Reetrant spin glass behaviour in the replica symmetric solution of the Hopfield neural network model

Jean-Pierre Naef, Andrew Canning

Classification
Physics Abstracts
64.60 — 75.10N — 87.30

Reentrant spin glass behaviour in the replica symmetric solution of the Hopfield neural network model

Jean-Pierre Naef * and Andrew Canning

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

(Received 18 November 1991, accepted 6 December 1991)

Abstract. — Numerical and analytical solutions at low temperature are presented for the replica symmetric order parameter equations of the Hopfield neural network model with interactions between all the spins. We find that these equations give weak reentrant spin glass behaviour at $T < 0.023$. This is similar behaviour to that found for the replica symmetric solution of the SK spin glass [4] although in that case the reentrant phase is much larger. It is also known from what is believed to be the true solution for the SK spin glass that this reentrant behaviour is unphysical so, by analogy, we believe it is unphysical for neural network models. Even so, the maximum value in replica theory of α_c, which measures the storage capacity of the Hopfield model, is to be found at $\alpha_c(T = 0.023) = 0.1382$. This is slightly higher than the zero temperature value of $\alpha_c(T = 0) = 0.1379$.

1. Introduction.

Since the work of Hopfield in 1984 [1], which showed the analogy between a simple neural network model and spin glass models in statistical mechanics, replica symmetric theory (see [2] for a review) has been applied to this model to find its thermodynamic properties [3]. Although this technique was found to give spin glass reentrant behaviour in the case of the SK spin glass model this behaviour was not seen in the Hopfield neural network model. We know, in the case of the SK spin glass, that this reentrant behaviour is unphysical because the symmetry broken solution can be calculated in this case [5]. This solution is believed to be exact and does not show any reentrant behaviour.

In the next section of this paper we will derive an equation which, when solved simultaneously with the order parameter equations derived by Amit *et al.* [3], gives the phase transition line between the memory phase and the pure spin glass phase. Using an appropriate numerical technique we will then solve these equations to find the phase transition line. In the same section we shall also present analytical solutions of these equations to first order in T.

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

2. Equations for the memory to spin glass phase boundary.

The order parameter equations for the Hopfield neural network model for solutions having a macroscopic overlap with only one of the stored patterns can be written as [3]

$$\begin{aligned} f_1(m,q,\alpha,T) &= \int \frac{\mathrm{d}z}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \tanh \beta(\sqrt{\alpha r(q)}z+m) - m = 0 \\ f_2(m,q,\alpha,T) &= \int \frac{\mathrm{d}z}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \tanh^2 \beta(\sqrt{\alpha r(q)}z+m) - q = 0 \end{aligned} \tag{1}$$

where the function $r(q)$ is given by

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

The physical interpretation of m and q (see Ref. [3]) is: m is analogous to magnetization and measures the overlap of the state of the system with a single pattern nominated for condensation and q is the Edwards Anderson order parameter which signifies spin glass ordering if it is non-zero when m is zero. In the Hopfield neural network model the phase transition line we seek corresponds, for T fixed, to the value of α where physical ferromagnetic solutions of the type $m \neq 0$, $q \neq 0$ are no longer present and only spin glass solutions of the type $m = 0$, $q \neq 0$ exist. For this model these ferromagnetic solutions disappear discontinuously. For solutions to be physical they must be minima of the free energy and real as well as solutions of equations (1). The desired phase transition line corresponds to the point at which, varying α, two real solutions of the form $m \neq 0$, $q \neq 0$ converge and then become complex. One of these real solutions is not a minimum of the free energy and is therefore not a physical solution. Thus, this is a bifurcation point in terms of the solutions of equations (1) and corresponds to the point in the solution space at which the determinant of the matrix of partial derivatives of the order parameter equations, with respect to the order parameters, becomes zero i.e.

$$f_3(m,q,\alpha,T) = \frac{\partial f_1}{\partial m}\frac{\partial f_2}{\partial q} - \frac{\partial f_1}{\partial q}\frac{\partial f_2}{\partial m} = 0. \tag{3}$$

This equation also has a more physical interpretation in that it can be shown to correspond to the maximum value of m as a function α for T fixed.

Simultaneously solving equations (1) and (3) will give us the phase transition line we seek $\alpha_c(T_M)$.

3. Results.

The three equations (1) and (3) were solved for a range of temperatures between 1 and 0 using a Newton Raphson algorithm in three variables. The phase diagram derived using this numerical method is shown in figure 1 where the spin glass to paramagnetic phase transition is also shown for completeness. The important new result for this phase diagram is the reentrant phase which can be seen most clearly on the blow-up. Below $T = 0.023$ the gradient of the phase transition becomes positive and there is reentrant behaviour from the memory phase to the pure spin glass phase. The maximum value of α_c is obtained at $\alpha_c(T_M = 0.023) = 0.1382$ which is slightly above the zero temperature value of $\alpha_c(T_M = 0) = 0.1379$. The replica symmetry breaking line is also shown, this being the line below which the replica symmetric solution for the memory phase is known to be incorrect. This is the equivalent of the Almeida-Thouless line in the case of the SK spin glass [6] and it cuts the phase transition line just above the point

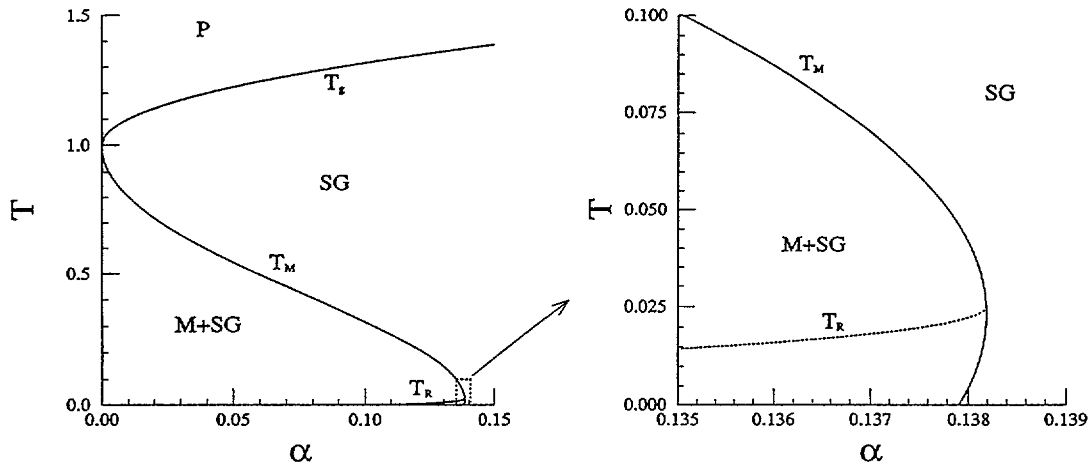

Fig.1. — Phase diagram for the Hopfield neural network model (originally presented in Ref. [3]) with low temperature blow-up. α is the number of patterns stored per spin. P, SG and M+SG, refer to the paramagnetic phase ($m = 0,\ q = 0$), pure spin glass phase ($m = 0,\ q \neq 0$) and the memory spin 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 line T_g and the memory states below the line T_M. T_R is the line below which replica symmetry is broken for the memory states. The reentrant memory to spin glass behaviour can clearly be seen on the low temperature blow-up.

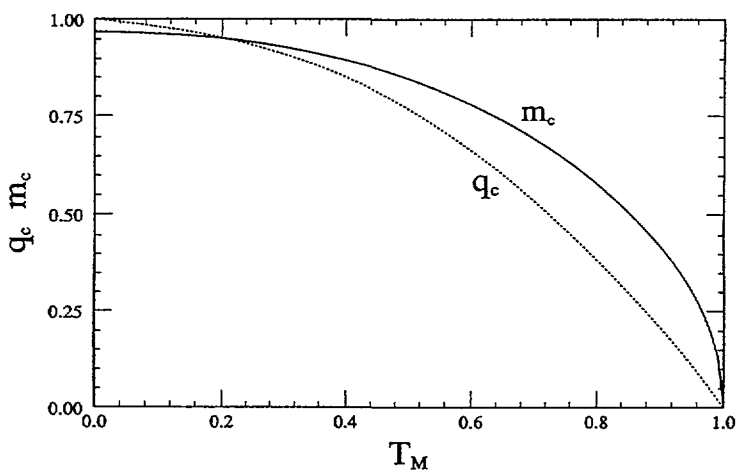

Fig.2. — Critical value of the order parameters $m_c(T_M)$ and $q_c(T_M)$ along the phase transition line.

at which its gradient changes sign. Figure 2 shows the magnetization $m_c(T_M)$ and spin glass order parameter $q_c(T_M)$ along the phase transition line. They both show a different behaviour from $\alpha_c(T_M)$ in that they always decreases as T increases. We also found theoretical solutions close to the $T = 0$ axis for equations (1) and (3) by expanding in T. To first order in T,

$$\begin{aligned} T_M &= -5.56 + 40.3\alpha_c(T_M) \\ m_c(T_M) &= m_c(0) - 0.023T_M \\ q_c(T_M) &= q_c(0) - 0.18T_M \end{aligned} \tag{4}$$

where $m_c(0) = 0.967$ and $q_c(0) = 1$. The first of these equations is written as a function of α_c rather than T_M so that it corresponds to the phase transition diagram (Fig. 1). It should be noted that this equation is different from that quoted in reference [3]. They found a gradient of -40 while we find a gradient of the same value but positive which signifies the reentrant behaviour. We can also see from these equations that as T increases it is only $\alpha_c(T_M)$ which increases while $m_c(T_M)$ and $q_c(T_M)$ decrease.

4. Discussion.

The main result of this work is to show that the replica solution of the Hopfield model has reentrant spin glass behaviour which seems characteristic of the replica symmetric calculation when applied to disordered systems. This behaviour has already been seen in the case of the bond diluted Hopfield model where the number of interactions per site grows more slowly than N (e.g. N^a, $0 < a < 1$) [7, 8]. In this case the reentrant behaviour is much more marked and the memory phase has the same properties as the ferromagnetic phase of the SK spin glass.

By analogy with spin glasses it is believed, for neural networks, that the true replica broken solution removes the reentrant phase and increases $\alpha_c(T_M)$ below the replica symmetry breaking line. This means that the maximum value of α_c found in replica theory will always be a lower bound for the true value. Thus our work gives, for the fully connected Hopfield model, a lower bound for the true storage capacity of $\alpha_c^{\max} \geq 0.1382$. This is still well below the value of $\alpha_c(T_M = 0) = 0.144$ [9] found by a one step broken symmetry calculation.

It is quite probable that all neural network phase diagrams based on the replica symmetric assumption will have reentrant spin glass behaviour. The magnitude of this reentrant behaviour may depend on the learning rule as well as the architecture of interactions. It is therefore important, for these models, to study the value of $\alpha_c(T)$ at all temperatures to find its maximum rather than just at zero temperature as most papers to date have done. This maximum is expected to be a lower bound for the true maximum.

Acknowledgements.

The authors would like to thank P. Wittwer for help in developing the numerical Newton Raphson algorithm used for solving the order parameter equations.

References

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

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

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

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

[5] Parisi G., *J. Phys.* **A13** (1980) L115; *J. Phys.* **A13** (1980) 1101; *J. Phys.* **A13** (1980) 1887; *Phys. Rev. Lett.* **50** (1983) 1946.

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

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

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

[9] Crisanti A., Amit D.J. and Gutfreund H., *Europhys. Lett.* **2** (1986) 337.