IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Generalization in a Hopfield network with noise

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1993 J. Phys. A: Math. Gen. 26 3983

(http://iopscience.iop.org/0305-4470/26/16/015)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.68
The article was downloaded on 01/06/2010 at 19:26

Please note that terms and conditions apply.

J. Phys. A: Math. Gen. 26 (1993) 3983–3993. Printed in the UK

Generalization in a Hopfield network with noise

P R Krebs and W K Theumann

Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, 91501-970 Porto Alegre, RS, Brazil

Received 15 February 1993, in final form 11 May 1993

Abstract. The generalization ability of the Hopfield model of neural networks trained with examples is studied in mean-field theory in the presence of synaptic noise. Although the latter improves the generalization ability for a finite number of concepts, it does not for a macroscopic number of them. Nevertheless, the network performance is still robust against synaptic noise. Numerical simulations are performed to verify the mean-field theory predictions.

1. Introduction

The main interest over the past in feedback neural networks has been in their storage capacity and ability to retrieve a set of learned patterns [1, 2]. A further attractive issue is the generalization ability (and related aspects, i.e., categorization [3–5]), that is, the spontaneous emergence of features that were not originally built into the network in the training stage. Candidates for this goal are networks with hierarchical patterns, as long as the learning rules that are employed do not already contain the specific features desired as an output that appeared in earlier works [6]. Generalization has been the subject of intensive studies in the context of single-layer feed-forward neural networks [7–12].

The creation of a representation for concepts in simple noiseless feedback neural networks, described by the Hopfield model trained with examples has been studied in recent works [13–15]. It has been shown that the spurious mixture states always present in these networks play a crucial role in extracting meaningful information from the activity patterns to which the network is exposed in the training stage.

It turns out that a minimum number of examples should be taught to the network before it starts to generalize. It was found, in mean-field theory (MFT), that the generalization error of the noiseless network drops discontinously when a critical number of examples is presented [14]. This is in contrast to the behaviour found for single-layer feed-forward networks [8–11]. Early numerical simulations [15] failed to confirm this behaviour.

In biological systems some neurons may spontaneously become active without external stimulus. In order to mimic this, a stochastic behaviour may be introduced in an artificial neural network by defining a parameter β as an inverse temperature (T^{-1}). The natural question that arises is: how does the temperature (synaptic noise) affect the generalization ability? The introduction of temperature allows to compensate for an overemphasis on the examples in the training stage, as shown in the context of feed-forward networks [9, 10], but not yet exploited for attractor neural networks.

In the Hopfield model the states of the neurons are represented by Ising spins, $S_i = \pm 1, i = 1, ..., N$. Training the network with examples means that during the learning

0305-4470/93/163983+11$07.50 © 1993 IOP Publishing Ltd

stage, a set of s examples $\{\xi_i^{\mu\nu}\}$, $\nu = 1, ..., s$ of each concept $\{\xi_i^{\mu}\}$, $\mu = 1, ..., p$ is presented and stored in the network using the Hebb learning rule

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p}\sum_{\nu=1}^{s}\xi_i^{\mu\nu}\xi_j^{\mu\nu} \qquad (i \neq j). \tag{1}$$

The examples are assumed to be statistically independent and equally distributed random variables according to the probability distribution

$$P(\xi_i^{\mu\nu}) = b_1\delta(\xi_i^{\mu\nu} - 1) + b_2\delta(\xi_i^{\mu\nu} + 1) \tag{2}$$

where $b_1 = \frac{1}{2}(1 + \xi_i^{\mu}b)$, $b_2 = \frac{1}{2}(1 - \xi_i^{\mu}b)$ and $0 \leq b \leq 1$. The components of the concepts are statistically independent and chosen as $\xi_i^{\mu} = \pm 1$, with equal probability. This choice provides the simplest form for generalization of hierarchically organized patterns. The parameter b measures the correlation between the examples and the concepts so that the lower b the more difficult is the extraction of meaningful information from the examples.

Since the J_{ij} are symmetric, an energy function can be defined as

$$H = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}J_{ij}S_iS_j \tag{3}$$

which governs the dynamic retrieval process. The generalization performance is characterized by the overlap between a concept ξ_i^{μ} and a minimum (or near-minimum) state S_i of H

$$m^{\mu} = \frac{1}{N}\sum_{i=1}^{N}\xi_i^{\mu}S_i \qquad \mu = 1, \ldots, p\,. \tag{4}$$

Creating a representation for concepts amounts to generate a set of finite overlaps m^{μ}; the larger their size the greater the generalization ability of the network. Associated with these overlaps is the generalization error, defined as the Hamming distance

$$\epsilon^{\mu} = \tfrac{1}{2}(1 - m^{\mu}) \tag{5}$$

between a state S_i and a given concept. In the case of statistically independent concepts, as we have in this paper, one may focus on any one of them, say $\mu = 1$, in which case the generalization error is simply $\epsilon \equiv \epsilon^1$.

In this paper we study the effect of thermal noise on the generalization ability. The paper is organized as follows. In section 2 we consider the effect of temperature in MFT, in the cases $\alpha = p/N = 0$ and $\alpha \neq 0$, following standard procedures for attractor neural networks [2, 16, 17]. Explicit phase diagrams are obtained. In section 3 we present the results of numerical simulations for $\alpha \neq 0$ and compare them with the predictions of MFT and with former numerical work at $T = 0$ [15]. In section 4 we summarize our results.

2. Effect of noise in mean-field theory

Following Fontanari [14] we obtain the thermodynamic properties from the free energy associated to the Hamiltonian

$$H = -\frac{1}{2N}\sum_{\mu\nu}\left(\sum_i \xi_i^{\mu\nu}S_i\right)^2 + \sum_{i\mu}h^{\mu}\xi_i^{\mu}S_i \tag{6}$$

where the term including the field h^μ is introduced to compute the overlap m^μ in the limit $h^\mu \to 0$, according to

$$m^\mu = \left.\frac{\partial f}{\partial h^\mu}\right|_{h^\mu=0}. \tag{7}$$

2.1. *Finite number of concepts*

Due the simplicity of this case, in which $\alpha = 0$ we present just the final result obtained for the free energy density

$$f = \frac{1}{2}\sum_{\mu\nu}(m^{\mu\nu})^2 - \frac{1}{\beta}\left\langle \ln 2\cosh\left[\beta\left(\sum_{\mu\nu} m^{\mu\nu}\xi^{\mu\nu} - \sum_{\mu} h^\mu \xi^\mu\right)\right]\right\rangle \tag{8}$$

where $\langle\cdots\rangle$ denotes averages over the examples and over the concepts, in this order. The overlap with the examples, defined as

$$m^{\mu\nu} = \frac{1}{N}\sum_{i=1}^{N} \xi_i^{\mu\nu} S_i, \tag{9}$$

is obtained from the saddle-point equation

$$m^{\mu\nu} = \left\langle \xi^{\mu\nu} \tanh\left(\beta \sum_{\rho\sigma} m^{\rho\sigma}\xi^{\rho\sigma}\right)\right\rangle \tag{10}$$

and the overlaps with the concepts (generalization overlaps) obtained from (7) are given by

$$m^\mu = \left\langle \xi^\mu \tanh\left(\beta \sum_{\rho\sigma} m^{\rho\sigma}\xi^{\rho\sigma}\right)\right\rangle. \tag{11}$$

We consider a particular class of solutions of the form $m^{\mu\nu} = m^{1\nu}\delta_{\mu 1}$, with $m^{1\nu}$ given by [14]

$$m^{1\nu} = m^{11}\delta_{1\nu} + (1 - \delta_{1\nu})m_{s-1} \tag{12}$$

that has a bias for associative memory with any of the examples of a given concept (here $\mu = 1$). This kind of solution ensures that any other network behaviour will be a spontaneous property of the network and not a consequence of the particular class of chosen solutions. Note that any example can be selected as the first one, $\nu = 1$, since this solution is s degenerate. The equations for m^{11} and m_{s-1} are

$$m^{11} = \langle \xi^{11} \tanh[\beta(m^{11}\xi^{11} + m_{s-1}x_{s-1})]\rangle \tag{13}$$

$$m_{s-1} = \frac{1}{s-1}\langle x_{s-1} \tanh[\beta(m^{11}\xi^{11} + m_{s-1}x_{s-1})]\rangle \tag{14}$$

where $x_{s-1} = \sum_{\nu>1}\xi^{1\nu}$ is a random variable that follows a binomial probability distribution given by

$$P(x_{s-1}) = \begin{pmatrix} s-1 \\ k \end{pmatrix} b_1^k b_2^{s-1-k} \tag{15}$$

with $k = \frac{1}{2}(x_{s-1} + s - 1)$. We keep this distribution in place of the Gaussian appropriate for large s [14] in order to have reliable results also for small s. The averages in (13) and

(14) over ξ^{11}, x_{s-1} and ξ^1, in this order, yield

$$m^{11} = \frac{1}{2}\sum_{k=0}^{s-1}\binom{s-1}{k}\left[P_1(k)\tanh(\beta\Lambda_+) + P_1(k-1)\tanh(\beta\Lambda_-)\right] \tag{16}$$

$$m_{s-1} = \frac{1}{2(s-1)}\sum_{k=0}^{s-1}\binom{s-1}{k}(2k-s+1) \times\left[P_1(k)\tanh(\beta\Lambda_+) - P_1(k-1)\tanh(\beta\Lambda_-)\right] \tag{17}$$

where

$$P_1(k) = \left(\frac{1+b}{2}\right)^{k+1}\left(\frac{1-b}{2}\right)^{s-1-k} + \left(\frac{1+b}{2}\right)^{s-1-k}\left(\frac{1-b}{2}\right)^{k+1} \tag{18}$$

$$\Lambda_\pm = m^{11} \pm m_{s-1}(2k-s+1)\,. \tag{19}$$

Similarly, the generalization overlap m^1 that follows from (11) is given by

$$m^1 = \frac{1}{2}\sum_{k=0}^{s-1}\binom{s-1}{k}\left[P_2(k)\tanh(\beta\Lambda_+) + P_2(s-k-1)\tanh(\beta\Lambda_-)\right] \tag{20}$$

where

$$P_2(k) = \left(\frac{1+b}{2}\right)^{k+1}\left(\frac{1-b}{2}\right)^{s-1-k} - \left(\frac{1+b}{2}\right)^{s-1-k}\left(\frac{1-b}{2}\right)^{k+1}\,. \tag{21}$$

Equation (20) yields then the generalization error $\epsilon = (1-m^1)/2$.

Solving equations (13)–(21) for several temperatures and evaluating the corresponding free energy in each case, in order to make a stability analysis, yields the phase diagram shown in figure 1, for two values of b.

The retrieval phase (R) corresponds to asymmetric solutions in which the network retrieves a particular example so that $m^{11} \neq m_{s-1}$. For low temperatures $m^{11} \approx 1$, $m_{s-1} \approx b^2$ and $m_1 \approx b$, corresponding to a generalization error $\epsilon \approx (1-b)/2$. In the retrieval phase there is also the competing symmetric solution $m^{11} = m_{s-1}$, which is the

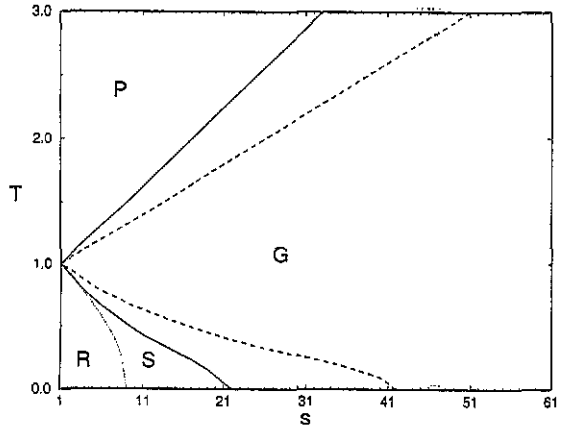

Figure 1. Phase diagram for s examples of a finite number of concepts ($\alpha = 0$) and two values of the correlation parameter: $b = 0.25$ (full lines) and $b = 0.20$ (dashed lines). The doted line separates the global stability regions for retrieval of examples (R) and symmetric mixture states (S) for $b = 0.25$. The generalization (G) and paramagmetic (P) phases are shown.

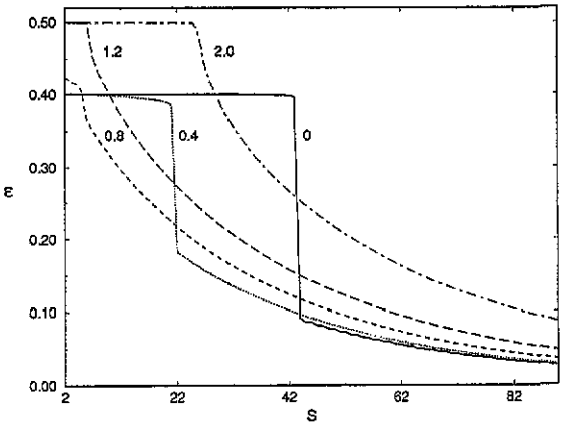

Figure 2. Generalization error as a function of the number of examples for $\alpha = 0$, $b = 0.2$ and several synaptic noise levels (values of T indicated on the curves).

more stable solution in the region S to the right of the dotted line in figure 1, for $b = 0.25$, while the retrieval solution becomes more stable to the left of that line. There exists a line of first-order phase transitions defining a critical number of examples s_c above which the network starts to generalize. Our $T = 0$ results of $s_c = 22$, for $b = 0.25$, and $s_c = 42$, for $b = 0.20$, are consistent with figure 2 of [14].

In the generalization phase (G) the network is not able to recover a particular example but it rather recognizes the common feature of the examples in the form of symmetric mixture states, $m^{11} = m_{s-1}$. Were it not for the simultaneous sharp increase of the overlap with the concept associated with those examples (which is what is meant by generalization) as shown in figure 2, one would have a state of perfect confusion. Our results for the generalization error at $T = 0$ is in agreement with figure 1 of [14].

Increasing the temperature we reach a paramagnetic phase (P) where the network neither generalizes nor retrieves the examples ($m^{11} = m_{s-1} = m^1 = 0$). The P–G transition is of second order with a phase boundary given by

$$T_0(b) = 1 + (s-1)b^2 \tag{22}$$

and the solutions in the generalization phase are the symmetric $m^{\mu\nu} = m_s\delta_{\mu 1}, \forall\nu$, with

$$m_s = (1/s)\langle x_s \tanh(\beta m_s x_s)\rangle \tag{23}$$

where

$$x_s = \sum_{\nu=1}^{s} \xi^{1\nu}. \tag{24}$$

Note, incidentally, as one would expect, that increasing b (i.e., the correlation between examples and concepts), the generalization phase grows at the expense of the paramagnetic and retrieval phases. Note also that the generalization performance of the network first improves with an increasing synaptic noise, in that a smaller number of examples has to be presented, and that generalization becomes harder when the noise level is above $T_0(b)$.

2.2. *Macroscopic number of concepts*

We deal here with the case where $\alpha = p/N$ is finite in the thermodynamic limit $N \to \infty$. In this case the network should create an extensive number of representations having access only to a finite number of examples. In order to compute the free energy density it is

necessary to resort to the replica method [17]. Allowing only for the examples $\{\xi^{l\nu}\}$ to condense, and assuming replica symmetry (RS), the average free energy density becomes through standard calculations

$$f = \frac{1}{2}\sum_{\nu}(m^{l\nu})^2 + \frac{\alpha r C}{2} - \frac{\alpha}{\beta}\ln G(q) - \frac{1}{\beta}\int_{-\infty}^{\infty} \mathrm{D}z\,\langle \ln[2\cosh(\beta\Delta)]\rangle \tag{25}$$

where

$$\ln G(q) = -\tfrac{1}{2}[(s-1)\ln(1-C(1-b^2)) + \ln(1-C(1-b^2+sb^2)) - \frac{\beta q s(1-C(1-b^2)(1-b^2+sb^2))}{(1-C(1-b^2))(1-C(1-b^2+sb^2))}] \tag{26}$$

in which

$$\Delta = z\sqrt{\alpha r} + \sum_{\nu} m^{l\nu}\xi^{l\nu} - h^l\xi^l \tag{27}$$

$$C = \beta(1-q) \tag{28}$$

$$\mathrm{D}z = \frac{\mathrm{d}z}{\sqrt{2\pi}}\exp(-z^2/2)\,. \tag{29}$$

The order parameters are obtained from the saddle-point equations (taking $h^\mu = 0$)

$$m^{l\nu} = \left\langle \int_{-\infty}^{\infty} \mathrm{D}z\,\xi^{l\nu}\tanh\left[\beta\left(\sqrt{\alpha r}z + \sum_{\nu} m^{l\nu}\xi^{l\nu}\right)\right]\right\rangle \tag{30}$$

$$q = \left\langle \int_{-\infty}^{\infty} \mathrm{D}z\,\tanh^2\left[\beta\left(\sqrt{\alpha r}z + \sum_{\nu} m^{l\nu}\xi^{l\nu}\right)\right]\right\rangle \tag{31}$$

$$r = sq\frac{[1-C(1-b^2)(1-b^2+sb^2)]^2 + (s-1)b^4}{[1-C(1-b^2)]^2[1-C(1-b^2+sb^2)]^2} \tag{32}$$

and, from (7)

$$m^l = \left\langle \int_{-\infty}^{\infty} \mathrm{D}z\,\xi^l\,\tanh\left[\beta\left(\sqrt{\alpha r}z + \sum_{\nu} m^{l\nu}\xi^{l\nu}\right)\right]\right\rangle. \tag{33}$$

Since the generalization ability is characterized by a symmetric solution ($m^{\mu\nu} = m_s\delta_{\mu 1}$), equations (30) and (31) become

$$m_s = \frac{1}{s}\left\langle \int_{-\infty}^{\infty} \mathrm{D}z\,x_s\tanh[\beta(\sqrt{\alpha r}z + m_s x_s)]\right\rangle \tag{34}$$

$$q = \left\langle \int_{-\infty}^{\infty} \mathrm{D}z\,\tanh^2\beta(\sqrt{\alpha r}z + m_s x_s)\right\rangle. \tag{35}$$

The averages over the examples and the concepts yield

$$m_s = \frac{1}{2s}\int_{-\infty}^{\infty} \mathrm{D}z\sum_{k=0}^{s}\binom{s}{k}(2k-s)P_+(k)\tanh(\beta\Delta_s) \tag{36}$$

$$q = \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{D}z\sum_{k=0}^{s}\binom{s}{k}P_+(k)\tanh^2(\beta\Delta_s) \tag{37}$$

where

$$P_{\pm}(k) = \left(\frac{1+b}{2}\right)^{k}\left(\frac{1-b}{2}\right)^{s-k} \pm \left(\frac{1+b}{2}\right)^{s-k}\left(\frac{1-b}{2}\right)^{k} \tag{38}$$

and

$$\Delta_s = \sqrt{\alpha r}z + m_s(2k-s)\,. \tag{39}$$

The overlap with the concepts becomes

$$m^1 = \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{D}z \sum_{k=0}^{s} \binom{s}{k} P_{-}(k)\tanh(\beta\Delta_s)\,. \tag{40}$$

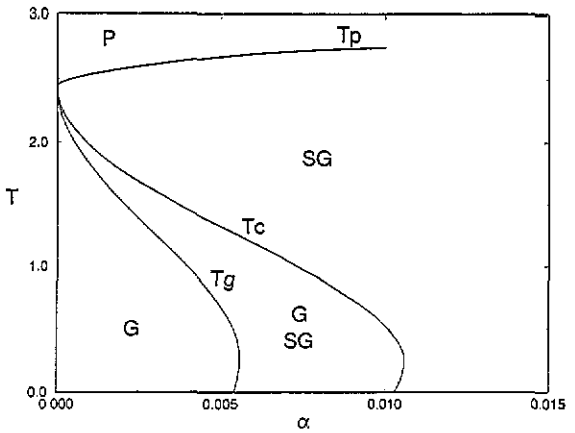

Figure 3. Finite-α phase diagram for $b = 0.4$ and $s = 10$ examples. A globally stable spin-glass (SG) phase appears above T_c, and a region of locally or globally stable generalization states (G) exists above or below T_g, respectively.

By solving these equations we obtain the (α, T) phase diagram shown in figure 3 for correlation parameter $b = 0.4$ and $s = 10$ examples. There exists a critical temperature T_p above which $m_s = 0$ and $q = 0$, characterizing the paramagnetic phase (P). Below this temperature and above the curve T_c there is a spin glass phase (SG) where $m_s = 0$ and $q \neq 0$, in which the solutions have no correlation neither with the concepts nor with the examples. Here the network is in a state of total confusion. A finite overlap m_s characterizing a G phase appears discontinuously at the phase boundary T_c. However, the SG solution has the lower free energy until the lower phase boundary T_g is reached, where the G phase is the most stable one, in which the states are strongly correlated with the concepts. In this region the order parameters are $m_s \neq 0$ and $q \neq 0$. An interesting feature is the reentrant shape in the spin glass phase below a small temperature also found recently by Naef and Canning [18], within a RS calculation. The reentrance does not appear with replica symmetry-breaking (RSB), in which case the lowest part of the phase boundary is moved slightly to the right [19].

Our result, figure 3, at $T = 0$ for the critical $\alpha_c \simeq 0.01$, beyond which generalization is no loger possible, is in agreement with Fontanari's work [14]. Indeed, the ratio $\alpha_c/\alpha_0 \simeq 0.07$ is the same as that of figure 4 of [14], where $\alpha_0 \approx 0.138$ is the storage capacity of the standard Hopfield model.

The role of temperature for finite α is different from that found in the $\alpha = 0$ case. Now it is no longer possible to control the desired creation of concepts just by tuning the temperature for a fixed number of examples. The spin-glass phase that competes with the generalization

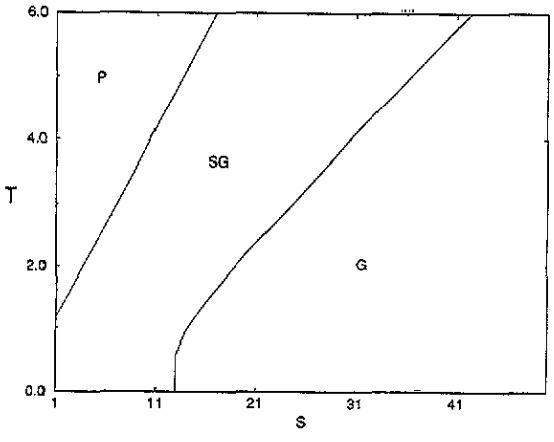

Figure 4. Phase diagram for $\alpha = 0.03125$ and $b = 0.5$ showing the paramagnetic (P), spin-glass (SG) and the generalization (G) phases.

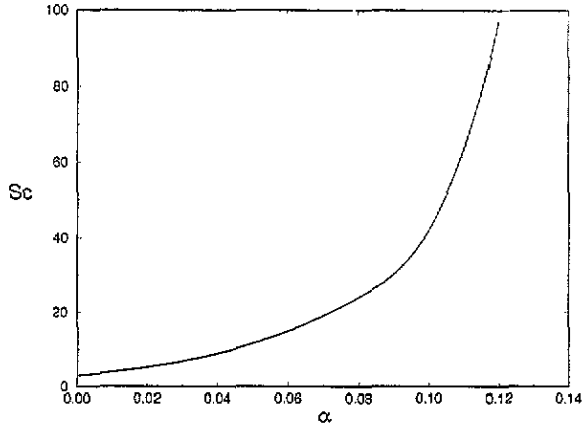

Figure 5. MFT results for the critical number of examples as a function of α for $b = 0.6$.

phase is stabilized by the Gaussian noise due to the uncondensed examples $\{\xi^{\mu\nu}\}, \mu > 1$. Nevertheless, the network ability to generalize is still fairly robust against synaptic noise. On the other hand, the presentation of a sufficiently large number of examples is still crucial in creating a representation for the concepts, as shown in figure 4. There is a critical number of exemples, s_c, that has to be reached for generalization to take place, for a given α and b. At $T = 0$, this value is $s_c = 13$ which, together with the ratio $\alpha_c/\alpha_0 \simeq 0.22$, is in agreement with the results of figure 4 of Fontanari's work [14]. A direct check of the results of our equations against the curves of figure 4 of [14], for $b = 0.2$ and 0.4 was also carried out. Our analytical results for the generalization error can be found below.

The generalization error decreases with the number of examples, for $s > s_c$, according to a power-law behaviour, in consistency with the result of Miranda [15], who also finds that the critical number of examples grows exponentially with the noise level in the examples (as compared to the concepts) shown to the network and grows linearly with α. In contrast we find, for $T = 0$ and $\alpha = 0.05$, that $s_c \sim b^{-\gamma}$ with $\gamma \simeq 2.99$, and that the critical number of examples varies with α in a way shown in figure 5, instead of the linear law $s_c \sim \alpha$ [15]. These results follow from fits on the numerical solution of the equations for $s = s_c$.

The calculations within this paper are restricted to RS because RSB has only a very small effect at very low noise level in a neural network with a Hebbian learning rule, equation (1), [17] and we are mainly interested in establishing the effects of a finite noise level. To check

on the replica method (even with standard RSB procedures [20] it has not been fully justified) we resort next to numerical simulations.

3. Numerical simulations

We verify the predictions of MFT discussed in the previous section by means of numerical simulations for finite α, at zero and finite temperature. For zero temperature the dynamic evolution of the system consists in updating the neurons at each time step according to

$$S_i(t+1) = \mathrm{sgn}[h_i(t)] \tag{41}$$

where $h_i(t) = \sum_{j \neq i} J_{ij} S_j(t)$ is the local field acting in the neuron i. This procedure decreases the system energy every time that one spin is updated so that $S_i(t+1)h_i(t+1) > 0$. Storage of the synaptic matrix J_{ij} is very expensive in terms of computer memory, so we store only the examples and the concepts. The system dynamics can be described in terms of the overlaps $m^{\mu\nu}$ which are calculated with the network in an initial state that coincides with an example and are updated at every time step by one spin-flip. For this purpouse it is convenient to express the Hamiltonian (3) in terms of the overlaps

$$H = -\frac{N}{2} \sum_{\mu\nu} (m^{\mu\nu})^2 + \frac{1}{2} ps. \tag{42}$$

Equation (42) is then used to perform the Monte Carlo simulations in standard way [21].

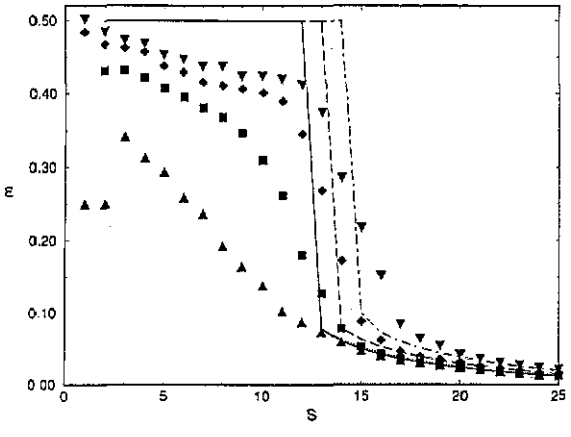

Figure 6. Generalization error curves obtained from simulations for $\alpha = 0.0312\,5$, $b = 0.5$ at $T = 0$ (triangles), $T = 0.4$ (squares), $T = 0.8$ (diamonds) and $T = 1.2$ (inverted triangles). The lines correspond to the analytical results for $T = 0$ (full line), $T = 0.4$ (dotted line), $T = 0.8$ (dashed line) and $T = 1.2$ (dot-dashed line).

The numerical results for the generalization error for finite α and various T are shown in figure 6, together with analytical results, for a network with $N = 2048$, $b = 0.5$ and $\alpha = 0.0312\,5$, setting the initial configuration as an example. Note first, that there is a spin glass regime that competes with generalization starting with the SG for $3 \leq s \leq s_c$, where s_c is the critical number of examples. The SG phase appears, quite early due to the relatively large value of α, where there is an exponentially large number of metastable states. Since these states are stable against single spin flips, the system becomes trapped in these states resulting in non-zero overlaps $m^{\mu\nu}$ and m^{μ}. For this reason the generalization error shown by the numerical results does not reach the theoretical equilibrium prediction, $\epsilon = 0.5$. This is also due to finite system size [22].

At $s \simeq s_c$ there is a transition to the generalization regime, which seems to follow the analytical prediction except for the very lowest noise levels. The critical generalization

error at $T = 0$, $\epsilon_c \simeq 0.08$, which is the lower value of ϵ on the jump, is in agreement with the results of figure 5 of [14]. An accurate estimate of the critical number of examples for generalization from numerical simulations is strongly dependent on system size, (cf [22]). The size of our network ($N = 2048$) is somewhat restrictive and that is even more so in Miranda's case, where $N = 512$ neurons. Finally, at $T = 0$ (upper triangles) there is a small plateau at $\epsilon = 0.25$ for $s \leq 2$ where the examples are retrieved.

4. Conclusions

In this paper we studied, by means of MFT and numerical simulations, the effects of synaptic noise on the generalization ability of the Hopfield model of an attractor neural network trained with examples of the patterns for which a representation is to be created, when the training takes place according to the Hebb learning rule. We confirm that symmetric mixture states play a relevant role in generalization in that they serve to extract the common features of the examples that lead to a representation for the concepts.

We find that synaptic noise has quite different effects depending on whether the number of concepts to be created is macroscopic or not. If only a finite number of concepts is considered, synaptic noise helps to generalize, up to a fairly high noise level, but this is no longer the case for a macroscopic number of patterns. Nevertheless, the network is still able to generalize in the presence of noise as long as the storage ratio α remains below the critical saturation value.

The finite-α phase diagrams are in accordance with what one expects. A better understanding of the role of synaptic noise, however, requires a study of the metastable states, which is currently being carried out and about which we expect to report in a forthcoming publication.

Acknowledgments

It is a pleasure to thank J F Fontanari for stimulating discussions and his continued interest during the course of this work. We thank the Supercomputing Center of the Universidade Federal do Rio Grande do Sul (CESUP-UFRGS) for use of the Cray YMP-2E. This work was supported, in part, by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos (FINEP) and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS).

References

[1] Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554
[2] Amit D J 1989 *Modeling Brain Function* (Cambridge: Cambridge University Press) p 272
Hertz J, Krogh A and Palmer R 1991 *Introduction to the Theory of Neural Computation* (Reading, MA: Addison-Wesley)
[3] Denker J, Schwartz D, Wittner B, Solla S, Howard R, Jackel L and Hopfield J J 1987 *Complex Syst.* **1** 877
[4] Patarnello S and Carnevali P 1987 *Europhys. Lett.* **4** 503
Carnevali P and Patarnello S 1987 *Europhys. Lett.* **4** 1199
[5] Tishby N, Levin E and Solla S 1989 *Proc. Int. Joint Conf. on Neural Networks*
[6] Bacci S, Mato G and Parga N 1990 *Neural Networks and Spin Glasses* ed W K Theumann and R Köberle (Singapore: World Scientific) p 219; 1990 *J. Phys. A: Math. Gen.* **23** 1801
Gutfreund H 1988 *Phys. Rev.* A **37** 570
[7] Del Giudice P, Franz S and Virasoro M A 1989 *J. Physique* **50** 121
[8] Hansel D and Sompolinsky H *Europhys. Lett.* **11** (1990) 687

[9] Györgyi G and Tishby N 1990 *Neural Networks and Spin Glasses* ed W K Theumann and R Köberle (Singapore: World Scientific) p 3
[10] Györgyi G 1990 *Phys. Rev. Lett.* **64** 2957
[11] Seung H S, Sompolinsky H and Tishby N 1992 *Phys. Rev.* A **45** 6056
[12] Meir R and Fontanari J F 1992 *J. Phys. A: Math. Gen.* **25** 1149
[13] Fontanari J F and Meir R 1989 *Phys. Rev.* A **40** 2806
[14] Fontanari J F 1990 *J. Physique* **51** 2421
[15] Miranda E N 1991 *J. Physique*. **I 1** 999
[16] Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev.* A **32** 1007
[17] Amit D J, Gutfreund H and Sompolinsky H 1987 *Ann. Phys., NY* **173** (1987) 30
[18] Naef J-P and Canning A 1992 *J. Physique* I **2** 147
[19] Crisanti A, Amit D J and Gutfreund H 1986 *Europhys. Lett.* **2** 337
[20] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific) p 23
[21] Heermann D W 1986 *Computer Simulation Methods in Theoretical Physics* (Berlin: Springer) p 68
[22] Branchtein M C and Arenzon J J 1993 *Preprint* IF-UFRGS