

Macroscopic dynamics in separable neural networks

Chen Yong,* Wang Yinghai, and Yang Kongqing

Department of Physics, Lanzhou University, Lanzhou Gansu, 730000, China

(Received 18 September 2000; revised manuscript received 27 November 2000; published 16 March 2001)

Parallel dynamics of neural networks with separable coupling is given starting from Coolen-Sherrington theory. Away from saturation, it is shown that this parallel retrieve dynamics is equivalent to the sequential dynamics for finite temperature. But the finite-size effects were found to be governed by a homogeneous Markov process, not by a time-dependent Ornstein-Uhlenbeck process in sequential dynamics.

DOI: 10.1103/PhysRevE.63.041901

PACS number(s): 87.10.+e, 75.10.Nr, 02.50.-r

Ising spin models for neural networks have made increasingly significant contributions to the understanding of the information processing in biotic nervous systems following the pioneering work of Little [1], Hopfield [2], and Amit *et al.* [3]. Moreover, starting from the theory of equilibrium statistical mechanics of spin-glass-like systems, a survey of the properties of Hopfield model with symmetric synapses is presented by Amit *et al.* [3,4], for far and near saturation. A static analysis was found to be sufficient to investigate the networks with symmetric coupling. However, when one deals with networks with asymmetric connections, which are ubiquitous in real neurons of living systems [5], it is more important to consider the dynamics rather than the equilibrium properties. Recently, Coolen *et al.* developed a series of analytic sequential dynamical schemes for systems with separable synapses [6–8]. It is clear that there exist two types of deterministic dynamics, sequential and parallel, in the time evolution of states in neural systems. One may ask: (1) What form does a macroscopic description of these systems take for the synchronous case and are there differences between the two updating cases, synchronous and asynchronous? (2) Additionally, how do the finite-size effects in the parallel case compare with those in the sequential case presented by Castellanos *et al.* [9]?

Since the systems studied include the asymmetric conjunctions, a stochastic analysis has been employed. The use of Markovian dynamics has the advantage of providing a simple description of the stochastic processes and including their nonequilibrium properties [10]. For this reason, we take up Markov analysis and take Coolen-Sherrington (CS) theory as the base in our paper.

As usual in an Ising spin model of neural networks consisting of N neurons, the time-dependent state of the i th neuron in time t is described by $s_i(t) \in \{1, -1\}$. The network has p sets of patterns $\xi_i^\mu \in \{1, -1\}$ ($\mu = 1, 2, \dots, p$; $i = 1, 2, \dots, N$), which are embedded for the purpose of associative memory retrieval through the synaptic connections with the Hebb learning rule taken into account [11]. The separable interaction matrices \mathbf{J} was taken as [6,12]

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

where \mathbf{A} is a $p \times p$ matrix, used to represent all kinds of conjunctions. Then, the coupling synapses, both symmetric and asymmetric, can be represented by Eq. (1). Furthermore, the time evolution of the synchronous dynamics in these systems can be defined as

$$s_i(t + \Delta t) = \text{sgn} \left(\sum_{j=1}^N J_{ij} s_j(t) \right), \quad (2)$$

where Δt is the length of each time step. In Eq. (2), the threshold of every nerve cell was set as a constant, zero, and the updating function is a step function $\text{sgn}(x)$ that has the value 1 for $x \geq 0$ and -1 for $x < 0$.

Now, the microscopic state probability $P_t(\mathbf{s})$ denotes the probability of finding the entire system of N neurons in a compound state $\mathbf{s} = (s_1, s_2, \dots, s_N)$ at time t . It is well-known that the continuous time version of Markov process can be described by

$$\frac{dP_t(\mathbf{s})}{dt} = \sum_{\mathbf{s}' \neq \mathbf{s}} w(\mathbf{s}' \rightarrow \mathbf{s}) P_t(\mathbf{s}') - \sum_{\mathbf{s}' \neq \mathbf{s}} w(\mathbf{s} \rightarrow \mathbf{s}') P_t(\mathbf{s}) \quad (3)$$

where $w(\mathbf{s}' \rightarrow \mathbf{s})$ represents the density of unit time for transition probability (it is equal to transition rate), which the states of the system change from \mathbf{s}' to \mathbf{s} , in the interval between t and $t + \Delta t$. For a parallel evolutionary process, the usual form of the transition rate can be written as by [12,13]

$$w(\mathbf{s}' \rightarrow \mathbf{s}) = \prod_{i=1}^N \frac{1}{2} \{1 + s_i \tanh[\beta h_i(\mathbf{s}')]\}, \quad (4)$$

where $h_i(\mathbf{s}') \equiv \sum_j J_{ij} s_j'$ is the local field of a stochastic alignment of the spins and $\beta \equiv 1/T$ denotes a measure of the inverse magnitude of the amount of noise affecting the neurons, or we may say that β^{-1} acts as the role of temperature in analogy to a thermodynamic spin system.

This paper focuses, in general, more on the problem of macroscopic features than the microscopic details of a network. Toward this end, a set of linear order variables Ω_μ is introduced [6,14], defined by

$$\Omega_\mu(\mathbf{s}) = \frac{1}{N} \sum_j \xi_j^\mu s_j \quad P_t(\boldsymbol{\Omega}) = \sum_{\mathbf{s}} P_t(\mathbf{s}) \delta[\boldsymbol{\Omega} - \boldsymbol{\Omega}(\mathbf{s})] \quad (5)$$

$$\mu = 1, 2, \dots, p,$$

*Electronic address: ychen@lzu.edu.cn

where $\mathbf{\Omega}$ is pattern overlaps, a macroscopic description of the process, which measure the similarity between the state of this system and each of the P stored patterns, and $P_t(\mathbf{\Omega})$ denotes the probability of finding $\mathbf{\Omega}$ in time t . Using the above definitions and the Eq. (3), it is easy to get

$$\begin{aligned} \frac{dP_t(\mathbf{\Omega})}{dt} &= \sum_{\mathbf{s}} \sum_{\mathbf{s}' \neq \mathbf{s}} w(\mathbf{s}' \rightarrow \mathbf{s}) P_t(\mathbf{s}') \delta(\mathbf{\Omega} - \mathbf{\Omega}(\mathbf{s})) \\ &\quad - \sum_{\mathbf{s}} \sum_{\mathbf{s}' \neq \mathbf{s}} w(\mathbf{s} \rightarrow \mathbf{s}') P_t(\mathbf{s}) \delta(\mathbf{\Omega} - \mathbf{\Omega}(\mathbf{s})) \end{aligned} \quad (6)$$

Obviously, the above master equation can be simplified to the equivalent expression

$$\begin{aligned} \frac{dP_t(\mathbf{\Omega})}{dt} &= \sum_{\mathbf{s}} \sum_{\mathbf{s}'} w(\mathbf{s} \rightarrow \mathbf{s}') P_t(\mathbf{s}) \{ \delta[\mathbf{\Omega} - \mathbf{\Omega}(\mathbf{s}')] \\ &\quad - \delta[\mathbf{\Omega} - \mathbf{\Omega}(\mathbf{s})] \}. \end{aligned} \quad (7)$$

Following the CS theory [7], one introduces any function $\Phi(\mathbf{\Omega})$ and its average is $\langle \Phi(\mathbf{\Omega}) \rangle_t \equiv \int d\mathbf{\Omega} P_t(\mathbf{\Omega}) \Phi(\mathbf{\Omega})$. Its time differential is

$$\begin{aligned} \langle \Phi(\mathbf{\Omega}) \rangle_t &= \sum_{\mathbf{s}} \sum_{\mathbf{s}'} w(\mathbf{s} \rightarrow \mathbf{s}') P_t(\mathbf{s}) [\Phi(\mathbf{\Omega}(\mathbf{s}')) - \Phi(\mathbf{\Omega}(\mathbf{s}))] \\ &= \sum_{\mathbf{s}} \sum_{\mathbf{s}'} w(\mathbf{s} \rightarrow \mathbf{s}') P_t(\mathbf{s}) \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\mu_1=1}^n \sum_{\mu_2=1}^n \dots \\ &\quad \times \sum_{\mu_p=1}^n [\mathbf{\Omega}(\mathbf{s}') - \mathbf{\Omega}(\mathbf{s})]^{\mu_1 + \mu_2 + \dots + \mu_p} \\ &\quad \times \frac{\partial^n \Phi[\mathbf{\Omega}(\mathbf{s})]}{\partial \Omega_{\mu_1} \partial \Omega_{\mu_2} \dots \partial \Omega_{\mu_p}} \end{aligned} \quad (8)$$

where there exists $\sum_{i=1}^p \mu_i = n$. Inserting the unit operator $\int d\mathbf{\Omega} \delta(\mathbf{\Omega} - \mathbf{\Omega}(\mathbf{s}))$ and then performing partial integrations yields

$$\begin{aligned} \frac{dP_t(\mathbf{\Omega})}{dt} &= - \sum_{\mathbf{s}} \sum_{\mathbf{s}'} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\mu_1=1}^n \sum_{\mu_2=1}^n \dots \\ &\quad \times \sum_{\mu_p=1}^n [\mathbf{\Omega}(\mathbf{s}') - \mathbf{\Omega}(\mathbf{s})]^{\mu_1 + \mu_2 + \dots + \mu_p} \\ &\quad \times \frac{\partial^n w(\mathbf{s} \rightarrow \mathbf{s}') P_t(\mathbf{s}) \delta(\mathbf{\Omega} - \mathbf{\Omega}(\mathbf{s}))}{\partial \Omega_{\mu_1} \partial \Omega_{\mu_2} \dots \partial \Omega_{\mu_p}} \end{aligned} \quad (9)$$

Briefly, Eq. (9) is in the form of a Kramers-Moyal-like expansion for the master equation (7) for the macroscopic state (pattern overlaps) probability.

The properties of pattern overlaps are now investigated by considering only the lowest-order term in the expansion of Eq. (9). Firstly, Eq. (9) can be written as

$$\begin{aligned} \frac{dP_t(\mathbf{\Omega})}{dt} &= - \sum_{\mu} \frac{\partial}{\partial \Omega_{\mu}} \left\{ \sum_{\mathbf{s}} \sum_{\mathbf{s}'} [\mathbf{\Omega}(\mathbf{s}') - \mathbf{\Omega}(\mathbf{s})] w(\mathbf{s} \rightarrow \mathbf{s}') \right. \\ &\quad \left. \times P_t(\mathbf{s}) \delta(\mathbf{\Omega} - \mathbf{\Omega}(\mathbf{s})) \right\} + \sum_{m \geq 2} \mathcal{O} \left(NP \left(\frac{2}{N} \right)^m \right). \end{aligned} \quad (10)$$

Then, in the limit of far away from saturation and the pre-supposition of finite temperature, Eq. (10) becomes

$$\begin{aligned} \frac{dP_t(\mathbf{\Omega})}{dt} &= - \sum_k \frac{\partial}{\partial \Omega_k} \left\{ \sum_{\mathbf{s}} \left[\sum_{\mathbf{s}'} [\mathbf{\Omega}(\mathbf{s}') - \mathbf{\Omega}(\mathbf{s})] w(\mathbf{s} \rightarrow \mathbf{s}') \right] \right. \\ &\quad \left. \times P_t(\mathbf{s}) \delta(\mathbf{\Omega} - \mathbf{\Omega}(\mathbf{s})) \right\} \end{aligned} \quad (11)$$

Next, the following simplification is introduced, $\Delta_i = \frac{1}{2} [1 - s_i \tanh(\beta h_i(\mathbf{s}))]$ and $\Delta'_i = \frac{1}{2} [1 + s_i \tanh(\beta h_i(\mathbf{s}))]$, corresponding to $w(s_i \rightarrow -s_i)$ and $w(s_i \rightarrow s_i)$ respectively, in the Eq. (4). According to the definitions of the order parameters $\mathbf{\Omega}$ in Eq. (5) and the coupling matrix \mathbf{J} in Eq. (1), the stochastic local field $h_i(\mathbf{s})$ of Eq. (4) can be written as

$$h_i(\mathbf{s}) = \frac{1}{N} \sum_j J_{ij} s_j = \sum_{\mu, \nu} \xi_i^{\mu} A_{\mu\nu} \left(\frac{1}{N} \sum_j \xi_j^{\nu} s_j \right) = \xi_i \cdot \mathbf{A} \mathbf{\Omega}.$$

Then, the transition rate of a single neuron becomes

$$\Delta_i = \frac{1}{2} (1 - s_i \tanh[\beta \xi_i \cdot \mathbf{A} \mathbf{\Omega}])$$

$$\Delta'_i = \frac{1}{2} (1 + s_i \tanh[\beta \xi_i \cdot \mathbf{A} \mathbf{\Omega}])$$

Therefore, for the terms of $[\dots]$ on the right side of Eq. (11), from the definitions in Eq. (4) and Eq. (5), we get

$$\begin{aligned} &\sum_{\mathbf{s}'} [\mathbf{\Omega}(\mathbf{s}') - \mathbf{\Omega}(\mathbf{s})] w(\mathbf{s} \rightarrow \mathbf{s}') \\ &= - \frac{2}{N} \left(\sum_{i=1}^N \frac{\xi_i s_i (\Delta_i / \Delta'_i)}{1 + \Delta_i / \Delta'_i} \right) \left[\prod_{j=1}^N \left(1 + \frac{\Delta_j}{\Delta'_j} \right) \right] \left(\prod_k \Delta'_k \right) \\ &= - 2 \frac{1}{N} \sum_{i=1}^N \xi_i s_i \Delta_i = - \mathbf{\Omega}(\mathbf{s}) + \frac{1}{N} \sum_i \xi_i \tanh(\beta \xi_i \cdot \mathbf{A} \mathbf{\Omega}) \end{aligned} \quad (12)$$

Substituting the above relation into Eq. (11) yields

$$\begin{aligned} \frac{d}{dt} P_t(\mathbf{\Omega}) &= - \sum_{\mu} \frac{\partial}{\partial \Omega_{\mu}} \left\{ P_t(\mathbf{\Omega}) \left(- \mathbf{\Omega}(\mathbf{s}) + \frac{1}{N} \right. \right. \\ &\quad \left. \left. \times \sum_{i=1}^N \xi_i \tanh(\beta \xi_i \cdot \mathbf{A} \mathbf{\Omega}) \right) \right\}. \end{aligned} \quad (13)$$

Finally, on the basis of Markov process theory [15], the deterministic Liouville form of Eq. (13) is

$$\frac{d}{dt}\Omega_\mu = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N \xi_i^\mu \tanh(\beta \xi_i \cdot \mathbf{A}\Omega) - \Omega_\mu \right) \quad (14)$$

with the initial value can be set as $\Omega(0) = \Omega_0 = \frac{1}{N} \sum_{i=1}^N \xi_i s_i(0)$.

Comparing Eq. (14) with the deterministic time-evolution equations of sequential dynamics [6,7], beyond our expectation, the results are completely identical. The reasons for this equivalent result in both cases are found in the premises of the continuous Markov process and the omitting the higher-order terms in the Fokker-Planck-type approach to the Master equations Eq. (7). From the Lindeberg continuous condition for a Markov process [15], the difference between the probabilities of contiguous microscopic states goes to zero faster, as the time step Δt goes to zero. To the extent that there only exist one or several updating spins in one time step for the parallel dynamics, it can be treated approximately as the sequential dynamics. As a result, the macroscopic parallel dynamics is the same expression of macroscopic order variables as sequential dynamics in the case of being far away from saturation and finite of temperature.

Moreover, considering the limit of zero temperature (or $\beta \rightarrow \infty$), the transition rate Eq. (4) become $w(\mathbf{s}' \rightarrow \mathbf{s}) = \prod_{i=1}^N \frac{1}{2} \{1 + s_i \text{sgn}[h_i(\mathbf{s}')] \}$. It is apparent that this equation is not suitable to figure the parallel process. But for the case of sequential process, the transition rate is easy to be described by $w(s_i \rightarrow -s_i) = \frac{1}{2} \{1 - s_i \text{sgn}[h_i(\mathbf{s})] \}$. Consequently, the above solution for parallel dynamics, while being far away from saturation, cannot be extended to the limit of $T \rightarrow 0$ or $\beta \rightarrow \infty$.

Another interrelated important topic is finite-size effects in networks [9,16]. In other words, the problem is how large is a small system. For example, in numerical simulation, it is necessary to consider a system size of up to $N \approx 3 \times 10^4$ for calculating certain properties of the Hopfield model [17]. This puzzle of how to account for the finite size of the networks and extract useful information about the asymptotic, $N \rightarrow \infty$, limit from networks of only a few hundred to a few thousand neurons was investigated by Forrest [18], Kanter and Sompolinsky [19]. Furthermore, in the asynchronous evolutionary case, Castellanos *et al.* presented a more detail analytical exploration [9], that the effects are governed by a time-dependent Ornstein-Uhlenbeck process for cases far away from saturation. In the following context, such effects are discussed for systems in a synchronous evolutionary circumstance.

Applying the analogical technical route in Eq. (12), it is easy to obtain the quadric term in the right-hand side of Eq. (9) as

$$\begin{aligned} & \sum_{\mathbf{s}'} [\mathbf{\Omega}(\mathbf{s}') - \mathbf{\Omega}(\mathbf{s})]^2 w(\mathbf{s} \rightarrow \mathbf{s}') \\ &= \frac{4}{N^2} \left\{ \left[\sum_{i=1}^N \frac{\xi_i s_i - \xi_i \tanh[\beta h_i(\mathbf{\Omega})]}{2} \right]^2 \right. \\ & \quad \left. + \sum_{i=1}^N \frac{\xi_i^2 \{1 - \tanh^2[\beta h_i(\mathbf{\Omega})]\}}{4} \right\} \\ &= \left[\mathbf{\Omega} - \frac{1}{N} \sum_{i=1}^N \xi_i \tanh[\beta h_i(\mathbf{\Omega})] \right]^2 \\ & \quad + \frac{1}{N^2} \sum_{i=1}^N \xi_i^2 \{1 - \tanh^2[\beta h_i(\mathbf{\Omega})]\}. \end{aligned} \quad (15)$$

It is now straightforward to obtain the following Fokker-Planck-type equation by substituting Eq. (15) into Eq. (9):

$$\begin{aligned} \frac{dP_t(\mathbf{\Omega})}{dt} &= \sum_k \frac{\partial}{\partial \Omega_k} \left\{ P_t(\mathbf{\Omega}) \left[\Omega_k(\mathbf{s}) \right. \right. \\ & \quad \left. \left. - \frac{1}{N} \sum_{i=1}^N \xi_i \tanh(\beta \xi_i \cdot \mathbf{A}\Omega) \right] \right\} \\ & \quad + \frac{1}{2} \sum_{k,l} \frac{\partial^2}{\partial \Omega_k \partial \Omega_l} \left\{ P_t(\mathbf{\Omega}) \left[\right. \right. \\ & \quad \left. \left. - \left(\mathbf{\Omega} - \frac{1}{N} \sum_{i=1}^N \xi_i \tanh(\beta \xi_i \cdot \mathbf{A}\Omega) \right)^2 \right. \right. \\ & \quad \left. \left. - \frac{1}{N^2} \sum_{i=1}^N \xi_i^2 [1 - \tanh^2(\beta \xi_i \cdot \mathbf{A}\Omega)] \right] \right\}. \end{aligned} \quad (16)$$

In there, the diffusion parameter is different from the one in sequential dynamics [9]. This means that the parallel dynamics investigated here differ from the sequential dynamics of fluctuations or the internal noise distribution.

Following the approach of Castellanos *et al.* [9,16], the new rescaled variable \mathbf{q} and its probability distribution function $P(\mathbf{q})$ are defined as

$$\begin{aligned} \mathbf{q}(t) &= \sqrt{N} [\mathbf{\Omega}(t) - \mathbf{\Omega}^*(t)], \\ P_t(\mathbf{q}) &= \int d\mathbf{\Omega} P_t(\mathbf{\Omega}) \delta[\mathbf{q} - \sqrt{N}(\mathbf{\Omega} - \mathbf{\Omega}^*)], \end{aligned} \quad (17)$$

where $\mathbf{\Omega}^*$ is the deterministic solution of the Liouville equation (14), and Eq. (17) means that the order vector $\mathbf{\Omega}$ can be resolved into the sum of a deterministic terms $\mathbf{\Omega}^*$ and fluctuating terms with the latter terms vanishing in the limit of $N \rightarrow \infty$. Moreover, from the central limit theorem, the fluctuating term can be scaled as $N^{-1/2}$ [12,20].

In the limit of $N \rightarrow \infty$ and with the help of Eq. (16), the following Fokker-Planck-type equation of rescaled variables is deduced:

$$\begin{aligned} \frac{dP_i(\mathbf{q})}{dt} = & - \sum_k \frac{\partial}{\partial q_k} \{P_i(\mathbf{q})F_k(\mathbf{q},t)\} \\ & + \frac{1}{2} \sum_{k,l} \frac{\partial^2}{\partial q_k \partial q_l} \{P_i(\mathbf{q})D_{kl}(\mathbf{q},t)\}. \end{aligned} \quad (18)$$

Here the drift factor $F_k(\mathbf{q},t)$ is given by

$$\begin{aligned} \mathbf{F}(\mathbf{q},t) = & \beta \langle \xi(\xi \cdot \mathbf{A}\mathbf{q}) [1 - \tanh^2(\beta \xi \cdot \mathbf{A}\mathbf{\Omega}^*)] \rangle_{\xi} - \mathbf{q} \\ & + \lim_{N \rightarrow \infty} \sqrt{N} \left\{ \frac{1}{N} \sum_i \xi_i \tanh(\beta \xi_i \cdot \mathbf{A}\mathbf{\Omega}^*) - \mathbf{\Omega}^* \right\} \end{aligned} \quad (19)$$

and $\langle g(\xi) \rangle_{\xi} = \lim_{N \rightarrow \infty} (1/N) \sum_k g(\xi_k)$ with $\xi_k = (\xi_k^1, \dots, \xi_k^p)$. The last term on the right-hand side of Eq. (19) depicts finite-size corrections to the flow field of pattern overlaps. Similarly, the diffusion factor can be given by

$$\begin{aligned} \mathbf{D}(\mathbf{q},t) = & - \langle \xi^2 [1 - \tanh^2(\beta \xi \cdot \mathbf{A}\mathbf{\Omega}^*)] \rangle_{\xi} \\ & + \left\{ [\beta \langle \xi(\xi \cdot \mathbf{A}\mathbf{q}) [1 - \tanh^2(\beta \xi \cdot \mathbf{A}\mathbf{\Omega}^*)] \rangle_{\xi} - \mathbf{q}] \right. \\ & \left. + \lim_{N \rightarrow \infty} \sqrt{N} \left[\frac{1}{N} \sum_i \xi_i \tanh(\beta \xi_i \cdot \mathbf{A}\mathbf{\Omega}^*) - \mathbf{\Omega}^* \right] \right\}^2. \end{aligned} \quad (20)$$

Obviously, unlike sequential dynamics, the finite-size effects in parallel dynamics are governed by a homogeneous Markov process. More intensive and detailed work, including verification of the numerical simulation, is ongoing.

In short, the macroscopic dynamics of neural networks were found to yield similar results for two updating ways, synchronous and asynchronous, for cases far from saturation of stored capacity and finite temperature. But differences found between two cases in the fluctuations of pattern overlaps are given in Eq. (16). Moreover, far from saturation, the finite-size effects in parallel dynamics were described by a homogeneous Markov process, but not by the time-dependent Ornstein-Uhlenbeck process in sequential dynamics.

-
- [1] W.A. Little, *Math. Biosci.* **19**, 101 (1974).
[2] J.J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554 (1982).
[3] D.J. Amit, H. Gutfreund, and H. Sompolinsky, *Phys. Rev. A* **32**, 1007 (1985).
[4] D.J. Amit, H. Gutfreund, and H. Sompolinsky, *Ann. Phys. (N.Y.)* **173**, 30 (1987).
[5] J. C. Eccles, *The Understanding of the Brain* (McGraw-Hill, New York, 1977).
[6] A.C.C. Coolen and Th.W. Ruijgrok, *Phys. Rev. A* **38**, 4253 (1988); A.C.C. Coolen and D. Sherrington, *J. Phys. A* **25**, 5493 (1992); A.C.C. Coolen, *Europhys. Lett.* **16**, 73 (1991).
[7] A.C.C. Coolen and L.G.V.M. Lenders, *J. Phys. A* **25**, 2577 (1992).
[8] A.C.C. Coolen and D. Sherrington, *Phys. Rev. Lett.* **71**, 3886 (1993); A.C.C. Coolen and D. Sherrington, *Phys. Rev. E* **49**, 1921 (1994); S.N. Loughton and A.C.C. Coolen, *ibid.* **51**, 2581 (1995).
[9] A. Castellanos, A.C.C. Coolen, and L. Viana, *J. Phys. A* **31**, 6615 (1998).
[10] H. Haken, *Rev. Mod. Phys.* **47**, 67 (1975).
[11] D. C. Hebb, *The Organization of Behavior: A Neurophysiological Theory* (Wiley, New York, 1957).
[12] M. Shiino, *J. Stat. Phys.* **59**, 1051 (1990).
[13] P. Peretto and J.J. Niez, *IEEE Trans. Syst. Man Cybern.* **SMC-16**, 73 (1986); O. Bernier, *Europhys. Lett.* **16**, 531 (1991); A.C.C. Coolen and D. Sherrington, *J. Phys. A* **25**, 5493 (1992).
[14] G. Parisi, *J. Phys. A* **13**, 1101 (1980); G. Parisi, *Phys. Rev. Lett.* **50**, 1946 (1983).
[15] G. W. Gardiner, *Handbook of Stochastic Book for Physics, Chemistry, and the Natural Sciences* (Springer-Verlag, New York, 1983).
[16] N. Brunel and V. Hakim, *Neural Comput.* **11**, 1621 (1999).
[17] G.A. Kohring, *J. Stat. Phys.* **59**, 1077 (1990); G.A. Kohring, *J. Phys. A* **23**, 1137 (1990).
[18] B.M. Forrest, *J. Phys. A* **21**, 245 (1988).
[19] I. Kanter and H. Sompolinsky, *Phys. Rev. A* **35**, 380 (1987).
[20] E. Scacciatelli and B. Tirozzi, *J. Stat. Phys.* **67**, 981 (1992).