Journal of Statistical Physics, Vol. 59, Nos. 3/4, 1990

Stochastic Analyses of the Dynamics of Generalized Little–Hopfield–Hemmen Type Neural Networks

Masatoshi Shiino[1]

Received July 11, 1989; revision received January 2, 1990

Stochastic analyses are conducted of model neural networks of the generalized Little–Hopfield–Hemmen type, in which the synaptic connections with linearly embedded p sets of patterns are free of symmetric ones, and a Glauber dynamics of a Markovian type is assumed. Two kinds of approaches are taken to study the stochastic dynamical behavior of the network system. First, by developing the method of the nonlinear master equation in the thermodynamic limit $N \to \infty$, an exact self-consistent equation is derived for the time evolultion of the pattern overlaps which play the role of the order parameters of the system. The self-consistent equation is shown to describe almost completely the macroscopic dynamical behavior of the network system. Second, conducting the system-size expansion of the master equation for the N-body probability distribution of the Glauber dynamics makes it possible to analyze the fluctuations. In the course of the analysis, the self-consistent equation for the pattern overlaps is derived again. The main result of the rigorous fluctuation analysis is that as far as the fluctuations are concerned, the time course of the pattern overlap fluctuations behaves independently of the fluctuations in the remaining modes of the system's macrovariables, in accordance with the self-determining property of the macroscopic motion of the pattern overlaps for neural networks with linear synaptic couplings.

KEY WORDS: Stochastic neural networks; asymmetric connection; Glauber dynamics; pattern overlaps; nonlinear master equation; fluctuations; system-size expansion; nonequilibrium phase transitions.

1. INTRODUCTION

The study of neural networks has been receiving much interest from physicists in the field of statistical mechanics.[1–6] Assemblies of interacting neurons which assume the two states firing and resting are well represented

[1] Department of Applied Physics, Faculty of Science, Tokyo Institute of Technology, Oh-okayama, Meguro-ku, Tokyo, Japan.

0022-4715/90/0500-1051$06.00/0

by spin systems with exchange interactions, and the process of associative memory retrieval is described by relaxational spin dynamics toward certain fixed points corresponding to the embedded patterns in the configuration space.[2,7] From such a statistical mechanical point of view, the Hopfield model of neural networks which deals with the case of symmetric connections has been extensively investigated and the theory of equilibrium statistical mechanics of spin-glass-like systems has been successfully applied.[5] This means that a static analysis suffices to explore the network properties in the case of symmetric connections. By contrast, when one deals with neural networks with asymmetric connections,[8-23] which are ubiquitous in real neurons of living systems,[24] it becomes quite important to consider the dynamics itself rather than the statics of the networks, due to the fact that the system then is of a dynamic nature.[12-22]

We are interested in the dynamical behavior of Hopfield-type neural networks with asymmetric synaptic connections. To explore the dynamical properties of the networks, it is necessary to first formulate the dynamics governing the time evolution of the system. There are, in general, a variety of ways of formulating dynamics, such as synchronous[1,25] or asynchronous,[2,26,27] deterministic[1,2] or stochastic,[1,26,28] Markovian[1,27] or non-Markovian,[14,17-19] etc. We take up the Glauber dynamics of Markovian type, which is considered to have the advantage that it not only provides us with a simple description of stochastic dynamics, but also enables us to deal with nonequilibrium phase transitions,[29-31] as is discussed below.

Glauber dynamics[32] was originally devised and developed for the study of the dynamics of ferromagnetic spin systems and it is well known that its mean-field version is capable of describing the qualitative nature of the dynamical behavior of ferromagnetic phase transitions in an intuitive as well as rigorous way.[33] As will be shown later, the extension of the Glauber dynamics to such a neural network system as endowed with asymmetric interactions is formal and quite easily done under the condition that the parameter representing the magnitude of external noise corresponds to temperature. The outcome of the extension is the capability of describing nonequilibrium phase transitions. Glauber dynamics with asymmetric interactions in general no longer ensures that the resulting dynamical equation for the order parameters yields a fixed-point-type attractor. In a previous paper,[21] I reported that asymmetric synaptic connections, if appropriately chosen, give rise to limit-cycle-type oscillations in the order parameters, ensuring associative memory retrieval of temporal pattern sequences. I found that as the magnitude of noise is varied, a variety of types of bifurcation in general can occur, including Hopf bifurcation. Since the occurrence of Hopf bifurcation in an infinite-particle system is considered to provide the concept of nonequilibrium phase transitions,[21,34] it follows that using Glauber dynamics, we can systematically discuss the behavior of the nonequilibrium phase transitions, which can be viewed as a natural extension of the concept of thermodynamic phase transitions. In short, Glauber dynamics with asymmetric connections provides interesting problems which bear upon two important points: associative memory retrieval of temporal pattern sequences and nonequilibrium phase transitions.

The aim of the present paper is to get insights into the dynamical structure of the Glauber dynamics capable of exhibiting nonequilibrium phase transitions, for the purpose of exploring the dynamical behavior of the stochastic neural network model. Noting that in a system of infinitely many degrees of freedom such as in a thermodynamic one, its dynamical properties are almost sufficiently described by only a few macrovariables in the form of the macroscopic motions and their fluctuations,[35–43] we are concerned with applying this idea to our neural network system.

In the analysis of the static properties of neural networks with symmetric connections, the macrovariables of physical relevance are known to be the so-called pattern overlaps.[4] To deal with the macroscopic motion of the pattern overlaps of asymmetric neural networks, we develop the method of the nonlinear master equation.[31,34,42–45] It will be noted that a master equation describing Markovian dynamics of a system composed of a finite number of particles is usually a linear equation and accordingly, in general, has nothing to do with bifurcation phenomena, owing to the fact that the usual type of H-theorem ensures the approach to a unique equilibrium state of a system.[46–48] The nonlinear master equation studied in the present paper is deduced from the underlyinig linear master equation for the Glauber dynamics by taking the thermodynamic limit $N \to \infty$ with the use of the mean-field character of the synaptic couplings between the neurons. Thereby, the nonlinear master equation in general is no longer expected to exhibit ergodic behavior and it becomes capable of displaying bifurcation phenomena.[21,34,42–44] From the nonlinear master equation, we can quite easily derive a self-consistent equation describing the time evolution of the pattern overlaps of the neural networks with asymmetric couplings which are linear in the embedded patterns. In this process, it is shown that the number of macrovariables necessary to describe the macroscopic behavior of the system is considerably reduced from 2^p to p and a self-determining property of the macroscopic motion of the pattern overlaps follows. It also follows that the p macrovariables, i.e., the pattern overlaps, are entitled to be the order parameters of the system even in the dynamical sense.

Then the problems arise of analyzing the order parameter fluctuations and of discussing the critical fluctuations accompanied by the occurrence of nonequilibrium phase transitions involving, as a typical case, Hopf bifurcation. While the behavior of the order parameter fluctuations in systems undergoing thermodynamic phase transitions have been extensively studied,[49,50] there has been relatively little activity in the study of their counterparts in systems exhibiting nonequilibrium phase transitions[29–31] except for the investigations of model chemical reaction systems[51] and of laser systems.[41,52]

So we are concerned in the present paper with the construction of a foundation for fluctuation analysis in neural network systems, based on which particular problems of fluctuation behavior such as critical fluctuations will be readily studied. In our problem of analyzing fluctuations associated with the pattern retrieval dynamics, it is of particular concern to investigate the extent to which the fluctuations in the pattern overlaps behave independently of the fluctuations in the remaining modes of the system's macrovariables.

In our fluctuation analysis, we employ the method of system-size expansion to extract the dynamics of the fluctuations in the form of the Fokker–Planck equation, which is obtained in the central limit scaling. The formalism of our system-size expansion, however, differs a little from the usual one in that the starting master equation assumes an explicit form of N-body equation written for the N-neuron system, although our system-size expansion follows substantially the spirit of what was devised originally by van Kampen[35] and, in fact, an alternative formulation using the usual recipe of the expansion is also available to obtain the same result. Coolen and Ruijgrok[20] used a similar method to ours of the expansion to obtain, in lowest order of the expansion, the time evolultion equation for the macroscopic pattern overlaps of the present model. However, an attempt to try to describe the fluctuations by simply extending their expansion up to higher orders turns out to fail, because of insufficiency in the number of macrovariables used to specify the state of the system. This situation is closely related to the point of our primary concern of how the pattern overlap fluctuations behave. Our system-size expansion of the master equation is based on the choice of a sufficient set of macrovariables, i.e., sublattice magnetizations, for the system-size expansion to make sense.

The present paper is organized as follows. In Section 2 we begin with the description of model neural networks of the generalized Little–Hopfield–Hemmen type to formulate the Glauber dynamics. We present the master equation describing the time evolution of the N-body probability distribution for states of N neurons. Introducing the concept of sublattice and empirical probability, we develop the method of the nonlinear master equation in the thermodynamic limit $N \to \infty$ to derive a self-consistent equation for the time evolution of the pattern overlaps. A brief report of our results was previously published elsewhere. In Section 3 we lay the foundations for the fluctuation analysis of the pattern overlaps to be studied in the next section. On the basis of the N-body master equation, a system-size expansion of the reduced probability distribution for the variables of sublattice magnetizations is develop to derive the Fokker–Planck equation for the fluctuations. Section 4 is devoted to the description of the main results of the fluctuation analysis of the present neural network system. By integrating out modes other than the pattern overlap fluctuations in the Fokker–Planck equation obtained in Section 3, we derive the time evolution equation for the pattern overlap fluctuations in a closed form for the fluctuations. In Section 5 we present a brief summary.

2. GLAUBER DYNAMICS IN THE NEURAL NETWORKS AND NONLINEAR MASTER EQUATION APPROACH

2.1. Model Neural Networks and Glauber Dynamics

The generalized Little–Hopfield–Hemmen model of neural networks[1,2,7] is a stochastic system of formal neurons[53] described by Ising spins ($S = \pm 1$), which are coupled to each other by synaptic connections corresponding to exchange interactions J_{ij}. The network has p sets of patterns which are embedded for the purpose of associative memory retrieval through the synaptic connections with the Hebb learning rule[54] taken into account.

Let $\{\xi_i^{(\nu)}\}$ ($\nu = 1, \ldots, p$) denote the νth pattern representing the νth particular state of neurons $\{\xi_1^{(\nu)}, \xi_2^{(\nu)}, \ldots, \xi_N^{(\nu)}\}$, with $\xi_i^{(\nu)}$ taking either $+1$ or -1. There is no need to impose restrictions on the choice of $\{\xi_i^{(\nu)}\}$, that is, the embedded patterns are free of the degree of randomness or correlations. The p sets of patterns are incorporated into defining synaptic couplings J_{ij} from neuron j to i with synaptic modifications of the Hebb type. We take the J_{ij} to be[20–22]

$$J_{ij} = \frac{1}{N} \sum_{\mu, \nu = 1}^{p} \xi_i^{(\mu)} a_{\mu\nu} \xi_j^{(\nu)} \tag{2.1}$$

where N refers to the total number of neurons and $(a_{\mu\nu})$ is a $p \times p$ matrix representing the connection strength. We omit the usual restriction of symmetry of $(a_{\mu\nu})$ to include asymmetric connections.

For the dynamics which governs the time evolution of the above system, we consider a Markov process described by the Glauber dynamics with one-spin flip. We assume the transition rate to be determined by the effective local field representing the weighted sum of the transmitted signals a neuron receives. Given a state of neurons $\{S\} = (S_1, S_2,..., S_N)$, the effective local field h_i to which the ith neuron S_i is subjected is defined by

$$h_i\{S\} = \sum_j J_{ij} S_j \tag{2.2}$$

Substituting Eq. (2.1), the h_i is rewritten as

$$h_i\{S\} = \sum_{\mu, \nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_j \frac{1}{N} \xi_j^{(\nu)} S_j = \sum_{\mu, \nu} \xi_i^{(\mu)} a_{\mu\nu} G^{(\nu)} \tag{2.3}$$

where we define the overlap $G^{(\nu)}$ of the instantaneous state of neurons with the embedded patterns $\{\xi_i^{(\nu)}\}$ $(\nu = 1,..., p)$

$$G^{(\nu)} = \frac{1}{N} \sum_j \xi_j^{(\nu)} S_j \tag{2.4}$$

in which $G^{(\nu)} = 1$ implies complete memory retrieval with respect to the νth pattern, while $G^{(\nu)} = 0$ represents no correlation with the embedded pattern. Defining the transition rate $w(S_i \to -S_i)$ to be

$$w(S_i \to -S_i) = \tfrac{1}{2}(1 - S_i \tanh \beta h_i\{S\}) \tag{2.5}$$

the Glauber dynamics is described by the time evolution equation for the N-body probability distribution $P(S_1,..., S_N; t)$,

$$\begin{aligned} \frac{\partial}{\partial t} P(S_1,..., S_N; t) &= -\sum_i w(S_i \to -S_i)\, P(S_1,..., S_i,..., S_N; t) \\ &\quad + \sum_i w(-S_i \to S_i)\, P(S_1,..., -S_i,..., S_N; t) \end{aligned} \tag{2.6}$$

Here, β represents a measure of the inverse magnitude of the external noise affecting the neurons, and we may say that $1/\beta$ plays the role of temperature in analogy to thermodynamic spin systems. We note that when $(a_{\mu\nu})$ is symmetric, Eq. (2.6) ensures the existence of an equilibrium distribution of Gibbs type bearing a striking resemblance to thermodynamic equilibrium state. In the case of asymmetric $(a_{\mu\nu})$, one cannot expect, in general, to have such a thermodynamic potential and thus is led to deal directly with the dynamics based on Eq. (2.6).

2.2. Nonlinear Master Equation and Self-Consistent Equation for the Pattern Overlaps

To proceed further with Eq. (2.6), we introduce the concept of sublattice[7,55] to take advantage of the present model with separability and mean-field character in the coupling scheme of Eq. (2.1). We divide the system of N neurons into at most 2^p sublattices according to the endowment of the p embedded patterns.

Let H^p denote the p-dimensional hypercube of ± 1 coordinates ($H^p = \{+1, -1\}^p$). Defining, for each $\xi \in H^p$, the sublattice $\Omega(\xi)$ such that

$$\Omega(\xi) = \{\text{site } i \mid \xi_i^{(\mu)} = \xi^{(\mu)},\ \mu = 1,..., p\} \tag{2.7}$$

we partition the whole system Ω into 2^p disjoint sublattices

$$\Omega = \bigcup_{\xi \in H^p} \Omega(\xi) \tag{2.8}$$

Denoting by $|\Omega(\xi)|$ the size of the set $\Omega(\xi)$, i.e., the number of elements of $\Omega(\xi)$, we define the rate of appearance $R(\xi)$ of the pattern component $\xi \in H^p$:

$$R(\xi) = \frac{|\Omega(\xi)|}{N} \tag{2.9}$$

We are interested in the behavior of our neural networks in the large-N limit. In the thermodynamic limit $N \to \infty$, one expects the ratio to the total number $|\Omega(\xi)|$ of the number of sites with spin up ($S_i = 1$) in $\Omega(\xi)$ to tend to a certain nonfluctuating quantity $p(t \mid \xi)$:

$$\frac{|\{\text{site } i \in \Omega(\xi) \mid S_i(t) = 1\}|}{|\Omega(\xi)|} \xrightarrow[N \to \infty]{} p(t \mid \xi) \tag{2.10}$$

since the law of large numbers holds. The $p(t \mid \xi)$ represents the empirical probability of finding $+1$ spin at time t in sublattice $\Omega(\xi)$. Since

$$\begin{aligned} G^{(\nu)} &= \frac{1}{N} \sum_j \xi_j^{(\nu)} S_j \\ &= \sum_{\xi \in H^p} R(\xi)\, \xi^{(\nu)} \frac{\sum_{j \in \Omega(\xi)} S_j}{|\Omega(\xi)|} \end{aligned} \tag{2.11}$$

the overlap $G^{(\nu)}$ also converges to a nonfluctuating quantity in the limit $N \to \infty$:

$$\begin{aligned} G^{(\nu)} \xrightarrow[N \to \infty]{} g^{(\nu)} &= \sum_{\xi \in H^p} r(\xi)\, \xi^{(\nu)} \{(+1)\, p(t \mid \xi) + (-1)[1 - p(t \mid \xi)]\} \\ &= \sum_{\xi \in H^p} r(\xi)\, \xi^{(\nu)} [2p(t \mid \xi) - 1] \end{aligned} \tag{2.12}$$

where we define $\lim_{N\to\infty} R(\boldsymbol{\xi}) = r(\boldsymbol{\xi})$. Accordingly, it turns out that the effective local field h_i at each site in $\Omega(\boldsymbol{\xi})$ also assumes in this limit a nonfluctuating value $h(\boldsymbol{\xi})$:

$$h(\boldsymbol{\xi}) = \sum_{\mu,\nu} \xi^{(\mu)} a_{\mu\nu}\, g^{(\nu)}$$

$$= \sum_{\mu,\nu} \xi^{(\mu)} a_{\mu\nu} \sum_{\boldsymbol{\eta}\in H^p} r(\boldsymbol{\eta})\, \eta^{(\nu)} [2p(t\,|\,\boldsymbol{\eta}) - 1] \tag{2.13}$$

Since the transition rates $w(S_i \to -S_i)$ and $w(-S_i \to S_i)$ in Eq. (2.6) then involve only the single variable S_i, the probability distribution $P(S_1,\dots, S_N, t)$ will get formally decoupled into N seemingly independent probability distributions, which, though, in reality remain coupled with each other only through the nonfluctuating local field $h(\boldsymbol{\xi})$ determined by themselves. In other words, the stochastic behavior of each spin S_i will then be governed by a single-body master equation of identical form which is obtained from Eq. (2.6) by setting $N=1$ and replacing $h_i\{S\}$ in the transition rate (2.5) by the nonfluctuating quantity $h(\boldsymbol{\xi})$. When one notes that the whole system is partitioned into sublattices $\{\Omega(\boldsymbol{\xi})\}$, each of which consists of a macroscopic number $[O(N)]$ of spins subjected to a common local field $h(\boldsymbol{\xi})$ which is expressed in terms of $p(t\,|\,\boldsymbol{\xi})$'s, it turns out that the macroscopic dynamical behavior of the system is described by the time evolution of the empirical probability $p(t\,|\,\boldsymbol{\xi})$ defined in Eq. (2.10).

In fact, we can write down the following self-consistent nonlinear master equation for the $p(t\,|\,\boldsymbol{\xi})$'s, which corresponds to the above-mentioned single-body master equation:[21]

$$\frac{d}{dt} p(t\,|\,\boldsymbol{\xi}) = -\frac{1}{2}[1 - \tanh \beta h(\boldsymbol{\xi})]\, p(t\,|\,\boldsymbol{\xi})$$

$$+\frac{1}{2}[1 + \tanh \beta h(\boldsymbol{\xi})][1 - p(t\,|\,\boldsymbol{\xi})]$$

$$= -p(t\,|\,\boldsymbol{\xi}) + \frac{1}{2}[1 + \tanh \beta h(\boldsymbol{\xi})] \tag{2.14}$$

with $h(\boldsymbol{\xi})$ determined self-consistently through Eq. (2.13). This equation, which is exact in the thermodynamic limit $N\to\infty$, will be rederived in another way using the method of system-size expansion in the next section. Unlike the usual linear master equations, the above nonlinear master equation, which shares a common property of yielding broken ergodicity with nonlinear Fokker–Planck equations developed for the study of phase transitions in stochastic systems of coupled nonlinear oscillators,[34,42–44] is capable of exhibiting bifurcation phenomena of its solutions which represent nonequilibrium phase transitions.

The dynamical structure of the above nonlinear master equation will be made clearer when we extract the dynamics of the pattern overlaps. In other words, combining Eqs. (2.12) and (2.14), we can obtain another self-consistent equation describing the time evolution of the overlaps $g^{(\nu)}(t)$ in a p-dimensional phase space. Differentiating Eq. (2.12) with respect to t and substituting Eq. (2.14), one obtains[20,21]

$$\frac{d}{dt} g^{(\nu)} = -g^{(\nu)} + \sum_{\xi \in H^p} r(\xi) \xi^{(\nu)} \tanh\left(\beta \sum_{\mu,k} \xi^{(\mu)} a_{\mu k}\, g^{(k)}\right)$$
$$\nu = 1,..., p \tag{2.15}$$

Here we have made a considerable reduction in the number of macrovariables of the differential equations, from 2^p in Eq. (2.14) to p in Eq. (2.15). We may say that the macrovariables of pattern overlaps have a self-determining property. A self-consistent equation for the pattern overlaps analogous to Eq. (2.15) was obtained by Buhmann and Schulten[15] for a simpler model of neural networks capable of temporal memory retrieval. The existence of such a self-determining property of the pattern overlaps, which was first found by Choy and Sherrington[56] in the case of symmetric connections with $p = 2$, is attributed to the nature of the synaptic couplings, where J_{ij} is linear in $(1/N) \sum_{\mu,\nu}^{p} \xi_i^{(\mu)} a_{\mu\nu} \xi_j^{(\nu)}$.[17]

Comparing Eqs. (2.14) and (2.15), we may also say that the motion of the $p(t \mid \xi)$ in 2^p-dimensional phase space is slaved by that of the pattern overlaps $g^{(\nu)}(t)$ in a p-dimensional phase space such that the dynamics of $g^{(\nu)}$, which is governed by the self-consistent equation (2.15), drives the time evolution of the $p(t \mid \xi)$ through the effective local field $h(\xi)$. It is further noted that the nonlinear equation (2.14) is linear in $p(t \mid \xi)$ except for the terms involving the hyperbolic tangent function, which plays the role of a driving force. Accordingly, it turns out that the macroscopic dynamical behavior of the neural network system including bifurcation phenomena is exhaustively described only by the dynamics of the pattern overlaps. The self-determining property of the macrovariables of the pattern overlaps together with their slaving property implies that they will be entitled to be the order parameters of the present neural network system at least in a macroscopic sense.

On the basis of the equation for the pattern overlaps, we studied previously the occurrence of nonequilibrium phase transitions in connection with associative memory retrieval of temporal pattern sequence.[21] Since setting $p = 2$ suffices to observe the effect of asymmetry on the dynamical behavior of the neural networks, we conducted a bifurcation analysis of Eq. (2.15) with $p=2$ and obtained the conditions of $(a_{\mu\nu})$ for the occurrence of limit-cycle-type solutions. We found that neural networks with asymmetric connections are characterized by the occurrence of a rich variety of bifurcation phenomena, including Hopf bifurcation, in contrast to the case of symmetric connections.

3. SYSTEM-SIZE EXPANSION AND SUBLATTICE MAGNETIZATION FLUCTUATIONS

We now turn to the study of the fluctuations of macrovariables describing the macroscopic motion of the system. Since the pattern overlaps $g^{(\nu)}$ have turned out to be the order parameters of the system and hence their dynamical behavior plays a crucial role in determining the structure of nonlinear dynamics of the network system as well as in featuring the network properties of learning and associative memory retrieval, the study of the pattern overlap fluctuations is of considerable importance. In particular, considering the fact that the p variables of the pattern overlaps $g^{(\nu)}$ constitute a sufficient set of relevant macrovariables with respect to the macroscopic behavior of the present system, which is in reality of 2^p-dimensional character, our primary concern is with the questions: Does the same thing also hold for the fluctuations? If not, to what extent do the pattern overlap fluctuations behave independently of the remaining degrees of freedom? In this and the following sections we are concerned with such problems.

To analyze the behavior of fluctuations, it is necessary to extract the fluctuations from the microscopic dynamics governing the system. Van Kampen[35–37] devised a system-size expansion method to investigate the time course of fluctuations in macrovariables of physical relevance, within the framework of Markovian dynamics. The method has been developed and applied by many authors[38,41,51,52,57] to investigate fluctuation phenomena in nonequilibrium systems such as chemical reaction ones. In this section we develop the system-size expansion method to deal with the behavior of the fluctuations in the neural network system, in which the N-body master equation (2.6) is explicitly formulated as governing its dynamical evolution, unlike the usual case where the system-size dependence appears only through a parameter characterizing the transition rates with the assumption of extensivity.[35–38]

Since the system-size expansion analysis involving only the pattern overlaps does not work in the present system, it is necessary to consider a sufficient set of macrovariables for the fluctuation analysis to be possible. We choose here a set of 2^p variables $\{M(\xi), \xi \in H^p\}$, which are linear functions of the empirical probabilities $p(\xi)$ and are defined below. They will be referred to, for convenience, as sublattice magnetizations, in analogy to magnetic spin systems.

Let $1_\xi(\cdot)$ ($\xi \in H^p$) be the indicator function of the set $\Omega(\xi)$:

$$\begin{aligned} 1_\xi(j) &= 1 \qquad \text{for} \quad j \in \Omega(\xi) \\ &= 0 \qquad \text{for} \quad j \notin \Omega(\xi) \end{aligned} \tag{3.1}$$

we define $M(\xi, \{S\})$ as

$$M(\xi, \{S\}) = \frac{1}{N} \sum_{j=1}^{N} 1_\xi(j) S_j = \frac{|\Omega(\xi)|}{N} \frac{\sum_{j \in \Omega(\xi)} S_j}{|\Omega(\xi)|}, \qquad \xi \in H^p \tag{3.2}$$

On the basis of the master equation (2.6), we consider the time evolution of the reduced probability distribution for the variables $M(\xi, \{S\})$, which is defined as

$$P_M(\{M\}, t) = \sum_{\{S\}} \prod_{\xi \in H^p} \delta(M(\xi) - M(\xi, \{S\}))\, P(\{S\}, t) \tag{3.3}$$

In dealing with the master equation (2.6), we make an assumption, only for the sake of simplicity of description, that the self-excitation J_{ii} is allowed in the contribution to the local field h_i, so that the summation in (2.2) is taken over all the N neurons. This assumption will turn out not to affect the final results on the behavior of the fluctuations.

From the master equation (2.6), we obtain the time derivative of the $P_M(\{M\}, t)$:

$$\begin{aligned}
&\frac{\partial}{\partial t} P_M(\{M\}, t) \\
&\quad = -\sum_i \sum_{\{S\}} \frac{1}{2} \prod_\xi \delta(M(\xi) - M(\xi, \{S\})) \\
&\qquad \times \left[1 - S_i \tanh\left(\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_\xi \xi^{(\nu)} \frac{\sum_j 1_\xi(j) S_j}{N}\right)\right] P(\{S\}, t) \\
&\qquad + \sum_i \sum_{\{S'\}} \frac{1}{2} \prod_\xi \delta\left(M(\xi) - \left[M(\xi, \{S'\}) - \frac{2 \cdot 1_\xi(i) S'_i}{N}\right]\right) \\
&\qquad \times \left[1 - S'_i \tanh\left(\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_\xi \xi^{(\nu)} \frac{\sum_j 1_\xi(j) S'_j}{N}\right)\right] \\
&\quad = \sum_i \sum_{\{S\}} \frac{1}{2} \left\{\prod_\xi \delta\left(M(\xi) - M(\xi, \{S\}) + \frac{2 \cdot 1_\xi(i) S_i}{N}\right)\right. \\
&\qquad \left. - \prod_\xi \delta(M(\xi) - M(\xi, \{S\}))\right\} \\
&\qquad \times \left[1 - S_i \tanh\left(\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_\xi \xi^{(\nu)} M(\xi, \{S\})\right)\right] P(\{S\}, t)
\end{aligned}$$

$$
=\sum_i \sum_{\{S\}} \frac{1}{2}\left[\sum_{k=1}^{\infty} \frac{1}{k!}\left(\sum_{\boldsymbol{\eta}} \frac{2\cdot 1_{\boldsymbol{\eta}}(i) S_i}{N} \frac{\partial}{\partial M(\boldsymbol{\eta})}\right)^k \prod_{\xi} \delta(M(\xi)-M(\xi,\{S\}))\right]
$$
$$
\times\left[1-S_i \tanh\left(\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)} M(\xi,\{S\})\right)\right] P(\{S\}, t) \tag{3.4}
$$

Here we used, in the first line of the above equation, the change of variables $\{S\} \to \{S'\}$: $S_j' = S_j$ $(j \neq i)$, $S_i' = -S_i$, noting that

$$
M(\xi,\{S\}) = \frac{1}{N}\left[\sum_{j=1}^{N} 1_{\xi}(j) S_j' - 2\cdot 1_{\xi}(i) S_i'\right]
$$
$$
= M(\xi,\{S'\}) - \frac{2\cdot 1_{\xi}(i) S_i'}{N} \tag{3.5}
$$

and

$$
h_i(S_1 \cdots -S_i \cdots S_N) = h_i(S_1' \cdots S_i' \cdots S_N')
$$
$$
= \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)} M(\xi,\{S'\}) \tag{3.6}
$$

The third line is a result of the Taylor expansion of the part involving the variation of the δ function, giving rise to the $1/N$ expansion of the master equation (2.6). We note that when the assumption of the self-excitation is removed, the term

$$
\left\{1-S_i \tanh\left[\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)} M(\xi,\{S\})\right]\right\}
$$

in Eq. (3.4) should be replaced by

$$
1-S_i \tanh\left[\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)}\left(M(\xi,\{S\}) - \frac{1_{\xi}(i) S_i}{N}\right)\right]
$$
$$
= 1 - S_i \sum_{l=0}^{\infty} \frac{1}{l!}\left[\sum_{\boldsymbol{\eta}} -\frac{1_{\boldsymbol{\eta}}(i) S_i}{N} \frac{\partial}{\partial M(\boldsymbol{\eta},\{S\})}\right]^l
$$
$$
\times \tanh\left[\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)} M(\xi,\{S\})\right] \tag{3.7}
$$

We proceed to calculate expansion (3.4) up to $k=2$ i.e., $O(1/N)$. The $k=1$ term L_1 in Eq. (3.4) can be written, after first performing the summation over i and noting $S_i^2=1$, as

$$L_1 = \sum_{\{S\}} \sum_{\boldsymbol{\eta}} \frac{\partial}{\partial M(\boldsymbol{\eta})} \prod_{\xi} \delta(M(\xi) - M(\xi, \{S\}))$$
$$\times \left\{ \sum_i \frac{1_{\boldsymbol{\eta}}(i) S_i}{N} - \sum_i \frac{1_{\boldsymbol{\eta}}(i)}{N} \tanh\left[\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)} M(\xi, \{S\}) \right] \right\} P(\{S\}, t)$$
$$= \sum_{\{S\}} \sum_{\boldsymbol{\eta}} \frac{\partial}{\partial M(\boldsymbol{\eta})} \prod_{\xi} \delta(M(\xi) - M(\xi, \{S\}))$$
$$\times \left\{ M(\boldsymbol{\eta}, \{S\}) - \frac{|\Omega(\boldsymbol{\eta})|}{N} \tanh\left(\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)} M(\xi, \{S\}) \right) \right\} P(\{S\}, t) \tag{3.8}$$

Further, changing the order of taking the summations over $\{S\}$ and $\boldsymbol{\eta}$, one obtains

$$L_1 = \sum_{\boldsymbol{\eta}} \frac{\partial}{\partial M(\boldsymbol{\eta})} \sum_{\{S\}} \prod_{\xi} \delta(M(\xi) - M(\xi, \{S\}))$$
$$\times \left\{ M(\boldsymbol{\eta}, \{S\}) - \frac{|\Omega(\boldsymbol{\eta})|}{N} \tanh\left[\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)} M(\xi, \{S\}) \right] \right\}$$
$$\times P(\{S\}, t)$$
$$= \sum_{\boldsymbol{\eta}} \frac{\partial}{\partial M(\boldsymbol{\eta})} \left\{ M(\boldsymbol{\eta}) - r(\boldsymbol{\eta}) \tanh\left[\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\nu)} M(\xi) \right] \right\}$$
$$\times P_M(\{M\}, t) \tag{3.9}$$

Here we assumed for simplicity that the dependence of $|\Omega(\boldsymbol{\eta})|/N$ on N can be discarded. This assumption will be reasonable when we confine ourselves to the situation in which the embedded patterns $\{\xi_i^{(\mu)}\}$ are of a kind of regular structure such that $|\Omega(\boldsymbol{\eta})|/N = r(\boldsymbol{\eta}) + O(1/N)$ in the limit of large N. Since the residual of $O(1/N)$ turns out not to contribute to the fluctuations studied below under the central limit scaling, we are then allowed to set $|\Omega(\boldsymbol{\eta})|/N = r(\boldsymbol{\eta})$ from the beginning. As a result of the above assumption, L_1 becomes an N-independent term, which survives after the limit $N \to \infty$ is taken.

The next leading term with N dependence of $O(1/N)$, which we denote by L_2, is the one with $k = 2$ in the expansion (3.4). We have from Eq. (3.4)

$$L_2 = \sum_i \sum_{\{S\}} \frac{1}{4} \left\{ \sum_{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2} \frac{4 \cdot 1_{\boldsymbol{\eta}_1}(i)\, 1_{\boldsymbol{\eta}_2}(i)}{N^2} \frac{\partial^2}{\partial M(\boldsymbol{\eta}_1)\, \partial M(\boldsymbol{\eta}_2)} \prod_{\xi} \delta(M(\xi) - M(\xi, \{S\})) \right\}$$
$$\times \left\{ 1 - S_i \tanh\left(\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_{\xi} \xi^{(\mu)} M(\xi, \{S\}) \right) \right\} P(\{S\}, t)$$

$$
\begin{aligned}
&= \sum_{\{S\}} \sum_{\boldsymbol{\eta}} \frac{1}{N^2} \frac{\partial^2}{\partial M(\boldsymbol{\eta})^2} \prod_{\boldsymbol{\xi}} \delta(M(\boldsymbol{\xi}) - M(\boldsymbol{\xi}, \{S\})) \\
&\quad \times \left\{ \sum_i 1_{\boldsymbol{\eta}}(i) - \sum_i 1_{\boldsymbol{\eta}}(i) S_i \tanh\left(\beta \sum_{\mu\nu} \xi_i^{(\mu)} a_{\mu\nu} \sum_{\boldsymbol{\xi}} \xi^{(\nu)} M(\boldsymbol{\xi}, \{S\}) \right) \right\} P(\{S\}, t) \\
&= \frac{1}{N} \sum_{\boldsymbol{\eta}} \frac{\partial^2}{\partial M(\boldsymbol{\eta})^2} \sum_{\{S\}} \prod_{\boldsymbol{\xi}} \delta(M(\boldsymbol{\xi}) - M(\boldsymbol{\xi}, \{S\})) \\
&\quad \times \left\{ \frac{|\Omega(\boldsymbol{\eta})|}{N} - \frac{1}{N} \left[\sum_i 1_{\boldsymbol{\eta}}(i) S_i \right] \tanh\left[\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\boldsymbol{\xi}} \xi^{(\nu)} M(\boldsymbol{\xi}, \{S\}) \right] \right\} \\
&\quad \times P(\{S\}, t) \\
&= \frac{1}{N} \sum_{\boldsymbol{\eta}} \frac{\partial^2}{\partial M(\boldsymbol{\eta})^2} \left\{ r(\boldsymbol{\eta}) - M(\boldsymbol{\eta}) \tanh\left[\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\boldsymbol{\xi}} \xi^{(\nu)} M(\boldsymbol{\xi}) \right] \right\} \\
&\quad \times P_M(\{M\}, t)
\end{aligned} \tag{3.10}
$$

In the absence of the self-excitation, one has an additional term for L_2 which originates from the multiplication of $k = 1$ and $l = 1$ terms in the respective expansions (3.4) and (3.7):

$$
\frac{1}{N} \sum_{\boldsymbol{\eta}} \frac{\partial}{\partial M(\boldsymbol{\eta})} \left(M(\boldsymbol{\eta}) \left\{ \frac{\partial}{\partial M(\boldsymbol{\eta})} \tanh\left[\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\boldsymbol{\xi}} \xi^{(\nu)} M(\boldsymbol{\xi}) \right] \right\} \right) P_M(\{M\}, t) \tag{3.11}
$$

The sum of L_1 and L_2 constitutes the system-size expansion up to $O(1/N)$ of the master equation, leading to the Fokker–Planck equation for the reduced probability density $P_M(\{M\}, t)$:

$$
\begin{aligned}
\frac{\partial}{\partial t} P_M(\{M\}, t) = \sum_{\boldsymbol{\eta}} \frac{\partial}{\partial M(\boldsymbol{\eta})} &\left\{ M(\boldsymbol{\eta}) - r(\boldsymbol{\eta}) \right. \\
&\left. \times \tanh\left[\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\boldsymbol{\xi}} \xi^{(\nu)} M(\boldsymbol{\xi}) \right] \right\} P_M(\{M\}, t) \\
&+ \frac{1}{N} \sum_{\boldsymbol{\eta}} \frac{\partial^2}{\partial M(\boldsymbol{\eta})^2} \left\{ r(\boldsymbol{\eta}) - M(\boldsymbol{\eta}) \right. \\
&\left. \times \tanh\left[\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\boldsymbol{\xi}} \xi^{(\nu)} M(\boldsymbol{\xi}) \right] \right\} P_M(\{M\}, t)
\end{aligned} \tag{3.12}
$$

In the thermodynamic limit $N \to \infty$, the second term on the rhs of

Eq. (3.12) vanishes and then Eq. (3.12) reduces to the one describing the macroscopic motion of the sublattice magnetizations $\bar{M}(\boldsymbol{\eta}, t)$ $[=\lim_{N\to\infty} M(\boldsymbol{\xi}, \{S(t)\})]$

$$\frac{d}{dt}\bar{M}(\boldsymbol{\eta}, t) = -\bar{M}(\boldsymbol{\eta}, t) + r(\boldsymbol{\eta}) \tanh\left[\beta \sum_{\mu,\nu} \eta^{(\mu)} a_{\mu\nu} \sum_{\boldsymbol{\xi}} \xi^{(\nu)} \bar{M}(\boldsymbol{\xi}, t)\right] \tag{3.13}$$

When one notes the relation between the sublattice magnetizations (3.2) and the empirical probabilities (2.10) in the limit $N \to \infty$,

$$\bar{M}(\boldsymbol{\eta}, t) = r(\boldsymbol{\eta})[2p(t \mid \boldsymbol{\eta}) - 1]$$

the nonlinear master equation (2.14) readily follows from Eq. (3.13). Since the pattern overlap $g^{(\nu)}$ in the limit $N \to \infty$ is given by

$$g^{(\nu)}(t) = \sum_{\boldsymbol{\xi}} \xi^{(\nu)} \bar{M}(\boldsymbol{\xi}, t) \tag{3.14}$$

its time evolution is readily obtained as

$$\begin{aligned} \frac{d}{dt} g^{(\nu)}(t) &= \sum_{\boldsymbol{\xi}} \xi^{(\nu)} \frac{d}{dt} \bar{M}(\boldsymbol{\xi}, t) \\ &= -g^{(\nu)}(t) + \sum_{\boldsymbol{\xi}} r(\boldsymbol{\xi})\, \xi^{(\nu)} \tanh\left[\beta \sum_{\mu,k} \xi^{(\mu)} a_{\mu k}\, g^{(k)}(t)\right] \end{aligned} \tag{3.15}$$

which is consistent with the result of the analysis using the nonlinear master equation.

We proceed to extract the fluctuation behavior from the Fokker–Planck equation (3.12). To this end, we use the well-known recipe of splitting $M(\boldsymbol{\eta})$ into its average $m(\boldsymbol{\eta}, t)$ and fluctuations $z(\boldsymbol{\eta})/\sqrt{N}$ for N large.[35–38, 43] Setting

$$\begin{aligned} M(\boldsymbol{\eta}) &= m(\boldsymbol{\eta}, t) + \frac{1}{\sqrt{N}} z(\boldsymbol{\eta}) \\ P_M(\{M\}, t)\, d\{M\} &= Q(\{z\}, t)\, d\{z\} \end{aligned} \tag{3.16}$$

and noting

$$\begin{aligned} N^{-2^{p}-1} \frac{\partial P_M}{\partial t} &= \frac{\partial Q}{\partial t} - \sqrt{N} \sum_{\boldsymbol{\eta}} \frac{\partial Q}{\partial z(\boldsymbol{\eta})} \frac{dm(\boldsymbol{\eta}, t)}{dt} \\ \frac{\partial}{\partial M(\boldsymbol{\eta})} &= \sqrt{N} \frac{\partial}{\partial z(\boldsymbol{\eta})} \end{aligned} \tag{3.17}$$

we obtain

$$\frac{\partial Q}{\partial t}-\sqrt{N}\sum_{\boldsymbol{\eta}}\frac{\partial Q}{\partial z(\boldsymbol{\eta})}\frac{dm(\boldsymbol{\eta},t)}{dt}$$
$$=\sum_{\boldsymbol{\eta}}\sqrt{N}\frac{\partial}{\partial z(\boldsymbol{\eta})}\left\{m(\boldsymbol{\eta},t)+\frac{1}{\sqrt{N}}z(\boldsymbol{\eta})\right.$$
$$\left.-r(\boldsymbol{\eta})\tanh\left[\beta\sum_{\mu\nu}\eta^{(\mu)}a_{\mu\nu}\sum_{\boldsymbol{\xi}}\xi^{(\nu)}\left(m(\boldsymbol{\xi},t)+\frac{z(\boldsymbol{\xi})}{\sqrt{N}}\right)\right]\right\}Q(\{z\},t)$$
$$+\sum_{\boldsymbol{\eta}}\frac{\partial^2}{\partial z(\boldsymbol{\eta})^2}\left\{r(\boldsymbol{\eta})-\left(m(\boldsymbol{\eta},t)+\frac{z(\boldsymbol{\eta})}{\sqrt{N}}\right)\right.$$
$$\left.\times\tanh\left[\beta\sum_{\mu\nu}\eta^{(\mu)}a_{\mu\nu}\sum_{\boldsymbol{\xi}}\xi^{(\nu)}\left(m(\boldsymbol{\xi},t)+\frac{z(\boldsymbol{\xi})}{\sqrt{N}}\right)\right]\right\}Q(\{z\},t) \quad (3.18)$$

With regard to $O(\sqrt{N})$, we have

$$-\sum_{\boldsymbol{\eta}}\frac{\partial Q}{\partial z(\boldsymbol{\eta})}\frac{dm(\boldsymbol{\eta},t)}{dt}$$
$$=\sum_{\boldsymbol{\eta}}\left\{m(\boldsymbol{\eta},t)-r(\boldsymbol{\eta})\tanh\left[\beta\sum_{\mu\nu}\eta^{(\mu)}a_{\mu\nu}\sum_{\boldsymbol{\xi}}\xi^{(\nu)}m(\boldsymbol{\xi},t)\right]\right\}\frac{\partial Q}{\partial z(\boldsymbol{\eta})} \quad (3.19)$$

When we choose $m(\boldsymbol{\eta},t)$ such that $m(\boldsymbol{\eta},t)$ is a solution to Eq. (3.13) of $\bar{M}(\boldsymbol{\eta},t)$'s, the above equation turns out to be automatically satisfied.

Concerning the term of $O(1)$, it follows that

$$\frac{\partial Q(\{z\},t)}{\partial t}$$
$$=\sum_{\boldsymbol{\eta}}\frac{\partial}{\partial z(\boldsymbol{\eta})}\left(z(\boldsymbol{\eta})-r(\boldsymbol{\eta})\sum_{l}\left\{\frac{\partial}{\partial g^{(l)}}\tanh\left[\beta\sum_{\mu\nu}\eta^{(\mu)}a_{\mu\nu}g^{(\nu)}\right]\right\}\left[\sum_{\boldsymbol{\xi}}\xi^{(l)}z(\boldsymbol{\xi})\right]\right)$$
$$\times Q(\{z\},t)$$
$$+\sum_{\boldsymbol{\eta}}\frac{\partial^2}{\partial z(\boldsymbol{\eta})^2}\left[r(\boldsymbol{\eta})-m(\boldsymbol{\eta},t)\tanh\left(\beta\sum_{\mu\nu}\eta^{(\mu)}a_{\mu\nu}g^{(\nu)}\right)\right]Q(\{z\},t) \quad (3.20)$$

with

$$g^{(\nu)}(t)=\sum_{\boldsymbol{\xi}}\xi^{(\nu)}m(\boldsymbol{\xi},t)$$

This equation is a linear Fokker–Planck equation with the drift terms linear in the variables $z(\boldsymbol{\eta})$ and describes the time course of the sublattice magnetization fluctuations, which are scaled according to Eq. (3.16) in the limit $N \to \infty$.

Even without the assumption of self-excitation on the local fields, we see that the same result follows, since the additional term in L_2 arising from the removal of the self-excitation turns out only to yield a term of $O(1/\sqrt{N})$ instead of that of $O(1)$, and accordingly it does not contribute to the fluctuations under the central limit scaling of Eq. (3.16).

We also note that the remark about the neglect of the term of $O(1/N)$ in L_1, on the basis of the assumption of the N dependence of $|\Omega(\mathbf{\eta})|/N$ is legitimate, since the residual of $O(1/N)$ in $|\Omega(\mathbf{\eta})|/N$ is observed to give rise to a contribution of $O(1/\sqrt{N})$ as well.

4. TIME EVOLUTION OF THE PATTERN OVERLAP FLUCTUATIONS

Our next problem is to investigate the fluctuations in the pattern overlaps $G^{(\nu)}$ on the basis of the linear Fokker–Planck equation (3.20). The fluctuations $y^{(\nu)}$ of the pattern overlaps are defined through

$$G^{(\nu)} = g^{(\nu)} + \frac{1}{\sqrt{N}} y^{(\nu)}, \qquad \nu = 1,..., p \tag{4.1}$$

Since

$$G^{(\nu)} = \sum_{\xi \in H^p} \left[m(\xi, t) + \frac{1}{\sqrt{N}} z(\xi) \right] \xi^{(\nu)} \tag{4.2}$$

$y^{(\nu)}$ can be written as

$$y^{(\nu)} = \sum_{\xi \in H^p} \xi^{(\nu)} z(\xi) \tag{4.3}$$

We attempt to obtain the time evolution of the probability distribution for the p variables $\{y^{(\nu)}\}$ from the linear Fokker–Planck equation for the 2^p variables $\{z(\xi)\}$. For this purpose, we perform the reduction of the macrovariables from $\{z\}$ to $\{y\}$ in (3.20). In other words, we define the reduced probability distribution for the pattern overlap fluctuations:

$$\tilde{Q}(\{\bar{y}\}, t) = \int d\{z\} \prod_{\nu=1}^{p} \delta\left(\bar{y}^{(\nu)} - \sum_{\xi} \xi^{(\nu)} z(\xi) \right) Q(\{z\}, t). \tag{4.4}$$

Differentiation of $\tilde{Q}$ with respect to t and substitution of Eq. (3.20) yields

$$
\begin{aligned}
\frac{\partial}{\partial t}\tilde{Q}(\{\bar{y}\}, t)
&= \sum_{\boldsymbol{\eta}} \int d\{z\} \prod_{\nu} \delta\left(\bar{y}^{(\nu)} - \sum_{\boldsymbol{\xi}} \xi^{(\nu)} z(\boldsymbol{\xi})\right) \\
&\quad \times \frac{\partial}{\partial z(\boldsymbol{\eta})} \left\{ z(\boldsymbol{\eta}) - r(\boldsymbol{\eta}) \sum_{l} J_l^{(1)}(\boldsymbol{\eta}) \left[\sum_{\boldsymbol{\xi}} \xi^{(l)} z(\boldsymbol{\xi}) \right] \right\} Q(\{z\}, t) \\
&\quad + \sum_{\boldsymbol{\eta}} \int d\{z\} \prod_{\nu} \delta\left(\bar{y}^{(\nu)} - \sum_{\boldsymbol{\xi}} \xi^{(\nu)} z(\boldsymbol{\xi})\right) \frac{\partial^2}{\partial z(\boldsymbol{\eta})^2} J^{(2)}(\boldsymbol{\eta})\, Q(\{z\}, t)
\end{aligned} \tag{4.5}
$$

with

$$
\begin{aligned}
J_l^{(1)}(\boldsymbol{\eta}) &= \frac{\partial}{\partial g^{(l)}} \tanh\left(\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu}\, g^{(\nu)}\right) \\
J^{(2)}(\boldsymbol{\eta}) &= r(\boldsymbol{\eta}) - m(\boldsymbol{\eta}, t) \tanh\left(\beta \sum_{\mu\nu} \eta^{(\mu)} a_{\mu\nu}\, g^{(\nu)}\right)
\end{aligned} \tag{4.6}
$$

The first term on the rhs of Eq. (4.5) involving $\partial/\partial z(\boldsymbol{\eta})$ (denoted by ⓘ) can be rewritten, by using Eq. (4.3), as

$$
\begin{aligned}
\text{ⓘ} = -\sum_{\boldsymbol{\eta}} \int d\{z\} &\left\{ \left(\sum_{\alpha=1}^{p} \eta^{(\alpha)} \frac{\partial}{\partial y^{(\alpha)}} \right) \prod_{\nu} \delta(\bar{y}^{(\nu)} - y^{(\nu)}) \right\} \\
&\times \left\{ z(\boldsymbol{\eta}) - r(\boldsymbol{\eta}) \sum_{l} J_l^{(1)}(\boldsymbol{\eta})\, y^{(l)} \right\} Q(\{z\}, t)
\end{aligned} \tag{4.7}
$$

We make a change of variables from $\{z\}$ to $\{y, y^c\}$ to perform the above integral:

$$
\int \cdots d\{z\} = \int \cdots \mathscr{J}(\{y\}, \{y^c\})\, d\{y^c\}\, d\{y\} \tag{4.8}
$$

with y^c denoting a set of appropriately chosen $2^p - p$ variables and $\mathscr{J}$ the Jacobian associated with the transformation. We note that using the transformed variables, we can rewrite Eq. (4.4) as

$$
\begin{aligned}
\tilde{Q}(\{\bar{y}\}, t) &= \int \mathscr{J}(\{y\}, \{y^c\})\, d\{y^c\}\, d\{y\} \prod_{\nu} (\bar{y}^{(\nu)} - y^{(\nu)})\, Q(\{y\}, \{y^c\}, t) \\
&= \int d\{y^c\}\, Q(\{\bar{y}\}, \{y^c\}, t)\, \mathscr{J}(\{\bar{y}\}, \{y^c\})
\end{aligned} \tag{4.9}
$$

Then we have

$$
\begin{aligned}
\text{①} = & -\sum_{\alpha} \int \mathscr{I}(\{y\}, \{y^c\})\, d\{y^c\}\, d\{y\} \left[\frac{\partial}{\partial y^{(\alpha)}} \prod_{\nu} \delta(\bar{y}^{(\nu)} - y^{(\nu)}) \right] \\
& \times \left[\sum_{\boldsymbol{\eta}} \eta^{(\alpha)} z(\boldsymbol{\eta}) - \sum_{\boldsymbol{\eta}} \eta^{(\alpha)} r(\boldsymbol{\eta}) \sum_{l} J_l^{(1)}(\boldsymbol{\eta})\, y^l \right] Q(\{y\}, \{y^c\}, t) \\
= & \sum_{\alpha} \int d\{y^c\}\, d\{y\} \prod_{\nu} \delta(\bar{y}^{(\nu)} - y^{(\nu)}) \frac{\partial}{\partial y^{(\alpha)}} \left\{ \left[y^{(\alpha)} - \sum_{\boldsymbol{\eta}} \eta^{(\alpha)} r(\boldsymbol{\eta}) \sum_{l} J_l^{(1)}(\boldsymbol{\eta})\, y^{(l)} \right] \right. \\
& \left. \times Q(\{y\}, \{y^c\}, t)\, \mathscr{I}(\{y\}, \{y^c\}) \right\} \\
= & \sum_{\alpha} \int d\{y^c\} \frac{\partial}{\partial \bar{y}^{(\alpha)}} \left(\bar{y}^{(\alpha)} - \sum_{\boldsymbol{\eta}} \eta^{(\alpha)} r(\boldsymbol{\eta}) \sum_{l} J_l^{(1)}(\boldsymbol{\eta})\, \bar{y}^{(l)} \right) Q(\{\bar{y}\}, \{y^c\}, t) \\
& \times \mathscr{I}(\{\bar{y}\}, \{y^c\}) \\
= & \sum_{\alpha} \frac{\partial}{\partial \bar{y}^{(\alpha)}} \left\{ \bar{y}^{(\alpha)} - \sum_{l} \left[\sum_{\boldsymbol{\eta}} r(\boldsymbol{\eta})\, \eta^{(\alpha)} J_l^{(1)}(\boldsymbol{\eta}) \right] \bar{y}^{(l)} \right\} \tilde{Q}(\{\bar{y}\}, t)
\end{aligned}
\tag{4.10}
$$

The second term on the rhs of Eq. (4.5) involving $\partial^2/\partial z(\boldsymbol{\eta})^2$, which we denote by ⑪, can also be integrated by almost the same procedure as in ①. Employing integration by parts twice, we obtain

$$
\begin{aligned}
\text{⑪} = & \sum_{\boldsymbol{\eta}} \int \mathscr{I}\, d\{y^c\}\, d\{y\} \left[\left(\sum_{\alpha} \eta^{(\alpha)} \frac{\partial}{\partial y^{(\alpha)}} \right)^2 \prod_{\nu} \delta(\bar{y}^{(\nu)} - y^{(\nu)}) \right] \\
& \times [J^{(2)}(\boldsymbol{\eta})\, Q(\{y\}, \{y^c\}, t)] \\
= & \sum_{\boldsymbol{\eta}} \int d\{y^c\}\, d\{y\} \prod_{\nu} \delta(\bar{y}^{(\nu)} - y^{(\nu)}) \left(\sum_{\alpha} \eta^{(\alpha)} \frac{\partial}{\partial y^{(\alpha)}} \right)^2 \\
& \times J^{(2)}(\boldsymbol{\eta})\, Q(\{y\}, \{y^c\}, t)\, \mathscr{I}(\{y\}, \{y^c\}) \\
= & \sum_{\boldsymbol{\eta}} \int d(y^c\} \left(\sum_{\alpha} \eta^{(\alpha)} \frac{\partial}{\partial \bar{y}^{(\alpha)}} \right)^2 J^{(2)}(\boldsymbol{\eta})\, Q(\{\bar{y}\}, \{y^c\}, t)\, \mathscr{I}(\{\bar{y}\}, \{y^c\}) \\
= & \sum_{\boldsymbol{\eta}} \left(\sum_{\alpha} \eta^{(\alpha)} \frac{\partial}{\partial \bar{y}^{(\alpha)}} \right)^2 J^{(2)}(\boldsymbol{\eta})\, \tilde{Q}(\{\bar{y}\}, t) \\
= & \sum_{\alpha,\gamma} \left[\sum_{\boldsymbol{\eta}} \eta^{(\alpha)} \eta^{(\gamma)} J^{(2)}(\boldsymbol{\eta}) \right] \frac{\partial^2}{\partial \bar{y}^{(\alpha)}\, \partial \bar{y}^{(\gamma)}} \tilde{Q}(\{\bar{y}\}, t)
\end{aligned}
\tag{4.11}
$$

Summing up ① and ⑪ and omitting the bar over y's, we finally obtain the time evolution equation of $\tilde{Q}(\{y\}, t)$:

$$\frac{\partial}{\partial t}\tilde{Q}(\{y\},t)=-\sum_{\alpha}\frac{\partial}{\partial y^{(\alpha)}}\left\{-y^{(\alpha)}+\sum_{l}\left[\sum_{\boldsymbol{\eta}}r(\boldsymbol{\eta})\,\eta^{(\alpha)}J_l^{(1)}(\boldsymbol{\eta})\right]y^{(l)}\right\}\tilde{Q}(\{y\},t)$$
$$+\sum_{\alpha,\gamma}\left[\sum_{\boldsymbol{\eta}}\eta^{(\alpha)}\eta^{(\gamma)}J^{(2)}(\boldsymbol{\eta})\right]\frac{\partial^2}{\partial y^{(\alpha)}\,\partial y^{(\gamma)}}\tilde{Q}(\{y\},t) \qquad (4.12)$$

with $J_l^{(1)}(\boldsymbol{\eta})$ and $J^{(2)}(\boldsymbol{\eta})$ given by (4.6).

This is a p-variable linear Fokker–Planck equation for the pattern overlap fluctuations $\{y^{(\nu)}\}$. Although the coefficients of the second derivatives contain through $J^{(2)}(\boldsymbol{\eta})$ not only the average pattern overlaps $\{g^{(\nu)}\}$, but also the average sublattice magnetizations $\{m(\boldsymbol{\eta},t)\}$, the above equation is observed to assume a closed form of the p variables of $\{y^{(\nu)}\}$, as far as the fluctuations are concerned. In other words, the time evolution of the pattern overlap fluctuations behaves independently of the fluctuations in the remaining modes of the system's macrovariables. Consequently, when one wants to know the time evolution of the pattern overlap fluctuations $\{y^{(\nu)}\}$, it suffices to get information on the macroscopic motion of the sublattice magnetizations $\{m(\boldsymbol{\eta},t)\}$ as well as the initial conditions on the $\{y^{(\nu)}\}$. One need not know more detailed information, such as on the $z(\boldsymbol{\xi})$. This situation will be a consequence of the fact that the reduced macrovariables of the pattern overlaps have the self-determining property mentioned in the previous section.

Under the central limit scaling (3.16) or (4.1), the probability density $\tilde{Q}(\{y\},t)$ takes a Gaussian form and thus is determined only by its first and second moments. On the basis of the linear Fokker–Planck equation (4.12), the time evolution equation for those moments can be easily obtained. Defining

$$\langle y^{(\nu)}\rangle_t=\int d\{y\}\,y^{(\nu)}\tilde{Q}(\{y\},t) \qquad (4.13a)$$

$$\langle y^{(\mu)}y^{(\nu)}\rangle_t=\int d\{y\}\,y^{(\mu)}y^{(\nu)}\tilde{Q}(\{y\},t) \qquad (4.13b)$$

$$K_{lm}=\sum_{\boldsymbol{\eta}}r(\boldsymbol{\eta})\,\eta^{(l)}J_m^{(1)}(\boldsymbol{\eta})-\delta_{lm}$$
$$=\frac{\partial}{\partial g^{(m)}(t)}\sum_{\boldsymbol{\eta}}r(\boldsymbol{\eta})\,\eta^{(l)}\tanh\left[\beta\sum_{\mu\nu}\eta^{(\mu)}a_{\mu\nu}\,g^{(\nu)}(t)\right]-\delta_{lm} \qquad (4.13c)$$

$$D_{lm}=\sum_{\boldsymbol{\eta}}\eta^{(l)}\eta^{(m)}J^{(2)}(\boldsymbol{\eta})$$
$$=\sum_{\boldsymbol{\eta}}\eta^{(l)}\eta^{(m)}\left\{r(\boldsymbol{\eta})-m(\boldsymbol{\eta},t)\tanh\left[\beta\sum_{\mu\nu}\eta^{(\mu)}a_{\mu\nu}\,g^{(\nu)}(t)\right]\right\} \qquad (4.13d)$$

it is straightforward to obtain from Eq. (4.12)

$$\frac{d}{dt}\langle y^{(l)}\rangle_t = \sum_{m=1}^{p} K_{lm}\langle y^{(m)}\rangle_t \tag{4.14a}$$

$$\frac{d}{dt}\langle y^{(l)}y^{(m)}\rangle_t = \sum_{n=1}^{p} K_{lm}\langle y^{(m)}y^{(n)}\rangle_t + \sum_{n=1}^{p} K_{mn}\langle y^{(l)}y^{(n)}\rangle_t + 2D_{lm} \tag{4.14b}$$

In the present context of the fluctuation analysis, Eq. (4.14a) turns out to be redundant, because one can set $\langle y^{(l)}\rangle = 0$ $(l = 1,..., p)$ owing to the definition in Eq. (4.1). Equation (4.14b) can be rewritten in a matrix form as

$$\frac{d}{dt}\sigma_y = K\sigma_y + \sigma_y K^T + 2D \tag{4.15}$$

with

$$(\sigma_y)_{lm} = \langle y^{(l)}y^{(m)}\rangle_t$$

Although this type of equation is very familiar in the usual system-size expansion analysis of fluctuations[35–38] such as for stochastic model systems of chemical reactions, the above equation for the pattern overlap fluctuations differs from its counterparts appearing in those systems in that the complete form of the self-determining property does not hold with respect to the fluctuations. In other words, as is shown in Eq. (4.13d), D cannot be determined solely by the p-dimensional macroscopic flow of the pattern overlaps $\{g^{(\nu)}\}$, but requires the 2^p-dimensional flow of $\{m(\boldsymbol{\eta}, t)\}$.

Note, however, that if the system is in a stationary state where the transient dies out, the order parameter fluctuations then exhibit a complete form of self-determining property, as is discussed below.

Suppose one is concerned with a stationary state of equilibrium type corresponding to a fix-point-type attractor of the macroscopic dynamical equations (3.13) or (3.15). Owing to the slaving property of the pattern overlaps, the equilibrium values $\{m(\boldsymbol{\eta}, t=\infty)\}$ turn out to be determined by the order parameters $\{g^{(\nu)}\}$ in the form

$$m(\boldsymbol{\eta}, t=\infty) = r(\boldsymbol{\eta}) \tanh\left(\beta \sum_{\mu,\nu} \eta^{(\mu)} a_{\mu\nu}\, g^{(\nu)}(t=\infty)\right) \tag{4.16}$$

Then it follows that the pattern overlap fluctuations $\{y^{(\nu)}\}$ in the stationary state are governed only by the macroscopic motion of the pattern overlaps $\{g^{(\nu)}\}$ through Eq. (4.12) or (4.15). In other words, the order parameter fluctuations recover the self-determining property as in the case of macroscopic dynamics. The above situation also applies to a stationary state exhibiting limit-cycle-type oscillatory behavior in the neural network system. In this case, the stationary time-periodic motion of $\{m(\boldsymbol{\eta}, t)\}$ also turns out to be solved uniquely in terms of stationary periodic functions of $\{g^{(\nu)}(t)\}$. Consequently, the knowledge of $\{g^{(\nu)}(t)\}$ completely determines the dynamical behavior of the pattern overlap fluctuations governed by Eq. (4.12) or (4.15). In short, we may say that the stationarity condition allows the pattern overlap fluctuations to recover the self-determining property and that the macrovariables of the pattern overlaps become worthy of being called order parameters in their true sense.

Finally, from the viewpoint of the self-determining property of the pattern overlap fluctuations, we remark on the result of a dynamical analysis made by Buhmann and Schulten(15) for a different system of neural networks capable of temporal retrieval of pattern sequences. They considered a system in which the bivariate variables representing states of neurons were assumed to take values either 1 or 0, unlike Ising spins of the present system, and further a kind of orthogonality condition was imposed among the embedded patterns. The specific assumptions are likely to have made their system remarkably simplified in such a way that the starting master equation describing the dynamics of the networks is written down only in terms of the pattern overlaps. Accordingly, in their theory there is no need to prepare deliberately variables other than the pattern overlaps, unlike in the present study, in conducting the fluctuation analysis. Then the pattern overlap fluctuations will have the complete form of the self-determining property as obtained using a usual recipe of the system-size expansion analysis. If one removes the assumption of the orthogonality condition, such a simplification is no longer expected to follow, as was suggested in their paper, and the problems of the fluctuations then will have to be studied along the line of the present analysis. For this reason, our present approach will be of more general character and will have potential applicability in exploring fluctuation phenomena in general types of stochastic neural network systems, including their case.

5. SUMMARY

We have conducted stochastic analyses of the Glauber dynamics in generalized Little–Hopfield–Hemmen-type neural networks. By employing two kinds of approaches to the analyses of the starting master equation, we have elucidated the dynamical structure of the time evolution equations for both the macroscopic average and fluctuations of the pattern overlaps in the present neural network system.

First we presented the method of the nonlinear master equation in describing the macroscopic dynamical behavior of the system under the thermodynamic limit $N \to \infty$. The time evolution equation governing the macroscopic motion of the pattern overlaps has been quite easily and transparently obtained using the nonlinear master equation for the empirical probabilities. We have found that the macrovariables of the pattern overlaps play the role of the order parameters, and the nonlinear dynamical behavior of the system including phenomena of the nonequilibrium phase transitions are exhaustively described by the macroscopic dynamical equations of the patern overlaps.

The method of the nonlinear master equation studied in the present paper will have wide applicability in the study of neural networks of similar structure to the present one, such as a system with a discrete-time version of the Glauber dynamics[58,59] or a system described by bivariate variables taking values either 1 or 0. Details of such applications as well as of the dynamical analysis of the present system conducted from the viewpoint of nonequilibrium phase transitions will be published elsewhere, although a preliminary result was previously reported.[21]

Another main result of our stochastic approaches concerns the fluctuation analysis using the method of system-size expansion. Taking up, as a sufficient set of macrovariables to specify the state of the systems, the sublattice magnetizations, we have incorporated the extraction of the reduced probability distribution for the sublattice magnetizations into a power series expansion in system size N of the N-body master equation for the Glauber dynamics. From the expansion up to $O(1/N)$ together with the central limit scaling of the fluctuations, the time evolution of the pattern overlap fluctuations has been extracted in a closed form for the fluctuations. It has been found that the pattern overlap fluctuations behave independently of the fluctuations in the remaining modes of the system's macrovariables, although they remain dependent on the macroscopic motions of the remaining modes (sublattice magnetizations). We have shown, however, that when the system is in a stationary state, the pattern overlap fluctuations recover the complete form of the self-determining property.

On the basis of the time evolution equation for the fluctuations, one can investigate particular problems of fluctuation phenomena, such as the behavior of critical fluctuations associated with nonequilibrium phase transitions, which are exhibited in a variety of ways by the macroscopic dynamical equation of the pattern overlaps. Application to those problems will be studied elsewhere.

REFERENCES

1. W. A. Little, *Math. Biosci.* **19**:101 (1974).
2. J. J. Hopfield, *Proc. Natl. Acad. Sci. USA* **79**:2554 (1982).
3. P. Peretto, *Biol. Cybernet.* **50**:51 (1984).
4. D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Phys. Rev. A* **32**:1007 (1985); *Ann. Phys.* **173**:30 (1987).
5. J. L. van Hemmen and I. Morgenstern, eds., *Heidelberg Colloquium on Glassy Dynamics* (Springer, Berlin, 1987).
6. E. Domany, *J. Stat. Phys.* **51**:743 (1988).
7. J. L. van Hemmen, *Phys. Rev. Lett.* **49**:409 (1982); *Phys. Rev. A* **34**:3435 (1986).
8. G. Parisi, *J. Phys. A* **19**:L675 (1986).
9. S. Shinomoto, *Biol. Cybern.* **57**:197 (1987).
10. A. Crisanti and H. Sompolinsky, *Phys. Rev. A* **36**:4922 (1987).
11. R. Kree and A. Zippelius, *Phys. Rev. A* **36**:4421 (1987).
12. S. Amari, *IEEE Trans.* **C-21**:1197 (1972).
13. E. Harth, *IEEE Trans. Syst. Man Cybern.* **SMC-13**:782 (1983).
14. H. Sompolinsky and I. Kanter, *Phys. Rev. Lett.* **57**:2861 (1986).
15. J. Buhmann and K. Schulten, *Europhys. Lett.* **4**:1205 (1987). J. Buhmann and K. Schulten, in *Neural Computers* (NATO ASI Series, Vol. F41).
16. I. Tsuda, E. Koerner and H. Shimizu. *Prog. Theor. Phys.* **78**:51 (1987).
17. U. Riedel, R. Kuhn, and J. L. van Hemmen, *Phys. Rev. A* **38**:1105 (1988).
18. D. Kleinfield, *Proc. Natl. Acad. Sci. USA* **83**:9469 (1986).
19. D. Kleinfield and H. Sompolinsky, *Biophys. J.* **54**:1039 (1988).
20. A. C. C. Coolen and Th. W. Ruijgrok, *Phys. Rev. A* **38**:4253 (1988).
21. M. Shiino, H. Nishimori, and M. Ono, *J. Phys. Soc. Jpn.* **58**:763 (1989).
22. H. Nishimori, T. Nakamura, and M. Shiino, *Phys. Rev. A*, March 15 (1990).
23. T. Fukai, *J. Phys. A* **23**:249 (1990).
24. J. C. Eccles, *The Understanding of the Brain* (McGraw-Hill, New York, 1977).
25. L. Personnaz, I. Guyon, and G. Dreyfus, *Phys. Rev. A* **34**:4217 (1986).
26. P. Peretto and J. J. Niez, *IEEE Trans. Syst. Man Cybern.* **SMC-16**:73 (1986).
27. H. Rieger, M. Schreckenberg, and J. Zittartz, *Z. Phys. B* **72**:523 (1988).
28. D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, *Cognitive Sci.* **9**:147 (1985).
29. G. Nicolis and I. Prigogine, *Self-organization in Non-equilibrium Systems* (Wiley, New York, 1977).
30. H. Haken, *Synergetics, an Introduction* (Springer, Berlin, 1978); *Rev. Mod. Phys.* **47**:67 (1975).
31. Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer, Berlin, 1984).
32. R. J. Glauber, *J. Math. Phys.* **4**:294 (1963).
33. M. Suzuki and R. Kubo, *J. Phys. Soc. Jpn.* **24**:51 (1968).
34. M. Shiino, *Phys. Lett.* **111A**:396 (1985).
35. N. G. van Kampen, *Can. J. Phys.* **39**:551 (1961).
36. N. H. van Kampen, *Adv. Chem. Phys.* **34**:245 (1976).
37. N. G. van Kampen, *Stochastic Process in Physics and Chemistry* (North-Holland, Amsterdam, 1981).
38. R. Kubo, K. Matsuo, and K. Kitahara, *J. Stat. Phys.* **9**:51 (1973).
39. H. Mori, *Prog. Theor. Phys.* **33**:423 (1965).
40. K. Kawasaki, *Ann. Phys.* **61**:1 (1970).
41. K. Hepp and E. H. Lieb, *Helv. Phys. Acta* **46**:573 (1973).
42. D. A. Dawson, *J. Stat. Phys.* **31**:29 (1983).

43. M. Shiino, *Phys. Rev. A* **36**:2393 (1987).
44. M. Shiino, *Phys. Lett.* **112A**:302 (1985).
45. M. Malek Mansour and G. Nicolis, *J. Stat. Phys.* **13**:197 (1975).
46. F. Schlogl, *Z. Phys.* **243**:303 (1971); **244**:199 (1971).
47. H. Risken, *The Fokker–Planck Equation* (Springer, Berlin, 1984).
48. J. W. Clark, *Phys. Rep.* **158**:91 (1988).
49. P. C. Hohenberg and B. I. Halperin, *Rev. Mod. Phys.* **49**:435 (1977).
50. S.-K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, 1976).
51. K. Tomita and H. Tomita, *Prog. Theor. Phys.* **51**:1731 (1974).
52. K. Tomita, T. Todani, and H. Kidachi, *Physica* **84A**:350 (1976).
53. W. A. McCulloch and W. Pitts, *Math. Biophys.* **5**:115 (1943).
54. D. C. Hebb, *The Organization of Behavior; A Neurophysiological Theory* (Wiley, New York, 1957).
55. J. L. van Hemmen and R. Kuhn, *Phys. Rev. Lett.* **57**:913 (1986).
56. T. C. Choy and D. Sherrington, *J. Phys. C* **17**:739 (1984).
57. M. Suzuki, *Prog. Theor. Phys.* **55**:1064 (1976).
58. T. Fukai and M. Shiino, *Phys. Rev. Lett.* **19**:1465 (1990).
59. M. Shiino and T. Fukai, *J. Phys. Soc. Jpn.*, to appear.