

Associative Memory by Hopfield NN with Chaos Injection

Yoko Uwate Yoshifumi Nishio

Department of Electrical Engineering

Tokushima University, Japan

E-mail: {uwate,nishio}@ee.tokushima-u.ac.jp

Tohru Ikeguchi

Department of Information

and Computer Sciences

Saitama University, Japan

E-mail: tohru@nls.ics.saitama-u.ac.jp

Abstract—Several people point out that the Hopfield Neural Network (abbr.NN) with chaos injection gains the good performance for solving traveling salesman problems, which is one of combinatorial optimization problems. In this study, we investigate the performance of the intermittency chaos injected to the Hopfield NN working as an associative memory. By computer simulations, the rate and the speed of the convergence to an embedded pattern are evaluated. Furthermore, in order to confirm the reason of the good ability of intermittency chaos, we carry out the same simulation using the time series produced by the Markov chain model. Simulated results show that the Markov chain model is good enough to gain similar performance of the intermittency chaos.

I. INTRODUCTION

Intermittency chaos [1] is deeply related to *the edge of chaos* [2] and many people 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 NN (Hopfield NN) [3] solving combinatorial optimization problems to avoid trappings of the solutions into a local minimum. If we choose connection weights between neurons appropriately according to given problems, we can obtain a good solution by the energy minimization principle. However, the solutions are often trapped into a local minimum and do not reach the global minimum. In order to avoid this critical problem, several people proposed the method adding some kinds of noise to the Hopfield NN. Hayakawa *et al.* pointed out that chaos near the three-periodic window of the logistic map gains the best performance for solving traveling salesman problems (TSP) [4]. The authors have investigated the performance of the Hopfield NN solving combinatorial optimization problems when intermittency chaos, which is the phenomenon on the edge of chaos, is inputted to the neurons as noise [5]–[7]. By computer simulations, the intermittency chaos has been confirmed to gain better performance to escape out of local minima than fully-developed chaos. However, only the results on the combinatorial optimization problems can not conclude that the intermittency chaos has an excellent ability to escape out of local minima of the Hopfield NNs.

In this study, the performance of the Hopfield NN working as an associative memory is evaluated when intermittency chaos is inputted as noise. 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. By computer simulations, the rate and the speed of the convergence to an embedded pattern are evaluated. Furthermore, in order to confirm the reason of the good ability of intermittency chaos, we carry out the same simulation using the time series produced by the Markov chain model that have already proposed by the authors [8]. Simulated results show that the Markov chain model is good enough to gain similar performance of the intermittency chaos.

II. 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 neurons and M stored binary patterns is defined by the following equation.

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{ij} x_i x_j - \sum_{i=1}^N \theta_i x_i \quad (1)$$

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.

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

$$\theta_i = 0. \quad (3)$$

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

$$x_i(t+1) = g \left(\sum_{j=1}^N w_{ij} x_j(t) + \theta_i + \beta z_i \right) \quad (4)$$

where g is a sigmoidal function defined as follows:

$$g(x) = \frac{1}{1 + \exp \left(-\frac{x}{\varepsilon} \right)} \quad (5)$$

z_i is the intermittency chaos or the fully-developed chaos, and β limits the amplitude of the injected time series. Figure 1 shows a conceptual neuron model for this NN.

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

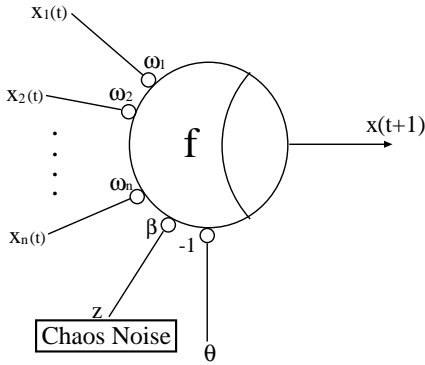


Fig. 1. Neuron model.

III. CHAOS NOISE

In this section, we describe chaos noise injected to the Hopfield NN. The logistic map is used to generate chaos noise:

$$\hat{z}_{im}(t+1) = \alpha \hat{z}_{im}(t)(1 - \hat{z}_{im}(t)). \quad (6)$$

Varying parameter α , Eq. (6) behaves chaotically via a period-doubling cascade. When we inject chaos noise to the Hopfield NN, we normalize \hat{z}_{im} by Eq. (7).

$$z_{im}(t+1) = \frac{\hat{z}_{im}(t) - \bar{z}}{\sigma_z} \quad (7)$$

Where \bar{z} is the average of $\hat{z}(t)$, and σ_z is the standard deviation of $\hat{z}(t)$. Figure 2 shows an example of the time series of the intermittency chaos near the three-periodic window. As we can see from the figure, the intermittency chaos 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.

For the comparison, we also carry out computer simulation for the case that fully-developed chaos and uniform random noise in Figs. 3 and 4 are injected to the Hopfield NN.

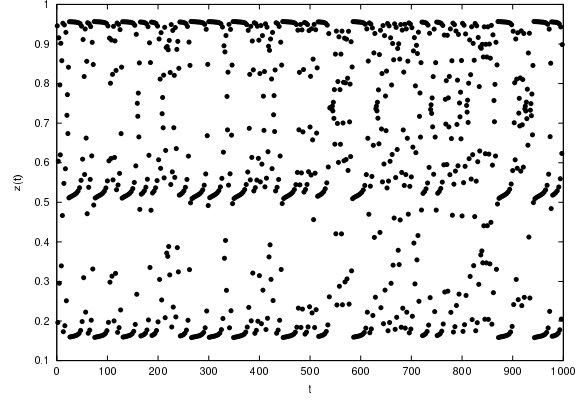


Fig. 2. Intermittency chaos ($\alpha = 3.8274$).

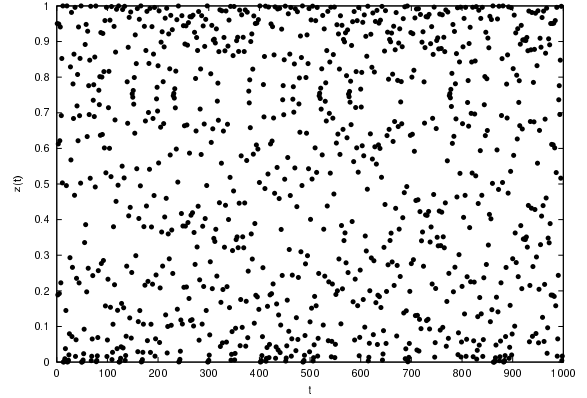


Fig. 3. Fully-developed chaos ($\alpha = 4.0000$).

IV. SIMULATED RESULTS

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 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.

A. Convergence Speed

Typical results for 8 stored patterns when the initial Hamming distance is fixed as $d_H = 150$ is shown in Fig. 5. The horizontal axis is time and the vertical axis is the Hamming distance. Figure 5 (a) shows the case of intermittency chaos injection, (b) shows the case of fully-developed chaos injection, (c) shows the case of uniform random noise injection, and (d) shows the case of no noise. In these figures, we can see that the case of no noise does not achieve the convergence and that the cases of noise injections succeed in converging to one of the embedded patterns. Especially, we confirm that

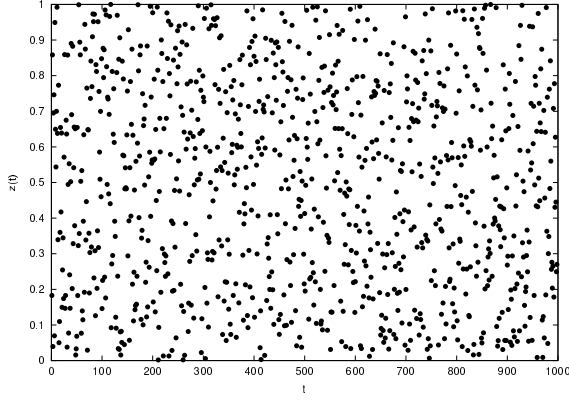


Fig. 4. Uniform random noise.

the intermittency chaos can converge quicker than the fully-developed chaos and the uniform random noise.

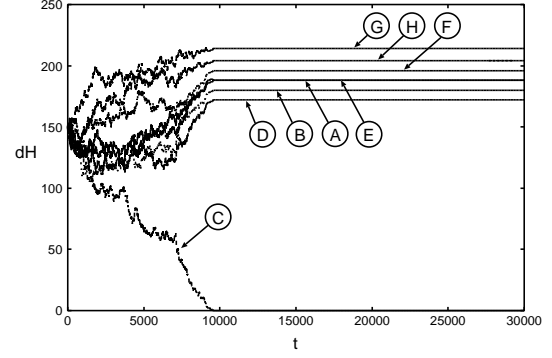
Next, we carry out the simulations when the input patterns are divided into two groups with different initial Hamming distances. Typical results for 10 stored patterns are shown in Fig. 6. The initial Hamming distance of one group of 5 patterns are fixed is $d_H = 190$, and the initial Hamming distance of the other group is fixed as $d_H = 140$. In these figures, we can see the similar results to the 8 patterns case of the good performance of the intermittency chaos and that the cases of noise injections succeed in converging to one of the embedd patterns in the small Hamming distance group.

B. Markov Chain Modeling

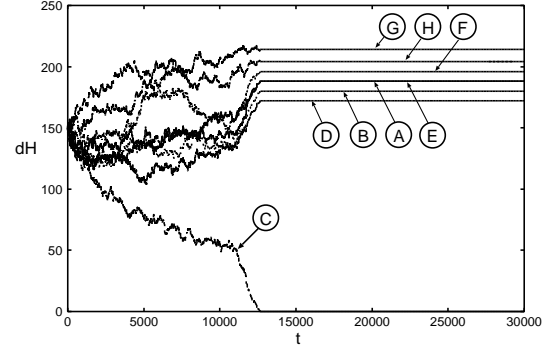
In order to clear up the reason of the good ability of the intermittency chaos, we have proposed the modeling method of the intermittency chaos by using the Markov chain model [8]. This Markov chain model was made by imitating some characteristics of the intermittency chaos; “existing laminar parts and burst parts” and “distribution of period of laminar part”. If the Markov chain model got the similar performance to the intermittency chaos, we can explain the reason of the good ability of the intermittency chaos by using these characteristics.

We explain how to make the Markov chain model. 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. 7. 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. We consider that this is the most distinguished feature of the intermittency chaos.

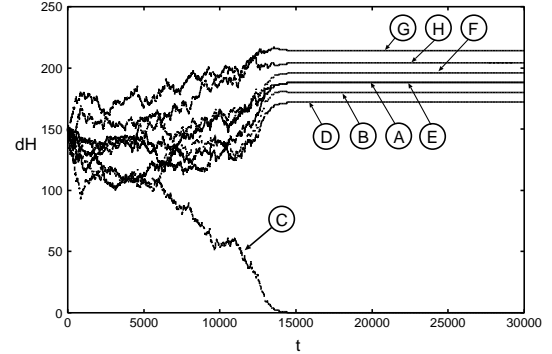
In order to model the above-mentioned feature of the intermittency chaos, we proposed the Markov chain as shown



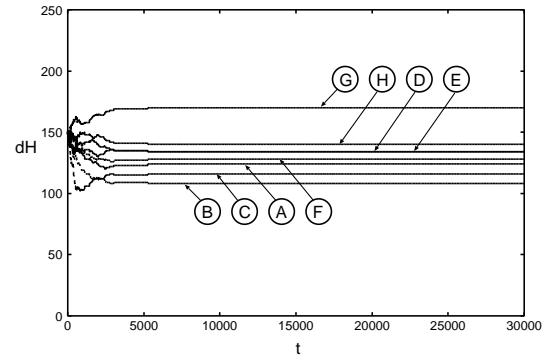
(a) Intermittency chaos



(b) Fully-developed chaos

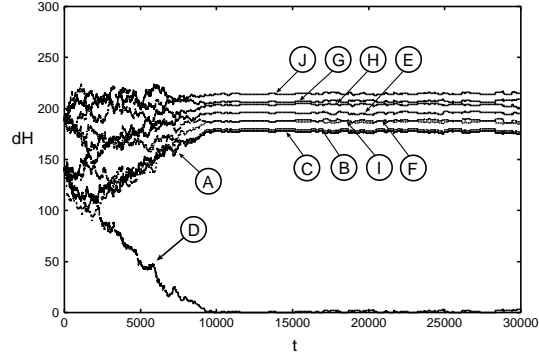


(c) Uniform random noise

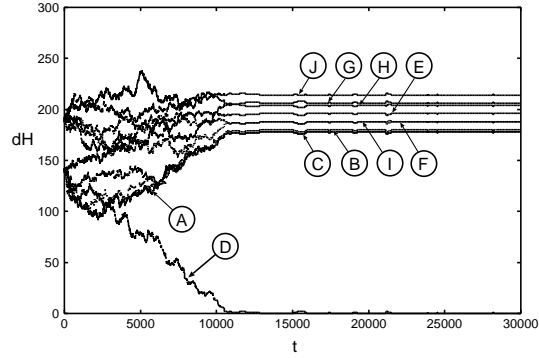


(d) No noise

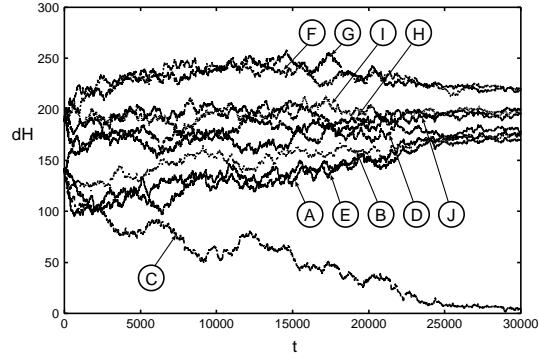
Fig. 5. Speed of convergence (8 patterns). Initial Hamming distance of all patterns $d_H = 150$.



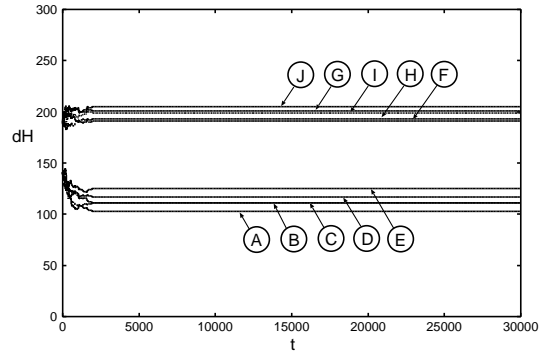
(a) Intermittency chaos



(b) Fully-developed chaos



(c) Uniform random noise



(d) No noise

Fig. 6. Speed of convergence (10 patterns). Initial Hamming distance of one group of 5 patterns is $d_H = 190$ and initial Hamming distance of the other group is $d_H = 140$.

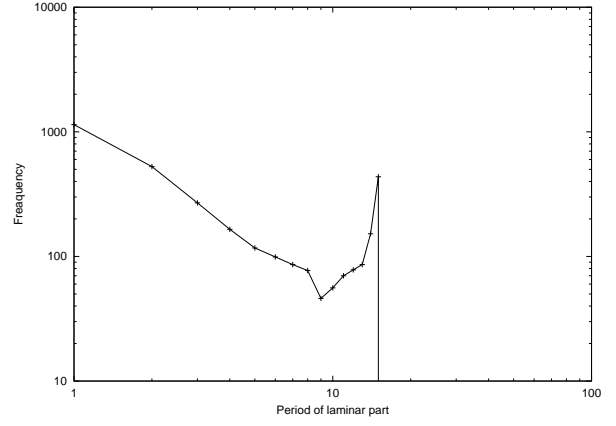


Fig. 7. Distribution of period of laminar part. (Intermittency chaos for $\alpha=3.827940$.)

in Fig. 8. In this Markov chain, the state S_0 corresponds to the burst part and the states S_1, S_2, \dots, 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

$$P(S_{k+1}|S_k) + P(S_0|S_k) = 1 \quad (0 \leq k < L) \quad (8)$$

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.

$$Q(S_0) = \sum_{l=0}^{L-1} P(S_0|S_l)Q(S_l) + Q(S_L) \quad (9)$$

$$Q(S_k) = P(S_k|S_{k-1})Q(S_{k-1}) \quad (0 < k \leq L) \quad (10)$$

$$\sum_{k=0}^L Q(S_k) = 1. \quad (11)$$

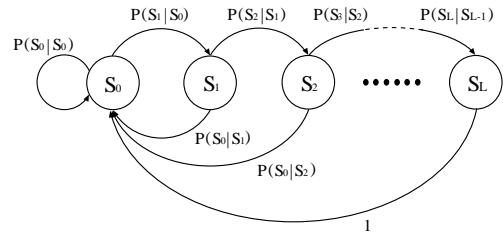


Fig. 8. Markov chain.

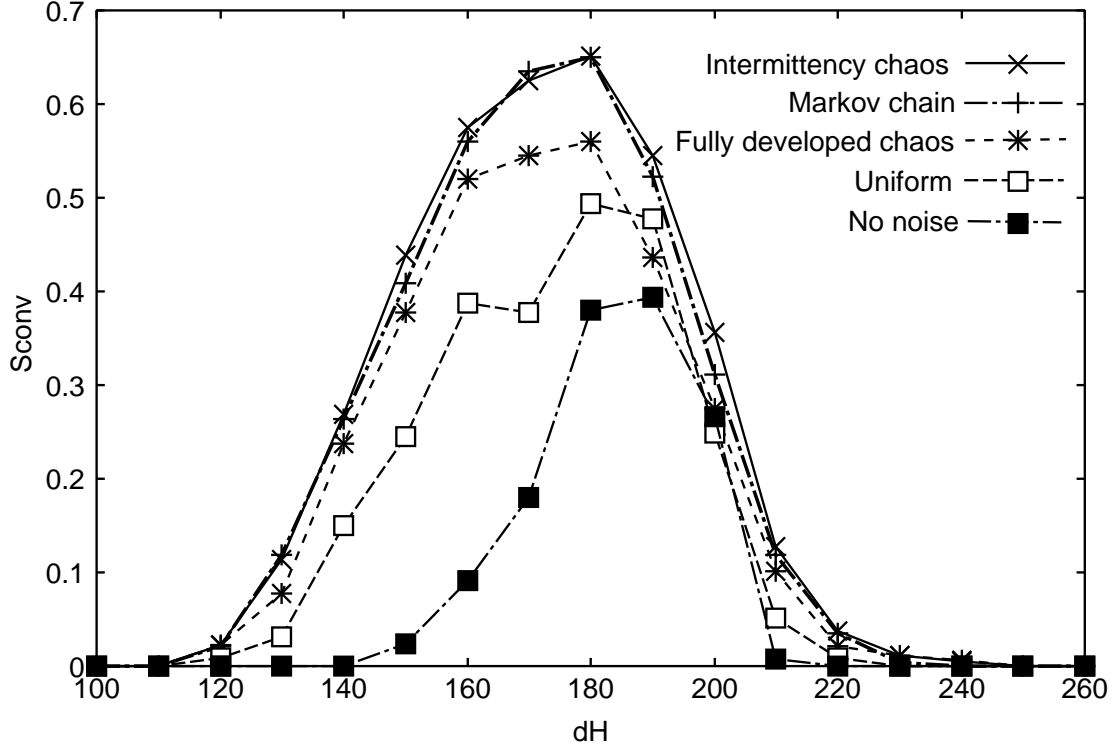


Fig. 11. Convergence speed (S_{conv}).

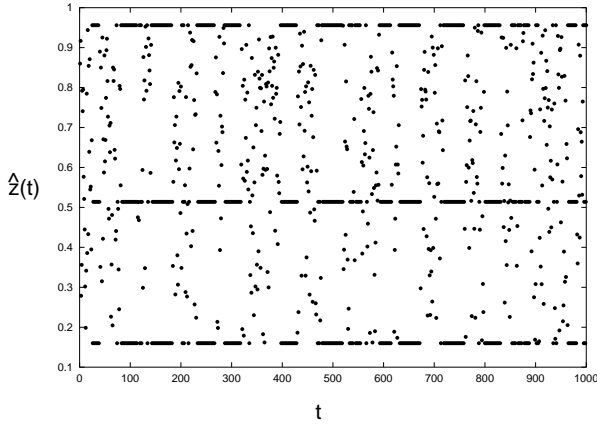


Fig. 9. Time series of the Markov chain model for $L=15$.

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. (9) and (10).

Figure 9 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 probabilities 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 time series, we

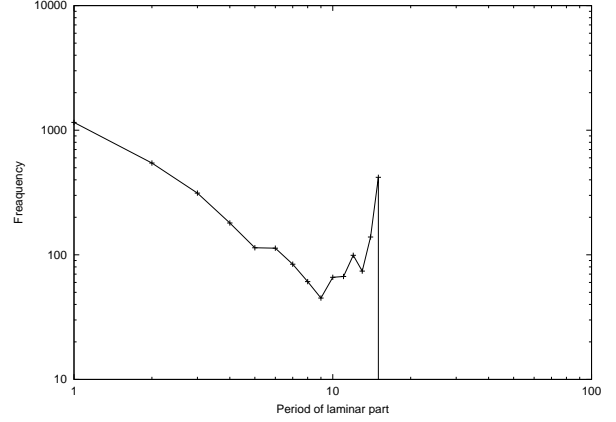


Fig. 10. Distribution of period of laminar part. (Markov chain model for $L=15$.)

counted the period of the laminar part. The result is shown in Fig. 11. We can say that the distribution is very close to that in Fig. 8.

C. Evaluation Method

In order to investigate the performance of the network under difficult conditions, we propose S_{conv} to evaluate convergence speed more precisely. S_{conv} is defined by Eq. (12).

$$S_{conv} = 1 - \frac{\min[T_{conv}, N_{max}]}{N_{max}} \quad (12)$$

where N_{max} is the given upper limit of the iteration number of the simulation. Namely, if the network does 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 5 stored patterns is shown in Fig. 11. For this simulation, the amplitude of the injected noise is $\beta = 50.0$. 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.

The results show that the intermittency chaos and the fully-developed chaos have much better performance than the case of no noise. Furthermore, we consider that the convergence speed of the intermittency chaos is quicker than the fully-developed chaos in wide range. And, the Markov chain model is good enough to gain similar performance of the intermittency chaos.

V. CONCLUSIONS

In this study, the performance of the Hopfield NNs working as an associative memory has been evaluated when the intermittency chaos is inputted as noise. By computer simulations, the rate and the speed of the convergence to an embedded pattern were investigated. We confirmed that the noise injection succeeds in converging to one of the embedded patterns, and the intermittency chaos is better than fully-developed chaos.

REFERENCES

- [1] Y. Pomeau and P. Manneville, "Intermittent Transition to Turbulence in Dissipative Dynamical Systems," *Comm. 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," *Physical Review E*, vol. 51, no. 4, pp. 2693-2696, 1995.
- [5] T. Ueta, Y. Nishio and T. Kawabe, "Comparison between Chaotic Noise and Burst Noise on Solving Ability of Hopfield Neural Networks," *Proc. NOLTA'97*, vol. 1, pp. 409-412, Nov. 1997.
- [6] Y. Uwate, Y. Nishio, T. Ueta, T. Kawabe and T. Ikeguchi, "Solving Ability of Hopfield Neural Network with Chaotic Noise and Burst Noise for Quadratic Assignment Problem," *Proc. ISCAS'02*, vol. 3, pp. 465-468, May 2002.
- [7] 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*, vol. 1, pp. 267-270, Oct. 2002.
- [8] Y. Uwate, Y. Nishio, A. Ushida, "Markov Chain Modeling of Intermittency Chaos and its Application to Hopfield NN for QAP," *Proc. ECCTD'03*, vol. 2, pp. 57-60, Sep. 2003.