

Article

External Stimuli on Neural Networks: Analytical and Numerical Approaches

Evaldo M. F. Curado , Nilo B. Melgar  and Fernando D. Nobre *

Centro Brasileiro de Pesquisas Físicas and National Institute of Science and Technology for Complex Systems, Rua Xavier Sigaud 150, Urca, Rio de Janeiro 22290-180, Brazil; evaldo@cbpf.br (E.M.F.C.); nilobarrantes@gmail.com (N.B.M.)

* Correspondence: fdnobre@cbpf.br; Tel.: +55-(21)-2141-7513

Abstract: Based on the behavior of living beings, which react mostly to external stimuli, we introduce a neural-network model that uses external patterns as a fundamental tool for the process of recognition. In this proposal, external stimuli appear as an additional field, and basins of attraction, representing memories, arise in accordance with this new field. This is in contrast to the more-common attractor neural networks, where memories are attractors inside well-defined basins of attraction. We show that this procedure considerably increases the storage capabilities of the neural network; this property is illustrated by the standard Hopfield model, which reveals that the recognition capacity of our model may be enlarged, typically, by a factor 10^2 . The primary challenge here consists in calibrating the influence of the external stimulus, in order to attenuate the noise generated by memories that are not correlated with the external pattern. The system is analyzed primarily through numerical simulations. However, since there is the possibility of performing analytical calculations for the Hopfield model, the agreement between these two approaches can be tested—matching results are indicated in some cases. We also show that the present proposal exhibits a crucial attribute of living beings, which concerns their ability to react promptly to changes in the external environment. Additionally, we illustrate that this new approach may significantly enlarge the recognition capacity of neural networks in various situations; with correlated and non-correlated memories, as well as diluted, symmetric, or asymmetric interactions (synapses). This demonstrates that it can be implemented easily on a wide diversity of models.

Keywords: neural networks; models of single neurons; artificial intelligence; nonlinear dynamical systems

PACS: 07.05.Mh, 87.85.dq, 87.19.lj, 87.18.Sn, 05.90.+m



Citation: Curado, E.M.F.; Melgar, N.B.; Nobre, F.D. External Stimuli on Neural Networks: Analytical and Numerical Approaches. *Entropy* **2021**, *23*, 1034. <https://doi.org/10.3390/e23081034>

Academic Editors: Viola Folli, Giorgio Gosti and Edoardo Milanetti

Received: 8 July 2021

Accepted: 5 August 2021

Published: 11 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Although the area of neural networks (NNs) has experienced impressive developments in the last few decades [1–4], essential characteristics and reactions of the brain are still far from being satisfactorily replicated in these models. This is an essential direction to pursue, since NNs were initially introduced in order to reproduce some of the primary functions of the brain [5], and their first main practical application was pattern recognition [6]. So far, NNs have been more successful in artificial intelligence applications, such as information and image processing, than in emulating the reactions of living beings in nature.

The remarkable works of Darwin and Wallace on evolutionary theory [7–9] were fiercely debated throughout the 19th century, but are now considered fundamental to our understanding the world. They are accepted by all scientists as a starting point for studying everything related to living beings, and all organs of living creatures were, of course, molded by evolution. In particular, the brain itself grew and developed under the action of

natural selection; it works in accordance with its evolutionary history. But what is signified by the evolutionary history of the brain? It means the development, from generation to generation, from species to species, of a complex system with fast responses to external stimuli—a fundamental condition of survival. A living being needs to react, in many cases instantaneously, to escape from a predator, to catch prey, to avoid an accident, and many other situations. This implies that the nervous systems of animals were developed as fundamental tools, able to react promptly to changes in the environment.

After millions of years of the development of nervous systems, which generally increased in size and complexity, a concentrated region of nervous cells appeared—the primitive brain—which quickly enlarged its size; this was a breakthrough in the evolution of living beings. However, a further duty was yet to be fully developed. Living beings had to react quickly and more efficiently to changes in the environment—this task was at the origin of the brain and is recorded “au fer et au feu” in the way that it works. This stimuli-dependence of the brain is behind all of its assignments, the old ones (essentially instinctive reactions), as well as the biologically newer ones (e.g., those related to cognition). During the whole of its evolutionary history, pattern recognition was a basic attribute of the brain, allowing living beings to react appropriately. NN models should, therefore, incorporate some of the main features related to this; below we outline, briefly, the main procedure for pattern recognition in living beings.

- (a) Living beings accumulate memories (past input patterns), that are stored in some way, e.g., using Hebb’s rule [10,11].
- (b) Without external stimuli, they do not recognize any pattern, and remain in a noisy state.
- (c) In the presence of an external stimulus that is associated with some stored pattern, they recognize that pattern; if the external stimulus has no relation to any stored pattern, nothing is recognized.
- (d) If the external stimulus associated with some memory disappears, the effectivity of the recognized pattern decreases, becoming essentially null after some time delay. This induces a return to a noisy state.
- (e) In line with a common feature in nature, whenever an external stimulus abruptly changes its pattern, the living being quickly adjusts to the new recognized pattern and away from the old one.
- (f) These steps are followed repeatedly in the presence of each new stimulus, or sequences of stimuli.

Biologically realistic pattern-recognition NN models should reproduce the actions sketched above (at least, most of them), by following the route of new concepts in complex systems [12–15]. These systems essentially live at the chaos–order or ordered–disordered borderlines, and effective NN models should be defined at these borderlines; chaotic, ordered, or disordered regimes can never yield appropriate proposals for NNs.

Along these lines, we will introduce a stimuli-dependent neural-network model, and its effectiveness in pattern recognition will be demonstrated; certainly, similar scenarios can be constructed for other brain process. Stimuli will be defined in the form of an external local field, which has already shown its effectiveness in the suppression of chaos in a recurrent fully connected NN model [16]. Here, we will illustrate how this works using the standard Hopfield model. In the next section we briefly review the paradigmatic Hopfield model, define the class of attractor neural networks (ANNs), and analyze some of their limitations. In Section 3, we present our model, which incorporates several of the points sketched above; the procedure will be applied to the Hopfield model, for which analytical calculations are feasible. We also discuss the new concepts embodied in this approach, emphasizing its pattern storage capacity and showing this is substantially larger than the capacity of ANN models. Analytical calculations within a mean-field approximation, with special attention given to its zero-temperature ($T = 0$) limit, are carried out in Section 4. In Section 5, we present results from numerical simulations of uncorrelated patterns at $T = 0$, comparing some of these with the analytical results. The significant increase in pattern

storage capacity, as compared to ANN models, is indicated by these numerical simulations. The ability of the model to react promptly to changes in the external stimuli—particularly in replicating actions (b)–(e) described above, which are crucial for living beings—is verified in Section 6. The good performance of the present proposal in the case of previously stored correlated patterns is reported in Section 7. In Section 8, we introduce site dilution into the model, revealing that it also functions well for asymmetric diluted synapses. In Section 9, we discuss our main results, emphasizing some of the advantages with respect to previous NN models, and propose potential applications; finally, in Section 10, we present our conclusions.

2. The Hopfield Model and Attractor Neural Networks

A breakthrough in the NN area occurred in the beginning of the 1980s, with the proposal of a model due to Hopfield [17] (known in the literature as the Hopfield model), which has attracted the attention of many during the last few decades [1–3]. Imitating ideas from random magnetic models, particularly those from the Ising spin-glass model [4,18,19], one makes the analogies, *neurons* \leftrightarrow *spins* and *synapse intensities* \leftrightarrow *coupling constants*. In this way, the Hopfield model is defined by means of a Hamiltonian,

$$\mathcal{H}(t) = - \sum_{(i,j)} J_{ij} \sigma_i(t) \sigma_j(t), \quad (1)$$

with $\sigma_i(t) = \pm 1$ ($i = 1, 2, \dots, N$) representing the state of the i -th neuron at time t [5], activated or at rest. Moreover, J_{ij} stands for the intensity of the synapses between neurons i and j (considered as symmetric), whereas $\sum_{(i,j)}$ denotes a sum over all distinct pairs of neurons, corresponding to a fully connected neural network. The intensity of the synapses J_{ij} is expressed in terms of p stored memories $\{\xi_i^\mu\}$, according to Hebb's rule [10,11],

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad (i \neq j). \quad (2)$$

These memories should remain fixed along the whole time evolution (i.e., are quenched variables) and are assumed to be orthogonal on average [for finite N there may occur overlaps between memories of $O(1/\sqrt{N})$]. They obey the equation

$$P(\xi_i^\mu) = \frac{1}{2} \delta(\xi_i^\mu - 1) + \frac{1}{2} \delta(\xi_i^\mu + 1). \quad (3)$$

The definition of the Hamiltonian above has attracted the interest of many physicists, particularly because statistical-mechanics techniques could be applied [20–23]. However, an immediate question concerns the assumption of symmetric couplings $\{J_{ij}\}$, which is a serious obstacle to an appropriate emulation of the brain, since it is well-known that real synapses violate this symmetry, i.e., $J_{ij} \neq J_{ji}$. In this case no Hamiltonian can be defined.

The Hopfield model has helped to create the class of ANNs; systems where the fundamental states become stored memories. Each such memory defines a nontrivial phase space and possesses its own basin of attraction. In ANNs, an initial pattern that has a significant overlap with one stored memory—that is, belonging to its basin of attraction—evolves by means of an appropriate dynamics (stochastic or not) to a lower energy state, i.e., to the stored memory. This leads to the recognition of the pattern [1–4,20–40]. However, once a pattern is recognized, *the system stays in its resultant state forever*, even if the input pattern has acted only at the beginning of the process. Due to this aspect, it is easy to see that, of the items listed in the previous section, few are consistently fulfilled.

Three basic limitations that are common for ANNs are listed and discussed below. (i) They do not operate at the chaos–order border, as any complex system should [12–15], but rather on the ordered side; (ii) They present a maximum limit in their storage capacity; and (iii) They are inefficient when reacting to changes in the external environment.

The ordered side, where a multiplicity of low-energy states appears, is the main ingredient of an ANN; as a proposal for modeling one of the most complex systems in nature, restriction (i) means that the ANNs are very far from achieving an appropriate approach to the brain. The constraint (ii) comes from the fact that the memories correspond to ground states, each one having its own basin of attraction, so that the size of the basins reduces as the number of stored memories increases. This severely limits their number, even if these memories are not correlated. In general, the storage capacity of ANNs, with N neuronal units, may be expressed as $p = \alpha N$, with α characterized by threshold values (α_c); in most cases α_c is much smaller than one. In the standard Hopfield model, it is well-known that $\alpha_c \simeq 0.14$ [22], so that the system becomes unable to recognize any further pattern for $\alpha > 0.14$. The limitation (iii) is related to the above-mentioned fact that, once a pattern is recognized, the system stays in its corresponding state; this feature prevents quick reactions to external changes. In order to consider such reactions in the standard Hopfield model, for $\alpha < \alpha_c$, one should change the initial state $\{\sigma_i(0)\}$, which may lead to a jump in phase space to a different basin of attraction. As will be shown later, for the model we define next, reactions that follow changes in the external environment occur naturally and smoothly.

Modifications to the Hopfield model, such as modifying coupling constants to be asymmetric (in which case there is no Hamiltonian) [41–43], correlating or coupling pairs of memories [44–51], and introducing dilution into the synapses [3,43,52–62], yield minor changes, but largely fail to satisfy the requirements for an appropriate description of the brain [steps (b)–(e) of the previous section], and do not overcome the drawbacks described in items (i)–(iii) above. Asymmetric coupling constants, correlations among memories, and dilution may lead to fixed points or cycles in the dynamics, which appear on the ordered side of the chaos–order dichotomy. These ingredients could also lead to a chaotic state [63,64]; in both situations (ordered or disordered states) one is not at the border of chaos and order, and thus the situation is inappropriate for emulating a complex system such as the brain.

3. Biologically Motivated Model: The Relevance of External Stimuli

Reacting quickly to changes in the neighboring environment is so important to life that many reactions are written into our DNA. Being the most primitive ones; they are commonly called instinctive reactions. The new-born mammal, which sucks anything that enters its mouth, illustrates an instinctive reaction, fundamental to the first weeks of any mammals' life. Rapid and involuntary movements to escape from predators represent provide additional examples, among many other instinctive reactions in animals. Clearly, the brain was forged by evolution for the task of quickly analyzing any external stimulus and triggering a muscle reaction if necessary. With evolution, further types of reactions to external stimuli, non-instinctive reactions, were developed in the nervous systems of many animals. These animals can take some time to analyze the stimulus, recognize it, and then, react.

Prompt reactions to external stimuli should be ubiquitous for any brain activity; besides, of course, the important task of pattern recognition. During pattern recognition, in short, the system is able to store a certain number of patterns (memories), which come from previous experiences. If an external stimulus related to one of these patterns is presented, the system should be able to recognize it as one of the stored patterns.

The typical framework of using ANNs for pattern recognition generates a topography, in an appropriate mathematical space, where memories are at the bottom of the valleys and external stimuli are (initially) located on the mountain slopes, giving birth to the recognition process. An appropriate choice of dynamics should lead the external stimulus from the mountain slopes to the bottom of the valley, associating it with the corresponding memory.

In these types of ANN models, the influence of an external stimulus on a specific neuron belonging to the NN can be essentially split into two contributions; namely, a signal, connected with the external stimulus and correlated to a stored pattern (memory), as well

as noise, produced by all stored memories that are not related to the external stimulus. This external stimulus yields the initial state of neuronal activity that, by evolving according to some internal process, leads the system to recognize the memory associated with the initial stimulus. As the number of stored patterns (memories) increases, the width of the noise distribution also increases, so that when the system attains a sufficiently large number of stored memories, this width becomes of the order of the signal; once this occurs, the NN is not able to recognize any further stored pattern. The ability of an ANN to recognize stored memories fails because the width of the noise distribution (roughly) cancels out the signal from the external initial stimulus.

In this paper, our main assumption is that the long-term evolution of living beings has allowed them to calibrate the external-stimulus influence, so that the noise produced by memories that are not correlated with the external stimulus is canceled as much as possible. Roughly, we can express the influence (h_i) due to other neurons ($j \neq i$) on a particular neuron i , in the presence of an external stimulus correlated with some stored memory, in terms of the following contributions [65],

$$h_i = \text{signal} + \text{noise} + \text{external stimulus}. \quad (4)$$

If the external stimulus succeeds in canceling the noise contribution, the signal should remain, allowing the recognition of the stimulus whenever it is associated to some stored pattern. The external stimulus does not merely correspond to an initial state, which is responsible for starting the process of recognition as happens in standard NNs [1–4], but it will rather be present during the whole process of recognition. In this framework, the initial state $\{\sigma_i(0)\}$ becomes irrelevant; the fundamental point concerns the process of tuning the external-stimulus intensity in order to cancel precisely (or approximately) the noise term. Operationally, in a given computational NN, the first two contributions on the right-hand-side of Equation (4) result from the model construction [e.g., from the definition of the intensity of synapses in Equations (1) and (2)], whereas the last contribution corresponds to an extra term acting on each neuron i , introduced in order to attenuate the noise contribution. From now on, we refer to this scheme as a stimulus-dependent neural network (SDNN).

In principle, this framework may be implemented in many NN models, but here it will be illustrated for the Hopfield model, due to its simplicity and potential for producing analytical results. In this case, the local field h_i at time t becomes

$$h_i(t) = \sum_{j \neq i}^N J_{ij} \sigma_j(t) + \kappa \eta_i, \quad (5)$$

where, in the first term on the r.h.s., we have the usual contribution due to $(N - 1)$ neurons, whereas the second term corresponds to the external stimulus acting on neuron i . This later contribution should remain fixed during the whole time evolution, with $\eta_i = \pm 1$ and κ ($\kappa \geq 0$) depicting its intensity, to be considered in the recognition process. The main idea is to maintain the influence of the external pattern during the whole process, and not only as an initial state, as happens in typical ANNs.

We will consider two typical situations concerning this external stimulus: (i) It should present an overlap with one specific stored memory (e.g., memory ρ , $1 \leq \rho \leq p$). It should be orthogonal to all other memories and should lead to an external pattern being correlated with a specific memory; (ii) It should be orthogonal to all stored memories, possess no overlap with any memory, and correspond to an external pattern not correlated with any stored memory. Case (i) can be expressed by assuming that the set $\{\eta_i\}$ obeys the probability distribution

$$P(\eta_i) = \gamma \delta(\eta_i - \xi_i^{\mu}) + (1 - \gamma) \delta(\eta_i + \xi_i^{\mu}), \quad (6)$$

where $1/2 < \gamma \leq 1$ for $\mu = \rho$ and $\gamma = 1/2$ for $\mu \neq \rho$. The particular case $\gamma = 1$ means that the set of external-stimulus signs $\{\eta_i\}$ yields the same pattern as the memory ρ , covering common real situations where one should recognize an external pattern that coincides precisely with the stored one. On the other hand, case (ii) is characterized by

$$P(\eta_i) = \frac{1}{2}\delta(\eta_i - \zeta_i^\mu) + \frac{1}{2}\delta(\eta_i + \zeta_i^\mu), \quad (1 \leq \mu \leq p). \quad (7)$$

Let us investigate the local field of Equation (5), by separating the contribution of memory ρ from those of other memories ($\mu \neq \rho$) in Equation (2); one obtains,

$$h_i(t) = \frac{1}{N} \sum_{j \neq i}^N \zeta_i^\rho \zeta_j^\rho \sigma_j(t) + \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu \neq \rho}^p \zeta_i^\mu \zeta_j^\mu \sigma_j(t) + \kappa \eta_i. \quad (8)$$

Comparing the equation above with Equation (4), one immediately identifies each of the contributions on its r.h.s.: the first one is the signal activated when the neuron configuration $\{\sigma_i(t)\}$ is roughly the same as the stored memory $\{\zeta_i^\rho\}$; the second one represents the noise, induced by other memories distinct from ρ ; and the third contribution corresponds to the external stimulus. In the case $\gamma = 1$, if we choose $\sigma_i(t) = \eta_i = \zeta_i^\rho$, then

$$h_i(t) = \frac{N-1}{N} \zeta_i^\rho + \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu \neq \rho}^p \zeta_i^\mu \zeta_j^\mu \zeta_j^\rho + \kappa \zeta_i^\rho, \quad (9)$$

which shows that one may calibrate κ appropriately, allowing the external stimulus to roughly cancel the noise term, in order to favor the signal contribution. This would allow the system to recognize the submitted external pattern. Therefore, κ must not be so small, since it will not be able to cancel the noise term, nor very large, in which case it may dominate the local field, forcing an alignment with the external stimulus, whether correlated with a particular memory or not. Finding an optimal value for κ is a fundamental point of this framework.

Next, we estimate, approximately, the optimal value of κ , by focusing on the noise contribution in Equation (9),

$$h_i^{\text{noise}}(t) = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu \neq \rho}^p \zeta_i^\mu \zeta_j^\mu \zeta_j^\rho. \quad (10)$$

Considering that there are no correlations between patterns μ and ρ , and that patterns associated with different sites ($j \neq i$) are independent, each of these contributions take the values ± 1 with equal probability. For large enough values of N and p the average over the variables $\{\zeta\}$ is,

$$\langle h_i^{\text{noise}}(t) \rangle_{\zeta} \simeq 0, \quad (11)$$

whereas its associated variance yields

$$\langle (h_i^{\text{noise}}(t))^2 \rangle_{\zeta} = \left\langle \frac{1}{N^2} \sum_{j \neq i}^N \sum_{\mu \neq \rho}^p \sum_{k \neq i}^N \sum_{v \neq \rho}^p \zeta_i^\mu \zeta_j^\mu \zeta_j^\rho \zeta_i^v \zeta_k^v \zeta_k^\rho \right\rangle_{\zeta} \simeq \frac{p}{N} = \alpha. \quad (12)$$

Hence, the width of the noise distribution is approximately $\langle (h_i^{\text{noise}}(t))^2 \rangle_{\zeta}^{1/2} \sim \sqrt{\alpha}$, so that for independent patterns, a good choice for κ should roughly be $\sqrt{\alpha}$; the numerical simulations to be presented later on display good agreement with this choice. This procedure could indicate pattern recognition for values of p that are much higher than those restricted by the upper limit of the Hopfield model. However, in those models where analytical calculations for the optimal value of κ are not feasible, it should still be possible to estimate this value numerically; in fact, this is an easy task, as will be shown later on.

From Equation (8), with the reminder that the couplings $\{J_{ij}\}$ are symmetric in the Hopfield model, one can define a Hamiltonian at time t ,

$$\begin{aligned} \mathcal{H}(t) = & - \sum_{i=1}^N h_i(t) \sigma_i(t) = - \frac{1}{N} \sum_{(i,j)} \xi_i^\rho \xi_j^\rho \sigma_i(t) \sigma_j(t) \\ & - \frac{1}{N} \sum_{(i,j)} \sum_{\mu \neq \rho}^p \xi_i^\mu \xi_j^\mu \sigma_i(t) \sigma_j(t) - \kappa \sum_{i=1}^N \eta_i \sigma_i(t), \end{aligned} \tag{13}$$

where $\sum_{(i,j)}$ denotes sums over all distinct pairs of neurons, corresponding to a fully connected NN.

In the simulations to be presented later, we essentially focused on two parameters, namely, the macroscopic superposition (or overlap) of a neuron state $\{\sigma_i(t)\}$ with a given pattern $\{\xi_i^\mu\}$ at time t ,

$$m_\mu(t) = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \sigma_i(t) \quad (\mu = 1, 2, \dots, p), \tag{14}$$

as well as the overlap of this neuron state with the special case where the external-stimuli signs $\{\eta_i\}$ are orthogonal to all stored memories [i.e., case (ii) in Equation (7) with $\gamma = 1/2$ for all memories], defined as

$$m_\perp(t) = \frac{1}{N} \sum_{i=1}^N \eta_i \sigma_i(t). \tag{15}$$

The quantity above represents a macroscopic superposition with an external pattern, not stored and orthogonal to all stored patterns, for which $\eta_i = \pm 1$ with equal probability. Frequently, the set $\{m_\mu(t)\}$ is considered to contain the components of a p -dimensional vector, $\vec{m}(t) = (m_1(t), m_2(t), \dots, m_p(t), \dots, m_p(t))$, and we will give a special emphasis to $m_\rho(t)$, as associated with a single condensed pattern and identified in the first term of the Hamiltonian in Equation (13). All other components $m_\mu(t)$ ($\mu \neq \rho$) appear in the noise contribution of the Hamiltonian.

4. Analytic Calculations

In this section we perform analytical calculations, which essentially correspond to a mean-field approach, along the lines of Refs. [21,22]. We assume that the system defined by the Hamiltonian in Equation (13) attains, after a sufficiently long time, well-defined thermal equilibrium states (for finite temperatures, $T > 0$), together with its zero-temperature limit ($T \rightarrow 0$). For a given realization of the quenched disorder ($\{\xi_i^\mu\}, \{\eta_i\}$), one may define a partition function $\mathcal{Z} \equiv \mathcal{Z}(\{\xi_i^\mu\}, \{\eta_i\})$, so that the free energy per neuron becomes

$$- \beta f = \lim_{N \rightarrow \infty} \frac{1}{N} \left\langle \ln \mathcal{Z}(\{\xi_i^\mu\}, \{\eta_i\}) \right\rangle_{\eta, \xi}, \tag{16}$$

where $\beta = 1/T$ (we work in units $k_B = 1$). Above, $\langle \dots \rangle_{\eta, \xi}$ indicates quenched averages on $\{\eta_i\}$ and $\{\xi_i^\mu\}$, which, according to Equations (3) and (6), should be carried out with the average over the $\{\eta_i\}$ taken before the one over $\{\xi_i^\mu\}$. Following standard procedure, we apply the replica method to calculate the free energy [1–4],

$$- \beta f = \lim_{N \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{Nn} \left(\left\langle \mathcal{Z}^n(\{\xi_i^\mu\}, \{\eta_i\}) \right\rangle_{\eta, \xi} - 1 \right), \tag{17}$$

where $\mathcal{Z}^n(\{\xi_i^\mu\}, \{\eta_i\})$ corresponds to the partition function of n independent replicas, for a given realization of the disorder.

Then, one assumes that the total number of stored patterns is an extensive quantity, expressed as $p = \alpha N$; in the standard Hopfield model, this holds up $\alpha_c \simeq 0.14$ [22], so that

pattern recognition does not work for $\alpha > 0.14$. At the above-mentioned equilibrium, one may perform the usual averages, so the quantity of Equation (14) leads to

$$m_\mu = \left\langle \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \langle \sigma_i \rangle \right\rangle_{\eta, \xi} \quad (\mu = 1, 2, \dots, p), \tag{18}$$

whereas, for the case where the external stimulus is orthogonal to all stored memories, the superposition in Equation (15) yields

$$m_\perp = \left\langle \frac{1}{N} \sum_{i=1}^N \eta_i \langle \sigma_i \rangle \right\rangle_{\eta, \xi}, \tag{19}$$

with $\langle \sigma_i \rangle$ corresponding to local magnetizations.

Lengthy (but well-established) calculations for the free energy of Equation (17) are outlined in the Appendix A, where we have split the patterns $\{\xi_i^\mu\}$ into two sets; namely, s condensed patterns and $(p - s)$ non-condensed ones. From now on, we will investigate the recognition of a single pattern $\mu = \rho$ (i.e., $s = 1$), meaning that the external stimulus presents a superposition with memory ρ , corresponding to case (i) of Equation (6), for which one has the overlapping

$$m_\rho = \left\langle \frac{1}{N} \sum_{i=1}^N \xi_i^\rho \langle \sigma_i \rangle \right\rangle_{\eta, \xi}. \tag{20}$$

Moreover, the free energy at the end of the Appendix A [cf. Equation (A7)] becomes

$$f = \frac{\alpha}{2} + \frac{1}{2} m_\rho^2 + \frac{\alpha\beta}{2} r(1 - q) + \frac{\alpha}{2\beta} \left[\ln(1 - \beta + \beta q) - \frac{\beta q}{1 - \beta + \beta q} \right] - \frac{1}{\beta} \int \mathcal{D}z \langle \ln [2 \cosh \beta(z\sqrt{\alpha r} + m_\rho \xi^\rho + \kappa \eta)] \rangle_{\eta, \xi}, \tag{21}$$

where

$$\int \mathcal{D}z \dots \equiv \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \dots, \tag{22}$$

and the averages $\langle \dots \rangle_{\eta, \xi}$ should be considered over the single condensed pattern ξ^ρ only. Two further parameters appear in the above free energy, namely, the Edwards–Anderson one (which becomes relevant only for very low temperatures),

$$q = \left\langle \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle^2 \right\rangle_{\eta, \xi}, \tag{23}$$

and a parameter that measures the noise produced by $(p - 1)$ non-condensed patterns,

$$r = \frac{1}{\alpha N^2} \sum_{\mu \neq \rho} \left\langle \left(\sum_{i=1}^N \xi_i^\mu \langle \sigma_i \rangle \right)^2 \right\rangle_{\eta, \xi}. \tag{24}$$

In deriving the free energy in Equation (21), as well as the parameters in Equations (23) and (24), we have assumed the replica-symmetry ansatz: in full replica space, both parameters depend on two replica indices, representing matrix elements, as can be seen in the Appendix A. In this ansatz, one assumes that all off-diagonal matrix elements are equal, given by Equations (23) and (24) [1–4,18].

The equilibrium solution comes from the derivatives of the free-energy with respect to the parameters above, leading to the saddle-point equations,

$$m_\rho = \int \mathcal{D}z \langle \xi^\rho \tanh[\beta(z\sqrt{\alpha r} + m_\rho \xi^\rho + \kappa \eta)] \rangle_{\eta, \xi'} \tag{25}$$

$$q = \int \mathcal{D}z \langle \tanh^2[\beta(z\sqrt{\alpha r} + m_\rho \xi^\rho + \kappa \eta)] \rangle_{\eta, \xi'} \tag{26}$$

$$r = \frac{q}{(1 - \beta + \beta q)^2}. \tag{27}$$

For case (ii) in Equation (7), where the external stimulus is orthogonal to all memories, there is no condensed pattern, and the parameter m_\perp of Equation (19) may be calculated from

$$m_\perp = \frac{1}{\beta N} \left\langle \frac{\partial}{\partial \kappa} \ln \mathcal{Z}(\{\xi_i^\mu\}, \{\eta_i\}) \right\rangle_{\eta, \xi} = \frac{1}{\beta N} \frac{\partial}{\partial \kappa} \left\langle \ln \mathcal{Z}(\{\xi_i^\mu\}, \{\eta_i\}) \right\rangle_{\eta, \xi} = -\frac{\partial f}{\partial \kappa}, \tag{28}$$

where we have assumed, as usual, that the derivative commutes with the average operations, since these latter are essentially expressed by sums and integrals. It is important to stress that the averages above should be considered over the orthogonal patterns, for which $m_\rho = 0$ and $\eta_i = \pm 1$ with equal probability. In this way,

$$m_\perp = \int \mathcal{D}z \langle \eta \tanh[\beta(z\sqrt{\alpha r} + \kappa \eta)] \rangle_\eta. \tag{29}$$

Limit $T \rightarrow 0$

Throughout this work, a special emphasis will be given to limit $T \rightarrow 0$ (or $\beta \rightarrow \infty$); in this case, $q = 1 + O(T)$, so that the internal energy per neuron becomes

$$u = \frac{1}{2} \alpha (1 - r) - \frac{1}{2} m_\rho^2 - \kappa m_\perp. \tag{30}$$

Additionally, Equations (25)–(29) lead to

$$m_\rho = \gamma \operatorname{erf}\left(\frac{m_\rho + \kappa}{\sqrt{2\alpha r}}\right) + (1 - \gamma) \operatorname{erf}\left(\frac{m_\rho - \kappa}{\sqrt{2\alpha r}}\right), \tag{31}$$

$$r = \frac{1}{(1 - C)^2}; \quad C = \sqrt{\frac{2}{\pi \alpha r}} \left\{ \gamma \exp\left[-\frac{(m_\rho + \kappa)^2}{2\alpha r}\right] + (1 - \gamma) \exp\left[-\frac{(m_\rho - \kappa)^2}{2\alpha r}\right] \right\}, \tag{32}$$

with the case where the external stimulus is orthogonal to all stored memories yielding

$$m_\perp = \operatorname{erf}\left(\frac{\kappa}{\sqrt{2\alpha r}}\right), \tag{33}$$

where $\operatorname{erf}(x)$ denotes the error function. Notice that in the particular limit $\gamma = 1$, Equations (31) and (32) become

$$m_\rho = \operatorname{erf}\left(\frac{m_\rho + \kappa}{\sqrt{2\alpha r}}\right), \tag{34}$$

and

$$r = \frac{1}{(1 - C)^2}; \quad C = \sqrt{\frac{2}{\pi \alpha r}} \exp\left[-\frac{(m_\rho + \kappa)^2}{2\alpha r}\right], \tag{35}$$

which, for $\kappa = 0$, recover the zero-temperature equations of Ref. [22].

The above zero-temperature equations may be solved numerically for given values of γ , α and κ ; in Figure 1 we illustrate the particular case $\gamma = 1$, with curves for the overlaps m_ρ and m_\perp resulting from the solution of Equations (33)–(35). The overlaps m_ρ and m_\perp versus α (typical values of κ), or versus κ (typical values of α), are presented in Figure 1a,b and Figure 1c,d, respectively. Here, we again emphasize that two different

situations for the external stimulus are being considered, i.e., a fully correlated external stimulus (plots for m_ρ) and an external stimulus orthogonal to all memories (plots for m_\perp). One notices certain ranges of κ and α where multiple solutions for these quantities appear, characterizing first-order phase transitions. Some typical cases are illustrated in the corresponding insets, showing the coexistence of more than one solution; in these cases, the precise locations of the discontinuities may be computed through Maxwell constructions, where one equates the internal energy of Equation (30) for both solutions. For the purpose of pattern recognition, the quantities m_ρ and m_\perp are expected to vary smoothly, so that discontinuities, where one may have large variations on either one of these overlaps, for infinitesimal changes of given parameters, do not correspond to common situations in natural systems. Hence, for the rest of this work we will concentrate on solutions where the parameters m_ρ and m_\perp , associated with two different types of external stimuli, vary smoothly on either one of the parameters κ or α ; more particularly, we will focus on the role played by κ in the non-retrieval region of the standard Hopfield model, i.e., $\alpha > 0.14$, along which one has $m_\rho = m_\perp = 0$ for $\kappa = 0$. We call the attention to those plots for higher values of κ (or α) in Figure 1, focusing on the relevant regions for pattern recognition ($\kappa \simeq \sqrt{\alpha}$), around which one notices significant values for the overlapping m_ρ (typically, $m_\rho > 0.8$).

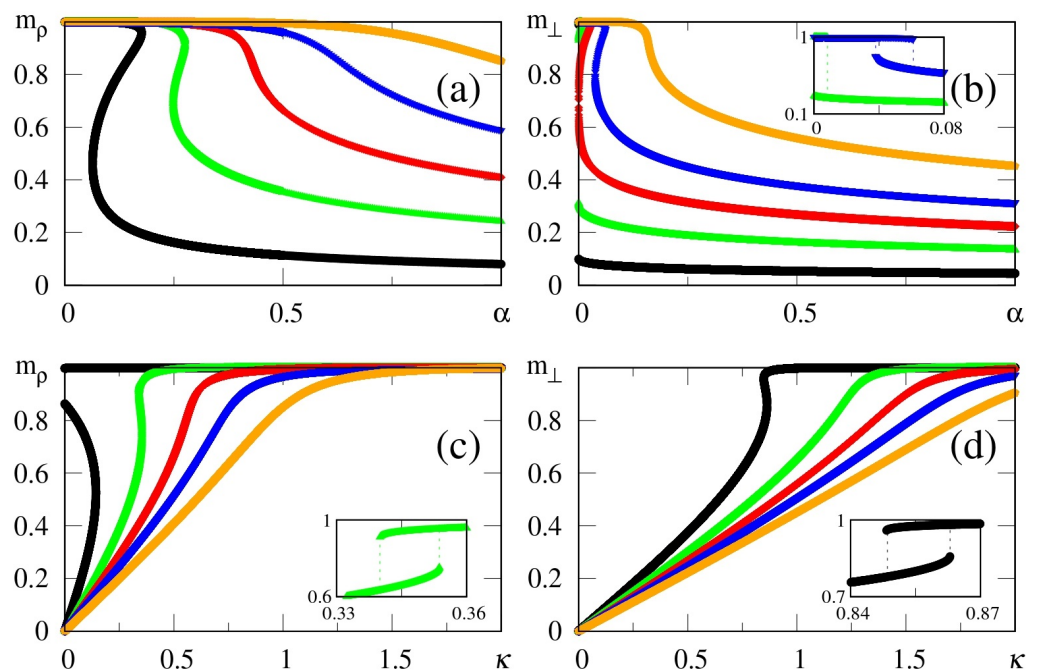


Figure 1. (Color online) Results from mean-field approximation in the zero-temperature limit, for the overlaps m_ρ [Equation (34)] and m_\perp [Equation (33)], are exhibited for the case $\gamma = 1$: m_ρ and m_\perp are shown versus α (typical values of κ) in panels (a,b), and represented versus κ (typical values of α) in panels (c,d), respectively. In (a,b) one has $\kappa = 0.1, 0.3, 0.5, 0.7$, and 1.0 (from bottom to top), whereas in (c,d) one has $\alpha = 0.1, 0.3, 0.5, 0.7$, and 1.0 (from top to bottom). For smaller values of κ and α , one notices multiple solutions for both m_ρ and m_\perp , typical of first-order phase transitions. Insets illustrate some cases where two solutions coexist for certain intervals of these parameters: (b) $\kappa = 0.3$ and 0.7 ; (c) $\alpha = 0.3$; (d) $\alpha = 0.1$.

5. Numerical Simulations: Recognition of Uncorrelated Patterns

We studied the SDNN, defined by the local field $h_i(t)$ of Equations (5)–(8), through zero-temperature numerical simulations, for $N = 10^4$ neurons and $p = \alpha N$ stored memories, considering several choices for the parameters α and γ . At first, we set couplings and external stimuli according to Equations (2), (3) and (6); being quenched variables, these quantities were kept fixed along each time evolution. The initial neuron configuration was

chosen at random, i.e., $\sigma_i(0) = \pm 1$ with equal probability; then, these dynamical variables were updated in a sequential way, through the zero-temperature dynamics,

$$\sigma_i(t+1) = \text{sign}[h_i(t)]; \quad (\forall i), \quad (36)$$

and each time unit corresponded to the operation above on a single neuron. Following this procedure, the nearest local minimum of energy is attained after a sufficiently long time t^* ($t^* \simeq 10^2 N$, i.e., each neuron is visited 10^2 times), with macroscopic quantities presenting slight fluctuations for $t > t^*$. In our analysis, we have emphasized the overlaps of Equations (19) and (20); at a given local minimum of energy, i.e., for $t > t^*$, these quantities can be calculated, and according to Equations (3) and (6), the average over $\{\eta_i\}$ should be taken before the one over $\{\xi_i^H\}$; for these averaging procedures, each simulation was repeated 10^3 times.

Throughout this section, we restrict ourselves to the recognition of uncorrelated patterns; the satisfying performance of the present proposal in the case of correlated patterns will be shown later. In what follows, we will be particularly interested in the role played by κ in the non-retrieval region of the standard Hopfield model, where $m_\rho = m_\perp = 0$ for $\kappa = 0$. The intensity of the external stimulus $\kappa > 0$ yields two distinct situations concerning the external pattern; namely, $1/2 < \gamma \leq 1$ for $m_\rho(\kappa)$ and $\gamma = 1/2$ for $m_\perp(\kappa)$. According to this, sufficiently large values of κ lead to $m_\rho \rightarrow 2\gamma - 1$ and $m_\perp \rightarrow 1$. However, these two particular limits (κ small or large) are not appropriate for pattern recognition; on the other hand, a well-calibrated value of κ , to be called hereafter κ_c , chosen in such a way to cancel (or diminish as much as possible) the effects of the noise contribution in Equation (8), as discussed in Section 3, represents the crucial point of this new framework. Values $\kappa \ll \kappa_c$ will not be sufficient to cancel the noise term, whereas for $\kappa \gg \kappa_c$, the external field will dominate the total local field and there will be no pattern recognition at all; in this later case, the local field will simply reproduce the external stimulus, whether associated with a memory or not. Here, we propose to monitor the absolute value of the difference between the overlaps of Equations (19) and (20), which correspond to two different types of external stimuli,

$$\Delta m = |m_\rho - m_\perp| = \left\langle \frac{1}{N} \sum_{i=1}^N |\xi_i^\rho - \eta_i| \langle \sigma_i \rangle \right\rangle_{\eta, \xi}, \quad (37)$$

in order to identify the optimal value $\kappa = \kappa_c$. One notices that only the orthogonal (i.e., non-stored) memories will contribute to Δm , so that the appropriate choice of κ_c should correspond to its maximum value, $\max(\Delta m)$.

Results from numerical simulations are shown in Figure 2, where the quantities defined in Equations (19), (20) and (37) are plotted versus κ , for $\gamma = 1.0$ and typical values of α . In all cases, κ_c is clearly identified by means of $\max(\Delta m)$, and since the storage capacity p grows with α , the variance of the noise contribution also increases. The analytical approximate result $\kappa_c \simeq \sqrt{\alpha}$ [cf. Equation (12)] shows good agreement with the numerical estimates of Figure 2. As another consequence of this increase, one finds a tendency towards diminishing the difference Δm , as well as the magnitude of the values $\max(\Delta m)$. In general, computing extremely accurate estimates for κ_c is not a central aim in the present scheme, since slight variations around the values of $\max(\Delta m)$ provide equally good results. This is directly related to fluctuations in the noise contribution, shown in Section 3. It should be stressed that all values of α considered in Figure 2 are greater (or much greater) than the critical value $\alpha_c \simeq 0.14$, above which the standard Hopfield NN does not recognize any pattern. Moreover, one can observe the limits $m_\rho = m_\perp = 0$ ($\kappa = 0$), as well as $m_\rho \rightarrow 2\gamma - 1$ and $m_\perp \rightarrow 1$ ($\kappa \gg \kappa_c$).

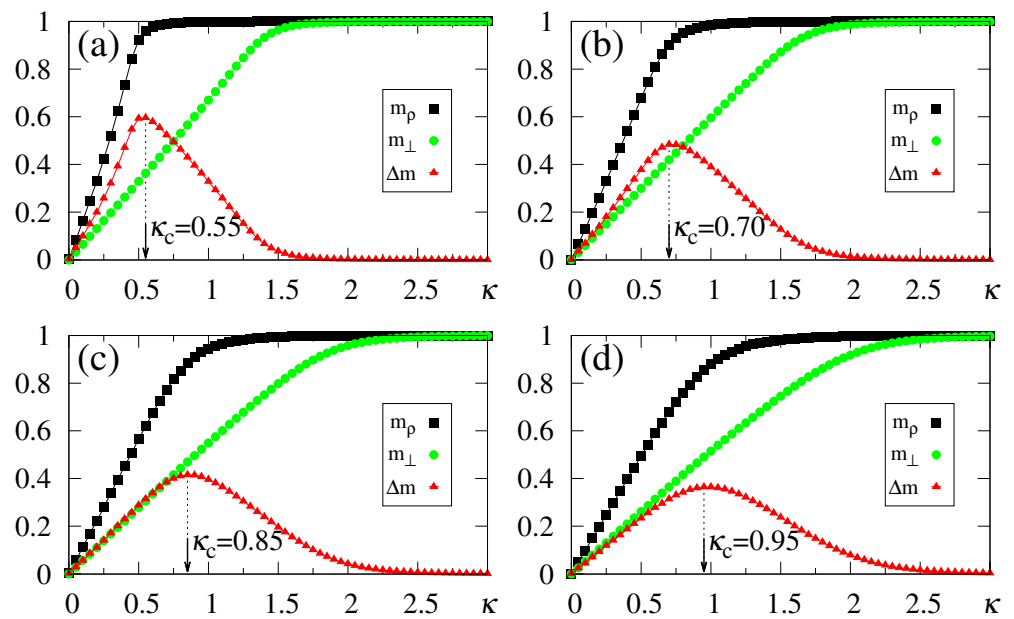


Figure 2. (Color online) Results from numerical simulations for the overlaps m_ρ (black squares) [cf. Equation (20)], m_\perp (green circles) [cf. Equation (19)], and the modulus of their difference, Δm (red triangles) [cf. Equation (37)], are plotted versus κ , for $\gamma = 1.0$, and typical values of α : (a) $\alpha = 0.4$; (b) $\alpha = 0.6$; (c) $\alpha = 0.8$; (d) $\alpha = 1.0$. In each case, the maximum value of Δm yields the best choice for κ , denoted by κ_c . The two overlaps shown represent, respectively, macroscopic superpositions of a neuron state with a stored pattern ρ (m_ρ), and with non-stored patterns orthogonal to the stored one (m_\perp). The lines interpolating the symbols are guides for the eye.

An interesting effect appears in Figure 3, where we present results from numerical simulations for the quantities defined in Equations (19), (20) and (37) for $\gamma = 0.9$, and the same values of α used in Figure 2. For $\kappa < \kappa_c$, all plots are qualitatively similar to those of Figure 2 and, in particular, very close estimates for κ_c (or even equal in some cases) were obtained; one should notice that although the analytical result $\kappa_c \simeq \sqrt{\alpha}$ of Equation (12) was derived for the case $\gamma = 1$, it holds for $\gamma = 0.9$ as a good approximation. Nevertheless, the limits $m_\rho \rightarrow 2\gamma - 1 = 0.8$ and $m_\perp \rightarrow 1$ are fulfilled for $\kappa \gg \kappa_c$, so that their plots cross ($m_\rho = m_\perp$) at some value of κ , yielding $\Delta m = 0$. To the right of this crossing point one has that $m_\perp > m_\rho$, signaling that the external stimulus starts dominating the total local field, as one approaches a non-relevant limit for the SDNN. In this region, pattern recognition begins deteriorating, and for $\kappa \gg \kappa_c$ the local field will simply reproduce the external stimulus, whether associated with a memory or not. The optimal value κ_c , associated with the maximum difference between overlaps m_ρ and m_\perp , represents the appropriate one, allowing the recognition of a stored pattern, for $\alpha > 0.14$; it cannot be too large, otherwise it prevents the external stimulus from dominating the other terms of the local field in Equation (8). This subtle balance is the heart of this approach.

One of the main virtues of the SDNN, illustrated in Figures 2 and 3, concerns the range of α values, above the critical value of the standard Hopfield model, that still allow significant pattern recognition capabilities. Considering, as an illustrative example, the case $\alpha = 1$ [cf. Figures 2d and 3d], one finds $\kappa_c \simeq 0.95$ in both cases $\gamma = 1.0$ and $\gamma = 0.9$, and the macroscopic superposition with a given pattern ρ presents the values $m_\rho \simeq 0.9$ [Figure 2d] and $m_\perp \simeq 0.7$ [Figure 3d]. It should be noted that superpositions with memories $\mu \neq \rho$ are still present, being of order $1/\sqrt{N}$. This enlargement in the pattern recognition capacity is directly related to the fact that the SDNN expresses no basins of attraction, although it creates a huge and single basin of attraction when $\kappa \simeq \kappa_c$. However, by setting $\kappa = 0$, this basin of attraction disappears and the memories do not occupy any volume in phase space,

for any $\alpha > \alpha_c$. As will be shown later, within this framework recognition works well, even for values of $\alpha \gg 1$.

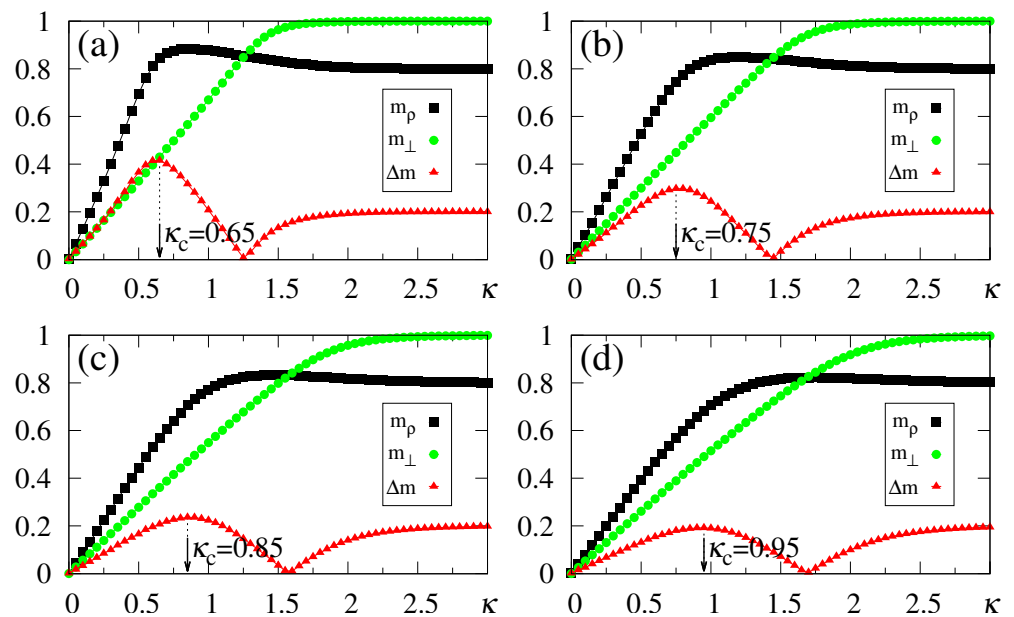


Figure 3. (Color online) Results from numerical simulations for the overlaps m_ρ (black squares) [cf. Equation (20)], m_\perp (green circles) [cf. Equation (19)], and the modulus of their difference, Δm (red triangles) [cf. Equation (37)], are plotted versus κ , for $\gamma = 0.9$, and typical values of α : (a) $\alpha = 0.4$; (b) $\alpha = 0.6$; (c) $\alpha = 0.8$; (d) $\alpha = 1.0$. In each case, the maximum value of Δm yields the best choice for κ , denoted by κ_c . The two overlaps shown represent, respectively, macroscopic superpositions of a neuron state with a stored pattern ρ (m_ρ), and with non-stored patterns orthogonal to the stored one (m_\perp). The lines interpolating the symbols are guides for the eye.

As defined in Equation (6), the parameter γ measures the disturbance associated with recognition of a given stored pattern ρ . Therefore, as γ decreases from the value of one, we should find dissimilarities between the external stimulus and the stored pattern, so that pattern recognition fails for some $\gamma < 1$. We illustrate this aspect in Figure 4, where we present a situation close to the limit of disturbance for the recognition of the stored pattern, $\gamma \simeq 0.74$. One notices that there is not a clear optimal choice for κ in this case, so that one can not distinguish a stored pattern from a new pattern that is not stored and is orthogonal to other stored patterns. This feature of the model was investigated in Ref. [65], where this limit was verified, and it was shown that it is not possible to find an optimal choice of κ for $\gamma \lesssim 0.74$. Hence, in this approach, we can deform the stored pattern close to 25% and the SDNN will still recognize it, even for values of α greater than the critical limit of the Hopfield model. This peculiarity demonstrates an impressive performance compared with other NN models that are studied in the literature.

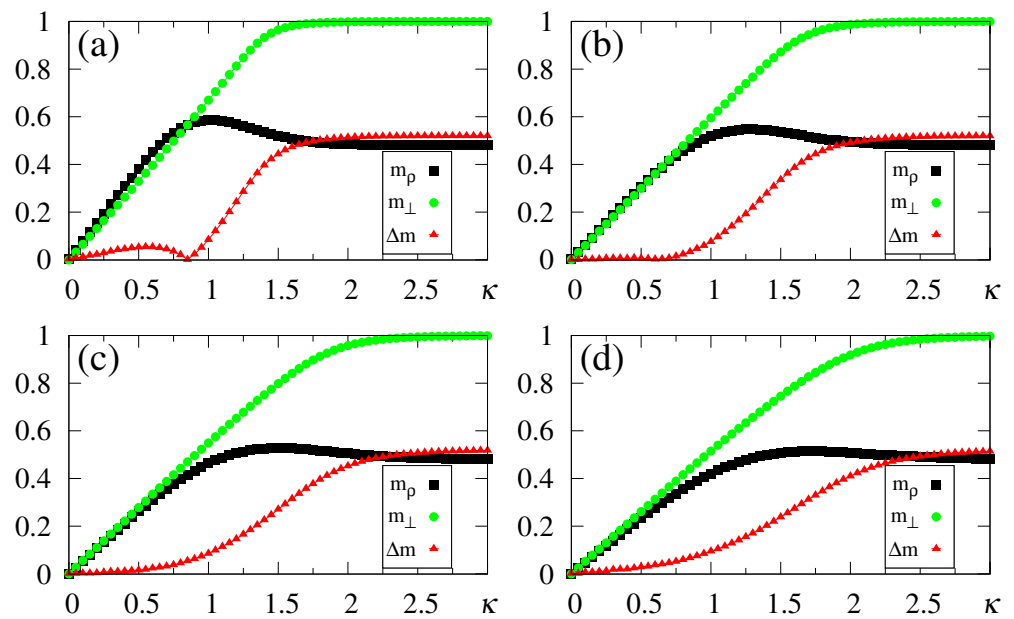


Figure 4. (Color online) Results from numerical simulations for the overlaps m_ρ (black squares) [cf. Equation (20)], m_\perp (green circles) [cf. Equation (19)], and the modulus of their difference, Δm (red triangles) [cf. Equation (37)], are plotted versus κ , for $\gamma = 0.74$, and typical values of α : (a) $\alpha = 0.4$; (b) $\alpha = 0.6$; (c) $\alpha = 0.8$; (d) $\alpha = 1.0$. The two overlaps shown represent, respectively, macroscopic superpositions of a neuron state with a stored pattern ρ (m_ρ), and with non-stored patterns orthogonal to the stored one (m_\perp). The lines interpolating the symbols are guides for the eye.

We have also carried out simulations that consider much higher values of α , with the external pattern corresponding to ρ , i.e., $\gamma = 1$, as shown in Figure 5. One sees that there is no well-defined limit for the storage capacity, showing that the procedure described above holds for considerably large values of α . Particularly, as presented in Figure 5c, this technique was easily implemented for $\alpha = 16$; in this case we acquire $\kappa_c \simeq 3.3$, together with $m_\rho \simeq 0.7$, whereas the difference Δm remains quite significant (revealing a maximum value $\Delta m \simeq 0.1$), and still allowing recognition. One should notice that the value of α used in Figure 5c is more than 10^2 times larger than the threshold of the standard Hopfield model ($\alpha_c \simeq 0.14$) [22]. Analyzing together the results of Figures 1 and 5 (all of them for $\gamma = 1$), two characteristics deserve discussion, as follows. (i) The analytical approximate result (derived for $\gamma = 1$ and $N \rightarrow \infty$), $\kappa_c \simeq \sqrt{\alpha}$ [cf. Equation (12)], shows a good agreement with most numerical estimates (typically within 10% error), although the discrepancies increase for the larger α values. One sees that the analytical calculation yields overestimates for κ_c , with respect to the numerical results. These discrepancies are mostly due to the approximations leading to Equation (12), as well as to the finite sizes used in the simulations. (ii) Considering even larger values for α , a saturation effect occurs, in the sense that $m_\perp \rightarrow m_\rho$ from below, yielding $\Delta m \rightarrow 0$. This characteristic is illustrated in Figure 5d, where we consider $\alpha = 32$, which shows that the curves for m_ρ and m_\perp become very close, leading to a flat behavior in Δm . In this regime, clear identification of κ_c becomes difficult, explaining the increase in the discrepancy between analytical and numerical results. We verified that such a saturation occurs, typically, for $\alpha \gtrsim 40$. Nevertheless, we can state that for relatively high values of α , as illustrated in Figure 5b,c, the SDNN is able to achieve pattern recognition with no difficulties.

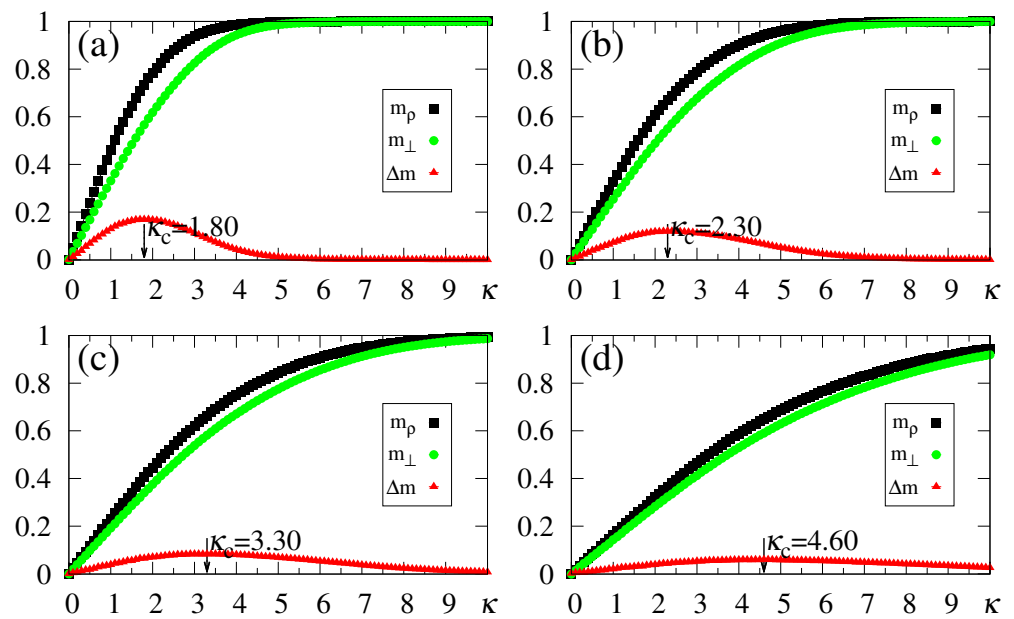


Figure 5. (Color online) Results from numerical simulations for the overlaps m_ρ (black squares) [cf. Equation (20)], m_\perp (green circles) [cf. Equation (19)], and the modulus of their difference, Δm (red triangles) [cf. Equation (37)], are plotted versus κ , for $\gamma = 1.0$, and typical high values of α : (a) $\alpha = 4$; (b) $\alpha = 8$; (c) $\alpha = 16$; (d) $\alpha = 32$. In each case, the maximum value of Δm yields the best choice for κ , denoted by κ_c . The two overlaps shown represent, respectively, macroscopic superpositions of a neuron state with a stored pattern ρ (m_ρ), and with non-stored patterns orthogonal to the stored one (m_\perp). The lines interpolating the symbols are guides for the eye.

In Figure 6, we compare data from numerical simulations of $N = 10^4$ neurons (symbols) with results from analytical calculations in the zero-temperature limit (full lines), considering typical choices for α and γ . The analytical results for m_\perp follow from Equations (33) and (32), whereas those for m_ρ were obtained from Equations (31) and (32); the corresponding numerical estimates were computed from Equations (19) and (20), respectively, with $\Delta m = |m_\rho - m_\perp|$ given by Equation (37). In all cases one notices a good agreement between the two approaches, with small discrepancies between them (at most of the order 0.05), which become more significant for small values of κ , as well as around the best choice κ_c , as expected, signaling the most relevant region for pattern recognition. However, the results from the two approaches essentially coincide when the external stimulus becomes very large, since in this limit the fluctuations in pattern recognition disappear. In Figure 6a, we present results for $\gamma = 0.74$ —i.e., very close to the limit of disturbance for the recognition of a stored pattern, as illustrated in Figure 5—so that there is not a well-defined value for κ_c . However, κ_c may be obtained clearly from Figure 6b–d, and one notices in these cases that the analytical estimates (given by the maximum of Δm) are always overestimates with respect to the numerical ones [even in the cases $\gamma = 0.9$ of Figure 6b,c]. This is in agreement with previous discussions referring to the approximations that led to Equation (12). The above-mentioned discrepancies are essentially due to finite-size effects, since analytical and numerical results are expected to coincide for $N \rightarrow \infty$, in which limit the mean-field approximation becomes exact for a fully connected network [1–4].

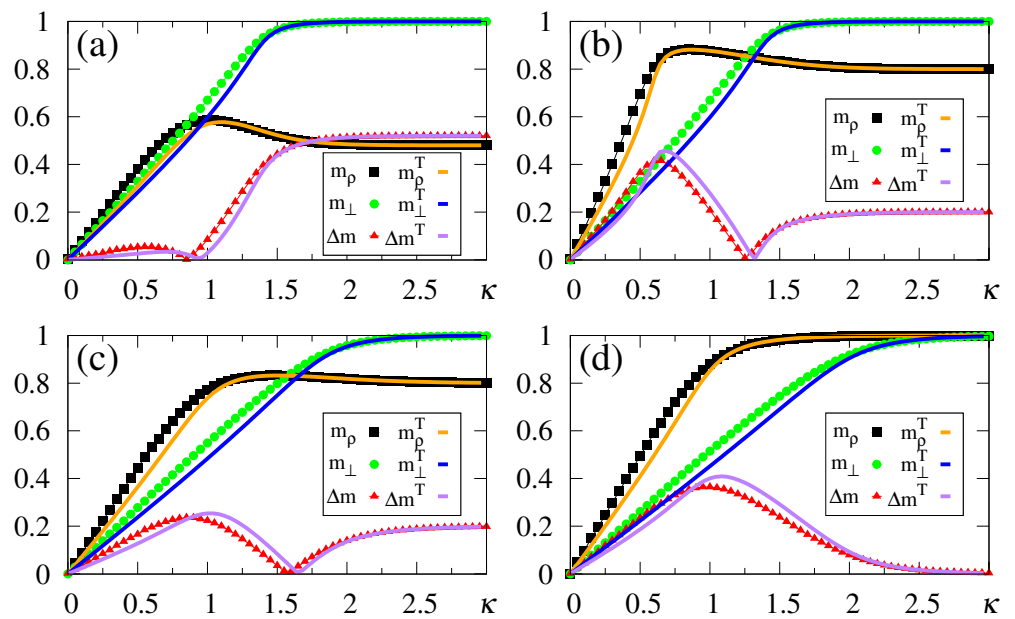


Figure 6. (Color online) Zero-temperature results from analytical calculations (full lines) are plotted versus κ and compared with those from numerical simulations (symbols) for typical values of α and γ : (a) $\alpha = 0.4$ and $\gamma = 0.74$; (b) $\alpha = 0.4$ and $\gamma = 0.9$; (c) $\alpha = 0.8$ and $\gamma = 0.9$; (d) $\alpha = 1.0$ and $\gamma = 1.0$. To distinguish from the numerical data, in each case the upper index T stands for theoretical results. The theoretical overlaps of Equation (33) (blue full line) and Equation (31) (brown full line) are compared with the numerical estimates for m_{\perp} (green circles) [cf. Equation (19)] and m_{ρ} (black squares) [cf. Equation (20)], respectively, whereas the theoretical modulus of the difference between these quantities (purple full line) is compared with the numerical Δm (red triangles) [cf. Equation (37)].

6. Time Evolution of the Macroscopic Superposition under Changes in the External Pattern

One well-known characteristic of living beings, fundamental to their survival, consists in their ability to react to changes in the environment. In this section, we show how the SDNN adjusts itself, being capable of replacing a previously recognized pattern with a new one, according to changes in the external stimulus. Let us consider a typical real situation where at a time t_0 we present a pattern associated with one stored memory, say memory ρ ; then, at a later time t_1 , we withdraw it and present a different pattern associated with memory ν . In order to illustrate the effectiveness of the SDNN in responding to these changes, we now allow the set of external-stimulus signs to change in time, i.e., $\{\eta_i\} \equiv \{\eta_i(t)\}$. Furthermore, we assume abrupt changes in these variables at two given times, $t = t_0$ and $t = t_1$, so that the probability distribution in Equation (6) is replaced by the time-dependent set

$$P(\eta_i(t)) = \gamma_1 \delta(\eta_i(t) - \zeta_i^{\rho}) + (1 - \gamma_1) \delta(\eta_i(t) + \zeta_i^{\rho}); \quad (t_0 \leq t \leq t_1), \tag{38}$$

and

$$P(\eta_i(t)) = \gamma_2 \delta(\eta_i(t) - \zeta_i^{\nu}) + (1 - \gamma_2) \delta(\eta_i(t) + \zeta_i^{\nu}); \quad (t > t_1), \tag{39}$$

with $1/2 \leq \gamma_1, \gamma_2 \leq 1$. The actions at times t_0 and t_1 correspond, respectively, to the presentation of pattern ρ , and to its replacement by the new pattern ν , giving rise to the question of whether the SDNN will respond adequately to these modifications.

In Figure 7, we illustrate how the SDNN reacts to the above-mentioned changes, by plotting the time-dependent macroscopic superpositions [cf. Equation (14)] $m_{\rho}(t)$ (black circles) and $m_{\nu}(t)$ (dashed green line) versus time, for $\alpha = 0.8$, while considering four different intensities of the external stimulus, in increasing order, as shown in Figure 7a–d. Up to $t_0 = 5.0 \times 10^4$, no external pattern is presented ($\kappa = 0$), and no memory is recognized, as expected. Then, a pattern with 80% ($\gamma_1 = 0.8$) superposition with the stored pattern $\{\zeta_i^{\rho}\}$

is presented, leading to the onset of the macroscopic superposition $m_\rho(t)$ ($0 \leq m_\rho(t) \leq 2\gamma_1 - 1$). Similarly, up to $t_1 = 10^5$, the macroscopic superposition $m_v(t)$ remains zero. It becomes nonzero only for $t > t_1$, due to the abrupt change in the stimulus at $t = t_1$, where pattern ρ is replaced by the new pattern, presenting a superposition $\gamma_2 = 1.0$ with stored pattern $\{\zeta^v\}$. Accordingly, Figure 7 presents another important attribute of the SDNN—that is, its reactions to this change at $t = t_1$, resulting in a decrease in $m_\rho(t)$, together with a growth in $m_v(t)$ ($0 \leq m_v(t) \leq 2\gamma_2 - 1$) for $t > t_1$.

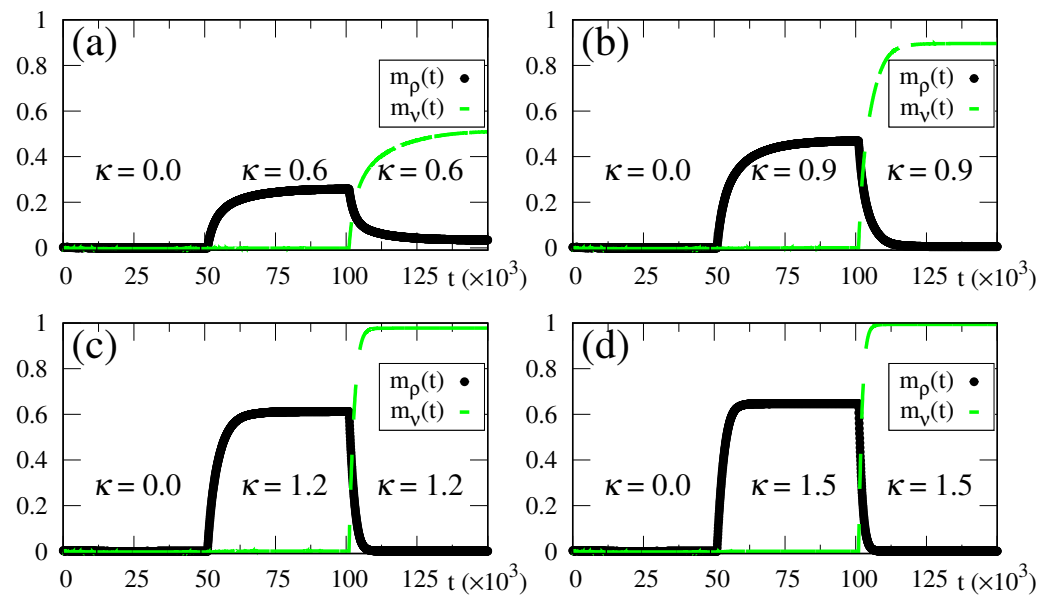


Figure 7. (Color online) Results from numerical simulations showing the time evolution of two different macroscopic superpositions of states of neurons with memories ρ and v [$m_\rho(t)$ and $m_v(t)$], defined according to Equation (14), for $\alpha = 0.8$. The external stimulus changes in time, following Equations (38) and (39), where the two relevant times are $t_0 = 5.0 \times 10^4$ and $t_1 = 10^5$. Hence, for $t < t_0$ there is no stimulus, i.e., $\kappa = 0.0$, and at $t = t_0$ a stimulus with an intensity κ appears with a superposition $\gamma_1 = 0.8$ with stored pattern $\{\zeta^\rho\}$; then, at $t = t_1$ it changes abruptly, presenting a superposition $\gamma_2 = 1.0$ with stored pattern $\{\zeta^v\}$. Four different intensities of the external stimulus are considered for $t > t_0$: (a) $\kappa = 0.6$; (b) $\kappa = 0.9$; (c) $\kappa = 1.2$; (d) $\kappa = 1.5$. The macroscopic superposition $m_v(t)$ (dashed green line) remains zero for $t < t_1$, whereas $m_\rho(t)$ (black circles) becomes nonzero for $t_0 < t < t_1$.

Some interesting effects are shown in Figure 7, as discussed next. (i) The overlaps $m_\rho(t)$ and $m_v(t)$ approach a plateau whose height increases for larger values of κ , in agreement with the results presented in Figures 2–6. (ii) The response time to the changes considered, either in the growth or reduction in the overlaps, is directly related to the intensity of the external stimulus κ : larger (smaller) values of κ lead to smaller (larger) response times. This type of behavior is also in agreement with reactions of living beings. In this way, one notices that for $\kappa = 0.6$ the long-time limits for the overlaps have not been fully attained, as one can see by comparing the plot of $m_v(t)$ of Figure 7a with the long-time result of Figure 2c at $\kappa = 0.6$ (both for $\alpha = 0.8$ and $\gamma = 1.0$). One should note that the time interval for the action of the external stimulus in all panels of Figure 7 is 5.0×10^4 (for both patterns ρ and v), which is smaller than the time used in the simulations of Figures 2–6, $t^* \simeq 10^2 N$ (i.e., $t^* \simeq 10^6$), and thus sufficient for the SDNN to reach its nearest local minimum of energy. Due to this, for the smaller values of κ [cf., e.g., Figure 7a], both overlaps $m_\rho(t)$ and $m_v(t)$ have not yet reached their long-time limits. However, for larger values of κ , much smaller times are required for the overlaps $m_\rho(t)$ and $m_v(t)$ to approach their long-time limits; particularly, in Figure 7d, one notices that these overlaps have attained their maximum values; namely, $m_\rho(t) \rightarrow 2\gamma_1 - 1$ and $m_v(t) \rightarrow 2\gamma_2 - 1$. (iii) All panels of Figure 7 are

illustrative, and exhibit the ability of the SDNN to react under changes in the environment. However, the most relevant interval for pattern recognition should be close to $\kappa = 0.9$, as considered in Figure 7b, according to the values of κ_c estimated in Figures 2c and 3c.

One should note that reactions can also be studied in ANNs, as in the standard Hopfield model for $\alpha < \alpha_c$. In these cases, this may be achieved by changing the initial state $\{\sigma_i(0)\}$, which may lead to jumps in phase space among different basins of attraction. The results presented above show, clearly, that the SDNN modifies the recognized pattern, in a smooth way, according to changes in the external pattern, which is a common characteristic of living beings. In many real situations, living beings need to react to any new external pattern presented; this model behaves precisely in this way.

7. Recognition of Correlated Patterns

It is also interesting to investigate the performance of this framework when correlated patterns are stored. Let us suppose an external stimulus that presents a maximum overlap with one particular stored memory, let us say memory ρ , i.e., $\gamma = 1$ in Equation (6). Then, we consider another pattern, θ , that presents a correlation b ($0 \leq b \leq 1$) with pattern ρ ; this means that $\xi_i^\rho = \xi_i^\theta$ for a fraction b of indices i , leading to

$$\frac{1}{N} \sum_{i=1}^N \xi_i^\rho \xi_i^\theta = b - (1 - b) = 2b - 1 \quad (\rho \neq \theta). \quad (40)$$

Above, the limits $b = 0$ and $b = 1$ correspond to anti-correlated and fully correlated patterns, respectively, whereas the most interesting situations occur for $0 < b < 1$.

In Figure 8, we present results from the simulations of two patterns, $\{\xi_i^\rho\}$ and $\{\xi_i^\theta\}$, with a correlation parameter $b = 0.8$ between them; all other patterns are uncorrelated among themselves, as well as with these two. One may notice that the macroscopic superpositions, m_ρ and m_θ , attain the expected saturation limits for sufficiently large values of κ , i.e., $m_\rho \rightarrow 1$ (due to its maximum overlap with the external stimulus) and $m_\theta \rightarrow 0.6$ [following cf. Equation (40)]. Similarly to the previous situations investigated with uncorrelated patterns, the optimal value κ_c increases for increasing values of α . Around their corresponding optimal values κ_c , the system recognizes both patterns, with $m_\rho \gtrsim 0.8$ and m_θ already close to its saturation limit, $m_\theta \simeq 0.6$, for all values of α considered, whereas all memories $\mu \neq \rho, \theta$ present overlaps that are essentially zero (in fact, of order $1/\sqrt{N}$). It is important to stress that even for $\alpha = 1.0$ one finds recognition of both stored patterns, indicating that pattern θ presents a significant overlap with the external stimulus, as a direct consequence of its correlation with the stored pattern ρ , demonstrating an appropriate feature of the SDNN model when dealing with correlated patterns. As another interesting result, one should call attention to the reduction of the estimates of κ_c , as a consequence of correlations; this aspect is revealed by comparing Figure 2d ($\kappa_c = 0.95$) and Figure 8d ($\kappa_c = 0.70$), both for $\gamma = 1$ and $\alpha = 1$, for which one typically notices a 25% decrease on the value of κ_c .

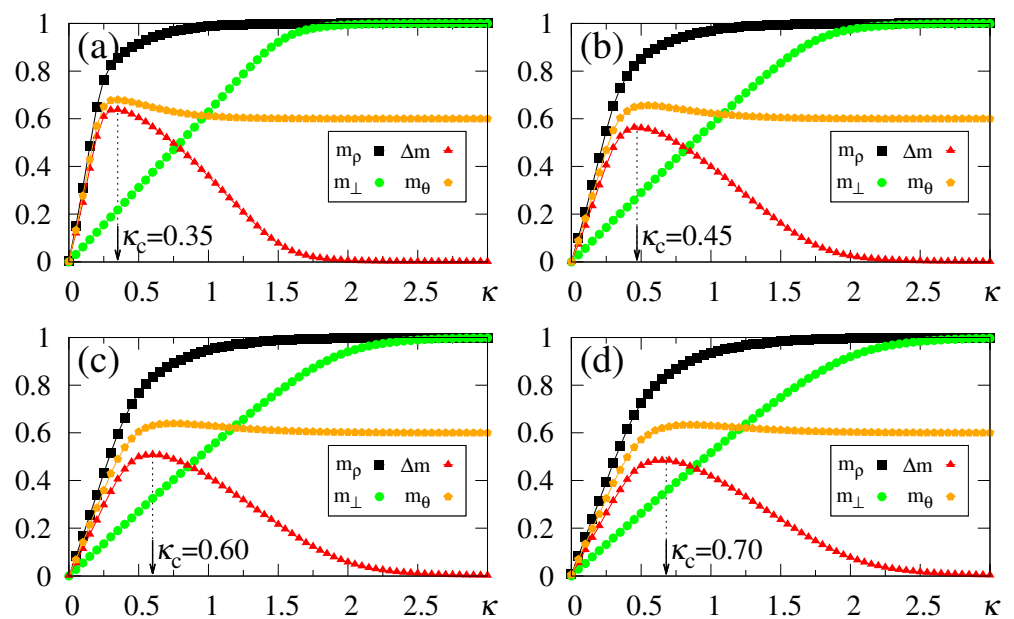


Figure 8. (Color online) Results from numerical simulations of two correlated patterns ($\{\xi_i^\rho\}$ and $\{\xi_i^\theta\}$), with between them present a correlation parameter $b = 0.8$ [cf. Equation (40)], are plotted versus κ . The pattern ρ is fully correlated with the external stimulus ($\gamma = 1$), and data for typical values of α are shown: (a) $\alpha = 0.5$; (b) $\alpha = 0.7$; (c) $\alpha = 0.9$; (d) $\alpha = 1.0$. The macroscopic superpositions with memories ρ and θ (m_ρ and m_θ) are represented by black squares and brown pentagons, respectively; the green circles are data for the macroscopic superposition m_\perp [cf. Equation (19)], computed with a different external stimulus, orthogonal to all stored memories, whereas the red triangles stand for the modulus of the difference, $\Delta m = |m_\rho - m_\perp|$. In each case, the maximum value of Δm yields the best choice for κ , denoted by κ_c . The lines interpolating the symbols are guides for the eye.

Results from the numerical simulations of several patterns—more precisely, $\{\xi_i^\rho\}$, and three other patterns correlated with $\{\xi_i^\rho\}$ —are exhibited in Figure 9. The remaining $(p - 4)$ patterns are uncorrelated with these four and among themselves. Pattern $\{\xi_i^\rho\}$ presents a maximum overlap with the external stimulus ($\gamma = 1$), whereas the three correlated ones follow Equation (40) with $b = 0.7, 0.8$, and 0.9 . All macroscopic superpositions attain the expected saturation limits for sufficiently large values of κ , i.e., $m_\rho \rightarrow 1$, whereas the three correlated patterns approach their corresponding $(2b - 1)$ values, according to Equation (40). Notice that $\kappa_c \simeq 0.2$ is sufficient to recover all four patterns, with $m_\rho \gtrsim 0.9$ and the other three macroscopic superpositions already very close to their saturation limits, for all values of α considered; such significant values appear as direct consequences of the correlations with pattern $\{\xi_i^\rho\}$. A curious aspect of Figure 9 concerns the fact that the optimal value κ_c increases much slower with α , when compared with previous cases investigated, for both uncorrelated patterns (see, e.g., Figures 2 and 3) and two correlated patterns (cf. Figure 8). Besides small variations in κ_c , the magnitude of κ_c is diminished by increasing the number of correlated patterns, as can be seen by comparing, e.g., Figure 2d ($\kappa_c = 0.95$), Figure 8d ($\kappa_c = 0.70$), and Figure 9d ($\kappa_c = 0.21$), all of them for $\gamma = 1$ and $\alpha = 1$. The present simulations suggest that, for a sufficiently large number of correlated patterns, κ_c should converge to a small finite value $\kappa_c \gtrsim 0$, for α fixed; by increasing α , a slow increase in the values of κ_c should occur. This later result is in agreement with the recognition of similar patterns that is performed by living beings, where once one constituent of a given group is recognized, all similar members of the group are also recognized immediately, requiring little external stimulus for this task. The results of Figures 8 and 9 illustrate additional important features of the SDNN model, presenting characteristics very similar to those of living beings in the recognition of correlated patterns.

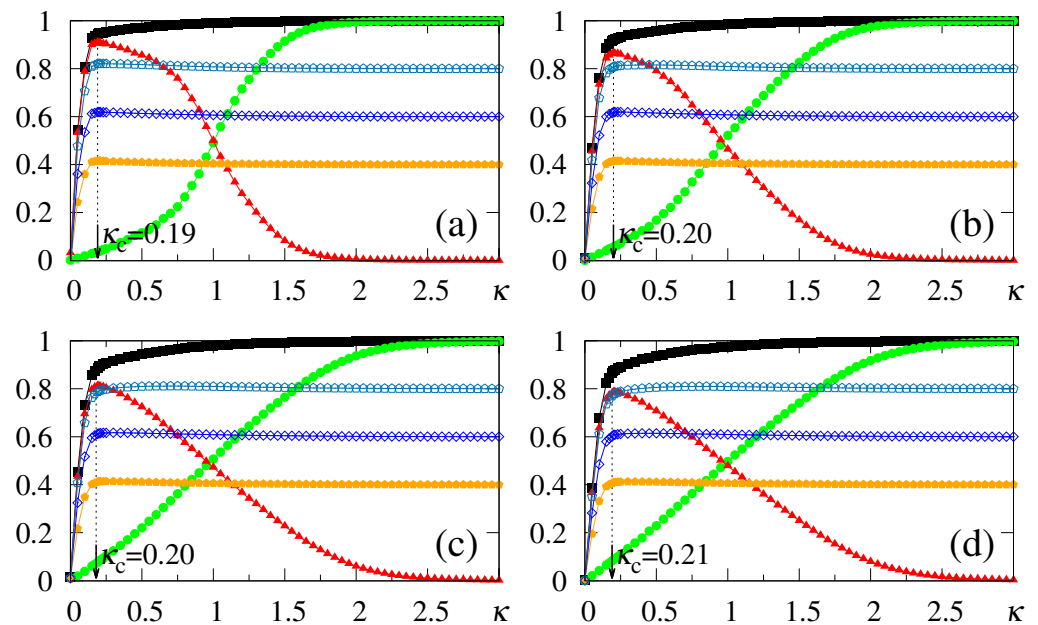


Figure 9. (Color online) Results from numerical simulations for three patterns, correlated with pattern $\{\xi_i^0\}$, are plotted versus κ . The pattern ρ is fully correlated with the external stimulus ($\gamma = 1$) and data for typical values of α are shown: (a) $\alpha = 0.5$; (b) $\alpha = 0.7$; (c) $\alpha = 0.9$; (d) $\alpha = 1.0$. The macroscopic superposition with memory ρ , m_ρ , is represented by black squares, whereas those for the three correlated patterns, for values $b = 0.7, 0.8$, and 0.9 , with respect to pattern $\{\xi_i^0\}$ [cf. Equation (40)], are depicted by brown pentagons, open blue diamonds, and open green pentagons, respectively (increasing values of b from bottom to top). The green circles are data for the macroscopic superposition m_\perp [cf. Equation (19)], computed with a different external stimulus, orthogonal to all stored memories, whereas the red triangles stand for the modulus of the difference, $\Delta m = |m_\rho - m_\perp|$. In each case, the maximum value of Δm yields the best choice for κ , denoted by κ_c ; one notices that κ_c still increases with α , although much slower than in Figure 8. The lines interpolating the symbols are guides for the eye.

We have also tested how the SDNN behaves by considering several correlated patterns in the regime of large α values, as shown in Figure 10. Similarly to the experiments for Figures 8 and 9, the estimates of κ_c decrease for correlated patterns. This effect may be verified by comparing the results of Figure 10 (four correlated patterns) with those of Figure 5 (uncorrelated patterns), e.g., in the case $\alpha = 4$ [Figure 10b ($\kappa_c = 1.0$) and Figure 5a ($\kappa_c = 1.8$)], as well as $\alpha = 16$ [Figure 10d ($\kappa_c = 2.6$) and Figure 5c ($\kappa_c = 3.3$)], all for $\gamma = 1$. This aspect is directly related to the gaps between the curves for m_ρ and m_\perp , as can be seen from the corresponding above-mentioned plots. In these cases, one notices that the maximum value of Δm is typically doubled in Figure 10b,d, when contrasted to those of Figure 5a,c, respectively, making it easier to compute the values of κ_c . Such an increase in the maximum of Δm , together with the reduction in the values of κ_c , should yield an enlargement in the total number of memories, since the range of α values for pattern recognition is expanded; these results indicate that the saturation effect observed in Figure 5 should occur for even larger values of α , when one introduces correlations among patterns. Consequently, the introduction of such an important ingredient for real systems, i.e., correlations among patterns, is expected to improve, even further, the storage capacity of the SDNN.

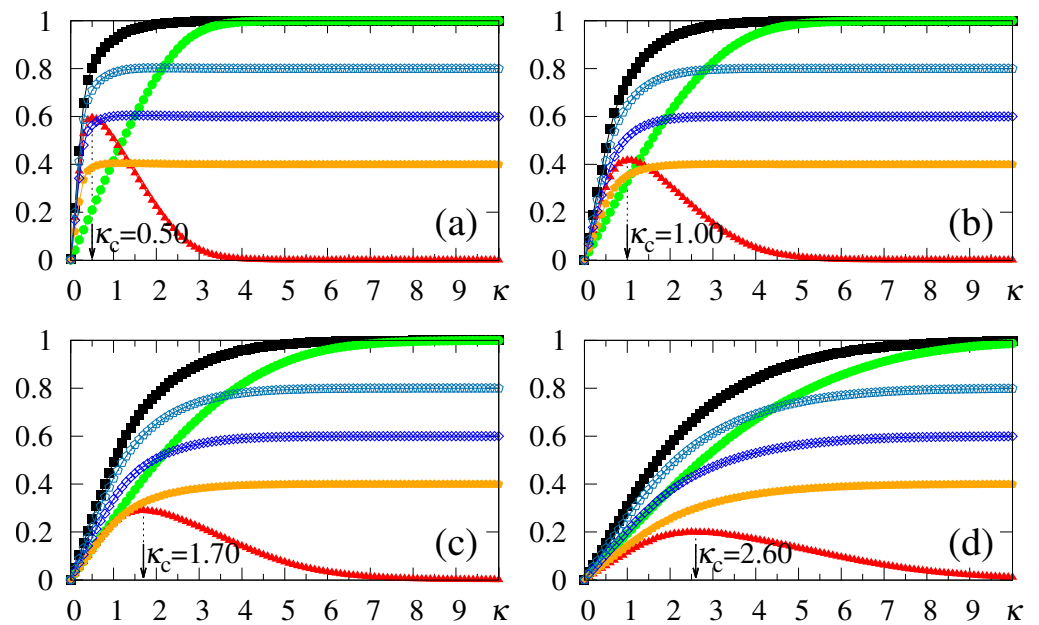


Figure 10. (Color online) Results from numerical simulations for three patterns, correlated with pattern $\{\zeta_i^\rho\}$, are plotted versus κ . The pattern ρ is fully correlated with the external stimulus ($\gamma = 1$) and data for typical values of α are shown: (a) $\alpha = 2$; (b) $\alpha = 4$; (c) $\alpha = 8$; (d) $\alpha = 16$. The macroscopic superposition with memory ρ , m_ρ , is represented by black squares, whereas those for the three correlated patterns, for values $b = 0.7, 0.8$, and 0.9 , with respect to pattern $\{\zeta_i^\rho\}$ [cf. Equation (40)], are depicted by brown pentagons, open blue diamonds, and open green pentagons, respectively (increasing values of b from bottom to top). The green circles are data for the macroscopic superposition m_\perp [cf. Equation (19)], computed with a different external stimulus, orthogonal to all stored memories, whereas the red triangles stand for the modulus of the difference, $\Delta m = |m_\rho - m_\perp|$. In each case, the maximum value of Δm yields the best choice for κ , denoted by κ_c ; one notices that κ_c still increases with α . The lines interpolating the symbols are guides for the eye.

8. Pattern Recognition in a Diluted Neural Network

In this section we show that the SDNN also performs well for diluted models; we illustrate this by analyzing its performance on the Hopfield model with non-symmetric synapse dilution. Let us then consider synaptic couplings in the modified form [52,53],

$$J_{ij} = \frac{C_{ij}}{N(1-d)} \sum_{\mu=1}^p \zeta_i^\mu \zeta_j^\mu, \tag{41}$$

where $\{C_{ij}\}$ are independent random variables following the probability distribution,

$$P(C_{ij}) = (1-d) \delta(C_{ij} - 1) + d \delta(C_{ij}) \quad (C_{ij} \neq C_{ji}). \tag{42}$$

The variables $\{C_{ij}\}$ represent the asymmetric dilution of the couplings and $d \in [0, 1[$ is a dilution parameter, defined as the fraction of the total number of connections that have been eliminated, i.e., a macroscopic dilution. Notice that the interactions J_{ij} are not symmetric now, since for each neuron pair $[\sigma_i(t)$ and $\sigma_j(t)]$, C_{ij} and C_{ji} are independent variables with $C_{ij} \neq C_{ji}$; consequently, no Hamiltonian can be defined. The two extremum values for the parameter d correspond to $d = 0$ (undiluted limit), whereas its maximum value comes from Equation (41), leading to $(1-d) \sim 1/N$.

Next, to understand better the effects on κ_c due to the dilution of synapses, let us estimate its value within the molecular-field approach; the procedure is similar to the one carried out in Section 3, through an analysis of the noise contribution. Therefore, the local field on neuron i is given by

$$h_i(t) = \sum_{j \neq i}^N J_{ij} \sigma_j(t) + \kappa \eta_i, \tag{43}$$

with $\{J_{ij}\}$ now given by Equation (41). Then, we assume that at a time t the neuron states $\{\sigma_i(t)\}$, as well as the external stimulus $\{\eta_i\}$, coincide with pattern $\{\xi_i^\rho\}$, leading to,

$$\begin{aligned} h_i(t) &= \sum_{j \neq i}^N \sum_{\mu=1}^p \frac{C_{ij}}{N(1-d)} \xi_i^\mu \xi_j^\mu \xi_j^\rho + \kappa \xi_i^\rho, \\ &= \sum_{j \neq i}^N \frac{C_{ij}}{N(1-d)} \xi_i^\rho + \sum_{j \neq i}^N \sum_{\mu \neq \rho}^p \frac{C_{ij}}{N(1-d)} \xi_i^\mu \xi_j^\mu \xi_j^\rho + \kappa \xi_i^\rho. \end{aligned} \tag{44}$$

As before, the sum over memories μ was split into two contributions; namely, a signal produced by pattern $\mu = \rho$, and a noise due to memories $\mu \neq \rho$,

$$h_i^{\text{noise}}(t) = \sum_{j \neq i}^N \sum_{\mu \neq \rho}^p \frac{C_{ij}}{N(1-d)} \xi_i^\mu \xi_j^\mu \xi_j^\rho. \tag{45}$$

Assuming that there are no correlations between patterns μ and ρ , and that patterns associated with different sites ($j \neq i$) are independent, each memory contribution takes the values ± 1 with equal probability. Noting that $\{C_{ij}\}$ are also independent variables, the noise random variables $\{h_i^{\text{noise}}(t)\}$ do not yield a Gaussian for a finite N , but they should become Gaussian distributed for $N \rightarrow \infty$. Hence, considering large enough values of N and p , one acquires for the average over the variables $\{\xi\}$,

$$\langle h_i^{\text{noise}}(t) \rangle_{\xi} \simeq 0, \tag{46}$$

whereas

$$(h_i^{\text{noise}}(t))^2 = \frac{1}{[N(1-d)]^2} \sum_{j \neq i}^N \sum_{\mu \neq \rho}^p \sum_{k \neq i}^N \sum_{v \neq \rho}^p C_{ij} C_{ik} \xi_i^\mu \xi_j^\mu \xi_j^\rho \xi_i^v \xi_k^v \xi_k^\rho. \tag{47}$$

Since $\langle C_{ij}^2 \rangle = \langle C_{ij} \rangle = 1 - d$, the associated variance with respect to the variables $\{\xi\}$ becomes

$$\langle (h_i^{\text{noise}}(t))^2 \rangle_{\xi} \simeq \frac{(N-1)(p-1)}{N^2(1-d)} \simeq \frac{p}{N(1-d)} = \frac{\alpha}{1-d}. \tag{48}$$

Similarly to the discussion carried out in Section 3, one should choose κ_c equal to the width of the noise distribution,

$$\kappa_c = \langle (h_i^{\text{noise}}(t))^2 \rangle_{\xi}^{1/2} \simeq \sqrt{\frac{\alpha}{1-d}} \tag{49}$$

in order to cancel (approximately) the noise contribution. According to this, a dilution parameter $0 < d < 1$ yields an increase in the analytical approximate value for κ_c , as compared to the undiluted case $d = 0$; this effect is in agreement with numerical simulations, as will be shown below. The limit $d \rightarrow 1 - N^{-1}$, i.e., a fully diluted system, is not relevant for living beings and will not be investigated here; in this case one should have $\kappa_c \rightarrow \infty$.

In what follows, for simplicity, we illustrate the effects of synapse dilution for uncorrelated patterns, considering $\gamma = 1.0$ in Equation (6), i.e., an external pattern coinciding precisely with the stored one. Results from numerical simulations are shown in Figures 11 and 12, with the quantities defined in Equations (19), (20) and (37) being plotted versus κ , for decreasing values of the dilution parameter d , and two typical choices of α , namely, $\alpha = 0.5$ [Figure 11] and $\alpha = 1.0$ [Figure 12]. One notices that the limits $m_\rho = m_\perp = 0$ ($\kappa = 0$) still hold; furthermore, for sufficiently high values of κ , the saturation

limits $m_\rho \rightarrow 2\gamma - 1$ and $m_\perp \rightarrow 1$ should be approached, even for high dilutions. One important aspect in these figures concerns the fact that κ_c is clearly identified by means of $\max(\Delta m)$, even for the largest value of d considered [$d = 0.7$ in Figures 11a and 12a]. Moreover, for fixed choices of the dilution parameter d , κ_c increases with α , as already observed for the undiluted case, whereas for fixed α , one notices that κ_c decreases by decreasing d ; these results are in agreement with the theoretical estimate of Equation (49). However, the discrepancies between the numerical and analytical estimates for κ_c are larger for $0 < d < 1$, when compared to those for the undiluted limit ($d = 0$); whereas the relative discrepancies in the undiluted limit are typically of the order 10%, for larger dilutions, e.g., $d = 0.7$, one may have relative discrepancies slightly larger than 20%. Similarly to the large α limit of the undiluted case [see, e.g., Figure 5] the curves for Δm become flatter for high dilutions, leading to such larger discrepancies. Furthermore, as before, the analytical result of Equation (49) represents overestimates with respect to the numerical ones, for all cases shown in Figures 11 and 12. Increasing values of κ_c with dilution is qualitatively expected, since synapse dilution diminishes the possible paths for transmitting given information, and so, larger intensities of the external field should be necessary in the pattern-recognition process.

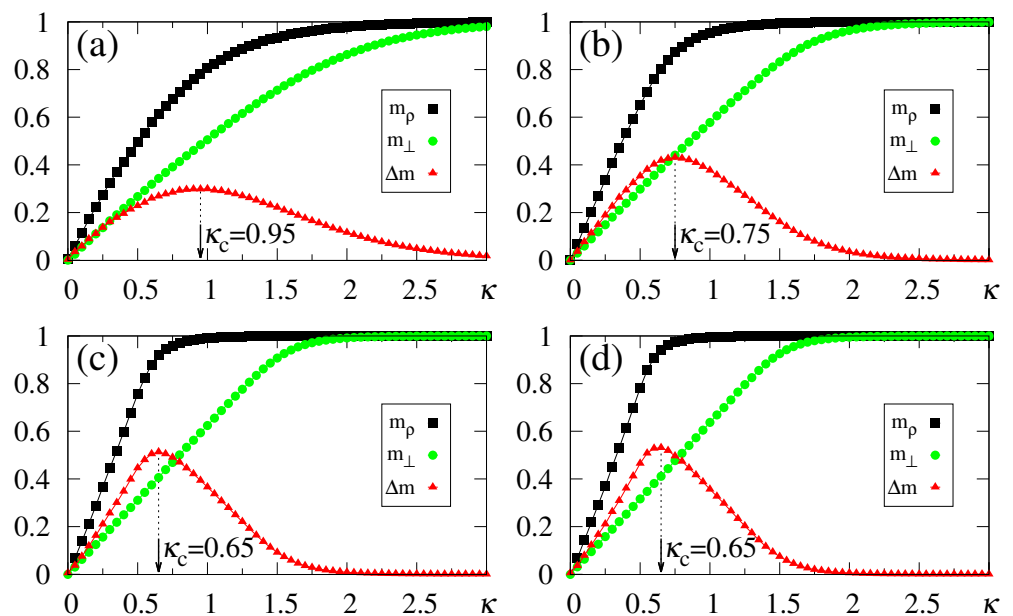


Figure 11. (Color online) Results from numerical simulations for the overlaps m_ρ (black squares) [cf. Equation (20)], m_\perp (green circles) [cf. Equation (19)], and the modulus of their difference, Δm (red triangles) [cf. Equation (37)], are plotted versus κ , for $\gamma = 1.0$, $\alpha = 0.5$, and typical values of the dilution parameter d : (a) $d = 0.7$; (b) $d = 0.4$; (c) $d = 0.1$; (d) $d = 0.0$. In each case, the maximum value of Δm yields the best choice for κ , denoted by κ_c . The two overlaps shown represent, respectively, macroscopic superpositions of a neuron state with a stored pattern ρ (m_ρ), and with non-stored patterns orthogonal to the stored one (m_\perp). The lines interpolating the symbols are guides for the eye.

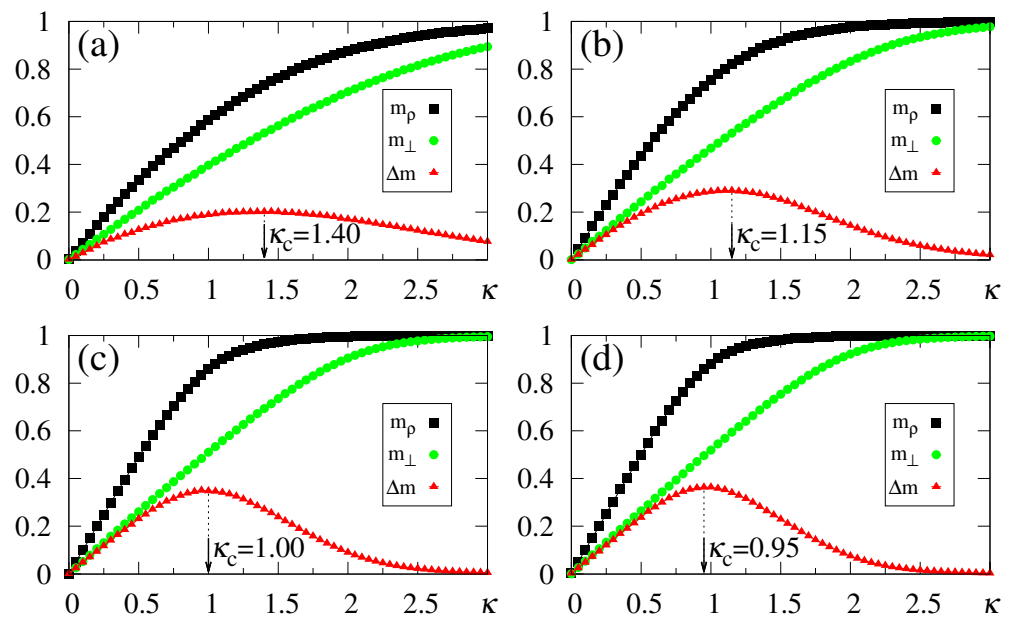


Figure 12. (Color online) Results from numerical simulations for the overlaps m_ρ (black squares) [cf. Equation (20)], m_\perp (green circles) [cf. Equation (19)], and the modulus of their difference, Δm (red triangles) [cf. Equation (37)], are plotted versus κ , for $\gamma = 1.0$, $\alpha = 1.0$, and typical values of the dilution parameter d : (a) $d = 0.7$; (b) $d = 0.4$; (c) $d = 0.1$; (d) $d = 0.0$. In each case, the maximum value of Δm yields the best choice for κ , denoted by κ_c . The two overlaps shown represent, respectively, macroscopic superpositions of a neuron state with a stored pattern ρ (m_ρ), and with non-stored patterns orthogonal to the stored one (m_\perp). The lines interpolating the symbols are guides for the eye.

9. General Discussion and Potential Applications

In contrast to standard ANNs, where memories are associated with minima in a metaphorical energy landscape, in the present SDNN model a given minimum appears only after a stimulus correlated with a stored pattern is presented. Such a stimulus is inspired by the form of a random magnetic field, commonly used on models of magnetism, acting independently on each neuron; it is expressed as $\kappa\eta_i$ [$\eta_i = \pm 1$; $\kappa \geq 0$] and kept active during the entirety of the recognition process.

The crucial part of the scheme is to calibrate the intensity of the external stimulus κ in order to cancel out, as much as possible, the noise contribution due to other patterns, which are not correlated with the external pattern presented. We developed a technique for calculating this optimal value (referred to as κ_c), which consists in estimating the maximum gap between: i) the macroscopic superpositions of a neuron state with a stored pattern ρ (m_ρ), and ii) those with non-stored patterns orthogonal to the stored one (m_\perp). More specifically, the maximum of $\Delta m = |m_\rho - m_\perp|$ yields κ_c , and it was shown that, around this value, the overlap m_ρ attained significant values, making the model appropriate for pattern recognition.

We showed that the SDNN considerably increases the capability of the NN to recognize previously stored patterns. The proposal was illustrated through the inclusion of the additional contribution $\kappa\eta_i$ to the standard Hopfield model, for which the number of stored patterns is expressed as $p = \alpha N$ (N representing the total number of neurons), and is known to be unable to recognizing any further patterns for $\alpha \gtrsim 0.14$. Taking into account that analytical calculations may be performed for the Hopfield model, we compared analytical and numerical results, showing good agreement between the two approaches. In contrast to the standard Hopfield case, we found no threshold value for α . Rather, a saturation effect for sufficiently large α was observed, in the sense that the two overlaps, m_ρ and m_\perp , become very close, resulting in $\Delta m \rightarrow 0$. In spite of this, we verified a significant increase

in the recognition capacity of the neural network, so that for this specific application, considering no correlations between stored patterns, the range of possible values of α is enlarged, typically, by a factor 10^2 .

The impressive performance of the present proposal was also demonstrated in situations designed to mimic the common daily tasks of living beings, as described next. (i) Its ability to react promptly to changes in the external environment, reproducing a fundamental characteristic of living beings which need to react quickly to newly presented external patterns. (ii) Its recognition of correlated patterns, showing that correlations lead to a decrease in the optimal values of κ_c , together with an increase in the maxima of Δm ; as a direct consequence, the numerical simulations indicate that the range of α values become enlarged, due to these correlations. These results are in agreement with the recognition of groups of patterns performed by living beings; once one constituent of a given group is recognized, all similar members of this group are identified immediately, and this process requires only a small external stimulus. (iii) Its correct functioning for both asymmetric and diluted synapses, showing that when dilution is included, higher values of the external stimulus are necessary.

It is important to stress that the features of the present SDNN can be implemented both computationally (i.e., in software), as well as on devices. Computationally, there are many modern algorithms that use various gates (such as AND or XOR), and which implement Hebb-like rules (see, e.g., Refs. [66–69]). In line with these algorithms, the synapse values $\{J_{ij}\}$ may be updated whenever new memories are added to the network. In the present proposal, synapses are responsible for the first two terms in the right-hand-side of Equations (8) and (9), so that by changing their values, these two contributions are modified. To introduce the external-pattern contribution [third term of Equations (8) and (9)], one should add the signal of the external stimulus $\kappa\eta_i$ to each neuron i . Hence, neurons will be updated in a sequential way, through the zero-temperature dynamics, $\sigma_i(t+1) = \text{sign}[h_i(t)]$. In a given device, Hebb's rule can be implemented in several ways by means of neuromorphic engineering, which is a recently developed area of research (see, e.g., Refs. [70,71]).

10. Conclusions

Based on the common behavior of living beings, we proposed a new neural-network framework, in which an external stimulus exerts a strong influence on the pattern-recognition process. This external stimulus, introduced in the form of a random field, remains active during the whole recognition process, considerably increasing the capability of the neural network to recognize previously stored patterns. In contrast to more-common attractor neural networks in the absence of an external field, memories are not attractors inside basins of attraction, and basins can be generated for external stimuli that present significant macroscopic superpositions with stored memories.

Finally, it is important to mention that this procedure may be implemented upon a large diversity of neural-network models, utilizing both analytical and numerical investigations. Moreover, its potential application in other processes, distinct from pattern recognition, is very appealing. The present proposal should help to shorten the wide gap between the performance of many of these NN models and the common characteristics of real living beings, improving their performance and, in particular, leading to a considerable increase in their recognition and reaction capabilities. Furthermore, the application of the presented theoretical concepts, both computationally (i.e., in software), as well as on physical devices, is a very promising prospect, and we hope that this task is taken up in the near future.

Author Contributions: Conceptualization, E.M.F.C. and F.D.N.; methodology, E.M.F.C., N.B.M. and F.D.N.; numerical studies and plots, N.B.M.; formal analysis, E.M.F.C., N.B.M. and F.D.N.; discussions, E.M.F.C., N.B.M. and F.D.N.; original draft preparation, E.M.F.C. and F.D.N.; writing—review and editing, E.M.F.C., N.B.M. and F.D.N. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge partial financial support from CNPq, CAPES and FAPERJ (Brazilian funding agencies).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors thank Constantino Tsallis for useful comments.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Mean-Field Equations

In this Appendix A, we detail some of the analytic calculations of Section 4; in particular, those leading to the free energy of Equation (21). We follow the lines of Refs. [21,22], where one assumes that the total number of memories p increases linearly with N , by introducing $\alpha = p/N$ as a finite number. In order to calculate the average over the quenched disorder, one applies the replica method, as defined in Equation (17), where $\mathcal{Z}^n(\{\xi_i^\mu\}, \{\eta_i\})$ corresponds to the partition function of n independent replicas for a given realization of the disorder [1–4]. One has that

$$\langle \mathcal{Z}^n \rangle_{\eta, \xi} = \left\langle \text{Tr}_{\{\sigma^a\}} \left[\prod_{a=1}^n \exp \beta \left(\frac{1}{N} \sum_{(i,j)} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \sigma_i^a \sigma_j^a + \sum_i \kappa \eta_i \sigma_i^a \right) \right] \right\rangle_{\eta, \xi}, \quad (\text{A1})$$

where $\sum_{(i,j)}$ denotes a sum over all distinct pairs of neurons, corresponding to a fully connected neural network; in this case, the mean-field approach becomes exact in the thermodynamic limit. Moreover, a tags the n replicas ($a = 1, 2, \dots, n$) and $\text{Tr}_{\{\sigma^a\}}$ stands for a trace over the neuron states, for each of the n replicas.

Now, using the Gaussian integral

$$\exp(\lambda a^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \exp[-x^2/2 + \sqrt{2\lambda}ax], \quad (\text{A2})$$

one transforms the neuron-pair contributions (i, j) into single-neuron ones, by introducing integration variables m_μ^a . Splitting the p patterns into two sets, namely, condensed ($\mu = 1, 2, \dots, s$), and non-condensed ($\mu = s + 1, s + 2, \dots, p$) patterns, one may calculate the averages over the non-condensed patterns ($\mu > s$),

$$\begin{aligned} \langle \mathcal{Z}^n \rangle_{\eta, \xi} &= e^{-\beta p n / 2} \left\langle \text{Tr}_{\{\sigma^a\}} \int \prod_{a, \mu} \frac{dm_\mu^a}{\sqrt{2\pi}} \exp \left(- \sum_{\substack{\mu, a \\ (\mu > s)}} \frac{(m_\mu^a)^2}{2} + \sum_{\substack{i, \mu \\ (\mu > s)}} \ln \left(\cosh \sqrt{\frac{\beta}{N}} \sum_a m_\mu^a \sigma_i^a \right) \right) \right. \\ &\quad \left. \times \exp \left(- \sum_{\substack{\mu, a \\ (\mu \leq s)}} \frac{(m_\mu^a)^2}{2} + \sum_{\substack{\mu, a \\ (\mu \leq s)}} \sqrt{\frac{\beta}{N}} m_\mu^a \sum_i \xi_i^\mu \sigma_i^a + \beta \kappa \sum_{i, a} \eta_i \sigma_i^a \right) \right\rangle_{\eta, \xi}. \quad (\text{A3}) \end{aligned}$$

Assuming that for non-condensed patterns $m_\mu^a = O(1)$, one uses the approximation $\ln[\cosh x] \approx \frac{1}{2}x^2$ and integrates over m_μ^a for $\mu > s$, leading to

$$\begin{aligned} \langle \mathcal{Z}^n \rangle &= e^{-\beta p n / 2} \text{Tr}_{\{\sigma^a\}} \exp \left\{ - \frac{p}{2} \text{Tr} \ln[(1 - \beta)\mathbf{I} - \beta \mathbf{q}] \right\} \\ &\quad \times \left\langle \int \prod_{\mu, a} \left(\frac{N\beta}{2\pi} \right)^{1/2} dm_\mu^a \exp \left[\beta N \left(- \frac{1}{2} \sum_{\mu, a} (m_\mu^a)^2 + \frac{1}{N} \sum_{i, \mu, a} m_\mu^a \xi_i^\mu \sigma_i^a + \frac{\kappa}{N} \sum_{i, a} \eta_i \sigma_i^a \right) \right] \right\rangle_{\eta, \xi} \quad (\text{A4}) \end{aligned}$$

where the change $m_\mu^a \rightarrow (N\beta)^{1/2} m_\mu^a$ was used. The matrix \mathbf{q} is defined by the elements $q_{ab} = \frac{1}{N} \sum_i \sigma_i^a \sigma_i^b - \delta_{ab}$, and from now on, the index μ applies to condensed states ($\mu = 1, 2, \dots, s$). The integrals over the set $\{m_\mu^a\}$ are evaluated by means of the steepest-

descent method; introducing r_{ab} as Lagrange multipliers for the non-diagonal elements $\{q_{ab}\}$, one has

$$f = \frac{\alpha}{2} + \lim_{n \rightarrow 0} \frac{1}{n} g_n(\{m_\mu^a\}, \{q_{ab}\}, \{r_{ab}\}), \quad (\text{A5})$$

with the functional (in full replica space) $g_n(\{m_\mu^a\}, \{q_{ab}\}, \{r_{ab}\})$ given by

$$g_n(\{m_\mu^a\}, \{q_{ab}\}, \{r_{ab}\}) = \frac{1}{2} \sum_{\mu,a} (m_\mu^a)^2 + \frac{\alpha}{2\beta} \text{Tr} \ln[(1-\beta)\mathbf{I} - \beta\mathbf{q}] + \frac{\alpha\beta}{2} \sum_{\substack{a,b \\ (a \neq b)}} r_{ab} q_{ab} - \frac{1}{\beta} \left\langle \ln \left[\text{Tr}_{\{\sigma^a\}} \exp \left(\frac{\alpha\beta^2}{2} \sum_{a \neq b} r_{ab} \sigma^a \sigma^b + \beta \sum_{\mu,a} m_\mu^a \xi^\mu \sigma^a + \beta\kappa \sum_a \eta \sigma^a \right) \right] \right\rangle_{\eta, \xi}. \quad (\text{A6})$$

Now, one assumes the replica-symmetry ansatz [4], by imposing all off-diagonal elements of the matrix \mathbf{q} (as well as their corresponding Lagrange multipliers) to be equal, i.e., $q_{ab} = q$ and $r_{ab} = r$ ($a \neq b$). Considering the limit $n \rightarrow 0$, the free energy in Equation (A5) may be written as

$$f = \frac{\alpha}{2} + \frac{1}{2} \sum_{\mu=1}^s (m_\mu)^2 + \frac{\alpha\beta}{2} r(1-q) + \frac{\alpha}{2\beta} \left[\ln(1-\beta+\beta q) - \frac{\beta q}{1-\beta+\beta q} \right] - \frac{1}{\beta} \int \mathcal{D}z \left\langle \ln \left[2 \cosh \beta \left(z\sqrt{\alpha r} + \sum_{\mu=1}^s m_\mu \xi^\mu + \kappa \eta \right) \right] \right\rangle_{\eta, \xi}, \quad (\text{A7})$$

where

$$\int \mathcal{D}z \dots \equiv \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \dots \quad (\text{A8})$$

References

1. Amit, D.J. *Modelling Brain Function: The World of Attractor Neural Networks*; Cambridge University Press: Cambridge, UK, 1989.
2. Peretto, P. *An Introduction to the Modelling of Neural Networks*; Cambridge University Press: Cambridge, UK, 1992.
3. Hertz, J.A.; Krogh, A.; Palmer, R.G. *Introduction to the Theory of Neural Computation*; CRC Press: Miami, FL, USA, 2018.
4. Nishimori, H. *Statistical Physics of Spin Glasses and Information Processing*; Oxford University Press: Oxford, UK, 2001.
5. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [\[CrossRef\]](#)
6. Pitts, W.; McCulloch, W.S. How we know universals the perception of auditory and visual forms. *Bull. Math. Biol.* **1947**, *9*, 127–147. [\[CrossRef\]](#)
7. Darwin, C.R.; Wallace, A.R. On the tendency of species to form varieties and on the perpetuation of varieties and species by natural means of selection. *J. Proc. Linn. Soc. Lond. Zool.* **1858**, *3*, 45–62. [\[CrossRef\]](#)
8. Darwin, C.R. *On the Origin of Species*; John Murray: London, UK, 1859.
9. Beddall, B.G. Wallace, Darwin, and the theory of natural selection: A study in the development of ideas and attitudes. *J. Hist. Biol.* **1968**, *1*, 261–323. [\[CrossRef\]](#)
10. Hebb, D.O. *The Organization of Behavior: A Neuropsychological Theory*; Wiley: New York, NY, USA, 1949.
11. Freeman, W.J. The Hebbian paradigm reintegrated: Local reverberations as internal representations. *Behav. Brain Sci.* **1995**, *18*, 631. [\[CrossRef\]](#)
12. Schroeder, M. *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*; W. H. Freeman: New York, NY, USA, 1991.
13. Manneville, P. *An Introduction to Nonlinear Dynamics and Complex Systems*; Imperial College Press: London, UK, 2004.
14. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*; Springer: New York, NY, USA, 2009.
15. Cencini, M.; Cecconi, F.; Vulpiani, A. *Chaos: From Simple Models to Complex Systems*; World Scientific Publishing: Singapore, 2010.
16. Rajan, K.; Abbott, L.F.; Sompolinsky, H. Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E* **2010**, *82*, 011903. [\[CrossRef\]](#)
17. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2554–2558. [\[CrossRef\]](#)
18. Mezard, M.; Parisi, G.; Virasoro, M.A. *Spin Glass Theory and Beyond*; World Scientific: Singapore, 1987.
19. Curado, E.M.F.; Meunier, J.-L. Spin-glass in low dimensions and the Migdal-Kadanoff Approximation. *Phys. A* **1988**, *149*, 164–181. [\[CrossRef\]](#)

20. Peretto, P. Collective properties of neural networks: A statistical physics approach. *Biol. Cybern.* **1984**, *50*, 51–62. [[CrossRef](#)] [[PubMed](#)]
21. Amit, D.J.; Gutfreund, H.; Sompolinsky, H. Spin-glass models of neural networks. *Phys. Rev. A* **1985**, *32*, 1007–1018. [[CrossRef](#)] [[PubMed](#)]
22. Amit, D.J.; Gutfreund, H.; Sompolinsky, H. Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.* **1985**, *55*, 1530–1533. [[CrossRef](#)] [[PubMed](#)]
23. Mezard, M.; Nadal, J.P.; Toulouse, G. Solvable models of working memories. *J. Phys.* **1986**, *47*, 1457–1462. [[CrossRef](#)]
24. Abu-Mostafa, Y.S.; St. Jacques, J.-M. Information capacity of the Hopfield model. *IEEE Trans. Inf. Theory* **1985**, *31*, 461–464. [[CrossRef](#)]
25. McEliece, R.J.; Posner, E.C.; Rodemich, E.R.; Venkatesh, S.S. The capacity of the Hopfield associative memory. *IEEE Trans. Inf. Theory* **1987**, *33*, 461–482. [[CrossRef](#)]
26. Fontanari, J.; Köberle, R. Information storage and retrieval in synchronous neural networks. *Phys. Rev. A* **1987**, *36*, 2475–2477. [[CrossRef](#)] [[PubMed](#)]
27. Chiueh, T.-D.; Goodman, R.M. High-capacity exponential associative memories. *IEEE Int. Conf. Neural Netw.* **1988**, *1*, 153–160.
28. Buhmann, J.; Divko, R.; Schulten, K. Associative memory with high information content. *Phys. Rev. A* **1989**, *39*, 2689–2692. [[CrossRef](#)]
29. Penna, T.J.P.; de Oliveira, P.M.C. Simulations with a large number of neurons. *J. Phys. A* **1989**, *22*, L719–L721. [[CrossRef](#)]
30. Tamas, G. *Physical Models of Neural Networks*; World Scientific Publishing: Singapore, 1990.
31. de Oliveira, P.M.C. *Computing Boolean Statistical Models*; World Scientific Publishing: Singapore, 1991.
32. Aiyer Sreeram, V. B.; Nirajan, M.; Fallside, F. A theoretical investigation into the performance of the Hopfield model. *IEEE Trans. Neural Netw.* **1990**, *1*, 204–215. [[CrossRef](#)]
33. Clark, J.W. Neural network modelling. *Phys. Med. Biol.* **1991**, *36*, 1259–1317. [[CrossRef](#)]
34. Sherrington, D. Neural networks: The spin glass approach. *N.-Holl. Math. Libr.* **1993**, *51*, 261–291.
35. Chapeton, J.; Fares, T.; LaSota, D.; Stepanyants, A. Efficient associative memory storage in cortical circuits of inhibitory and excitatory neurons. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E3614–E3622. [[CrossRef](#)]
36. Alemi, A.; Baldassi, C.; Brunel, N.; Zecchina, R. A three-threshold learning rule approaches the maximal capacity of recurrent neural networks. *PLoS Comput. Biol.* **2015**, *11*, e1004439. [[CrossRef](#)] [[PubMed](#)]
37. Da Silva, I.N.; Spatti, D.H.; Flauzino, R.A.; Liboni, L.H.B.; dos Reis Alves, S.F. *Artificial Neural Networks: A Practical Course*; Springer International Publishing: Cham, Switzerland, 2017.
38. Folli, V.; Leonetti, M.; Ruocco, G. On the maximum storage capacity of the Hopfield model. *Front. Comput. Neurosci.* **2017**, *10*, 144. [[CrossRef](#)] [[PubMed](#)]
39. Kim, D.-H.; Park, J.; Kahng, B. Enhanced storage capacity with errors in scale-free Hopfield neural networks: An analytical study. *PLoS ONE* **2017**, *12*, e0184683. [[CrossRef](#)] [[PubMed](#)]
40. Gosti, G.; Folli, V.; Leonetti, M.; Ruocco, G. Beyond the Maximum Storage Capacity Limit in Hopfield Recurrent Neural Networks. *Entropy* **2019**, *21*, 726. [[CrossRef](#)]
41. Derrida, B.; Nadal, J.P. Learning and Forgetting on Asymmetric Diluted Neural Networks. *J. Stat. Phys.* **1987**, *48*, 993–1009. [[CrossRef](#)]
42. Stariolo, D.A.; Curado, E.M.F.; Tamarit, F.A. Distribution of eigenvalues of ensembles of asymmetrically diluted Hopfield matrices. *J. Phys. A Math. Gen.* **1996**, *29*, 4733–4739. [[CrossRef](#)]
43. Folli, V.; Gosti, G.; Leonetti, M.; Ruocco, G. Effect of dilution in asymmetric recurrent neural networks. *Neural Netw.* **2018**, *104*, 50–59. [[CrossRef](#)] [[PubMed](#)]
44. Tsodyks, M.V.; Feigel'man, M.V. The enhanced storage capacity in neural networks with low activity level. *Europhys. Lett.* **1988**, *6*, 101–105. [[CrossRef](#)]
45. Gutfreund, H. Neural networks with hierarchically correlated patterns. *Phys. Rev. A* **1988**, *37*, 570–577. [[CrossRef](#)] [[PubMed](#)]
46. Krauth, W.; Mézard, M. Storage capacity of memory networks with binary couplings. *J. Phys.* **1989**, *50*, 3057–3066. [[CrossRef](#)]
47. Fontanari, J.F. Generalization in a Hopfield network. *J. Phys.* **1990**, *51*, 2421–2430. [[CrossRef](#)]
48. Fontanari, J.; Theumann, W. On the storage of correlated patterns in Hopfield's model. *J. Phys.* **1990**, *51*, 375–386. [[CrossRef](#)]
49. Erichsen, R., Jr.; Theumann, W.K. Mixture states and storage with correlated patterns in Hopfield's model. *Int. J. Neural Syst.* **1991**, *1*, 347–353. [[CrossRef](#)]
50. Cugliandolo, L.F.; Tsodyks M.V. Capacity of networks with correlated attractors. *J. Phys. Math. Gen.* **1994**, *27*, 741–756. [[CrossRef](#)]
51. Löwe, M. On the storage capacity of Hopfield models with correlated patterns. *Ann. Appl. Probab.* **1998**, *8*, 1216–1250. [[CrossRef](#)]
52. Sompolinsky, H. The theory of neural networks: The Hebb rule and beyond. In *Heidelberg Colloquium on Glassy Dynamics*; van Hemmen, J.L., Morgenstern, I., Eds.; Springer: Berlin, Germany, 1987; pp. 485–527.
53. Derrida, B.; Gardner, E.; Zippelius, A. An Exactly Solvable Asymmetric Neural Network Model. *Europhys. Lett.* **1987**, *4*, 167–173. [[CrossRef](#)]
54. Treves, A.; Amit, D.J. Metastable states in asymmetrically diluted Hopfield networks. *J. Phys. A Math. Gen.* **1988**, *21*, 3155–3169. [[CrossRef](#)]
55. Derrida, B. Distribution of the activities in a diluted neural network. *J. Phys. A Math. Gen.* **1989**, *22*, 2069–2080. [[CrossRef](#)]
56. Evans, M.R. Random dilution in a neural network for biased patterns. *J. Phys. A Math. Gen.* **1989**, *22*, 2103–2118. [[CrossRef](#)]

57. Derrida, B. Dynamics and Optimal Storage Capacity of Neural Networks. *Phys. Scr.* **1989**, T29, 24–30. [[CrossRef](#)]
58. Tamarit, F.A.; Stariolo, D.A.; Curado, E.M.F. Diluted-neural-network model with higher-order interactions. *Phys. Rev. A* **1991**, *43*, 7083–7086. [[CrossRef](#)]
59. Kohler, H.M.; Widmaier, D. Sign-constrained linear learning and diluting in neural networks. *J. Phys. A Math. Gen.* **1991**, *24*, L495–L502. [[CrossRef](#)]
60. Kree, R.; Zippelius, A. Asymmetrically diluted neural networks. In *Models of Neural Networks I*; Domany, E., van Hemmen, J.L., Schulten, K., Eds.; Springer: Berlin, Germany, 1991; pp. 193–212.
61. da Silva, C.R.; Tamarit, F.A.; Lemke, N.; Arenzon, J.J.; Curado, E.M.F. Generalization in a diluted neural network. *J. Phys. A Math. Gen.* **1995**, *28*, 1593–1602. [[CrossRef](#)]
62. da Silva, C.R.; Tamarit, F.A.; Curado, E.M.F. Diluted neural network with refractory periods. *Phys. Rev. E* **1997**, *55*, 3320–3327. [[CrossRef](#)]
63. Sompolinsky, H.; Crisanti, A.; Sommers, H.-J. Chaos in random neural networks. *Phys. Rev. Lett.* **1988**, *61*, 259–262. [[CrossRef](#)]
64. Wainrib, G.; Touboul, J. Topological and dynamical complexity of random neural networks. *Phys. Rev. Lett.* **2013**, *110*, 118101. [[CrossRef](#)]
65. Curado, E.M.F.; Melgar, N.B.; Nobre, F.D. Neural Network Under External Stimulus: Improving Storage Capacity and Reactions. *Phys. A* **2021**, *564*, 125507. [[CrossRef](#)]
66. Widrow, B.; Kim, Y.; Park, D.; Perin, J.K. Nature's Learning Rule: The Hebbian-LMS Algorithm. In *Artificial Intelligence in the Age of Neural Networks and Brain Computing*; Kozma, R., Alippi, C., Choe, Y., Morabito, F.C., Eds.; Academic Press: New York, NY, USA, 2018; pp. 1–30.
67. Available online: <https://medium.com/analytics-vidhya/hebb-network-c38596e1a7a1> (accessed on 7 August 2021).
68. Molter, C.; Salihoglu, U.; Bersini, H. Introduction of a Hebbian unsupervised learning algorithm to boost the encoding capacity of Hopfield networks. In Proceedings of the 2005 International Joint Conference on Neural Networks, Montreal, QC, Canada, 31 July–4 August 2005; pp. 1552–1557.
69. Tolmachev, P.; Manton, J.H. New Insights on Learning Rules for Hopfield Networks: Memory and Objective Function Minimisation. *arXiv* **2020**, arXiv:2010.01472v1.
70. Ziegler, M.; Riggert, C.; Hansen, M.; Bartsch, T.; Kohlstedt, H. Memristive Hebbian plasticity model: Device requirements for the emulation of Hebbian plasticity based on memristive devices. *IEEE Trans. Biomed. Circuits Syst.* **2015**, *9*, 197–206. [[CrossRef](#)] [[PubMed](#)]
71. Hansen, M.; Zahari, F.; Kohlstedt, H.; Ziegler, M. Unsupervised Hebbian learning experimentally realized with analogue memristive crossbar arrays. *Sci. Rep.* **2018**, *8*, 8914. [[CrossRef](#)] [[PubMed](#)]