Gardner-Derrida neural networks with correlated patterns

1991 J. Phys. A: Math. Gen. 24 L565

(http://iopscience.iop.org/0305-4470/24/10/013)

J. Phys. A: Math. Gen. 24 (1991) L565–L571. Printed in the UK

LETTER TO THE EDITOR

Gardner–Derrida neural networks with correlated patterns

W K Theumann and Rubem Erichsen Jr

Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, 91500 Porto Alegre, RS, Brazil

Received 5 February 1991

Abstract. The storage properties of an optimal neural network with correlated patterns is studied allowing for a finite fraction of errors determined by the Gardner–Derrida cost function. A discontinuity in the probability distribution of the local stabilities is seen as a drastic decrease in the minimal fraction of errors. There is also an enlargement of the domain in which the replica symmetric solution is stable, allowing for higher storage capacities.

Learning in attractor neural networks amounts to organizing the space of network states into basins of attraction of preassigned memories [1]. A crucial role in organizing the network space is played by the synaptic matrix $\{J_{ij}\}$ between a pair of neurons i and j. Considering the J_{ij}s as dynamical variables, which need not be explicitly prescribed in terms of the stored patterns $\xi_i^\mu = \pm 1$ $(i = 1, \ldots, N;\ \mu = 1, \ldots, p)$ as in the Hopfield approach [1, 2], Gardner [3] has shown how to calculate the storage capacity (i.e. the maximum ratio $\alpha = p/N$) for the optimal network. This is a network in which all the stored patterns, which are fixed points of the dynamics, have finite (non-zero) basins of attraction that are guaranteed when the local stabilities

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

with the field h_i^μ at site i, satisfy the inequalities

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

for every i and μ, where κ is a positive constant. The constraint

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

is usually imposed to account for the freedom in κ due to an overall scaling.

Gardner and Derrida [4] have shown how to go beyond the storage capacity of the optimal network allowing for a minimal fraction of violations of (2) that is independent of their size and determined by the cost function

$$E = \sum_{i,\mu} \theta(\kappa - \Delta_i^\mu) \tag{4}$$

where $\theta(x) = 1$, for $x \geq 0$, and zero otherwise.

The minimal fraction of errors, $f_{\min}$, is related to the minimum average 'energy' E_0 by $pf_{\min} = E_0$, where

$$E_0 = \lim_{h\to\infty} \frac{\mathrm{d}}{\mathrm{d}h} \langle \ln Z(h) \rangle_{\{\xi_i^\mu\}} \tag{5}$$

0305-4470/91/100565+07$03.50

in which

$$Z = \langle \exp[-hE(\{J_{ij}\})] \rangle_{\{J_{ij}\}} \tag{6}$$

is the canonical partition function. The averages $\langle \ldots \rangle$ are either over the probability distribution of the stored patterns $\{\xi_i^\mu\}$ or over the continuous dynamical variables $\{J_{ij}\}$ satisfying the constraint (3).

When the replica method [5] is used to do the average over the quenched $\{\xi_i^\mu\}$ the overlap [3, 4]

$$q^{\alpha\beta} = \frac{1}{N} \sum_{j \neq i} J_{ij}^\alpha J_{ij}^\beta \qquad \alpha \neq \beta \tag{7}$$

appears between every pair of replicas α and β. Assuming replica symmetry, where $q^{\alpha\beta} = q$ for all $\alpha \neq \beta$, the critical storage capacity α_c is reached through a single ground-state configuration in the limit $q \to 1$. It is relevant to know the boundary of the stability region in the (α, κ) plane against replica-symmetry breaking, and this is now known for random patterns with a non-zero minimal fraction of errors [4].

Correlations between patterns are known to increase considerably the storage capacity of a network, albeit at the expense of the information content [3, 6]. In this note we consider the minimal fraction of errors and the storage capacity of networks that are ruled by the Gardner–Derrida [4] cost function using biased patterns with the statistically independent probability distribution [3, 6]:

$$p(\xi_i^\mu) = \tfrac{1}{2}(1+a)\delta(\xi_i^\mu - 1) + \tfrac{1}{2}(1-a)\delta(\xi_i^\mu + 1) \tag{8}$$

for each i and μ, where $a \in [-1, 1]$. This accounts for effective correlations $\langle \xi_i^\mu \xi_i^\nu \rangle = a^2$, for $\mu \neq \nu$. We restrict ourselves to the replica symmetric theory in which there is a second 'order' parameter $M = M^\alpha$ for all α, where

$$M^\alpha = \frac{1}{\sqrt{N}} \sum_{j \neq i} J_{ij}^\alpha . \tag{9}$$

Our results are as follows. First, there is a discontinuity in $M(\kappa)$ for a non-zero minimal fraction of errors. This leads always to a discontinuity in the ground-state energy E_0 and to a lower minimal fraction of errors, excluding thus a first-order transition in M. Note that h is an inverse temperature variable in equation (6) and in the zero temperature limit in (5) a first-order transition would require the same energy for two phases with different M. Second, on a somewhat more basic level, a discontinuity also appears in the distribution

$$\rho(\Lambda) = \langle \delta(\Lambda - \Delta_i^\mu) \rangle_{\{J, \xi\}} \tag{10}$$

of local stabilities [7] in which the average is over both $\{J_{ij}\}$ and the distribution of stored patterns. This distribution is of interest in itself since it belongs to a universality class discussed recently [8]. Third, the discontinuity in $M(\kappa)$ leads to a larger domain of stability for the replica symmetric solution in the (α, κ) plane, with a considerable increase in the storage capacity α for a given κ, at fixed a. In the following we only present and discuss our results which are based on now standard calculations [4, 8, 9].

The natural variable for the network near saturation in the limit $h \to \infty$ and $q \to 1$ is [4]

$$x \equiv [2h(1-q)]^{1/2} \tag{11}$$

which remains finite and is given by the solution of the equation

$$\alpha(\kappa, a)\left\{\tfrac{1}{2}(1+a)\int_{\tilde{\kappa}_--x}^{\tilde{\kappa}_-} Dz(z-\tilde{\kappa}_-)^2+\tfrac{1}{2}(1-a)\int_{\tilde{\kappa}_+-x}^{\tilde{\kappa}_+} Dz(z-\tilde{\kappa}_+)^2\right\}=1 \tag{12}$$

only for $\alpha > \alpha_c$, i.e., beyond the critical storage capacity of the errorless network of Gardner [3], determined by

$$\alpha_c(\kappa, a)\left\{\tfrac{1}{2}(1+a)\int_{-\infty}^{\tilde{\kappa}_-} Dz(z-\tilde{\kappa}_-)^2+\tfrac{1}{2}(1-a)\int_{-\infty}^{\tilde{\kappa}_+} Dz(z-\tilde{\kappa}_+)^2\right\}=1 \tag{13}$$

in which

$$Dz \equiv \frac{dz}{\sqrt{2\pi}}\exp(-z^2/2) \tag{14}$$

and

$$\tilde{\kappa}_\pm = (\kappa \pm aM)/\sqrt{1-a^2}. \tag{15}$$

Equation (12) is solved together with the equation for M,

$$(1+a)\int_{\tilde{\kappa}_--x}^{\tilde{\kappa}_-} Dz(z-\tilde{\kappa}_-) = (1-a)\int_{\tilde{\kappa}_+-x}^{\tilde{\kappa}_+} Dz(z-\tilde{\kappa}_+) \tag{16}$$

and, for simplicity, we assume a zero threshold in the state updating rule [3].

If $\alpha \leqslant \alpha_c$, we have $x=\infty$ and, accordingly,

$$f_{\min}=0 \tag{17}$$

recovering the Gardner results. On the other hand, for $\alpha > \alpha_c$ we find

$$f_{\min}=\tfrac{1}{2}(1+a)H(x-\tilde{\kappa}_-)+\tfrac{1}{2}(1-a)H(x-\tilde{\kappa}_+) \tag{18}$$

in which x is given by the solution of equation (12) and

$$H(y) \equiv \int_y^\infty Dz. \tag{19}$$

For the normalized probability distribution of local stabilities, equation (10), we obtain [8, 9]

$$\rho(\Lambda)=\tfrac{1}{2}(1+a)\tilde{\rho}_+(\Lambda)+\tfrac{1}{2}(1-a)\tilde{\rho}_-(\Lambda) \tag{20}$$

where

$$\tilde{\rho}_\pm(\Lambda)=\frac{1}{\sqrt{2\pi(1-a^2)}}\exp[-(\Lambda \mp aM)^2/2(1-a^2)]$$
$$\times[\theta(\Lambda-\kappa)+\theta(\kappa-x-\Lambda)]+[H(\tilde{\kappa}_\mp-x)-H(\tilde{\kappa}_\mp)]\delta(\Lambda-\kappa) \tag{21}$$

where $\tilde{\kappa}_\pm$ is given by (15). Besides the sharp gap $\kappa-x<\Lambda<\kappa$ and the Gaussian tail at extreme negative values of Λ, also present in the case of random patterns ($a=0$) [9], one has now shifts in the Gaussian and in the complementary error functions $H(y)$, due to the overlap between stored patterns. There is the same shift of $\pm aM$ and a 'renormalization' with $(1-a^2)^{-1/2}$ in both parts of $\tilde{\rho}_\pm(\Lambda)$. In this way, the distribution of local stabilities belongs to one of the universality classes discussed by Abbott and Kepler [8]. If there is a discontinuity in M, as we will see next, these shifts will be discontinuous and so will be $\rho(\Lambda)$. These discontinuities vanish in the low-activity limit $a \to 1$.

Incidentally, the minimal fraction of errors is recovered through the relationship [10]

$$f_{\min} = \int_{-\infty}^{\kappa} \rho(\Lambda)\, d\Lambda. \tag{22}$$

To illustrate analytically the discontinuity in M, consider the solution of equation (16) for small a. To lowest order we find

$$M = \left\{ \kappa \left[\mathrm{erf}\left(\frac{\kappa}{\sqrt{2}}\right) - \mathrm{erf}\left(\frac{\kappa - x}{\sqrt{2}}\right) \right] + \sqrt{\frac{2}{\pi}} [\exp(-\kappa^2/2) - \exp(-(\kappa - x)^2/2)] \right\} (\mathrm{D}(\kappa, x))^{-1} \tag{23}$$

in which the denominator is

$$\mathrm{D}(\kappa, x) = \mathrm{erf}\left(\frac{\kappa}{\sqrt{2}}\right) - \mathrm{erf}\left(\frac{\kappa - x}{\sqrt{2}}\right) - \sqrt{\frac{2}{\pi}}\, x \exp[-(\kappa - x)^2/2] \tag{24}$$

and where

$$\mathrm{erf}\left(\frac{y}{\sqrt{2}}\right) = 2 \int_0^y \mathrm{D}z \tag{25}$$

is the error function. For fixed and finite x, that is when $f_{\min} \neq 0$ for $\alpha > \alpha_c$, $\mathrm{D}(\kappa, x)$ starts being positive for small κ and changes sign at intermediate values while the numerator remains positive, leading to a divergent discontinuity. The behaviour for fixed a can be obtained from the numerical solution of (16) and the result is shown in figure 1 for $\alpha = 3 > \alpha_c(\kappa = 0)$ and four values of a. Note that there are always two branches. Similar results are obtained for $\alpha < \alpha_c(\kappa = 0)$ if $\kappa > \kappa_c(\alpha)$, the value of κ on the critical storage curve (cf figure 3, below).

The reader is reminded that $J_{ij}^{\alpha}/\sqrt{N}$ is the expansion parameter in Gardner's procedure and that M^{α} may become very large. If the J_{ij} were random and the learning rule is the perceptron algorithm discussed by Gardner [3], or an improved version of

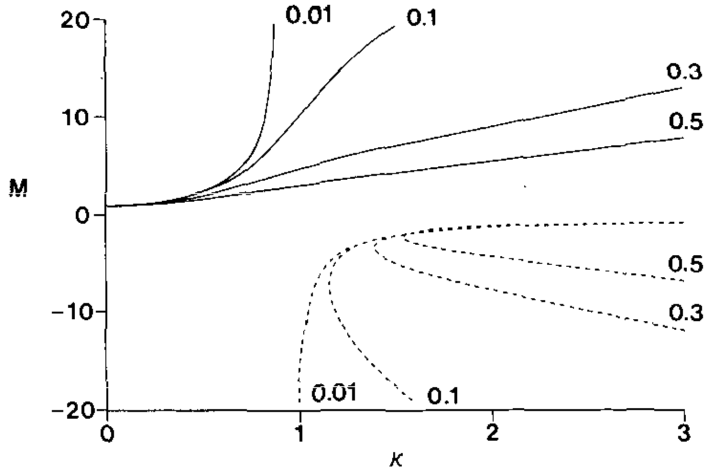

Figure 1. Order parameter M, equation (9), as a function of κ for $\alpha = 3 > \alpha_c(\kappa = 0)$, the limiting critical storage capacity, and $a = 0.01$, 0.1, 0.3 and 0.5.

it [11], M^α would be of O(1) but, otherwise, as in the case of correlated patterns, it does not have to be so.

The minimal fraction of errors, given by (18), is shown in figure 2, where the full curves for non-zero a correspond to the upper branch in figure 1. For reference, $f_{\min}$ for unbiased patterns is also shown there. Of course, only the lower branch can be thought as a true minimal fraction of errors. The reason for the rapid drop in $f_{\min}$ is the increasingly large aM in equations (18) and (19) as a becomes vanishingly small, with the limiting behaviour $f_{\min} \doteq \frac{1}{2}(1-a)$ already for $\kappa \sim 2$. Thus, there is a discontinuity in the limit $a \to 0$.

We consider next the critical line for replica symmetry breaking. Following Gardner and Derrida [4] in adapting the de Almeida and Thouless analysis for the spin-glass problem [12], we find that the replica symmetric solution is unstable and the expression for the minimal fraction of errors unjustified when

$$\frac{x}{\sqrt{2\pi}}\{\tfrac{1}{2}(1+a)\exp[-\tfrac{1}{2}(\tilde{\kappa}_- - x)^2] + \tfrac{1}{2}(1-a)\exp[-\tfrac{1}{2}(\tilde{\kappa}_+ - x)^2]\}$$
$$> \left\{\tfrac{1}{2}(1+a)\tilde{\kappa}_- \int_{\tilde{\kappa}_- - x}^{\tilde{\kappa}_-} \mathrm{D}z(\tilde{\kappa}_- - z) + \tfrac{1}{2}(1-a)\tilde{\kappa}_+ \int_{\tilde{\kappa}_+ - x}^{\tilde{\kappa}_+} \mathrm{D}z(\tilde{\kappa}_+ - z)\right\}. \qquad (26)$$

The critical lines for replica symmetry breaking (RSB) for various values of a are shown in figure 3 where on the left are the curves for the critical storage capacity of the optimal network ($f_{\min}=0$) and the locus of $f_{\min}=0.05$, extrapolating the replica symmetry results. The branches in the full curves correspond to the solutions for M that yield the true lower $f_{\min}$ of figure 2. These lines are always below the unphysical solutions, shown in the short broken curves, that correspond to the other branch of M, and also here one has a discontinuity in the critical line for replica symmetry breaking as $a \to 0$, leading to a sizeable increase of the domain where the replica calculations of this work are applicable.

The dramatic increase in storage capacity for a given κ beyond a certain minimum near the critical line is possible only with an increasingly large minimal fraction of

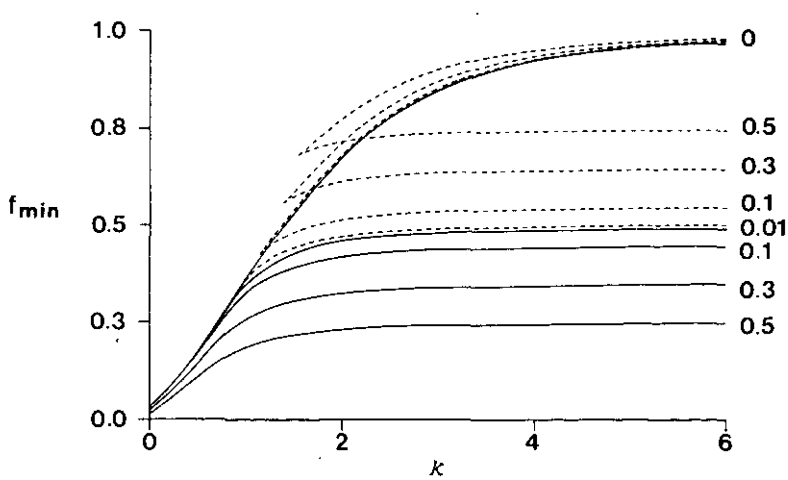

Figure 2. Minimal fraction of errors, in full curves, as a function of κ for $\alpha = 3 > \alpha_c(\kappa = 0)$ and $a = 0$, 0.01, 0.1, 0.3 and 0.5. The branch in the broken curves corresponds to the lower part in figure 1.

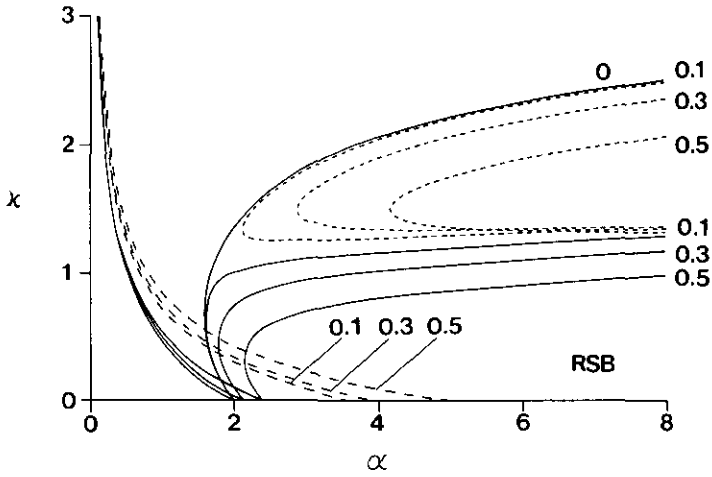

Figure 3. Critical storage capacity $\alpha_c(\kappa)$ for the optimal network and critical lines for replica symmetry breaking (RSB) shown in the full curves, for $a = 0.1$, 0.3, 0.5 and $a = 0$, for reference. There is a gap that is filled in by the unphysical solution in the previous figures shown in the short broken curves. The locus of $f_{\min} = 0.05$ is also shown (in the long broken curves) extrapolating the replica symmetry results.

errors. The lines of constant $f_{\min} > 0.05$, not shown for clarity in figure 3, progress towards larger α with increasing $f_{\min}$. However, due to the tail in $\rho(\Lambda)$ for for negative Λ, an increase in the storage capacity through a violation of (2) does not guarantee an attractor and one may have to include noise in order to have a network that is useful for retrieval [9]. The effects of noise with correlated patterns will be discussed elsewhere.

Preliminary results for the first-step hierarchical replica-symmetry breaking scheme of Parisi [13, 14], in terms of two overlaps, q_0 and q_1, their conjugate variables and the additional order parameter m, indicate that the only solution of the saddle-point equations [3] seems to be the trivial one: $q_0 = q_1$ and $m = 0$ of the replica symmetric theory [15], when $a = 0$, justifying a further study of this point.

It should also be interesting to construct the synaptic matrix for correlated patterns by means of a generalized perceptron learning rule. This and other issues will be considered in future work.

This work was supported in part by CNPq—Conselho Nacional de Desenvolvimento Científico e Technológico and FINEP—Financiadora de Estudos e Projetos, Brazil.

References

[1] Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554
[2] Hopfield J J 1984 *Proc. Natl Acad. Sci. USA* **81** 3088
[3] Gardner E 1988 *J. Phys. A: Math. Gen.* **21** 257
[4] Gardner E and Derrida B 1988 *J. Phys. A: Math. Gen.* **21** 271
[5] Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **32** 1972
[6] Amit D J, Gutfreund H and Sompolinsky H 1987 *Phys. Rev.* A **35** 2293
[7] Kepler T B and Abbott L F 1988 *J. Physique* **49** 1657
[8] Abbott L F and Kepler T B 1989 *J. Phys. A: Math. Gen.* **22** 2031

[9] Amit D J, Evans M R, Horner H and Wong K Y M 1990 *J. Phys. A: Math. Gen.* **23** 3361
[10] Griniasty M and Gutfreund H 1990 *Preprint*
[11] Abbott L F and Kepler T B 1989 *J. Phys. A: Math. Gen.* **22** L711
[12] de Almeida J R and Thouless D J 1978 *J. Phys. A: Math. Gen.* **11** 983
[13] Parisi G 1980 *J. Phys. A: Math. Gen.* **13** L115; 1101
[14] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (World Scientific: Singapore)
[15] Erichsen Jr R and Theumann W K unpublished