Projection-operator approach to overlap dynamics in a Hopfield network

Toyonori Munakata and Yasuyuki Nakamura
Department of Applied Mathematics and Physics, Kyoto University, Kyoto 606, Japan
(Received 23 November 1992)

The Glauber dynamics of the Hopfield model is studied with use of the Mori-Zwanzig projection-operator formalism for irreversible processes and exact evolution equations for the overlaps are derived.

PACS number(s): 87.10.+e, 05.40.+j, 03.20.+i

The Hopfield model [1] for neural networks with symmetric connection J_{ij} between two neurons i and j has been extensively studied from the standpoint of (equilibrium) statistical mechanics of phase transitions in disordered spin systems [2]. Typical information which can be afforded by the theory is the number of the patterns which can be stored and retrieved in the network and the rate of errors in the retrieved pattern. Recently dynamical behavior of the model has gathered considerable interest in connection with the speed of the association process and the size and structure of the basins of attraction [3–5]. These questions can only be answered within *a nonequilibrium statistical dynamic treatment* of the system, which is the main concern in this paper.

Basic evolution of the Hopfield model, consisting of N formal neurons (or Ising spins) is usually governed by the Glauber dynamics [6]

$$\frac{\partial p(\mathbf{S},t)}{\partial t}=-(1/\Gamma)\sum_{i=1}^{N}\sum_{S_i'=\pm 1} S_iS_i' f(-S_i'|h_i)p(\mathbf{S}(S_i'),t)$$
$$=L_G p(\mathbf{S},t)\ , \tag{1}$$

where $p(\mathbf{S},t)$ denotes the N-body probability distribution at time t with $\mathbf{S}=(S_1,\ldots,S_N)$ and $\mathbf{S}(S_i')=(S_1,\ldots,S_{i-1},S_i',S_{i+1},\ldots,S_N)$. The Γ is the cycle time in which each of the N neurons updates once on average. The function $f(S_i|h_i)$ represents the probability for the ith spin to flip into the state S_i in the field h_i, which is given at temperature T by

$$f(S_i|h_i)=\frac{\exp(h_iS_i/T)}{\exp(h_i/T)+\exp(-h_i/T)}\ , \tag{2}$$

and the field h_i is determined by $N-1$ spins as

$$h_i=\sum_{j\,(\neq i)} J_{ij}S_i\ , \tag{3}$$

when there is no external field. The connection J_{ij} is expressed in terms of the p embedded patterns $\xi^{(\mu)}=\{\xi_i^{(\mu)}\}$ $(i=1,\ldots,N)$ by the Hebbian rule [1,6],

$$J_{ij}=(1/N)\sum_{\mu=1}^{p}\xi_i^{(\mu)}\xi_j^{(\mu)}(1-\delta_{ij})\ . \tag{4}$$

With regard to a retrieval process the most important quantity is the overlap, defined by [7]

$$\hat{m}_\mu\equiv\left[\frac{1}{N}\right]\sum_{i=1}^{N}\xi_i^{(\mu)}S_i\equiv\left[\frac{1}{N}\right]\hat{M}_\mu\ , \tag{5}$$

which quantitatively measures similarity between the state $\mathbf{S}$ of the system and the pattern μ. The retrieval or association process is the one in which one of the $\hat{m}_\mu(t)$ approaches one with the other $p-1$ overlaps remaining small (typically of order $N^{-1/2}$). The equilibrium distribution $p_{\text{eq}}(\mathbf{S})$, which satisfies $L_G p_{\text{eq}}=0$, is given by

$$p_{\text{eq}}(\mathbf{S})=Z^{-1}\exp\left[(2NT)^{-1}\sum_{\mu}\hat{M}_\mu^2\right]\ , \tag{6}$$

where Z is the partition function of the system. When the number of the patterns p is much smaller than that of neurons N, as is often the case, mapping the Glauber dynamics to the overlap dynamics belongs to a typical problem of deriving a reduced description and this is most effectively performed by utilizing some kind of projection-operator method, which has been playing an important role in irreversible statistical mechanics [8–10].

In this Brief Report we apply the Mori-Zwanzig projection-operator formalism to the Glauber dynamics (1) in order to investigate dynamics of the retrieval process. Before proeceeding to the details of the calculation we briefly present the main framework of the formalism in a form suitable for later application. Let us assume that the probability distribution function $p(\mathbf{x},t)$ is governed by

$$\frac{\partial p(\mathbf{x},t)}{\partial t}=Lp(\mathbf{x},t)\ , \tag{7}$$

where L is a general time-independent linear operator. Although the original formulation is given for the Hamiltonian system [9] and a stochastic system described by a Fokker-Planck operator [10], the L in Eq. (7) can be the Glauber operator L_G, (1) where $\mathbf{x}$ denotes $\mathbf{S}$ and integration over $\mathbf{x}$ in Eq. (8) below should be interpreted to be a summation over $\mathbf{S}$. The operator adjoint to L is defined by

$$\int d\mathbf{x}\,f(\mathbf{x})Lg(\mathbf{x})=\int d\mathbf{x}\,g(\mathbf{x})\Lambda f(\mathbf{x}) \tag{8}$$

for arbitrary functions $f(\mathbf{x})$ and $g(\mathbf{x})$. The operator Λ is seen to control time evolution of dynamical variables, i.e.,

$$\frac{dA(t)}{dt}=\Lambda A(t)\quad\text{or}\quad A(t)=\exp[\Lambda t]A\ . \tag{9}$$

To see the implication of Eq. (9) we consider that at time $t=0$ we know that the system is located in the state space at $\mathbf{x}_0$, that is, $p(\mathbf{x},t=0)=\delta(\mathbf{x}-\mathbf{x}_0)$. Then the expectation of an arbitrary variable $A(\mathbf{x})$ at time t is calculated as

$$\begin{aligned}\langle A(t)\rangle_{x_0} &= \int d\mathbf{x}\, A(\mathbf{x}) \exp[Lt]\delta(\mathbf{x}-\mathbf{x}_0) \\ &= \exp[\Lambda t] A(\mathbf{x}_0) , \end{aligned} \tag{10}$$

where now the Λ is an operator working in the space $\mathbf{x}_0$. Let us consider collective dynamics of a set of dynamical variables $A_i(\mathbf{x})$ $(i=1,\ldots,I)$. Mutual correlation among the variables $\{A_i(\mathbf{x})\}$ is best extracted by introducing a projection operator P into a space spanned by $\{A_i(\mathbf{x})\}$, defined by [9,10]

$$PB(\mathbf{x}) = \sum_{i,j} (B(\mathbf{x}), A_i(\mathbf{x}))(\Xi^{-1})_{i,j} A_j(\mathbf{x}) , \tag{11}$$

where the inner product $(f(\mathbf{x}),g(\mathbf{x}))$ of two dynamical variables f and g denotes the equilibrium average $\int d\mathbf{x} f(\mathbf{x})g(\mathbf{x})p_{\text{eq}}(\mathbf{x})$ and $\Xi_{ij}\equiv(A_i(\mathbf{x}), A_j(\mathbf{x}))$. Applying the operator P on Eq. (9) one obtains what is called a generalized Langevin (GL) equation [9,10],

$$\frac{dA_i(t)}{dt} = \sum_j \omega_{ij} A_j(t) - \int_0^t ds\, \Psi_{ij}(t-s) A_j(s) + f_i(t) , \tag{12}$$

where the frequency and damping matrices are given by

$$w_{ij} = \sum_k (\Lambda A_i, A_k)(\Xi^{-1})_{kj} , \tag{13}$$

$$\Psi_{ij}(t) = -\sum_k (\Lambda f_i(t), A_k)(\Xi^{-1})_{kj} , \tag{14}$$

and

$$\begin{aligned} f_i(t) &= \exp[t(1-P)\Lambda]\{(1-P)\Lambda A_i\} \\ &\equiv \exp[t(1-P)\Lambda] f_i . \end{aligned} \tag{15}$$

Usually calculation of the damping kernel $\Psi(t)$ is prohibitively difficult. However, by including all the relevant variables in the set $\{A_i(\mathbf{x})\}$ we may have a rather good description of the collective dynamics of the system even if we neglect the damping kernel entirely. For dense gases and liquids [11], for example, by choosing as $\{A_i(\mathbf{x})\}$ the density in a μ space, $f_{rp}(\mathbf{x})=\sum_n \delta(\mathbf{r}-\mathbf{r}_n)\delta(\mathbf{p}-\mathbf{p}_n)$, where the suffix i of $A_i(\mathbf{x})$ now becomes continuous parameters $(\mathbf{r},\mathbf{p})$ and $\mathbf{x}=(\mathbf{r}_1,\ldots,\mathbf{r}_N,\mathbf{p}_1,\ldots,\mathbf{p}_N)$, we have a kinetic equation for $f_{rp}(t)$ in which the matrix ω gives a Vlasov (mean-field) term and the kernel Ψ represents effects of collisions. From many examples of applications of Eq. (12) to various systems [12,9], it is reasonable to call the GL equation (12) without the kernel the mean-field approximation. Effects of damping could be approximately taken into account as $\Psi(t)\cong\Psi(0)\psi(t)$ where $\psi(t)$ is chosen to be a simple function with some parameter(s) determined by sum-rule arguments. Below we mainly consider the mean-field approximation with a brief comment on the damping kernel.

Now we turn to the problem of overlap dynamics. From Eqs. (7) and (1) the operator adjoint to L_G, Eq. (1), is given by

$$\Lambda_G g(\mathbf{S}) = \Gamma^{-1} \sum_i f(-S_i|h_i)[g(\mathbf{S}(-S_i)) - g(\mathbf{S})] . \tag{16}$$

Since we are interested in dynamics of retrieval we take as the collective variables $\{A_i(\mathbf{S})\}$ the quantity

$$g_{\mathbf{M}}(\mathbf{S}) = \prod_{\mu=1}^{p} \Delta(\hat{M}_\mu - M_\mu) \equiv \Delta(\hat{\mathbf{M}} - \mathbf{M}) , \tag{17}$$

where $\Delta(n)$ is defined for integer n with $\Delta(n)=1$ for $n=0$ and 0 otherwise. $g_{\mathbf{M}}(\mathbf{S})$ represents the event in which the overlap $\hat{M}_\mu(\mu=1,\ldots,p)$ takes the value M_μ for each μ. The ensemble average $\langle g_{\hat{\mathbf{M}}}(\mathbf{S})\rangle$, which expresses the probability of observing the overlap $\mathbf{M}=(M_1,\ldots,M_p)$ in the equilibrium state, the correlation $\Xi(\mathbf{M},\mathbf{M}')\equiv\langle g_{\mathbf{M}}(\mathbf{S})g_{\mathbf{M}'}(\mathbf{S})\rangle$ and its inverse are given by

$$\langle g_{\mathbf{M}}(\mathbf{S})\rangle = z^{-1} \exp\left[\sum_\mu M_\mu^2/2NT\right] \Omega_N(\mathbf{M}) , \tag{18}$$

$$\begin{aligned} \Xi(\mathbf{M},\mathbf{M}') &= \Delta(\mathbf{M}-\mathbf{M}')\langle g_{\mathbf{M}}(\mathbf{S})\rangle , \\ \Xi^{-1}(\mathbf{M},\mathbf{M}') &= \Delta(\mathbf{M}-\mathbf{M}')/\langle g_{\mathbf{M}}(\mathbf{S})\rangle , \end{aligned} \tag{19}$$

where $\Omega_N(\mathbf{M})=\sum_{\mathbf{S}} g_{\mathbf{M}}(\mathbf{S}) = T_r g_{\mathbf{M}}(\mathbf{S})$ is the number of the events, among the 2^N possible microscopic spin configurations, in which $g_{\mathbf{M}}(\mathbf{S})=1$. From Eqs. (16) and (17)

$$\begin{aligned} \Lambda_G g_{\mathbf{M}}(\mathbf{S}) = \Gamma^{-1} \sum_i f(-S_i|h_i) \\ \times \prod_\mu [\Delta(\hat{M}_\mu(i) - \xi_i^{(\mu)} S_i - M_\mu) \\ - \Delta(\hat{M}_\mu(i) + \xi_i^{(\mu)} S_i - M_\mu)] , \end{aligned} \tag{20}$$

where we have introduced $\hat{M}_\mu(i)$ by

$$\hat{M}_\mu(i) = \sum_{j(\neq i)} \xi_j^{(\mu)} S_j , \tag{21}$$

and the field h_i, Eq. (3), is expressed in terms of $\hat{M}_\mu(i)$ as

$$h_i = N^{-1} \sum_\mu \xi_i^{(\mu)} \hat{M}_\mu(i) \equiv N^{-1} \xi_i \cdot \hat{\mathbf{M}}(i) , \tag{22}$$

where ξ_i and $\hat{\mathbf{M}}(i)$ are p-dimensional vectors $(\xi_i^{(1)},\ldots,\xi_i^{(p)})$ and $[\hat{M}^{(1)}(i),\ldots,\hat{M}^{(p)}(i)]$. The first step to obtain

$$\begin{aligned} \omega_{\mathbf{M},\mathbf{M}'} &\equiv \sum_{\mathbf{M}''} (\Lambda_G g_{\mathbf{M}}, g_{\mathbf{M}''}) \Xi^{-1}(\mathbf{M}'',\mathbf{M}') \\ &= (\Lambda g_{\mathbf{M}}, g_{\mathbf{M}'})/\langle g_{\mathbf{M}'}\rangle \end{aligned}$$

[see Eqs. (13) and (18)] consists in calculating $D_{\mathbf{M},\mathbf{M}'}\equiv(\Lambda g_{\mathbf{M}}, g_{\mathbf{M}'})$. From Eqs. (20) and (6) we see that, setting $\gamma\equiv\exp[\sum_\mu M_\mu'^2/2NT]$,

$$D_{\mathbf{M},\mathbf{M}'}/\gamma=\sum_i T_r\left[\prod_\mu \Delta(M'_\mu-M_\mu-2\xi_i^{(\mu)}S_i)f(-S_i|h_i)\Delta(\hat{M}_\mu(i)-\xi_i^{(\mu)}S_i-M_\mu)\right.$$
$$\left.-\prod_\mu \Delta(M'_\mu-M_\mu)f(-S_i|h_i)\Delta(\hat{M}_\mu(i)+\xi_i^{(\mu)}S_i-M_\mu)\right] . \tag{23}$$

We take the trace over $S_i=\pm1$ and then over the remaining spin variables to obtain

$$D_{\mathbf{M},\mathbf{M}'}/\gamma=\sum_i\left[\Delta(\mathbf{M}'-\mathbf{M}-2\xi_i)f\left[-1|h_i=N^{-1}\sum_\mu\xi_i^{(\mu)}(M_\mu+\xi_i^{(\mu)})\right]\Omega_{N-1}(\mathbf{M}+\xi_i)\right.$$
$$+\Delta(\mathbf{M}'-\mathbf{M}+2\xi_i)f\left[1|h_i=N^{-1}\sum_\mu\xi_i^{(\mu)}(M_\mu-\xi_i^{(\mu)})\right]\Omega_{N-1}(\mathbf{M}-\xi_i)$$
$$-\Delta(\mathbf{M}'-\mathbf{M})f\left[-1|h_i=N^{-1}\sum_\mu\xi_i^{(\mu)}(M_\mu-\xi_i^{(\mu)})\right]\Omega_{N-1}(\mathbf{M}-\xi_i)$$
$$\left.-\Delta(\mathbf{M}'-\mathbf{M})f\left[1|h_i=N^{-1}\sum_\mu\xi_i^{(\mu)}(M_\mu+\xi_i^{(\mu)})\right]\Omega_{N-1}(\mathbf{M}+\xi_i)\right] . \tag{24}$$

From Eqs. (19) and (24) we see that our next task is to calculate

$$p(\xi_i;\mathbf{M})\equiv\Omega_{N-1}(\mathbf{M}-\xi_i)/\Omega_N(\mathbf{M}) , \tag{25}$$

which appears in

$$\omega_{\mathbf{M},\mathbf{M}'}=D_{\mathbf{M},\mathbf{M}'}/[\Gamma\gamma\Omega_N(\mathbf{M}')] . \tag{26}$$

From the definition (25) $p(\xi_i;\mathbf{M})$ stands for the probability of $S_i=1$ under the condition $\sum_j S_j\xi_j=\mathbf{M}$. Without loss of generality we take $i=1$. First we consider the case $p=1$, i.e., there is only one pattern. From the condition $\sum_j S_j\xi_j^{(1)}=M_1$ we have $S_1+\sum_{j(\neq1)}\xi_1^{(1)}\xi_j^{(1)}S_j=M_1\xi_1^{(1)}$. From the law of large number, half of $\xi_1^{(1)}\xi_j^{(1)}$ take the value 1 and we relabel these j to yield $\sum_{i=1}^{N/2}S_i-\sum_{i=N/2+1}^{N}S_i=M_1\xi_1^{(1)}$ or

$$\sum_{i=1}^{N/2}S_i=M_1\xi_1^{(1)}/2 . \tag{27}$$

Since the probabilistic variables $S_1,\ldots,S_{N/2}$ are equivalent to each other we have from Eq. (27)

$$p(\xi_1^{(1)};M_1)N/2-[1-p(\xi_1^{(1)};M_1)]N/2=M_1\xi_1^{(1)}/2 ,$$

leading to

$$p(\xi_1^{(1)};M_1)=\frac{(1+M_1\xi_1^{(1)}/N)}{2}=\frac{(1+m_1\xi_1^{(1)})}{2} . \tag{28}$$

If there are two patterns, $p=2$, we have an additional condition $S_1+\sum_{j(\neq1)}\xi_1^{(2)}\xi_j^{(2)}S_j=M_2\xi_1^{(2)}$. Using the law of large number again we note that half of the coefficients $\xi_1^{(2)}\xi_j^{(2)}$ $(j\le N/2)$ take the value 1, which are relabeled again to have $\sum_{i=1}^{N/4}S_i-\sum_{i=N/4+1}^{N/2}S_i=M_2\xi_1^{(2)}/2$.

From Eq. (27) and the above we have

$$\sum_{i=1}^{N/4}S_i=[M_1\xi_1^{(1)}+M_2\xi_1^{(2)}]/4 , \tag{29}$$

which gives

$$p(\xi_1;\mathbf{M})\times\frac{N}{4}-[1-p(\xi_1;\mathbf{M})]\times\frac{N}{4}=[M_1\xi_1^{(1)}+M_2\xi_1^{(2)}]/4 ,$$

thus

$$p(\xi_1;\mathbf{M})=\left[1+\sum_{i=1}^{2}m_i\xi_1^{(i)}\right]\Big/2 . \tag{30}$$

So long as p is kept finite as N becomes large, we can continue the arguments above to obtain generally

$$p(\xi_i;\mathbf{M})=\left[1+\sum_{\mu=1}^{p}m_\mu\xi_i^{(\mu)}\right]\Big/2 . \tag{31}$$

The mean-field term $\sum_j\omega_{ij}A_j$ in Eq. (12) is now ready to calculate, from Eqs. (26) and (31) leading to

$$\sum_{\mathbf{M}'}\omega_{\mathbf{M},\mathbf{M}'}g_{\mathbf{M}'}=\Gamma^{-1}\sum_i\left[g_{\mathbf{M}+2\xi_i}f\left[-1|h_i=N^{-1}\sum_\mu\xi_i^{(\mu)}(M_\mu+\xi_i^{(\mu)})\right]p(\xi_i;\mathbf{M}+2\xi_i)-(\mathbf{M}\to\mathbf{M}-2\xi_i)\right]$$
$$+\Gamma^{-1}\sum_i\left[g_{\mathbf{M}}f\left[1|h_i=N^{-1}\sum_\mu\xi_i^{(\mu)}(M_\mu+\xi_i^{(\mu)})\right][1-p(\xi_i;\mathbf{M})]-(\mathbf{M}\to\mathbf{M}-2\xi_i)\right] , \tag{32}$$

where $(\mathbf{M}\to\mathbf{M}-2\xi_i)$ means that we replace all the $\mathbf{M}$ appearing on the left-hand side by $\mathbf{M}-2\xi_i$. The remaining task is to represent the result in terms of the overlap $\{m_\mu\}$ rather than $\{M_\mu\}$. Since $\sum_{\mathbf{M}} g_{\mathbf{M}}=1$ we define $g(\mathbf{m})\equiv N^p g_{\mathbf{M}}$ which satisfies $\int d\mathbf{m}\, g(\mathbf{m})=1$. After multiplying N^p on both sides of Eq. (32) we use the relation

$$v(\mathbf{m})-v\left[\mathbf{m}-2\frac{\xi_i}{N}\right]=\sum_\mu\left[\frac{2\xi_i^{(\mu)}}{N}\right]\frac{\partial v(\mathbf{m})}{\partial m_\mu} -\sum_{\mu\nu}\left[\frac{1}{2}\right]\left[\frac{2\xi_i^{(\mu)}}{N}\right]\left[\frac{2\xi_i^{(\nu)}}{N}\right]\times\frac{\partial^2 v(\mathbf{m})}{\partial m_\mu\partial m_\nu}+O(N^{-3})$$

to derive finally

$$\frac{\partial g(\mathbf{m},t)}{\partial t}=(N\Gamma)^{-1}\left\{\sum_{i,\mu}\xi_i^{(\mu)}\left[\frac{\partial}{\partial m_\mu}\right]\times\left[\mathbf{m}\cdot\xi_i-\tanh\left[\frac{\mathbf{m}\cdot\xi_i}{T}\right]\right]\times g(\mathbf{m},t)+O\left[\frac{1}{N}\right]\right\}, \tag{33}$$

where the order N^{-1} correction is given by

$$N^{-1}\sum_{i,\mu,\nu}2\xi_i^{(\mu)}\xi_i^{(\nu)}\frac{\partial}{\partial m_\mu}\left[[1-\tanh(\mathbf{m}\cdot\xi_i)]\frac{\partial}{\partial m_\nu}\times\left[\frac{(1+\mathbf{m}\cdot\xi_i)g(\mathbf{m},t)}{4}\right]\right]. \tag{34}$$

Here we give two comments on the GL Eq. (12) with regard to its application to overlap dynamics. The first one is concerned with the damping kernel. For the one-pattern case ($p=1$), which is equivalent to infinite-range ferromagnetic system [6], we have calculated the kernel $\Psi(t=0)$ to the lowest order in N^{-1} to find that it exactly vanishes. This strongly suggests that $\Psi(t)$ vanishes for $t\geq 0$ and $p\geq 1$ to the lowest order in N^{-1}. Secondly, since the random force $f_{\mathbf{M}}(t)$, which is given as $f_i(t)$ in Eq. (12), is orthogonal to $\{g_{\mathbf{M}}\}$, we can safely neglect it in discussing the probability distribution function based on the GL Eq. (12) [10].

Now from Eq. (33) it is seen that the average $\langle m_\mu(t)\rangle\equiv\int d\mathbf{m}\, m_\mu g(\mathbf{m},t)$ follows the equation

$$\frac{d\langle m_\mu(t)\rangle}{dt}=\Gamma^{-1}[-\langle m_\mu(t)\rangle+\langle\langle\xi^{(\mu)}\tanh(\langle\mathbf{m}(t)\rangle\cdot\xi/T)\rangle\rangle], \tag{35}$$

where $\langle\langle\ \rangle\rangle$ denotes the average over the patterns according to the probability law $p(\xi^{(\mu)}=1)=\frac{1}{2}$. In deriving Eq. (35) we have replaced $\langle\tanh(\mathbf{m}\cdot\xi/T)\rangle$ by $\tanh(\langle\mathbf{m}\rangle\cdot\xi/T)$, which is consistent with our mean-field approximation in the limit $N\to\infty$ because in this limit we have only convection of probability and no conduction (no fluctuations), Eq. (33).

Equation (35) has been derived, to the authors' knowledge, by two methods. One is based on the notion of sublattice magnetization [3,13–15] and the other is based on a path-integral formulation of spin dynamics [5]. The sublattice idea [3] is elegant enough to be applied to the more general Hopfield model. The merit of our mean-field approach is that it is based on a general method of irreversible statistical mechanics and thus sheds some light on the implication of the overlap-dynamics equations (33) and (35). The remaining important problem is the case of extensively many patterns, that is, finite $\alpha\equiv p/N$. Overlap dynamics in this case is at present mainly investigated by computer simulations, except for a few theoretical works [4,5,15,16]. With inclusion of some additional variables in the set of collective variables $\mathbf{A}$, it is hoped that overlap dynamics for a finite α case could be handled with the mean-field approximation developed in this paper.

[1] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).
[2] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985); **35**, 2293 (1987).
[3] U. Reidel, R. Kühn, and J. L. van Hemmen, Phys. Rev. A **38**, 1105 (1988).
[4] H. Horner, D. Bormann, M. Frick, H. Kinzelbach, and A. Schmidt, Z. Phys. B **76**, 381 (1989).
[5] H. Rieger, M. Schreckenberg, and J. Zittartz, Z. Phys. B **72**, 523 (1988).
[6] D. J. Amit, *Modeling Brain Function* (Cambridge University Press, Cambridge, England, 1989).
[7] Later we will introduce m_μ (M_μ), which is not a dynamical variable (a function of **S**) like $\hat{m}_\mu$ ($\hat{M}_\mu$) but a number.
[8] R. Zwanzig, Phys. Rev. **124**, 983 (1961).
[9] H. Mori, Prog. Theor. Phys. **33**, 423 (1965).
[10] H. Mori and H. Fujisaka, Progr. Theor. Phys. **49**, 764 (1973).
[11] A. Z. Akcasu, and J. J. Duderstadt, Phys. Rev. **188**, 479 (1969).
[12] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Benjamin, London, 1975).
[13] A. C. C. Coolen, and Th. W. Ruijgrok, Phys. Rev. A **38**, 4253 (1988).
[14] M. Shiino, H. Nishimori, and M. Ono, J. Phys. Soc. Jpn. **58**, 763 (1989).
[15] J. L. Van Hemmen and R. Kühn, in *Models of Neural Networks*, edited by E. Domany, J. L. van Hemmen, and K. Schulten (Springer, Berlin, 1991).
[16] S. Amari and K. Maginu, Neural Networks **1**, 63 (1988).