PAPER *Special Section on Selected Papers from the 16th Workshop on Circuits and Systems in Karuizawa*

Markov Chain Modeling of Intermittency Chaos and Its Application to Hopfield NN

Yoko UWATE[†a)], ***Student Member***, **Yoshifumi NISHIO**[†b)], ***Member***, ***and*** **Akio USHIDA**[††c)], ***Fellow***

SUMMARY In this study, a modeling method of the intermittency chaos using the Markov chain is proposed. The performances of the intermittency chaos and the Markov chain model are investigated when they are injected to the Hopfield Neural Network for a quadratic assignment problem or an associative memory. Computer simulated results show that the proposed modeling is good enough to gain similar performance of the intermittency chaos.

key words: *intermittency chaos, burst noise, Markov chain, neural network, QAP, associative memory*

1. Introduction

Intermittency chaos [1] is deeply related to *the edge of chaos* [2] and many researchers suggest that such a behavior between order and chaos gains better performance for various kinds of information processing than fully developed chaos. One good example of this is an application of chaos to the Hopfield Neural Networks (Hopfield NN) [3] solving combinatorial optimization problems to avoid trappings of the solutions into a local minimum. Hayakawa et al. pointed out the chaos near the three-periodic window of the logistic map gains the best performance for solving traveling salesman problems (TSP) [4]. However, the reason why the intermittency chaos exhibits such a good performance has not been clarified. Therefore, it is very important to make simpler models of the good characteristics of the intermittency chaos and to investigate their detailed properties.

In this study, we propose a modeling method of the intermittency chaos obtained from the logistic map by using the Markov chain. Various people have already proposed the Markov chain modelings of chaotic systems [5]–[7]. The modelings have been successfully applied to the chaos-based spread spectrum communication systems for the purposes of the noise cleaning of chaotic sequences [5] and the analytical estimation of the performance [6]. Further, the modeling has been extended to generate more complex nonlinear phenomena such as self-similarity [7]. These modelings are effective in the sense that the models could generate almost all the phenomena observed from the original chaotic system. However, it is not appropriate to reveal the reasons of the good performance of the intermittency chaos. Therefore, in this study, we pay our attentions only on the distribution of the lengths of the laminar parts and the burst parts, which seems to be the most distinguished feature of the intermittency chaos. The proposed modeling using the Markov chain is completely different from those in the references on the point that each state in the Markov chain is not a quantized value (or an interval) of a variable but a behavior of successive orbits. As a result, the model becomes very simple and enhances the feature of the intermittency chaos.

In order to confirm that the proposed model keeps the good property of the intermittency chaos, we investigate the performances of the intermittency chaos and the Markov chain model when they are injected to the Hopfield NN for a quadratic assignment problem (QAP) or an associative memory, which are known as representative applications of the Hopfield NN. Computer simulated results show that the proposed modeling is good enough to gain similar performance of the intermittency chaos.

2. Intermittency Chaos

We consider the logistic map to generate chaotic time series;

$$\hat{z}(t+1) = \alpha\hat{z}(t)(1-\hat{z}(t)). \tag{1}$$

Varying the control parameter α, Eq. (1) behaves chaotically via a period-doubling cascade. Further, it is well known that

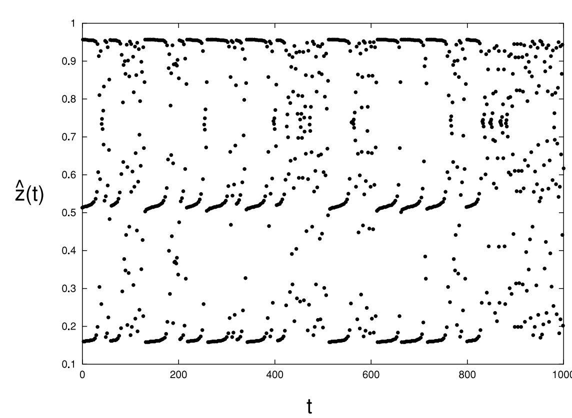


Fig. 1 Intermittency chaos for α=3.827940.

Manuscript received June 30, 2003.
Manuscript revised September 16, 2003.
Final manuscript received November 10, 2003.

[†]The authors are with the Department of Electrical and Electronic Engineering, Tokushima University, Tokushima-shi, 770-8506 Japan.
[††]The author is with the Department of Mechanical-Electronic Engineering, Tokushima Bunri University, Sanuki-shi, 769-2193 Japan.
a) E-mail: uwate@ee.tokushima-u.ac.jp
b) E-mail: nishio@ee.tokushima-u.ac.jp
c) E-mail: ushida@fe.bunri-u.ac.jp

the map produces intermittent bursts just before periodic-windows appear. Figure 1 shows an example of the intermittency chaos near the three-periodic window. As we can see from the figure, the chaotic time series could be divided into two phases; laminar part of periodic behavior with period 3 and burst part of spread points over the invariant interval. As increasing α, the ratio of the laminar parts becomes larger and finally the three-periodic window appears.

3. Markov Chain Modeling

In this section, we model the intermittency chaos by using the Markov chain.

At first, we distinguish the laminar part and the burst part of the intermittency chaos. Because we treat only the intermittency chaos near the three-periodic window, we regard three successive sequences starting from a point whose value is 0.9444 or more as one-period of the laminar part. Other points are regarded as the burst part.

In order to make the Markov chain model precisely, we counted the period of the laminar parts. The frequency of each period of the laminar part during 100000 iterations of the logistic map is shown in Fig. 2. We can see that the curve does not obey any simple scaling rules. Namely, the period of the laminar part is bounded and the maximum value of the period takes a peak. The intermittency chaos occurs just before a tangent bifurcation. The movement of the laminar part of the intermittency chaos is shown in Fig. 4 of [1]. If the control parameter of a given map is fixed, the width of the channel between the map and the bisector is decided. Since the width is finite, the number of the iteration to pass through the channel is bounded. We consider that this is the most distinguished feature of the intermittency chaos.

In order to model the above-mentioned feature of the intermittency chaos, we propose the Markov chain as shown in Fig. 3. In this Markov chain, the state S_0 corresponds to the burst part and the states $S_1, S_2, \cdots, S_L$ correspond to the laminar parts where the subscript k of S_k indicates the period of the continuing laminar part and L is the maximum period of the laminar part. In the state S_0, three points whose values are uniformly spread over the interval [0.160, 0.956] are generated. In the state S_k ($k \neq 0$), three successive points {0.956, 0.160, 0.514} corresponding to the three-periodic points of the logistic map are generated. The conditional probabilities $P(S_k|S_l)$ means the transition probability from the state S_l to the state S_k, and

$$P(S_{k+1}|S_k) + P(S_0|S_k) = 1 \quad (0 \leq k < L) \tag{2}$$

must be satisfied.

If we denote the stationary probability for the state S_k as $Q(S_k)$, the transition probabilities satisfy the following equations.

$$Q(S_0) = \sum_{l=0}^{L-1} P(S_0|S_l)Q(S_l) + Q(S_L) \tag{3}$$

$$Q(S_k) = P(S_k|S_{k-1})Q(S_{k-1}) \quad (0 < k \leq L) \tag{4}$$

$$\sum_{k=0}^{L} Q(S_k) = 1. \tag{5}$$

We derive the stationary probabilities of the Markov chain from the simulated data of the intermittency chaos by counting the number of the corresponding state. Further, the transition probabilities are calculated from the stationary probabilities by using Eqs. (3) and (4).

Figure 4 shows an example of time series obtained from the Markov chain model for $L = 15$. The transitions between the states are decided by using a random function according to the obtained transition probablities and the values in the burst parts are also given by using a random function. In order to check the statistical property of the obtained

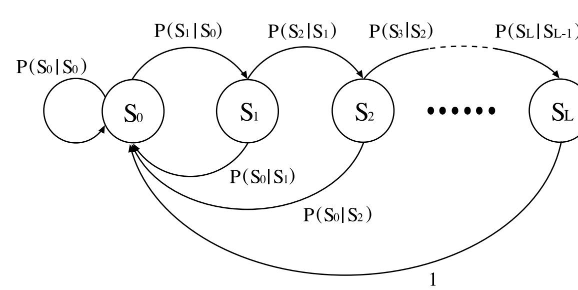


Fig. 3 Markov chain.

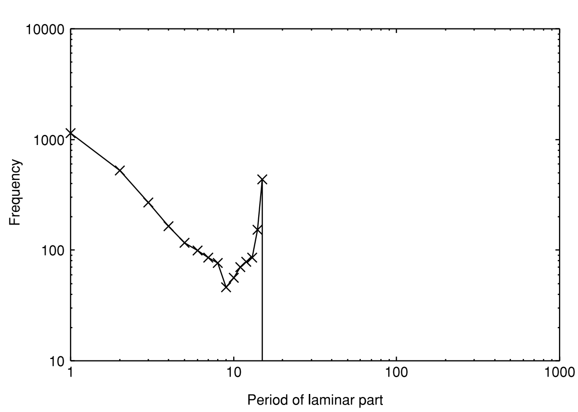


Fig. 2 Distribution of period of laminar part (Intermittency chaos for α=3.827940).

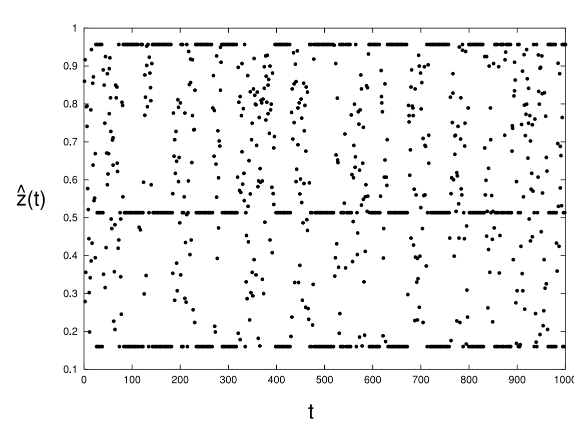


Fig. 4 Time series of the Markov chain model for L=15.

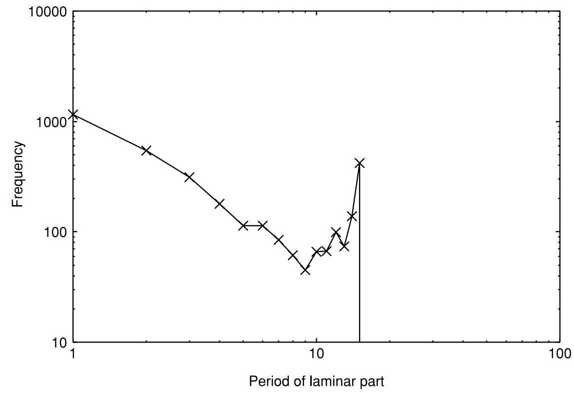


Fig. 5 Distribution of period of laminar part (Markov chain model for $L = 15$).

Table 1 Statistical properties of intermittency chaos and Markov chain model.

	Ratio of laminar part		Average period of laminar part	
L	Chaos	Markov	Chaos	Markov
7	0.4215	0.4156	3.1495	3.0789
9	0.4714	0.4580	3.8271	3.7139
11	0.5049	0.5049	4.3486	4.2729
13	0.5348	0.5289	4.9732	4.9216
15	0.5558	0.5524	5.4389	5.2984
17	0.5798	0.5700	5.9185	5.7266
19	0.5905	0.6083	6.5688	6.5358
21	0.6251	0.6313	7.0888	7.1835
23	0.6422	0.6424	7.4430	7.3444
25	0.6445	0.6427	7.4853	7.4764
30	0.6878	0.6804	9.2968	8.8704
40	0.7261	0.7196	11.5463	11.1663
50	0.7655	0.7580	13.6257	12.8656
70	0.8240	0.8167	19.7186	19.2385
100	0.8563	0.8532	24.8647	24.4863

time series, we counted the period of the laminar part. The result is shown in Fig. 5. We can say that the result is very close to that in Fig. 2.

Further, we produce the Markov chain models for various sizes of the maximum period L of the laminar part. The comparison of the statistical properties of the original intermittency chaos and the corresponding Markov chain model is summarized in Table 1. We can confirm that the property of the Markov chain models is similar to that of the original intermittency chaos.

4. Application to Hopfield NN for QAP

In this section, as the first example of applications of the Markov chain model of the intermittency chaos, we investigate the performance of the Hopfield NN for a QAP when the intermittency chaos and the Markov chain model are injected to the network in order to avoid the local minimum trapping problem.

4.1 Solving QAP with Hopfield NN

A QAP is expressed as follow: given two matrices, distance matrix C and flow matrix D, and find the permutation $\mathbf{p}$ which corresponds to the minimum value of the objective function $f(\mathbf{p})$ in Eq. (6).

$$f(\mathbf{p}) = \sum_{i=1}^{N_Q} \sum_{j=1}^{N_Q} C_{ij} D_{p(i)p(j)} \tag{6}$$

where C_{ij} and D_{ij} are the (i, j)-th elements of C and D, respectively, $p(i)$ is the i-th element of vector $\mathbf{p}$, and N_Q is the size of the problem. There are many real applications which are formulated by Eq. (6).

Because QAP is very difficult, it is almost impossible to solve the optimum solutions in larger problems. The largest problem which is solved by deterministic methods may be only 20 in recent study. Further, computation time is very long to obtain the exact optimum solutions. Therefore, it is usual to develop heuristic methods which search nearly optimal solutions in reasonable time. For solving N_Q-element QAP by the Hopfield NN, $N_Q \times N_Q$ neurons are required and the following energy function is defined to fire (i, j)-th neuron at the optimal position:

$$E_Q = -\frac{1}{2} \sum_{i,m=1}^{N_Q} \sum_{j,n=1}^{N_Q} w_{im;jn} x_{im} x_{jn} + \sum_{i,m=1}^{N_Q} \theta_{im} x_{im}. \tag{7}$$

The neurons are coupled each other with the synaptic connection weight. Suppose that the weight between (i, m)-th neuron and (j, n)-th neuron and the threshold of the (i, m)-th neuron are described by:

$$\begin{aligned} w_{im;jn} &= -2\left\{A(1-\delta_{mn})\delta_{ij} + B\delta_{mn}(1-\delta_{ij}) + \frac{C_{ij}D_{mn}}{q}\right\} \\ \theta_{im} &= -(A+B) \end{aligned} \tag{8}$$

where A and B are positive constants, and δ_{ij} is Kronecker's delta. The states of the $N_Q \times N_Q$ neurons are asynchronously updated due to the following difference equation:

$$x_{im}(t+1) = g\left(\sum_{j,n=1}^{N_Q} w_{im;jn} x_{jn}(t) + \theta_{im} + \beta_Q z_{im}(t)\right) \tag{9}$$

where $g(\cdot)$ is a sigmoidal function

$$g(x) = \frac{1}{1 + \exp\left(-\dfrac{x}{\varepsilon}\right)}, \tag{10}$$

$z_{im}(t)$ is the intermittency chaos or the Markov chain model, and β_Q limits the amplitude of the injected time series. Note that we normalize $\hat{z}_{im}$ by Eq. (11) before the injection.

$$z_{im}(t+1) = \frac{\hat{z}_{im}(t) - \bar{z}}{\sigma_z} \tag{11}$$

where $\bar{z}$ is the average of $\hat{z}(t)$, and σ_z is the standard deviation of $\hat{z}(t)$.

Further, we use the method suggested by Sato et al. (1.1 in [8]) to decide firing of neurons.

4.2 Simulated Results for QAP

The problem used here is chosen from the site QAPLIB which collects the bench mark problems [9]. We carried out computer simulations for the problems with various sizes. The results for "Nug12" are shown in this section. The global minimum of this problem is known as 578. The parameters of the Hopfield NN are fixed as $A = 1.0$, $B = 1.0$, $q = 100$ and $\varepsilon = 0.02$ and the amplitude of the injected time series is fixed as $\beta_Q = 0.6$. The number of updating the network N_{it} is 10000.

In order to evaluate the performance precisely, we neglect the once-appeared-solutions and use the two functions *Depth_1* and *Depth_2* proposed in our previous study [10]. Both functions have been introduced to evaluate finding a lot of nearly optimal solutions. The difference between them is the treatment of bad solutions. The function *Depth_1* evaluates all obtained solutions as good factor, even if the energies of the solutions are much worse than the optimal energy. The concept of this function is based on the idea that the network which finds a lot of bad solutions is better than the network does not find any solutions at all. On the other hand, *Depth_2* evaluates bad solutions as bad factor. In order to eliminate the effect of bad solutions, we not only set up a threshold but give a penalty according to the energy. The concept of this function is based on the idea that finding a lot of bad solutions hinders finding good solutions.

4.2.1 *Depth_1*

The first function *Depth_1* is defined as

$$Depth\_1 = \sum_{k=0}^{n} \{f(\mathbf{p}_k) - D_\infty\}^2 \tag{12}$$

where D_∞ is a constant which is large enough to include the energies of all solutions, n is the number of the accepted solutions and the energy $f(\mathbf{p}_k)$ is calculated by Eq. (6) using the permutation $\mathbf{p}_k$ corresponding to the k-th solution.

The calculated result of *Depth_1* is shown in Fig. 6. The horizontal axis is the maximum period of the laminar part L. We confirm that both the intermittency chaos and the Markov chain model exhibit similar tendency such that *Depth_1* decreases as the maximum period L of the laminar part increases. However, the *Depth_1* for the Markov chain model is not as large as those for the intermittency chaos.

4.2.2 *Depth_2*

The function *Depth_2* is defined as follows:

$$Depth\_2 = \sum_{k \in \mathbf{k}_g}^{n} \{f(\mathbf{p}_k) - D_{th}\}^2 - \sum_{k \notin \mathbf{k}_g}^{n} \{f(\mathbf{p}_k) - D_{th}\}^2$$
$$\text{where} \quad \mathbf{k}_g = \{k \mid f(\mathbf{p}_k) \leq D_{th}\}. \tag{13}$$

The calculated result of *Depth_2* is shown in Fig. 7.

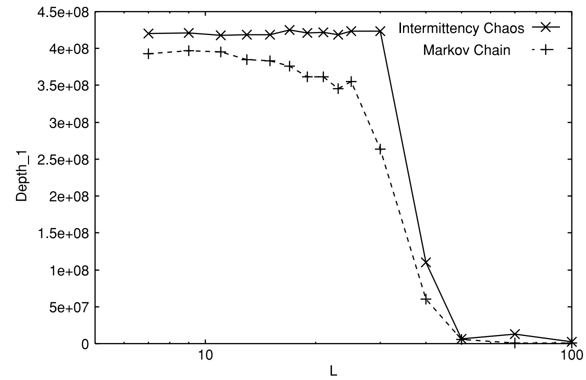


Fig. 6 Result of *Depth_1* ($D_\infty = 1000$).

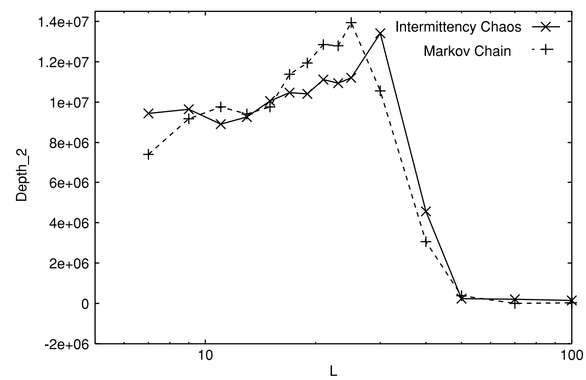


Fig. 7 Result of *Depth_2* ($D_{th} = 812$).

The Markov chain model gains similar performance to the intermittency chaos. This result shows that the Markov chain model keeps the good characteristic of the intermittency chaos to find a lot of nearly optimal solutions of QAP.

5. Application to Hopfield NN for Associative Memory

In this section, as the second example, we investigate the convergence performance of the Hopfield NN working as an associative memory when the intermittency chaos and the Markov chain model are injected to the network as a noise. Similarly to the case of QAP, we have shown that the injection of the intermittency chaos to the Hopfield NN for an associative memory could avoid the local minimum trapping problem and could improve the rate and the speed of the convergence to a stored pattern [11].

5.1 Hopfield NN Working as Associative Memory

Associative memory is a system which returns a stored pattern that is similar to a presented pattern. Noisy patterns can be corrected or distorted patterns can be recognized by a well-constructed associative memory.

The Hopfield NN is used as an associative memory by exploiting the property that the network has multiple stable states. Namely, if the parameters of the network can be decided in such a way that the patterns to be stored become stable states of the network, the network produces a stored pattern that is similar to an input pattern.

The energy function of the Hopfield NN with N_A neurons and M stored binary patterns is defined by the following equation.

$$E_A = -\frac{1}{2}\sum_{i=1}^{N_A}\sum_{j=1}^{N_A} w_{ij}x_i x_j - \sum_{i=1}^{N_A}\theta_i x_i \tag{14}$$

where w_{ij} is the weight between i-th neuron and j-th neuron, and θ_i is the threshold of the i-th neuron. They are given as follows.

$$w_{ij} = \begin{cases} \sum_{m=1}^{M}(2x_{mi}-1)(2x_{mj}-1) & (i \neq j) \\ 0 & (i = j) \end{cases} \tag{15}$$

$$\theta_i = 0. \tag{16}$$

The states of the neurons are asynchronously updated due to the following difference equation:

$$x_i(t+1) = g\left(\sum_{j=1}^{N_A} w_{ij}x_j(t) + \theta_i + \beta_A z_i\right) \tag{17}$$

where $g(\cdot)$ is a sigmoidal function, z_i is the intermittency chaos or the Markov chain model, and β_A limits the amplitude of the injected time series.

In this application, firing of neurons is decided by the output value of more than 0.5.

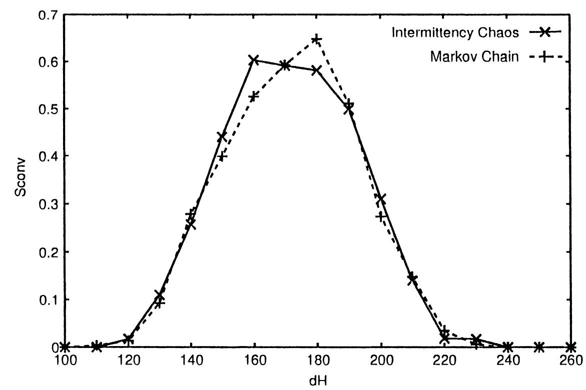
5.2 Simulated Results for Associative Memory

In the computer simulations, we consider the Hopfield NN with 400 neurons. In order to investigate the performance of the network under difficult conditions, we prepare an input binary pattern at random and produce several binary patterns to be stored whose distances from the input pattern are the exactly same. The Hamming distance d_H is used to evaluate the convergence of the network. Namely, the convergence time T_{conv} is defined as the iteration number of the network when the Hamming distance between the output of the network and one of the stored patterns becomes zero. Further, we propose S_{conv} to evaluate convergence speed more precisely. S_{conv} is defined by Eq. (18).

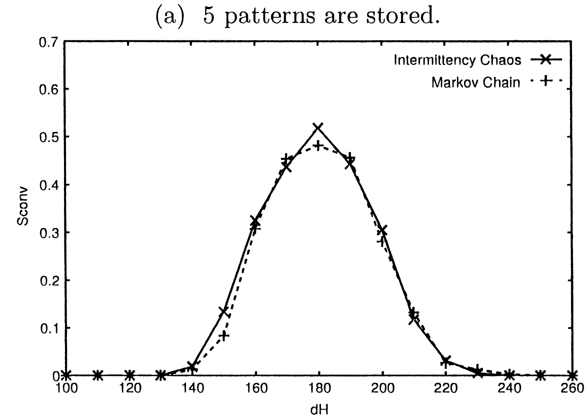
$$S_{conv} = 1 - \frac{\min[\, T_{conv},\ N_{max}\,]}{N_{max}} \tag{18}$$

where N_{max} is the given upper limit of the iteration number of the simulation. Namely, if the network dose not converge to any stored patterns during the given iteration, the value of S_{conv} is zero.

Computer simulations are carried out for various conditions. Typical results for two cases of 5 stored patterns and 8 stored patterns are shown in Fig. 8. For this simulation, the intermittency chaos for $\alpha = 3.827940$ and the corresponding Markov chain model with $L = 15$ are injected. The amplitude of the injected time series is $\beta_A = 50.0$, the parameter of the Hopfield NN is $\varepsilon = 0.02$, and the maximum iteration number of the network is fixed as $N_{max} = 10000$. The horizontal axis is the Hamming distance between the input pattern and the stored patterns and the vertical axis is the average value of S_{conv} in 100 trials.



(a) 5 patterns are stored.



(b) 8 patterns are stored.

Fig. 8 Simulated results of S_{conv}.

The Markov chain model gains similar performance to the intermittency chaos. This result shows that the Markov chain model keeps the good property of the intermittency chaos, which can improve the rate and the speed of the convergence of the Hopfield NN as an associative memory.

6. Conclusions

In this study, a modeling method of the intermittency chaos using the Markov chain has been proposed. The performances of the intermittency chaos and the Markov chain model were investigated when they were injected to the Hopfield NN for a QAP or an associative memory. Computer simulated results showed that the proposed modeling was good enough to gain similar performance of the intermittency chaos.

References

[1] Y. Pomeau and P. Manneville, "Intermittent transition to turbulence in dissipative dynamical systems," Commun. Math. Phys., vol.74, pp.189–197, 1980.

[2] C.G. Langton, "Computation at the edge of chaos: Phase transitions and emergent computation," Physica D, vol.42, pp.12–37, 1990.
[3] J.J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," Proc. Natl. Acad. Sci. USA, vol.81, pp.3088–3092, 1984.
[4] Y. Hayakawa and Y. Sawada, "Effects of the chaotic noise on the performance of a neural network model for optimization problems," Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip Top. vol.51, no.4, pp.2693–2696, 1995.
[5] H. Dedieu and M.J. Ogorzalek, "Overview of nonlinear noise reduction algorithms for systems with known dynamics," Proc. NOLTA'98, vol.3, pp.1297–1300, 1998.
[6] R. Rovatti, G. Mazzini, and G. Setti, "A tensor approach to higher order expectations of quantized chaotic trajectories; part I and II," IEEE Trans. Circuits Syst. I, Fundam. Theory Appl., vol.47, no.11, pp.1571–1596, 2000.
[7] R. Rovatti and G. Mazzini, "Tensor function analysis of quantized chaotic piecewise-affine pseudo-Markov systems; I and II," IEEE Trans. Circuits Syst. I, Fundam. Theory Appl., vol.49, no.2, pp.137–162, 2002.
[8] K. Sato, T. Ikeguchi, M. Hasegawa, and K. Aihara, "An optimization method for quadratic assignment problems by chaotic dynamics and its characterization by Lyapunov dimensions," IEICE Technical Report, NLP64-13, 2001.
[9] R.E. Berkard, S.E. Karisch, and F. Rendl, "QAPLIB—A quadratic assignment problem library," http://www.opt.math.tu-graz.ac.at/qaplib
[10] Y. Uwate, Y. Nishio, T. Ueta, T. Kawabe, and T. Ikeguchi, "Performance of chaos noise injected to Hopfield NN for quadratic assignment problems," Proc. NOLTA'02, pp.267–270, Oct. 2002.
[11] Y. Uwate, Y. Nishio, and T. Ikeguchi, "Associative memory by Hopfield NN with chaos injection," IEICE Technical Report, NLP2003-44, 2003.

Yoko Uwate was born in Tokushima, Japan, in 1980. She received the B.E. degree from Tokushima University, Tokushima, Japan, in 2003. She is currently working towards M.E. degree at the same university. Her research interests include complex phenomena in chaotic circuits and neural networks. Ms. Uwate is a student member of the IEEE.

Yoshifumi Nishio received the B.E., M.E. and Ph.D. degrees in Electrical Engineering from Keio University, Yokohama, Japan, in 1988, 1990 and 1993, respectively. In 1993, he joined the Department of Electrical and Electronic Engineering at Tokushima University, Tokushima, Japan, where he is currently an Associate Professor. From May 2000 he spent a year in the Laboratory of Nonlinear Systems (LANOS) at the Swiss Federal Institute of Technology Lausanne (EPFL) as a visiting professor. His research interests include analysis and application of chaos in electrical circuits, analysis of synchronization in nonlinear circuits, development of analytical methods for nonlinear circuits and theory and application of cellular neural networks. Dr. Nishio is a member of the IEEE.

Akio Ushida received the B.E. and M.E. degrees in Electrical Engineering from Tokushima University in 1961 and 1966, respectively, and the Ph.D. degree in Electrical Engineering form University of Osaka Prefecture in 1974. He was an Associate Professor from 1973 to 1980 at Tokushima University. From 1980 to 2003, he was a Professor in the Department of Electrical and Electronic Engineering at Tokushima University. Since 2003, he has been with the Department of Mechanical-Electronic Engineering, Tokushima Bunri University. From 1974 to 1975 he spent one year as visiting scholar at the Department of Electrical Engineering and Computer Sciences at the University of California, Berkeley. His current research interests include numerical methods and computer-aided analysis of nonlinear systems. Dr. Ushida is a member of IEEE.