Competition between Hopfield and symmetry transform interactions in a neural net

1991 J. Phys. A: Math. Gen. 24 4445

(http://iopscience.iop.org/0305-4470/24/18/028)

J. Phys. A: Math. Gen. 24 (1991) 4445-4457. Printed in the UK

Competition between Hopfield and symmetry transform interactions in a neural net

M R Evans, D J Wallace and C Zhan

Physics Department, University of Edinburgh, Mayfield Rd, Edinburgh EH9 3JZ, UK

Received 29 November 1990, in final form 12 June 1991

Abstract. We consider a neural network proposed by Coolen and Kuijk that has interactions composed of two competing elements. The first is the Hopfield interaction for recall of a set of stored patterns. The second is interactions between pairs of sites arranged so that the configuration undergoes a symmetry transformation at each parallel update. First we consider the effect of the symmetry-transform interactions alone. We show that for sequential updating the symmetry transformation is no longer carried out faithfully, but rather the spin configuration tends to a symmetry invariant. In order to understand how the retrieval phase of the Hopfield model is disrupted by the symmetry-transform interactions we perform a replica symmetric analysis. We demonstrate that the symmetry-transform interactions generate a noise very similar to that of random external fields on the memory states. The phase diagram suggests the possibility of symmetry-invariant recognition for an extensive number of patterns and an optimal value for the symmetry-transform interaction strength. We present numerical simulations of the model under parallel dynamics to confirm these predictions.

1. Introduction

The Hopfield model has been shown to be a robust model for pattern recognition [1-4] although it is not optimal with regard to storage capacity [5] or for the purpose of providing transformation invariant pattern recognition. In the model the pattern is considered as an N-bit vector $\{\xi_i^\mu\}$ where ξ_i^μ is the value of the spin in pattern μ at spin i. The robustness of the model stems from the co-operative nature of the recall process that has each spin interacting with all the others through synaptic connections. This gives the model a tolerance to various noises such as temperature [2], synaptic clipping and synaptic destruction [4]. Patterns are stored by giving the synaptic connections between spins values according to the Hopfield rule:

$$H_{ij} = \frac{1}{N} \sum_{\mu=1}^{P} \xi_i^\mu \xi_j^\mu. \tag{1}$$

The recall is performed by presenting a configuration $\{S_i\}$ to the network. If this configuration has macroscopic overlap with a pattern 1 then the order parameter m associated with this overlap is non-zero where

$$m = \frac{1}{N} \sum_{i=1}^{N} \xi_i^1 \langle S_i \rangle. \tag{2}$$

0305-4470/91/184445+13$03.50

The field received at a site i is then composed of a signal and a noise

$$h_i = \sum_{j=1}^{N} H_{ij} S_j \tag{3}$$

$$= \xi_i^\mu m + \frac{1}{N} \sum_{\mu>1, j\neq i} \xi_i^\mu \xi_j^\mu S_j. \tag{4}$$

The first, signal, term is linear in the overlap order parameter m and consequently the larger the value of m the more likely the pattern is to be recalled.

If we consider this process as adjudging whether a given configuration is correlated enough with a stored pattern then the measure of correlation used is the Hamming distance or number of incorrect sites. This ignores the possibility of the configuration being some simple transformation of a stored pattern. For example one would like to be able to recognize a reflected or translated version of a stored pattern. This corresponds to sculpting more complicated basins of attraction on the energy landscape. For the example of recognizing a stored pattern and its reflection as the same input one in fact requires a disjoint domain of attraction with two basins. One basin is the usual Hopfield model basin of attraction comprising configurations a small Hamming distance away from the stored pattern. The second basin should contain configurations a small Hamming distance away from the reflected pattern. In order to sculpt a channel from the latter basin to the former, the Hopfield model must be modified.

Recently several schemes have been proposed to perform such a task. One proposal involves the use of dynamic connection strengths [6, 7]. Another option is to pre-process the image before presentation to the final network [8, 9]. A third possibility is not to have preprocessing but to have competing directions within the configurational flow produced by the dynamics of the network. The ideal that is sought with this approach is a network that symmetry-transforms the presented configuration until a macroscopic overlap with one of the stored patterns is found. At this point in the configurational flow the Hopfield interactions predominate and the pattern is recalled. In effect the network performs its own pre-processing.

To this end several mechanisms have been proposed. Dotsenko [10] has used modifiable thresholdings as the component of the model that causes the symmetry-transformation to occur. Coolen and Kuijk [11] have shown that the connections can be trained by example from pairs of configurations and their symmetry-transforms to perform the desired symmetry transformation.

In the present work we consider the latter case [11]. The total connection strengths are given by

$$J_{ij} = H_{ij} + T_{ij}. \tag{5}$$

The interactions that perform the transformation (T_{ij}) only couple a site and its image site under the transformation. We will use the term 'local' to refer to the fact that the lattice is fully connected so that the sites can be rearranged to leave a site and its image adjacent to each other. In this sense the symmetry transform interaction is local as opposed to H_{ij} where all sites have interactions with each other and the interaction is therefore long-range. Our aim in this paper is to demonstrate that this interaction scheme is effective for an extensive number of stored patterns. In order to provide a theoretical framework we will study how the local symmetry-transform interactions acts as a noise upon the long range Hopfield interactions. In particular we wish to discover whether the retrieval phase of the Hopfield model persists in the presence of

the symmetry transform interactions when an extensive number of patterns are stored. If this were the case, then it would suggest that symmetry invariant pattern recognition is possible for an extensive number of patterns.

In section 2 we examine the symmetry transform interaction and show that it is only for parallel dynamics that the symmetry transform is faithfully carried out; for sequential dynamics the configurational flow is towards a symmetry invariant. This suggests that parallel dynamics are more suitable for symmetry invariant pattern recognition. However sequential dynamics allows us to write down a configurational energy and in section 3 analyse, within the replica symmetric ansatz, how the retrieval phase of the Hopfield model is affected by the symmetry transformation interactions. We find that the overall effect is very close to that caused by a random external field at each site. The main result of the section is that, for a range of symmetry transform interaction strengths, there is a retrieval phase for an extensive number of stored patterns. In section 4 we investigate whether the potential for symmetry invariant recognition of an extensive number of patterns is realized under parallel dynamics. We present extensive numerical simulations which confirm that this is the case. These simulations suggest an optimal value for the symmetry transform interaction strength which agrees well with a theoretical estimate relying on the results of section 3.

2. The symmetry transform

In order that the symmetry transform interactions be as 'local' as possible, in the sense described in section 1, we choose a Z_2 symmetry. This will also give the convenient property of symmetric interactions $J_{ij}=J_{ji}$. In the transformation each spin is acted upon by an element π of the group. The spin is either mapped onto itself by the identity element so that $\pi(i)=i$ or mapped onto another spin $\pi(i)\neq i$. In the latter case the Z_2 constraint $\pi(\pi(i))=i$ ensures that the pair of spins at i and $\pi(i)$ are interchanged by the transformation. For simplicity we will set the number of spins mapped onto themselves to zero.

A simple interaction [11] that produces the above symmetry transformation is given by

$$T_{ij}=a\delta_{i,\pi(j)} \qquad i\neq j \tag{6}$$

$$=0 \qquad i=j. \tag{7}$$

Thus (6) couples sites to their image sites under the transformation. In order to see that a symmetry transformation is indeed performed by the interaction (6), consider the local field produced at a site i by this interaction at time t:

$$h_i(t)=\sum_{j\neq i} T_{ij}S_j(t)$$

$$=aS_{\pi(i)}(t). \tag{8}$$

The field at each site has the sign of the spin at the image of the site and so, at least at zero temperature under parallel dynamics, the transformation will be faithfully performed. For serial dynamics however, a single updating sweep allows the system to converge to a transformation invariant configuration. To see this, consider a pair of sites i and $\pi(i)$. If starting from time t, i is visited first in the updating sequence at time t_1, then the spin at i will be updated to $S_{\pi(i)}(t)$ so that $S_i(t_1)=S_{\pi(i)}(t)$. When

the site $\pi(i)$ is visited at a later time t_2, the local field will be $h_{\pi(i)}(t_2) = aS_i(t_2) = aS_i(t_1) = aS_{\pi(i)}(t)$ so that the spin at $\pi(i)$ will already be aligned to its local field. Both spins then end up taking the value $S_{\pi(i)}(t)$. In contrast if $\pi(i)$ is visited first both spins will end up taking the value of $S_i(t)$. Clearly the order of updating within the sweep determines the final configuration but whatever order is chosen the final configuration will be symmetry invariant. It is now apparent that serial and parallel updating give contrasting dynamic tendencies: equation (8) showed that parallel dynamics faithfully performs the symmetry-transformation; sequential dynamics drives the configuration into a symmetry invariant.

This point is rather interesting as it implies that the two different dynamics endow the model with rather different properties. Although we are not suggesting that this model is of particular biological relevance, we may recall that in the neural interpretation the two dynamics represent different degrees of synchronization. This may then be an indication that synchronization is a parameter that could be utilized in neural information processing.

3. Disruption of the retrieval phase

The main purpose of this paper is to demonstrate that symmetry invariant recognition is possible for an extensive number of patterns. A prerequisite for this is that the retrieval phase of the Hopfield model should persist in the presence of the symmetry transform interactions. Although we shall argue in section 4 that parallel dynamics are more suitable for invariant pattern recognition, a study of the model defined by random sequential dynamics will allow a quantitative analysis of when the retrieval phase is destroyed by the symmetry-transform interactions. In particular this will allow us to demonstrate the possibility of symmetry-invariant pattern recognition for an extensive number of stored patterns.

The full model with random sequential dynamics has configurational energy

$$E = -\tfrac{1}{2}\sum_{i\neq j} H_{ij}S_iS_j - \frac{a}{2}\sum_i S_iS_{\pi(i)}. \tag{9}$$

To begin the calculation we must discuss the relevant order parameters. A classification of sites relevant to the symmetry transformation of our nominated pattern will clarify the choice of order parameters.

1. The sites unchanged by the transformation:

$$i = \pi(i). \tag{10}$$

2. Pairs of sites that are interchanged by the transformation and which take the same value in the pattern:

$$i \neq \pi(i) \qquad \text{and} \qquad \xi_i^1 = \xi_{\pi(i)}^1. \tag{11}$$

3. Pairs of sites that are interchanged by the transformation and which take values of opposite sign in the pattern:

$$i \neq \pi(i) \qquad \text{and} \qquad \xi_i^1 = -\xi_{\pi(i)}^1. \tag{12}$$

We assume that the nominated patterns are random

$$P(\xi) = \tfrac{1}{2}(\delta(\xi-1) + \delta(\xi+1)) \tag{13}$$

and also disregard the possibility of sites being mapped onto themselves, so that there are no sites of class 1. However the equations can be easily generalized to include such a possibility. The overlap order parameters for our remaining two classes of sites are given respectively by

$$m_2 = \frac{1}{N}\sum_i \langle S_i\rangle(\xi_i^1 + \xi_{\pi(i)}^1) \tag{14}$$

$$m_3 = \frac{1}{N}\sum_i \langle S_i\rangle(\xi_i^1 - \xi_{\pi(i)}^1) \tag{15}$$

so that

$$m = \tfrac{1}{2}(m_2 + m_3). \tag{16}$$

The distinction between the two overlap parameters is that m_2 measures the overlap of the configuration with the pattern at sites where the pattern is symmetric; m_3 measures the overlap at sites where the pattern is antisymmetric.

We must also consider the ordering due to the symmetry transform interaction. For this purpose we introduce

$$g = \tfrac{1}{2}(g_2 + g_3) \tag{17}$$

where

$$g_2 = \frac{1}{N}\sum_i \langle S_i S_{\pi(i)}\rangle|\xi_i^1 + \xi_{\pi(i)}^1| \tag{18}$$

$$g_3 = \frac{1}{N}\sum_i \langle S_i S_{\pi(i)}\rangle|\xi_i^1 - \xi_{\pi(i)}^1|. \tag{19}$$

The order parameter g measures the symmetry invariance of the system. The two orders, overlap with a stored pattern and symmetry invariance, compete directly in sites of class 3, where the spins in pattern 1 at the site and image site are of opposite sign.

The mean field theory of the Hopfield model [2] is now a well trodden path. In appendix 1 we indicate how the calculation has to be modified to deal with the local symmetry transform interactions and derive zero temperature mean-field equations. The retrieval phase of the model is characterized by the existence of a solution of (40)–(44) of the form

$$m_2 > 0 \qquad m_3 > 0 \qquad m_2 > m_3$$

$$g_2 > 0 \qquad g_3 < 0 \qquad g > 0.$$

In the non-retrieval phase we only have a solution

$$m = m_2 = m_3 = 0$$

$$g_2 > 0 \qquad g_2 = g_3.$$

The non-retrieval phase has a large degree of symmetry under the transformation $i \to \pi(i)$; we shall refer to it as the symmetric phase to contrast with the usual spin glass non-retrieval phase, although the symmetry may not be total ($g < 1$). In figure 1 the phase boundary is plotted in the space of $a - \alpha$. At $a = 0$ we have $\alpha_c = 0.138$ recovering the usual Hopfield case. As a increases α_c decreases until at $a = 1.0$ we have $\alpha_c = 0$, showing that at this value of a there is no retrieval phase.

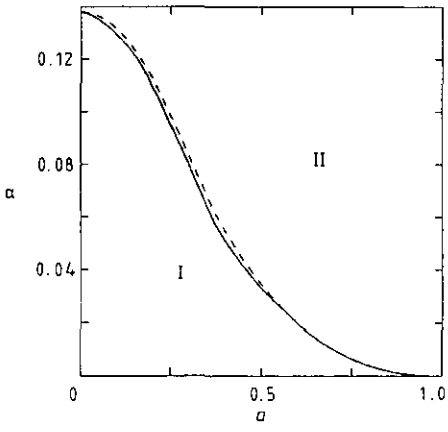

Figure 1. Phase diagram for the Hopfield model with: full curve, Z_2 symmetry interactions; dashed curve, random external fields. The retrieval (I) and non-retrieval (II) phases are described in the text.

In figure 2 we plot the overlap order parameters within the retrieval phase. One should note that $m_2 > m_3$, indicating that the quality of retrieval at sites where the pattern is symmetric is superior to the quality at sites where the pattern is antisymmetric. This also implies that there is some symmetry invariance in the phase ($g > 0$), however this effect is small ($g \ll 1$) so that the retrieval phase is not significantly altered from the Hopfield case. One may summarize figures 1 and 2 by saying that the two interaction components compete rather than co-operate.

3.1. The Hopfield model with random external fields

As we are interested in how the symmetry transform interactions act as a noise upon the Hopfield interactions it is useful to compare the phase diagram resulting from equations (40)–(44) to a phase diagram resulting from a simpler form of noise. Equation (8) shows that addition of symmetry interactions results in an additional component

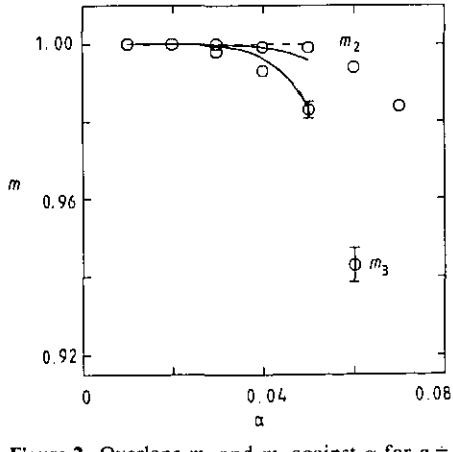

Figure 2. Overlaps m_2 and m_3 against α for $a = 0.4$; curves are as in figure 1. Symbols are results of simulations for network of 256 spins. Error bars are shown only when they are larger than the symbol size.

of magnitude a to the local field generated by the Hopfield interactions. The sign of this field component is determined by the direction of the spin at the image site. A simpler form of noise results if the signs of these field components are assigned randomly. For this model, which we shall refer to as the Hopfield model with random external fields, the configurational energy is given by

$$E = -\tfrac{1}{2}\sum_{i\neq j} H_{ij}S_iS_j - a\sum_i \zeta_i S_i \tag{20}$$

where each ζ is selected randomly according to the distribution

$$P(\zeta) = \tfrac{1}{2}(\delta(\zeta-1)+\delta(\zeta+1)). \tag{21}$$

The random fields are then quenched in contrast to the fields in (8) which are 'annealed' in character. The order parameters m_2 and m_3 now measure the overlap at sites where the pattern spin is in the same and opposite direction, respectively, as the random external field. We have

$$m_2 = \frac{1}{N}\sum_i (\xi_i+\zeta_i)\langle S_i\rangle \tag{22}$$

$$m_3 = \frac{1}{N}\sum_i (\xi_i-\zeta_i)\langle S_i\rangle. \tag{23}$$

The zero temperature mean-field equations are given in appendix 2.

The phase diagram for this model is also shown in figure 1. The non-retrieval phase is where the only solution to the equations gives $m_3 < 0$ indicating that the spins are aligned to the random external fields rather than to a stored pattern. If one ignores the physical difference in the non-retrieval phase then the two phase diagrams are strikingly similar. To investigate the extent of the similarity figure 2 compares the solutions of the zero temperature mean-field equations for the two models. The m_2 curves are almost identical, so much so that the random external field curve is obscured. For m_3 however, the random external field model gives a higher value than the symmetry-transform model. This suggests that at sites of class 3 the symmetry-transform interactions act as a stronger noise than random external fields on the Hopfield interactions. To summarize, it appears that the zero temperature mean-field equations for the Hopfield model with random external fields approximate rather well the far more complicated equations for the Hopfield model with symmetry interactions.

4. Parallel dynamics and invariant pattern recognition

In this section we will present numerical results of the performance of the model in recognition of transformed patterns. The analysis of the previous section provides a theoretical framework within which the numerical results may be understood as demonstrating the main finding of the paper: the model can recognize an extensive number of transformed patterns. Before presenting the simulations we will argue that parallel dynamics should be used for this purpose.

We showed in section 2 that using the symmetry transform under parallel and sequential dynamics gives considerably different configurational flows. If we consider the simple model of section 2, impose a transformed version of a stored pattern on the network and iterate, then for serial dynamics the direction of the configurational flow induced by the symmetry transform interaction is towards a symmetrized version

of the pattern. In terms of the overlap order parameters the transformed pattern is given by $m_2 = 1$, $m_3 = -1$ and $m = 0$; the symmetrized version is given by $m_2 = 1$, $m_3 = 0$ and $m = 0.5$. Thus starting from the transformed exact pattern we reach a configuration which is the pattern with 25% noise. This contrasts with parallel dynamics where after one time step the exact pattern would be obtained. In the light of this we consider the use of serial dynamics less suitable for invariant pattern recognition than the use of parallel dynamics.

However we have seen that for random sequential dynamics the retrieval phase persists in the presence of the symmetry transform interactions for an extensive number of stored patterns. In order to verify that this is also the case for parallel dynamics, we will present numerical simulations of the model. First we will use the results from section 3 to demonstrate the possibility of symmetry-invariant recognition for an extensive number of patterns under parallel dynamics, and to estimate the optimal symmetry transform interaction strength.

If parallel dynamics are chosen one can derive exact equations for the first time step of the iteration [14].

$$m_2(1) = \tfrac{1}{2}(1+m_2)\,\mathrm{erf}\left[\frac{m+a}{\sqrt{2\alpha}}\right] + \tfrac{1}{2}(1-m_2)\,\mathrm{erf}\left[\frac{m-a}{\sqrt{2\alpha}}\right] \tag{24}$$

$$m_3(1) = \tfrac{1}{2}(1+m_3)\,\mathrm{erf}\left[\frac{m-a}{\sqrt{2\alpha}}\right] + \tfrac{1}{2}(1-m_3)\,\mathrm{erf}\left[\frac{m+a}{\sqrt{2\alpha}}\right]. \tag{25}$$

In these equations the order parameters on the RHS are evaluated at $t = 0$. Starting from the symmetry-transformed pattern ($m_2 = 1$, $m_3 = -1$) we find

$$m_2(1) = m_3(1) = m(1) = \mathrm{erf}\left[\frac{a}{\sqrt{2\alpha}}\right]. \tag{26}$$

With a high enough value of a the pattern will be accurately reproduced after one parallel iteration. However if a is too large then at the second time step the transformed pattern will be reproduced.

In order to gauge the optimum value of a we can use the following guidelines:

$$a > \sqrt{2\alpha} \tag{27}$$

$$\alpha < \alpha_c(a) \tag{28}$$

where $\alpha_c(a)$ is the phase boundry plotted in figure 1. The first condition comes from the requirement that the transformed pattern be mapped accurately onto the pattern after one parallel update; the second condition comes from the requirement that the symmetry interactions should not disrupt the retrieval phase of the Hopfield model. Although we are considering parallel dynamics, the value of α_c for random sequential dynamics should give a good approximation as to when the retrieval phase is destroyed. To determine the optimal value of a from (27) and (28) one simply searches for the maximum value of $\alpha_c(a)$ subject to the constraint (27). These approximate arguments suggest an optimum value of $a = 0.4$ at which $\alpha_c = 0.06$. These numbers indicate that if we desire invariant pattern recognition, then although the maximum capacity of the network is reduced from the Hopfield case, we can still store an extensive number of patterns.

In figure 3 the results of numerical simulations of invariant pattern recognition using parallel dynamics are presented. The simulations were carried out on the Edinburgh Concurrent Supercomputer (a Meiko Computing Surface) in the OCCAM

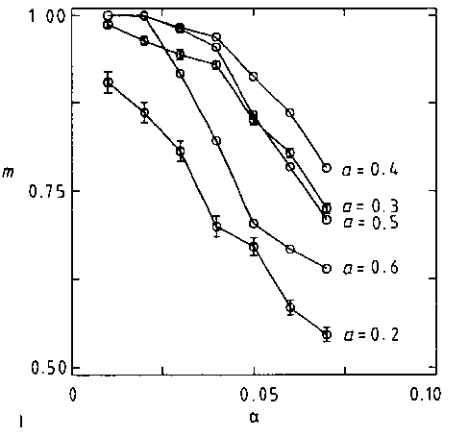

Figure 3. Simulation results for recognizing transformed patterns under parallel dynamics. Error bars are shown only when they are larger than the symbol size. The lines are drawn solely in order to guide the eye between simulation points for constant a.

language. The program can be run on any number of transputers with any number of spins. All the simulation results presented here used a network of 256 spins. Each point in the figures is an average over 25 sets of nominated patterns. In these simulations the transformed pattern was presented to the network and 20 parallel iterations were performed. The final overlaps were then calculated. The even number of iterations is convenient because if the symmetry transform interactions dominate and the configuration is symmetry transformed at each time step, then after an even number of iterations the configuration will return near to the transformed pattern, which has small overlap with the nominated pattern. We also found that 20 iterations were enough to ensure that a fixed point was reached if one existed.

Figure 3 indicates the existence of an optimal value of a because the curves rise and fall, to some extent, as a increases. At $a = 0.2$ one may note that even for very low α the recall of the transformed pattern is not good, indicating that the dynamic tendency to transform the pattern is not strong enough. As a is increased through 0.3 and 0.4 the retrieval quality at all α increases. For a equal to 0.5 the retrieval quality begins to fall. Finally for $a = 0.6$ the recall is excellent at $\alpha < 0.02$, but deteriorates rapidly as α increases. This reflects the disruption of the retrieval attractors at low α, for high a. The optimal value appears to be $a = 0.4$, which gives the highest retrieval quality at all α. This is in agreement with the prediction from (27), (28). However the retrieval quality m appears to deteriorate when α passes 0.05, which is lower than the α_c predicted from (27), (28).

5. Discussion

In this work we have investigated the Hopfield model with Z_2 symmetry transform interactions with two aims in mind. Firstly we have investigated how the local symmetry transform interaction acts as a noise upon the long range Hopfield interaction. Secondly we have demonstrated that symmetry invariant pattern recognition is possible with this simple model, for an extensive number of stored patterns.

The replica symmetric mean-field theory of section 3 is in fact a solution of a model with both long range and local interactions. We have found that in the ordered phase of the long range interactions (the retrieval phase) the local interactions act as a noise. The effect of this noise is very similar to imposing random external fields at each site. This supplements the results of Sompolinsky [4] where Gaussian noise in synapses was examined.

With regard to symmetry invariant pattern recognition it appears that sequential updating is unsuitable because the configurational flow induced by the symmetry transform interactions is towards a symmetry invariant configuration. However under parallel dynamics our simulations show that transformed versions of an extensive number of paterns may be recognized for a range of symmetry transform interaction strengths. Furthermore there is an optimal symmetry transform interaction strength that best balances the competition between the two interactions.

It would be of interest to discover whether symmetry invariant recognition is possible for an extensive number of patterns when a more sophisticated symmetry group, such as Z_3, is desired. However in this case the techniques employed in section 3 are no longer applicable because the interactions are no longer symmetric.

Acknowledgments

MRE acknowledges the award of a postgraduate studentship from the Science and Engineering Research Council. CZ thanks the British Council and the University of Edinburgh for financial support. The Edinburgh Concurrent Supercomputer is a collaborative project with Meiko Limited supported by major grants from the Department of Trade and Industry, the Computer Board and the SERC.

Appendix 1. Replica symmetric theory

In the mean-field theory we look for the non-ergodic phase with macroscopic overlap m with a stored pattern. We proceed as in [2] to use the replica trick [12] to average the free energy over the quenched variables $\{\xi_i^\mu\}$. We make a replica symmetric ansatz to obtain the free energy as

$$f=\frac{\alpha}{2}+\tfrac{1}{2}m^2+\frac{\alpha}{2\beta}\left[\ln(1-\beta+\beta q)-\frac{\beta q}{1-\beta+\beta q}\right]-\tfrac{1}{2}\alpha\beta qr$$
$$-\frac{1}{Nn\beta}\left\langle\!\!\left\langle \ln\left[\mathrm{Tr}_{S_i^\rho}\exp\left(\frac{\alpha\beta^2 r}{2}\sum_{i,\rho\neq\sigma}S_i^\rho S_i^\sigma+\beta m\sum_i \xi_i^1\sum_\rho S_i^\rho\right.\right.\right.\right.$$
$$\left.\left.\left.\left.+\frac{\beta a}{2}\sum_{i,\rho}S_i^\rho S_{\pi(i)}^\rho\right)\right]\right\rangle\!\!\right\rangle \tag{29}$$

where the double angular brackets stand for the quenched average over the set of $\{\xi_i^1\}$ and q is the replica symmetric Edwards–Anderson order parameter

$$q^{\rho\sigma}=\frac{1}{N}\sum_i\langle S_i^\rho S_i^\sigma\rangle$$
$$=q \qquad \text{for } \rho\neq\sigma. \tag{30}$$

It is only the final term of (29) that differs from the equivalent expression for the Hopfield model. We develop this term by factorizing over pairs of sites i, $\pi(i)$ and performing two Gaussian transforms to linearize the first term inside the exponential for each pair of sites. The configurational trace may then be taken to give the free energy as

$$f=\frac{\alpha}{2}+\tfrac{1}{2}m^2+\frac{\alpha}{2\beta}\left[\ln(1-\beta+\beta q)-\frac{\beta q}{1-\beta+\beta q}\right]+\tfrac{1}{2}\alpha\beta r(1-q)$$
$$-\frac{1}{2\beta}\left\langle\!\!\left\langle\int \mathrm{D}z\int \mathrm{D}z_\pi \ln[2\mathrm{e}^{\beta a}\cosh[\beta\sqrt{\alpha r}\,(z+z_\pi)+\beta m(\xi^1+\xi^1_\pi)]\right.\right.$$
$$\left.\left.+2\mathrm{e}^{-\beta a}\cosh[\beta(\sqrt{\alpha r}\,(z-z_\pi)+m(\xi^1-\xi^1_\pi))]\right]\right\rangle\!\!\right\rangle \tag{31}$$

where the Gaussian measure is denoted by

$$\mathrm{D}z\equiv\frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}. \tag{32}$$

The values taken by the order parameters appearing in (31) are given by the solution of the saddle point equations. In addition we have the order parameter associated with the symmetry transform interactions

$$g=-2\frac{\partial f}{\partial a}. \tag{33}$$

One finds the mean field equations to be

$$m_2=\int \mathrm{D}z\int \mathrm{D}z_\pi\frac{\mathrm{e}^{2\beta a}\sinh[\beta(\sqrt{\alpha r}\,(z+z_\pi)+2m)]}{\mathrm{e}^{2\beta a}\cosh[\beta(\sqrt{\alpha r}\,(z+z_\pi)+2m)]+\cosh[\beta(\sqrt{\alpha r}\,(z-z_\pi))]} \tag{34}$$

$$m_3=\int \mathrm{D}z\int \mathrm{D}z_\pi\frac{\sinh[\beta(\sqrt{\alpha r}(z-z_\pi)+2m)]}{\mathrm{e}^{2\beta a}\cosh[\beta\sqrt{\alpha r}\,(z+z_\pi)]+\cosh[\beta(\sqrt{\alpha r}\,(z-z_\pi)+2m)]} \tag{35}$$

$$q=\tfrac{1}{2}\int \mathrm{D}z\int \mathrm{D}z_\pi\frac{\mathrm{e}^{4\beta a}\sinh 2[\beta(\sqrt{\alpha r}\,(z+z_\pi)+2m)]+\sinh^2[\beta(\sqrt{\alpha r}\,(z-z_\pi))]}{(\mathrm{e}^{2\beta a}\cosh[\beta(\sqrt{\alpha r}(z+z_\pi)+2m)]+\cosh[\beta(\sqrt{\alpha r}(z-z_\pi))])^2}$$
$$+\tfrac{1}{2}\int \mathrm{D}z\int \mathrm{D}z_\pi\frac{\mathrm{e}^{4\beta a}\sinh^2[\beta(\sqrt{\alpha r}(z+z_\pi))]+\sinh^2[\beta(\sqrt{\alpha r}\,(z-z_\pi)+2m)]}{(\mathrm{e}^{2\beta a}\cosh[\beta(\sqrt{\alpha r}\,(z+z_\pi))]+\cosh[\beta(\sqrt{\alpha r}\,(z-z_\pi)+2m)])^2} \tag{36}$$

$$r=\frac{q}{(1-\beta+\beta q)^2} \tag{37}$$

$$g_2=\int \mathrm{D}z\int \mathrm{D}z_\pi\frac{\mathrm{e}^{2\beta a}\cosh\beta[(\sqrt{\alpha r}\,(z+z_\pi)+2m)]-\cosh[\beta\sqrt{\alpha r}\,(z-z_\pi)]}{\mathrm{e}^{2\beta a}\cosh\beta[(\sqrt{\alpha r}\,(z+z_\pi)+2m)]+\cosh[\beta\sqrt{\alpha r}\,(z-z_\pi)]} \tag{38}$$

$$g_3=\int \mathrm{D}z\int \mathrm{D}z_\pi\frac{\mathrm{e}^{2\beta a}\cosh\beta([\sqrt{\alpha r}\,(z+z_\pi))]-\cosh[\beta(\sqrt{\alpha r}\,(z-z_\pi)+2m)]}{\mathrm{e}^{2\beta a}\cosh\beta[(\sqrt{\alpha r}\,(z+z_\pi))]+\cosh[\beta(\sqrt{\alpha r}\,(z-z_\pi)+2m)]}. \tag{39}$$

In (34)-(39) the average over the independent quenched spin variables ξ^1 and ξ^1_π has been taken. The parameters with subscript 2 result from the contributions with $\xi^1=\xi^1_\pi$; the parameters with subscript 3 result from the contributions with $\xi^1=-\xi^1_\pi$.

Obtaining (40)-(44), the zero temperature limit of the mean-field equations (34-39), is a task requiring some patience. The strategy used is to examine the integrands in the equations and determine the regions of the $z-z_\pi$ plane where they do not vanish.

$$m_2 = \tfrac{1}{2}\operatorname{erf}\left[\frac{m+a}{\sqrt{2\alpha r}}\right] + \tfrac{1}{2}\operatorname{erf}\left[\frac{m-a}{\sqrt{2\alpha r}}\right] + \tfrac{1}{2}\int_{m-a/\sqrt{\alpha r}}^{m+a/\sqrt{\alpha r}} \mathrm{D}z \operatorname{erf}\left[\frac{2m}{\sqrt{2\alpha r}} - \frac{z}{\sqrt{2}}\right] \tag{40}$$

$$m_3 = \tfrac{1}{2}\operatorname{erf}\left[\frac{m+a}{\sqrt{2\alpha r}}\right] + \tfrac{1}{2}\operatorname{erf}\left[\frac{m-a}{\sqrt{2\alpha r}}\right] + \tfrac{1}{4}\operatorname{erf}^2\left[\frac{m-a}{\sqrt{2\alpha r}}\right] - \tfrac{1}{4}\operatorname{erf}^2\left[\frac{m+a}{\sqrt{2\alpha r}}\right] \tag{41}$$

$$\beta(1-q) = \frac{1}{\sqrt{2\pi\alpha r}}\left(\mathrm{e}^{-(m+a)^2/2\alpha r} + \mathrm{e}^{-(m-a)^2/2\alpha r}\right)\left(1 - \tfrac{1}{2}\operatorname{erf}\left[\frac{m+a}{\sqrt{2\alpha r}}\right] + \tfrac{1}{2}\operatorname{erf}\left[\frac{m-a}{\sqrt{2\alpha r}}\right]\right)$$
$$+ \frac{1}{2\sqrt{\pi\alpha r}}\operatorname{erf}\left[\frac{a}{\sqrt{\alpha r}}\right]\mathrm{e}^{-m^2/\alpha r} + \frac{1}{4\sqrt{\pi\alpha r}}\left(\operatorname{erf}\left[\frac{m+a}{\sqrt{\alpha r}}\right] + \operatorname{erf}\left[\frac{m-a}{\sqrt{\alpha r}}\right]\right)$$

$$r = (1-\beta(1-q))^{-2} \tag{42}$$

$$g_2 = \operatorname{erf}\left[\frac{m+a}{\sqrt{2\alpha r}}\right] - \operatorname{erf}\left[\frac{m-a}{\sqrt{2\alpha r}}\right] - \tfrac{1}{2}\operatorname{erf}^2\left[\frac{m-a}{\sqrt{2\alpha r}}\right] - \tfrac{1}{2}\operatorname{erf}^2\left[\frac{m+a}{\sqrt{2\alpha r}}\right]$$
$$+ 2\int_{2a/\sqrt{2\alpha r}}^{2(m+a)/\sqrt{2\alpha r}} \mathrm{D}z \operatorname{erf}\left[\frac{m+a}{\sqrt{\alpha r}} - \frac{z}{\sqrt{2}}\right] + 2\int_{-2a/\sqrt{2\alpha r}}^{2(m-a)/\sqrt{2\alpha r}} \mathrm{D}z \operatorname{erf}\left[\frac{m-a}{\sqrt{\alpha r}} - \frac{z}{\sqrt{2}}\right] \tag{43}$$

$$g_3 = \operatorname{erf}\left[\frac{m+a}{\sqrt{2\alpha r}}\right] - \operatorname{erf}\left[\frac{m-a}{\sqrt{2\alpha r}}\right] - \tfrac{1}{2}\operatorname{erf}^2\left[\frac{m-a}{\sqrt{2\alpha r}}\right] - \tfrac{1}{2}\operatorname{erf}^2\left[\frac{m+a}{\sqrt{2\alpha r}}\right] \tag{44}$$

where

$$\operatorname{erf}[x] = \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{d}z\, \mathrm{e}^{-z^2}. \tag{45}$$

In this process one finds that the symmetry $m_3(-a) = m_2(a)$; $g_3(-a) = g_2(a)$ apparent in (34)-(39), is broken in the $\beta \to \infty$ limit, because one must restrict the equations to the case $a > 0$. The $a = 0$ limit of these equations gives the equivalent equations for the Hopfield model [2].

Appendix 2. Random external field model zero temperature equations

The following zero temperature mean field equations for the random external field model are a straightforward extension of those of the Hopfield model [2].

$$m_2 = \operatorname{erf}\left[\frac{m+a}{\sqrt{2\alpha r}}\right] \tag{46}$$

$$m_3 = \operatorname{erf}\left[\frac{m-a}{\sqrt{2\alpha r}}\right] \tag{47}$$

$$\beta(1-q) = \frac{1}{\sqrt{2\pi\alpha r}}\left(\mathrm{e}^{-(m+a)^2/2\alpha r} + \mathrm{e}^{-(m-a)^2/2\alpha r}\right) \tag{48}$$

$$r = (1-\beta(1-q))^{-2} \tag{49}$$

References

[1] Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554
[2] Amit D J, Gutfreund H and Sompolinsky H 1987 *Ann. Phys., NY* **173** 30
[3] Bruce A D, Gardner E and Wallace D J 1987 *J. Phys. A: Math. Gen.* **20** 2909
[4] Sompolinsky H 1987 *Heidelberg Colloquium on Glassy Dynamics* ed J L Van Hemmen and I Morgestern (Heidelberg: Springer)
[5] Gardner E 1988 *J. Phys. A: Math. Gen.* **21** 257
[6] von der Malsburg C and Bienenstock E 1987 *Europhys. Lett.* **3** 1243
[7] Bienenstock E and von der Malsburg C 1987 *Europhys. Lett.* **4** 121
[8] Kree R and Zippelius A 1988 *J. Phys. A: Math. Gen.* **21** L813
[9] Mas J and Ramos E 1989 *J. Phys. A: Math. Gen.* **22** 3379
[10] Dotsenko V 1988 *J. Phys. A: Math. Gen.* **21** L783
[11] Coolen A C C and Kuijk F W 1989 *Neural Networks* **2** 495
[12] Kirkpatrick S and Sherrington D 1978 *Phys. Rev.* B **17** 4384
[13] Gardner E 1986 *J. Phys. A: Math. Gen.* **19** L1047
[14] Gardner E, Derrida B and Mottishaw P 1987 *J. Physique* **48** 741