Information space dynamics for neural networks

R. M. C. de Almeida and M. A. P. Idiart
Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, 91501-970 Porto Alegre, RS, Brazil
(Received 6 February 2002; published 21 June 2002)

We propose a coupled map lattice defined on a hypercube in M dimensions, the information space, to model memory retrieval by a neural network. We consider that both neuronal activity and the spiking phase may carry information. In this model the state of the network at a given time t is completely determined by a function $y(\vec{\sigma},t)$ of the bit strings $\vec{\sigma}=(\sigma_1,\sigma_2,\ldots,\sigma_M)$, where $\sigma_i=\pm 1$ with $i=1,2,\ldots,M$, that gives the intensity with which the information $\vec{\sigma}$ is being expressed by the network. As an example, we consider logistic maps, coupled in the information space, to describe the evolution of the intensity function $y(\vec{\sigma},t)$. We propose an interpretation of the maps in terms of the physiological state of the neurons and the coupling between them, obtain Hebb-like learning rules, show that the model works as an associative memory, numerically investigate the capacity of the network and the size of the basins of attraction, and estimate finite size effects. We finally show that the model, when exposed to sequences of uncorrelated stimuli, shows recency and latency effects that depend on the noise level, delay time of measurement, and stimulus intensity.

DOI: 10.1103/PhysRevE.65.061908 PACS number(s): 87.18.Sn, 05.45.Ra

I. INTRODUCTION

Coupled map lattices [1,2] present a wealth of different and interesting behaviors and are used as a tool to model nonlinear systems made of many interacting elements. In particular, it has been recently shown that coupled map lattices defined over the vertices of a hypercube in M dimensions may present short and long term memory of its exposition to external stimuli, as well as a dynamical mechanism to forgetting [3,4]. In this context, a map is associated with each hypercube vertex, labeled by a string of M bits, and the different states assumed by these maps represent different combinations of patterns that are simultaneously expressed by the system. The hypercube is then the information space of these models where couplings between different patterns are explicitly considered and may be monitored. The fact that each bit string may be mapped to an integer in the interval $[0,2^M-1]$ brings additional advantages to the numerical treatment of the evolution of such systems [5].

Neural networks are generally conceived considering interactions involving two neurons, the one that is firing, that is, releasing neurotransmitters, and the post synaptic neuron that is receiving these neurotransmitters. However, as neurons are not exactly touching each other in the sense that there is a small space—the synaptic cleft—between the presynaptic neuron and the receptors of the postsynaptic neuron, the synapse would be better described as a small region containing extracellular liquid, the receptors of the postsynaptic neuron, and the region of the presynaptic neuron from which the neurotransmitters are released. It is important to notice that other neurons can have axons or dendrites inside the influence region of a given synapse and axon-axon and dendrite-dendrite synapses are also possible. Therefore the postsynaptic response may in fact depend on the state of many neurons, instead of only two, being more exactly described as many-body interactions involving more neurons. Moreover, both the release and reception of neurotransmitters are strongly influenced by the local properties of the synaptic cleft, as an electrical potential or the concentration of neurotransmitters, their agonists, or antagonists [6,7]. In this sense, the effective intensity of the exchange of neurotransmitters may vary, depending on the state of the brain as a whole. Recent developments in neuroscience reveal that the biochemistry of the local extracellular medium, due to the presence of hormones and neuromodulators, may modulate the intensity of interactions among neurons, representing global interactions through other channels besides the synapses [8–12]. Regions in the brain may be recognized, where specific information processing takes place as, for example, speech centers or regions where the different sensory organs send their signals to. However, these centers are also intensely connected to other regions in the brain and may receive feedback. Brain activity and the emerging mind originate in this intricate exchange of information through synapses and hormone release and in being recursively and externally stimulated by both environment and signals coming from the body. The modeling of the brain by synapses only, and in fact, by one only homogeneously coupled neural net, is certainly too simple when the aim is to successfully describe the emergence of the mind or, less ambitiously, some specific cerebral function. It is important to consider multi-interactions and a dynamical modulation of these interactions, in the sense that it is the overall activity of the net that should define how these interactions are modulated.

Events happening in different time scales play different roles in the functioning of the brain. Specifically, when a neuron fires, it generally fires a train of pulses. The typical time scale for one pulse is 1 ms followed by an absolute refractory period of 1 ms, such that the maximum firing rate of a neuron could be of the order of 500 Hz. Temporal summation of excitatory postsynaptic potentials, that is, the action potential measured inside the postsynaptic neuron, is possible when the pulses occur in rapid succession, within 5 to 15 ms from one another. The modeling of neural networks by physicists generally considers discrete time evolution, where at each time step a neuron is either active or inactive,

in the sense that the information received from other neurons have been summed at the soma [13–16]. This time step should englobe the necessary time for the neuron to integrate the incoming signals both spatially and temporally, that is, the physicists' time step should be of the order of 15 ms. However, it has been pointed out by many authors that not only the coincidence in neurons activity during the integration time interval, but also the relative phase of pulse trains could be relevant for information processing [17–22]. These relative phases may represent additional variables that regulate the intensity of effective neurotransmitters exchange at a given time. Consequently correlations among spike trains of different neurons, averaged over the integration time, can be regarded as further dynamical variables of the system, also subject to evolution equations.

Furthermore, there are some indications that neural signals are integrated both temporally and spatially. For example, Mountcastle [23–25] proposed minicolumns as sets of neurons more intensely coupled between themselves than with other neurons. Analogously, minicolumns that interact more intensely between themselves than with other minicolumns form the so-called cortical columns. There is evidence of these spatially integrated structures beyond the sensory cortex [26–29], and they have been proposed as the processing units in a mammalian cortex [30,31].

What emerges from the scenario described above is a highly complex structure, with neurons integrating signals both spatially and temporally coming through nonlinear interactions involving many neurons, represented by synapses and other information diffusion channels, and responding to them in a nontrivial way. On the other hand, we observe that, although simple models for neural networks, such as the Hopfield model or the Perceptron and their derivations, have many unbiological features, they do present the possibility of recognizing what information is and how information processing takes place. In these models, initial states or inputs may be mapped to some given information and the result of letting the system evolve is the retrieved information or the output for a question. Some earlier works approached associative memory in attractor neural nets [32–36] and rule learning processes in layered nets [37,38], where multi-interactions have been explicitly considered. The result is that the information processing capability in the two instances is greatly enhanced by multi-interactions.

It would certainly be rather desirable that a more complex model for neural networks could keep the ability of following the information flux, incorporating features such as temporal and spatial signal integration together with the dynamical modulation of the synapses to emulate hormone and extracellular medium effects.

With this too ambitious goal in mind, we essay a first move in this direction by introducing a model for neural networks where we can recognize Hebb terms and Hebb-like terms for multi-interactions, as well as the modulation of the interactions by the global activity of the net, in an attempt to incorporate biologically based hypotheses to the information processing capabilities of previous models. The model starts from a different point of view in relation to classical neural network models, by proposing evolution equations in an ''information'' space from which the dynamics for the neurons can be derived. In Sec. II we define information space and present the model, in Sec. III and IV we define order parameters and discuss analytical results, and in Sec. V we present numerical simulations results and apply this model to short term human memory. Finally, in Sec. VI we discuss our findings and conclude.

II. THE MODEL

A. Experimental quantities and model variables

The definition of complexity is rather controversial, but in general we can state that more complex devices may discriminate subtler differences in stimuli, yielding richer response repertoires. When proposing a mathematical model, higher complexity implies that the model system states and external stimuli should present more possibilities, leading to different trajectories in the life history of the system. In very complex systems, small differences in the system state or in the external stimulus may imply completely different outcomes. Neurons are complex unities and we should expect that a whole network of neurons is even more complex. On the other hand, it is common wisdom that one should not complicate models more than necessary at the price of dealing with too much information to infer the relevant causes of a given phenomenon. The optimal equilibrium between complexity and simplicity, such that the model is tractable and the phenomenon is still present may be elusive and difficult to reach. Modeling is also an art.

Neural networks, as has been vastly investigated by physicists, consist of a network of simple unities that can in general assume two values. The connection between the mathematical models and real neural networks is made through the assumption that the value $+1$ of a binary variable S_i at a given time t should be associated with the experimental fact that the neuron is active. Since a neuron in physiological conditions is always spiking, activity has to be interpreted as a state where the neural firing rate exceeds the baseline firing rate.

Physicists then proceeded by assuming evolution equations for the individual neurons, that take into account the state of the network of these idealized neurons in previous times. Typically the interaction is considered to happen between each pair of neurons, describing independent synapses; that is, each interaction depends on the state of the two involved neurons only and on a predetermined synaptic parameter, fixed during a previous learning phase. The results are well described in the literature (for reviews see [13–16]) and, although very interesting, these models present serious limitations in reproducing the observed behavior of real neuronal networks. From this particular point of view, we can say that these systems are not responding differently for different stimuli and hence other, more complex models should be proposed.

There are at least two different directions to increase complexity in neural networks. The first is to assume multi-interactions, as has been proposed in some earlier works [32–38]. The results show an enhanced information processing performance but the models still lack biological features.

The second direction is to consider more complex unities, that is, a system of N neurons should have more possibilities to describe its state. In this work we propose a model that increases complexity in both directions.

We assume that to completely determine the state of the network at a given time one must know not only the firing rate of each neuron, but also their spiking relative phases in a time interval Δt, of order of 15 ms, during which a neuron can spike and recover several times. This is the relevant time scale for the discrete dynamics of our model and it is considered as the integration time scale, or the integration time that leads to cognitively meaningful states. The spikes have roughly the same shape, varying the cross membrane potential from to -65 mV up to 40 mV, then decreasing to -80 mV, and finally relaxing back to -65 mV during a time interval of about 2 ms. One way to model the state of a neuron is to consider a partition of the interval Δt in K equal slices, typically less than 2 ms, and assign a value 1 if the neuron spiked in that time slice and zero otherwise. The state of a neuron during a given integration time interval is then given by a sequence of K bits, similar to what is done in information theoretical analysis of spike trains [39,40]. As we shall see in what follows, in this model neuron states at a given time t determine the neuron states in the next time interval, labeled by $t+\Delta t$, implying that the dynamics of the model takes into account the exact spike times occurring in K previous time slices. This fact may be regarded as a form of temporal integration of neural activity.

Besides temporal integration, spatial integration may also be considered. Individual neurons subject to the same stimuli present large variability regarding whether and when the spikes occur, generating sources of noise that may degrade information processing. Summing over a set of neurons as well as different time slices could decrease this effect. Similar to the integration time interval Δt, an integrated processing unit may be defined as a set of intensely coupled neurons, in the spirit of the minicolumns that were proposed by Mountcastle [23–25]. We note that spatially integrated structures have been found in the somatosensory cortex as early as 1956 [25], in the visual cortex [26,27], and more recently their existence has been proposed beyond the sensory cortex [28,29]. To consider the evolution of these processing units, one may then consider the joint evolution of all neurons of the set, during a finite time interval. We model this joint temporal and spatial integration by considering that each unit is composed of ν neurons, such that the state $\tilde{S}_i(t)$ of the ith processing unit at time t is specified by the state of each one of the ν neurons at each time slice in the interval between $t-\Delta t$ and t, that is

$$\tilde{S}_i(t)=(s^i_{1,1},s^i_{1,2},\ldots,s^i_{1,K},s^i_{2,1},s^i_{2,2},\ldots,s^i_{2,K},\ldots,s^i_{\nu,1},s^i_{\nu,2},\ldots,s^i_{\nu,K}), \tag{1}$$

where $s^i_{j,\tau}=\pm 1$ indicates whether the jth neuron in processing unit i has spiked in the τth time subinterval.

Correlation functions among different units are defined as

$$\langle \tilde{S}_{i_1}(t)\tilde{S}_{i_2}(t)\cdots\tilde{S}_{i_m}(t)\rangle=\frac{1}{\nu K}\sum_{j=1}^{\nu}\sum_{\tau=1}^{K} s^{i_1}_{j,\tau}s^{i_2}_{j,\tau}\ldots s^{i_m}_{j,\tau}, \tag{2}$$

where $i_1, i_2, \ldots, i_m$ correspond to different processing units. As each bit $s^i_{j,\tau}=\pm 1$, the above equation implies that correlation functions with one or more repeated units are redundant. For example, the self-correlation function $\langle \tilde{S}^k_i\rangle$ is calculated as

$$\langle \tilde{S}^k_i\rangle=\frac{1}{\nu K}\sum_{j=1}^{\nu}\sum_{\tau=1}^{K}(s^i_{j,\tau})^k=\begin{cases}\langle \tilde{S}_i\rangle & \text{if } k \text{ is odd,}\\ 1 & \text{if } k \text{ is even.}\end{cases} \tag{3}$$

Consequently, the complete set of correlations involving only different units carry all information about the system.

For a network with N neurons, there are $M=N/\nu$ processing units, and M correlation functions involving only one unit. The firing rate, defined as the average number of spikes produced by the ν neurons during the integration time Δt, is therefore

$$r_i=\frac{1}{2\tau_0}[\langle \tilde{S}_i\rangle+1], \tag{4}$$

where $\tau_0=\Delta t/K$ is defined as the lasting time of a spike, and hence $0\leqslant r_i\tau_0\leqslant 1$ since $-1\leqslant\langle \tilde{S}_i\rangle\leqslant 1$.

For $m>1$ in Eq. (2), the correlation functions involve more units and carry information on the spiking relative phases. Observe that for the same values of $\langle \tilde{S}_1\rangle$ and $\langle \tilde{S}_2\rangle$, for example, there can be different values for $\langle \tilde{S}_1\tilde{S}_2\rangle$. In fact, there are 2^M-1 different correlation functions for $0<m\leqslant N$, and they may assume the discrete values $-1,-1+2/(K\nu),\ldots,-2/(K\nu),0,2/(K\nu),\ldots,1-2/(K\nu),1$, that is, $1/(K\nu)$ gives a scale for the correlation functions, such that when $K\nu\to\infty$ the correlation functions are continuous quantities. These correlation functions are well defined quantities that may be, in principle, experimentally measured. It is their evolution that we propose to model here and we do it in an indirect way, using an associated space that we call information space.

We begin by defining a given information pattern by a bit string $\vec{\sigma}=(\sigma_1,\sigma_2,\ldots,\sigma_M)$ of M bits ($\sigma_i=\pm 1$ for $1\leqslant i\leqslant M$). This bit string is a vector in an M-dimensional space but may also be mapped to the binary representation of an integer, here represented by σ, such that $0\leqslant\sigma\leqslant 2^M-1$. The idea is to somehow associate information patterns to con-

figurations of a neural network. As the bit strings representing the information patterns have M bits and to completely represent the configuration of a neural network we need $K\nu M$ bits, only when $K=1$ and $\nu=1$ the information patterns may be mapped to the instantaneous configuration of the net. In this case the processing units in the neural network are composed by single binary neurons ($\nu=1$) with nonsliced integration times ($K=1$). Here we work with more complex units, so this map is not straightforward.

In order to obtain the map between the configuration of the network and the information patterns, we consider that the network made of M processing units may simultaneously express different information patterns $\vec{\sigma}$ with intensities $y(\vec{\sigma},t)$ at a given time t. Given the expression intensity of all information patterns, one should be able to uniquely determine the state of the network that is accomplishing such a deed and vice versa. This is possible by prescribing the following map between the representation intensity function $y(\sigma,t)$ and the quantities representing the network activity and correlations:

$$a(t)=\sum_{\sigma=0}^{2^M-1} y(\sigma,t),$$

$$\langle \tilde{S}_i(t)\rangle a(t)=\sum_{\sigma=0}^{2^M-1} y(\sigma,t)\sigma_i,$$

$$\langle \tilde{S}_i(t)\tilde{S}_j(t)\rangle a(t)=\sum_{\sigma=0}^{2^M-1} y(\sigma,t)\sigma_i\sigma_j, \tag{5}$$

$$\vdots$$

$$\langle \tilde{S}_1(t)\tilde{S}_2(t)\cdots\tilde{S}_M(t)\rangle a(t)=\sum_{\sigma=0}^{2^M-1} y(\sigma,t)\sigma_1\sigma_2\cdots\sigma_M.$$

The number of correlation functions on the left-hand side must be equal to the number of averages in the information space that lays in the right-hand side of the above equations, which explains the reason why the bit-string length M in the information space must be equal to the number N/ν of processing units, where N is the total number of neurons in the net. The role played by K, the number of time slices in the integration time, is to approach the correlation functions to the continuous limit, ($K\nu\to\infty$). Observe that, given all experimental quantities, in the continuous limit, we can univocally determine $y(\sigma,t)$ up to a normalization constant $a(t)$. On the other hand, $a(t)$ can be viewed as an overall information activity of the network, as measured in the information space.

B. The dynamics

The dynamics is modeled by the way the information patterns interact with one another in the information space. That is, we propose a dynamics for the intensities $y(\sigma,t)$ as follows:

$$y(\sigma,t+1)=[1-a(t)]y(\sigma,t)\times\left[x(\sigma)+\frac{z}{a(t)}\sum_{i=1}^{M} y(\sigma^{(i)},t)\right], \tag{6}$$

where the integer $\sigma^{(i)}=\sigma+(1+\sigma_i)2^{i-2}$ is associated with the vertex neighbor to σ in the hypercube that has its ith bit, σ_i, flipped. z is a parameter of the model that regulates the coupling between a given information pattern and its neighbors in the information space. One can imagine couplings between information patterns with more bits flipped or some other neighborhood relation; this is certainly interesting but is beyond the scope of the present work. Observe also that the information activity $a(t)$ of the net modulates the coupling between neighboring sites in the information space: when the net is expressing a lot of different information and is too active the association between similar information is less intense.

To better appreciate the relevance of each term, observe that in some cases Eq. (6) may be regarded as a logistic map with an effective parameter $\lambda=\{x(\sigma)+[z/a(t)]\Sigma_{i=1}^{M} y(\sigma^{(i)},t)\}$. Depending on whether λ is less than, equal to, or greater than one, there may be attractors for $y(\sigma,t\to\infty)$ larger than zero, for some σ. There are two terms in the expression for λ: $x(\sigma)$, which does not depend on time and is a function of σ only, and a dynamically set term, which describes the coupling between different information patterns. Hence, the retrieving of a given pattern may or may not be stable depending on the value of $x(\sigma)$, and the first term in the square bracket of the right-hand side of Eq. (6) describes the difference between permanently learned and not learned patterns. On the other hand, the second term dynamically sets the possible values for the effective parameter λ: it describes how the state of the whole network influences the effective retrieving of a given pattern.

Equation (6) describes the evolution of the pattern intensities $y(\sigma,t)$, making the processing units to follow a given trajectory. In fact, there is an underlying dynamics for the neurons that may be made apparent by multiplying Eq. (6) by σ_i and summing over σ. We then arrive at

$$a(t+1)\langle \tilde{S}_i\rangle_{t+1}=\sum_{\sigma=0}^{2^M-1}\sigma_i[1-a(t)]y(\sigma,t)\times\left[x(\sigma)+\frac{z}{a(t)}\sum_{i=1}^{M} y(\sigma^{(i)},t)\right]. \tag{7}$$

The first term in the large square brackets, containing information about stored memories, may be related to Hebb-like learning rules for the synapses, while the second term is the highly nonlinear term that describes nonlearned synapses and other global connections between the units and consequently between the neurons in the net. Obviously to obtain the Hebb-like synapses we must conveniently define the function $x(\sigma)$. We choose $x(\sigma)$ to assume either one of two values:

$$x(\sigma)=\begin{cases} k_m & \text{if } \sigma \text{ is a learned pattern,} \\ k_v & \text{otherwise,} \end{cases} \tag{8}$$

where k_m and k_v are conveniently chosen to ensure, respectively, that σ is or is not a memory, that is, it may present an attractor state with an intensity $y(\sigma,t\to\infty)$ larger than zero.

Now, suppose that there is a set of P memorized patterns σ^μ, for $\mu=1,2,\ldots,P$. In case we want to reobtain the Hopfield model in the appropriate limits, we must consider that the patterns images on the hypercube, $\overline{\sigma^\mu}=(2^M-1)-\sigma^\mu$, are also stored [32]. The function $x(\sigma)$ may be written as

$$x(\sigma)=k_v+(k_m-k_v)\sum_{\mu=1}^{P}[\,\delta(\sigma-\sigma^\mu)+\delta(\sigma-\overline{\sigma^\mu})], \quad (9)$$

where $\delta(\sigma-\sigma^\mu)=1$ if $\sigma=\sigma^\mu$ and zero otherwise. The trick now is to write the δ functions as follows:

$$\delta(\sigma-\sigma^\mu)=\prod_{i=1}^{M}\frac{1+\sigma_i\sigma_i^\mu}{2}$$

$$=\frac{1}{2^M}\left[1+\sum_{i=1}^{M}\sigma_i\sigma_i^\mu+\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\sigma_i\sigma_i^\mu\sigma_j\sigma_j^\mu+\cdots+\sigma_1\sigma_1^\mu\sigma_2\sigma_2^\mu\cdots\sigma_M\sigma_M^\mu\right], \quad (10)$$

such that by using this expansion for both δ functions in Eq. (9), we may rewrite $x(\sigma)$ as

$$x(\sigma)=k_v+\frac{2P}{2^M}(k_m-k_v)\left[1+\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}J_{ij}^{(2)}\sigma_i\sigma_j+\sum_{i=1}^{M-3}\sum_{j=i+1}^{M-2}\sum_{k=j+1}^{M-1}\sum_{l=k+1}^{M}J_{ijkl}^{(4)}\sigma_i\sigma_j\sigma_k\sigma_l+\cdots+J_{12\ldots M}^{(M)}\sigma_1\sigma_2\cdots\sigma_M\right], \quad (11)$$

where the synaptic intensities $J^{(k)}$ describe multi-interactions involving k neurons and are given as

$$J_{ij}^{(2)}=\frac{1}{P}\sum_{\mu=1}^{P}\sigma_i^\mu\sigma_j^\mu,$$

$$J_{ijkl}^{(4)}=\frac{1}{P}\sum_{\mu=1}^{P}\sigma_i^\mu\sigma_j^\mu\sigma_k^\mu\sigma_l^\mu, \quad (12)$$

$$\vdots$$

$$J_{12\ldots N}^{(N)}=\frac{1}{P}\sum_{\mu=1}^{P}\sigma_1^\mu\sigma_2^\mu\cdots\sigma_N^\mu,$$

which are the same expressions presented in Refs. [32] and [37]. Here we have only even order synapses as a consequence of storing both patterns and their images. The expansion for $x(\sigma)$ proposed in Eq. (11) implies that the first term in the evolution equation Eq. (7) is regulated by Hebb-like terms. Using Eq. (11) in Eqs. (7) we obtain

$$\begin{aligned}a(t+1)\langle\tilde{S}_i\rangle_{t+1}=&\left[k_v+\frac{2P}{2^M}(k_m-k_v)\right]\sum_{\sigma=0}^{2^M-1}[1-a(t)]y(\sigma,t)\sigma_i+\frac{2P}{2^M}(k_m-k_v)\\&\times\sum_{\sigma=0}^{2^M-1}[1-a(t)]y(\sigma,t)\sum_{j=1}^{M-1}\sum_{k=j+1}^{M}J_{jk}^{(2)}\sigma_j\sigma_k\sigma_i+\frac{2P}{2^M}(k_m-k_v)\\&\times\sum_{\sigma=0}^{2^M-1}[1-a(t)]y(\sigma,t)\sum_{j=1}^{M-3}\sum_{k=j+1}^{M-2}\sum_{l=k+1}^{M-1}\sum_{m=l+1}^{M}J_{jklm}^{(4)}\sigma_j\sigma_k\sigma_l\sigma_m\sigma_i\vdots\\&+\sum_{\sigma=0}^{2^M-1}[1-a(t)]y(\sigma,t)\frac{z}{a(t)}\sum_{i}^{M}y(\sigma_i,t)\sigma_i.\end{aligned} \quad (13)$$

That can be rewritten as

$$a(t+1)\langle\tilde{S}_i\rangle_{t+1}=[1-a(t)]a(t)\Bigg\{\left[k_v+\frac{2P}{2^M}(k_m-k_v)\right]\langle\tilde{S}_i\rangle_t+\frac{2P}{2^M}(k_m-k_v)\sum_{j=1,j\neq i}^{M}J^{(2)}_{ij}\langle\tilde{S}_j\rangle_t+\frac{2P}{2^M}(k_m-k_v)$$
$$\times\sum_{j=1,j\neq i}^{M-1}\sum_{k=j+1,k\neq i}^{M}J^{(2)}_{jk}\langle\tilde{S}_i\tilde{S}_j\tilde{S}_k\rangle_t+\frac{2P}{2^M}(k_m-k_v)\sum_{j=1,j\neq i}^{M-2}\sum_{k=j+1,k\neq i}^{M-1}\sum_{l=k+1,l\neq i}^{M}J^{(4)}_{ijkl}\langle\tilde{S}_j\tilde{S}_k\tilde{S}_l\rangle_t$$
$$+\frac{2P}{2^M}(k_m-k_v)\sum_{j=1,j\neq i}^{M-3}\sum_{k=j+1,k\neq i}^{M-2}\sum_{l=k+1,l\neq i}^{M-1}\sum_{m=l+1,m\neq i}^{M}J^{(4)}_{jklm}\langle\tilde{S}_i\tilde{S}_j\tilde{S}_k\tilde{S}_l\tilde{S}_m\rangle_t\ldots\Bigg\}$$
$$+z\frac{[1-a(t)]}{a(t)}\sum_{\sigma=0}^{2^M-1}y(\sigma,t)\sigma_i\sum_{j}^{M}y(\sigma^{(j)},t). \tag{14}$$

In the above equations we can recognize three different classes of terms involving synaptic intensities. The first class contemplates the usual interactions, where other processing units act on the ith unit state through terms as $J^{(2)}_{ij}\langle\tilde{S}_j\rangle_t$. In the limit where $\nu=1$ and $K=1$ it is a usual, two-body, Hebb-like interaction, for other values of ν and K it represents a mean field action of unit j on unit i. The second class considers interactions as $J^{(4)}_{ijkl}\langle\tilde{S}_j\tilde{S}_k\tilde{S}_l\rangle_t$, which describes how the joint activity of three units (j,k,l) may act on a fourth (i) composing a fourth order coupling. These are mean field approximations to Hebb-like terms describing many-body interactions. The third class of terms, such as $J^{(2)}_{jk}\langle\tilde{S}_i\tilde{S}_j\tilde{S}_k\rangle_t$, describes how the interaction between other units (jk) influences the ith unit evolution, depending on the current state of $\tilde{S}_i$. This term can be interpreted as the consequences on $\langle\tilde{S}_i\rangle_{t+1}$ due to changes in the extracellular medium caused by the activation of synapses between other neurons in the network, for example.

In the discussion above, we chose to store both pattern and image to be able to recover the Hopfield model dynamics in the appropriate limits. We do not have to do so. In the results we present in Sec. III we do not make this assumption, since we think this is the more general case. We remark, however, that when images are not stored, the odd order synapses are also present and their learning rules are direct generalizations of the even order synapses learning rules.

The last term in Eq. (14) also deserves some comments. It is a highly nonlinear term and it can be shown that it may be decomposed in sums of products of correlation functions involving a different number of units. They do not contain Hebb-like synaptic intensities and are not modified in learning. They are interpreted as describing inbuilt relations, represented by z, between neurons that are specific of each brain center. In the particular model we propose here, these relations have been chosen to pair similar information patterns and to endow the network with content addressable memory capabilities. In other models, describing other devices, different relations between information patterns may be assumed. We also stress that the assumption of Eq. (6) is rather arbitrary. It has been chosen due to its memory device properties [4], due to the fact that it presents a sensible limit for $\nu=K=1$, and because it yields promising results. However, other equations could also present the same advantages. The point here is to propose an information space formalism with the adequate transformation to the network configuration space as a convenient tool to approach information processing by biological neuronal networks.

The correlation functions are hence relevant in determining the evolution of the system. In fact, for the evolution equations given in Eq. (6) to completely specify $y(\sigma,t+1)$ (up to the normalization constant), it is necessary to know all pattern intensities at time t. As there are 2^M intensities, the complete specification of the state of the net requires a phase space of 2^M dimensions, where Eq. (6) can be regarded as a master equation of a Markov process. Observe that the set of the quantities $a(t)$, $\langle\tilde{S}_i\rangle_t$, $\langle\tilde{S}_i\tilde{S}_j\rangle_t$, $\langle\tilde{S}_i\tilde{S}_j\tilde{S}_k\rangle_t,\ldots,$ $\langle\tilde{S}_1\tilde{S}_2\cdots\tilde{S}_M\rangle_t$ contains 2^M elements and allows us to determine the 2^M values of $y(\sigma,t)$ at a given time t and their subsequent evolution. In the next section we define order parameters that we use to investigate the behavior of the present model.

III. ORDER PARAMETERS

We first define the average overlap $\langle m^\mu\rangle_t$ of the network with the information pattern σ^μ at time t as an average over the information space where the weights are given by the relative intensity $y(\sigma,t)/a(t)$ with which the information σ is being expressed by the net at time t,

$$\langle m^\mu\rangle_t=\sum_{\sigma=0}^{2^M-1}\frac{y(\sigma,t)}{a(t)}m(\sigma^\mu,\sigma), \tag{15}$$

where the time dependence appears through the average implied in this equation and the specific overlap $m(\sigma^\mu,\sigma)$ is the usual overlap between the stored pattern σ^μ and the information σ,

$$m(\sigma^\mu,\sigma)=\frac{1}{M}\sum_{i=1}^{M}\sigma^\mu_i\sigma_i. \tag{16}$$

We note that σ_i may take the values ±1 so that $m(\sigma^\mu,\sigma)$ may assume values in the interval $[-1,1]$ and assumes 1, 0,

or -1 when σ and σ^μ are, respectively, equal, orthogonal, or images of one another in the information space. The specific overlaps can be written in terms of the Hamming distance H, defined as the number of different bits between two patterns:

$$m(\sigma^\mu,\sigma)=1-\frac{2H(\sigma^\mu,\sigma)}{M}. \tag{17}$$

In what follows we obtain stationary values and time evolution for the above defined quantities under different protocols. Depending on how the stored patterns are chosen and the prescription we use to run and initialize the time evolutions, we can gather different information about the performance of the net.

IV. STATIONARY SOLUTIONS

We consider a network with P sparsely stored patterns, such that memories are not first or second neighbors on the hypercube. We look for stationary solutions where $l\langle P$ of these patterns are simultaneously retrieved, in the sense that the intensities y of these patterns are greater than zero. In this case, the simplest stationary solution is

$$y(\sigma,t)=\begin{cases} y_0^\mu & \text{if } \sigma=\sigma^\mu \quad \text{for } \mu=1,2,\ldots,l, \\ y_1^\mu & \text{if } \sigma \text{ is first neighbor to a } \sigma^\mu, \\ \text{for} & \mu=1,2,\ldots,l, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Using this solution in Eq. (6) we get

$$y_0^\mu=(1-a)y_0^\mu\left[k_m+\frac{zMy_1^\mu}{a}\right], \tag{19}$$

$$y_1^\mu=(1-a)y_1^\mu\left[k_v+\frac{zy_0^\mu}{a}\right].$$

Assuming that $y_0^\mu>0$ and $y_1^\mu>0$ for all $\mu\leqslant l$, the above equations imply that

$$y_1^\mu=\frac{1}{M}\left[y_0^\mu-\frac{a(k_m-k_v)}{z}\right]. \tag{20}$$

Summing over μ we get

$$\sum_{\mu=1}^{l} y_1^\mu=\frac{1}{M}\left[\sum_{\mu=1}^{l} y_0^\mu-\frac{al(k_m-k_v)}{z}\right]. \tag{21}$$

We now can use that

$$a=\sum_{\mu=1}^{l} y_0^\mu+M\sum_{\mu=1}^{l} y_1^\mu \tag{22}$$

together with Eqs. (19) and (21) to obtain the following results:

$$\sum_{\mu=1}^{l} y_0^\mu=\frac{z+l(k_m-k_v)}{2z}\left[\frac{z+l(k_m+k_v)-2l}{z+l(k_m+k_v)}\right],$$

$$\sum_{\mu=1}^{l} y_1^\mu=\frac{1}{M}\,\frac{z-l(k_m-k_v)}{2z}\left[\frac{z+l(k_m+k_v)-2l}{z+l(k_m+k_v)}\right], \tag{23}$$

$$a=\left[\frac{z+l(k_m+k_v)-2l}{z+l(k_m+k_v)}\right].$$

The above equations imply the possibility of many different stationary solutions, depending on the individual values of y_0^μ and y_1^μ. For example, the activity a, given by the third equation of (23), can be written as a sum of any l individual activities $a_\mu=y_0^\mu+My_1^\mu>0$, provided they are compatible with positive values for y_1^μ yielded by Eq. (20).

The constraint of positive intensities imposes limits on the number of simultaneously retrieved memories l. To begin with, $a>0$, and hence

$$l<\frac{z}{2-(k_m+k_v)}. \tag{24}$$

Also, $\Sigma_{\mu=1}^{l} y_1^\mu$ must be positive, so that

$$l<\frac{z}{k_m-k_v}. \tag{25}$$

Considering $0<k_v<1$, the above conditions define a maximum number of simultaneously expressed memories,

$$l_{max}=\begin{cases} z/[2-(k_m+k_v)] & \text{if } k_v<k_m<1, \\ z/(k_m-k_v) & \text{if } k_m>1. \end{cases} \tag{26}$$

The interesting point here is an upper limit for l that may be larger than one, stating that the system may be expressing more than one previously memorized pattern simultaneously, which is a feature presented by short term memory in humans [41].

We have so far analyzed the existence of these stationary solutions. We turn our attention now to their stability. As there are many different solutions, with different stability conditions, we restrict ourselves to special cases when either $l=1$ or all retrieved memories are equally expressed, that is $y_0^\mu=y_0^l$ and $y_1^\mu=y_1^l$ for all $\mu\leqslant l$. In these cases, Eq. (6) is written as a fixed point equation of a logistic map, that is,

$$y_0^l=\frac{z+l(k_m+k_v)}{2l}y_0^l\left(1-\frac{2z}{z+l(k_m-k_v)}y_0^l\right), \tag{27}$$

where we can define the logistic parameter λ as

$$\lambda=\frac{z+l(k_m+k_v)}{2l}, \tag{28}$$

yielding the solution

$$\frac{2z}{z+l(k_m-k_v)}y_0^l=1-\frac{1}{\lambda} \tag{29}$$

that reproduces the solutions given by Eqs. (23) in the adequate limits. Imposing the limits for λ corresponding to stable fixed points of the logistic maps, $1<\lambda<3$, we have

$$\frac{z}{6-(k_m+k_v)}<l<\frac{z}{2-(k_m+k_v)}, \tag{30}$$

where the lower limit corresponds to the onset of bifurcations and the upper limits correspond to the existence of the stationary solutions, and it is the same condition obtained in Eq. (24).

Equations (23) and the above conditions consider that the neighbors of a retrieved pattern are expressed by the net, that is, $y_1^\mu>0$. It is also possible a solution with $y_1^\mu=0$. In this case, the stationary solution is

$$\sum_{\mu=1}^{l} y_0^\mu=1-\frac{1}{k_m} \tag{31}$$

and, when all y_0^μ are equal, it is stable provided $1<k_m<3$ and $k_m\leqslant(k_v+z/l)$.

We have explicitly used that only some of the memorized patterns and first neighbors are being expressed by the net, with every other pattern presenting zero intensity. This is reasonable, since $y_1^l\sim 1/M$ and we expect further neighbors to be even less disturbed. In this approximation, the value for the overlap with the recovered patterns may also be obtained, in the limit where the stored patterns may all be taken as orthogonal to each other. In this case, the average overlap with one of the recovered pattern $\mu\leqslant l$ is

$$\langle m^\mu\rangle=\langle m_l\rangle=\frac{y_0^l+My_1^l(1-2/M)}{l(y_0^l+My_1^l)}$$
$$=\frac{1}{l}-\frac{z-l(k_m-k_v)}{zlM}, \tag{32}$$

where it can be seen that the average overlap with in the retrieving solution increases with M, going to 1 as M goes to infinity. These solutions and their stability have been analyzed regarding the evolution equations as approximations to logistic maps. This assumption may be easily overruled by a further neighborhood effectively acting on the stored pattern, and other stable or nearly stable solutions cannot be discarded.

We can expect limits in the performance of any real tool. Here we can identify at least two different mechanisms that impose limits in the memory capacity of this model. The first one is explicit and originates in the fact that a should be greater than zero. When l increases, y_0^l decreases and memory recovery is less intense. Observe that the maximum number of simultaneously retrieved memories depends on the values of k_m and k_v and, more interestingly, increasing the coupling between patterns, z, l_{max} increases.

The second mechanism to limit the capacity of the net lays in the validity of the assumptions and has to do with the number P of stored patterns. When P is too large, it may not be possible to find information patterns such that $x(\sigma)=k_m$, while their first and second neighbors present $x(\sigma)=k_v$. Anyway, this mechanism suggests that the capacity of the net scales with the size of the information space, that is, with 2^M.

Having investigated stationary states, we must now look at dynamical features of the model, such as the stability of these solutions and the size of the basins of attraction, and their relations with the network size.

V. NUMERICAL SIMULATIONS

A. Simulations without noise

We first choose the values for k_m and k_v that define which patterns are memorized by the net and which are not. In this paper we shall consider $k_m=1.5$ and $k_v=0.5$. We also must choose the coupling parameter and we take $z=1$. For these parameters, Eq. (26) states that $l_{max}\to\infty$, but as $l\leqslant P$, this result should be restated as $l_{max}=P$. However, we note that for too large values of P, Eq. (26) is no longer valid. The stationary solutions as defined by Eqs. (23) take the values $y_0^1=1/3$ and $y_1^1=0.0$ when there is only one retrieved memory with overlap given by $\langle m_{l=1}\rangle=1.0$. We stress that these solutions are valid when the other information patterns intensities $y(\sigma,t)$ are zero or nearly so. This is only possible when memorized patterns are sparsely distributed on the hypercube, since nearby memorized patterns could be excited through the coupling with their neighbors.

Given a number M of processing units, there are 2^M vertices in the information space. This exponential increase in the number of the intensity function components strongly limit our calculations. Here we consider $M=12$, 14, and 16, and the calculations were performed in a Pentium II, 800 MHz personal computer using multispin coding techniques [42] to treat the integers representing the information patterns and to access their binary representation.

Initial states for the simulations are built as follows. An information pattern $\sigma_{initial}$ is generated by flipping h_0 randomly chosen bits of a randomly chosen memorized pattern σ^*. The initial configuration of the system is given by

$$y(\sigma,t=0)=\begin{cases}0.3 & \text{if } \sigma=\sigma_{initial},\\ 0.012/M, & \\ \text{if} & \sigma=\text{first neighbor of } \sigma_{initial},\\ 0.000001r & \text{otherwise,}\end{cases} \tag{33}$$

where r is a random number in the interval $[0,1]$. The initial configuration is represented in the information space by a peak in the intensity distribution located at a distance of h_0 bits from a randomly chosen memorized pattern. We then let the system evolve and we monitored the average overlap with the memorized pattern σ^*, $\langle m^*\rangle$, averaged over 100 samples during 100 time steps. We considered different values for the number P of memorized patterns and different values of h_0.

Consider first the time evolution of $\langle m^*\rangle$ for different values of h_0. Figure 1 shows the results for a system with $M=16$. Typically, for $P=1$, the system always converges to the stationary solution with a peak at the memorized pattern,

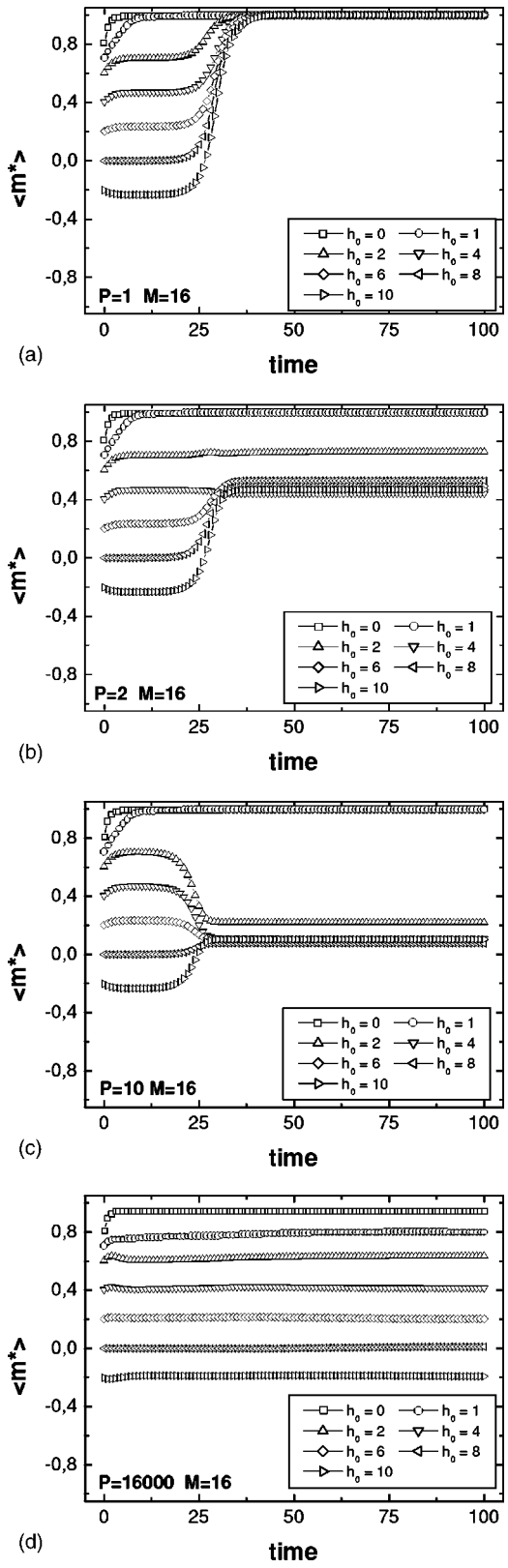

FIG. 1. Time evolution of the overlap $\langle m^* \rangle$ with the stored pattern σ^*, averaged over 100 samples. The initial condition is an intensity function peaked at an information pattern h_0 bits far from σ^*, for $M=16$ and (a) $P=1$, (b) $P=2$, (c) $P=10$, (d) $P=16\,000$.

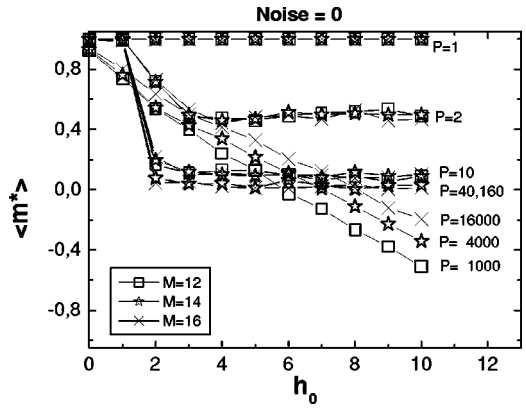

FIG. 2. Long time value of $\langle m^* \rangle$ for different values of P and network sizes of $M=12$, 14, and 16.

regardless of the initial condition, as shown in Fig. 1(a). For all cases, a converges to 1/3 corresponding to the calculated stationary solution. However, the convergence time to this stationary solution increases with h_0. For $P=2$, the final state depends on the initial conditions, as shown in Fig. 1(b). When the system starts too far from σ^*, the final configuration may either converge to a stationary solution around either one of the two memorized patterns, or to a mixed solution, presenting peaks at both memorized patterns, preserving the overlaps with the two stored patterns at roughly 1/2.

Figure 1(c) shows that for $P=10$ there are enough memorized patterns distributed on the hypercube for the final configuration to converge to patterns that are, on average, orthogonal to σ^* when $h_0 \geqslant 2$. As the number of memorized patterns increases still further, the chances that the initial condition is at or very near a memorized pattern other than σ^* also increases and the final configuration converges to these nearby memorized patterns. The final overlap then stays at roughly its initial value, as shown Fig. 1(d), for $P=16\,000$.

We now analyze what happens when $h_0=0$, that is, initially the intensity peak is localized right at a memorized pattern and we vary the number of stored memories P. Figure 2 presents the long time value of $\langle m^* \rangle$ for different values of P and network sizes of $M=12$, 14, and 16. For $P<10$, the performance is the same for every network size: in this limit the number of stored patterns is far from the percolation threshold on the hypercube ($P \ll 2^M/M$), and the long time behavior of $\langle m^* \rangle$ depends on P, but not on M. However, as P increases, the percolation threshold is reached earlier for smaller nets, and we find that analogous behaviors are presented by systems where the value of $P/2^M$ is the same.

Figures 1 and 2 show that the system behaves as an attractor neural network with content addressable memories. The size of the basins of attraction is roughly 2 bits for small P, and this limit is clearly due to the coupling assumed in Eq. (6). Longer range couplings, involving further neighbors in the information space, or larger values for the coupling constant z will have relevant effects in the size of the basins of attraction. This point is under investigation and will be pub-

lished elsewhere. The limit on the network load P depends on the performance indicator we use. In the simplest case of uncorrelated information patterns, with couplings between first neighbors in the information space, the limit for retrieving only one pattern at a time depends on the size of the region covered by the retrieving solution: when only one pattern and its first neighbors are noticeably excited, this limit scales with the percolation threshold of occupying the hypercube with balls formed by a vertex and its first neighbors.

We have considered noiseless equations and their stable solutions. We do not expect this to be the case of neuronal networks in a mammalian brain. Random stimuli both from other parts of the brains, as well as from the body and from the environment are continuously being received by the different regions of the brain. Moreover, individual neurons may present chaotic dynamics and do not work as completely deterministic units. In what follows, we consider the effect of a small random term introduced in the evolution equations.

B. Simulations with noise

The noise term is introduced in the evolution equations of the model, by adding a random term in Eq. (6),

$$y(\sigma,t+1)=[1-a(t)]\times\left\{y(\sigma,t)\left[x(\sigma)+\frac{z}{a(t)}\sum_{i=1}^{M}y(\sigma^{(i)},t)\right]+n(t)\right\}, \tag{34}$$

where $n(t)$ has probability $(1-p_R)$ of being zero and p_R of assuming a small value n_R. In this paper we considered $p_R=0.01$ and $n_R=10^{-4}$. Figure 3 shows the evolution of $\langle m^*\rangle$ starting with $h_0=0$, averaged over 10 samples during 8×10^4 steps for $M=14$ and different network loads P. The solutions are stable only for $P=1$ and 2. For larger P the systems tend to lose the initial memory as time passes, and the typical time for losing the memory depends on P. When P is not too large, increasing P implies a smaller memory decay time. However, for very large P, there is a stabilization at higher values of $\langle m^*\rangle$. This is so because the activity has attained a stable value by the excitation of nearby memories.

This result is particularly interesting. Two relevant features in human short term memory are that (i) it may keep simultaneously aroused several items and (ii) it decays with time. These features are presented by this model, as we have just shown, and to better illustrate its potential we consider in the next section the results for a simulation protocol where the laboratory procedures for performance measurement of short term memory are reproduced.

C. Short term memory performance

As nicely reviewed by Baddeley in his book *Human Memory* [41], short term memory in humans is measured by asking subjects to memorize stimuli, which can vary from a list of written known words to spoken unknown sounds, under different circumstances where the time delay between stimuli and the time delay for testing the memory, background sound or visual stimuli, and simultaneously performed tasks are varied. In the simplest example, subjects are asked to read words that successively appear on a screen. After a list of N items, the subject is asked to remember the words. The retrieving probability, defined as the relative frequency with which each word of the list is retrieved, is obtained after the test is repeated with different subjects. Plotting the retrieving probability as a function of the order number of the words, we can see peaks for the first word, which is called the latency effect, and for the last words in the list, which is known as the recency effect [43]. Typical investigations aim at measuring the number of items that humans can simultaneously remember within a short interval of time and the causes for memory loss. Two different mechanisms for short term memory loss seem to be in action: a natural decay with time, which should be intrinsic to the dynamics of the system, and interference effects, where the memory loss is due to another stimulus the subject has been exposed to.

The model for neural networks we propose here is able to present latency and recency effects and shows both memory loss mechanisms. To show that, we simulated the model

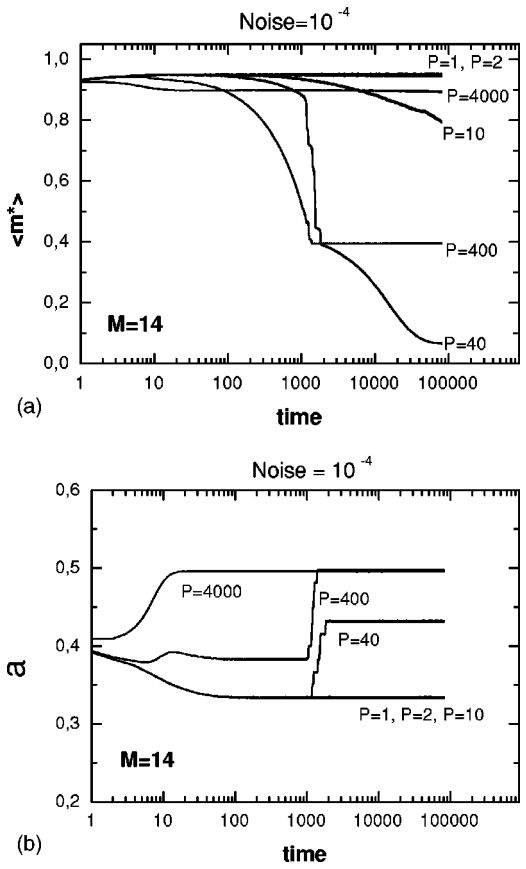

FIG. 3. Evolution of $\langle m^*\rangle$ and a with time, starting with $h_0=0$ and averaged over 10 samples during 8×10^4 steps for $M=14$ and a different network load P.

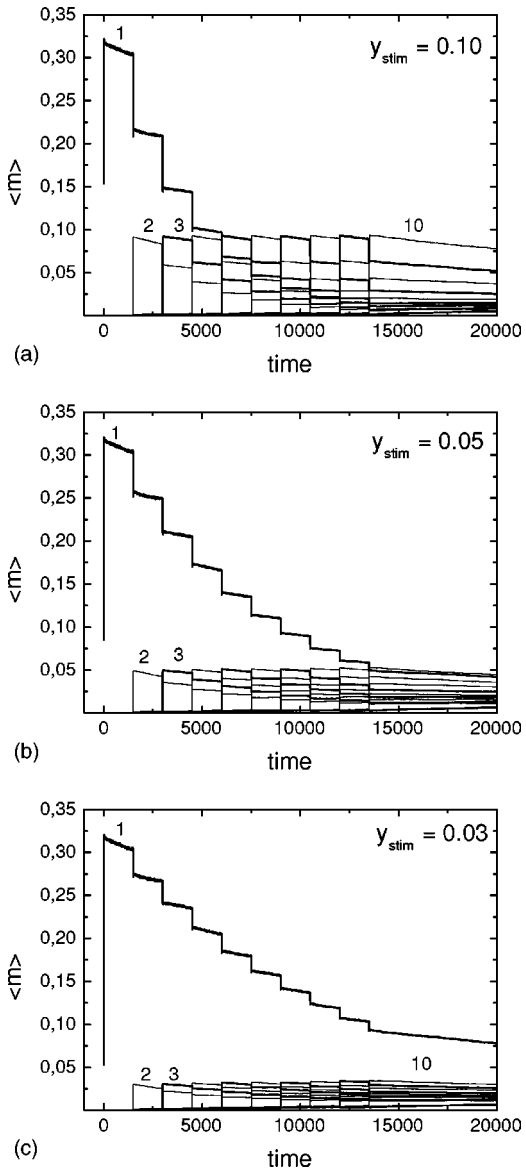

FIG. 4. Time evolution of the average overlap $\langle m(\mu)\rangle$ with time for $\mu=1,\ldots,10$. At intervals of 1500 time steps, y_{stim} is added to a randomly chosen memorized pattern intensity. (a) $y_{stim}=0.1$, (b) 0.05, and (c) 0.03.

equations with noise. We started with every intensity being zero and, after stabilizing the system, at intervals of T time steps we added a fixed value y_{stim} to a randomly chosen memorized pattern. We considered $P=20$, $T=1500$, $y_{stim}=0.1$, 0.05, and 0.03, with other parameters as in the preceding section. Figure 4 shows the time evolution of the average overlap $\langle m(\mu)\rangle$ with time for $\mu=1,\ldots,10$. The presentation of each stimulus is clearly seen by the discontinuity in the curves. When a new stimulus is presented, its overlap jumps from the background to a finite value. At this instant, previously shown patterns suffer a decrease in their overlap, which is interpreted as the interference mechanism for memory loss. Simultaneously, several patterns have their overlaps above the background what may be interpreted as being simultaneously remembered. Eventually these overlaps merge back in the background noise, and that can be interpreted that the associated pattern has been forgotten. After ten stimuli, the network is let to evolve without further perturbation and all overlaps decrease, eventually merging with the background: the decay mechanism for memory loss is clearly present. The decay mechanism in this model is caused by the noise, which here we take as being intrinsic of the dynamics of individual neurons and the exposition of the network to some environment.

When measured immediately after the presentation of the last stimulus, the lastly presented patterns have their overlap above the background, that is, the last stimuli are always remembered. This is the recency effect, which may disappear when the measurement is performed later after the presentation of the last stimulus due to the intrinsic decay.

The latency effect, that is, the high probability of remembering the first stimulus in the list, is present when the stimuli are not too intense. Everything happens as if the first stimulus had forced the network in a quasistable solution, with a large decaying time, which weak stimuli cannot strongly disturb. Strong enough stimuli, on the other hand, disrupt this solution and latency effects are not verified anymore.

As we mentioned, there are several different protocols and procedures to investigate short term memory, where the noise level, number of items in the list, waiting times, simultaneous tasks, correlation between the items in the list, etc., are varied. It would certainly be interesting to investigate whether and if yes, under which circumstances, the present model is able to reproduce the experimental results.

VI. DISCUSSION AND CONCLUSIONS

We have presented a model to simulate an associative memory device, and we indicated how the different terms in the model evolution equations could be realized by a network of neurons where multi-interactions and modulations of these interactions are taking place in such a manner that the many neuron correlation functions are also dynamic variables of the system. The model is very idealized in the sense that all possible interactions and modulations are present and it worked well as an associative device. It has the appeal of pointing in what direction all the wet machinery present in the brain may be acting to enhance its information processing abilities, and indicates the relative spiking phases together with an assembly of neurons as quantities that should be further monitored. On the other hand, the assumed interactions in the information space may be too simple and the incorporation of so many more dynamic variables (from M to 2^M), besides the desirable fact of enhancing information processing abilities, brings along the unwanted increase in the demands of computational resources to deal with realistic numbers of neurons.

The interactions in the information space are supported by synapses and modulations among real neurons. A careful analysis of what terms are possible and what are not in real

systems is a necessary step to obtain a more realistic, or less idealized, model of neuronal networks. However, in our opinion, the analysis should start from two different points of view. At one end, we should consider only a few, more relevant terms, supported by experimental evidence, where the first of these terms corresponds to the two neuron, Hebb synapses. At the other end, the analysis should start by cutting some terms from an ideal, optimal model and by studying its performance as an information processor. The present model could be the zeroth step of the investigation from the idealized system side.

We applied the model to describe short term memory in humans and could find recency and latency effects, and two different mechanisms of memory loss (decay and interference). In our opinion, these results are very encouraging, since they link the results of experiments with human performance and neuron dynamics.

The dynamics for the intensity function, as proposed by Eqs. (6) and (34), is certainly arbitrary. They have the virtue of presenting Hebb-like terms in memorization-dependent terms, but these terms could have different functional forms. The justification for this choice here is hence *a posteriori* since they yield sensible results. However, a thorough experimental investigation must be performed before we can state we have a first principle model for short term human memory. Also, the second kind of term in the dynamical equations, that is, the pairing between neighbor patterns in the hypercube is intended to describe associative memory, but different couplings are possible to describe other brain functions.

Nevertheless, we stress that a novel and strong point in this approach is the transformation from the neuron network configuration space to the information space, which is made possible by the consideration of more complex dynamical units.

ACKNOWLEDGMENTS

We acknowledge partial financial support from Brazilian agencies FAPERGS, CNPq, and CAPES. We thank J.A. Quillfeldt for fruitful discussions and valuable suggestions.

[1] K. Kaneko, Physica D **34**, 1 (1989).
[2] H. Chaté and P. Manneville, Prog. Theor. Phys. **87**, 1 (1992).
[3] M. C. Lagreca, R. M. C. de Almeida, and R. M. Zorzenon dos Santos, Physica A **289**, 144 (2001).
[4] M. C. Lagreca, K. Aquere, R. M. C. de Almeida, and R. Vianna (unpublished).
[5] R. M. C. de Almeida, N. Lemke, and I. A. Campbell, Eur. Phys. J. B **18**, 513 (2000).
[6] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience: Exploring the Brain* (Williams and Wilkins, Baltimore, 1996).
[7] M. J. Zigmond, F. E. Bloom, S. C. Landis, J. L. Roberts, and L. R. Squire, *Fundamental Neuroscience* (Academic Press, New York, 1999).
[8] A. Damásio, *Descartes' Error: Emotion, Reason, and the Human Brain* (Grosset/Putnam, New York, 1996).
[9] A. Damásio, *The Feeling of What Happens* (Harcourt Brace Jovanovitch, San Diego, 1999).
[10] J. LeDoux, *The Emotional Brain* (Simon and Schuster, New York, 1996).
[11] J.-D. Vincent, *Biologie des Passions* (Éditions Odile Jacob, Paris, 1999).
[12] G. M. Edelmann and G. Tononi, *A Universe of Consciousness: How Matter Becomes Imagination* (Perseus Books Group, New York, 2000).
[13] J. Herz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, New York, 1991).
[14] H. S. Seung, H. Sompolinsky, and N. Tishby, Phys. Rev. A **45**, 6056 (1992).
[15] D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, New York, 1989).
[16] T. L. H. Watkin, A. Rau, and M. Biehl, Rev. Mod. Phys. **65**, 499 (1993).
[17] C. M. Gray and W. Singer, Proc. Natl. Acad. Sci. U.S.A. **86**, 1698 (1989).
[18] E. Vaadia, I. Haalman, M. Abeles, H. Bergman, Y. Prut, H. Slovin, and A. Aertsen, Nature (London) **373**, 515 (1995).
[19] W. Singer, Neuron **24**, 49 (1999).
[20] P. N. Steinmetz, A. Roy, P. J. Fitzgerald, S. S. Hsiao, K. O. Johnson, and E. Niebur, Nature (London) **404**, 187 (2000).
[21] Z. Nádasdy, J. Physiol. (Paris) **94**, 505 (2000).
[22] M. Galarreta and S. Hestrin, Science **292**, 2295 (2001).
[23] V. B. Mountcastle, in *The Mindful Brain*, edited by G. M. Edelman and V.B. Mountcastle (MIT, Cambridge, MA, 1978), p. 1.
[24] V. B. Mountcastle, Brain **120**, 701 (1997).
[25] V. B. Mountcastle, *Perceptual Neuroscience: The Cerebral Cortex* (Harvard University, Cambridge, MA, 1998).
[26] D. Hubel and T. Wiesel, J. Physiol. (London) **160**, 106 (1962).
[27] D. Hubel and T. Wiesel, J. Physiol. (London) **195**, 215 (1968).
[28] P. S. Goldman-Rakic, Trends Neurosci. **7**, 425 (1984).
[29] P. Rakic, Science **241**, 170 (1988).
[30] G. L. Shaw, D. J. Silverman, and D. J. Pearson, Proc. Natl. Acad. Sci. U.S.A. **82**, 2364 (1985).
[31] M. Sardesai, C. Figge, M. Bodner, M. Crosby, J. Hansen, J. A. Quillfeldt, S. Landau, A. Ostling, S. Vuong, and G. L. Shaw, Biol. Cybern. **84**, 173 (2001).
[32] R. M. C. de Almeida and J. R. Iglesias, Phys. Lett. A **146**, 239 (1990).
[33] J. J. Arenzon, R. M. C. de Almeida, and J. R. Iglesias, J. Stat. Phys. **69**, 385 (1992).
[34] J. J. Arenzon and R. M. C. de Almeida, Phys. Rev. E **48**, 4060 (1993).
[35] R. M. C. de Almeida, P. M. C. de Oliveira, and T. J. P. Penna, in *Annual Reviews on Computational Physics*, edited by D. Stauffer (World Scientific, Singapore, 1994), Vol. 1, pp.

193–217.
[36] D. Bollé, J. Huyghebaert, and G. M. Shim, J. Phys. A **27**, 5871 (1994).
[37] E. Botelho, R. M. C. de Almeida, and J. R. Iglesias, J. Phys. A **28**, 1879 (1995).
[38] R. M. C. de Almeida and E. Botelho, Physica A **242**, 27 (1997).
[39] S. P. Strong, R. Koberle, R. R. de Ruyter van Steveninck, and W. Bialek, Phys. Rev. Lett. **80**, 197 (1998).
[40] F. Rieke, D. Warland, R. R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, MA, 1997).
[41] A. Baddeley, *Human Memory: Theory and Practice*, revised edition (Allyn and Bacon, Boston, 1998).
[42] P. M. C. de Oliveira, *Computing Boolean Statistical Models* (World Scientific, Singapore, 1991).
[43] L. Postman and L. W. Phillips, Q. J. Exp. Psycholo. **17**, 132 (1965).