Three-state neural network: From mutual information to the Hamiltonian

David R. Carreta Dominguez*
ESCET, Universidad Rey Juan Carlos, C. Tulipán, Mostoles, 28933 Madrid, Spain

Elka Korutcheva†
Departmento Física Fundamental, Universidad Nacional de Educación a Distancia, c/Senda del Rey No. 9, 28080 Madrid, Spain

(Received 17 December 1999; revised manuscript received 15 March 2000)

The mutual information, I, of the three-state neural network can be obtained exactly for the mean-field architecture, as a function of three macroscopic parameters: the overlap, the neural activity and the *activity-overlap*, i.e., the overlap restricted to the active neurons. We perform an expansion of I on the overlap and the activity-overlap, around their values for neurons almost independent of the patterns. From this expansion we obtain an expression for a Hamiltonian which optimizes the retrieval properties of this system. This Hamiltonian has the form of a disordered Blume-Emery-Griffiths model. The dynamics corresponding to this Hamiltonian is found. As a special characteristic of such a network, we see that information can survive even if no overlap is present. Hence the basin of attraction of the patterns and the retrieval capacity is much larger than for the Hopfield network. The extreme diluted version is analyzed, the curves of information are plotted and the phase diagrams are built.

PACS number(s): 87.10.+e, 64.60.Cn, 07.05.Mh

I. INTRODUCTION

The collective properties of neural networks, such as the storage capacity and the overlap with the memorized patterns, have been a subject of intensive research in the last decade [1,2]. However, more precise measures of their performance as an associative memory, as the information capacity and the basins of attraction of their retrieval states, have received comparatively less attention [3–6]. For some models as the sparse-code networks [7–9], or the three-state networks [10–12], where the patterns are not uniformly distributed, an information-theoretical approach [13–15] seems crucial.

Calculations of the Shannon mutual information (I) for the sparse-code network were made [16–18]. For low storage of patterns, a few time steps are needed to retrieve them [3,17]. However, for large storage, only *im*perfect retrieval is possible. The closer to saturation, the larger the time steps required to dynamical retrieval. So, first-time retrieval is not enough and it is interesting to study the information capacity of recurrent networks. To improve I for this recurrent network, a scheme, based on a self-control threshold mechanism, was proposed [19]. This Self-Control Neural Network (SCNN) is an adaptive scheme induced by the dynamics itself instead of imposing any external constraint on the activity of the neurons. Such procedure successfully increases both I and the basins of attraction of the patterns. Similar mechanisms can improve I for three-state low-activity networks [20], with diluted and fully connected architectures.

Here we propose a new method, based on direct use of the I calculated in the mean-field approximation, to obtain a Hamiltonian which maximizes I within a large range of values for the activity of the network.

A three-state neural network is defined by the use of a set of $\mu=1,\ldots,p$ *ternary* patterns, $\{\xi_i^\mu \in \{0,\pm 1\},\ i=1,\ldots,N\}$, which are independent random variables given by the probability distribution

$$p(\xi_i^\mu)=a\,\delta(|\xi_i^\mu|^2-1)+(1-a)\,\delta(\xi_i^\mu), \tag{1}$$

where a is the *activity* of the patterns ($\xi_i^\mu=0$ are the inactive states). A low-activity three-state neural network corresponds to the case where the distribution is not uniform, i.e., $a<2/3$. In the limit $a=1$ the binary Hopfield model is reproduced.

The information enclosed in a simple unit ξ_i^μ is given by the entropy of its probability, $H_{\mu i}=-a\,\ln(a/2)-(1-a)\,\ln(1-a)$. One can define as sparse a code whose fraction of active neurons is very small and tends to zero in the thermodynamic limit [3].

Besides the fact that ternary patterns are a step toward an analog neural model, they have the advantage that they can be generated with a bias but keep their symmetric distribution (both ± 1 states are considered active). An important question related to the three-state model is the measurement of the retrieval quality in the cases where this is imperfect. Although for strictly homogeneous ternary patterns, the Hamming distance can be used, for the cases where $a\neq 2/3$ the errors in retrieving the active states have different relevance (they contain more information than the inactive). To solve this problem, the conditional probability of neuron states given the pattern states [20], was used to obtain the mutual information I of the attractor neural network (ANN). This I is a function of three parameters: the overlap m, the neural-activity q, and the activity-overlap n.

We then expand the I around the values of the parameters when the neurons are independent on the patterns. This expansion gives us an expression that can be interpreted as a

*Electronic address: dcarreta@escet.urjc.es
†Permanent address: G. Nadjakov Inst. Solid State Physics, Bulgarian Academy of Sciences, 1784 Sofia, Bulgaria. Electronic address: elka@fisfun.uned.es

Hamiltonian, a function only of the neuron states and the synaptic couplings. This Hamiltonian is similar to the Blume-Emery-Griffiths [21–28] spin-1 model (BEG), but with random interactions. The BEG model, originally proposed to study He_3–He_4 mixtures, was later used to describe several systems, like memory alloys, fluid mixtures, microemulsions, etc., and displays a variety of new thermodynamic phases.

Some disordered BEG models have been recently studied [29–31], where either the exchange-interactions or the crystal-field are random variables. However, from our knowledge, no random biquadratic-interactions model has been treated up to this date.

We describe our model and the *I* measure used to evaluate the performance of the ANN in Sec. II. In Sec. III we derive the BEG Hamiltonian from *I*. After solving the thermodynamics for this model in Sec. IV, we present some results for the dynamics and the phase diagrams in Sec. V, comparing the results with previous works. We conclude in the last section with some comments about possible improvements of the network.

II. MODEL

As well as the pattern states, the neuron states are three-state variables, defined as

$$\sigma_i \in \{0,\pm 1\},\ i=1,\ldots,N. \tag{2}$$

They are coupled to the other neurons through synaptic interaction, the specific form of which will be obtained later, by construction. We will see they are of the Hebbian type, that is, the learning is local (the synapses depend only on the two neurons interacting). Moreover, the updating rule will be also obtained by construction, no supposition being done here except that the patterns have the same three-state symmetry as the neuron states.

The three-state patterns $\xi_i^\mu \in \{0,\pm 1\}, \mu=1,\ldots,p$, are independent identically distributed random variables (IIDRV) chosen according to the probability distribution in Eq. (1). There is no bias ($\langle \xi_i^\mu \rangle = 0$) nor correlation between patterns ($\langle \xi_i^\mu \xi_i^\nu \rangle = 0$), and $a=\langle |\xi_i^\mu|^2 \rangle$ is the activity of the patterns. An extensive load of patterns is achieved when $p=\alpha N$.

The mean-field networks have the property of being site-independent; that means, the correlations between different sites are negligable in the thermodynamic limit, $N\to\infty$. This implies that every macroscopic quantity satisfies the conditions of the law of large numbers (LLN), so they can be defined as an average on the probability distribution of a state in a single site. Alternatively, we may say that the probability distribution factories, $p(\{\sigma\},\{\xi\})=\Pi_i p(\sigma_i,\xi_i)$.

The task of retrieval is successful if the state of the neuron $\{\sigma_{it}\}$ matches (at least approximately) the pattern $\{\xi_i^\mu\}$. The measure of the quality of the retrieval we will use here is the mutual information, which is a function of the conditional distribution of the neurons given the patterns, $p(\sigma|\xi)$. It has been noted [19,20] that the order parameters which are needed to calculate $p(\sigma|\xi)$ are the thermodynamic limits of the standard overlap of the μth pattern with the neuron state,

$$m_N^\mu \equiv \frac{1}{aN}\sum_i \xi_i^\mu \sigma_i \to m = \left\langle \langle \sigma \rangle_{\sigma|\xi} \frac{\xi}{a} \right\rangle_\xi , \tag{3}$$

the neural activity,

$$q_{Nt} \equiv \frac{1}{N}\sum_i |\sigma_{it}|^2 \to q = \langle \langle \sigma^2 \rangle_{\sigma|\xi} \rangle_\xi , \tag{4}$$

and the so called *activity-overlap* [20],

$$n_{Nt}^\mu \equiv \frac{1}{aN}\sum_i^N |\sigma_{it}|^2 |\xi_i^\mu|^2 \to n = \left\langle \langle \sigma^2 \rangle_{\sigma|\xi} \frac{\xi^2}{a} \right\rangle_\xi . \tag{5}$$

In the expressions for the thermodynamic limits, lim $N\to\infty$ in the equations for, m,q,n, the index μ for the considered pattern were dropped out. The averages are over the conditional distribution, $p(\sigma|\xi)$, and over the pattern distribution, $p(\xi)$, in Eq. (1).

For the dynamics used in most work found in the literature [10,12,32], where the synapses used are of the Hopfield form $J_{ij}=\Sigma_\mu \xi_i^\mu \xi_j^\mu$, the parameter n_N^μ does not seem to play any role in the evolution of the network, independent of the architecture considered (diluted, layered or fully connected, for instance). In fact it doesn't appear in the expression for the Hamming distance. However, n_N^μ is necessary to define the mutual information of the network, as well as being necessary in computing the network's performance [33,34].

Knowing the conditional probability for each site and pattern, one can define the *Mutual Information I* [13,14], a theoretical information quantity used to measure the average amount of information that can be received by the user by observing the symbol (or the signal) at the output of a channel. We can regard the pattern as the input and the neuron states as the output of the channel, so the *I* is written as

$$I[\sigma;\xi] = S[\sigma] - \langle S[\sigma|\xi] \rangle_\xi ,$$

$$S[\sigma] \equiv -\sum_\sigma p(\sigma)\ln[p(\sigma)], \tag{6}$$

$$S[\sigma|\xi] \equiv -\sum_\sigma p(\sigma|\xi)\ln[p(\sigma|\xi)].$$

$S[\sigma]$ and $S[\sigma|\xi]$ are the entropy of the output and the conditional entropy of the output, respectively. The quantity $\langle S[\sigma|\xi] \rangle_\xi$ is also called the *equivocation term* of the $I[\sigma;\xi]$.

Using the conditional probability obtained in Ref. [20],

$$p(\sigma|\xi) = (s_\xi + m\xi\sigma)\,\delta(\sigma^2-1) + (1-s_\xi)\,\delta(\sigma),$$

$$s_\xi \equiv s + \frac{n-q}{1-a}\xi^2,\ s \equiv \frac{q-na}{1-a}, \tag{7}$$

the expressions for the entropies defined above are

$$S[\sigma] = -q\,\ln\frac{q}{2} - (1-q)\ln(1-q),$$

$$\langle S[\sigma|\xi] \rangle_\xi = aS_a + (1-a)S_{1-a},$$

$$S_a = -\frac{n+m}{2}\ln\frac{n+m}{2} - \frac{n-m}{2}\ln\frac{n-m}{2} - (1-n)\ln(1-n), \tag{8}$$

$$S_{1-a} = -s\,\ln\frac{s}{2} - (1-s)\ln(1-s).$$

Several threshold mechanisms have been proposed as a dynamical evolution of three-states, neural network [9]. Most of them are fixed-threshold, using a Hopfield Hamiltonian, together with a Hebbian-like synaptic interaction. Recently a Hopfield three-state network with a self-control (SC) mechanism was introduced, the threshold of which adapts itself according to the following dynamics [20]: $\theta_t = c(a)\Delta_t$. Here $c(a) = \sqrt{-2\,\ln(a)}$ is a function only of the pattern activity, while the variance of the cross-talk noise (due to the $p-1$ nonretrieved patterns) has the simple form $\Delta_t = \sqrt{\alpha q_{Nt}}$ for the diluted architecture [35]. That yields the best performance for low-activity patterns. In this paper we take the alternative approach of starting from the mutual information for the model, and in the results, both models are compared.

III. DERIVATION OF THE HAMILTONIAN

We search for a Hamiltonian which is symmetric in any permutations of the patterns ξ^μ, since they are not known during the retrieval process. This imposes that the retrieval of any pattern ξ^μ is weak, i.e., σ is almost independent of it. Then obviously the overlap $m^\mu \sim 0$. An expansion of I with $a = 1 = q$ around $m^\mu \sim 0$ yields the Hopfield Hamiltonian. If afterwards some particular overlap eventually becomes large, this should be a consequence of the network evolution.

However, for general a,q, this is not the only quantity which vanishes in this limit. The variable σ^2 is also almost independent of $(\xi^\mu)^2$, so that $n^\mu \sim q$. Hence, the parameter

$$l^\mu \equiv \frac{n^\mu - q}{1-a} = \langle \sigma^2 \eta^\mu \rangle, \eta^\mu \equiv \frac{(\xi^\mu)^2 - a}{a(1-a)}, \tag{9}$$

also vanishes when the states of the neurons and the patterns are independent.

We use this fact to look at the information close to the nonretrieval regime. An expansion of the expression for the I around $m^\mu = 0, l^\mu = 0$ gives

$$I^\mu \approx \frac{1}{2}\frac{a}{q}(m^\mu)^2 + \frac{1}{2}\frac{a(1-a)}{q(1-q)}(l^\mu)^2. \tag{10}$$

Since this expression gives the information for a single site i of a single pattern μ, $I(m^\mu, l^\mu) \equiv I^\mu$, it should be summed $I_{pN} = N\Sigma_\mu I^\mu$ to give the total information of the network. It is natural to associate this quantity with the opposite of the Hamiltonian, because the maximum of the information gives the minimal energy.

We suppose, as a further simplification of the model, that the neural activity is of the same order of the pattern activity, $q \sim a$. This is made by convenience, otherwise the statistical mechanics of the model will be difficult to compute. With this assumption, I from Eq. (10) depends in the same way on m^μ and l^μ. Substituting the expressions for these parameters, given by the definitions (3),(4) and (5) (before the thermodynamic limit), we obtain the following expression for the I:

$$\mathcal{H} = -I_{pN} \equiv \mathcal{H}_1 + \mathcal{H}_2, \tag{11}$$

where

$$\mathcal{H}_1 = -\frac{1}{2}\sum_{ij} J_{ij}\sigma_i\sigma_j \tag{12}$$

and

$$\mathcal{H}_2 = -\frac{1}{2}\sum_{i,j} K_{ij}\sigma_i^2\sigma_j^2 \tag{13}$$

are the quadratic and the biquadratic terms, respectively. The above expression for the Hamiltonian, obtained from the mutual information close to the nonretrieval regime, has the same form as of the BEG model [21]. We call our model the BEG Neural Network (BEGNN).

The interactions are randomly distributed, given by

$$J_{ij} = \frac{1}{a^2 N}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \tag{14}$$

and

$$K_{ij} = \frac{1}{N}\sum_{\mu=1}^{p} \eta_i^\mu \eta_j^\mu. \tag{15}$$

The first term of the Hamiltonian is the usual Hopfield model with the Hebbian rule given by Eq. (14). The second term, arising from the term depending on l^μ in Eq. (10), related to the activity-overlap, is also Hebbian-like, but is associated, as will be seen later, with the quadrupolar order of the system.

Note that the Hamiltonian formulation of the problem is only possible in the case of the fully connected neural network, where the interaction matrix is symmetric. In the next section we will present the dynamical formulation of the problem, which can be applied to the cases of asymmetric couplings [35], although no Hamiltonian exists.

As is well known, the phase diagram of the usual BEG model is very rich, showing different phases, depending on the sign and the strength of the biquadratic coupling constant. Without any disorder and for very negative biquadratic coupling constant, a quadrupolar phase, related to the quadrupolar moment $\langle \sigma^2 \rangle$ also appears, apart from the usual disordered and ferromagnetic phases [22–28]. However, our variables ξ_i^μ are quenched, so we have a disordered system. BEG models with disordered quadratic coupling have been recently studied [29–31], showing some new phases (spin-glass, quadrupolar spin-glass phases, etc.), but, from our knowledge, no disordered biquadratic BEG model has been studied up to this date.

IV. ASYMPTOTIC MACRODYNAMICS

A. Thermodynamics for $\alpha \to 0$

For the derivation of the asymptotic macrodynamics we will use a naive mean-field (MF) approach using the Hamiltonian Eqs. (12)–(13). Since the Hamiltonian is quadratic in the overlaps, we can linearize it, using Gaussian transformation, to obtain the partition function

$$Z=\mathrm{Tr}_{\{\sigma\}}e^{-\beta\mathcal{H}}$$
$$=\int\prod_{\mu}\left[D\Phi(\sqrt{\beta N}m^{\mu})D\Phi(\sqrt{\beta N}l^{\mu})\right]\prod_{i}\sum_{\sigma=\pm1,0}e^{\widetilde{\mathcal{H}_i}}, \tag{16}$$

where $D\Phi(z)\equiv dze^{-(z^2/2)}/\sqrt{2\pi}$, and $\beta=1/T$. The effective Hamiltonian is

$$\widetilde{\mathcal{H}_i}=h_i\sigma_i+\theta_i\sigma_i^2, \tag{17}$$

where the local fields are

$$h_i=\frac{1}{a}\sum_{\mu}^{p}\xi_i^{\mu}m^{\mu},$$
$$\theta_i=\sum_{\mu}^{p}\eta_i^{\mu}l^{\mu}. \tag{18}$$

After taking the trace over the spin variables, we apply a saddle-point integration and use the thermodynamic limit, to get the free energy in terms of the parameters m, l and q,

$$f=-\frac{T}{N}\ln Z=\frac{1}{2}(\mathbf{m}^2+\mathbf{l}^2)-T\langle\ln\tilde{Z}\rangle_{\xi}, \tag{19}$$

where the effective partition function is

$$\tilde{Z}=1+2e^{\beta\theta}\cosh(\beta h). \tag{20}$$

The fields h,θ are defined in Eq. (18), but the indices i can be dropped out. The saddle-point equations $\partial f/\partial m^{\mu}=0$ and $\partial f/\partial l^{\mu}=0$ lead to the following expressions for the stationary states:

$$m^{\mu}=\left\langle\frac{1}{a}\xi^{\mu}\overline{\sigma}\right\rangle_{\xi},$$
$$l^{\mu}=\langle\eta^{\mu}\overline{\sigma^2}\rangle_{\xi}, \tag{21}$$

where the angular brackets mean the average over the patterns, and the thermal averages of the states are

$$\overline{\sigma}\equiv F_{\beta}(h,\theta)=\frac{1}{\tilde{Z}}2e^{\beta\theta}\sinh(\beta h),$$
$$\overline{\sigma^2}\equiv G_{\beta}(h,\theta)=\frac{1}{\tilde{Z}}2e^{\beta\theta}\cosh(\beta h). \tag{22}$$

We remark that these Eqs. (22) are formally the same as for the $Q=3$-Ising model [12,33]; however, the expression for θ is both site-dependent and a function of the order parameters.

For zero-temperature, the behavior of the averages is

$$F_{\infty}(h,\theta)=\mathrm{sgn}(h)\Theta(|h|+\theta),$$
$$G_{\infty}(h,\theta)=\Theta(|h|+\theta), \tag{23}$$

where $\Theta(\ldots)$ is the step function. This naive thermodynamic approach is only valid if the number of patterns is not extensive. Whenever the ratio of patterns $\alpha=p/N$ is finite, the noncondensed overlaps yield another global order parameter.

This result, obtained from the naive MF theory, can be easily understood if we write the Hamiltonian in Eqs. (12),(13) in the form

$$\mathcal{H}=\frac{1}{2}\sum_i\widetilde{\mathcal{H}_i}=-\frac{1}{2}\sum_i(h_i\sigma_i+\theta_i\sigma_i^2),$$
$$h_i\equiv\sum_j J_{ij}\sigma_j, \tag{24}$$
$$\theta_i\equiv\sum_j K_{ij}\sigma_j^2.$$

From the first term of the Hamiltonian, its minimum arises once the neuron states are given by

$$\sigma_i=\mathrm{sgn}(h_i)G(h,\theta), \tag{25}$$

where $G(h,\theta)\in\{0,1\}$. This form of σ_i doesn't affect the second term of the Hamiltonian. On the other hand, replacing Eq. (25) in Eq. (24) we arrive at

$$\mathcal{H}=-\frac{1}{2}\sum_i(|h_i|+\theta_i)G(h_i,\theta_i). \tag{26}$$

Now it is clear that if a minimum of this Hamiltonian holds, the expression for $G(h,\theta)$ should be exactly the same as $G_{\infty}(h,\theta)$, given in Eq. (23).

B. Diluted dynamics

Alternatively to the thermodynamic approach, in the noise case, we can also start from the stochastic parallel dynamics [12,33],

$$p(\sigma_{i,t+1}|\{\sigma_t\})=\exp[\beta\widetilde{\mathcal{H}_i^t}]/\tilde{Z}, \tag{27}$$

where $\widetilde{\mathcal{H}_i^t}$ is given by Eq. (17) (in the time step t), and $\tilde{Z}$ by Eq. (20). Differently from the dynamics for the $(Q=3)$-Ising model [12,33], here the field $\theta=\theta(\{\sigma_j^2\})$ in the effective Hamiltonian is a function of the states in the previous time steps. The resulting noise-averaged states coincide with Eqs. (22) in the stationary regime. Again, here the field θ_i is site-dependent.

At $T=0$, in comparison with the $Q-Ising$ model [12,36], and with Eq. (25), the following deterministic parallel dynamics, which leads to the minimization of the effective Hamiltonian, is suggested,

$$\sigma_{i,t+1}=\mathrm{sgn}(h_i^t)\Theta(|h_i^t|+\theta_i^t), \tag{28}$$

where the local fields h_i^t,θ_i^t (associated with the variables $\sigma_{j,t},\sigma_{j,t}^2$, respectively) are given in the time step t. Such dynamics has the same form as the zero-temperature function F_{∞} in Eq. (23).

Because we are mainly interested in the retrieval properties of our network, we take an initial configuration whose retrieval overlaps are only macroscopic of order $O(1)$ for a

given pattern, let's say the first one. We singled out the term $\mu=1$ in the local fields of Eq. (18) in order to study the retrieval of the first pattern.

Supposing an initial configuration $\{\sigma_{i,t=0}\}$ as a collection of IIDRV with zero-mean and variance $q_{t=0}$, the fields $h_{t=0}$ and $\theta_{t=0}$ in the zeroth time step are given by

$$h_{t=0}=\frac{1}{a}\xi m_{t=0}+\omega_{t=0}; \quad \omega_{t=0}\equiv\sum_{\nu\geqslant 2}^{p}\frac{1}{a}\xi^{\nu}m^{\nu}_{t=0},$$
$$\theta_{t=0}=\eta l_{t=0}+\Omega_{t=0}; \quad \Omega_{t=0}\equiv\sum_{\nu\geqslant 2}^{p}\eta^{\nu}l^{\nu}_{t=0}, \tag{29}$$

where the indices $\mu=1$ were dropped, and the rest of the patterns one regarded as some additive noise. According to the Central Limit Theorem (CLT), they are independent Gaussian distributed [12,32], with zero mean and variance

$$\mathrm{Var}[\omega_{t=0}]=\frac{1}{a^2}\alpha q_{t=0}\equiv\Delta^2,$$
$$\mathrm{Var}[\Omega_{t=0}]=\frac{\Delta^2}{(1-a)^2}. \tag{30}$$

Although the dynamics for the parameters m_t, n_t and q_t in the first time step is a function of the initial step, the expression for the noises in the next steps evolves with time in a more complicated way than in Eqs. (30). In the extremely diluted synaptic case [35], however, the first time step describes the dynamics for every time step t. This allows us to calculate exactly its dynamics. Observe that the extreme diluted network has no possible Hamiltonian formulation, nevertheless, even if no feedback is present there, it can be supposed as an approximation of the fully connected network showing qualitatively similar behavior. From now on we will adopt this limiting case. The effective Hamiltonian now is to be considered as an energy potential [32,36].

Thus, in the asymptotic limit $N\to\infty$, the expression for the overlap $m_t=\lim_{N\to\infty} m^1_{Nt}$ becomes, after averaging over the pattern ξ,

$$m_{t+1}=\left\langle\frac{\xi\overline{\sigma_t}}{a}\right\rangle_{\xi}=\int D\Phi(y)\int D\Phi(z)F_{\beta}\left(\frac{m_t}{a}+y\Delta_t;\frac{l_t}{a}+z\frac{\Delta_t}{1-a}\right), \tag{31}$$

where the averages over ω,Ω on the brackets should be done with the Gaussian distributions, Eq. (30).

The neural activity is the thermodynamic limit of Eq. (4), which reads

$$q_t=\langle\overline{\sigma_t^2}\rangle_{\xi}=an_t+(1-a)s_t,$$
$$s_{t+1}\equiv\int D\Phi(y)\int D\Phi(z)G_{\beta}\left(y\Delta_t;-\frac{l_t}{1-a}+z\frac{\Delta_t}{1-a}\right). \tag{32}$$

Here s is the variable defined in Eq. (7) and the activity-overlap is given by

$$n_{t+1}=\left\langle\frac{\xi^2\overline{\sigma_t^2}}{a}\right\rangle_{\xi}=\int D\Phi(y)\int D\Phi(z)G_{\beta}\left(\frac{m_t}{a}+y\Delta_t;\frac{l_t}{a}+z\frac{\Delta_t}{1-a}\right). \tag{33}$$

The equation for l_t is obtained using the definition in Eq. (9), $l_t=(n_t-q_t)/(1-a)$. It is worth noting that the definitions of the parameters m,q,n in Eqs. (3) are the same as those in Eqs. (31)–(33), since the average over the conditional probability $p(\sigma|\xi)$ is equivalent to the average over the noise due to the $p-1$ remaining patterns $\{\boldsymbol{\xi}|\xi\}$. Equations (31)–(33) describe the macro-dynamics of the diluted BEGNN by adapting selfconsistently the threshold during the time-evolution of the system. With these equations we can calculate the mutual information from Eqs. (6)–(8).

V. PHASE DIAGRAM

In this section we present some explicit results for the BEGNN model. We first calculated the stable fixed points of Eqs. (31)–(33) for the asymptotic $N\to\infty$ network, and obtained the curves for the order parameters m,q,n and the *information* $i=I\alpha$ as a function of the load parameter α for two values of the activity a (Fig. 1). For small load ($\alpha<0.2$), the overlap remains close to $m\sim 1$ and the neural activity is $q\sim a$. When more patterns are stored in the network, i increases almost linearly, up to an optimal value, $i_{\mathrm{opt}}(\alpha_{\mathrm{opt}})$, after which i decreases to zero in $\alpha_{\max}$. The comparison is done with the self-control neural network ($SCNN$) model [19,20]. It is seen that for small activities ($a=0.3$), the BEGNN model gives worse results than with the SCNN model, with a smaller value for i, while for $a=0.6$ (close to the uniform distribution of patterns, $a=2/3$), the BEGNN performs better, an optimal value of the information $i\sim 0.15$, although it is attained for a smaller value of load, $\alpha\sim 0.2$. The reason for this behavior is that the third order parameter (related to the activity-overlap) is $l\sim 1$ for the BEGNN (SCNN) and $l\ll 1$ for the SCNN (BEGNN) with $a=0.6(a=0.3)$. Also the neural activity approaches $q\sim 2/3$ (so it goes away from $q=a$) sooner (with less patterns) for the BEGNN than for the SCNN. In our opinion this is due to the approximation used in the Hamiltonian derivation, $q=a$.

The behavior of the order parameters and the i with load, for the zero-temperature case, is presented for three different values of the activities (Fig. 2). The initial conditions are used where $m_0=10^{-6}, l_0=1, q_0=a$, such that there is almost no initial overlap. In this case there is always a sharp fall on the information for α not so much larger than α_{opt}. We see different behaviors depending on the activities.

The corresponding dynamical phase diagram is drawn in Fig. 3. Four possible phases are present: the retrieval $R(m\neq 0, l\neq 0, q\sim a)$ and $M(m\neq 0, l<0.5m, q\sim a)$ phases, the quadrupolar phase $Q(m=0, l\neq 0, q\sim a)$ and the zero phase $Z(m=0, l=0, q\sim a)$. The factor in the definition of the M phase is somewhat arbitrary, indicating the appearance of two different behaviors. The last phase Z, so called because

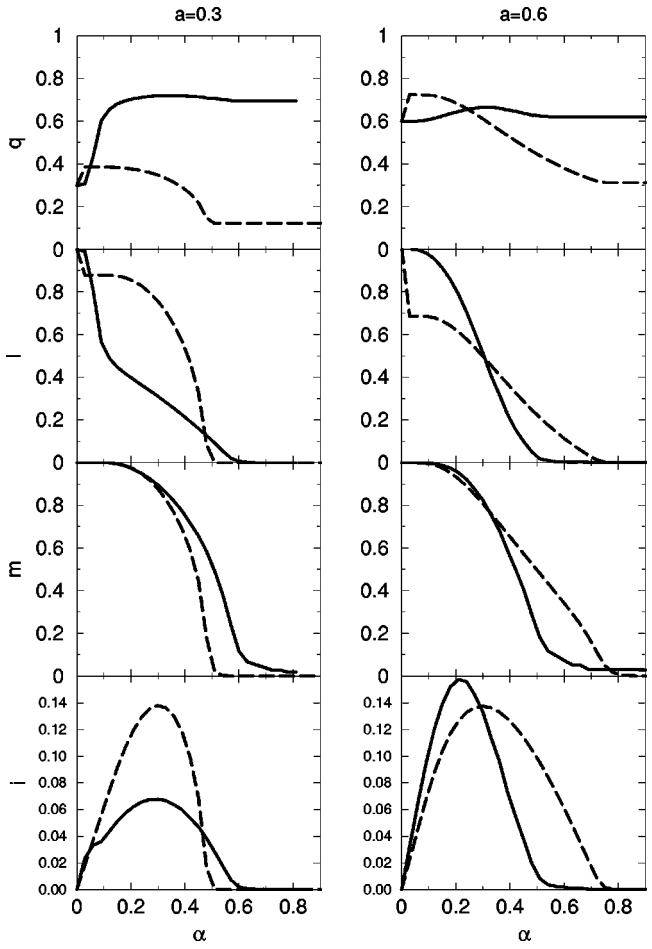

FIG. 1. The information $i=I\alpha$ and the order parameters m,l,q against α with activities $a=0.3$ (left) and $a=0.6$ (right). The temperature $T=0$ and the initial conditions are $m_0=1,l_0=1,q_0=a$. The continuous line is for the BEGNN while the dashed line is for the SCNN.

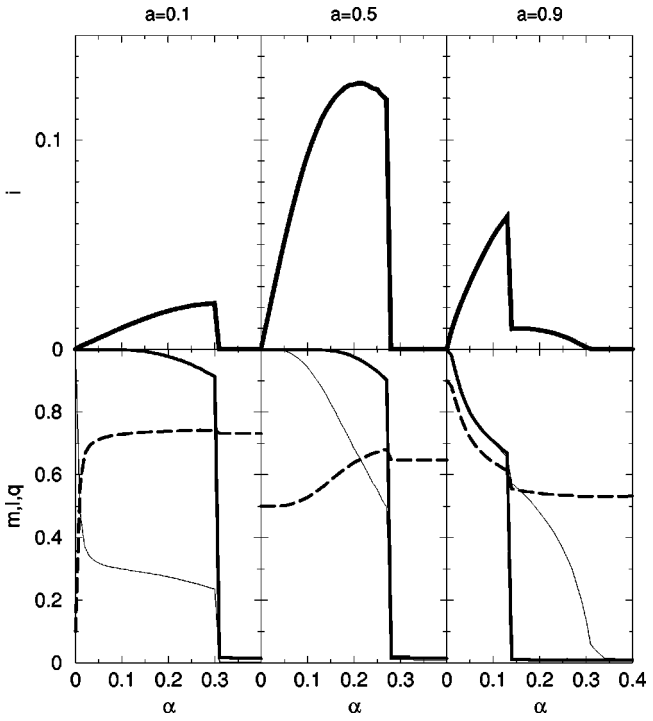

FIG. 2. The information i and the order parameters m (solid line), l (thin line) and q (dashed line) against α with activities $a=0.1$ (left), $a=0.5$ (center) and $a=0.9$ (right). The temperature $T=0$ and the initial conditions are $m_0=10^{-6},l_0=1,q_0=a$.

there is no information transmitted, is an analog to the self-sustained (S) activity phase of the $(Q=3)$-Ising ANN [12,33], since the parameter related to the spin-glass order is $q\neq 0$. We have not found any paramagnetic (P) phase, with all $(m=0,l=0,q=0)$ for the BEGNN. Note that the quadrupolar phase is a quite new phase, compared to the other NN models and is a special one for the BEGNN model. This phase is also present in the original BEG model [21], as well as in all its generalizations including disorder [29–31]. In the language of the neural network, the quadrupolar phase means that the active neurons (± 1) coincide with the active patterns, but the sign doesn't. It is seen in Fig. 2, for $a=0.9$, where the overlap goes to $m=0$ at $\alpha\sim 0.13$ (much before l, which goes to zero at $\alpha\sim 0.3$); this phase corresponds to nonzero information, although there is no retrieval overlap. The phase R appears for $a=0.5$, where both m and l are large and so is i. On the other hand, the phase M is observed for $a=0.1$, where the parameter l is much smaller than m. The phase transitions from R or M to Z are usually sharp, i.e., first-order, for $T=0$.

The behavior of the order parameters and the information with the temperature T for fixed activity $a=0.5$ is shown on Fig. 4. We observe an increase of i with the temperature, showing an optimal value for $T\sim 0.2$. Such an improvement

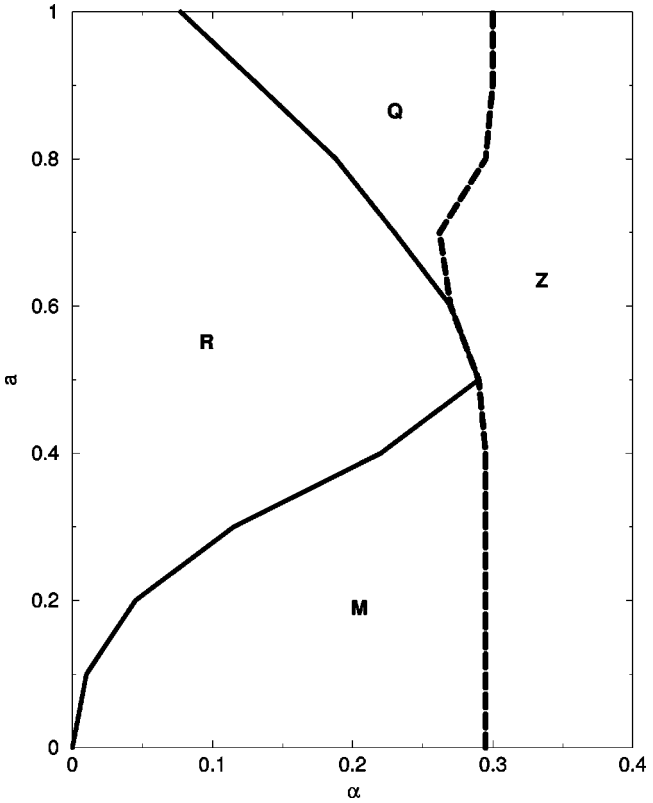

FIG. 3. The dynamical phase diagram $\alpha\times a$, for $T=0$ with initial conditions $m_0=10^{-6},l_0=1,q_0=a$. The different phases are explained in the text.

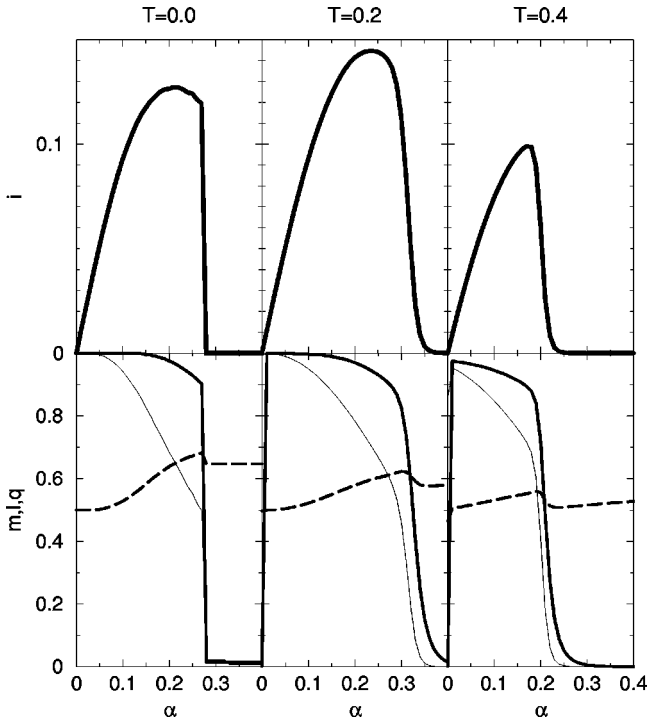

FIG. 4. The information i and the order parameters m (solid line), l (thin line) and q (dashed line) against α with temperatures $T=0.0$ (left), $T=0.2$ (center) and $T=0.4$ (right). The activity $a=0.5$ and the initial conditions are $m_0=10^{-6}, l_0=1, q_0=a$.

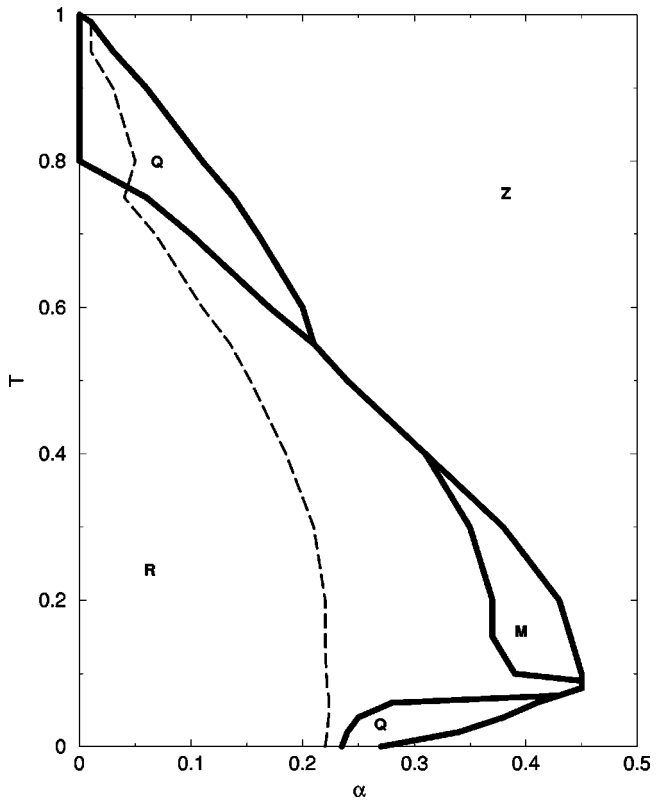

FIG. 5. The dynamical phase diagram $\alpha\times T$, for $a=0.7$ with initial conditions $m_0=10^{-6}, l_0=1, q_0=a$. The dashed line corresponds to the optimal information.

of a feeble signal with noise, similar to the stochastic resonance phenomena, appears also in other physical systems [37]. A further increase in temperature leads to decreasing the information of the model. We note that this behavior doesn't hold for $a\geqslant 2/3$, nevertheless there is still an increase of the storage capacity $\alpha_{\max}$. The last result is in agreement with other investigations of dynamical activity of real and model neurons, where the observed stochastic resonance disappears by increasing the amplitude of the external stimulus [38].

A cut of the phase diagram in the plane $T\times\alpha$ for a fixed value of the activity $a=0.7$ is shown in Fig. 5. The dashed line, which corresponds to the optimal case, $i_{\rm opt}(\alpha)$, is within either the phase R or Q. It is also interesting to observe that there are two separate Q-phase islands, for either small temperature T and large load α or large temperature T and small load α. The phase transitions become smoother with the temperature.

Finally, in Fig. 6 we present the evolution of the information and of the order parameters with the time t, for a given temperature $T=0.2$ and activity $a=0.7$, for two values of the load parameter α. As can be seen from this figure, for $\alpha=0.4$, which is close to the transition R-M, the change to the behavior of the order parameters needs more time steps than for $\alpha=0.2$. This is not strange due to the critical slowing down near the transition. However, an interesting new fact appears here: the parameters l_t and q_t have a fast melt down to a much smaller value, after which the network stays a long while with an almost zero overlap, and finally the BEGNN is able to retrieve the pattern quite well. For instance, for $\alpha=0.2$, l falls to $l\sim 0.6$ and m stays near $m\sim 0$ during the first $t\sim 20$ time steps, then they jump up to $l\sim 0.8, m\sim 0.9$, which means the memory pattern was (partially) attained. This result, caused by the instability of the Z-phase in this region, makes the BEGNN capacity much larger than that of the usual Hopfield model, in all its versions so far as we know.

The behavior of the continuous phase transitions can be analytically studied within the mean-field approximation by expanding Eqs. (31)–(33) for small values of the order parameters. A standard calculation, for example, for the transition line $QZ(m=0, l\ll 1)$ leads to the following expression:

$$l^{QZ}=\beta T_c^{QZ} l + \frac{(1-2a)\beta^2}{a^2\hat{a}^2} l^2 \cdot \left\langle e^{\beta\Omega}\cosh\beta\omega \frac{1-2e^{\beta\Omega}\cosh\beta\omega}{(1+2e^{\beta\Omega}\cosh\beta\omega)^3}\right\rangle_{\Omega,\omega}, \tag{34}$$

where the transition temperature between the phases Q and Z is

$$T_c^{QZ}=\frac{1}{a\hat{a}}\left\langle \frac{2e^{\beta\Omega}\cosh\beta\omega}{(1+2e^{\beta\Omega}\cosh\beta\omega)^2}\right\rangle_{\Omega,\omega}, \tag{35}$$

with

$$q^{QZ}=\left\langle \frac{2e^{\beta\Omega}\cosh\beta\omega}{1+2e^{\beta\Omega}\cosh\beta\omega}\right\rangle_{\Omega,\omega} + O(l^2). \tag{36}$$

Expanding the above expressions for a small value of the load rate and large temperatures, $\beta\sqrt{\alpha}\ll 1$, and calculating the averages over the noise up to the leading terms, one obtains the following equation for the transition line:

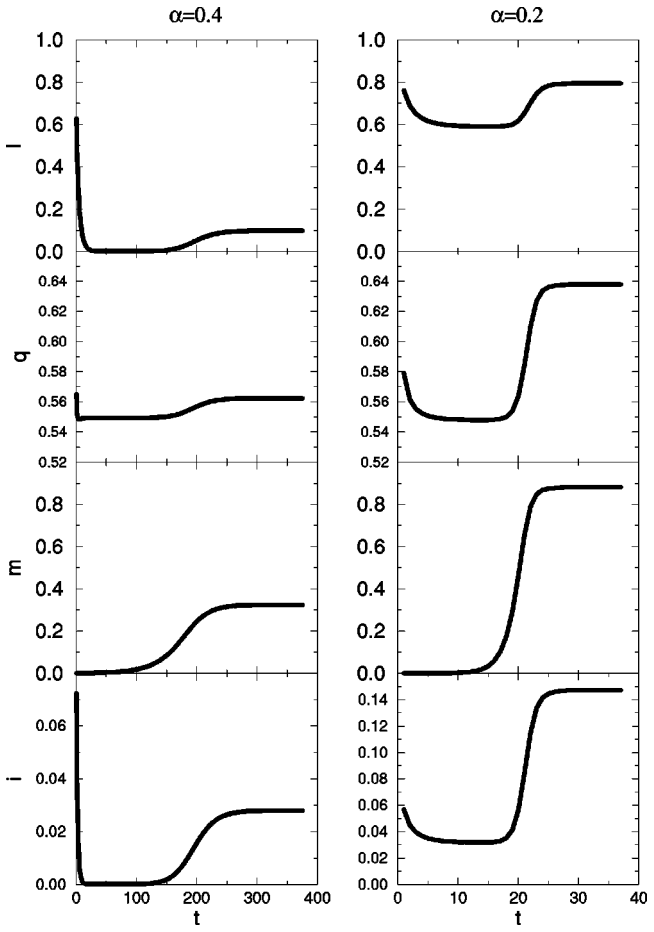

FIG. 6. The information i and the order parameters m,l,q, against the time t for temperature $T=0.2$ and activity $a=0.7$, with $\alpha=0.4$ (left) and $\alpha=0.2$ (right). The initial conditions are $m_0=10^{-6}, l_0=1, q_0=a$.

$$T_c=\frac{2}{9a\hat{a}}-\frac{1}{2}\frac{(1+\hat{a}^2)}{a\hat{a}}\alpha. \tag{37}$$

The last expression for T_c is in qualitative agreement with the previous results shown in Fig. 5.

Regarding the equation for the order parameter l, one can verify that in leading order,

$$l^{QZ}=\beta T_c^{QZ} l-\frac{1}{27}\frac{(1-2a)}{a^2\hat{a}^2}\beta^2 l^2+O(l^3,\alpha l^2). \tag{38}$$

By use of Eq. (37), it is seen that the quadratic term of the above expansion changes sign when the activity $a=0.5$, thus defining a tricritical line between the transition of second order ($a>0.5$) and of first order ($a<0.5$). Note that similar tricritical behavior has been described also in the other versions of the BEG model [22–31]. Similar analysis can also be performed for the other continuous transition between the different phases.

VI. CONCLUSIONS

In this paper we proposed a BEG-like Hamiltonian for a ternary neural network, the couplings of which arise from an expansion of its mean-field mutual information, I [20], resulting in a system evolving with a self-consistently adapting threshold. The stationary and dynamical equations for this model were obtained as functions of three order parameters, the overlap m, the neural activity q, and the activity-overlap n. Their solutions were explicitly calculated as functions of the variables: the pattern activity a, the load α and the temperature T. Only the extreme diluted version was studied here, and we hope this work motivates a detailed investigation of the properties of the fully connected BEGNN.

The comparison with the SCNN model [19,20], which is the best model known of the Hopfield-like NN, particularly for small pattern activity, allows a characterization of the BEGNN model. When the activity is near $a=2/3$, corresponding to the case of uniform ternary patterns, the BEGNN improves the information, compared with the SCNN model, while for small activities, it performs the worst, giving a smaller value for the information. We argue that this is due to the approximation $q\sim a$, used in the derivation of the Hamiltonian. We expect this paper stimulates analytical or simulation works which do not use such an approximation.

Improvement of the information content by increasing the noise, an effect similar to the stochastic resonance, is also observed for activities $a<2/3$, which is in agreement with results for real neurons.

There are four possible phases for the BEGNN, which were displayed in phase diagrams $a\times\alpha$ and $T\times\alpha$. In particular, a quadrupolar phase, Q, with $m=0, l\sim 1$, holds whenever the activity is large enough. This phase, known in the BEG literature, but new in an ANN context, carries out some nonzero information about the patterns even without any overlap m.

The phase transitions between the different phases are also investigated, showing sharp or continuous behavior, depending on the parameters. As the main result we obtained that, while the phase Z is not stable in a large range of the variables, the basin of attraction of the retrieval phase is increased with respect to the usual ternary neural network models. States with initial conditions, which have very small overlap, flow to final states with large overlap.

We believe that the BEGNN has a quite large range of applications for real systems. We also think that this way to obtain a Hamiltonian starting from a mean-field calculation of I, which yields an almost optimal retrieval dynamics, can be generalized to other spin systems, as the Q-Ising with $Q>3$ or the Potts models, for instance. Such a method, based on the maximization of the entropy, can be a universal approach to information systems.

Then, we expect that the same improvement should happen for analog neurons and for networks of binary *synapses*. It would be also interesting to investigate the case of local field for multi-neuron synapses, which comes up from higher order terms in the expansion of the mutual information, such that a better use of a network with fixed size is expected.

ACKNOWLEDGMENTS

We thank the workshop on ''Statistical Mechanics of Neural Networks,'' Max-Planck Institute, Dresden, for useful discussions. E.K. is financially supported by Spanish DGES Grant No. PB97-0076 and partly by Contract No. F-608 with the Bulgarian Scientific Foundation.

[1] P. Peretto, *An Introduction to the Modeling of Neural Networks* (Cambridge University Press, Cambridge, England, 1992).
[2] J. Hertz, A. Krogh, and R. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, Reading, MA, 1991).
[3] S. Amari, Neural Networks **2**, 451 (1989).
[4] C. Meunier, H. Yanai, and S. Amari, Network **2**, 469 (1991).
[5] C. J. Perez-Vicente, Europhys. Lett. **10**, 621 (1989).
[6] C. J. Perez-Vicente and D. J. Amit, J. Phys. A **22**, 559 (1989).
[7] D. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **35**, 2293 (1987).
[8] M. V. Tsodyks, Europhys. Lett. **7**, 203 (1988).
[9] M. Okada, Neural Networks **9**, 1429 (1996).
[10] J. S. Yedidia, J. Phys. A **22**, 2265 (1989).
[11] C. Meunier, D. Hansel, and A. Verga, J. Stat. Phys. **55**, 859 (1989).
[12] D. Bollé, G. M. Shim, B. Vinck, and V. A. Zagrebnov, J. Stat. Phys. **74**, 565 (1994).
[13] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[14] R. E. Blahut, *Principles and Practice of Information Theory*, (Addison-Wesley, Reading, MA, 1990) J. van der Lubbe, *Information Theory* (Cambridge University Press, Cambridge, England, 1997).
[15] M. Schluter, O. Kerschhaggl, and F. Wagner, Phys. Rev. E **60**, 2141 (1999).
[16] G. Palm, Biol. Cybern. **36**, 646 (1980).
[17] J. Nadal and G. Toulouse, Network Comput. Neural Syst. **1**, 61 (1990).
[18] F. Schwenker, F. T. Sommer, and G. Palm, Neural Networks **9**, 445 (1989).
[19] D. Dominguez and D. Bollé, Phys. Rev. Lett. **80**, 2961 (1998).
[20] D. Bollé, D. Dominguez and S. Amari, (unpublished); D. Bollé and D. Dominguez, e-print cond-mat/9912101; D. Bollé and G. Massolo, (unpublished).
[21] M. Blume, V. J. Emery, and R. B. Griffith, Phys. Rev. A **4**, 1071 (1971).
[22] H. Ez-Zahraouy, Phys. Scr. **51**, 310 (1995).
[23] K. P. Fittipaldi and T. Kaneyoshi, J. Phys. C **1**, 6513 (1989).
[24] B. S. Branco, Physica A **232**, 477 (1996).
[25] A. Z. Akhenyan and N. S. Ananikian, J. Phys. C **29**, 721 (1996).
[26] W. Hoston and A. N. Berker, Phys. Rev. Lett. **67**, 1027 (1991).
[27] A. Maritan *et al.*, Phys. Rev. Lett. **69**, 221 (1992).
[28] M. Keshin, C. Ekiz, O. Yalcin, Physica A **267**, 392 (1999).
[29] M. Sellitto, M. Nicodemi, and J. Arenzon, J. Phys. I **7**, 945 (1997).
[30] G. Schreiber, Eur. Phys. J. B **9**, 471 (1999).
[31] B. S. Branco, Phys. Rev. B **60**, 1033 (1999).
[32] D. Bollé and G. M. Shim, Phys. Rev. E **50**, 5043 (1994).
[33] D. Bollé, H. Rieger, and G. M. Shim, J. Phys. A **27**, 3411 (1994).
[34] G. M. Shim, K. Y. M. Wong, and D. Bollé, J. Phys. A **30**, 2637 (1997).
[35] B. Derrida, E. Gardner, and A. Zippelius, Europhys. Lett. **4**, 167 (1987).
[36] D. Bollé, G. M. Shim, and B. Vinck, J. Stat. Phys. **74**, 583 (1994).
[37] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998).
[38] D. Nozaki and Y. Yamamoto, Phys. Lett. A **243**, 281 (1998); X. Pei, L. Wilkens, and F. Moss, Phys. Rev. Lett. **77**, 4679 (1996).