Stability of oscillatory retrieval solutions in the oscillator neural network without Lyapunov functions

Satoki Uchiyama[1,*] and Hirokazu Fujisaka[2]
[1]*Department of Applied Mathematics, Hiroshima University, Higashi-Hiroshima 739-8527, Japan*
[2]*Department of Applied Analysis and Complex Dynamical Systems, Kyoto University, Kyoto 606-8501, Japan*
(Received 6 April 2001; revised manuscript received 26 September 2001; published 25 June 2002)

Constructing a Ginzburg-Landau map neural network, we analyze its storage capacity with an equilibrium theory of the self-consistent signal-to-noise analysis (SCSNA); however, the prediction does not consist with the simulation results just in the parameter region where the characteristic of the non-Lyapunov-function system gets enhanced [J. Phys. A **32**, 4623 (1999)]. It is expected that this inconsistency comes from the fact that the dynamics of retrieval and nonretrieval states governs the phase transition. Alternatively, we investigate its storage capacity with the help of the Amari-Maginu-Okada theory, a dynamical theory, for the stability analysis of dynamical states. We consequently found that the theory predicts dynamical states quite well especially in the region where the SCSNA breaks down, and that the phase diagram coincides quantitatively well with the simulation results.

DOI: 10.1103/PhysRevE.65.061912 PACS number(s): 87.10.+e, 05.70.Fh, 05.45.Xt, 07.05.Mh

I. INTRODUCTION

Neural networks without Lyapunov functions serve as fascinating systems that can exhibit nonrelaxation dynamics in relation to the weakly destabilized multistable structure of attractors. Generally, in contrast to the case of Hopfield-type models, oscillator neural networks have long characteristic time scales of retrieval dynamics, because oscillator variables move smoothly around in the wide range of continuous space. Hence, studying oscillator systems is advantageous for discovering a variety of nonrelaxation dynamics in long time scales.

The Noest model [1] has worked as the standard model of oscillator neural networks in the last decade. The Hopfield model [2] can be interpreted as a physical system constructed with Ising spins, whereas the Noest model corresponds to its generalization of having the constituent of XY spins. This system has a definite Lyapunov function, which guarantees its monotonous relaxation dynamics. Therefore, it is possible to analyze the property of its equilibrium states with the help of a static theory so far available, namely, the replica method [3] or self-consistent signal-to-noise analysis (SCSNA) [4]. These powerful tools have enabled us to derive many interesting results even for its variant models [5–11] that have these functions or something similar.

On the other hand, the theoretical study of oscillator neural networks without Lyapunov functions has not progressed greatly yet. Nonexistence of the functions means no guarantee of monotonic relaxation dynamics and eventually leads to quite complicated temporal evolutions. Static theories may pick up the static properties of dynamical systems [12,13]. However, it cannot describe the nonrelaxation dynamics, which is a prominent characteristic of such systems.

In order to deal with these systems properly it is indispensable to use dynamical theories, among which the Amari-Maginu-Okada theory (AMO, statistical neurodynamics theory) [14,15] has been found to be useful especially for the discrete dynamics of Hopfield-type models. For example, the retrieval dynamics with a nonmonotonic transfer function [16] and the case with sparse coding [17] have been analyzed successfully. The AMO theory can be extended easily to Noest-type models [18], and the resulting theory has been employed for its variants [19,20]. Nevertheless, in spite of its wide applicability, it has not worked yet to reveal the characteristic dynamics observed possibly in the oscillator neural networks without Lyapunov functions.

The *Ginzburg-Landau* (GL) *oscillator neural network*, namely, a neural network of oscillators each of which is described by the complex GL equation is one promising example of the system without Lyapunov functions. The differential dynamics of this system is obviously beyond the description with the Lyapunov function [21]. To our knowledge, the numerical simulations of these systems have not been extensively carried out yet, and its numerically obtained phase diagram has not been reported. The system does not have the mathematical form like the Noest model at all, so it seems impossible to extend either SCSNA [22] or the AMO theory to the analysis.

We have thus previously presented a method to reduce the coupled oscillator system into a coupled map system, and have derived the *GL map neural network* from the GL oscillator neural network [23,24]. The reduced system has the same mathematical form to that of the Noest model except for the presence of a parameter C_2, which measures the degree of nonexistence of Lyapunov functions. This readily encourages the use of SCSNA. However, because the dynamics of a reduced system does not obey the description of any Lyapunov function as well as the original GL oscillator system, SCSNA does not perfectly elucidate the results of numerical simulations [24,25]: Unless $C_2 \lesssim 1$, where ours is quite close to the Noest model, the retrieval-nonretrieval transition occurs beyond the applicability of equilibrium

*Email address: uchiyama@amath.hiroshima-u.ac.jp

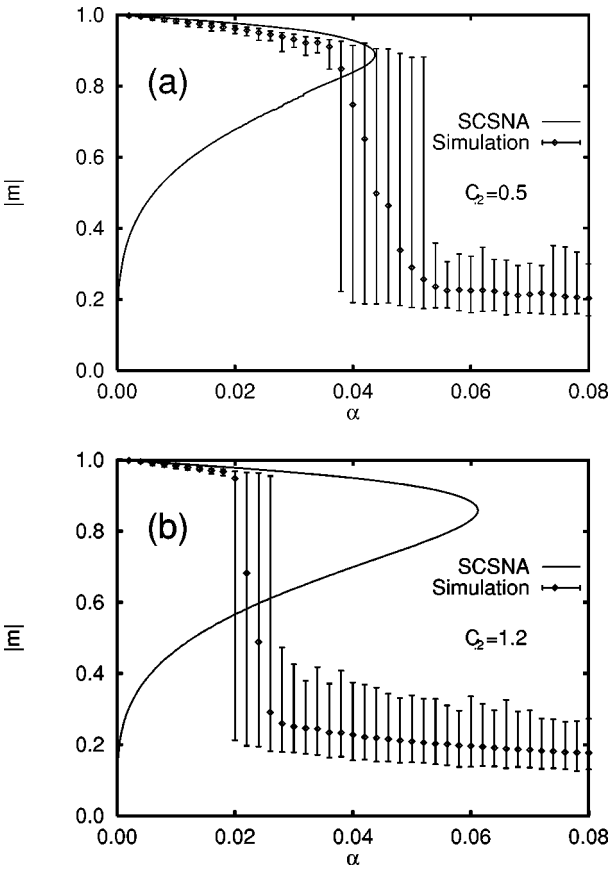

FIG. 1. The SCSNA for the m-α curves of our model with the corresponding numerical simulations (system size $N=2000$ and ensemble number 50): (a) $C_2=0.5$ and (b) 1.2. The symbol $\diamondsuit$ and the error bar stand for the mean value and range of $|m|$, respectively. For the detail, one should refer to Ref. [25] and the cases of $C_2=0.0$ and 1.0 therein. The definitions of α and m will be given also in Sec. III of the present paper.

statistical-mechanical approach, and the prediction of static theory eventually turns out to be meaningless. For the SCSNA results, see, e.g., Figs. 5(a) and 6 in Ref. [25]. See also Fig. 1. This SCSNA breaking, however, illuminates the existence of highly dynamical regions. Then we alternatively resorted to the AMO theory to improve the prediction in such regions. A fundamental aim of this paper is to study that.

The paper is organized as follows: In Sec. II, we briefly present the GL map neural network [24,25]. The AMO theory is expanded in Sec. III, particularly for the analysis of our model. We predict in Sec. IV the dynamical behavior with the theory, and compare it with that of the numerical simulations. Finally, we devote Sec. V to relating the results of AMO to SCSNA, and moreover to discuss the nonrelaxation dynamics with intermittent chaos.

II. MODEL

We will first briefly review the GL oscillator network as it is the fundamental model for this paper. Each uncoupled GL oscillator is represented by a complex variable W, and its time evolution is governed by the complex Ginzburg-Landau equation

$$\dot{W}=(1+iC_0)W-(1+iC_2)|W|^2W. \tag{2.1}$$

Real parameters C_0 and C_2 characterize the dynamical property of an element: The variable is necessarily attracted to a limit cycle $|W|=1$ with a frequency C_0-C_2. So we call this element a limit-cycle oscillator. Note here that the equation has no essential change under a rotational transformation $W\rightarrow W\exp(i\Theta)$, where Θ is a phase shift. That allows for an arbitrary choice of C_0. Connecting N GL oscillators through $N\times N$ random two-body interactions, we get the GL oscillator neural network [21]:

$$\dot{W}^{(j)}=(1+iC_0)W^{(j)}-(1+iC_2)|W^{(j)}|^2W^{(j)} + \sum_{k=1}^{N} D_{jk}(W^{(k)}-W^{(j)}) \quad (j=1,\ldots,N). \tag{2.2}$$

Here D_{jk}'s are complex coupling coefficients, and have random values with respect to indices j and k. The above system would work as an oscillator neural network with an appropriate set of couplings D_{jk}.

We previously proposed a coupled map system effectively equivalent to the above system of coupled differential equations [23,24]. Defining new phase variables $\theta^{(j)}$ through $\theta^{(j)}\equiv \arg W^{(j)}-C_2\ln|W^{(j)}|$, we can reduce the original system to the following simple one described by a set of N phases:

$$e^{i\theta_{t+1}^{(j)}}=f(h_t^{(j)},h_t^{(j)*}), \tag{2.3a}$$

$$h_t^{(j)}=\sum_{k=1}^{N} J_{jk}e^{i\theta_t^{(k)}} \quad (j=1,\ldots,N), \tag{2.3b}$$

$$f(h,h^*)=h|h|^{-1-iC_2}. \tag{2.3c}$$

Here we have used $C_0=C_2$ according to the rotational symmetry. The lower suffix t indicates discrete times ($t=0,1,2,\ldots$), and the upper j a site of oscillators. The state $\theta_t^{(j)}$ of jth oscillator is updated through a complex transfer function $f(,)$, i.e., the function of a complex local field $h_t^{(j)}$. We define the system such that all the elements update from t to $t+1$ simultaneously. The coupling coefficients J_{jk} have random static values as well as D_{jk}. An asterisk stands for complex conjugate. We call this the GL map neural network.

Now we consider how to embed $p=\alpha N$ random phase patterns $\boldsymbol{\phi_1},\ldots,\boldsymbol{\phi_p}$ into the couplings J_{jk}; or in other words, how to make the system such that the system state $\boldsymbol{\theta_t}=(\theta_t^{(1)},\ldots,\theta_t^{(N)})$ can be attracted to one of the embedded p patterns $\boldsymbol{\phi_\mu}=(\phi_\mu^{(1)},\ldots,\phi_\mu^{(N)})(\mu=1,\ldots,p)$, depending on its initial state. Actually, we are able to derive the matrix J from D [23,24], but in this paper, we treat J as a primarily given matrix by

$$J_{jk}=\frac{1}{N}\sum_{\mu=1}^{p}\exp[i(\phi_\mu^{(j)}-\phi_\mu^{(k)})] \quad (j,k=1,\ldots,N). \tag{2.4}$$

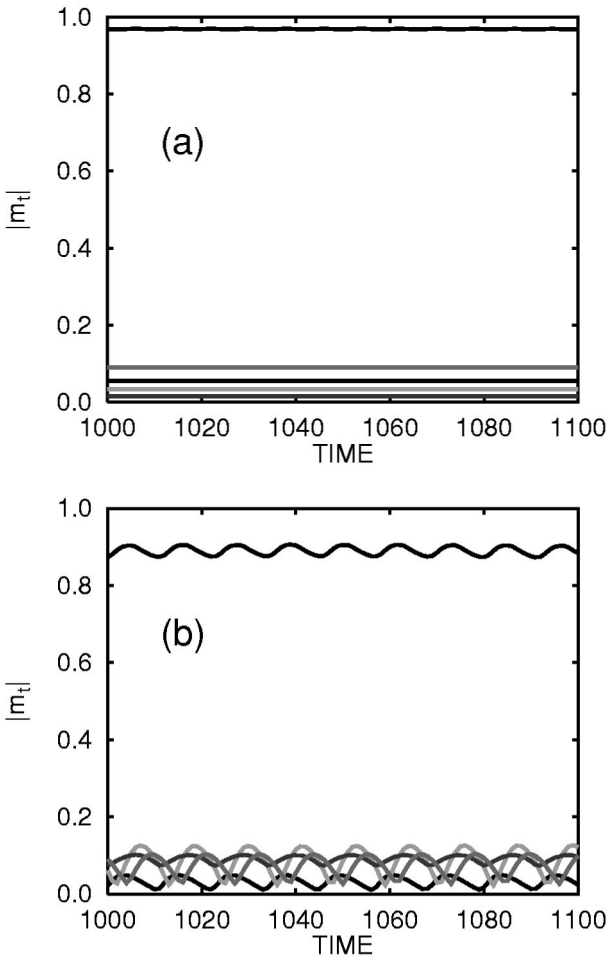

FIG. 2. Time evolutions of overlaps in numerical simulations for $N=1000$. The transient processes were removed here. Only five overlaps are drawn for simple illustration: Four represent $p-1$ nonretrieved patterns and the rest a retrieved one. Parameters are (a) $\alpha=0.035, C_2=0.85$ in the dark region of Fig. 5(a) of [25], and (b) $\alpha=0.040, C_2=1.05$ slightly inside the dark region of the same figure.

This form of the coupling is called the *generalized Hebb rule* [1,22]. In order to keep a simple mathematical perspective, ϕ is assumed to take random values in the interval $[-\pi,\pi)$. Then we can statistically neglect the correlations between patterns $\langle \phi_\mu^{(j)} \phi_\nu^{(j)} \rangle (\mu \neq \nu)$ and between sites $\langle \phi_\mu^{(j)} \phi_\mu^{(k)} \rangle (j \neq k)$. Under the assumption of random phase patterns, the validity of the embedding rule can be clarified with the help of the simple signal-to-noise analysis [25].

It is worth noting that a parameter C_2 plays a crucial role for the dynamical character of the present system. In case of $C_2=0$, the system reduces to the Noest model. Such a coincidence easily suggests the extension of the analytical methods developed for the Noest model ($C_2=0$) into the present case ($C_2 \neq 0$). Recall that C_2 comes from the coefficient of nonlinear term in the complex GL equation (2.1). So, changing it from zero to a finite value, we can continuously modify a Lyapunov-function system into a non-Lyapunov-function one. We can observe such a modification of the dynamical behavior in Fig. 2.

III. DYNAMICAL EQUATIONS OF MACROSCOPIC QUANTITIES

The main scope of this paper lies in studying how extensively our system can store many patterns in the limit of $N \to \infty$. We now will give some definitions for the scope. First, we define the overlap between a pattern and the current state of the system to be

$$m_t^\mu = \frac{1}{N} \sum_{j=1}^{N} \exp[i(-\phi_\mu^{(j)} + \theta_t^{(j)})] \quad (\mu = 1, \ldots, p). \tag{3.1}$$

Using these p macroscopic quantities, a retrieval state is characterized as both $m_t^\mu \equiv m_t \sim O(1)$ and $m_t^\nu \sim O(1/\sqrt{N})(\nu \neq \mu)$, in which the system retrieves the μth pattern, and a nonretrieval state as $m_t^\nu \sim O(1/\sqrt{N})(\nu = 1, \ldots, p)$. The difference between retrieval and nonretrieval states is just the appearance of an $O(1)$ overlap. We especially call this overlap the retrieval quality. Then we expect that, when the loading rate $\alpha \equiv p/N$ exceeds an upper limit, the system undergoes the retrieval-nonretrieval transition, and the retrieval quality m_t vanishes suddenly. This upper limit is termed the storage capacity α_c. The final goal of the paper is to derive the dynamical equations describing temporal evolution of macroscopic quantities based upon the statistical neurodynamics theory, and to determine the C_2 dependence of α_c.

We need dynamical equations of macroscopic quantities to theoretically predict the time series of m_t. Now we will give a brief derivation of the equations. Suppose that the system retrieves a pattern $\mu=1$ with the condition of fixed α and C_2. There is only one overlap of $m_t = O(1)$ that corresponds to the pattern $\mu=1$, namely, $m_t = m_t^1$. Thus N local fields are written as

$$h_t^{(j)} = e^{i\phi_1^{(j)}} m_t + z_t^{(j)} \quad (j = 1, \ldots, N).$$

Here we have defined $z_t^{(j)} = \Sigma_{\mu \neq 1}^{p} e^{i\phi_\mu^{(j)}} m_t^\mu$. This complex variable is calculated as the summation of the products of a random factor $e^{i\phi}$ and $O(1/\sqrt{N})$ overlaps. We regard it as noise in comparison with the first term of the above. The Amari-Maginu-Okada theory assumes its Gaussian form $N(0,\sigma_t)$ at every discrete time t. If the value of σ_t is known, it is easy to understand that the time evolution of retrieval quality obeys $m_{t+1} = \langle e^{-i\phi} f(e^{i\phi} m_t + z_t, \text{c.c.}) \rangle_{z,\phi}$, which is just the integral form of Eq. (3.1). For the detail, see Eq. (3.2) below. Moreover if σ_{t+1} was calculated only from σ_t and m_t, then we could obtain a closed set of the evolutions of two macroscopic quantities.

In order to arrive at the closed formula, we actually have to introduce auxiliary quantities as well as in the case of SCSNA formula. For $C_2=0$, the derivation of AMO equations has already been presented by Aoyagi and Kitano [18,19]. It is quite easy to extend the formulation to the case of $C_2 \neq 0$ along the line of their work. So we omit the detailed derivation in this paper. Resultingly, the AMO equations for GL map neural network are given as

$$m_{t+1}=\int_{-\infty}^{\infty} D\zeta f(m_t+\sigma_t\zeta), \tag{3.2}$$

$$U_{t+1}=\frac{1}{\sigma_t}\int_{-\infty}^{\infty} D\zeta\zeta^* f(m_t+\sigma_t\zeta), \tag{3.3}$$

$$\sigma_{t+1}^2=\alpha+|U_{t+1}|^2\sigma_t^2+\alpha\left(\sum_{\tau=t+1-n}^{t} Q_{t+1,\tau}\prod_{\kappa=\tau+1}^{t+1} U_\kappa+\text{c.c.}\right), \tag{3.4}$$

where

$$D\zeta\equiv\frac{1}{\pi}\exp(-\zeta^2)d\zeta,\quad \zeta=\xi+i\eta. \tag{3.5}$$

We have abbreviated $f(h,h^*)$ as $f(h)$ in the above. It should be emphasized that the right-hand side of Eq. (3.4) for noise variance includes the memory effect from the past n time steps. We call n the approximation order of statistical neurodynamics; the dynamical nature of the system properly determines the appropriate order. We have introduced, in Eq. (3.4), a macroscopic quantity Q as the cross correlation of outputs $e^{i\theta^{(j)}}$, namely, $Q_{t+1,\tau}\equiv\langle(e^{i\theta_t^{(j)}})^* e^{i\theta_{\tau-1}^{(j)}}\rangle_j$. Using another new macroscopic quantity $C_{t,\tau-1}\equiv\langle z_t^{(j)*} z_{\tau-1}^{(j)}\rangle_j$, i.e., the cross correlation of noises $z^{(j)}$, the definition of Q is explicitly written in the following integral form:

$$Q_{t+1,\tau}=\int\int\int_{-\infty}^{\infty} DaDbDc f^*(m_t+\sigma_t(d_0a+d_1c)) \times f(m_{\tau-1}+\sigma_{\tau-1}(d_0b+d_1c)). \tag{3.6}$$

Here we have put $d_0=\sqrt{1-C_{t,\tau-1}/(\sigma_t\sigma_{\tau-1})}$ and $d_1=\sqrt{C_{t,\tau-1}/(\sigma_t\sigma_{\tau-1})}$. By taking the memory terms of the past N steps into account, the noise correlation is calculated as well as the autocorrelation (3.4):

$$C_{t,\tau-1}=\alpha Q_{t,\tau-1}+U_t^* U_{\tau-1}C_{t-1,\tau-2} +\alpha\sum_{\chi=t-n+1}^{t-1} Q_{\chi,\tau-1}\prod_{\kappa=\chi+1}^{t} U_\kappa^* +\alpha\sum_{\chi=t-n+1}^{\tau-2} Q_{t,\chi}\prod_{\kappa=\chi+1}^{\tau-1} U_\kappa \quad (t-n+3\leqslant\tau\leqslant t, n\geqslant 3), \tag{3.7a}$$

$$C_{t,\tau-1}=\alpha Q_{t,\tau-1}+U_t^* C_{t-1,\tau-1}\quad(\tau=t-n+2, n\geqslant 2), \tag{3.7b}$$

$$C_{t,\tau-1}=0\quad(\tau=t-n+1, n\geqslant 1). \tag{3.7c}$$

We point out that, in comparison with the case $C_2=0$, only two major modifications for $C_2\neq 0$ appear in the last term of Eq. (3.4) and the third term of Eq. (3.7a). Using the above equations, one can determine all the quantities at $t+1$ recursively from the ones at $t, t-1, \ldots,$ and $t-n+1$. Equations (3.7b) and (3.7c) are useful only when the dynamics is in the early stage near $t=0$, since there is no information at $t=0$ about the past n steps. Thus we need special formulas to avoid that. Only two initial values of macroscopic quantities are necessary to start the AMO calculation, that is, m_0 $(0\leqslant|m_0|\leqslant 1)$ and σ_0 $(\sigma_0\geqslant 0)$. According to the simple estimation, initial variance σ_0 should be set to $\sqrt{\alpha}$. So we can only take the value of m_0 arbitrarily. [Initial values $Q_{0,\tau}(\tau=1,\ldots,n)$ are also necessary; however, because of Eqs. (3.2), (3.6), and (3.7c), we can put $Q_{0,\tau}=m_0^* m_\tau$ for $\tau=1,\ldots,n$.] These two bring forth the time series of all the AMO quantities.

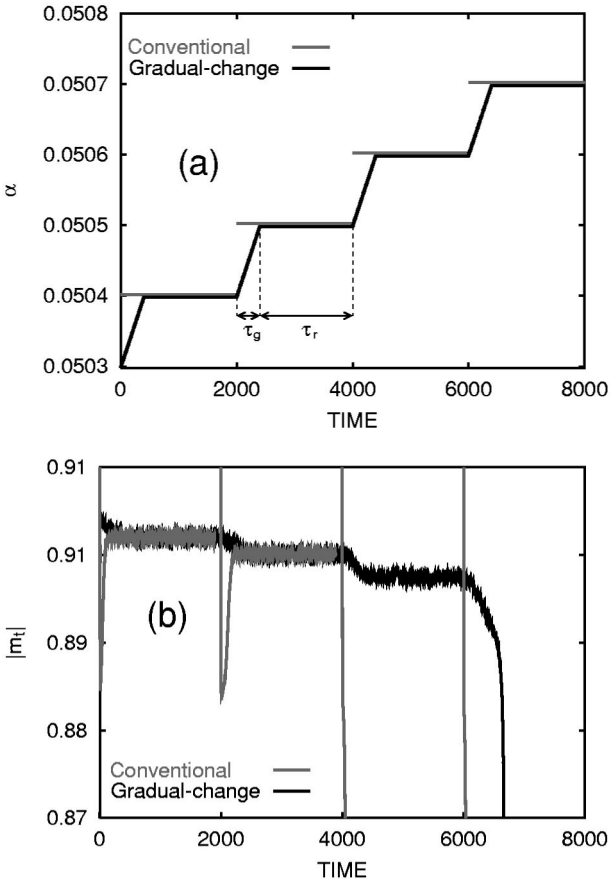

FIG. 3. Comparison between gradual-change algorithm (black lines) and conventional ones (gray) in the case of $n=8$ and $C_2=0.71$. $\tau_g=400$ and $\tau_r=1600$. (a) Profiles of α. A continuous profile corresponds with a gradual-change run, and a discontinuous one with a series of four conventional runs. Note that the time in the conventional ones must be reset to zero at the start of every run. (b) Responses of $|m_t|$ to the profiles. We notice the difference of $\alpha_c^{\text{conventional}}=0.0505$ with $\alpha_c^{\text{gradual}}=0.0506$. The use of long total interval $\tau_g+\tau_r=2000$ makes two nonretrieval solutions around $t=4000$ and 6000 look like vertical gray lines, respectively.

IV. RESULTS

A. Algorithm solving the AMO equations

Throughout this paper, we are interested in the dynamical stability of retrieval solutions, not the theoretical prediction of the whole retrieval process. So we just classify the system as in a retrieval state if the initial overlap $|m_0|\sim 1$ keeps a

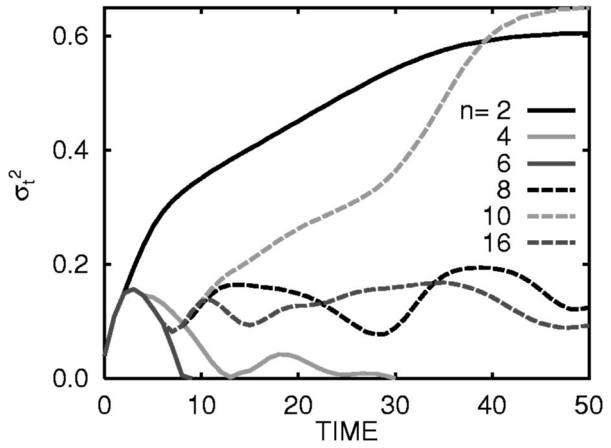

FIG. 4. Time series of σ_t^2 produced by the AMO dynamics. With fixed parameters $C_2=1.4$ and $\alpha=0.04$, we have calculated the series for $n=2$, 4, 6, 8, 10, and 16. Resulting negative σ_t^2 signifies the termination of the theoretical calculation at that time.

finite value after the transient dynamics; otherwise, as in a nonretrieval state where $|m_0|\sim 1$ decays to $m_t=0$. In the AMO theory, it is conventional to reset all the variables to start a run with newly given parameters. This algorithm makes sense for the study of the retrieval process; however, it may turn inappropriate in case of the system without the Lyapunov function. We have tentatively observed that our system almost always cannot be attracted to the AMO retrieval solutions in spite of their existence. According to preliminary results, such a tendency has gotten remarkable, especially for the dynamical retrieval solutions around the phase transition point. [See gray lines in Fig. 3(b).] Hence, the conventional algorithm does not necessarily work well to study the dynamical stability of the retrieval solutions. This results in the underestimation of α_c.

Here we propose an alternative: we try to determine proper transition points with gradually changing α or C_2. Let us explain the algorithm in more detail. First, we prepare a retrieval solution by the conventional algorithm. Then, we switch the algorithm to the alternative: changing a parameter gradually (gradual-change interval τ_g) and then keeping it for a while (relaxation interval τ_r), we observe whether a solution becomes destabilized or not. [See black lines in Fig. 3(b).] If not destabilized, we repeat this procedure without the reset of variables till the solution turns to be unstable. Typically, the changes per the gradual-change interval are $\Delta\alpha=10^{-6}$ and $\Delta C_2=10^{-5}$; the relaxation interval 10^2 steps, and the gradual-change one $10^2\sim 10^3$ steps. Hereafter we pay our attention only to the existence and stability of the AMO solution, and we do not touch on the basin of retrieval solutions theoretically.

B. Retrieval-nonretrieval transition

First of all we should answer the question of how many degrees n is necessary for the proper AMO description in the GL map neural network; however, we have not had any *a priori* answer to this question so far. Thus we look for the answer in the result of a numerical simulation. We observed

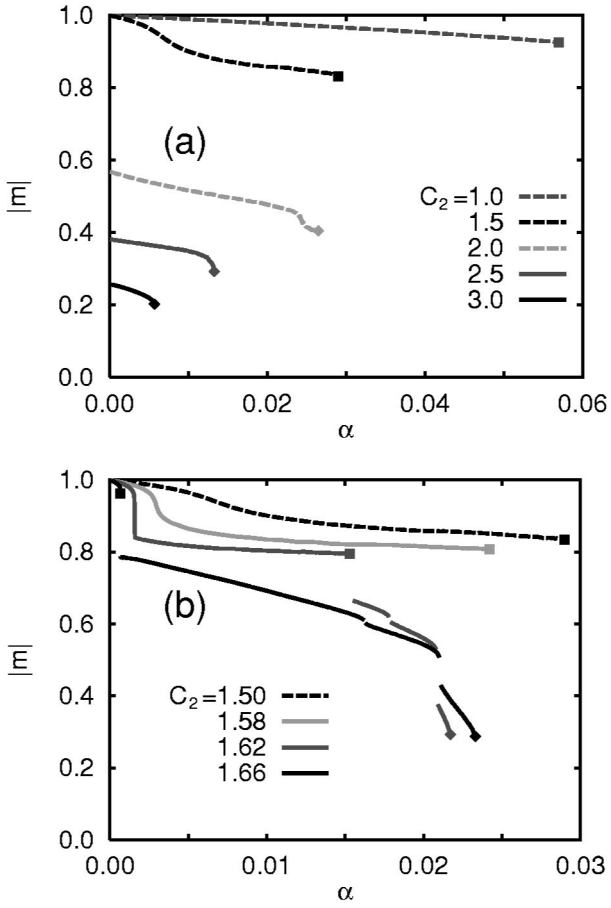

FIG. 5. Retrieval quality vs α predicted by AMO. The calculations were carried out with $n=8$ and representative values of C_2. The ordinate indicates the time average of the retrieval quality, in which the transients are removed. (a) $C_2=1.0$, 1.5, 2.0, 2.5, and 3.0. (b) $C_2=1.50$, 1.58, 1.62, and 1.66. These four values are in the range where the relation drastically changes as shown in (a). Symbols □ and ◇ stand for the positions of α_c and α_c^*, respectively.

the time series of σ_t^2 for various n's under the condition of fixed α and C_2 as in Fig. 4. The typical value of σ_t^2 is around 0.1 for the given parameters, provided that the system runs in a retrieval state. For an inappropriate n, the variance either grows up larger than the typical value, or goes to a negative value. The former indicates an undesirable relaxation to a nonretrieval solution despite the existence of a retrieval one, and the latter the breaking of the AMO calculation. As one can see in this figure, a small increase of the degree from $n=2$ does not necessarily yield the improvement of the approximation.

We eventually regard $n=8$ as the smallest meaningful degree: For $n<8$, the phase diagram becomes meaningless even in the qualitative comparison. Figure 5(a) shows the dependence of time averaged $|m_t|$ on α for several fixed C_2's. Figure 5(b) depicts the detailed behavior in the range of $1.5\leqslant C_2\leqslant 1.66$. Observing the branches for a fixed C_2, one can easily distinguish between a primary branch, which continuously starts with $|m|=1$ in $\alpha\to 0$, and subordinate ones. Thus let us redefine α_c as the transition point at which a primary branch changes into a subordinate one or $|m|=0$

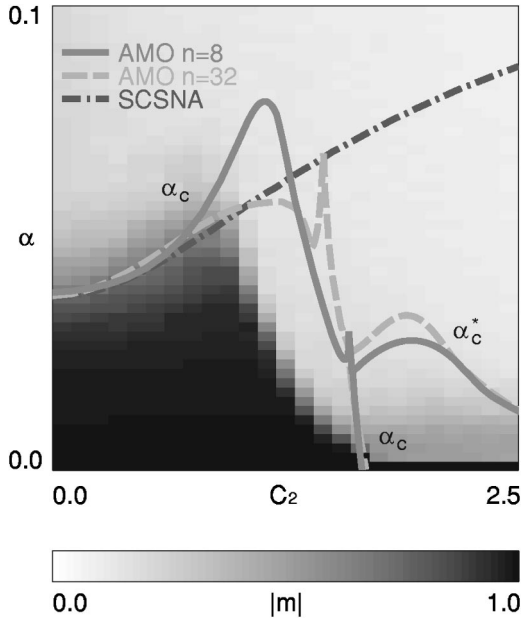

FIG. 6. Phase diagram predicted by the statistical neurodynamics. Solid curves mean the transition points for $n=8$, and dotted lines for $n=32$. The SCSNA result [25] is drawn for comparison. A gray scale indicates the numerical simulation result of retrieval qualities.

and, in addition, define α_c^* as the maximum transition point at which a subordinate one vanishes into $|m|=0$. The introduction of α_c^* will soon be necessary to know whether a theoretically obtained retrieval region consists of a set of primary branches or not. In general, the absolute values of overlaps at α_c^*'s are much smaller than those at α_c's as shown in Fig. 5(b). These two points for $n=8$ are plotted in Fig. 6 (solid curves). This figure also presents, with a gray scale, the ensemble-averaged time average of $|m_t|$ computed by numerical simulations. It is obvious that AMO shows the better prediction of α_c than SCSNA; on the other hand, the numerically obtained retrieval region does not seemingly trace out the boundary α_c^*. This may relate to the fact that, in respect of the simulation data, the basin of the retrieval solutions is smaller than that of trivially existing nonretrieval ones below α_c^*. The nonretrieval solutions ($|m_t|\sim 0$) reduce the contribution of retrieval ones ($|m_t|>0$) on the ensemble averaged $|m|$ and consequently decrease its value. In contrast to the case of α_c^*, numerical simulations reveal that the region below α_c has negligibly small basins to the trivially existing nonretrieval solutions and that such contradiction is not observed at all.

In order to investigate n dependence of the results, we calculated the critical points for $n=32$ in the same way. The result is plotted also in Fig. 6 with dotted curves. One can easily find that the storage capacity shifts to below the SCSNA prediction and approaches the simulation results, as n increases to 32. There is a parameter region where the dotted curve of $n=32$ is much lower than the solid of $n=8$ ($0.7 \lesssim C_2 \lesssim 1.3$). This simply implies the improved approximation by the increase as well as in case of $C_2=0$ [18,19]. On the other hand, the opposite region ($1.3 \lesssim C_2 \lesssim 2.1$) can be explained as follows: First, the increased n admits the oscillatory-retrieval solutions with longer periods to survive in the retrieval region. Such contributions appear near the peak of α_c^*, but they are destroyed in case of $n=8$ because of the shortage of the approximation degree. That finally results in the lower peak of α_c^* for $n=8$. How extensively does the peak height get enhanced with further increase of n, and how does the increase deform the $n=32$ sharp peak of α_c (on $C_2\sim 1.45$)? It seems quite difficult to answer these questions unless we rely on the calculation with extremely large n, namely, with enormous computation time.

Actually, there exists another difficulty in the calculation: We have made α_c^* peaks for $n=8$ and 32 by means of interpolation. Even the gradual-change algorithm cannot produce smooth boundaries just around these peaks. We attribute the difficulty to the riddled structure of the parameter region in which retrieval solutions periodically oscillate with extremely long periods; hence, we have failed to make the smoothness. We are not able to prepare enough data to draw the zigzag boundaries, though possibly fractal ones. So we have exploited the interpolations.

C. Oscillatory-retrieval phase

Although we have mentioned ''oscillatory-retrieval dynamics'' above, we have not still investigated it extensively. Now let us explore its characteristic. It is quite natural to classify the retrieval solutions into two categories: (1) the fixed-point-type retrieval solution and (2) the oscillatory one. In order to distinguish between these, we focus on the temporal behavior observed in the time series of $|m_t|$. We hereafter deal only with the time series after transients, if not mentioned.

One may define as an oscillatory retrieval the solution that has the finite time average of $|m_t|$ (for the sign of a retrieval state) and the finite amplitude of fluctuated $|m_t|$ (for the sign of an oscillatory one), but this definition is not practical. Bear in mind that the AMO calculation contains the numerical integrals of macroscopic quantities. The integrands have singular points that correspond to a zero local field. (For example, see spirals in Figs. 7 and 8 of Ref. [25].) This numerically results in imperfect convergence of the integrals and, eventually, in a time-dependent artificial noise to the essential dynamics. So when we observe AMO solutions, it is sometimes difficult to make a distinction of the intrinsic oscillation from the artificial ''fluctuation.'' Therefore, we employed the following criterion: The oscillatory solutions can be identified when its amplitude is greater than 0.01. Accordingly, together with the retrieval criterion of $|m|_{\min} \geqslant 0.05$, the AMO solutions are classified as

Stationary retrieval: $|m|_{\min} \geqslant 0.05$ and $(|m|_{\max} - |m|_{\min})/2 < 0.01$,

Oscillatory retrieval: $|m|_{\min} \geqslant 0.05$ and $(|m|_{\max} - |m|_{\min})/2 \geqslant 0.01$,

Nonretrieval: $|m|_{\min} < 0.05$.

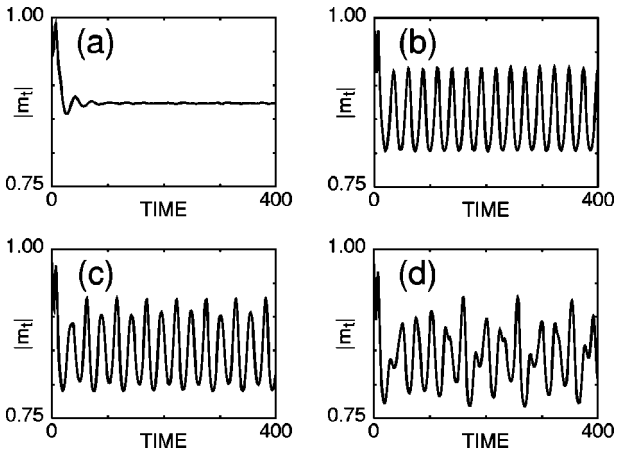

FIG. 7. Bifurcation behavior of a retrieval solution by AMO with $n=8$ and $C_2=1.5$: (a) $\alpha=0.015$, the fixed-point-type retrieval; (b) $\alpha=0.023$, the periodically oscillatory retrieval; (c) $\alpha=0.025$, the period doubling of (b); (d) $\alpha=0.028$, an oscillatory-retrieval solution with aperiodicity.

Here $|m|_{\min}\equiv\min_{t\in\mathbb{N}}(|m_t|)$, etc. The value 0.01 of the oscillation criterion seems too large, however, this works well to remove the artificially fluctuated solutions from the intrinsic oscillations.

On the other hand, we have adopted the value 0.05 as the lower limit of (nonaveraged) $|m_t|$ to judge whether the system keeps a retrieval state. This implies that, once $|m_t|$ goes below 0.05, it should be treated as a nonretrieval state and the calculation should be terminated at that time. It has been commonly reported that AMO successfully predicts the quantities only when the system has a finite overlap, i.e., the system in a retrieval state [15–19]. So we need not divide nonretrieval solutions into ''stationary type'' and ''oscillatory one.'' We should regard both just as nonretrieval type. It is incidentally remarked that, if we did not use this criterion, we would be bothered by the existence of the ''exotic'' AMO solutions in which $|m_t|$ alternately grows up to a finite value and relaxes to nearly zero.

Before discussing the diagram of dynamical phases, we first observed the bifurcation behavior from fixed-point-type retrieval to oscillatory retrieval. The retrieval processes for different α's are shown in Fig. 7, where $n=8$ and $C_2=1.5$. When α is small, a retrieval solution relaxes to the fixed-point type as usual [Fig. 7(a)]. One may perceive the narrow oscillation of $|m_t|$, but we regard this solution as the fixed-point type because its amplitude is less than 0.01. When α is increased slightly, this fixed-point-type retrieval solution undergoes a *Hopf bifurcation* and changes to a periodically oscillatory retrieval one [Fig. 7(b)]. With the further increase, one finds the *period-doubling bifurcation* as in Fig. 7(c). When α is increased furthermore, the dynamics of $|m_t|$ eventually shows an aperiodically oscillatory-retrieval solution as shown in Fig. 7(d). Since the system consists of identical N limit-cycle oscillators, the emergence of aperiodicity is quite nontrivial. Finally, as one can expect in neural network systems, the retrieval solution settles down to a nonretrieval one with a further large α.

It is worth stressing that this bifurcation behavior reminds us of the *period-doubling route to chaos* in the *logistic map*. To confirm this statement quantitatively, we made the bifurcation diagram by plotting the extrema of $|m_t|$ as shown in Fig. 8. These figures excellently display the period-doubling cascade till the first-order period-doubling points (the second branching points). The reason why the curves become blurred is explained with the matter of numerical convergence of integrals. Such blurred curves prevent us from determining accurately the second-order period-doubling points, which are indispensable for the tentative calculation of the *Feigenbaum ratio* [28]. Accordingly, it seems quite difficult to estimate the onset points of chaotically-oscillating retrieval solutions. Using the time-series data for Fig. 8, we analyzed the α dependence of the amplitude and frequency of the oscillatory retrieval solutions. The results are plotted in Figs. 9 and 10, which also suggests a part of its bifurcation phenomena.

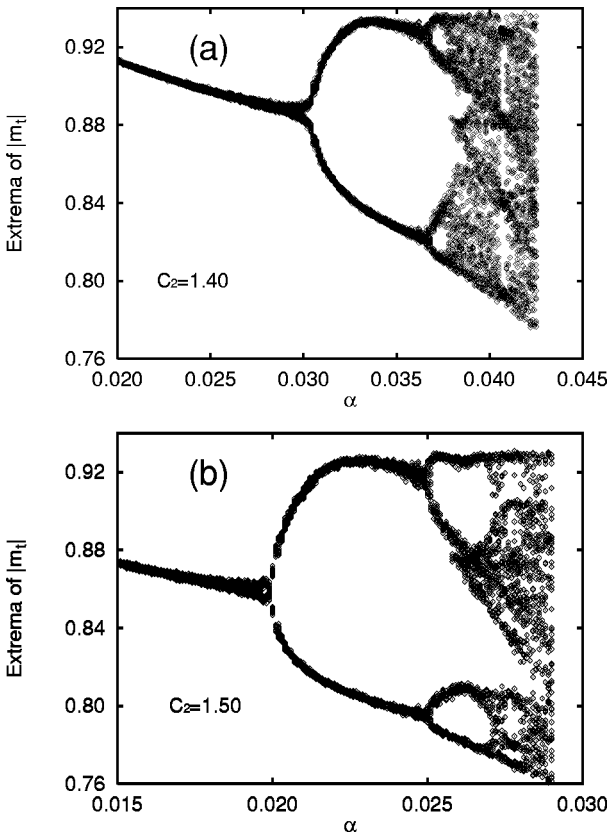

FIG. 8. Loading rate vs extrema in the time series of theoretically produced $|m_t|$ for $n=8$, (a) $C_2=1.4$, and (b) $C_2=1.5$. The transient processes are removed from the series. (All the results in Figs. 8, 9, and 10 are made from the same set of data.) A Hopf bifurcation point corresponds with the first bifurcation point and a period-doubling one with the second point in these figures. It is unavoidable that the technical problem of numerical convergence blurs theoretical curves. The right ends of the cascades imply the retrieval-nonretrieval transition points.

We explored extensively, with the $n=8$ approximation, the Hopf bifurcation points and period-doubling ones for the various choices of parameters. The result is summarized in Fig. 11. In generic retrieval phases, the areas of oscillatory-retrieval regions are created by the Hopf bifurcation line, and

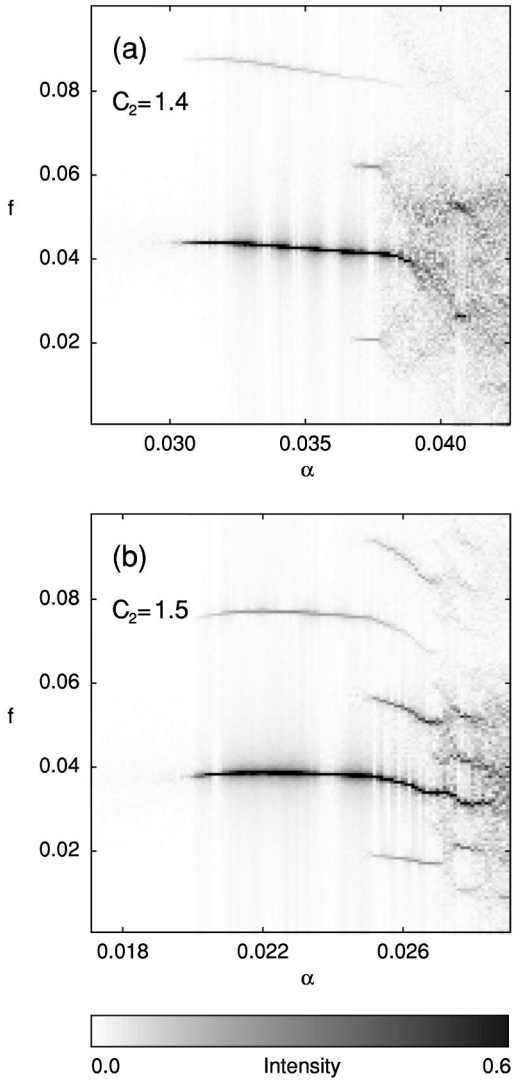

FIG. 9. The α-dependent characteristic of power spectra observed in oscillatory retrieval solutions. The spectra are analyzed from a set of the time series of $|m_t|$ in the interval from 0 to 1600. The intensities at zero frequency are not displayed to emphasize the detailed overall behavior.

are annihilated by the retrieval-nonretrieval transition line. Consequently, the complexity of retrieval dynamics gets enhanced as α approaches the storage capacity.

V. CONCLUDING REMARKS

We have employed the Amari-Maginu-Okada theory to investigate the (generic) retrieval regions of the GL map neural network. Our system has no Lyapunov function. Its retrieval and nonretrieval solutions may fluctuate in the course of time, and thus AMO theory is indispensable to describe the retrieval-nonretrieval transition. Actually, as one can read off in the comparison between Figs. 6 and 11, the regions of the AMO oscillatory-retrieval solutions correspond qualitatively well to the front of the numerically obtained SCSNA breaking regions. We conclude that dynamical solutions, which are definitely beyond the description by the static

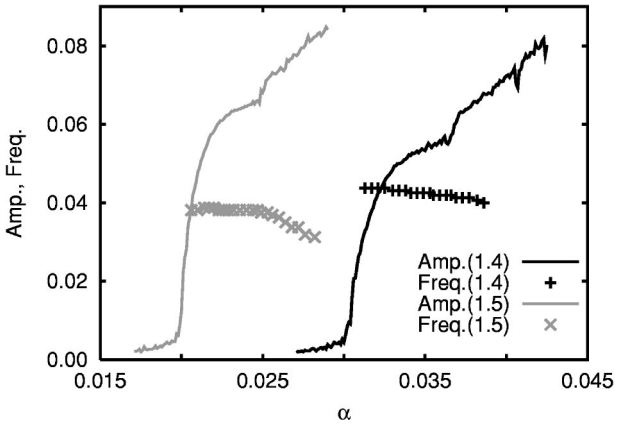

FIG. 10. The α dependence of the amplitude and frequency of the AMO oscillatory-retrieval solutions. A number in the parenthesis above means the value of C_2. We picked up the maximum and minimum values of $|m_t|$ from each series, and plotted (max $-$min)/2 as the amplitude of an oscillatory-retrieval solution. Moreover we defined the characteristic frequencies at which the intensities are greater than 0.6, i.e., nearly half of the maximum intensity, and plotted them, too. The contributions from zero frequency are not taken into account also in this figure. Note that we cannot define the characteristic frequencies in case of the fixed-point-type retrieval solutions and chaotically retrieval ones. Then, it is evident that the the frequency-α curves are shorter than the amplitude-α ones.

theory of SCSNA, governs the retrieval-nonretrieval transition essentially.

Apart from the qualitative agreement, a theoretically determined phase diagram does not converge to the simulation result so much, especially in the region of oscillatory retrieval. The reason may be due to the finite-size effect to the simulation results: Near a critical point, an oscillatory-retrieval solution becomes more sensitive to this effect than a

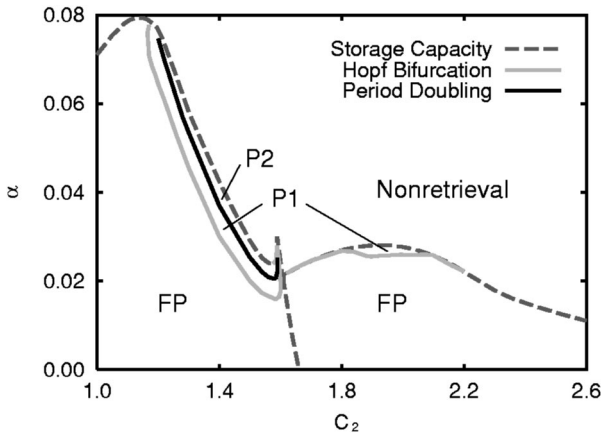

FIG. 11. Oscillatory-retrieval regions determined by AMO with $n=8$. FP refers to the region of fixed-point type [26,27]; $P1$ the periodically-oscillating retrieval ones as in Fig. 7(b); $P2$ the oscillatory-retrieval except P1 type, which includes the solutions shown in Figs. 7(c) and 7(d). The FP-$P1$ boundary and the $P1$-$P2$ one consist of the Hopf bifurcation points and the period-doubling points, respectively.

fixed-point-type one, because the dynamical solution has a number of instantaneous system states to suffer the effect. In order to verify that, we should try to make the simulation data for the larger systems in the future. As shown in Fig. 6, the increase of n does not monotonically improve the coincidence, in contrast to the case of $C_2=0$ [18,19]. So this discrepancy does not closely relate to the matter of approximation degree.

We have identified that the Amari-Maginu-Okada theory reveals the existence of chaotic retrieval solutions in our model; namely, when we observe a time-dependent statistical quantity in a deterministically-driven random system with many degrees of freedom, the quantity can exhibit a kind of chaos in the observation. This implies the generalization of the *collective chaos* found in the globally coupled GL map system [23]. They are open questions to explore the chaotic property of AMO solutions in detail, and to realize them in numerical simulations.

Finally, let us consider carefully the fact that the oscillatory-retrieval region faces the boundary of the retrieval-nonretrieval transition line. This means that we are able to slightly destabilize an oscillatory-retrieval state into a nonretrieval one near the boundary. It is quite fascinating to study what kind of dynamics the system shows there. As is well known in the case of a system with a few degrees of freedom, a weakly destabilized solution plays an important role in the appearance of intermittent chaos [28]. So we wonder if the same situation happens also in the field of neural networks. In other words, there exists the possibility that the multistable structure of neural system is connected with intermittent dynamics. We are studying this scope at present.

ACKNOWLEDGMENT

We thank Evan Lavoie for careful reading and a revision of the manuscript.

[1] A. J. Noest, in *Proceedings of the IEEE Conference on Neural Information Processing Systems, Neural and Synthetic, Denver, 1987*, edited by D. Z. Anderson (AIP, New York, 1987); Europhys. Lett. **6**, 469 (1988); Phys. Rev. A **38**, R2196 (1988).
[2] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 434 (1982); **81**, 3088 (1984).
[3] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. Lett. **55**, 1530 (1985); Ann. Phys. (N.Y.) **173**, 30 (1987).
[4] M. Shiino and T. Fukai, J. Phys. A **25**, L375 (1992); Phys. Rev. E **48**, 867 (1993).
[5] T. Matsuno, Y. Imada, K. Toko, and K. Yamafuji, J. Phys. Soc. Jpn. **63**, 4335 (1994).
[6] K. Park and M. Y. Choi, Phys. Rev. E **52**, 2907 (1995).
[7] T. Aonishi, Phys. Rev. E **58**, 4865 (1998); T. Aonishi, K. Kurata, and M. Okada, Phys. Rev. Lett. **82**, 2800 (1999).
[8] M. Yamana, M. Shiino, and M. Yoshioka, J. Phys. A **32**, 3525 (1999).
[9] T. Aoyagi and K. Kitano, Phys. Rev. E **55**, 7424 (1997).
[10] G. Jongen, D. Bollé, and A. C. C. Coolen, J. Phys. A **31**, L737 (1998); G. Jongen, J. Anemüller, D. Bollé, A. C. C. Coolen, and C. Pérez-Vicente, *ibid.* **34**, 3957 (2001).
[11] S. Kawaguchi, Prog. Theor. Phys. **104**, 709 (2000).
[12] T. Fukai and M. Shiino, Europhys. Lett. **26**, 647 (1994).
[13] T. Fukai and M. Shiino, Neural Comput. **7**, 529 (1995).
[14] S. Amari and K. Maginu, Neural Networks **1**, 63 (1988).
[15] M. Okada, Neural Networks **8**, 833 (1995).
[16] H. Nishimori and Ioan Opriş, Neural Networks **6**, 1061 (1993).
[17] K. Kitano and T. Aoyagi, J. Phys. A **31**, L613 (1998).
[18] T. Aoyagi and K. Kitano, Neural Comput. **10**, 1527 (1998).
[19] K. Kitano and T. Aoyagi, Phys. Rev. E **57**, 5914 (1998).
[20] T. Aoyagi and M. Nomura, Phys. Rev. Lett. **83**, 1062 (1999); M. Nomura and T. Aoyagi, J. Phys. A **33**, 8681 (2000).
[21] T. Aoyagi, Phys. Rev. Lett. **74**, 4075 (1995).
[22] K. Okuda (unpublished).
[23] S. Uchiyama and H. Fujisaka, Phys. Rev. E **56**, 99 (1997).
[24] S. Uchiyama, Ph.D. dissertation, Kyushu University, 2000.
[25] S. Uchiyama and H. Fujisaka, J. Phys. A **32**, 4623 (1999).
[26] It has been well known that as far as the fixed-point-type retrieval (FP) states are concerned, they must exhaustively be found by the SCSNA. We can nevertheless observe some FP regions predicted only by the AMO. For example, there exists such contradiction when we compare the upper α of FP region in Fig. 11 with $C_2=1.2$ and Fig. 1(b). However, it is not important to compare two theoretical predictions quantitatively. In this region, they are no longer accurate in numerical simulations, respectively. The corresponding SCSNA solutions are, in terms of dynamical systems, changed to unstable fixed points [25]. On the other hand, the AMO solutions are necessarily stable, but its accuracy is worse than that of the SCSNA [15] and this tendency may get enhanced around the region concerned.
[27] One may find that the parameter value used in Fig. 2(b), which realizes an oscillatory-retrieval solution as shown in the figure, contradictorily belongs to the region of FP in Fig. 11. It is mainly a matter of the insufficient approximation of the AMO with $n=8$. Comparing the result of AMO with $n=8$ and that with $n=32$, we are able to infer as follow: this insufficiency makes the position of the $P1+P2$ regions upper than the well-predicted one. We remark, as the secondary reason, the finite-size effect of numerical simulations, which are considered to eventually shift the effective parameter value of Fig. 2(b) into the oscillatory-retrieval region.
[28] H. G. Schuster, *Deterministic Chaos*, 3rd ed. (VCH, Weinheim, 1995); E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, England, 1994).