Correlated-data-driven dynamics in a neural network

Y. Nakamura and T. Munakata
Department of Applied Mathematics and Physics, Kyoto University, Kyoto 606, Japan
(Received 25 June 1993)

Effects of correlations among the embedded patterns on dynamics in a neural network are studied. Combined with a delay of signal transmission, the correlations are shown to give rise to a time sequence of patterns.

PACS number(s): 87.10.+e, 02.60.Cb

A number of neural network models have been proposed to explain some properties of the nervous system in terms of the formal two-state neurons [1]. In the Hopfield model [2] with N neurons, the symmetric connection J_{ij} between the neurons i and j is expressed as

$$J_{ij}=\frac{1}{N}\sum_{\mu}\xi_i^{\mu}\xi_j^{\mu}\ , \tag{1}$$

where ξ_i^{μ} $(=\pm 1)$ $(\mu=1,\ldots,p;\ i=1,\ldots,N)$ denotes the state of the neuron i in the μth pattern. The postsynaptic potential (PSP) or the field strength $h_i(t)$ at the neuron i is expressed as

$$h_i(t)=\sum_j J_{ij}S_j(t)=\sum_{\mu}\xi_i^{\mu}m_{\mu}(t)\ , \tag{2}$$

where the overlap $m_{\mu}(t)$ with the μth pattern is defined by

$$m_{\mu}(t)=\frac{1}{N}\sum_i \xi_i^{\mu}S_i(t)\ . \tag{3}$$

The collective behavior of the network with a symmetric connection such as Eq. (1) is represented as a relaxation process toward the local minimum of the (free) energy and the model works as content-addressable or associative memories [2].

Recently issues of temporal association in neural networks have gathered considerable attention [3–7]. To effect transitions between patterns, a certain amount of asymmetry of the connection J_{ij} is required. After the proposal suggested by Hopfield [2], some models have been put forth in which delay in signal transmission is introduced to yield, e.g.,

$$h_i(t)=\sum_{\mu}\xi_i^{\mu}m_{\mu}(t)+\varepsilon\sum_{\mu}\xi_i^{\mu+1}m_{\mu}(t-\tau)\ . \tag{4}$$

For later convenience we briefly consider the transition mechanism within the framework of asynchronous dynamics [6]. We suppose that the system has been in the pattern ν for $0\le t<\tau$. As for $m_{\mu}(t-\tau)$ in Eq. (4), we can set $m_{\mu}(t-\tau)=C_{\mu\nu}$ for $\tau\le t<2\tau$, where the correlation $C_{\mu\mu'}$ between the patterns μ and μ' is defined by

$$C_{\mu\mu'}=\frac{1}{N}\sum_i \xi_i^{\mu}\xi_i^{\mu'}\ . \tag{5}$$

If we consider a network in which the correlations vanish on average, that is,

$$\bar{C}_{\mu\mu'}=\frac{1}{N}\sum_i\langle\xi_i^{\mu}\xi_i^{\mu'}\rangle=\delta_{\mu\mu'}\ , \tag{6}$$

and neglect fluctuations in $C_{\mu\mu'}$ of order $1/\sqrt{N}$, we have $m_{\mu}(t-\tau)=\delta_{\mu\nu}$ for $\tau\le t<2\tau$. If ε is chosen to be larger than 1 and τ sufficiently longer than one cycle time [or one Monte Carlo time (MCT) in which N neurons are updated], we see that $m_{\nu+1}(t)$ increases until it becomes 1 at about $t\simeq\tau+t_{\mathrm{MCT}}$. During the time other overlaps $m_{\mu}(t)$ $(\mu\neq\nu+1)$ decrease to zero. Thus after the pattern ν, the pattern $\nu+1$ is retrieved, then the pattern $\nu+2$, and so on. It should be noted that once $m_{\nu+1}(t)$ becomes large the first term on the right-hand side (rhs) of Eq. (4) contributes to the stabilization of the pattern $\nu+1$.

The model (4) and the modifications thereof have been successful in producing a pattern sequence [3–5]. We note, however, that the correlations among patterns $\bar{C}_{\mu\mu'}$ $(\mu\neq\mu')$ are assumed to be zero, Eq. (6), in order to enable the network to retrieve the prescribed sequence $\nu,\nu+1,\ldots$. Putting the model (4) aside for a while, we consider what happens when there are strong correlations among patterns in a network (brain). If the network is in a state (pattern) A and the pattern B alone is strongly correlated with the pattern A, it is expected that a spontaneous transition, if it occurs, is to the state B. That is, one can conceive that the internal correlations among the embedded patterns (data) can give rise to spontaneous or data-driven transitions. The purpose of this Brief Report is to present and investigate a model which materializes the notion above.

Our model consists of the PSP given by

$$h_i(t)=\sum_{\mu}\varepsilon_{\mu}\xi_i^{\mu}m_{\mu}(t)+\sum_{\substack{\mu,\nu\\ \mu\neq\nu}}\varepsilon_{\mu}^{T}\xi_i^{\mu}m_{\mu}(t)m_{\nu}(t-\tau)\ . \tag{7}$$

The correlations $\bar{C}_{\mu\mu'}$ defined by Eq. (6) now take nonzero values. Qualitatively the second (transition) term on the rhs of Eq. (7) is interpreted as follows: When the network is in the pattern ν_1 for $0\le t<\tau$, $m_{\nu}(t-\tau)$ in Eq. (7) is $C_{\nu\nu_1}$ for $\tau\le t<2\tau$. In the summation over ν in Eq. (7) the largest contribution comes from the term $\nu=\nu_1$ since $C_{\nu\nu_1}=1$ for $\nu=\nu_1$. By $[\nu_1]$ let us denote the pattern $(\neq\nu_1)$ which is most strongly correlated with the pattern ν_1. Now for t slightly larger than τ $(t\gtrsim\tau)$, $m_{\mu}(t)$ in Eq. (7) is nearly equal to $C_{\mu\nu_1}$. Since μ must be different from ν_1, the largest contribution in the summation over μ comes

from the term $\mu=[\nu_1]$ and the PSP tends to drive the system toward the state $[\nu_1]$. Once this transition is initiated, the first term also begins supporting the transition. Of course this is a drastically simplified scenario and more quantitative argument is necessary to understand the transition mechanism in Eq. (7).

For the purpose we confine ourselves to the case where only three patterns A, B, and C are embedded in the network. The PSP is explicitly written as

$$\begin{aligned} h_i(t)=&[\varepsilon_A m_A+\varepsilon_A^T m_A(\bar{m}_B+\bar{m}_C)]\xi_i^A \\ &+[\varepsilon_B m_B+\varepsilon_B^T m_B(\bar{m}_A+\bar{m}_C)]\xi_i^B \\ &+[\varepsilon_C m_C+\varepsilon_C^T m_C(\bar{m}_A+\bar{m}_B)]\xi_i^C , \end{aligned} \tag{8}$$

where ε_μ and ε_μ^T ($\mu=A,B,C$) are assumed to be positive and m_A and $\bar{m}_A$ denote $m_A(t)$ and $m_A(t-\tau)$, respectively. Suppose that at $t=0$ the system is in a state A and the PSP at the neuron i becomes

$$h_i(0)=\varepsilon_A\xi_i^A+\varepsilon_B C_{AB}\xi_i^B+\varepsilon_C C_{AC}\xi_i^C , \tag{9}$$

where the transition term in Eq. (7) is regarded to be effective after $t=\tau$. If

$$\varepsilon_A \geq \varepsilon_B C_{AB}+\varepsilon_C C_{AC} , \tag{10}$$

the system remains in the state A until $t=\tau$. At $t=\tau$, the PSP is given by Eq. (8) by setting $m_\mu=C_{\mu A}=\bar{m}_\mu$. Explicitly it reads as

$$\begin{aligned} h_i(t)=&[\varepsilon_A+\varepsilon_A^T(C_{AB}+C_{AC})]\xi_i^A \\ &+[\varepsilon_B C_{AB}+\varepsilon_B^T C_{AB}(1+C_{AC})]\xi_i^B \\ &+[\varepsilon_C C_{AC}+\varepsilon_C^T C_{AC}(1+C_{AB})]\xi_i^C \\ \equiv& h_A\xi_i^A+h_B\xi_i^B+h_C\xi_i^C . \end{aligned} \tag{11}$$

We now require that

$$h_B > h_A \quad \text{and} \quad h_B > h_C . \tag{12}$$

Under the condition (12), if $\xi_i^A=\xi_i^B$, S_i ($=\xi_i^A$) is unchanged and if $\xi_i^B=\xi_i^C\neq\xi_i^A$, S_i flips from ξ_i^A to $-\xi_i^A=\xi_i^B$. Since we do not require $h_B > h_A+h_C$, which is too severe a condition to be satisfied, when $\xi_i^A=\xi_i^C\neq\xi_i^B$, it seems that the transition from $S_i=\xi_i^A$ to ξ_i^B is not assured. However, as explained above, in the process of asynchronous updating, $S_i(t)$ takes the value ξ_i^B except for the neuron site, where $\xi_i^A=\xi_i^C\neq\xi_i^B$ and the coalition of A and C patterns resists the transition to the state B. At this point let us express the fraction Z of the coalition sites in terms of the correlations. If X, Y, Z, and $(1-X-Y-Z)$ denote the fractions of neuron sites at which $\xi_i^A=\xi_i^B=\xi_i^C$, $\xi_i^A=\xi_i^B\neq\xi_i^C$, $\xi_i^A=\xi_i^C\neq\xi_i^B$, and $\xi_i^B=\xi_i^C\neq\xi_i^A$, respectively, we have

$$\begin{aligned} C_{AB}&=2(X+Y)-1 , \\ C_{AC}&=2(X+Z)-1 , \\ C_{BC}&=1-2(Y+Z) . \end{aligned} \tag{13}$$

These equations lead to

$$Z=(1+C_{AC}-C_{BC}-C_{AB})/4 . \tag{14}$$

From the discussions above we see that the overlap m_B increases from C_{AB} at time $t=\tau$ to $m_B=1-2Z=(1+C_{BC}+C_{AB}-C_{AC})/2$ at $t=\tau+t_{\mathrm{MCT}}$ and the system approaches the state B considerably as exemplified later by our numerical simulation. We now write the conditions to be satisfied in order to keep the state B after $t=\tau+t_{\mathrm{MCT}}$, Eq. (15), to effect the transition to the state C after $t=2\tau$, Eq. (16), to keep the state C after $t=2\tau+t_{\mathrm{MCT}}$, Eq. (17), and finally to keep the state C forever ($t>3\tau$), Eq. (18):

$$\varepsilon_B+\varepsilon_B^T+\varepsilon_B^T C_{AC} \geq \varepsilon_A C_{AB}+\varepsilon_A^T(C_{AB})^2+\varepsilon_A^T C_{AB}C_{AC}+\varepsilon_C C_{BC}+\varepsilon_C^T C_{BC}+\varepsilon_C^T C_{BC}C_{AB} , \tag{15}$$

$$\varepsilon_C C_{BC}+\varepsilon_C^T C_{AB}C_{BC}+\varepsilon_C^T C_{BC} > \varepsilon_A C_{AB}+\varepsilon_A^T C_{AB}+\varepsilon_A^T C_{AB}C_{BC} , \quad \varepsilon_B+\varepsilon_B^T C_{AB}+\varepsilon_B^T C_{BC} , \tag{16}$$

$$\varepsilon_C+\varepsilon_C^T C_{AB}+\varepsilon_C^T \geq \varepsilon_A C_{AC}+\varepsilon_A^T C_{AC}+\varepsilon_A^T C_{AC}C_{BC}+\varepsilon_B C_{BC}+\varepsilon_B^T C_{BC}C_{AB}+\varepsilon_B^T(C_{BC})^2 , \tag{17}$$

$$\varepsilon_C+\varepsilon_C^T C_{AC}+\varepsilon_C^T C_{BC} \geq \varepsilon_A C_{AC}+\varepsilon_A^T C_{AC}C_{BC}+\varepsilon_A^T C_{AC}+\varepsilon_B C_{BC}+\varepsilon_B^T C_{BC}C_{AC}+\varepsilon_B^T C_{BC} . \tag{18}$$

If the conditions (10), (12), and (15)–(18) are satisfied we would have the sequence $A\rightarrow B\rightarrow C$, although the transition may not be perfect as noted in connection with the requirement Eq. (12) [see Eq. (16) also].

We now turn to computer simulations of the model (7). The stochastic Glauber dynamics is employed, in which the probability for the neuron i to take ±1 at time $t+\Delta t$ is given by [8,9]

$$\mathrm{Prob}\{S_i(t+\Delta t)=\pm1\}=\{1\pm\tanh[h_i(t)/T]\}/2 . \tag{19}$$

When the temperature of the system T goes to zero, the Glauber dynamics is reduced to the threshold dynamics

$$S_i(t+\Delta t)=\mathrm{sgn}\{h_i(t)\} , \tag{20}$$

with $\mathrm{sgn}(x)=1$ for $x\geq0$ and $\mathrm{sgn}(x)=-1$ for $x<0$. Three patterns A, B, and C are embedded in the system consisting of $N=400$ neurons and the delay τ is taken to be $3t_{\mathrm{MCT}}$. When three correlations C_{AB}, C_{BC}, and C_{AC} are given, we solve Eq. (13) to obtain X, Y, and Z. We first produce the pattern, say A, by using N random numbers. Then the pattern B is determined by either taking $\xi_i^B=\xi_i^A$ for the X and Y regions and $\xi_i^B=-\xi_i^A$ for the remaining region. The pattern C is made in the same way.

Figure 1 shows the overlap dynamics for $T=0$ and the parameters $\varepsilon_A=0.6$, $\varepsilon_B=0.8$, $\varepsilon_C=1.0$, $\varepsilon_A^T=1.0$, $\varepsilon_B^T=1.9$, $\varepsilon_C^T=3.0$, $C_{AB}=0.4$, $C_{BC}=0.5$, and $C_{AC}=0.2$. Although the pattern B is not fully retrieved in the intermediate-time region, we see that a transition sequence $A\rightarrow B\rightarrow C$ is realized with the parameters given above, which satisfy all the requirements listed before. The asymptotic ($t=\infty$) value of each overlap is seen to be consistent with the correlations in our data. We performed a simulation on a system with three uncorrelated

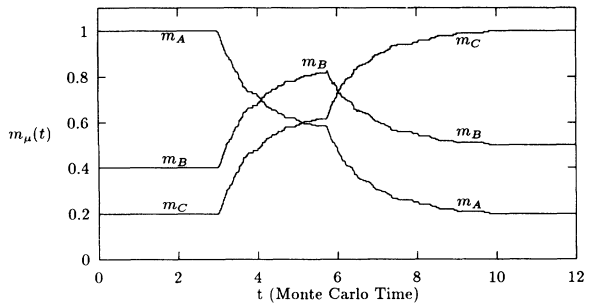

FIG. 1. Simulation results for the overlap dynamics. $T=0$, $C_{AB}=0.4$, $C_{BC}=0.5$, $C_{AC}=0.2$, $\varepsilon_A=0.6$, $\varepsilon_B=0.8$, $\varepsilon_C=1.0$, $\varepsilon_A^T=1.0$, $\varepsilon_B^T=1.9$, and $\varepsilon_C^T=3.0$.

patterns (other parameters are the same as above) and observed that all the overlaps remain constant, showing no sign of transitions. To see effects of finite temperature, we first change ε_B^T from 1.9 to 2.0, keeping other parameters as in Fig. 1. In this case the requirement Eq. (16) is not satisfied and, reflecting this, we notice in Fig. 2 that the sequence retrieval is incomplete. Now we change T from zero to 0.06 and the result is shown in Fig. 3, which reveals the familiar fact that fluctuations induced by temperature work positively for retrieval of patterns as long as it is not too high [8].

Here we give two comments on our model. The first one is concerned with the robustness of our scheme to the variation of the parameter values. To be concrete we change ε_B^T and ε_C^T with all the other parameters fixed as above. First we consider the case $T=0$. If we set $\varepsilon_C^T=3.0$ (the case of our simulation), ε_B^T should be in the range $1.9\leq\varepsilon_B^T<2.0$ in order to satisfy the conditions (10), (12), and (15)–(18). On the other hand, when we set $\varepsilon_C^T=4.0$, ε_B^T is allowed to be in a little wider range $2.483\leq\varepsilon_B^T<2.778$. When the temperature is nonzero, the range of ε_B^T for successful retrieve of the state C becomes large. For example, a retrieval process $A\rightarrow B\rightarrow C$ is confirmed for the case $\varepsilon_B^T=2.5$ and $\varepsilon_C^T=3.0$. However it is remarked that when T is very small (e.g., 0.01) one must wait a longer time for the retrieve of the state C and that when T is too large (e.g., 0.3) the state B is skipped, resulting in the retrieval process $A\rightarrow C$.

The second comment is on the number of patterns embedded in our system. When a group of correlated patterns in a sequence increases its members (in our simulations only three members A,B,C), the tuning of the parameters, including the correlations $\{C_{\mu\nu}\}$, naturally becomes severe. However, we can increase the number of patterns in our system by embedding many groups, each containing three patterns which are uncorrelated with patterns belonging to all the other groups. For example, in a system consisting of 60 neurons, we could embed three groups of correlated patterns $[(A_i,B_i,C_i), i=1,2,3]$.

Finally, we consider overlap dynamics theoretically in the thermodynamic limit $N\rightarrow\infty$ with the number of the pattern, p, kept finite ($p/N\rightarrow 0$) [10,11]. Our analysis is based on the notion of sublatice $I(\mathbf{x})$ and sublattice magnetization $m(\mathbf{x},t)$, which are due to van Hemmen and co-workers [10,6]. Given p binary patterns, we have N vectors $\xi_i=(\xi_i^1,\ldots,\xi_i^p)$ $(i=1,\ldots,N)$. Introducing the sublattice by

$$I(\mathbf{x})=\{i;\xi_i=\mathbf{x}\}\ ,\quad \mathbf{x}\in\{-1,1\}^p\ , \tag{21}$$

and the sublattice magnetization by

$$m(\mathbf{x},t)=\frac{1}{|I(\mathbf{x})|}\sum_{i\in I(\mathbf{x})}S_i(t)\ , \tag{22}$$

with $|I(\mathbf{x})|$ denoting the size of $I(\mathbf{x})$, one can easily verify that the PSP, Eq. (7), is expressed as

$$h_i(t)=\sum_{\mu,\mathbf{x}}\xi_i^\mu p_N(\mathbf{x})x_\mu m(\mathbf{x},t)\times\left\{\varepsilon_\mu+\varepsilon_\mu^T\sum_{\substack{\nu,\mathbf{y}\\ \nu\neq\mu}}p_N(\mathbf{y})y_\nu m(\mathbf{y},t-\tau)\right\}\ . \tag{23}$$

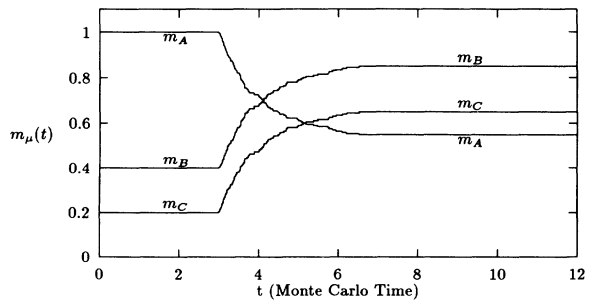

FIG. 2. Simulation results for the overlap dynamics. $\varepsilon_B^T=2.0$ and other conditions are the same as those for Fig. 1.

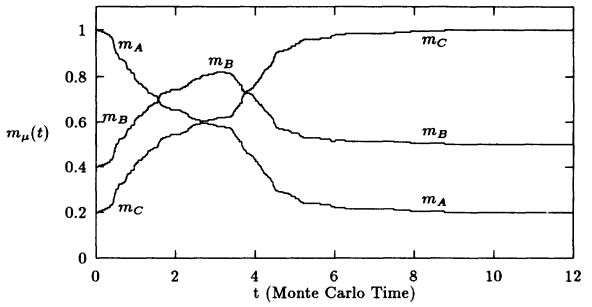

FIG. 3. Simulation results for the overlap dynamics. $T=0.06$ and other conditions are the same as those for Fig. 2.

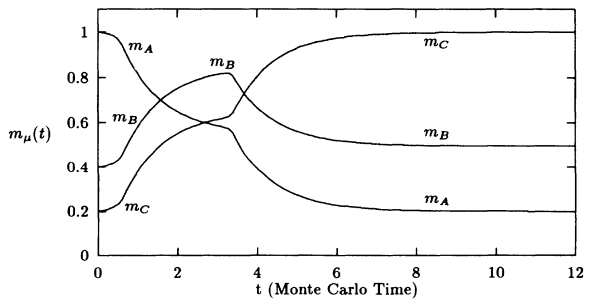

FIG. 4. Numerical solutions of Eq. (26). The conditions are the same as those for Fig. 3.

Here $p_N(\mathbf{x})=|I(\mathbf{x})|/N$ denotes the weight of the sublattice $I(\mathbf{x})$, which depends on the correlations among the patterns. Since $\xi_i=\mathbf{x}$ for all i in $I(\mathbf{x})$, the PSP depends on i only through the label $\mathbf{x}$ of the sublattice to which i belongs. Thus Eq. (23) can be written as

$$h(\mathbf{x},t)=\sum_{\mu,\mathbf{y}} x_\mu p_N(\mathbf{y}) y_\mu m(\mathbf{y},t) \times \left\{ \varepsilon_\mu+\varepsilon_\mu^T \sum_{\substack{\nu,\mathbf{z} \\ \nu\neq\mu}} p_N(\mathbf{z}) z_\nu m(\mathbf{z},t-\tau) \right\} . \quad (24)$$

It is well known that in the thermodynamic limit the sublattice magnetization $m(\mathbf{x},t)$ is governed by the following set of 2^p coupled differential equations [6]:

$$\frac{dm(\mathbf{x},t)}{dt}=-\{m(\mathbf{x},t)-\tanh[h(\mathbf{x},t)/T]\} . \quad (25)$$

Since the overlap $m_\mu(t)$ is expressed in terms of $m(\mathbf{x},t)$ as

$$m_\mu(t)=\sum_{\mathbf{x}} p(\mathbf{x}) x_\mu m(\mathbf{x},t) = \langle x_\mu m(\mathbf{x},t)\rangle ,$$

we finally obtain a set of p coupled equations for the overlaps,

$$\frac{dm_\mu(t)}{dt}=-m_\mu(t)+\left\langle x_\mu \tanh\left[\frac{1}{T} \sum_\nu x_\nu m_\nu(t) \left\{ \varepsilon_\nu+\varepsilon_\nu^T \sum_{\nu'(\neq\nu)} m_{\nu'}(t-\tau) \right\} \right] \right\rangle . \quad (26)$$

The second term on the rhs of Eq. (26) consists of 2^p terms, since there are 2^p possibilities for $\mathbf{x}$. We solve Eq. (26) with several values of T and the solutions to Eq. (26) are confirmed to be similar to our simulations. In Fig. 4 we show the numerical results for the overlaps obtained from Eq. (26). The parameters are the same with those for Fig. 3. From Figs. 3 and 4 it is seen that Eq. (26) reproduces our experimental results for $N=400$ and $p=3$ excellently.

In this Brief Report we investigated neural networks with correlated patterns. For the dynamics we proposed Eq. (7) as a model for sequential association. Concerning this model we obtain the following results.

(1) With the appropriate values for the parameters ε_μ and ε_μ^T, the sequence was generated and finally the system settled in a target pattern. Effects of temperature turned out to be supportive for the retrieval of the sequence.

(2) Our model, Eq. (7), does not itself induce transitions as stressed in connection with the discussions on Fig. 1. The model together with the correlations among the patterns make the sequence retrieval possible.

(3) Coupled nonlinear differential equations (26) were derived for the model (7). Here also, the correlations among the patterns is implicit. Numerical solutions to these equations are well correlated with the simulation results.

[1] W. A. Little, Math. Biosci. **6**, 101 (1974).
[2] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).
[3] H. Sompolinsky and I. Kanter, Phys. Rev. Lett. **57**, 2861 (1986).
[4] D. Kleinfeld, Proc. Natl. Acad. Sci. USA **83**, 9469 (1986).
[5] U. Riedel, R. Kühn, and J. L. van Hemmen, Phys. Rev. A **38**, 1105 (1988).
[6] R. Kühn and J. L. van Hemmen, in *Models of Neural Networks,* edited by E. Domany, J. L. van Hemmen, and K. Schulten (Springer-Verlag, Berlin, 1991).
[7] D. J. Amit, Proc. Natl. Acad. Sci. USA **85**, 2141 (1988).
[8] D. J. Amit, *Modeling Brain Function* (Cambridge University Press, Cambridge, 1989).
[9] P. Peretto, *An Introduction to the Modeling of Neural Networks* (Cambridge University Press, Cambridge, 1992).
[10] J. L. van Hemmen and R. Kühn, Phys. Rev. Lett. **57**, 913 (1988).
[11] T. Munakata and Y. Nakamura, Phys. Rev. E **47**, 3791 (1993).