Optimal storage of a neural network model: a replica symmetry-breaking solution

1993 J. Phys. A: Math. Gen. 26 L61

(http://iopscience.iop.org/0305-4470/26/2/007)

J. Phys. A: Math. Gen. 26 (1993) L61–L68. Printed in the UK

LETTER TO THE EDITOR

Optimal storage of a neural network model: a replica symmetry-breaking solution

R Erichsen Jr†‡ and W K Thuemann‡

† Departamento de Física, Universidade Federal de Santa Maria, Campus Camobí, 97119–900 Santa Maria, RS, Brazil
‡ Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, 90046–900 Porto Alegre, RS, Brazil

Received 28 May 1992, in final form 28 August 1992

Abstract. A break in replica symmetry is found above the critical line $\alpha_c(\kappa)$ for storage with a finite minimal fraction of errors, within the first stage of the Parisi scheme, for noiseless networks with continuous synapses. The replica symmetry-breaking solution yields the highest (most stable) minimal fraction of errors everywhere above $\alpha_c(\kappa)$.

The storage properties arising from the collective behaviour of a neural network are determined by the set $\{J_{ij}\}$ of synaptic connections between neurons i and j; $i, j = 1, 2, \ldots, N$. The patterns ξ^μ, $\mu = 1, \ldots, p$ are fixed points of the network dynamics with sizeable basins of attraction if the set of local stabilities, defined by [1]

$$\gamma_i^\mu \equiv \frac{1}{\sqrt{N}} \sum_j J_{ij} \xi_i^\mu \xi_j^\mu \tag{1}$$

obeys the inequalities

$$\gamma_i^\mu > \kappa \tag{2}$$

for all the patterns, in the whole network, in which κ is a non-negative parameter. The ξ_i^μ are statistically independent random variables that assume, if unbiased, the values $+1$ or -1 with equal probability.

The original approach by Gardner [1] considers the synapses as dynamical variables. Given a storage level $\alpha \equiv p/N$ below a critical capacity, there is a finite fractional volume in configuration space where all the inequalities (2) are obeyed, together with the overall spherical constraint

$$\sum_j J_{ij}^2 = N \tag{3}$$

in the case of continuous synapses. The critical storage capacity $\alpha_c(\kappa)$ is attained when the fractional volume shrinks to zero.

The behaviour of the optimal network in the regime $\alpha > \alpha_c$ has also been explored allowing for a finite fraction of violations of (2) [2–4]. This can be achieved by the introduction of cost functions that account for the fraction and extent of the violations. In the simplest case considered in this paper, the Gardner–Derrida cost function [2], that only penalizes violations regardless of their size, is given by

$$E_i(\{J_{ij}\}) = \sum_{\mu=1}^{p} \left[1 - \Theta\left(\xi_i^\mu \frac{1}{\sqrt{N}} \sum_j J_{ij}\xi_j^\mu - \kappa\right)\right] \qquad (4)$$

where $\Theta(x) = 1$ if $x \geqslant 0$ and zero otherwise. The main issue addressed by Gardner and Derrida [2] is the minimal fraction $f_{\min}$ of ill-stored patterns in each neuron for a given $\alpha > \alpha_c(\kappa)$. This amounts to finding the ground-state energy per pattern,

$$f_{\min} = -\frac{1}{p} \lim_{\beta\to\infty} \frac{\mathrm{d}}{\mathrm{d}\beta} \langle \ln \mathcal{Z}(\beta)\rangle \qquad (5)$$

in the large-p limit, where

$$\mathcal{Z}(\beta) = \sum_{\{J_{ij}\}} \exp[-\beta E_i(\{J_{ij}\})] \qquad (6)$$

is the partition function in which β is an inverse temperature and the angular brackets denote the configuration average with the probability distribution for the random patterns.

The configurational average which is performed by means of the replica method, via the relationship

$$\langle \ln \mathcal{Z}(\beta)\rangle = \lim_{n\to 0} \frac{1}{n}(\langle \mathcal{Z}^n(\beta)\rangle - 1) \qquad (7)$$

yields a stable replica-symmetric solution below a critical line $\alpha_c(\kappa)$, where $f_{\min} = 0$, and within a bounded region above this line where $f_{\min} > 0$. The replica-symmetric solution becomes unstable beyond this region [2].

It has already been pointed out by Gardner and Derrida that, for continuous J_{ij}, the space of solutions is connected whenever the mean fraction of errors in the stability relations (2) is zero, and that it *could* be disconnected if this fraction is positive. In terms of the overlaps

$$q^{\alpha\beta} = \frac{1}{N} \sum_{j\neq i} J_{ij}^\alpha J_{ij}^\beta \qquad \alpha \neq \beta \qquad (8)$$

between replicas α and β, this means that there should be a single valley around the replica-symmetric overlap, $q = q^{\alpha\beta}$, in the first case, while the solutions could be made up of many valleys with different $q^{\alpha\beta}$s, in the second case.

Because of the great interest and the wide use that is made of networks with the Gardner–Derrida algorithm with continuous J_{ij}, it is important to establish if there is in fact, a relationship between replica symmetry breaking (RSB), i.e. a multiple-valley solution in a disconnected phase space, and a *non*-zero mean fraction of errors, and the purpose of the present article is to clarify this point, conjectured originally by

Gardner and Derrida. It is also our aim to obtain an estimate of the size of the effect of RSB. The relevance of RSB has only been established, so far, for the case of discrete $J_{ij} = \pm 1$ [5].

It will be shown here that there is a broken-symmetry solution in the replica space everywhere beyond $\alpha_c(\kappa)$, and that this solution tends continuously to the replica-symmetric solution as the critical line is aproached from above.

We take the formal results of Gardner and Derrida for the J_{ij} phase-space calculation in the large-N limit [2], that yields

$$\langle \ln Z \rangle = \mathrm{extr} N G(q_{\alpha\beta}, \phi_{\alpha\beta}, \varepsilon_\alpha) \tag{9}$$

where

$$G(q_{\alpha\beta}, \phi_{\alpha\beta}, \varepsilon_\alpha) = \alpha G_1(q_{\alpha\beta}) + G_2(\phi_{\alpha\beta}, \varepsilon_\alpha) + \mathrm{i} \sum_{\alpha<\beta} \phi_{\alpha\beta} q_{\alpha\beta} \tag{10}$$

in which

$$G_1(q_{\alpha\beta}) = \ln \prod_\alpha \left[\mathrm{e}^{-\beta} \int_{-\infty}^{\kappa} \frac{\mathrm{d}\lambda_\alpha}{2\pi} + \int_{\kappa}^{\infty} \frac{\mathrm{d}\lambda_\alpha}{2\pi} \right] \times \int_{-\infty}^{\infty} \mathrm{d}x_\alpha \exp\left[\mathrm{i} \sum_\alpha x_\alpha \lambda_\alpha - \frac{1}{2} \sum_\alpha (x_\alpha)^2 - \sum_{\alpha<\beta} q_{\alpha\beta} x_\alpha x_\beta \right] \tag{11}$$

and

$$G_2(\phi_{\alpha\beta}, \varepsilon_\alpha) = \ln \left\{ \int \prod_\alpha \mathrm{d}J_\alpha \exp\left[\mathrm{i} \sum_\alpha \varepsilon_\alpha (J_\alpha^2 - 1) - \mathrm{i} \sum_{\alpha<\beta} \phi_{\alpha\beta} J^\alpha J^\beta \right] \right\}. \tag{12}$$

Here, indices α and β denote replicas, while $\phi_{\alpha\beta}$ and ε_α are Lagrange multipliers that introduce the spherical constraint and the overlaps (8) respectively. The parameters are to be determined by the saddle-point equations

$$\begin{aligned} \frac{\partial}{\partial q_{\alpha\beta}} G(q_{\alpha\beta}, \phi_{\alpha\beta}, \varepsilon_\alpha) &= 0 \\ \frac{\partial}{\partial \phi_{\alpha\beta}} G(q_{\alpha\beta}, \phi_{\alpha\beta}, \varepsilon_\alpha) &= 0 \\ \frac{\partial}{\partial \varepsilon_\alpha} G(q_{\alpha\beta}, \phi_{\alpha\beta}, \varepsilon_\alpha) &= 0. \end{aligned} \tag{13}$$

There are locally stable symmetric solutions $q = q^{\alpha\beta}$, for all α, β, for the physically interesting overlaps in the regime below $\alpha_c(\kappa)$, and in a restricted region above, bounded by a de Almeida–Thouless stability line beyond which the replica-symmetric solution becomes unstable.

Next, we take the first step of the Parisi RSB scheme [6, 7] in which the n replicas are divided into identical n/m groups of m replicas each, m being first an integer which is then continued to the range $0 \leqslant m \leqslant 1$ in the limit $n \to 0$. The matrix elements $q_{\alpha\beta}$ take the value q_1 if α and β belong to the same group of replicas and

q_0 otherwise. Correspondingly, $\varphi_{\alpha\beta}$ takes the values q_1' and q_0' in the first and the second case, respectively.

In the limit $n \to 0$ the solutions of the saddle-point equations for q_1', q_0' and $\epsilon = \epsilon_\alpha$ can be eliminated in favour of the remaining q_0, q_1 and m, leading to

$$G^{(1)}_{\mathrm{RSB}}(q_0, q_1, m) = \frac{1}{2} + \frac{1}{2}\ln 2\pi + \frac{1}{2m}\ln\left[1 + m\frac{q_1 - q_0}{1 - q_1}\right] + \frac{1}{2}\ln(1 - q_1)$$
$$+ \frac{1}{2}\frac{q_0}{1 - q_1 + m(q_1 - q_0)} + \frac{\alpha}{m}\int \mathcal{D}z \ln \Psi(q_0, q_1, m) \qquad (14)$$

where

$$\mathcal{D}z \equiv \frac{\mathrm{d}z}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2} \qquad (15)$$

$$\Psi(q_0, q_1, m) = \int \mathcal{D}y \Phi^m(q_0, q_1) \qquad (16)$$

and

$$\Phi(q_0, q_1) = \mathrm{e}^{-\beta} + (1 - \mathrm{e}^{-\beta}) H\left(\frac{\kappa - z\sqrt{q_0} - y\sqrt{q_1 - q_0}}{\sqrt{1 - q_1}}\right) \qquad (17)$$

in which

$$H(x) \equiv \int_x^\infty \mathcal{D}z. \qquad (18)$$

In the limit $\beta \to \infty$, to which we restrict ourselves in the following, we look for solutions with $q_1 \to 1$, through the variable $x = \sqrt{2\beta(1 - q_1)}$. The solutions are minima in q_0, q_1 and m [7]. Note that the replica-symmetric solution of Gardner and Derrida [2] is recovered when $q_0 = q_1$ and $m \to 0$. In order to reach a finite minimal fraction of errors

$$f_{\min} = -\frac{1}{\alpha}\left\{\lim_{\beta\to\infty}\frac{1}{\beta}\min G^{(1)}_{\mathrm{RSB}}(q_0, q_1, m)\right\} \qquad (19)$$

we need also that $m \to 0$, in such a way that βm is finite.

Thus, in the limits $q_1 \to 1$, $m \to 0$ and $\beta \to \infty$, (14) becomes

$$\frac{1}{\beta}G^{(1)}_{\mathrm{RSB}}(q_0, x, M) = \frac{1}{M}\ln\varphi + \frac{q_0}{x^2\varphi} + \frac{2\alpha}{M}\int \mathcal{D}z \ln \Psi(q_0, x, M) \qquad (20)$$

where $M = 2\beta m$,

$$\varphi = \varphi(q_0, x, M) = 1 + M(1 - q_0)/x^2 \qquad (21)$$

and

$$\Psi(q_0, x, M) = \mathrm{e}^{-M/2}\left[1 - H\left(\frac{\kappa - z\sqrt{q_0} - x}{\sqrt{1 - q_0}}\right)\right] + \frac{1}{\sqrt{\varphi}}\exp\left[-\frac{M(\kappa - z\sqrt{q_0})^2}{2x^2\varphi}\right]$$
$$\times\left[H\left(\frac{\kappa - z\sqrt{q_0} - x\varphi}{\sqrt{(1 - q_0)\varphi}}\right) - H\left(\frac{\kappa - z\sqrt{q_0}}{\sqrt{(1 - q_0)\varphi}}\right)\right] + H\left(\frac{\kappa - z\sqrt{q_0}}{\sqrt{(1 - q_0)}}\right). \qquad (22)$$

The numerical solutions to the saddle-point equations for q_0, q_1 and m are shown in figures 1–3, for various values of κ in the region above the critical line $\alpha_c(\kappa)$. Clearly, there is an RSB solution that departs continuously from the symmetric solution as α is increased, starting from $\alpha_c(\kappa)$. As $\alpha \to \alpha_c(\kappa)$, $q_0 \to q_1$ and $x \to \infty$, in accordance with the result of Gardner and Derrida, while the new variable M goes to infinity. On the other hand, an analytical result in the large-α limit yields

$$M = \sqrt{\frac{\ln \alpha}{c_1 \alpha}} \tag{23}$$

and

$$x = \frac{1}{c_2}\sqrt{\frac{c_1}{\alpha \ln \alpha}} \tag{24}$$

where

$$c_1 = c_1(q_0) = \frac{1}{2}\int \mathcal{D}z \left[1 - H\left(\frac{\kappa - z\sqrt{q_0}}{\sqrt{1-q_0}}\right)\right] \tag{25}$$

and

$$c_2 = \mathrm{e}^{-\kappa^2/2}. \tag{26}$$

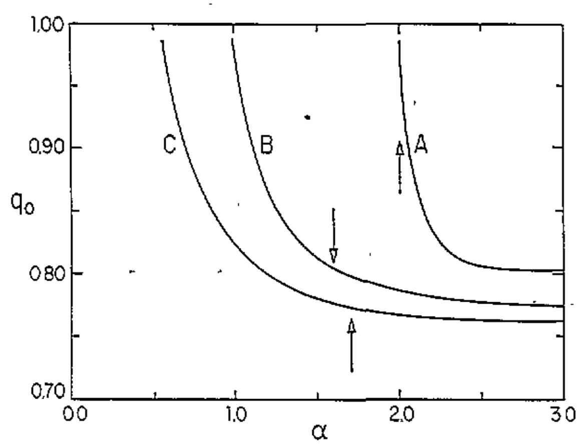

Figure 1. Saddle-point value of q_0 as a function of $\alpha = p/N$, for three values of κ (A, $\kappa = 0$; B, $\kappa = 0.5$; C, $\kappa = 1.0$). The arrow points to the value of α where the symmetric solution becomes unstable.

In figure 4 we display the numerical results for the minimal fraction of errors, $f_{\min}$, with RSB, which starts to grow continuously at $\alpha_c(\kappa)$, and we compare it with the symmetric solution. Thus, our results confirm the expectation, referred to earlier in this paper, that disconnected regions in phase space appear continuously with a growing fraction of errors. Although the appearance of these regions associated with

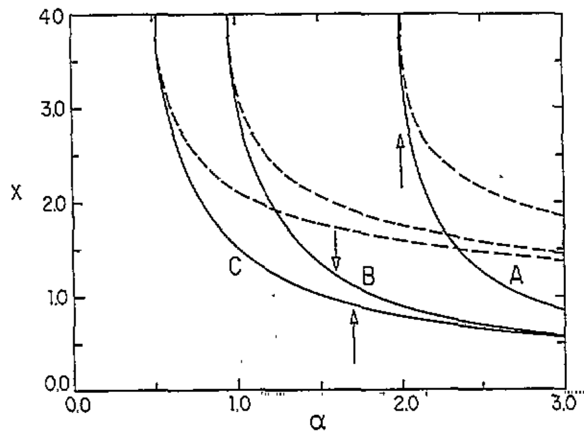

Figure 2. Saddle-point value of $x = [2\beta(1 - q_1)]^{1/2}$ as a function of α, for three values of κ (A, $\kappa = 0$; B, $\kappa = 0.5$; C, $\kappa = 1.0$). Full curves correspond to the RSB solution, while broken curves correspond to the symmetric solution. The arrows point to the values of α where the symmetric solution becomes unstable.

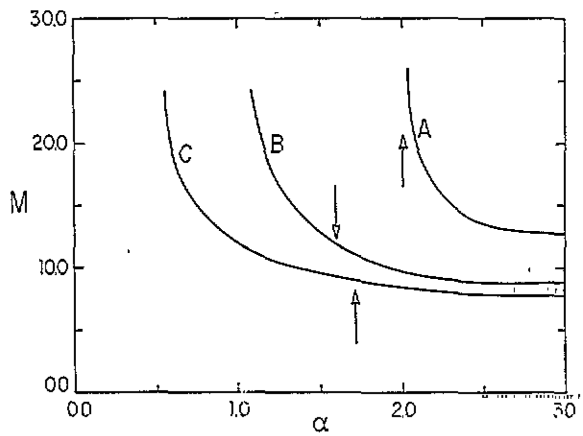

Figure 3. Saddle-point value of M as a function of α, for three values of κ (A, $\kappa = 0$; B, $\kappa = 0.5$; C, $\kappa = 1.0$). The arrows point to the values of α where the symmetric solution becomes unstable.

RSB requires a *higher* energy, that is a larger $f_{\min}$ than that for the symmetric case, we interpret the RSB solution, within the first step of the Parisi scheme, to be the correct solution. This is in spite of the statement demonstrated in spin-glass theory that, whenever there is a first-order transition with two (or more) stable solutions, one has to choose the one of *lowest* energy [8,9]. The situation here is somewhat similar to that of the Potts spin-glass, where thermodynamic consistency forces one to choose the solution of higher energy [8,10]. Had we chosen the lower energy (replica-symmetric) solution in our case, whenever it is stable below the de Almeida–Thouless stability limit, it would be thermodynamically impossible to reach the RSB solution beyond that limit, since there is a finite (free) energy difference between the two solutions. Rather, the transition from the replica-symmetric to the RSB mode

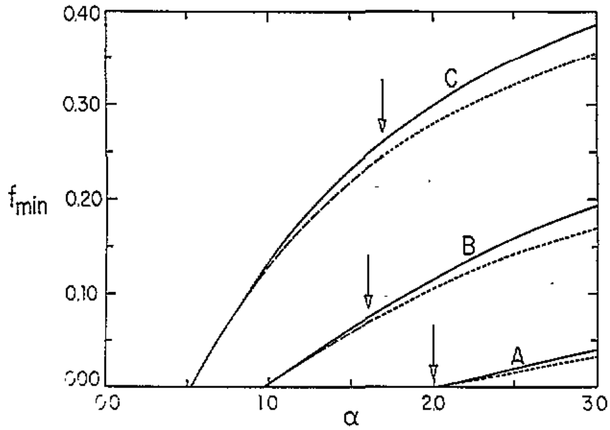

Figure 4. Minimal fraction of errors as a function of α, for the same values of κ as in the previous figures. The full curves represent the stable RSB solution, while the broken curves correspond to the replica-symmetric solution below the stability limit and the dotted curves to the continuation of this solution above.

takes place as soon as $f_{\min}$ starts to grow from zero.

It is worth noting that the RSB solution obtained in this work amounts to a finite break in replica symmetry while the usual replica-symmetric solution is stable to small symmetry-breaking perturbations.

Although we have not carried out the calculation for the second step of the Parisi RSB scheme, we expect the same qualitative results as those shown here within the first step.

In distinction to the case of discrete J_{ij}s, where RSB is necessary in order to yield a non-negative entropy near saturation [5], we need not report here on the effects of the RSB on the entropy for continuous J_{ij}s which can be negative. Rather, there are two interesting directions in which the present work can be extended. One is an analysis for finite β, to see when RSB may be disregarded, and the other is for correlated patterns that have been studied in a recent work [11]. These issues are being investigated at present.

To summarize, we have shown here that there is a break in replica symmetry above the critical line $\alpha_c(\kappa)$ in the network with continuous J_{ij} and that the corresponding multiple-valley picture appears as the more stable solution with a larger minimal fraction of errors than for the symmetric solution.

We acknowledge stimulating discussions with J F Fontanari, H Horner, M Griniasty and J-P Nadal. Part of this work was done while one of us (WKT) was visiting the Institute for Theoretical Physics, University of Heidelberg, Germany, under a FAPERGS (Fundação de Amparoà Pesquisa do Estado do Rio Grande do Sul, Brazil)-DAAD (Deutscher Akademischer Austausch Dienst) agreement. WKT thanks H Horner for his kind hospitality at Heidelberg and these agencies for support. This work was further supported in part by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil) and FINEP (Financiadora de Estudos e Projetos, Brazil). Most of the work of RE was undertaken at the Departamento de Física, Universidade Federal de Santa Maria.

References

[1] Gardner E 1988 *J. Phys. A: Math. Gen.* **1** 257
[2] Gardner E and Derrida B 1988 *J. Phys. A: Math. Gen.* **21** 271
[3] Amit D J, Evans M R, Horner H and Wong K Y M 1990 *J. Phys. A: Math. Gen.* **23** 3361
[4] Griniasty M and Gutfreund H 1991 *J. Phys. A: Math. Gen.* **24** 715
[5] Krauth W and Mézard M 1989 *J. Physique* **50** 3057
[6] Parisi G 1980 *J. Phys. A: Math. Gen.* **13** 1101
[7] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[8] Binder K and Young A P 1986 *Rev. Mod. Phys.* **58** 869
[9] Mottishaw P J and Sherrington D 1985 *J. Phys. C: Solid State Phys.* **18** 512
[10] Elderfield D and Sherrington D 1983 *J. Phys. C: Solid State Phys.* **16** L479
[11] Theumann W K and Erichsen Jr R 1991 *J. Phys. A: Math. Gen.* **24** L565