On the multi-neuron interaction model without truncating the interaction

1994 J. Phys. A: Math. Gen. 27 5871

(http://iopscience.iop.org/0305-4470/27/17/019)

J. Phys. A: Math. Gen. 27 (1994) 5871–5878. Printed in the UK

On the multi-neuron interaction model without truncating the interaction

D Bollé†‡, J Huyghebaert† and G M Shim†§

Institiuut voor Theoretische Fysica, K U Leuven, B-3001 Leuven, Belgium

Received 27 May 1994

Abstract. A replica-symmetric mean-field theory approach is presented to the multi-neuron interaction model introduced by de Almeida and Iglesias (1990 *Phys. Lett.* **146A** 239). Fixed-point equations are derived for the relevant order parameters of the model, extended to include biased patterns, without truncating the interaction. The capacity–bias and the temperature–capacity phase diagrams are discussed. Compared with the truncated version of the model, it is found that the capacity at zero temperature is infinite and that the retrieval states satisfy the de Almeida–Thouless stability condition.

1. Introduction

Many attempts have been made to improve the storage capacity and retrieval properties of neural networks by generalizing the Hopfield model in various directions (see, e.g., [1] for an overview).

In particular, networks with multi-neuron interactions have been looked at from this point of view [2–13]. It is well known that, if the two-neuron interaction of a Hopfield model of size N is replaced by a polynomial interaction of degree $p > 2$, then up to $O(N^{p-1})$ random patterns can be stored and retrieved.

Recently a model has been proposed by de Almeida and Iglesias [14] where all orders of neuron interactions are included in such a way that the Hamiltonian is proportional to the product of the Hamming distances between the network state and the embedded patterns. This system does not necessarily have a spin up–down symmetry, in contrast with the Hopfield model and its generalizations. However, this symmetry can always be imposed by adding the antipatterns to the original set of embedded patterns. Actually this possibility yields two extreme cases: storage of both the patterns and antipatterns (PAS) or storage of the patterns only (OPS). An extensive simulation of these cases [15] essentially reveals that for unbiased patterns a PAS network does not show a limit on the storage capacity. For the OPS network, however, the basins of attraction may become too small. This can be traced back to the existence of a spurious *central state* [1] in the direction of the sum vector of all embedded patterns. This state is built up out of positive overlaps of order $1/\sqrt{N\alpha}$ with all patterns and has a large basin of attraction.

A truncated version of the PAS model has been introduced and its retrieval performance has been compared with a generalized Hopfield model both by numerical simulations and

† Also at Interdisciplinair Centrum voor Neurale Netwerken, K U Leuven.

‡ E-mail address: FGBDA18@cc1.kuleuven.ac.be

§ E-mail address: FGBDA28@cc1.kuleuven.ac.be

0305-4470/94/175871+08$19.50

by a replica-symmetric mean-field analysis in the case of fourth-order corrections [16]. Also the strength of this correction term has been varied [17]. The phase diagrams for both models were discussed in detail. While the overall behaviour of the generalized Hopfield model qualitatively resembles the original Hopfield model, a strikingly different and very rich behaviour has been found for the PAS model, depending on the strength of the fourth-order term.

In this paper, we consider the original model of de Almeida and Iglesias allowing for an extensive number of embedded patterns. The patterns may be biased. We present a replica-symmetric mean-field analysis without truncating the interaction. This is realized by observing that the product of the terms in the Hamiltonian related to the non-condensed patterns can be described in the thermodynamic limit by the mean-square random overlap.

Independently of the bias we find that the critical storage capacity is infinite at zero temperature and that the replica-symmetric solution is stable against replica-symmetry breaking (RSB). In order to compare the retrieval performance of the full (i.e. non-truncated) with the truncated model we study the temperature–capacity and the capacity–bias phase diagram, and the behaviour of the overlap.

The rest of the paper is organized as follows. In section 2 the model is defined. Section 3 presents a replica-symmetric mean-field theory analysis. Section 4 discusses the solutions of the fixed-point equations and the corresponding capacity–bias and temperature–capacity phase diagrams. Also the behaviour of the overlap as a function of the relevant parameters is considered. A comparison is made with the truncated model. The concluding remarks are given in section 5.

2. The model

Let us consider a network consisting of N Ising neurons $\{\sigma_i\}$ interacting via the Hamiltonian [14]

$$H = N \prod_{\mu=1}^{p} \left[1 - \left(\frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \sigma_i \right)^2 \right] \tag{1}$$

where the $\{\xi_i^\mu\}, i = 1, \ldots, N, \mu = 1, \ldots, p$ is a collection of independent identically distributed random variables with p the number of embedded patterns.

The Hamiltonian (1) is proportional to the product of the Hamming distances between the network state and the embedded patterns. It is positive-definite, only being zero when the network state exactly coincides with one of the patterns. This means that the configurations corresponding to the embedded patterns are always the global minima. The first non-trivial (quadratic) term in an expansion of this Hamiltonian with respect to the $\{\sigma_i\}$ corresponds to the Hopfield model.

This model stores both the patterns and antipatterns and is therefore refered to as the PAS model. It has been studied in detail for low loading (finite p) of unbiased patterns in [1, 14]. In contrast with the Hopfield model the symmetric mixture states with more than one pattern are always unstable. Numerical simulations discussed in [15] found no limit on the storage ratio $\alpha = p/N$.

For biased patterns, however, these simulations [15] show that the PAS model has quite a limited storage capacity and a very small relative size of basin of attraction. In view of the results on the Hopfield model for neurons with low levels of activity [18] this should not be surprising. In fact, it is easy to show that for biased patterns the Hamiltonian (1) does not allow for Mattis-type solutions, which are most desirable as an associative memory.

To overcome these problems we consider the following Hamiltonian:

$$H = N(1-a^2)\prod_{\mu=1}^{p}(1-m_\mu^2) \tag{2}$$

where the $\{\xi_i^\mu\}$ is a collection of independent identically distributed random variables taking the values ± 1 with probabilities $\frac{1}{2}(1\pm a)$ and where

$$m_\mu = \frac{1}{N}\sum_i \frac{(\xi_i^\mu - a)\sigma_i}{1-a^2} \tag{3}$$

are the scaled (modified) overlaps between the network state and the embedded patterns involving the bias parameter a. In the sequel we will always use the overlap parameter (3).

The Hamiltonian (2) allows for Mattis-type solutions. Furthermore, the modification and scaling of the overlap is necessary to keep the essential property that the configurations corresponding to one of the stored patterns still have a zero Hamiltonian.

3. Mean-field theory

In the following we consider the network described by the Hamiltonian (2) allowing for an extensive number of biased patterns p. As usual we are interested in the situation where only a finite number l of the patterns, say $\mu = 1,\ldots,l$, are condensed in the network. The others have an overlap of at most order $O(1/\sqrt{N})$.

To perform the standard mean-field analysis [19] it is essential to observe that the product of the terms in (2) related to the non-condensed patterns may be described in the thermodynamic limit by the mean-square random overlap

$$\prod_{\mu=l+1}^{\alpha N}(1-m_\mu^2) \approx \exp\left(-\sum_{\mu=l+1}^{\alpha N} m_\mu^2\right) \equiv \exp(-\alpha s)\,. \tag{4}$$

Within the replica-symmetric approximation, we obtain the free energy

$$\begin{aligned}\beta f = {} & (1-a^2)\sum_{\mu=1}^{l}\tilde{m}_\mu m_\mu + \tfrac{1}{2}\tilde{q}(1-q) + \alpha(1-a^2)\tilde{s}s + \beta(1-a^2)\mathrm{e}^{-\alpha s}\prod_{\mu=1}^{l}(1-m_\mu^2)\\ & + \frac{\alpha}{2}\left[\log[1-2\tilde{s}(1-q)] - \frac{2\tilde{s}q}{1-2\tilde{s}(1-q)}\right]\\ & - \int \mathrm{D}z \left\langle\!\!\left\langle \log\left\{2\cosh\left[\sum_{\mu=1}^{l}\tilde{m}_\mu(\xi^\mu - a) + \sqrt{\tilde{q}}\,z\right]\right\}\right\rangle\!\!\right\rangle. \end{aligned} \tag{5}$$

Here the m_μ are the overlaps (3) with their conjugate variables $\tilde{m}_\mu$, q is the Edward–Anderson order parameter with its conjugate variable $\tilde{q}$, and s is the mean-square random overlap for the non-condensed patterns (cf equation (4)) with its conjugate variable $\tilde{s}$. The notation $\langle\!\langle\ \rangle\!\rangle$ stands for the average over the distribution of the condensed patterns and Dz is the Gaussian measure, $\mathrm{D}z = \mathrm{d}z\,\exp\left(-\frac{1}{2}z^2\right)/\sqrt{2\pi}$.

These variables are to be self-consistently determined by the solution of the fixed-point equations which minimize the free energy (5). These fixed-point equations read

$$m_\mu = \int \mathrm{D}z \left\langle\!\!\left\langle \frac{\xi^\mu - a}{1-a^2}\tanh\left[\tilde{\beta}\left(\sum_{\nu=1}^{l}\frac{m_\nu}{1-m_\nu^2}(\xi^\nu - a) + \sqrt{\alpha r}\,z\right)\right]\right\rangle\!\!\right\rangle \tag{6}$$

$$q = \int \mathrm{D}z \left\langle\!\!\left\langle \tanh^2 \left[\tilde{\beta} \left(\sum_{\nu=1}^{l} \frac{m_\nu}{1-m_\nu^2} (\xi^\nu - a) + \sqrt{\alpha r}\, z \right) \right] \right\rangle\!\!\right\rangle \tag{7}$$

$$s = \frac{1}{1-a^2} \frac{1-\tilde{\beta}(1-q)^2}{\left[1-\tilde{\beta}(1-q)\right]^2} \tag{8}$$

where

$$\tilde{\beta} = 2\beta \mathrm{e}^{-\alpha s} \prod_{\mu=1}^{l} (1-m_\mu^2) \qquad r = \frac{q}{[1-\tilde{\beta}(1-q)]^2} .$$

At this point we note that the magnitude $(\tilde{\beta}/\beta)\sqrt{\alpha r}$ of the Gaussian noise which is essentially caused by the non-condensed patterns is not the same as the mean-square random overlap. The reason is that the Hamiltonian is not quadratic with respect to the overlaps of the non-condensed patterns. The magnitude of the Gaussian noise depends explicitly not only on q but also on m_μ in such a way that the larger the overlap, the smaller the noise becomes. Furthermore, this magnitude decreases for increasing α, in contrast with the Hopfield model. There the contributions to the noise from the non-condensed patterns are additive so that increasing α means increasing the noise. However, in the model we are considering these contributions are multiplicative in such a way that the strength of the noises is exponentially small for large α.

The local stability of the replica-symmetric solution against breaking may be determined by examining the sign of the eigenvalues of the matrix of the transverse fluctuations in the replicon space [20]. It turns out that the solution is stable if the following expression, related to the de Almeida–Thouless (AT) replicon eigenvalue, is positive:

$$\lambda \equiv \left[1-\tilde{\beta}(1-q)\right]^2 - \alpha\tilde{\beta}^2 \int \mathrm{D}z \left\langle\!\!\left\langle \cosh^{-4} \left[\tilde{\beta} \left(\sum_{\nu=1}^{l} \frac{m_\nu}{1-m_\nu^2} (\xi^\nu - a) + \sqrt{\alpha r}\, z \right) \right] \right\rangle\!\!\right\rangle . \tag{9}$$

In the following section we study the fixed-point equations (6)–(8) for arbitrary temperature.

4. Retrieval properties and phase diagrams

At zero temperature the fixed-point equations for a Mattis-type solution $m_\mu = m\delta_{\mu,\nu}$ (for some ν) reduce to two equations

$$m = \frac{1}{2} \sum_{\eta=\pm 1} \mathrm{erf}\left(\frac{m(1-a\eta)}{(1-m^2)\sqrt{2\alpha r}} \right) \tag{10}$$

$$r = (1-a^2)s = \frac{1}{(1-C)^2} \tag{11}$$

where

$$C = \frac{1}{\sqrt{2\pi\alpha r}} \sum_{\eta=\pm 1} (1+a\eta) \exp\left(-\frac{m^2(1-a\eta)^2}{(1-m^2)^2 2\alpha r} \right) . \tag{12}$$

It is interesting to note that there appears a solution $m = 1$ in addition to a trivial solution $m = 0$ for all α. The $m = 1$ solution is the global minimum of the Hamiltonian for all α since its energy is zero, i.e. the lower bound of the Hamiltonian. An indication for the stability of this solution against RSB is given by the entropy

$$S_0 = -\frac{\alpha}{2} \left[\log(1-C) + \frac{C}{1-C} \right] . \tag{13}$$

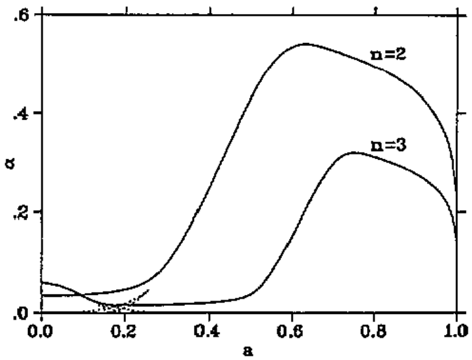

Figure 1. The α–a diagram for the $n = 2, 3$ symmetric states of the full model. The meaning of the broken curves is explained in the text.

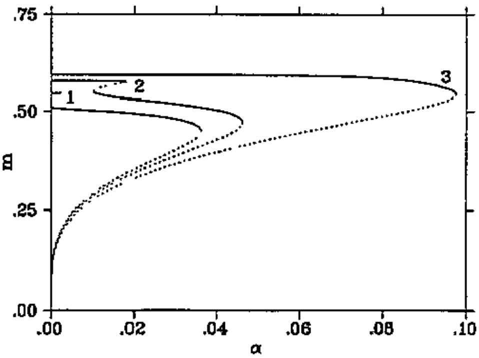

Figure 2. The m–α diagram for the $n = 2$ symmetric state of the full model with $a = 0.15$ (1), $a = 0.2$ (2) and $a = 0.3$ (3). The full curves represent stable solutions; the broken curves are the unstable branches.

As long as $\alpha > 0$ and $C > 0$, S_0 is negative and replica symmetry is broken. The solution $m = 1$, however, yields $S_0 = 0$. Furthermore, it leads to $\lambda = 1$ (cf equation (9)) and hence it is stable. (Of course one cannot exclude the possibility of a first-order transition in the replicon space.) An explanation of this fact may be given by observing that if one of the overlaps becomes unity, the effect from the non-condensed patterns disappears because of the very structure of the Hamiltonian. We conclude that the storage capacity is infinite, independent of the value of the bias, which is also consistent with the simulation for unbiased patterns [15].

In the truncated model the storage capacity remains finite and the replica-symmetric solutions are unstable against RSB. This is also the case in the Hopfield model. It suggests that the contributions from the higher-order interactions are very important for the associativity as well as for the stability of the replica-symmetric solutions.

Next we consider the symmetric mixture states of the form $m_\mu = m_n \sum_{\nu=1}^{n} \delta_{\mu\nu}$ having the same overlap with n patterns. Figure 1 shows the storage capacity for the symmetric mixture states for $n = 2, 3$. The overall properties of these mixture states are somewhat different from those of the Hopfield model. The symmetric mixture state involving two patterns has a finite storage capacity as a goes to zero, in contrast with the corresponding mixture state for the Hopfield model. In the region marked by the broken curves there appear other $n = 2, 3$ mixture states with different overlap. This can be clearly understood from figure 2 where, as an illustration, the overlap is shown for the $n = 2$ symmetric state. Indeed we see the appearance of two turning points for certain bias regions. In fact, the broken curves in figure 1 represent the turning point at smaller capacities. Finally we remark that although the mixture states for unbiased patterns are unstable at $\alpha = 0$ [1], there appear stable mixture states for finite α. A similar general behaviour has been found in Potts-glass neural networks [21].

Let us now turn to the properties of the model at finite temperatures. The paramagnetic solution characterized by $m = q = 0$ exists for any α and a with a well defined free energy. This is also true for the truncated model. In contrast, the paramagnetic solution in the Hopfield model has an ill-defined free energy if T is less than $1 - a^2$. However, the paramagnetic solution violates the AT stability condition (9) if the temperature is less than

$$T_g = 2(1 + \sqrt{\alpha}) \exp\left[-(1 - a^2)^{-1} \sqrt{\alpha}(1 + \sqrt{\alpha})\right] . \qquad (14)$$

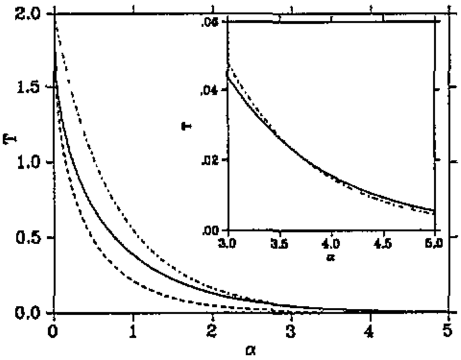

Figure 3. The T–α phase diagram for the full model with $a = 0$. The full curve denotes the retrieval transition curve, the broken curve indicates the thermodynamic transition and the chain curve the spin-glass transition.

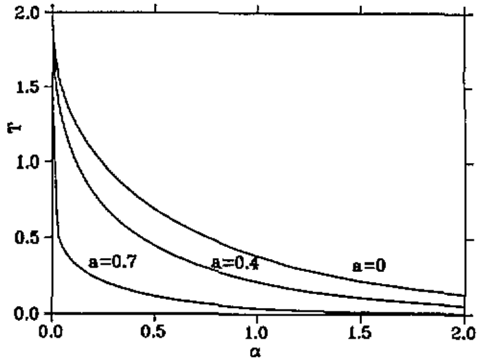

Figure 4. The T–α retrieval transition curve for the full model with $a = 0, a = 0.4$, and $a = 0.7$.

Below T_g the spin-glass solution appears continuously. Interestingly, near $\alpha = 0$, the transition temperature T_g behaves differently depending on whether the patterns are biased or not. It decreases as $\sqrt{\alpha}$ for biased patterns ($a > 0$), while it decreases linearly for unbiased patterns ($a = 0$). For large α it approaches zero exponentially in both cases. In the truncated model (where only the unbiased patterns are considered and the temperature is scaled by a half) the spin-glass transition temperature is found to be

$$T_g' = (1 + \sqrt{\alpha})(1 - \sqrt{\alpha} - \alpha) \tag{15}$$

which goes to zero at $\alpha = (3 - \sqrt{5})/2$. We remark that, except for the trivial scaling factor 2, the expression for T_g' corresponds to the first two terms of the expansion of T_g (cf (14)) for small α.

The Mattis retrieval state at finite temperature is found numerically by solving the fixed-point equations (6)–(8) for several values of a. For $a = 0$ the result is shown as the full curve in figure 3. In this figure we also display the thermodynamic transition line (broken curve) below which the retrieval state becomes the global minimum of the free energy. Finally the spin-glass transition temperature is also presented (chain curve). We remark that the retrieval and spin-glass transition curves cross at $\alpha = 3.66$. The biased case $a > 0$ is illustrated in figure 4 by drawing retrieval transition curves for several values of a. As expected, the thermodynamic and spin-glass transitions curves show a similar behaviour (cf figure 3).

It is certainly interesting to compare the retrieval performance of the full model with that of the truncated model. This is done by studying, at $a = 0$, both the retrieval transition temperature as a function of α and the overlap as a function of α and T. The results are collected in figures 5–7.

The T–α retrieval region is shown in figure 5. For $\alpha < 2.49$ the truncated model allows retrieval up to slightly higher temperatures; for larger values of α the full model does better. However, this result needs to be qualified when looking at the overlap-temperature diagram in figure 6 for different values of the capacity. For low capacity there is no difference between both models on the scale of the figure. For $\alpha = 1$ the overlap is always bigger for the full model (full curve). This stays that way for an overlap in the crossing region. This effect becomes even more dramatic by looking at the value of the overlap,

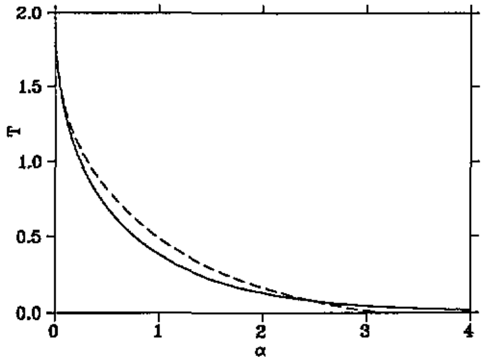

Figure 5. The T–α retrieval transition curve for the full model (full curve) and the truncated model (broken curve) with $a = 0$.

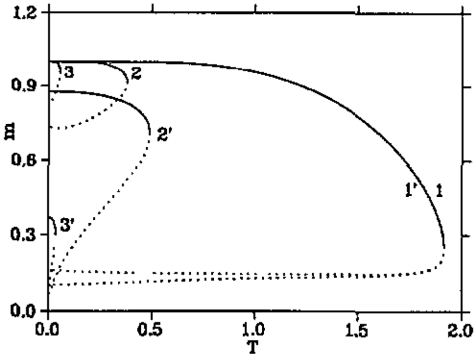

Figure 6. The m–T diagram at $\alpha = 0.001, \alpha = 1$ and $\alpha = 2.8$ for the retrieval state of the full model (1)–(3) and the truncated model (1′)–(3′) with $a = 0$. The full curves represent stable solutions, the broken curves are the unstable branches.

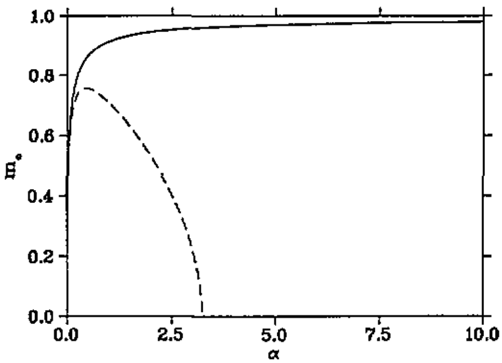

Figure 7. The m_c–α diagram at the retrieval transition temperature for the full model (full curve) and the truncated model (broken curve) with $a = 0$.

m_c, at the transition temperature of the retrieval region as a function of α, as displayed in figure 7.

5. Concluding remarks

We have considered the thermodynamic properties of the multi-neuron interaction model with an extensive number of stored patterns. The replica-symmetric mean-field theory of the model, extended to include biased patterns, and without any truncation of the interaction has been presented. This is essentially realized by observing that the product of the terms in the Hamiltonian related to the non-condensed patterns can be described in the thermodynamic limit by the mean-square random overlap.

The capacity–bias and the temperature–capacity phase diagrams, and the behaviour of the overlap as a function of both the temperature and the capacity have been discussed in comparison with the analogue quantities for the truncated model and the Hopfield model.

It is found that for all bias values the replica-symmetric retrieval states are stable against RSB and the storage capacity is infinite at zero temperature. The free energy of the retrieval state at zero temperature is zero, which is the lower bound for the Hamiltonian. For zero bias the full model leads to a bigger overlap than the truncated model, especially for higher values of the capacity. The latter allows for retrieval up to slightly higher values of the temperature for a capacity $\alpha < 2.49$, essentially by giving up some retrieval quality.

The spin-glass transition temperature decreases for small α, linearly for unbiased patterns and as a square root for biased patterns, and it becomes lower than the retrieval transition temperature for large α, a behaviour that is also seen in the truncated model.

These results clearly suggest that the contribution from the higher-order interactions is important for the retrieval performance of the network.

Acknowledgments

This work has been supported in part by the Research Fund of the K U Leuven (grant OT/91/13). DB and JH thank the Belgian National Fund for Scientific Research and the Inter-University Institute for Nuclear Sciences for financial support.

References

[1] Arenzon J J, de Almeida R M C and Iglesias J R 1992 *J. Stat. Phys.* **69** 385
[2] Peretto P and Niez J J 1986 *Biol. Cybern.* **54** 53
[3] Personnaz L, Guyon L and Dreyfus G 1987 *Europhys. Lett.* **8** 863
[4] Gardner E 1987 *J. Phys. A: Math. Gen.* **20** 3453
[5] Abbott L F and Arian Y 1987 *Phys. Rev.* A **36** 5091
[6] Baldi P and Venkatesh S S 1987 *Phys. Rev. Lett.* **58** 913
[7] Kanter I 1988 *Phys. Rev.* A **38** 5972
[8] Horn D and Usher M. 1988 *J. Physique* **49** 389
[9] Kohring G A 1990 *J. Physique* **51** 145
[10] Bajpai R P and Shlukla P 1990 *Phys. Rev.* A **42** 5006
[11] Kürten K E 1990 *Statistical Mechanics of Neural Networks* ed L Garrido (Berlin: Springer)
[12] Tamarit F A, Stariolo D A, and Curado E M F 1991 *Phys. Rev.* A **43** 7083
[13] Wang L and Ross J 1991 *Phys. Rev.* A **44** R2259
[14] de Almeida R M C and Iglesias J R 1990 *Phys. Lett.* **146A** 239
[15] Arenzon J J, de Almeida R M C, Iglesias J R, Penna T J P and de Oliveira P M C 1992 *J. Physique I* **2** 55
[16] Arenzon J J, de Almeida R M C, Iglesias J R, Penna T J P and de Oliveira P M C 1993 *Physica* **197A** 1
[17] Arenzon J J and de Almeida R M C 1993 *Phys. Rev.* E **48** 4060
[18] Amit D J, Gutfreund H and Sompolinsky H 1987 *Phys. Rev.* A **35** 2293
[19] Amit D J, Gutfreund H, and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 1530; 1987 *Ann. Phys., N Y* **173** 30
[20] de Almeida J R L and Thouless D 1978 *J. Phys. A: Math. Gen.* **11** 983
[21] Bollé D, Cools R, Dupont P and Huyghebaert J 1993 *J. Phys. A: Math. Gen.* **26** 549