L) \; . \tag{6}$$

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

$$\frac{D\sigma_i}{Dt} = - \frac{\partial E}{\partial s_i} \; , \quad (1 \le i \le N) \tag{7}$$

where the operator, $D\bullet/Dt$, 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

$$\frac{D\sigma_i(t)}{Dt} = -\lambda_i \sigma_i(t) + \sum_{j=1}^{N} w_{ij} s_j(t) \quad . \tag{8}$$

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

$$\frac{D\sigma_i(t)}{Dt} = \frac{\sigma_i(t+h) - \sigma_i(t)}{h} \quad , \tag{9}$$

where h is the time division interval for the t -axis. Making use of eq.(9), our dynamic equation (8) reads

$$\sigma_i(t+h) = \left(1-\lambda_i h\right)\sigma_i(t) + h \sum_{j=1}^{N} w_{ij} s_j(t) = \left(1-\lambda_i h\right)\sigma_i(t) + h\,\sigma^*_i(t) \quad , \tag{10}$$

where $\sigma^*_i(t)$ is defined by

$$\sigma^*_i(t) = \sum_{j=1}^{N} w_{ij} s_j(t) \quad . \tag{11}$$

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 τ 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 ($1 \le i \le N$). The symmetry of eqs.(10) and (11) under $s_i(t) \to -s_i(t)$ and $\sigma_i(t) \to -\sigma_i(t)$ 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 $N=1$, $\lambda_i = h = 1$, and $\gamma_i = 0$, in which the updating rule can be derived as

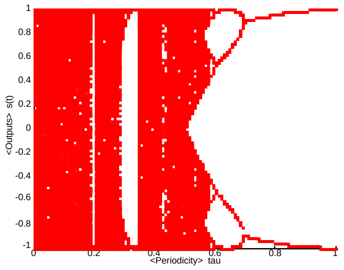
$$\sigma(t+1) = \sin\left(\frac{\pi}{2}\frac{\sigma(t)}{\tau}\right) = s(t) \quad . \tag{12}$$

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

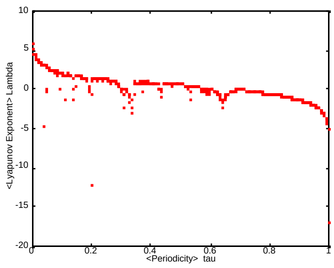
$$p\left(\sigma\right) = \int_{-1}^{1} d\theta \, \delta\left(\sigma - \sin\left(\frac{\pi}{2}\frac{\theta}{\tau}\right)\right) p\left(\theta\right) = \frac{2\tau}{\pi} \frac{1}{\sqrt{1-\sigma^2}} \sum_n p\left(\tau \frac{2}{\pi}\left((-1)^n \sin^{-1}\sigma + n\pi\right)\right) \tag{13}$$

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

$$\lambda \sim \log\left(\frac{\pi}{2\tau}\right) - \log 2 \quad . \tag{14}$$



(a) Bifurcation diagram



(b) Lyapunov exponent

Fig.1 Bifurcation diagram and the Lyapunov exponent.

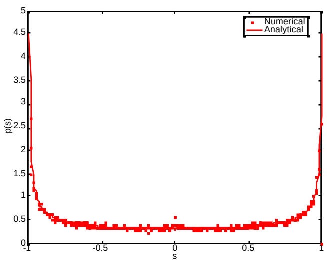


Fig.2 An example of the invariant measure (solid line: Analytical (eq.(13)), dots: Numerical) . Here $\tau = 10^{-2}$.

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

$$\frac{D\tau}{Dt} = \kappa\,(1-\tau)\quad, \qquad (15)$$

where κ 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 τ_0 (<<1), or a sufficiently small positive number so as to induce a chaotic dynamics[22] as follows,

$$\text{if } t = 0 \vee \frac{\|\sigma(t+h) - \sigma(t)\|}{h} < \varepsilon \longrightarrow \tau(t) = \tau_0 \ (0 < \tau_0 << 1) \quad , \qquad (16)$$

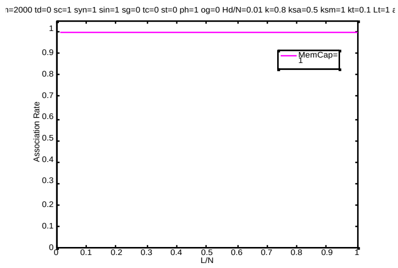
where $\vee$ is for the logical OR operation. At the same time, s_i and σ_i are to be set to θ_i to start the association process.

3. RESULTS

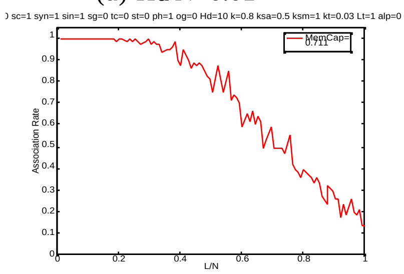
We shall show a few examples of the dynamic behaviour of the present model in a chaotic associative mode with eqs.(10) and (11). Hereafter h, κ, λ_i, τ_0 and ε are set to 0.5, 0.8, 1, 10^{-4}, and 10^{-10}, respectively, if not mentioned below[10-13]. The embedded pattern vectors were randomly selected 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 θ_i, 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 $e_i^{(s)}$ (1≤s≤L). The directional cosine, $\zeta^{(s)}$, of the initial input vector $\theta_i = s_i(0)$ with respect to a target pattern $e_i^{(s)}$ can be evaluated in terms of

$$\zeta^{(s)} = \frac{1}{N}\sum_{i=1}^{N} e_i^{(s)}\theta_i = \frac{N - H_d}{N} = 1 - \frac{H_d}{N} \qquad (17)$$

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 α_c~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].



(a) Hd/N=0.01



(b) Hd/N=0.1

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

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

$$p(s) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\delta(s - s_i(t)) \quad . \qquad (18)$$

$$p(N_e) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \delta\left(N_e - \left(\sigma_i(t) - e_i^{(r)}\right)\right). \qquad (19)$$

where $e_i^{(r)}$ is a target pattern such that $\Sigma e_i^{(r)} \theta_i = N - H_d$. The output and the error distributions are given in Figs.4 and 5, respectively. Figure 4 may resembles of Fig.2 derived from eq.(13). Then, from Fig.5, one may confirm that the error has a tendency to be concentrated around $N_e \sim 0$ even for a relatively large loading rate L/N.

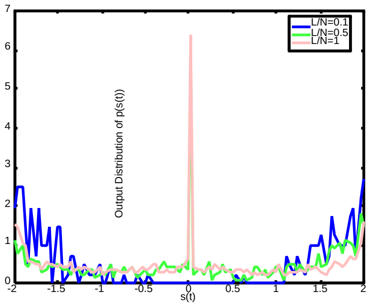


Fig.4 The output distribution defined by eq.(18) with $s_i(t) = \sin\left(\frac{\pi}{2} \frac{\sigma_i(t)}{\tau(t)}\right)$. Here $H_d/N = 0.01$.

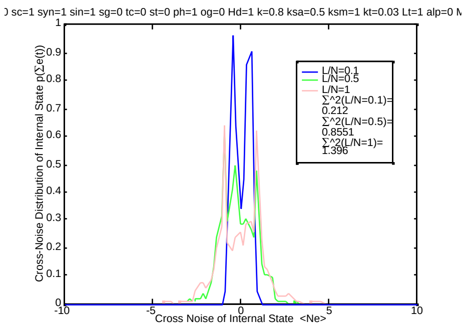


Fig.5 The cross-noise distribution defined by eq.(19) with $s_i(t) = \sin\left(\frac{\pi}{2} \frac{\sigma_i(t)}{\tau(t)}\right)$. Here $H_d/N = 0.01$.

The probability of the success of the output of s_i such that $\mathrm{sign}(s_i) = e_i^{(r)}$, where $e_i^{(r)}$ is a component of the target pattern to be retrieved, $\mathrm{Prob}\{\mathrm{sign}(s_i) = e_i^{(r)}\}$ defined by

$$\mathrm{Prob}\{\mathrm{sign}(s_i) = e_i^{(r)}\} = \int_{\mathrm{sign}(s_i) = e_i^{(r)}} dN_e\, p(N_e) \qquad (20)$$

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.

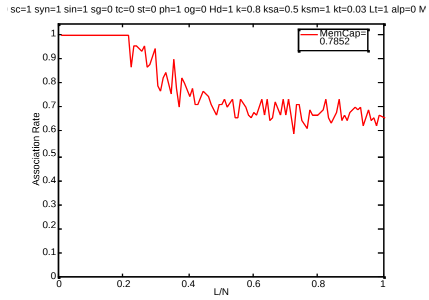


Fig.6 The dependence of the probability of the success defined by eq.(21) with on the loading rate L /N. Here $H_d/N = 0.01$.

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 ~1 with beyond the previous finding derived from the partial reverse dynamics with the discrete time proposed by Morita et al [7], in which ~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

[1] I. Tsuda: Neurocomputers and Attention I.(eds. A.V. Holden and V.I.Kryukou, Manchester Univ. Press, 1991)405.
[2] I. Tsuda: Neural Networks, 5(1992)313.
[3] S. Nara, P. Davis and H. Totsuji: Neural Networks, 6(1993)963.
[4] K. Aihara, T. Takabe and M. Toyoda: Phys. Lett. A144 (1990)333.
[5] K. Nakamura and M.Nakagawa: J. Phys. Soc. of Jpn.62 (1993)2942.
[6] T. Kasahara and M. Nakagawa: Electronics and Communications in Japan Part III-Fundamentals,78(1995)1.
[7] M. Morita : Neural Networks, 6(1993)115.
[8] Hiro-F. Yanai and S. Amari: Proc. of ICNN'93, San Francisco,(1993)1385.
[9] M. Shiino and T. Fukai: Phys. Rev. E48(1993)867.
[10] M. Nakagawa: Proc.of ICONIP'94, Seoul, 1(1994) 609.
[11] M. Nakagawa: Proc. of ICDC'94,Tokyo, 2(1995)603.
[12] M. Nakagawa: J. Phys.Soc. Jpn.64(1995)1023.
[13]T. Kasahara and M. Nakagawa: J. Phys. Soc. Jpn.64(1995)4964.
[14]M. Nakagawa: IEICE Trans. on Fundamentals E78-A(1995)412.
[15] M. Nakagawa: J. Phys.Soc.Jpn.64(1995)3112.
[16]M. Nakagawa: Proc. of ICNN'95, Australia(1995)3028.
[17] M. Nakagawa: J. Phys. Soc. Jpn.66(1997)263.
[18]M. Nakagawa: J. Phys. Soc. Jpn. 65(1996)1859.
[19]M. Nakagawa: Proc. of ICNN'96, 2(1996)862.
[20]T. Tanaka and M. Nakagawa: IEICE Trans. on Fundamentals J79-A(1996, in Japanese)1826.
[21] M. Nakagawa: J. Phys. Soc. Jpn.68(1999)2457.
[22] M.Nakagawa: Chaos and Fractals in Engineering (World Scientific,1999).
[23] M.Nakagawa: J. Phys.Soc.Jpn. **71**(2002)2316.