Phase diagrams and the instability of the spin glass states for the diluted Hopfield neural network model

Andrew Canning, Jean-Pierre Naef

Classification
Physics Abstracts
64.60 — 75.10N — 87.30

Phase diagrams and the instability of the spin glass states for the diluted Hopfield neural network model

Andrew Canning and Jean-Pierre Naef (*)

Département de Physique Théorique, Université de Genève, CH-1211 Genève 4, Switzerland

(Received 17 February 1992, accepted in final form 7 May 1992)

Abstract. — The replica-symmetric order parameter equations derived in [2, 4] for the symmetrically diluted Hopfield neural network model [1] are solved for different degrees of dilution. The dilution is random but symmetric. Phase diagrams are presented for $c = 1, 0.1, 0.001$ and $c \to 0$, where c is the fractional connectivity. The line T_c where the memory states become global minima (having lower free energy than the spin glass states) is also found for different values of c. It is found that the effect of dilution is to destabilize the spin glass states and the line T_c is driven rapidly towards the line T_M; the phase transition line where the memory states first stabilize. All the results are derived in the context of replica symmetry so are expected to be incorrect in certain parts of the phase diagram (the error increasing the further below the replica symmetry breaking lines we are) but our results do suggest in a general way that dilution, even in small quantities, increases the stability of the memory states with respect to the spin glass states.

1. Introduction.

In the study of attractor neural networks (for a review plus references see [5]) the main objectives have been to store as many patterns as possible with as little error as possible and also to avoid unwanted spurious states which crowd the phase space and reduce the basins of attraction of the stored states. In the case of the standard Hopfield model [8] these spurious states show up very clearly in the calculation and are analogous to spin glass states found in the SK spin glass model [9]. In replica-symmetric theory these spin glass states remain global minima of the free energy down to very low values of α (α is the number of stored patterns divided by the system size). By analogy with spin glasses it is also expected that their number will increase exponentially with the system size N. If we consider ourselves sitting close to α_c (the critical α value below which memory states are stable), as we increase N, the number of stored states only increases linearly with the system size. This means that the spin glass states

(*) *Present address:* Département de Physique Théorique, Université de Lausanne, CH-1015 Lausanne, Switzerland.

will quickly swamp the phase space. If the size of basins of attraction associated with the spin glass states are of the same order of magnitude as those of the memory states then the basins of attraction of the memory states will occupy a vanishingly small fraction of the phase space in the large N limit. Therefore, these spurious states will severely limit the operation of the network and in particular the size of the basins of attraction of the memory states.

There are also other spurious stable states which have overlaps with many patterns (mixture states). These states always have higher free energies than the memory states and stabilize at much lower values of α than the memory states. Thus we consider their effect on the performance of the system to be less of a problem than the spin glass states. They will be discussed in more detail in the separate section on mixture states.

Recently Watkin and Sherrington [6] have found that in the limit of very high symmetric dilution (e.g. of order N^a connections per site where $0 < a < 1$), the spin glass states are not stable for values of α and T, at which states associated with the stored patterns (memory states) are stable in contrast to the fully connected case. This means that there are regimes on the α, T phase diagram where only the memory states are stable which implies they must have very large basins of attraction. They also found that, in the limit of low symmetric connectivity, the order parameter equations for the memory states are equivalent to those of an SK spin glass. This was previously pointed out in reference [3] although in this paper it was wrongly thought that the spin glass states remained stable below α_c.

In this paper we will study the effect of gradual dilution on the phase diagram of random symmetrically diluted neural networks. In this way we will see how the phase diagram is gradually driven to that of the SK spin glass and more importantly we will see how dilution affects the stability of the spin glass states below α_c.

In the next three sections of this paper we will present the order parameter equations for the dilute Hopfield model along with numerical and, where possible, analytical solutions. Replica symmetry breaking and the stability of the mixture states will be discussed in sections five and six and the stability of the spin glass states below α_c will be discussed in section seven. In the last two sections of the paper we will present a brief discussion of the relationship between the low symmetric connectivity model and the SK spin glass model [9] along with our final conclusions and conjectures.

2. The model and its order parameter equations.

The model is based on the standard Hopfield model [1] with random but symmetric dilution of the bonds. We therefore consider a system of N Ising spins where the Hamiltonian is given by

$$H = -\sum_{ij} J_{ij} S_i S_j, \tag{1}$$

the sum being over all i and j. The interactions are chosen to be

$$J_{ij} = \frac{c_{ij}}{cN} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \tag{2}$$

where c_{ij} is 1 with probability c and 0 with probability $1 - c$. c is thus the connectivity of the system. We also choose $J_{ii} = 0$ and c_{ij} is symmetric. The patterns to be stored $\xi^\mu = \pm 1$, are random. Thus choosing $c = 1$ will give us the standard Hopfield model and the limit $c \to 0$ will give us the model studied in references [3, 6], which is closely related to the SK spin glass [9] (see Sect. 8). The replica-symmetric order parameter equations for the symmetrically diluted

model were calculated in [4] by noting the equivalence between synaptic noise and dilution and also separately in [2, 3] using a different technique (in these papers the solutions of the order parameter equations were only studied at zero temperature, for finite c, and at all temperatures for the limit $c \to 0$). The replica-symmetric order parameter equations for the symmetrically diluted model with no external field are

$$\begin{aligned} m^\nu &= \ll \int \frac{dz}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) \xi^\nu \tanh\beta\left(\sqrt{\alpha(r+(1-c)q)}z + \mathbf{m}.\xi\right) \gg \\ q &= \ll \int \frac{dz}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) \tanh^2\beta\left(\sqrt{\alpha(r+(1-c)q)}z + \mathbf{m}.\xi\right) \gg \\ r &= \frac{cq}{(1-C)^2}, \end{aligned} \tag{3}$$

where C is given by

$$C = \beta(1-q), \tag{4}$$

and the free energy per site f is given by

$$\begin{aligned} \frac{f}{c} &= \frac{\mathbf{m}^2}{2} + \frac{C\alpha r}{2} + \frac{\alpha c}{2}\left(1 + \frac{\ln(1-C)}{\beta} - \frac{q}{1-C}\right) - \frac{\alpha\beta(1-c)(1-q)^2}{4} \\ &\quad -\frac{1}{\beta} \ll \int \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \ln\left[2\cosh\beta(\sqrt{\alpha(r+(1-c)q)}z + \mathbf{m}.\xi)\right] \gg . \end{aligned} \tag{5}$$

The physical interpretation of the order parameters is given by

$$\begin{aligned} m^\nu &= \frac{1}{N}\sum_{i=1}^{N} \ll \xi_i^\nu < S_i > \gg \\ q &= \frac{1}{N}\sum_{i=1}^{N} \ll < S_i >^2 \gg \\ r &= \frac{1}{\alpha}\sum_{\mu=s+1}^{p} \ll (m^\mu)^2 \gg . \end{aligned} \tag{6}$$

α is a measure of the number of patterns stored per connection per site and is given by

$$\alpha = \frac{p}{Nc}, \tag{7}$$

which should be distinguished from the more standard definition $\alpha = \frac{p}{N}$ used in [4, 8]. $\ll \gg$ represents the quenched average over all the possible choices of the p sets of patterns. m^ν is a measure of the overlap of the system with the set of s patterns which are nominated for condensation s being finite. q is the Edwards Anderson order parameter and r is a measure of the overlap of the system with the infinite set of $p - s$ patterns which are not nominated for condensation. A fuller explanation of the techniques used in deriving these equations can be found in [5] where the calculation for the fully connected model is explained in detail.

3. The paramagnetic to spin glass phase boundary.

The spin glass phase corresponds to solutions of the order parameter equations such that $m^\nu = 0, \forall\nu$ and $q \neq 0$. Solutions of this form only exist below a critical temperature $T_g(\alpha)$

and the value of q always changes continuously across the phase boundary for all choices of c. Thus solving the order parameter equations (3) to first order in q will give us the line T_g. This gives a quartic in T

$$T^4 - 2T^3 + (1-\alpha)T^2 + 2\alpha(1-c)T + \alpha(c-1) = 0, \quad c \neq 0. \tag{8}$$

The largest root of this equation gives T_g while the other three roots do not correspond to physical solutions. In the case $c \to 0$ the phase boundary is the same as that for the SK spin glass [3, 6] with appropriate choices of the first and second moments of the bond distribution (this is discussed in more detail in Sect. 8). It should be noted that the limit $c \to 0$ does not correspond to the case where there are no interactions but to the case where the number of interactions rises more slowly than the system size. This would, for example, correspond to systems with of order N^a interactions per site where $0 < a < 1$ or systems with $\log N$ interactions per site. The line T_g for the model $c \to 0$ along with solutions of equation (8) for $c = 1, 0.1$ and 0.001 are shown in figure 3.

In general, the solutions $T_g(c, \alpha)$ of equation (8) are very complicated functions of c and α but there are a few cases where the quartic easily factorizes. These are $c = 1$ which, of course, is the result for the fully connected model $T_g = 1 + \sqrt{\alpha}$ and also $c = \frac{1}{2}$ where we find

$$T_g(c = \frac{1}{2}, \alpha) = \frac{1 + \sqrt{[1 + 2(\alpha + \sqrt{\alpha^2 + 2\alpha})]}}{2} \tag{9}$$

Looking at figure 3 we can see that the effect of dilution on the phase diagrams is to decrease T_g, except at its end points ($T_g = 1, \alpha = 0$) and ($T_g \to \sqrt{\alpha}, \alpha \to \infty$) which remain fixed for all values of c. The line T_g, for any finite c, is always enveloped above by the line $1 + \sqrt{\alpha}$, which is the result for the fully connected model, and below by the line $T_g = 1$ for $\alpha < 1$ and $T_g = \sqrt{\alpha}$ for $\alpha \geq 1$.

4. The transition to magnetic ordering.

There exists a critical value of the temperature T_M below which the order parameter equations (3) have solutions of the form $\mathbf{m}^\nu = (m\ 0\ 0 \ldots\ 0\ 0)$, $m \neq 0$. These correspond to stable states (which we shall call memory states) associated with the patterns we have tried to store in the system. In the solution space of the order parameter equations this transition corresponds to the line at which, for increasing α, two real solutions of the form $m \neq 0$ meet and become complex. Only one of these two solutions is a minimum of the free energy and thus a physical solution. The line we seek is therefore a bifurcation point in the solution space of the order parameter equations (3) and corresponds to the point at which the determinant of the matrix of first derivatives of the order parameter equations with respect to the order parameters is zero. This extra condition has to be solved along with the original three order parameter equations to give $T_M(\alpha_c)$. We solved these equations using a Newton Raphson algorithm at $c = 1, 0.1$ and 0.001 the results being shown in figure 3. A more detailed explanation of this approach to finding the transition line T_M can be found in reference [12] where it was applied to the fully connected Hopfield model to show up the reentrant spin glass behaviour.

The phase transition line T_M is first order (except in the limit $c \to 0$ where it becomes second order) so unlike the spin glass phase boundary there are no small parameters which can be expanded around a generic point to analytically determine the phase boundary. However, close to $T_M = 0$ we can expand the order parameter equations along the phase transition line to find

the critical values of the order parameters. This gives, to first order in T_M, (they are written in the form $x = x_0 + x_1 T_M$)

$$\begin{aligned} m &= m_0 + m_0 C_0 \frac{y_1}{y_0} T_M \\ q &= 1 - C_0 T_M \end{aligned} \tag{10}$$

where y was used to parameterize the order parameter equations and is given by

$$y = \frac{m}{\sqrt{2\alpha(r + (1-c)q)}}. \tag{11}$$

y_1 in equations (10) is a function of c, C_0 and y_0, given by,

$$y_1 = -\frac{y_0[c + (1-c)(1-C_0)^2](1-C_0)(1-C_0-2y_0^2)}{2\{c[4y_0^2 + (1+2y_0^2)(1-C_0-2y_0^2)] + 4(1-c)(1-C_0)^3(1-C_0-2y_0^2)\}}. \tag{12}$$

Solving the order parameter equations to first order in T_M also gives the phase transition line

$$T_M = \frac{\alpha - \alpha_c}{\alpha_c C_0} \tag{13}$$

The subscript 0 represents the solutions at $T_M = 0$ except for the variable α where we use the same notation as Amit [8] which is $\alpha_c (= \alpha_0)$. The values of α_c, m_0, C_0 and y_0 vary with c and can only be determined numerically (except in the limit $c \to 0$). Figures 1 and 2 show the values of these four parameters for c between 0 and 1.

At all values of c the phase transition lines tend to $T_M = 1$ in the limit of small α so we can also expand the order parameter equations in $t = 1 - T_M$ to find analytical solutions close to

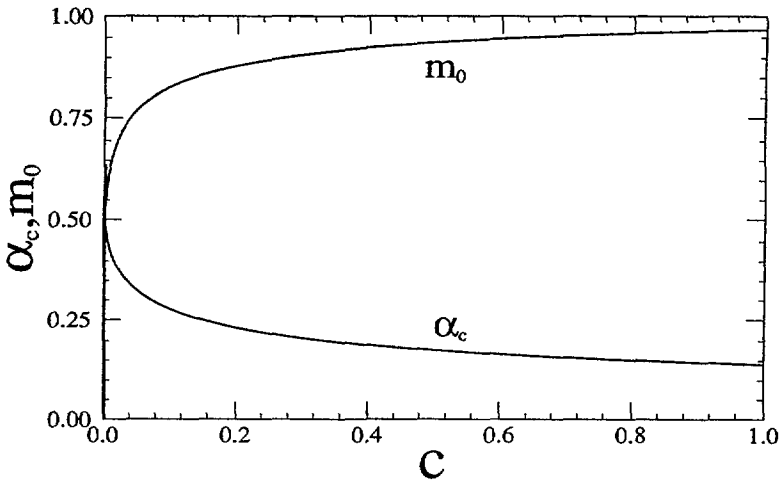

Fig.1. — Critical values of the storage capacity α and the overlap m at zero temperature (m_0) for values of c (the connectivity) between 0 and 1. m_0 ranges from 0 in the limit $c \to 0$ to 0.967 at $c = 1$. It drops almost vertically for values of m_0 below 0.5. α_c varies from $\frac{2}{\pi}$ in the limit $c \to 0$ to 0.138 at $c = 1$.

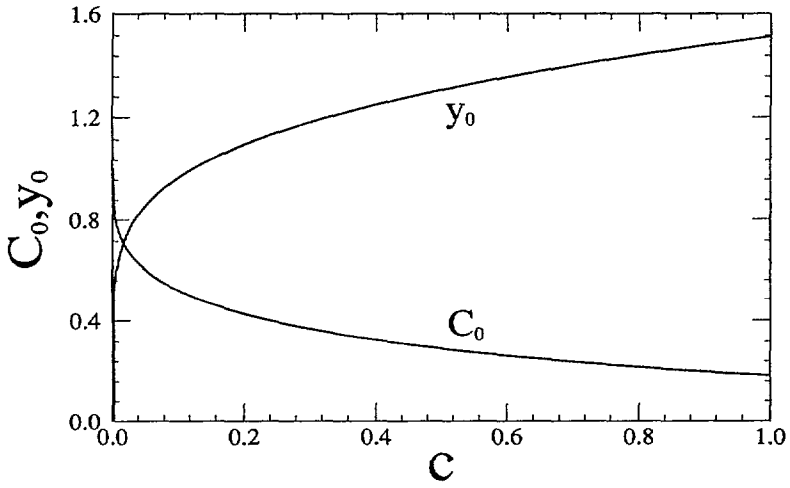

Fig.2. — Critical values of C and y (see text, Sect. 4 for definitions) at zero temperature for values of c (the connectivity) between 0 and 1. y_0 ranges between 0 in the limit $c \to 0$ to 1.51 at $c = 1$. C_0 varies from 1 in the limit $c \to 0$ to 0.180 at $c = 1$.

the point $T_M = 1, \alpha = 0$ on the phase transition line. To first order in t this gives

$$\begin{aligned} m^2 &= 1.78t \\ q &= 2.19t \\ \frac{1}{r} &= 0.645\frac{t}{c}, \end{aligned} \tag{14}$$

with only r varying with c. The phase transition line is given by

$$T_M = 1 - 1.95\sqrt{\alpha c}. \tag{15}$$

Thus we can explicitly see that the effect of decreasing c is to increase T_M, for fixed α, which is a result that was also borne out by the numerical solutions for T_M (see Fig. 3).

Looking at figure 3 we can see that the effect of dilution on the memory phase is to increase its size both in the temperature and α directions but the value of the overlap m decreases on the phase boundary T_M. The phase transition becomes second order in the limit $c \to 0$. This can be seen on the plot of m_0 against c at zero temperature (see Fig. 1). In fact for small c we can obtain an analytic expression to smallest order in c for $m_0(c)$

$$m_o(c) \simeq 2^{\frac{2}{3}}\sqrt{\left(\frac{3}{\pi}\right)}c^{\frac{1}{6}}. \tag{16}$$

As can be seen from the graph of m_0 against c, the value of m_0 holds up very well as c decreases (at $c = 0.1$, $m_0 = 0.83$) before dropping rapidly to zero. This is also reflected by the $\frac{1}{6}$ power in the leading order term of the expansion for m_0 (see above equation).

Although m decreases as c decreases on the phase transition line T_M, we found that for fixed α and T, within the memory phase, the value of m increased as c was decreased. Therefore, in terms of accuracy of storage per connection per neuron, the more dilute a network is the better its performance is.

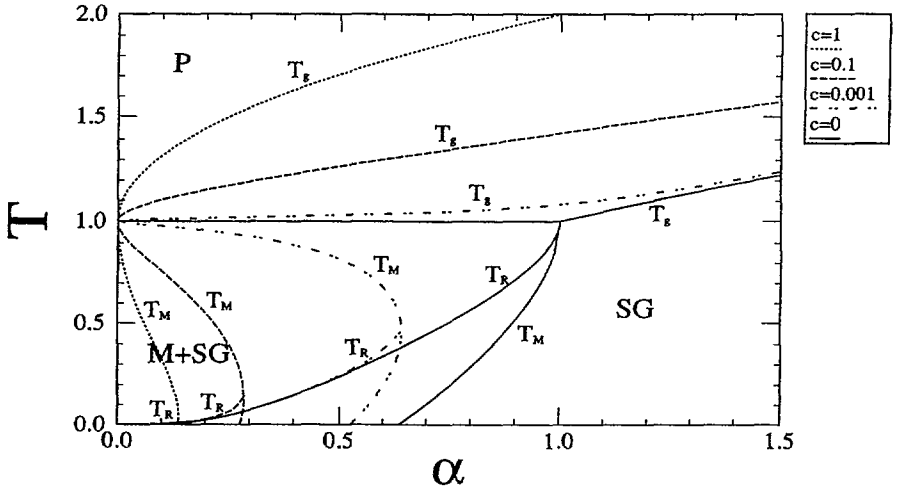

Fig.3. — Phase diagrams for the Hopfield neural network model in replica-symmetric theory with connectivity $c = 1, 0.1, 0.001$ and $c = 0$ which represents the limit $c \to 0$. α is the number of patterns stored per interaction per spin. P, SG and M+SG, refer to the paramagnetic phase ($m = 0$, $q = 0$), pure spin glass phase ($m = 0$, $q \neq 0$) and the memory spin glass phase where both the memory states ($m \neq 0$, $q \neq 0$) and the spin glass states ($m = 0$, $q \neq 0$) are stable. The spin glass states appear below the lines T_g and the memory states below the lines T_M. T_R is the line below which replica symmetry is broken for the memory states. For the model $c \to 0$ only the memory states are stable to the left of the line T_M. The phase diagrams for the models c = 1 and $c \to 0$ were first presented in references [8] and [2] respectively.

Another interesting feature of the memory phases for each value of c (see Fig. 3) is that they all show reentrant spin glass behaviour below a certain non-zero temperature similar to that found for the replica-symmetric phase diagram for the SK spin glass [9]. This is also seen in the leading order expansion for the phase transition line T_M around $T_M = 0$ (see Eq. (13)) where the gradient is positive for all values of c. This reentrant behaviour has previously been studied for the fully connected model in reference [12]. Like the SK spin glass this reentrant behaviour is thought to be a result of the replica-symmetric assumption and the true phase diagram (replica broken) is expected not to exhibit this property. This will be discussed in more detail in section 6 on replica symmetry breaking.

5. Mixture states.

Mixture states are states for which several of the m^ν are non-zero. As was found for the fully connected model [8] the mixture states were solutions of the order parameter equations below a certain critical line $T(\alpha)$. The values of α at which these states appeared as solutions were always found to be lower than the α_c for the memory states, except in the limit $c \to 0$ where all the mixture states appeared as solutions at $\alpha = \frac{2}{\pi}$ This is the same value of α_c at which the memory states stabilize. In fact we found that expanding to first order in c at zero temperature

$$\alpha_{1c} = \frac{2}{\pi}(1 - 1.89c^{\frac{1}{3}})$$

$$\alpha_{3c} = \frac{2}{\pi}(1 - 6.92c^{\frac{1}{3}}), \tag{17}$$

where α_{1c} and α_{3c} are the values of α below which there are solutions of the order parameter equations corresponding to the memory states and states having the same magnitude of overlap with three of the patterns respectively. As can be seen from these equations it is only in the limit $c \to 0$ that α_{1c} and α_{3c} become the same. Similar expressions exist for all the other overlap states. We did not look at the stability of the overlap states but we imagine that, like the fully connected model [8], only the solutions with an odd number of overlaps with the nominated patterns are stable and that the free energy of the mixture states is always higher than that of the memory states. It was shown in reference [6] in the case $c \to 0$ at zero temperature that even though states in which three patterns have an equal overlap appear at $\alpha_{3c} = \frac{2}{\pi}$ they are unstable and do not become true minima of the free energy until $\alpha = 0.123$. We believe this quantitative behaviour will be true for all states having an odd number of overlaps with the patterns and for all c values.

6. Replica symmetry breaking.

All the results we have presented so far in this paper are within the replica-symmetric assumption (see [7] for a fuller explanation of this approach and the physical interpretation of replica symmetry breaking). In the case of the spin glass phase it is known that the whole phase has broken replica symmetry so that the values of the free energies etc. calculated from this assumption can be expected to be in error but the size of the error is expected to be smaller the closer we are to the phase boundary. In fact, the true spin glass states in replica broken theory do not have the same free energies. In the case of the memory states there is a line analogous to the De Almeida Thouless line [10] for the SK spin glass below which replica symmetry is broken. Using a similar approach this line, which we shall call T_R, can also be calculated for the dilute Hopfield neural network and is given by

$$\frac{\beta^2 \alpha r}{q} \int \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \operatorname{sech}^4 \beta[\sqrt{\alpha(r + (1-c)q)}\, z + m] = 1. \tag{18}$$

This expression has to be solved simultaneously along with the order parameter equations to determine T_R: the line of replica symmetry breaking. This was done for the same values of c as the phase transition lines and is shown on the phase diagrams (see Fig. 3). It can be seen from this figure that even though the region of broken replica symmetry is very small for the fully connected model, it grows very rapidly as c decreases. In fact in the limit $c \to 0$ the whole of the spin glass to memory phase boundary lies below it and is therefore in error. Thus replica symmetry breaking becomes more and more important as we dilute the system and we can expect the error in α_c calculated in replica-symmetric theory to become larger and larger.

It is interesting to note that, to within the accuracy of our numerical solutions, the replica symmetry breaking line always meets the line T_M at the point where the gradient changes sign (see Fig. 3). Thus at this point the gradient of the line T_M is infinite and so, by analogy with the SK spin glass, we may imagine that the full replica broken solution is a vertical line (or close to this) below the line T_R. If this conjecture is correct then the true α_c at zero temperature (or very close to it) can always be found for these Hopfield type models by finding the point at which the line T_M meets the line T_R. The only other possible scenario, assuming reentrant behaviour is excluded, is for the memory phase transition line below T_R to move out to higher α_c causing it to have an inflection point.

7. The stability of the spin glass states below T_M.

We now turn to what are perhaps the most interesting results in this paper. It was shown in [6] that in the limit $c \to 0$ the spin glass states are not stable below the line T_M in contrast to the fully connected model. In this section we study their stability below the line T_M for finite values of c.

There exists a critical line $T_c(\alpha)$ below which the memory states become global minima of the free energy surface. In the case of the fully connected model [8] this line is well below the line T_M (at zero temperature the memory states are global minima for $\alpha < 0.051$ and $\alpha_c = 0.138$). The line $T_c(\alpha)$ is found by equating the free energies (see Eq. (5)) of the spin glass and memory states which are physical solutions of the order parameter equations. Figure 4 shows the lines $T_c(\alpha)$ and $T_M(\alpha)$ for the same three values of c ($c = 1, 0.1$ and 0.001) as the phase diagrams.

We can see clearly from figure 4 that dilution stabilizes the memory states at the cost of the spin glass states. For the case $c = 1$ in more than 50% of the memory phase the spin glass states have a lower free energy than the memory states. For the case $c = 0.001$ this drops to only a few percent. If we imagine, for our free energy surface, that the depth of minima is proportional to their width then this implies that dilution may enlarge the basins of attraction associated with the memory states. This means that dilution may have a positive effect on the storage properties (per connection) of the system although, simulations would be required to validate this conjecture.

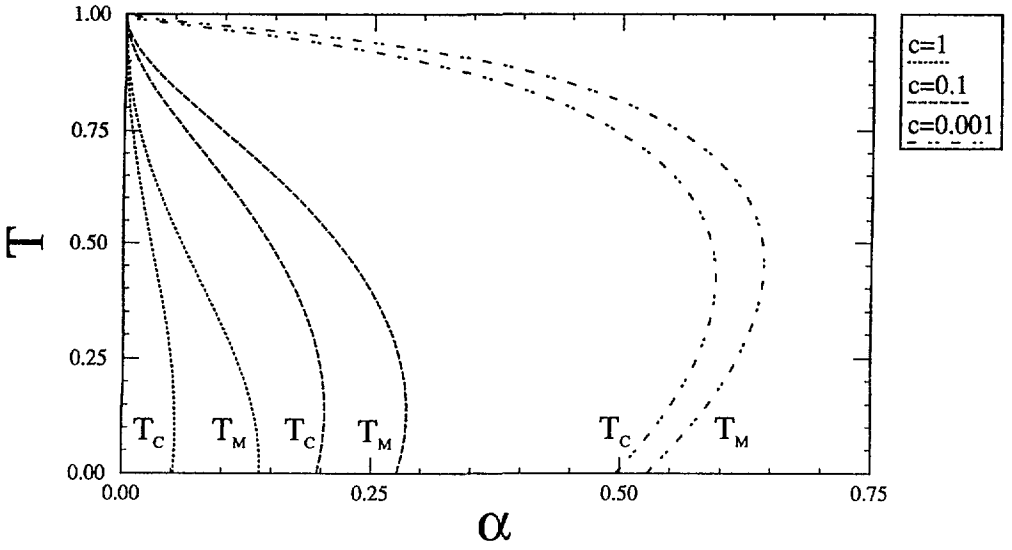

Fig.4. — Phase boundary lines T_M and T_c for connectivities $c = 1, 0.1$ and 0.001. To the left of the line T_M the memory states are stable and to the left of the line T_c they become the global minima of the free energy. It can clearly be seen on these three graphs that as c decreases the line T_c approaches the line T_M.

8. The limit $c \to 0$ and the SK spin glass.

Taking the limit $c \to 0$ on the order parameter equations (3) causes $r \to 0$ and the equations in m^ν and q for a single memory state become the same as those for the SK spin glass [9] with $J_0 = 1$ and $J^2 = \alpha$. $\frac{J_0}{N}$ and $\frac{J^2}{N}$ are the first and second moments of the Gaussian distribution which determines the interactions for the SK spin glass. This equivalence of the two sets of equations can be made clearer by performing a Mattis like [11] gauge transformation

$$\begin{aligned} S_i^{'} &= \xi_i^\nu S_i \\ J_{ij}^{'} &= \xi_i^\nu J_{ij} \xi_j^\nu , \end{aligned} \tag{19}$$

on the neural network equations, which means that the overlap order parameter m^ν now becomes the magnetization m and the first and second moments of $J_{ij}^{'}$ are $\frac{1}{N}$ and $\frac{\alpha}{N}$ respectively. The full replica order parameter equations without the replica-symmetric assumption, are also the same for both models so the true replica broken phase diagram is the same for both models [3, 6]. This means that the true form of the line T_M, in this case, is vertical at $\alpha = 1$ removing the reentrant spin glass behaviour (see Fig. 3). It should be noted that this is only a formal equivalence between the order parameter equations for the SK spin glass and the $c \to 0$ neural network. The two models are themselves not equivalent since the neural net has $2p$ stable states associated with the p patterns (only above T_R and at $\alpha < 1$) as well as stable mixture states at lower α values while the SK spin glass has only two stable states associated with the $\pm m$ solutions of the order parameter equations.

A more physically motivated understanding of how the phase transition of the neural net is driven towards that of the SK spin glass comes from studying the correlations around loops of interactions for the neural net. In the case of the Hopfield neural network, because of the way the bonds are chosen (see Eq. (2)), the interactions around a closed loop of any length are correlated. For the expectation values of the loop products this gives

$$N\overline{J_{ij}J_{ji}} = N^2\overline{J_{ij}J_{jk}J_{ki}} = . \quad = \alpha, \tag{20}$$

whereas in the case of the SK model, correlations of this type are not present as the interactions are chosen independently from a Gaussian distribution (except in the case of the two site loop since the interactions are symmetric). It is these non-zero loop products which drive the memory and spin glass phase transitions in the Hopfield neural network away from those of the SK spin glass. When we dilute the system we reduce the values of these loop products and in the limit $c \to 0$ all these loop products, except the first one, tend to zero and the phase transition is of the same type as the SK spin glass. This effect can be followed explicitly from the expression for r derived for dilute models in references [2, 3]. In these papers r is a sum of terms written in the form of loop products so we can explicitly see how by dilution $r \to 0$ and the phase transition is driven towards that of the SK spin glass.

A more rigorous and general derivation of the order parameter equations for structured neural networks will be given in [14] using the techniques developed in [13]. The dilute models we studied in this paper and in references [2, 3] are a subclass of the larger class of solvable structured neural network models which will be studied in [14]. In this paper the role of these loop products in determining the behaviour of the system will come out naturally from the expression for the order parameter r.

9. Conclusions.

The main result of this paper is to show that in replica theory even small amounts of dilution reduce the stability of the spurious spin glass states while stabilizing the memory states. We expect this result to hold for the true replica broken solution of the symmetrically diluted model although we await either numerical simulations or replica broken analytic solutions to confirm this. It would also be interesting to see if this effect is found in other types of attractor networks using different types of learning rules.

All the results given in this paper are expressed in terms of the concentration of bonds but, as pointed out by Sompolinsky [4], there is an equivalence between synaptic noise and synaptic dilution. This means that all our results also describe a system with synaptic noise of zero mean and variance $\Delta^2 = \frac{\alpha}{N}(1-c)$.

We have also shown in this paper how gradual dilution of the Hopfield neural network drives its phase diagram for the memory states towards that of an SK spin glass. All the models studied in this paper were found to have reentrant spin glass behaviour in the replica-symmetric solution and we found that the replica symmetry breaking line T_R always meets the line T_M (to within the numerical accuracy of our solutions) where it is vertical. We therefore postulate, by analogy with the SK spin glass, that the true shape of the line T_M below T_R is vertical or at least very close to vertical. The true maximum value of α_c would then be given by its value at the point where the line T_R meets the line T_M in replica-symmetric theory. This is also the maximum value of α on the phase boundary T_M in replica-symmetric theory.

References

[1] Hopfield J.J., (1982) *Proc. Natl. Acad. Sci. U.S.A* **79**, 2544.

[2] Canning A. and Gardner E., (1988) *J. Phys.* **A21**, 3275.

[3] Canning A., (1988) *Neural Networks from Models to Applications*, Proc. Neuro'88 Conf., E.S.P.C.I, Paris., Edited by Personnaz L. and Dreyfus G. (I.D.S.E.T. Paris) 326.

[4] Sompolinsky H., (1986) The Theory of Neural Networks: The Hebb Rule and Beyond. Heidelberg Colloquium on Glassy Dynamics., (June 1986), Springer-Verlag.

[5] Amit D.J. (1989) *Modeling Brain Function*, Cambridge University Press.

[6] Watkin T.L.H. and Sherrington D., (1991) *Europhys. Lett.*, **14**, 791

[7] Mezard M., Parisi G. and Virasoro M. A. (1986), *Spin Glass Theory and Beyond*, World Scientific Lecture Notes in Physics Vol. 9.

[8] Amit D.J., Gutfreund H. and Sompolinsky H., (1985) *Phys. Rev. Lett.* **55**, 1530.
(1987) *Ann. Phys.* **173**, 30.

[9] Sherrington D. and Kirkpatrick S., (1975) *Phys. Rev. Lett.* **32**, 1792.
(1978) *Phys. Rev.* **B17**, 4384.

[10] De Almeida J.R.L., and Thouless D.J., (1978) *Phys. Rev.* **A11**, 983.

[11] Mattis D.C., (1976) *Phys. Lett.* **A36**, 421.

[12] Naef J-P. and Canning A., (1992) *J. Phys. I France* **2**, 247.

[13] Canning A., (1992) Preprint, Submitted to *J. Phys.* **A**.

[14] Canning A., In preparation