RAPID COMMUNICATIONS

PHYSICAL REVIEW E VOLUME 51, NUMBER 4 APRIL 1995

Effects of the chaotic noise on the performance of a neural network model for optimization problems

Yoshinori Hayakawa, Atsushi Marumoto, and Yasuji Sawada
Research Institute of Electrical Communication, Tohoku University, Katahira, Aoba-ku, Sendai 980-77, Japan
(Received 5 July 1994; revised manuscript received 30 January 1995)

We studied effects of chaos added to the dynamics of a neural network model. By numerical simulations, we found the neural network with forcing by chaotic noise operated very efficiently to solve an optimization problem. We also showed that short time correlation of chaos was relevant to the dynamics of the network and it could work effectively for global minima search.

PACS number(s): 05.45.+b, 82.20.Mj, 05.40.+j

Chaotic neural networks have recently received attention due to the potential capability for information processing [1–5]. The idea of chaotic neural networks which operate chaotically was discussed by Freeman [1] in the study of the olfactory system, and a model of a chaotic single neuron was proposed by Aihara *et al.* [3]. Since then, it has been known that chaos may play a nontrivial and unexpected role in interconnected chaotic neurons [3–5]. For instance, desirable results were reported in associative memory search and optimization tasks by using such a model by Nozawa [5]. He derived a discrete time neural network model by taking Euler's difference of a typical Hopfield model. In the model, each neuron behaves as a chaotic oscillator represented by a one-dimensional map which is sensitive to the control parameters. When each map was globally coupled with an appropriate synaptic weight as in the ordinary Hopfield model, he showed that the best solution of a ten city traveling salesman problem was found with probability of more than 90% within 1000 iterative steps. However, behavior of such networks is, in general, quite complex, and the dynamics and mechanism of their operation still have been open problems. Therefore, in order to clarify the effects of chaos, it is important at this stage to study a simpler model in which chaotic oscillators can be separated from other components and in which conventional behavior as neural networks can be recovered by changing parameters.

Combinational optimization problems can be solved with neurocomputers in which the dynamics is described by relaxation processes with a Lyapunov (energy) function. It is known that trapping by spurious minima in the energy function is the origin of difficulty in this kind of problem for efficient operation, especially in a large size problem. One of the naive methods to avoid the trapping is to excite the state with some additive noise large enough to kick it out from such local minima. A number of techniques and algorithms have been proposed so far for this purpose [6]. Most of these works studied what is the best scheduling for temporal variation of white noise to lead the system to the best solution. In addition, motion driven by colored or correlated noise sequences in a multistable system has also been studied in the fields of chemistry and biophysics. When the noise has an appropriate correlation, as found in chaos, some nontrivial behaviors such as symmetry breaking of motion and extraordinarily large mobility compared to the case with white noise [7,8] are reported. Anomalous effects of colored noise on the barrier crossing rate in potential energy have also been studied to obtain a physical understanding of the chemical reaction [9]. The results from these works suggest that correlated time series may accelerate the activation of the network from metastable states, and this helps the network to escape from spurious local minima. In this paper, based on this idea, we introduce a simple chaotic neural network model by using chaotic noise generators instead of thermal noise, and study the effect of chaotic noise on the performance in solving a traveling salesman problem (TSP) as a typical combinational optimization.

First we introduce a chaotic neural network model as an extension of the Hopfield and Tank implementation of the analog neural network [10]. The connection between neurons is basically represented by a modified Hopfield model, and noise generators are attached to neural units as illustrated in Fig. 1. Let u_i^n be the state of neuron i at time n and V_i^n be the output from the neuron i. Neurons are interconnected by synapses of weight $W_{i,j}$, and the states of the neurons are updated by the following scheme with time difference Δt:

$$u_i^{n+1}=\Delta t\left(\sum_j W_{ij}V_j+h\right)+(1-\Delta t)u_i^n, \tag{1}$$

$$V_i^n=f(u_i^n+A\,\eta_i^n), \tag{2}$$

where A is a multiplier to the noise η_i^n whose amplitude is normalized to be unity, and f is a threshold function defined as

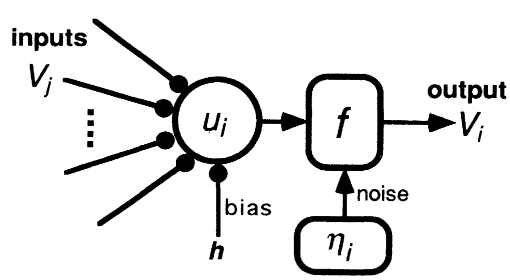

FIG. 1. Schematic diagram of model neuron with a noise generator.

$$f(y)=\frac{1}{1+\exp(-y)}\ . \tag{3}$$

Note that a generator of η_i is assigned to each neuron and they are independent of each other. The performance of the network may depend on the nature of the noise, and there can be several varieties of noise source η including chaotic oscillators. Here we study the dynamics of the network by adding the following three types of noise.

Case 1: Uniformly distributed random numbers. As a noise, we used first a sequence of uniformly distributed random numbers $x_i \in [0,1]$ which have no correlation between successive values, i.e.,

$$\langle x_i^n x_j^m \rangle = Q\,\delta_{i,j}\delta_{n,m}\,,$$

where Q is a constant. In this case, the dynamics of the network behaves as a so-called Gaussian machine [11], and Q is associated with a temperature parameter in the Gaussian machine.

Case 2: Logistic map. The following map is used as noise:

$$x_i^{n+1}=ax_i^n(1-x_i^n), \tag{4}$$

where $a \in [0,4]$ is the control parameter of map. In the present study we used the same values of a for all neurons. Thus all the noise generators are operated with the same rule but with different initial conditions.

Case 3: Logistic map with shuffling. In this case, the time series of the logistic map during 1000 iterative steps is stored into a computer memory, and selected one by one at random. Therefore the sequences of the shuffled noise have no time correlation, while other properties of x_i are the same as for case 2.

In order to make a quantitative comparison, raw data generated from those noise sources are normalized by the following procedure:

$$\eta_i^n=\frac{x_i^n-\langle x\rangle_n}{\sigma_x}\ , \tag{5}$$

where σ_x is the standard deviation of x from its average over n, i.e., $\sigma_x^2=\langle(x^n-\langle x\rangle_n)^2\rangle_n$.

As is well known, the dynamics of the network without noise sources can be expressed in terms of the relaxation processes, if synaptic weight $W_{i,j}$ is symmetric and if there is no feedback, in the potential energy defined by

$$E=-\tfrac{1}{2}\sum_{i,j} W_{i,j}V_iV_j-h\sum_i V_i+\sum_i \int_{1/2}^{V_i} f^{-1}(v)dv. \tag{6}$$

From an appropriate presentation of the energy function E for the TSP, a set of synaptic weight can be obtained. In this study, we used the TSP energy function in Ref. [10]: Using x and y as the index for cities and k and l for the order of visiting,

$$W_{i,j}=-c_0\delta_{x,y}(1-\delta_{k,l})-c_1\delta_{k,l}(1-\delta_{x,y})-c_2 - c_3 d_{x,y}(\delta_{k,l+1}+\delta_{k,l-1}), \tag{7}$$

where $\delta_{i,j}$ is Kronecker's delta, and $d_{x,y}$ is the distance between the cities indexed with x and y. For instance, for an N city problem, one can assign $x=i$ div$N$, $k=i$ mod$N$, $y=j$ div$N$, and $l=j$ mod$N$. In the following analysis, the configuration of cities is also identical to that reported in Ref. [10] with ten cities in two-dimensional space. Thus we had 100 neural units and 100^2 synapse connections. Due to the added noise, the dynamics of the network is no longer of the simple relaxation type in the potential energy given by Eq. (6). However, as long as the external noise is small, one expects that the state of the network in output-vector space senses the force due to the gradient of the energy of Eq. (6) and additive fluctuating force. The main purpose of the present work is to examine the effect of the additive fluctuating force.

Parameter tuning is one of the important issues to improve the performance of such kinds of networks. In fact, total performance for finding the best solution strongly depends on the set of coefficients in E, bias h, amplitude of externally applied noise A, and the properties (generators) of noise. Since this study is intended to clarify the role of chaotic noise for the dynamics of the optimizing device, we fix in the following analysis some of the parameters as $c_0=c_1=c_3=5$, $c_2=1.5$, $h=15$, and $A=3$ and compare results by changing the nature of the noise sources. Every noise source has the same average amplitude, whereas the color and distribution might be different.

In this model, each neural unit has two separated dynamics with different time scales: One is the fast dynamics produced by iterations of noise generators, and the other is the slower one represented by an ordinary discrete neural network model. Thus the ratio of these time scales can be considered as another tunable parameter. In our simulations, we took $\Delta t=0.1$ in every case. Note that, if this ratio is too large (i.e., Δt is very small), case 2 will be reduced to case 3 because of the chaotic nature of the output.

We confirmed that all the network models defined above could find solutions of the TSP. However, the states u_i of neural units fluctuate with a large deviation due to the external force noise as seen in Fig. 2. In our preliminary analysis, when the network worked efficiently to visit solutions, the power spectrum of the fluctuation of u_i^n had a $1/f^{\alpha}$-like property, where the exponent α was about 1.5. The detailed characteristics of the time series seem to depend on models and parameters. Looking at Fig. 2, although the behavior of u_i seems to be rather complex, the states of u_i can still be separated into two domains with an appropriate threshold value (e.g., about -3) and one can distinguish which unit is firing and which is not. Thus, if the operation is successful, the representation of the solution is obtained at every iterative step. Due to the dynamical properties of the network, evaluation of performance should be carried out by averaging over time. To do that, we have tested the visiting frequency $P_{\rm OS}$ at the optimal solutions which is defined by

$$P_{\rm OS}=(\text{number of steps staying at the optimal solution})/(\text{total steps}).$$

In this paper, we calculated $P_{\rm OS}$ for the last 1000 steps of 2000 iterative steps of computer runs. $P_{\rm OS}$ can be considered as an index of the escaping efficiency from local minima.

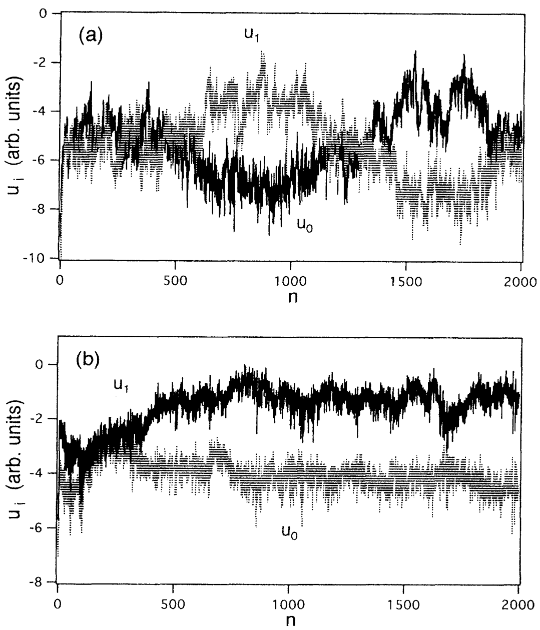

FIG. 2. Typical time evolution of neuron states u_i near the best solution (a) in case 1 (white noise) and (b) in case 2 (chaotic noise). All parameters are the same in both except for noise generators. In case 2, once the optimal solution is found, u_i stays there more persistently compared to case 1.

First, we checked the performance of the model with white noise generators (case 1) as a reference for later discussions. For 100 independent trials starting from random initial conditions, the obtained $P_{\rm OS}$ was about 0.003 with the values of the parameters mentioned above. With more careful parameter tuning of h and A, we obtained at most $P_{\rm OS}=0.49$.

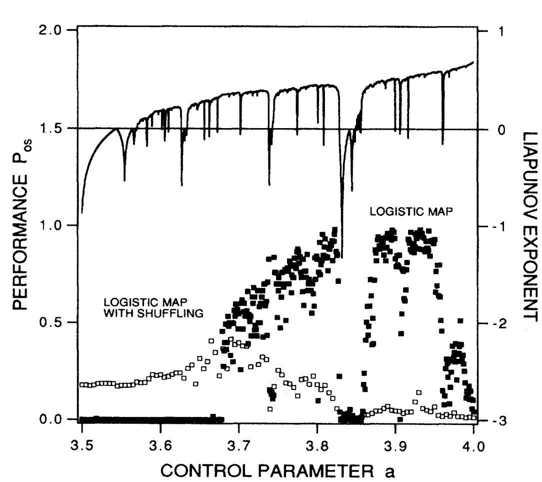

FIG. 3. Performance of network versus the control parameter for the logistic map (black rectangles) and the logistic map with shuffling (white rectangles). Each point is obtained by 30 independent runs of the simulation started from different initial conditions. For comparison, the Lyapunov exponent of the logistic map is also shown by a solid line.

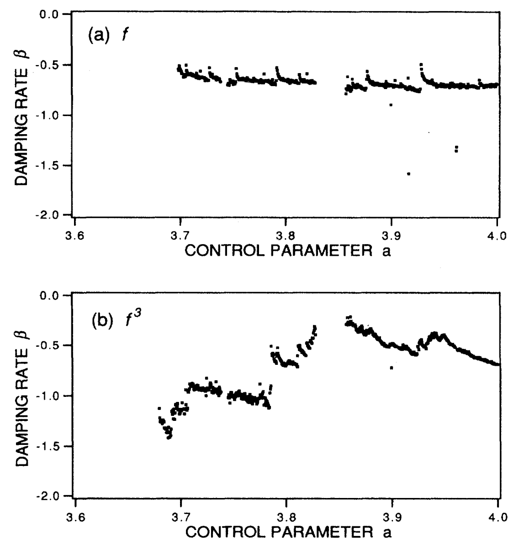

FIG. 4. Decay rate β of the histogram $n(l) \sim \exp(\beta l)$ of events in which positive sequence y_i successively occurs l times in the "logistic noise generator" for (a) the logistic map, and (b) the threefold iterated logistic map by changing the control parameter a. Data are shown only for nonperiodic states.

Next, the efficiency of computing in case 2 would be sensitive to the characteristics of the output signal from the maps. In Fig. 3, $P_{\rm OS}$ is plotted by using the logistic map as a noise source versus a (black rectangles) with the corresponding Lyapunov exponent of the logistic map (solid line). It is remarkable that below the band merging point ($a_{\rm BM} \simeq 3.68$), even in the parameter regime where the map already generates chaotic sequences, the performance cannot be improved at all compared to that for white noise. However, once two bands merge, the performance drastically increases and reaches about 100% except in the windows of the logistic map. By numerical evidence, the network works most effectively near the largest period-3 window. The performance remains high for the parameter value a greater than that for the period-3 window, and gradually decreases when a approaches 4.

The results of $P_{\rm OS}$ obtained are not trivial. As seen in Fig. 3, while there seems to be no correlation between the performance and the Lyapunov exponent of the map, some characteristics of chaotic noise which depend on the control parameter a certainly have a significant effect on the total behavior of the network.

To understand what factors are relevant in "logistic noise" for such a remarkable improvement of the operation, we preserved the distribution of inputs but erased the correlations by shuffling the time series (case 3). In Fig. 3, the obtained performance versus control parameter a is shown by white rectangles. We observe that the annihilation of correlation remarkably decreased the performance at any control parameter above the band merging point $a_{\rm BM}$, while it leads to better results below it. These observations clearly indicate that correlation of the noise plays an important role for the network system to find the best solution.

In order to characterize the time series more specifically, we examined the series of noise from the logistic mapping in terms of the following symbolic analysis. Taking a new variable defined by $y_i = x_i - \langle x_i \rangle$, we count the run length l of the positive sequences in y_i and make its histogram $n(l)$, which decays generally exponentially as

$$n(l) \sim \exp(\beta l). \tag{8}$$

Here the decay rate β is a measure of homogeneity and persistence of noise. When there is no correlation between successive events, the problem can be reduced to an ideal coin tossing, i.e., $\beta = -\ln 2 = -0.693\ldots$. The dependence of β on a is shown in Fig. 4 for the logistic map f and for the threefold iterated map f^3. While the decay rate of correlation for f barely changes form the case of independent coin tossing, β for f^3 largely depends on a, where the qualitative tendency of β, if we exclude periodic states, resembles the performance plots as shown in Fig. 3. We also studied further iterated maps and found again that the slow decay of correlation there could improve the performance, though not as much as that for f^3. Therefore the distribution function of these sojourn times is, most likely, crucial for the network performance. It is clear that a strongly correlated sequence of noise helps to kick the state out of local minima. It is noted that the correlation of the chaotic noise decreased P_{OS} below a_{BM}. The reason for this would be due to the fact that the sign of y_i alternates (quasi)periodically and that the effect of the noise is essentially annihilated. The fact that P_{OS} has a maximum in case 3 is related to the distribution of the noise produced from the logistic map. The distribution of the noise is another issue for the network performance, but it is not the subject of the present paper.

There have been more complex chaotic networks so far proposed for the purpose of improving efficiency to solve the optimization problem [4,5]. Some characteristics of them are, for example, as follows. (i) The individual neural unit (map) behaves as a chaotic oscillator whose behavior is very sensitive to a control parameter. (ii) The parameter of the map is also a dynamical variable. These features of the network create new aspects and possibly additional improvement of the efficiency in the optimization tasks. However, it is harder to analyze and comprehend the role of chaos in these systems, as the energylike function is not defined there. The present research was intended to clarify the role of chaotic noise in the global dynamics of a simple conventional neural network designed to solve optimization problems. We showed that the efficiency of the solution can be raised by adding chaotic noise as high as those of more complex systems, and clarified the role of chaotic noise in a simple neural network.

We would like to acknowledge helpful discussions with M. Sano, T. Hondo, Y. Hayakawa, and S. Sato.

[1] W. J. Freeman, Biol. Cybern. **56**, 139 (1987).
[2] I. Tsuda, E. Koerner, and H. Shimizu, Prog. Theor. Phys. **78**, 51 (1987).
[3] K. Aihara, T. Takabe, and M. Toyota, Phys. Lett. A **144**, 333 (1990).
[4] M. Inoue and A. Nagaoshi, Phys. Lett. A **158**, 373 (1991).
[5] H. Nozawa, Chaos **2**, 377 (1992); H. Nozawa, Institute of Electronics, Information and Communication Engineers Technical Report No. NLP92-36, 1992 (unpublished) (Japanese).
[6] P. J. M. van Laarhoven and E. H. L. Aarts, *Simulated Annealing: Theory and Applications* (Reidel, Dordrecht, 1987).
[7] M. O. Magnasco, Phys. Rev. Lett. **71**, 1477 (1993).
[8] T. Hondou, J. Phys. Soc. Jpn. **63**, 2014 (1994); T. Hondou and Y. Sawada (unpublished).
[9] P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[10] J. J. Hopfield and D. W. Tank, Biol. Cybern **52**, 141 (1985).
[11] Y. Akiyama, A. Yamashita, M. Kajiura, Y. Anzai, and H. Aiso, Institute of Electronics, Information and Communication Engineers Technical Report No. MBE88-183, 1988 (unpublished) (Japanese); Y. Akiyama, A. Yamashita, M. Kajiura, Y. Anzai, and H. Aiso (unpublished).