Physica A 197 (1993) 1–13
North-Holland

SDI: 0378-4371(93)E0054-I

PHYSICA A

Fourth order interactions in neural networks

J.J. Arenzon[a], R.M.C. de Almeida[a], J.R. Iglesias[a], T.J.P. Penna[b] and P.M.C. de Oliveira[b]

[a]*Instituto de Física, Universidade Federal do Rio Grande do Sul, C.P. 15051, 91501-970 Porto Alegre, RS, Brazil*[1]
[b]*Instituto de Física, Universidade Federal Fluminense, C.P. 100296, 24020 Niterói, RJ, Brazil*[2]

Received 1 February 1993

We present analytical and numerical results for a truncated version of the RS model for neural networks and also for the generalization of the Hopfield model that considers multineuron interactions. Static properties are analyzed in the replica mean-field approach and dynamical aspects are studied through numerical simulations using the multispin coding algorithm. The load capacity, the sizes of the basins of attraction and the mean convergence time are presented in order to compare the performance of both models as associative memories and to get insight into the shape of their energy landscape.

1. Introduction

Ising-like neuron models have been investigated in order to understand properties of real nervous systems [1,2]. These models consider connected arrays of binary spins and present a non-trivial behavior that is analytically and numerically studied with the tools provided by statistical mechanics.

Networks with N infinite range interacting binary spins ($S_i = \pm 1$) associated to the state of the neurons (active or inactive) are considered to describe learning, storage, and retrieval of information. N-dimensional vectors $S = (S_1, \ldots, S_N)$ represent the possible configurations of the network and the stored information is associated to P of these states (the patterns), denoted by the vectors ξ^μ, $\mu = 1, \ldots, P$. The network load is measured by the parameter

[1] E-mail: ARENZON@IF1.UFRGS.BR
[2] E-mail: GFITJPP@BRUFF.BITNET

$\alpha = P/N$ and its critical value, α_c, is the maximum α for which the network state, for a given energy function $E(S)$, has non-zero superposition with one of the embedded patterns. Different models consider diverse prescriptions for this energy function and for implementing as their minima the P nominated patterns. The performance of a model can then be measured by its storage capacity and its ability in recalling the patterns, e.g., the maximum allowed noise in an initial configuration and the time needed by the net to evolve and stabilize at, or near, one of the P memories.

The Hopfield model with the Hebb learning rule [2] describes satisfactorily an associative memory when the stored patterns are uncorrelated, provided that $\alpha < \alpha_c \approx 0.138$. Beyond this value, the retrieval ability of the network is destroyed. By introducing different learning rules or different energy functions one may attempt to bypass these and other difficulties. For example, the storage capacity can be enhanced by considering multisynapses between the neurons [3,4]; this idea has a biological motivation: axon–axon–dendrite connections, for instance, are relatively common in real nervous systems and can be described as third order synapses (see [4] and references therein). More intricate connections, involving more than two axons, may exist in the brain. However, since binary synapses are highly dominant, higher order terms should be considered as corrections to those terms. As stressed in ref. [4], this feature may play an essential role in the functioning of central nervous systems of superior vertebrated organisms.

Several works have already contemplated the idea of multispin interactions by generalizing the Hopfield model and Hebb learning rule by a monomial of degree $k > 2$ in the Ising spins [3]. Alternatively, the RS model [5] simultaneously considers several orders of interactions, besides the second order Hopfield term, with a simple underlying idea: the energy of a given configuration is proportional to the product of the Hamming distances between the net state and each one of the stored patterns. Although the original RS model yields an ideal phase space, where the only minima of the energy are the stored patterns (and symmetric states in the case of correlated stored information) [5,6], all orders of interactions up to P are present. Nevertheless, higher order terms are less significant, indicating that they may be neglected under adequate circumstances. In this paper we present a truncated version of the RS model (TRS), and study the effect of the first correction to the Hopfield term. We also investigate the effect of a Hopfield-like, fourth order correction and compare both prescriptions.

The paper is organized as follows: section 2 defines the models, sections 3 and 4 present the analytical and simulation results respectively and finally in section 5 we summarize and discuss our conclusions.

2. The models

2.1. The RS model and its truncated version

The RS energy function [5] for a network of N spins is

$$E = N \prod_{\mu=1}^{P} \left(\frac{1}{2N} \sum_{i=1}^{N} (\xi_i^\mu - S_i)^2 \right); \tag{1}$$

this energy function is proportional to the product of the Hamming distances between the network state S and the μth pattern $\boldsymbol{\xi}^\mu$. From eq. (1) it is clear that $E(S) \geqslant 0$ and $E(S) = 0$ if $S = \boldsymbol{\xi}^\mu$, for any μ. It means that, no matter how large is α, the patterns are always global minima of E. Considering the overlaps m_μ between the state of the net S and the pattern $\boldsymbol{\xi}^\mu$, given by

$$m_\mu = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu S_i \,, \tag{2}$$

the energy function, eq. (1), can be written as

$$E = N \prod_{\mu=1}^{P} (1 - m_\mu) \,. \tag{3}$$

This energy function is not invariant under the transformation $m_\mu \to -m_\mu$, therefore the antipatterns $-\boldsymbol{\xi}^\mu$ are not stored. Nevertheless, for every pattern one can also consider its antipattern, by teaching the network both $\boldsymbol{\xi}^\mu$ and $-\boldsymbol{\xi}^\mu$, for every μ. These extreme cases, when patterns and antipatterns are stored (PAS) or when only the patterns are stored (OPS), present strikingly different behaviors and a complete discussion of the phase space landscape in the $\alpha \to 0$ limit as well as simulation results can be found in refs. [5,6].

In the following we concentrate on the PAS case, that reduces to the Hopfield model for uncorrelated patterns and low load parameters: in this case the higher order terms are zero in the thermodynamic limit. The energy function can be rewritten in terms of the overlaps m_μ and it reads

$$E = N \prod_{\mu=1}^{P} (1 - m_\mu^2) \,. \tag{4}$$

The multineuron interaction feature of this equation becomes evident when it is displayed as

$$E = N \left(1 - \sum_{\mu_1} m_{\mu_1}^2 + \sum_{\mu_1 < \mu_2} m_{\mu_1}^2 m_{\mu_2}^2 + \cdots + (-1)^P \sum_{\mu_1 < \cdots < \mu_P} m_{\mu_1}^2 \cdots m_{\mu_P}^2 \right), \tag{5}$$

for a network storing P patterns and the corresponding P antipatterns. Notice that although the first non-trivial term is the Hopfield energy function, the higher order ones are different from any previous model because they contain mixed memory terms. Now we define an energy function, by neglecting the constant zeroth order one and considering the next M terms. This energy function, renormalized by a factor 1/2, reads

$$E = \frac{N}{2} \sum_{l=1}^{M} (-1)^l \sum_{\mu_1 < \cdots < \mu_l} m_{\mu_1}^2 m_{\mu_2}^2 \cdots m_{\mu_l}^2 . \tag{6}$$

Eq. (6) defines a model for neural networks, which can be regarded as the Hopfield model plus correction terms. These terms should prove themselves relevant where the Hopfield model fails: when correlated patterns are considered or the load parameter α is higher than its critical value $\alpha_c \simeq 0.138$. In what follows we consider only the first correction to the Hopfield term ($M = 2$). Eq. (6) can then be rewritten as

$$E = -\frac{1}{2} \sum_{i,j} J_{ij} S_i S_j + \frac{1}{2} \sum_{i,j,k,l} J_{ijkl}^{\mathrm{TRS}} S_i S_j S_k S_l . \tag{7}$$

The learning rule for the second order coupling J_{ij} is the Hebb prescription. The fourth order synapses J_{ijkl}^{TRS}, on the other hand, may be implemented through the following rule:

$$J_{ijkl}^{\mathrm{TRS}} := J_{ijkl}^{\mathrm{TRS}} + \frac{1}{N^2} J_{ij} \xi_k^{P+1} \xi_l^{P+1} , \tag{8}$$

where $\boldsymbol{\xi}^{P+1}$ is the $(P+1)$th pattern to be taught to the net. Eq. (8) can be seen as the multisynapses described in the introduction: the last term is the action of two axons upon a binary synapse. Remark that while binary connections are symmetric, i.e. $J_{ij} = J_{ji}$, the fourth order ones do not present the full symmetry under all possible indices interchange. Anyway, these couplings can be rewritten in a symmetric form [7]

$$J'_{ijkl} = \tfrac{1}{6}(J_{ijkl} + J_{kjil} + J_{ljki} + J_{ikjl} + J_{klij} + J_{ilkj}) \tag{9}$$

and, for the symmetric couplings J'_{ijkl}, the energy (7) is a Liapunov function for the dynamics

$$S_i(t+1) = \mathrm{sgn}\left(\sum_j J_{ij} S_j(t) - \sum_{j,k,l} J'_{ijkl} S_j(t)\, S_k(t)\, S_l(t) \right) . \tag{10}$$

This allows us to use the tools of statistical mechanics.

In the following sections we investigate this model both in the replica mean-field approach and through numerical simulations.

2.2. The generalized Hopfield model

In order to compare performances, we also contemplate a generalization of the Hopfield model (GH) that considers multispin interactions, namely

$$E = -\tfrac{1}{2}N\left(\sum_{\mu} m_{\mu}^{2} + \sum_{\mu} m_{\mu}^{k}\right), \qquad k>2, \tag{11}$$

where k is an integer. In this case the learning rule is

$$J^{\mathrm{H}}_{i_1\ldots i_k} := J^{\mathrm{H}}_{i_1\ldots i_k} + \frac{1}{N^{k-1}}\,\xi^{P+1}_{i_1}\cdots\xi^{P+1}_{i_k}. \tag{12}$$

Again, the differences between both models are in the fourth order connections. This system has a completely different static and dynamical behavior from the one defined by eq. (6). Previous works [3] considered always a unique term of order k. Here we address the problem of having a kth order term as a correction to the original second order Hopfield model. This higher order correction does not change qualitatively the behavior of the net. Also, since the lowest order interaction l in the energy function defines α_c (P scales with N^{l-1}), the presence of the second order term implies that the maximum number of patterns that can be stored is proportional to N, as in the TRS model. This result can be easily obtained by a simple signal-to-noise analysis [9] and is confirmed by the forthcoming calculations and simulations.

3. Mean field theory

3.1. The TRS model

The mean field analysis is performed by means of the standard techniques introduced by Amit et al. [8]. In the GH model, the cross-talk noise is governed by the second order term while the kth order one contributes only to the signal term. On the other hand, in the TRS model, there is a contribution from the higher order terms to the overall noise due to microscopically overlapping patterns. Truncating at fourth order and rewriting the TRS energy function, eq. (6), as

$$E = -\tfrac{1}{2}N\sum_{\mu} m_{\mu}^{2} - \tfrac{1}{4}N\sum_{\mu} m_{\mu}^{4} + \tfrac{1}{4}N\left(\sum_{\mu} m_{\mu}^{2}\right)^{2}, \tag{13}$$

the free energy per neuron can be obtained in the replica trick context and after performing the $n \to 0$ limit and taking the replica symmetric ansatz:

$$f = -\tfrac{1}{2}(1-y)\sum_{\mu} m_{\mu}^2 - \frac{1}{4}\sum_{\mu} m_{\mu}^4 - \tfrac{1}{4}y^2 + \sum_{\mu} t_{\mu} m_{\mu} + \frac{\alpha}{2\beta}\ln[1-\beta(1-y)(1-q)]$$
$$-\frac{1}{2}\,\frac{\alpha q(1-y)}{1-\beta(1-y)(1-q)} + \tfrac{1}{2}\alpha\beta r(1-q) - \frac{1}{\beta}\left\langle\!\left\langle \ln 2\cosh\beta\,(\boldsymbol{t}\cdot\boldsymbol{\xi} + \sqrt{\alpha r}\,z)\right\rangle\!\right\rangle , \tag{14}$$

where y is introduced to linearize the last term in eq. (13) and the variables t_{μ}, q, and r are usually introduced to linearize the non-linear terms. The symbol $\langle\!\langle\ \rangle\!\rangle$ stands for two averages: over the finite number of patterns that may condense and over the Gaussian variable z, related to the infinite microscopic overlapping memories. The saddle point equations are

$$\boldsymbol{m} = \left\langle\!\left\langle \boldsymbol{\xi}\tanh\beta\,(\boldsymbol{\xi}\cdot\boldsymbol{t} + \sqrt{\alpha r}\,z)\right\rangle\!\right\rangle , \tag{15}$$

$$t_{\mu} = (1-y)m_{\mu} + m_{\mu}^3 , \tag{16}$$

$$q = \left\langle\!\left\langle \tanh^2\beta\,(\boldsymbol{t}\cdot\boldsymbol{\xi} + \sqrt{\alpha r}\,z)\right\rangle\!\right\rangle , \tag{17}$$

$$r = q\left(\frac{1-y}{1-\beta(1-y)(1-q)}\right)^2 , \tag{18}$$

$$y = \sum_{\mu} m_{\mu}^2 + \alpha\,\frac{1-\beta(1-y)(1-q)^2}{[1-\beta(1-y)(1-q)]^2} . \tag{19}$$

We investigate solutions to the saddle point equations that present a macroscopic overlap m with one of the memories ($m_{\mu} = m\delta_{1\mu}$) at $T = 0$, which will be compared to the simulation results.

The $T = 0$ ($\beta \to \infty$) limit of the saddle point equations is

$$m = \operatorname{erf}\left(\frac{t}{\sqrt{2\alpha r}}\right) , \tag{20}$$

$$t = (1-y)m + m^3 , \tag{21}$$

$$r = \left(\frac{1-y}{1-C(1-y)}\right)^2 , \tag{22}$$

$$y = m^2 + \frac{\alpha r}{(1-y)^2} , \tag{23}$$

$$C = \sqrt{\frac{2}{\alpha\pi r}}\,\exp\left(\frac{-t^2}{2\alpha r}\right) , \tag{24}$$

where $q = 1$ and $C \equiv \beta(1-q)$. These equations are numerically solved for several values of α and the results are compared with the simulation whose details will be presented later. The numerical solution and simulation results are shown on fig. 1a and present a continuous transition ($m \sim |\alpha - \alpha_c|^{1/2}$ as $\alpha \to \alpha_c^-$) from the retrieval phase to a zero magnetization one. The critical value for the load parameter, α_c is obtained by expanding the above equations for small m:

$$\alpha_c = \left(1 + \sqrt{\frac{2}{\pi}}\right)^2 . \tag{25}$$

3.2. The GH model

Here we consider fourth order corrections ($k = 4$) in eq. (11) and follow the same procedure as in the previous section: the equations to be solved are the same as in the original Hopfield model [8] except for t_μ. They read

$$m = \langle\!\langle\, \xi \tanh \beta\, (\xi \cdot t + \sqrt{\alpha r}\, z) \rangle\!\rangle\,, \tag{26}$$

$$t_\mu = m_\mu + 2m_\mu^3\,, \tag{27}$$

$$q = \langle\!\langle \tanh^2\beta\, (t \cdot \xi + \sqrt{\alpha r}\, z) \rangle\!\rangle\,, \tag{28}$$

$$r = \frac{q}{[1-\beta(1-q)]^2}\,. \tag{29}$$

As in the standard Hopfield model, we also found a first order transition between the retrieval (ordered) phase and the spin glass phase with the critical value for the load parameter $\alpha_c = 1.556$, obtained with replica symmetry. The value of m at the criticality is 0.936.

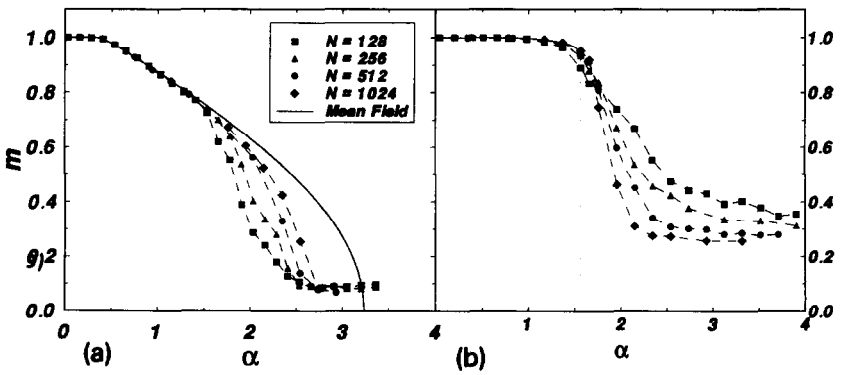

Fig. 1. Final overlap versus α for (a) the TRS model showing the continuous transition and (b) for the GH model showing the first order transition. The full curves are the mean field predictions.

4. Simulation results

In this section we present results obtained from zero temperature simulations of the dynamics of the previously presented models for network sizes of $N = 128$, 256, 512 and 1024 neurons. The steps are: an initial state with overlap m_o with one of the embedded memories is generated and a spin flip test is sequentially applied: a spin is flipped whenever this lowers the energy. This procedure is repeated until a stable fixed point is reached and several quantities are measured as, for instance, the final overlap m_f with the chosen memory and the number of steps required to reach the final state (convergence time).

In order to save computer memory and decrease computation time, we use the multispin coding algorithm proposed by Penna and de Oliveira. For further details on this algorithm we refer the interested reader to references [10].

The maximum value of α for which the network stabilizes at, or very near, an embedded pattern ($m_f \simeq 1$) gives information about the loading limit of the network. For an initial overlap $m_o = 1$ with a given memory, we measured the final overlap with that pattern for several values of α. The results are shown on fig. 1 where we can see fundamental differences between the behavior of both models. The transition is continuous for the TRS model, that is, the retrieval quality, measured by the final value of the overlap m_f, decreases monotonically to a small value as α increases. The magnetization is not zero for $\alpha > \alpha_c$, in contradiction to mean field predictions, but stabilizes around 0.1. This residual magnetization for the (generalized) Hopfield model is approximately 0.2. Also, the comparison with mean field calculations, presented in fig. 1, shows that the simulation yields a different α_c. This difference is probably due to the replica symmetry instability at $T = 0$, what is supported by the negative entropy at zero temperature obtained with the replica symmetry ansatz.

On the other hand, for the GH model with fourth order correction terms, there is a jump in the retrieval parameter at α_c, and the retrieval quality is reasonable (near 1) up to $\alpha \simeq \alpha_c$: the transition is discontinuous, as in the standard Hopfield model. Again, the discrepancy from the predicted value for α_c is caused by the replica symmetry instability at $T = 0$.

We also investigated the evolution of the net for initial overlaps different from one, $m_o < 1$, in order to measure the size of the basins of attraction (fig. 2). For the GH case, the bigger is α, the smaller are the basins, qualitatively reproducing the results of the original Hopfield model [11]. Observe that curves with different values of N but same α superpose, allowing a finite-size scaling analysis [11] to obtain $m_c(\alpha)$, the minimum critical overlap for which the memories can be successfully recalled. In the TRS case, however, the basins seem to be independent of α because curves with different load parameters α superpose. In this case, $m_c \sim 0.1$ for all $\alpha < \alpha_c$, implying big

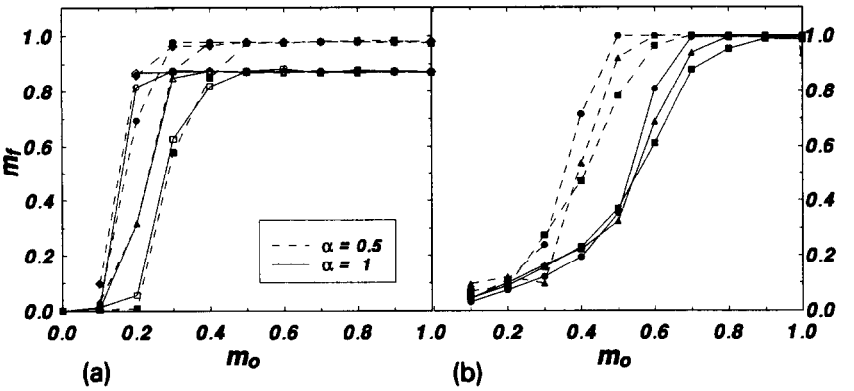

Fig. 2. Final overlap versus the initial superposition m_o for both models, (a) TRS and (b) GH. The basins of attraction are bigger for the TRS model but the retrieval fidelity is lower. The symbols (both blank and filled) size correspond to the legend of fig. 1.

basins of attraction. The comparison between the performance of these two models seems to put in evidence a compromise between the size of the basins of attraction and the retrieval quality, when the network load α increases.

Mean convergence time $\langle T \rangle$, that is, the average number of steps (whole network updating) required to reach a stable state, and the corresponding dispersion are supposed to be related to the irregularity of phase space around the memories [12]. In some models, long convergence times and large dispersions are related to the number of spurious states: when there are many metastable states around the true stored information, there are also many paths with different lengths leading the stimulus to the information. Besides increasing the average convergence time, this diversity of lengths also implies in larger dispersions σ. However, this interpretation is only valid when the stored pattern is at, or very near, the minimum of the basin of attraction, that is, the mean final retrieval quality (m_f) is near 1. On fig. 3b the mean convergence time is shown for the generalized Hopfield model: if the initial state is out of the basin of attraction, the necessary time to reach a stable state grows as the network increases, while inside the basins just one or two steps are enough.

In the TRS case, shown on fig. 3a, the mean convergence time is independent of the initial overlap: for any initial overlap, the necessary number of steps to reach the stationary state is roughly the same. In order to understand this peculiar behavior we remark that for large values of α the state at the bottom of the basin of attraction in the TRS model can be measurably different from the embedded pattern (what does not happen in the GH model below α_c) due to the continuous transition: the final overlap m_f for the TRS model decreases continuously from one to small values as α increases. In this

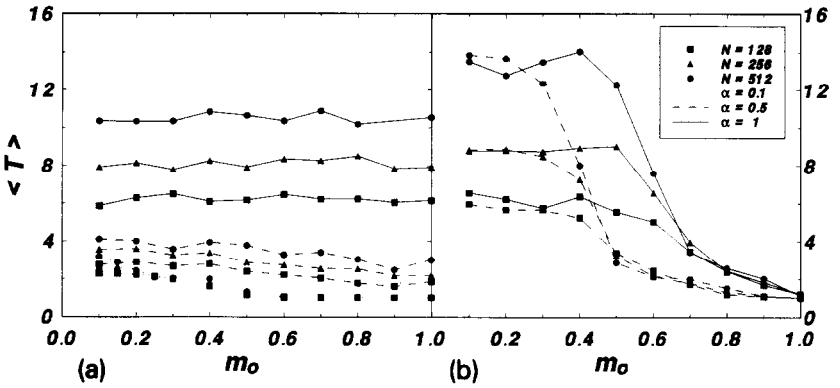

Fig. 3. Mean convergence time versus the initial overlap. In the TRS model (a) there is no dependence on the initial overlap while in the GH case (b) the convergence time depends on if the initial state is in or out of the basin of attraction.

case, when initial states with overlap m_0 with the chosen pattern are generated, they lay on the surface of a hyper-sphere of radius proportional to $1 - m_0$ centered at the pattern. On average, the distance in the phase space from these initial states to the bottom of the basin of attraction is equal to the distance from the memory to the energy minimum, independently of m_0. Consequently, the convergence time does not depend on m_0. For small values of α, on the other hand, the retrieval quality is good ($m_f \approx 1$) and the argument does not hold anymore: indeed, a small decrease in the convergence time is observed as m_0 increases.

This geometry also implies that non-zero dispersions around the mean

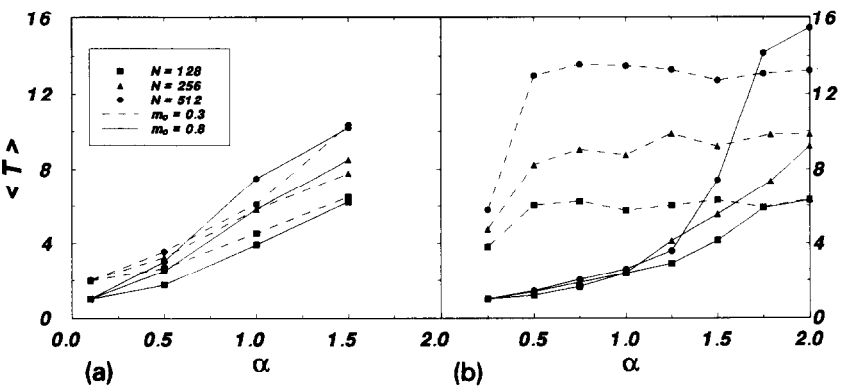

Fig. 4. Mean convergence time versus α. In both models the curves for $m_0 = 0.3$ and 0.8 merge signaling either (a) independence of the initial conditions (TRS) or (b) $\alpha > \alpha_c$ (GH).

convergence time are not always signaling the existence of metastable spurious states: whenever the stored pattern is measurably distant from the bottom of the basin of attraction ($m_f \ll 1$) a non-zero dispersion in the convergence time will be detected due to different distances from initial states to the energy minimum. In summary, for continuous transitions the existence of metastable spurious states is not assured by non-vanishing dispersions around the mean convergence time alone.

The mean convergence time is also plotted versus the load parameter α, figs. 4a and b, for two values of the initial overlap m_o. In the GH case, for $m_o = 0.3$, $\langle T \rangle$ becomes independent of α unless α is low enough. For higher values of α, this initial overlap is too small and the initial state is no longer inside the basin of attraction. However, for $m_o = 0.8$, the convergence time is low until we reach α_c and both curves (for $m_o = 0.3$ and 0.8) merge, indicating that the basins of attraction (and the stability) of memories are destroyed. In the TRS model, the differences between both curves exist only for small α and, as α increases, both curves merge confirming the independence of $\langle T \rangle$ on the initial overlap.

5. Discussion and conclusions

We compared the effect of two different fourth order corrections to the standard Hopfield model by considering two models – GH and TRS – and investigated their performances both through numerical simulation and analytically in the replica symmetric ansatz. The GH model considers a Hopfield-like fourth order term while the TRS energy function is obtained by truncating the RS model at the fourth order. Both corrections improved the load capacity of the net but acted over different aspects, as it is discussed in what follows.

(a) The GH model presents a discontinuous transition at $\alpha_c \simeq 1.556$. This critical value was obtained in the replica symmetric ansatz and is confirmed by numerical simulations. The overall behavior is qualitatively similar to the Hopfield model: for $\alpha < \alpha_c$ the retrieving quality is good ($m_f \simeq 1$) and the size of the basins of attraction decreases with α. Consequently, as α_c is larger due to the correction term, the basins of attraction may become too small.

(b) The TRS model presents a continuous transition at $\alpha_c \simeq 3.232$. However, this higher value cannot be taken as an improvement in comparison to the GH model because here the transition is continuous: for $\alpha > 1$ the final overlap m_f is lower than 0.9 and the retrieving quality is not reliable any more. On the other hand, the size of the basins of attraction are independent of α, as shown by the simulations. Also the convergence time $\langle T \rangle$ increases with α

and, as a consequence of the continuous transition, for high values of the load parameter, $\langle T \rangle$ does not depend on the initial overlap m_0.

As a further conclusion, we found that for transitions where the final overlap m_f decreases continuously to small values, non-zero dispersions around the average convergence time do not necessarily indicate the existence of spurious states near the memories: when m_f is measurably different from one, the initial m_0-overlapping states lay at different distances from the actual minimum (which is not the memory) and consequently may present diverse convergence times to stabilize at the bottom of the basin of attraction. To obtain information about the landscape around the bottom of the basin we must create an initial state with effective overlap with it, not with the embedded pattern.

Finally, the fourth order synapses introduced here are meant to be corrections to the Hopfield model and it would be interesting to study their relative importance with respect to the original second order connections. This can be implemented by considering weights in the fourth order terms. Particularly the TRS model, which shows qualitatively diverse behavior in respect to the original Hopfield model, presents a crossover as the relative weight varies from zero to one. However, these investigations are beyond the scope of this paper and will be presented elsewhere.

Acknowledgements

We acknowledge M.A.P. Idiart and A.T. Bernardes for useful discussions and L. Viana for a critical reading of the manuscript and helpful suggestions. We also thank one of the referees for pointing out to us the symmetrization of the fourth order connections in the TRS model. Work partially supported by Brazilian agencies CNPq, FINEP, FAPERGS and FAPERJ.

References

[1] D.J. Amit, Modeling Brain Function (Cambridge Univ. Press, New York, 1989).
[2] J.J. Hopfield, Proc. Natl. Acad. Sci. USA 79 (1982) 2554.
[3] E. Gardner, J. Phys. A 20 (1987) 3453;
L.F. Abbott and Y. Arian, Phys. Rev. A 36 (1987) 5091;
D. Horn and M. Usher, J. Phys. (Paris) 49 (1988) 389;
G.A. Kohring, J. Phys. 51 (1990) 145;
F.A. Tamarit, D.A. Stariolo and E.M.F. Curado, Phys. Rev. A 43 (2) (1991) 7083.
[4] P. Peretto and J.J. Niez, Biol. Cybern. 54 (1986) 53.
[5] R.M.C. de Almeida and J.R. Iglesias, Phys. Lett. A 146 (1990) 239;
J.J. Arenzon, R.M.C. de Almeida and J.R. Iglesias, J. Stat. Phys. 69 (1992) 385.

[6] J.J. Arenzon, R.M.C. de Almeida, J.R. Iglesias, T.J.P. Penna and P.M.C. de Oliveira, J. Phys. (Paris) I 1 (1992) 55;
T.J.P. Penna, P.M.C. de Oliveira, J.J. Arenzon, R.M.C. de Almeida and J.R. Iglesias, Int. J. Mod. Phys. C 2 (1991) 711.
[7] J.L. van Hemmen, Phys. Rev. Lett. 49 (1982) 409.
[8] D.J. Amit, H. Gutfreund and H. Sompolinsky, Phys. Rev. A 32 (1985) 1007;
D.J. Amit, H. Gutfreund and H. Sompolinsky, Ann. Phys. NY 173 (1987) 30.
[9] V. Deshpande and C. Dasgupta, J. Stat. Phys. 64 (1991) 755.
[10] T.J.P. Penna and P.M.C. de Oliveira, J. Phys. A 22 (1989) L719.
P.M.C. de Oliveira, Computing Boolean Statistical Models (World Scientific, Singapore, 1991).
[11] B.M. Forrest, J. Phys. A 21 (1988) 245.
[12] G. Kohring, J. Phys. A 23 (1990) 2237;
I. Kanter, Phys. Rev. A 40 (1990) 2611.