Associative memories: randomness, continuity and biological plausibility

Enrique Carlos Segura Meccia
School of Computing, Information Systems and Mathematics, South Bank University
103 Borough Road, London SE1 0AA, UK
and
Departamento de Computacion, Universidad de Buenos Aires
segurae@sbu.ac.uk

Abstract

A new, biologically plausible model of associative memory is presented. First, a historical perspective of the more relevant improvements to the basic Little-Hopfield model is given. Then, we introduce a stochastic system with graded response neurons and a network consisting of a non countable number of neurons organized in a continuous metric space. We do this by casting the retrieval process of an analog Hopfield model [7] into the framework of a diffusive process governed by the Fokker-Plank (F-P) equation. This model has the ability to escape spurious memories and, at the same time, is continuous in neural transfer function, topology and time scale. However, it requires the use of path integrals on functional, infinite dimensional spaces, thus turning very difficult any further analytical treatment. Then we resign the continuous topological description of the state space, unifying the graded response units model [7] and the stochastic approach, and obtaining a complete description of the retrieval process at both the microscopic, individual neuron level and the macroscopic level of time evolution of the probability density function over the space of all possible activation patterns.

Keywords: neural networks, associative memory, dynamical systems, Hopfield model, Fokker-Planck equation.

1 Introduction

Since the Little [10,11] and the Hopfield [6] models were formulated to describe the computational ability of an ensemble of simple processing units, appeared to be necessary to reconcile those formal descriptions (provided by discrete models inspired in an Ising system) with a series of biological evidences:

i) One of the strongest objections to the plausibility of the discrete Hopfield model [6] was that a two-state representation of the neural output is, from a biological point of view, an oversimplification and that it is necessary to describe relevant neural activity by firing rates, rather than merely by the presence or the absence of an individual spike.

ii) Little and Hopfield models are deterministic: its dynamics is "dissipative", in the sense that the energy is monotonically decreasing. Hence spurious, uncalled attractor states (whose number increases with that of memories stored in the synaptic weight matrix) cannot be escaped if the initial state is inside its basin attraction: they are surrounded by energy barriers. This deterministic conception of neural dynamics, apart from being impractical, is not very biologically plausible: most biologists currently consider that noise and randomness are almost universal in living systems. Then, if neuronal dynamics is stochastic, neurons can make transitions into states which are opposed to the direction of their presynaptic potential (PSP) [1], due to several factors as the actual difference between the PSP and the threshold and the level of the noise.

iii) Truly continuum of the neural tissue: while the empirical evidence always shows patterns of activity or quiescence involving patches with finite sizes, the ferromagnetic approach suggests systems with punctual, discrete processing units (with no dimensions). In spite of this simplification all the discrete models have been remarkably successful in providing descriptions of emergent processing abilities that correspond to stylized facts concerning basic elementary cognitive processes.

iv) Little and Hopfield models are also discrete in time. This means that its dynamics proceeds step by step and hence it can only be formally described by difference equations. This, besides the problem of logical plausibility (a discrete notion of time is counter-intuitive) prevents the analysis of the dynamical properties of the system by means of strong mathematical tools, which only are applicable for continuous, differential equations dynamics.

Upon time, different solutions were proposed in order to improve the original models concerning each of the aspects listed above:

i) In a later paper, Hopfield [7] aims at a more realistic model by replacing bistable neurons by graded response devices. In either case the retrieval process is again guaranteed by the nature of the matrix of synaptic efficacies.

ii) A number of authors [5,12] have considered more realistic pictures of the neuron response by assuming that the transition between the two stable states of individual neurons is affected by a random field representing thermal fluctuations. They introduced noisy dynamics into the discrete Hopfield model via the Glauber formalism. This allows to control the level of noise by means of a unique parameter $\beta = 1/T$, where T is called the "temperature". In fact, for low values of T, the noise is not too high, hence the system behaves quasi-deterministically and the spurious states persist. On the other hand, for high T, the dynamics is purely ergodic, so there are no attractors at all, either spurious or not. But for some medium range of values of T, it is possible to destabilize the spurious attractors while, at the same time, getting few errors in the retrieval of stored memories.

iii) In [13,14] an extension of the Little-Hopfield model to an infinite continuous dimensional state space is presented. With the only assumption that the synaptic matrix T is symmetric and with non-negative diagonal elements, several results are derived that generalize well known properties of discrete, Ising-type models.

iv) The graded response model by Hopfield [7] solves, at the same time, the problem of continuity of time scale, describing system dynamics by a system of non linear ordinary differential equations (as many equations as units in the system). The extension proposed in [13,14] also obeys to a continuous time dynamics, with the difference that it is described by a unique integro-differential equation over an infinite-dimensional euclidean space.

Our aim is to provide, in the same model, answers to several of these deficiencies simultaneously. A generalization of the non deterministic, finite temperature Glauber dynamics [3] to the case of graded response neurons is attempted (section 4) for the case in which the network consists of a non countable number of neurons organized in a continuous metric space. We do this by casting the retrieval process of a Hopfield model with graded response neurons, into the framework of a diffusive process governed by the Fokker-Plank (F-P) equation. We thus provide a description of the transitional regime that prevails during the retrieval process, that is currently disregarded. That approach formally leads to a functional F-P description of the retrieval dynamics but turns very difficult any further analytical treatment. So, in the last part of the paper (section 5), we restrict the diffusive description to the model with continuous transfer functions and continuous time scale, but in a discrete state space. In other words, we unify the graded response units model [7] and the stochastic approach, obtaining a complete and analytically rigorous description of the retrieval process at both the microscopic, individual neuron level and the macroscopic level of time evolution of the probability density function over the space of all possible activation patterns. In other words, we give a simultaneous answer to objections i), ii) and iv).

2 Associative memory. The original Little-Hopfield model

The problem of associative memory is that of storing a set of p patterns ξ^μ in such a way that when presented a new element ζ as input, the system output is the ξ^μ that most resembles ζ.

In both the Little [10,11] and the Hopfield [6] model, each ξ^μ $(\mu = 1,2,...,p)$ belongs to the set $\{1,-1\}^N$ of all N-tuples whose elements can take on the values 1 (active neuron) or -1 (inactive neuron).

The dynamics of the network is

$$S_i(t+1) = sgn \sum_j T_{ij} S_j(t) \qquad 1 \leq i \leq N$$

where $S_i(t)$ stands for the state of the i-th unit of the system at time t and

$$sgn(x) = \begin{cases} 1 & if \quad x \geq 0 \\ -1 & if \quad x < 0 \end{cases}$$

The stability condition for an element ξ is

$$\xi_i = sgn \sum_j T_{ij} \xi_j \qquad 1 \leq i \leq N$$

The main contribution of Hopfield (not present in Little's proposal), is the introduction of a Lyapunov function:

$$H[S] = -\frac{1}{2} \sum_{ij} T_{ij} S_i S_j$$

In [6] it is proven that if the ξ^μ are generated pseudoorthogonally, its number p doesn't exceed a critical value p_c and the weight matrix is computed following the Hebb rule [4], i.e. as

$$T_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$$

then the system has the property of associative memory.

As for the energy function H, and provided T is symmetric, it decreases monotonically as the system evolves, being its minima the stored ξ^μ (called *attractors* of the dynamics).

In the case of orthogonal memories, the storage capacity increases: it is possible to store and retrieve without errors exactly N memories.

3 Bridging the gaps

3.1 Continuous transfer functions: Hopfield'84

The gap reported in i) (see the Introduction) was bridged by Hopfield himself. In [7] he introduces the following dynamics:

$$\frac{\partial V_i}{\partial t} = -V_i + g_\sigma \left(\sum_j T_{ij} V_j \right) \qquad 1 \leq i \leq N$$

where g_σ is a one-parameter continuous function, non decreasing and odd, also satisfying $\lim\limits_{x \to \pm\infty} g_\sigma(x) = \pm 1$ and $\lim_{\sigma \to \infty} g_\sigma(x) = sgn(x) \quad \forall x \neq 0$

Therefore the elements $V = (V_1, ..., V_N)$ now belong to $[-1,1]^N$. The stability condition is rewritten as:

$$V_i = g_\sigma \left(\sum_j T_{ij} V_j \right) \qquad 1 \leq i \leq N$$

and the energy function is extended as:

$$H[V] = -\frac{1}{2} \sum_{ij} T_{ij} V_i V_j + \sum_i \int_0^{V_i} g_\sigma^{-1}(V) dV$$

The properties of this model are similar to those of the discrete one: its dynamics minimizes the H defined above, whose minima are the stored memories. Moreover, when $\sigma \to \infty$, the attractors tend to be located on the vertices of the hypercube $[-1,1]^N$; in fact, they coincide with those produced by the discrete dynamics [6] for the same T. The existence of spurious attractors is known also in this case.

3.2 Stochastic models

As for the objection ii) in the Introduction, several stochastic versions of the discrete Hopfield model have been proposed [5,12]. In those approaches, which are based on that of Glauber [3], the probability distribution of the state of the i-th processing unit S_i at time $n+1$ is given by

$$P(S_i = \pm 1) = \frac{1}{1+\exp(\mp 2\beta h_i)}$$

As usual $h_i^{\zeta} = \sum_{i=1}^{N} T_{ij}\zeta_j$ is the local field on site i when all neurons have activities labelled by ζ at time n. The parameter β represents the inverse temperature (we set the Boltzmann constant equal to 1) and T is the matrix of synaptic efficacies. This allows to control the level of noise, as described in the introduction. Choosing a proper value for β, it is possible to avoid both deterministic behavior (high β) and completely random walk (low β), thus escaping spurious states while retrieving stored memories with few errors.

The activity pattern of the system is defined through a Markov process that can be formulated as a *Master equation* [1]:

$$P(\xi, n+1) = P(\xi, n) + \sum_{\zeta \neq \xi} [W_\beta(\xi \mid \zeta)P(\zeta, n) - W_\beta(\zeta \mid \xi)P(\xi, n)] \tag{1}$$

where

$$W_\beta(\xi \mid \zeta) = \prod_{i=1}^{N} \frac{1}{\exp\{-\beta \xi_i h_i^{\zeta}\} + 1} \tag{2}$$

and n represents a discrete time. This equation reflects the temporal evolution of the probability distribution associated to the discrete Hopfield model, i.e. $P(\xi, n)$ is the probability of the system to be in an excitation pattern ξ at the n-th step of the evolution process.

3.3 A continuous model of the neural tissue

Now we ask if it is possible to extend the original, discrete Little-Hopfield model in another sense: to provide a continuous topological description of the neural cortex.

We assume that $v(x,t)$ describes the activity of a point-like neuron located in x at time t. This pattern of activity evolves according to:

$$\frac{\partial v(x,t)}{\partial t} = -v(x,t) + g_\sigma\left(\int_\Omega W(x,y)v(y,t)dy\right) \tag{3}$$

with $v(x,t) : \Omega \times R_{\geq 0} \to R$, $\Omega \subset X$. X is a metric space, Ω a compact domain, g_σ a *sigmoid* function, i.e. $g_\sigma \in C^1(R)$, non decreasing and odd and satisfying $\lim_{x\to\pm\infty} g_\sigma(x) = \pm \mathrm{V_M}$, $\lim_{\sigma\to\infty} g_\sigma(x) = sgn(x)$ $\forall x \neq 0$, $| g_\sigma(x) | < \min\{\mathrm{V_M}, \sigma x\}$ and $g'_\sigma(0) = \sigma$.

If we call S the set of all possible states $v(x)$(patterns of activity) of the system, a solution $v(x,t)$ fulfilling (3) is a trajectory in S.

From now on we assume, without loss of generality, that $\mathrm{V_M} = 1$. As for $W : \Omega \times \Omega \to R$, we assume it is continuous almost everywhere (a.e.) in order to warrant that the integral is well defined. As a natural extension of the discrete case we introduce the *local field* on (or net input to) the neuron located in x when the state of the system is $v(y,t)$:

$$h_t^v(x) \triangleq \int_\Omega W(x,y)v(y,t)dy$$

In the particular case $t = 0$, we write $h^v(x) \triangleq h_0^v(x)$. Note that h^v is linear in v.

Given an initial condition $v_0^\mu(x) \triangleq v^\mu(x,0)$ and the solution $v(x,t)$ of (3) that is associated to it, we say that $v^\mu(x)$ is a *memory* or an *attractor* iff:

1) v^μ is an equilibrium point, i.e. $v^\mu(x) = g_\sigma\left(\int_\Omega W(x,y)v^\mu(y)dy\right)$ *a.e.*

2) For every $t_0 \geq 0$ and v_0 a different initial condition that corresponds to the solution v, there exists $\delta(t_0) > 0$ such that if $\| v^\mu - v_0 \| < \delta$ then $\| v^\mu(\cdot,t) - v(\cdot,t) \| \to 0$ when $t \to \infty$.
Thus, attractors are stationary solutions of (3). Except when indicated, we assume S=$L^2(\Omega)$ and moreover $|\Omega| < \infty$ (Ω has finite Lebesgue measure).

We define the *energy* of the system at time t_0 as:

$$H\left[v(\cdot,t_0)\right] = -\frac{1}{2}\int_\Omega\int_\Omega W(x,y)v(x,t_0)v(y,t_0)dxdy + \int_\Omega\int_0^{v(x,t_0)} g_\sigma^{-1}(s)dsdx$$

where v is viewed as a function of x. Thus, each v in S has an energy $H(v)$ associated to it. This is a clear extension of what has been proposed in [7] for the energy of the (discrete) model with neurons with graded activation functions (Hopfield'84). The above definition is justified by the following:
Theorem 1: If W is symmetric, then H is monotonically decreasing with t and it reaches its minima at states $v_{t_e}(x) = v(x,t_e)$ such that

$$\left[\frac{\partial v}{\partial t}(x,t)\right]_{t_e} = 0 \tag{4}$$

a.e.in Ω, i.e. given a solution $v(x,t)$ the minima of H coincide with the equilibrium points of the system. The reciprocal is not necessarily true: from the previous theorem it does not follow that if a solution $v(x,t)$ of (3) satisfies condition (4) for some t^*, then $v(x,t^*)$ is an attractor. For example, the trivial solution $v \equiv 0$ satisfies it for all t, but as we will soon see its stability or instability depends on the slope σ of g_σ at the origin. In general, the possibility to construct non-trivial memories strongly depends on such parameter.
Theorem 2: If $\sigma < \frac{1}{\mathrm{M}|\Omega|}$, being M such that $|W(x,y)| \leq \mathrm{M}$, then the unique stationary solution of (3) is $v \equiv 0$.

Besides the preceding condition $\sigma\mathrm{M}|\Omega| \geq 1$, other requirements have to be fulfilled in order to warrant the effective existence of non trivial solutions.
Theorem 3: if for some $\varepsilon > 0$ holds that $W(x,y) \geq 0$ when $|x-y| < \varepsilon$, then

$$\lim_{\sigma\to\infty} \max_{v\in\{attractors\},x\in\Omega} \min\{1-v(x), 1+v(x)\} = 0$$

In other words, when $\sigma \to \infty$, the attractors approach the asymptotic bounds of g_σ, i.e. ± 1.

Many other results can be proven regarding this model. In particular, if the memories are orthogonal and the synaptic operator is constructed following the Hebb rule, the following results hold:
- The number of memories is not bounded.
- Necessary and sufficient conditions for the memories and the zero to be stable, are derived in terms of the relation between parameters of the transfer function.
- All memories have a basin of attraction with the same positive radius.
- If a spurious memory vanishes at some point of the space, then its basin of attraction has zero radius in the L^2 norm and it is a saddle point of the dynamics.

Also for the sake of biological plausibility, it is possible to:
- impose a "finite resolution" to the stored memories by limiting the minimum size of the activity patches.
- modify the equation that governs the evolution of the system, in order to some graded transition taking place between patches with firing neurons and patches without activity.

As pointed above, this proposal provides continuity, at the same time, in three levels: topology of the space of activity patterns, input-output response of individual neurons and time scale. Only objection iii) remains unanswered, i.e. the system remains deterministic and hence unable to escape spurious, metastable states.

4 Continuity + randomness: a unified approach

Until now we where mainly concerned with what is, thermodynamically speaking, a zero temperature dynamics, i.e. a deterministic law of evolution. In order to enhance the system introduced in the previous section with a stochastic dynamics, we must formulate a finite temperature version of it, in the same fashion as the stochastic versions [5,12] of the basic discrete model (see section 3.2). In our case, ξ and ζ are elements in the (normed) state space S and $W_\beta : \mathrm{S} \times \mathrm{S} \longmapsto [0,1]$ is the transition probability matrix. Being S continuous infinite dimensional it is no longer possible to construct $W_\beta(\xi \mid \zeta)$ from point transition probabilities. This problem can be circumvented with a proper definition of W_β. Using Feynmann path integrals [2], it is possible to obtain an expression for the time evolution of the probability distribution P which has the form of a Fokker-Planck equation in an infinite dimensional normed space:

$$\frac{\partial P(\xi,t)}{\partial t} = -a_1(\xi,P) + \frac{1}{2} a_2(\xi,P) \tag{5}$$

with $a_\nu(\xi,P) = \int_{\mathrm{S}} \{W_\beta(\xi;r)P(\xi,t)\}^{(\nu)}(r,\ldots,r)Dr, \quad (r,\ldots,r) \in \mathrm{S}^\nu$, where $\{W_\beta(\xi;r)P(\xi,t)\}^{(\nu)}$ stands for the ν-th order derivative of $\{W_\beta(\xi;r)P(\xi,t)\}$ at ξ applied to the element $(r,\ldots,r)$ in $\mathrm{S} \times \ldots \times \mathrm{S} = \mathrm{S}^\nu$ (being this ν-th order derivative a linear application from S^ν onto R). For the case $S = R$ we get the well-known one-dimensional Fokker-Planck equation:

$$\frac{\partial P(\xi,t)}{\partial t} = -\frac{\partial}{\partial \xi}\{a_1(\xi)P\} + \frac{1}{2}\frac{\partial^2}{\partial \xi^2}\{a_2(\xi)P\}$$

with $a_\nu(\xi) = \int_{\mathrm{S}} r^\nu W(\xi;r)dr$

The interpretation of the Fokker-Plank equation is simple. The probability density $P(\xi,t)$ is defined on the space of all possible activation patterns $\xi(x)$ and describes the evolution involved in the "retrieval" process of some particular activation pattern starting from a given arbitrary configuration. The transition probability matrix $W_\beta(\xi;r)$ governs this evolution through the "moments" involved in the Fokker-Plank equation. Upon retrieval, the (stationary) probability density reached when t goes to infinity is expected to be peaked in the neighborhood of the recalled memory. This final activation pattern is however not expected to be δ-like due precisely to the finite temperature fluctuations. A pure deterministic case can only be recovered by letting β go to infinity.

This model gives, in fact, an answer to the four objections posed in the introduction: it unifies stochasticity and the three levels of continuity. However, the analytical difficulties involved in this approach are clear: the introduction of path integrals to define the jump moments $a_\nu(\xi,P)$ over the space S, turn very complex any further analysis concerning the actual dynamics of the system, especially equilibrium solutions and its stability. So, it is necessary to resign some of the features of the system in order to get it more analytically tractable.

5 Resigning topological continuity, winning tractability

It is easy to note that the analytical difficulties of the model presented in the previous section are a consequence of the infinite dimensionality of its configuration space. But it is possible to get back to the finite dimensional state space. In fact, our derivation of the finite temperature dynamics for the continuous system is a natural extension of the master equation (1) based on the transition matrix (2). Therefore, it could have been applied to the analog Hopfield model [7] in a direct way, since that case requires a simple multivariate formulation of (5):

$$\frac{\partial P(\xi,t)}{\partial t} = -\sum_{i=1}^{N}\frac{\partial}{\partial \xi_i}\{a_1^i(\xi)P\} + \frac{1}{2}\sum_{i=1,j=1}^{N}\frac{\partial^2}{\partial \xi_i \partial \xi_j}\{a_2^{ij}(\xi)P\}$$

being now a_1^i the usual first order moment for the i-th coordinate, a_2^{ij} the element (i,j) of the covariance matrix and N the number of processing units.

Now the task is, from a theoretical viewpoint, to show that the Hopfield model with graded response activation can be extended to the case in which the evolution law is non deterministic, in the same way as the Hopfield discrete model is extended as to include finite temperature effects through the Glauber dynamics. It can be proven that this leads to a multidimensional Langevin process describing the process at a microscopic level, in the form

$$\dot{\xi}_i = -\xi_i + g_\gamma(h_i^\xi) + L_i(t) \qquad i = 1, \ldots, N$$

where $L_i(t)$ is a stochastic process whose parameter (the same for all units) plays the roll of a temperature, while the time evolution of the probability density distribution is governed by the multivariate Fokker-Planck equation operating over the space of all possible activation patterns. It can also be shown (the work is in progress) that the equilibrium distribution in the sense of ensemble averages (that is, for a large number of units) is Gaussian and that the system reaches it from any continuous initial distribution. In other words, the task is to find the equations representing the evolution of the probability distribution associated to a Hopfield model with a large enough number of graded response neurons with stochastic noise.

In sum, this neural network model of associative memory unifies the two historically more relevant enhancements to the basic Little-Hopfield discrete model: the graded response units approach [7] and the stochastic, Glauber inspired model with a random field representing thermal fluctuations. In other words, it is dynamically stochastic and retains continuity at the transfer function and the time scale levels. As for the topological continuity, although it is actually discrete, its equilibrium distribution and stability are valid in the sense of ensemble averages, i.e., for a large, number of units, hence in a virtually infinite dimensional state space.

In practical terms, the present unified model has two remarkable features: i) Greater biological plausibility, resulting from the conjunction of the graded response transfer function for individual units and the stochastic, noisy dynamics of the whole system; ii) ability to escape spurious memories, due to the form of equilibrium probability distributions, that are assymptotically stable and assign maximum probability to stored memories.

6 Conclusions

We have presented a new approach to the problem of enhancing a neural network of associative memory with as more biologically plausible features as possible.

Basically, four main gaps in that plausibility were listed. We presented a historical perspective of the more important contributions made in the sense of extending the basic Little-Hopfield model as to give answer to those different deficiencies.

Then, we introduced a general stochastic system with graded response neurons and a network consisting of a non countable number of neurons organized in a continuous metric space. We did this by casting the retrieval process of a Hopfield model with graded response neurons, into the framework of a diffusive process governed by the Fokker-Plank (F-P) equation. This model has the ability to escape spurious memories and, at the same time, is continuous in neural transfer function, topology and time scale.

However, it made necessary the use of path integrals on functional, infinite dimensional spaces, thus turning very difficult any further analytical treatment. Therefore, we restricted the diffusive description to the model with continuous transfer functions and continuous time scale, but in a discrete state space. In other words, we unified the graded response units model [7] and the stochastic approach, obtaining a description of the retrieval process at both the microscopic, individual neuron level and the macroscopic level of time evolution of the probability density function over the space of all possible activation patterns.

References

[1] D. J. Amit, “Modeling Brain Function”, Cambridge, Cambridge University Press, 1989.

[2] R. P. Feynman and A. R. Hibbs; “Quantum Mechanics and Path Integrals”; New York, McGraw-Hill, 1965.

[3] R. J. Glauber; “Time-dependent Statistics of the Ising Model”; Journal of Mathematical Physics 4, 294-307 (1963).

[4] D. O. Hebb; “The Organization of Behavior: A Neuropsychological Theory”; New York, Wiley, 1949.

[5] G. E. Hinton and T. J. Sejnowsky; “Optimal Perceptual Inference”; in Proc. IEEE Conf. Computer Vision and Pattern Recognition (Washington, 1983); pp. 448-453, New York, IEEE, 1983.

[6] J. J. Hopfield,“Neural networks and physical systems with emergent collective computational abilities”, Proc. Natl. Acad. Sci. 79, 2554-2558 (1982).

[7] J. J. Hopfield, “Neurons with graded response have collective computational properties like those of two-state neurons”, Proc. Natl. Acad. Sci. 81, 3088-3092 (1984).

[8] N. G. Van Kampen; “Stochastic Processes in Physics and Chemistry”; Amsterdam, Elsevier, 1997.

[9] T. Kohonen; “Self organization and associative memory”; Berlin, Springer-Verlag, 1989.

[10] W. A. Little, “The Existence of Persistent States in the Brain”, Mathematical Biosciences 19, 101-120 (1974).

[11] W. A. Little, “Analytic Study of the Memory Storage Capacity of a Neural Network”, Mathematical Biosciences 39, 281-290 (1978).

[12] P. Peretto; “Collective Properties of Neural Networks: A Statistical Physics Approach”; Biological Cybernetics 50, 51-62.

[13] E. C. Segura; “Memorias Asociativas en Espacios de Dimension Infinita”; Ph.D. Thesis (in spanish), University of Buenos Aires, 1999.

[14] E. C. Segura and R. P. J. Perazzo; “Associative memories in infinite dimensional spaces”; Neural Processing Letters, vol. 12, no. 2, 129-144.