Thermodynamic properties of fully connected Q-Ising neural networks



1994 J. Phys. A: Math. Gen. 27 3411

(http://iopscience.iop.org/0305-4470/27/10/018)

J. Phys. A: Math. Gen. 27 (1994) 3411–3426. Printed in the UK

# Thermodynamic properties of fully connected $Q$-Ising neural networks

D Bollé†§, H Rieger‡‖ and G M Shim†¶

† Instituut voor Theoretische Fysica and Interdisciplinair Centrum voor Neurale Netwerken, B-3001 Leuven, Belgium

‡ Institut für Theoretische Physik, Universität Köln, D-50937 Köln, Germany

Received 16 December 1993

**Abstract.** The thermodynamic and retrieval properties of fully connected $Q$-Ising networks are studied in the replica-symmetric mean-field approximation. In particular, capacity–gain parameter and capacity–temperature phase diagrams are derived for $Q = 3, 4$ and $Q = \infty$ and different distributions of the stored patterns. Furthermore, the optimal gain function is determined in order to obtain the best performance. Where appropriate, the results are compared with the diluted and layered versions of these models.

## 1. Introduction

Neural networks with multi-state neurons are able to store and retrieve grey-toned patterns [1], which are useful, for instance, in the context of pattern recognition of pictures containing different grey levels. Special cases, including $Q$-Ising-type models, were studied for the first time in [2–4] and have been the subject of intensive research since then (see [5, 6] for an overview).

The $Q$-Ising model reveals a very rich function structure for the gain parameter of the input–output relation. The parallel dynamics of extremely diluted [3, 5, 7] and layered feedforward [8] versions of this model have been solved exactly at arbitrary temperatures and for arbitrary $Q$. The structure of the capacity–gain parameter and the capacity–temperature diagrams has been determined for $Q = 3, 4$ and $Q = \infty$.

Concerning the fully connected $Q$-Ising network, the overlap dynamics in the case of low loading [9] has been studied for arbitrary gain parameters, $b$, especially for $Q = 3$. In particular, it has been demonstrated that the retrieval state with perfect recall only appears in a restricted interval for $b$. This interval shrinks to zero as $Q \to \infty$. For extensive loading of uniform patterns, i.e. patterns taking equidistant states, and with the gain parameter fixed to be $b = \frac{1}{2}$, the equilibrium properties have been discussed in [1] on the basis of replica-symmetric mean-field theory. It has also been shown [1] that the storage capacity decreases with the number of grey levels as $Q^{-2}$, which has been confirmed in recent extensive numerical simulations [10]. A different approach to the $Q$-state problem has been presented in [11]. Optimal Gardner capacities for multi-state and analogue neurons have been determined [12–14] and their generalization abilities have been derived [15].

§ e-mail: FGBDA18@cc1.kuleuven.ac.be

‖ e-mail: rieger@thp.uni-koeln.de

¶ e-mail: FGBDA28@cc1.kuleuven.ac.be

0305-4470/94/103411+16$19.50  

Very recently, it has been demonstrated [16] that these networks exhibit a number of metastable retrieval phases growing as $Q^2$ for low loading and living on restricted intervals for $b$. Furthermore, it has been shown that the transitions between these phases are usually first order, and that there are strong hysteresis effects, as exemplified by a capacity–gain parameter diagram for $Q = 4$.

In this paper, we consider the fully connected $Q$-Ising neural network with arbitrary gain parameter and we investigate explicitly both its thermodynamic and retrieval properties for $Q = 3, 4$ and $Q = \infty$.

The rest of this paper is organized as follows. In section 2 the model is defined from a dynamical point of view. Section 3 briefly recalls the replica-symmetric mean-field approximation and presents the relevant fixed-point equations for general $Q$. In section 4 these equations are studied in detail for zero temperature and $Q = 3$ (section 4.1), $Q = 4$ (section 4.2) and $Q = \infty$ (section 4.3) for different distributions of the patterns. In particular, the function storage capacity for the gain parameter and the optimal gain parameter leading to the smallest Hamming distance are discussed. The results turn out to be significantly different for odd and even $Q$. Section 5 analyses the capacity–temperature diagrams and determines the thermodynamic properties of the $Q = 3$ (section 5.1) and $Q = \infty$ (section 5.2) models. They are compared with the diluted and layered versions of these models. Section 6 presents the concluding remarks. Finally, the appendix contains the specific fixed-point equations for the different values of $Q$ treated in the paper.

## 2. The model

We consider a network of $N$ neurons which can take values in the set of equidistant states

$$S_Q = \{s_k = -1 + 2(k-1)/(Q-1), k = 1, \ldots, Q\}. \tag{1}$$

In this network, we want to store $p = \alpha N$ patterns $\{\xi^\mu, \mu = 1, \ldots, \alpha N\}$ that are supposed to be independent and identically distributed random variables with zero mean and variance $A$. The latter is a measure for the activity of the patterns.

Given a configuration $\sigma = (\sigma_1, \ldots, \sigma_N)$, the local field $h_i$ of neuron $i$ is

$$h_i(\sigma) = \sum_{j \neq i} J_{ij}\sigma_j \tag{2}$$

where $J_{ij}$ are the synaptic couplings given by the well known Hebb rule

$$J_{ij} = \frac{1}{NA}\sum_{\mu=1}^{\alpha N} \xi_i^\mu \xi_j^\mu \qquad \text{for } i \neq j,\ J_{ii} = 0. \tag{3}$$

The neurons are updated asynchronously according to the transition probability

$$\Pr(\sigma_i' = s_k|\sigma) = \frac{\exp[-\beta\epsilon_i(s_k|h_i(\sigma))]}{\sum_{l=1}^{Q}\exp[-\beta\epsilon_i(s_l|h_i(\sigma))]}. \tag{4}$$

Here the inverse temperature $\beta = T^{-1}$ measures the noise level, and the energy potential $\epsilon_i(s|h)$ is taken [1, 4] to be

$$\epsilon_i(s|h) = -hs + bs^2 \qquad b > 0. \tag{5}$$

At zero temperature, $\sigma_i'$ takes the value $s_k$, leading to the minimum of the energy potential. This is equivalent to using an input–output relation

$$\sigma_i' = g[h_i(\boldsymbol{\sigma})]$$

$$g(x) = \sum_{k=1}^{Q} s_k[\theta(b(s_{k+1} + s_k) - x) - \theta(b(s_k + s_{k-1}) - x)] \tag{6}$$

with $s_0 = -\infty$ and $s_{Q+1} = \infty$. For finite $Q$ this input–output relation has a step-like shape and the parameter $b$ controls the steepness of the steps. For $Q = \infty$, the input–output function (6) becomes the piecewise linear function

$$g(x) = \begin{cases} \operatorname{sign}(x) & \text{if } |x| > 2b \\ \dfrac{x}{2b} & \text{otherwise.} \end{cases} \tag{7}$$

The slope of the linear part is given by $(2b)^{-1}$. In general, as $b$ goes to zero, the input–output relation reduces to that of the Ising-type network, independent of $Q$.

In the following, we present a detailed study of the properties of these fully connected networks as a function of $T$ and $b$ for different values of $A$.

## 3. Replica-symmetric mean-field theory

The long-time behaviour of the network under consideration is governed by the Hamiltonian

$$H = -\tfrac{1}{2}\sum_{i\neq j} J_{ij}\sigma_i\sigma_j + b\sum_i \sigma_i^2. \tag{8}$$

Using standard techniques [17] it is straightforward to show that the free energy within the replica-symmetry approximation is given [1] by

$$f(\beta) = \tfrac{1}{2}A\sum_{\mu=1}^{s} m_\mu^2 + \frac{\alpha q C}{2(1-C)^2} + \frac{\alpha}{2\beta}\left[\ln(1-C) + \frac{C}{1-C}\right] - \frac{1}{\beta}\left\langle\!\!\left\langle \int \mathrm{D}z\, \ln \mathrm{Tr}_\sigma \left\{ \exp\left[\beta(z\sqrt{\alpha r} + \sum_\mu m_\mu \xi^\mu - \tilde{b}\sigma)\sigma\right]\right\}\right\rangle\!\!\right\rangle \tag{9}$$

with

$$r = \frac{q}{(1-C)^2} \tag{10}$$

$$\tilde{b} = b - \frac{\alpha}{2}\frac{C}{1-C}. \tag{11}$$

Here $s$ denotes the number of condensed patterns, $\langle\!\langle \cdots \rangle\!\rangle$ indicates the average over the patterns $\{\xi^\mu\}$ and $\mathrm{D}z$ is the Gaussian measure

$$\mathrm{D}z = \mathrm{d}z\,(2\pi)^{1/2}\exp(-z^2/2). \tag{12}$$

In (9)–(11) $m_\mu$ represents the macroscopic overlap between pattern $\mu$ and the network state, $q$ is the Edwards–Anderson order parameter with its conjugate variable $r$ (the mean-square random overlap with the non-condensed patterns), $\tilde{b}$ is the effective gain parameter, and $C$ is the 'susceptibility' that is proportional to the fluctuation of the 'magnetization', i.e. $C = \beta(\langle\langle\sigma^2\rangle - \langle\sigma\rangle^2\rangle)$.

At this point, we note that this result (9) is the same as equation (11) in [1], except that the gain parameter $b$, in the energy potential, is now arbitrary instead of being fixed to be $\frac{1}{2}$. As a consequence, the effective gain parameter $\tilde{b}$ can be negative, implying that the input–output function reduces to that of Ising-type neurons. Furthermore, the variables $m_\mu, C, \tilde{b}$ correspond to the variables $m_\mu/C, \chi, \tilde{U}$ in [1].

The phase structure of the network is determined by that solution of the fixed-point equations for the order parameters

$$m_\mu = \frac{1}{A}\left\langle\!\!\left\langle \int \mathrm{D}z\, \xi^\mu \langle\sigma(z)\rangle \right\rangle\!\!\right\rangle \tag{13}$$

$$q = \left\langle\!\!\left\langle \int \mathrm{D}z\, \langle\sigma(z)\rangle^2 \right\rangle\!\!\right\rangle \tag{14}$$

$$C = \frac{1}{\sqrt{\alpha r}}\left\langle\!\!\left\langle \int \mathrm{D}z\, z\langle\sigma(z)\rangle \right\rangle\!\!\right\rangle \tag{15}$$

which maximizes $-\beta f(\beta)$. Here

$$\langle\sigma(z)\rangle = \frac{\mathrm{Tr}_\sigma\, \sigma \exp\left[\beta\sigma\left(\sum_\mu m_\mu \xi^\mu + \sqrt{\alpha r}z - \tilde{b}\sigma\right)\right]}{\mathrm{Tr}_s \exp\left[\beta s\left(\sum_\mu m_\mu \xi^\mu + \sqrt{\alpha r}z - \tilde{b}s\right)\right]}. \tag{16}$$

In the following sections we discuss these equations for $Q = 3, 4$ and $Q = \infty$ models.

## 4. Retrieval properties at zero temperature

### *4.1.* $Q = 3$

Let us consider a three-state network with patterns taking the values $\pm 1$ with probability $A/2$, and 0 with probability $(1 - A)$. For a Mattis retrieval state, say $m_\mu = m\delta_{\mu 1}$, the fixed-point equations (13)–(15) have the particular form given in the appendix (equations (A.1)–(A.5)). These fixed-point equations have been studied numerically. For $\tilde{b} \leqslant 0$ the corresponding equations can be further reduced by introducing the variable $x = m/\sqrt{2\alpha r}$. One arrives at

$$\sqrt{2\alpha} = \frac{\mathrm{erf}(x)}{x} - \frac{2}{\sqrt{\pi}}(A\mathrm{e}^{-x^2} + 1 - A) \tag{17}$$

together with, in view of (11), the following condition

$$b \leqslant b_0 = \sqrt{\frac{\alpha}{2\pi}}(A\mathrm{e}^{-x^2} + 1 - A). \tag{18}$$

Equation (17) only has solutions for $A \geqslant \frac{1}{3}$. For $A = 1$, i.e. for Ising patterns, equation (17) reduces to its analogue in the Hopfield model [17] and the maximum possible value of $\alpha$ is $\alpha_0 \approx 0.138$ for any $b \leqslant b_0 \approx 0.0151$. This decreases as $A$ gets smaller. For uniformly distributed patterns ($A = \frac{2}{3}$) we get $\alpha_0 \approx 0.0209$ for $b \leqslant b_0 \approx 0.0276$.

The $\alpha$–$b$ phase diagrams for the retrieval state ($m \neq 0$) are shown in figure 1 for $A = 1$ and $A = \frac{2}{3}$. The retrieval state disappears discontinuously at $\alpha_c$, denoted by the (thin) full curve. For uniformly distributed patterns there are two types of retrieval state (with $\tilde{b} > 0$). In region I $r$ is of order O(1), while in region II it is of order O(10). The retrieval region II does not appear for $A < \frac{1}{3}$, in contrast with the extremely diluted version of the model [3]. In fact, except for the absence of this region II, the phase diagram for $A = 0.01$ looks very similar to that for the diluted version, i.e. region I in figure 2($b$) in [3]. The retrieval state becomes the global minimum of the free energy below the thick full curve. There is a sharp drop at $b = \frac{1}{2}$ when entering the paramagnetic phase (the latter is indicated in the figure with a broken curve). Numerically, it is extremely hard to detect a substantial re-entrance. Here, we note that the free energy of the retrieval state at $\alpha = 0$ is $A(b - \frac{1}{2})$, while that of the paramagnetic state is always zero.

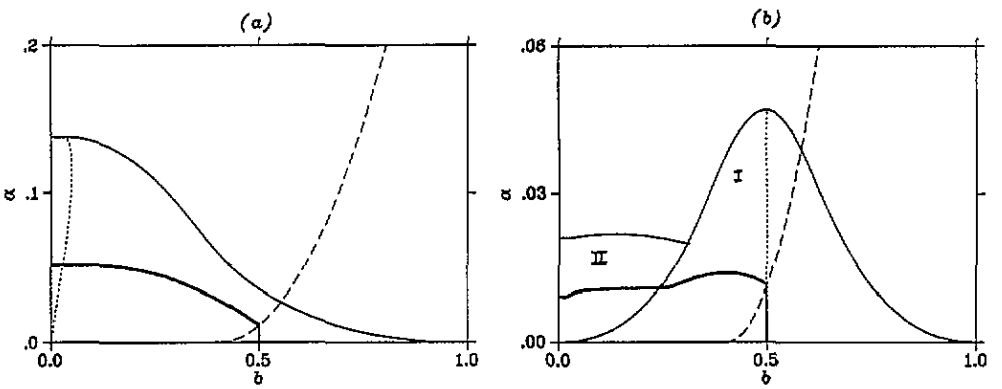


**Figure 1.** $Q = 3$ $\alpha$–$b$ phase diagrams at $T = 0$ for $A = 1$ and $A = \frac{2}{3}$, ($(a)$ and $(b)$, respectively). The (thin) full curve represents the maximal storage capacity $\alpha_c$, the thick full curve the thermodynamic transition of the retrieval state, the broken curve the spin–glass transition, and the dotted curve the optimal gain parameter. In $(b)$, there are two retrieval regions: in region I $r \approx$ O(1), while in region II $r \approx$ O(10).

The quality of the retrieval state is determined by the Hamming distance between the embedded pattern and the microscopic state of the network, i.e.

$$d_H(\xi^\mu, \sigma) = N^{-1} \sum_i (\xi_i^\mu - \sigma_i)^2.$$

This quantity depends not only on the overlap itself but also on the activity of the neurons $\langle\langle(\sigma^2)\rangle\rangle$. Figure 1 also shows the dotted curve as the best value for the gain parameter $b$ to obtain the minimal Hamming distance for a given value of $\alpha$.

Concerning the spin–glass phase we find that, for $\tilde{b} \leqslant 0$, $q = 1$ and $r = (1 + \sqrt{2/\pi\alpha})^2$, with the condition $\alpha > 2\pi b^2$. For positive $\tilde{b}$, the corresponding fixed-point equations can again be reduced into one equation in the variable $y = \tilde{b}/\sqrt{2\alpha}r$:

$$\sqrt{2\alpha} = \frac{b - 2y\mathrm{e}^{-y^2}/\sqrt{\pi}}{y\sqrt{E(y)} + \mathrm{e}^{-y^2}/(2\sqrt{\pi E(y)})} \tag{19}$$

where $E(y) = 1 - \mathrm{erf}(y)$. The right-hand side of (19) is lower bounded, giving a minimum value of $\alpha$. This minimum is zero for $b$ smaller than $\sqrt{2/\pi e} \approx 0.484$. Comparing the free energies of the spin–glass and the paramagnetic phase, we find the transition line displayed in figure 1. This line starts at $b \approx 0.405$ and the spin–glass solution disappears discontinuously at this boundary. We remark that this phase boundary is independent of $A$.

### 4.2. $Q = 4$

We expect the thermodynamic properties of a network consisting of neurons that are able to take on the zero state ($\sigma_i = 0$) to be significantly different from those of a network in which this state is forbidden for the neurons. This is due to the fact that in the latter a paramagnetic phase at zero temperature has to be absent. Although $m = q = r = 0$ is a solution of the saddle-point equations below, it does not have any physical meaning at zero temperature for even $Q$. This difference should, of course, become less significant for a higher number of possible states. Therefore, for completeness, we consider in this section the $Q = 4$ case.

According to (1), the neurons, as well as the patterns, can take on the values $-1, -1/3, +1/3, +1$. In analogy to the three-state network, where the parameter $A$ measures the weight of the extreme states $\pm 1$ in the pattern distribution, we consider a situation in which the patterns can take on the value $\pm 1$ with probability $\tilde{A}/2$ and $\pm 1/3$ with probability $(1 - \tilde{A})/2$ where $\tilde{A} = (9A - 1)/8$. The fixed-point equations for a retrieval state with $\tilde{b} > 0$ are given by (A.9)–(A.11) of the appendix. To be able to interpret the different phases occurring we note in passing that for the limiting case $\alpha \to 0$ one has the following retrieval solutions depending on the gain parameter $b$:

$$\begin{array}{ll} b < 1/4 & m = 3\Delta \\ 1/4 < b < 3\Delta/4 & m = 3\Delta \text{ and } 1 \\ 3\Delta/4 < b < 3/4 & m = 1 \text{ and } \Delta \\ 3/4 < b & m = \Delta \end{array} \tag{20}$$

with $\Delta = (1 + 2\tilde{A})/(1 + 8\tilde{A}) \in [1/3, 1]$. Hence, we will get a much richer phase diagram than for $Q = 3$.

For $\tilde{b} \leqslant 0$ the corresponding equations can again be reduced further by the introduction of the variable $x = m/\sqrt{2\alpha r}$ to

$$\sqrt{2\alpha} = \frac{1}{Ax}\left\{\tilde{A}\,\mathrm{erf}(x) + \frac{1-\tilde{A}}{3}\,\mathrm{erf}(x/3)\right\} - \frac{2}{\sqrt{\pi}}\left\{\tilde{A}\exp(-x^2) + (1-\tilde{A})\exp(-(x/3)^2)\right\} \tag{21}$$

which yields, in analogy to equation (17), the condition for $b$

$$b \leqslant b_0 = \sqrt{\frac{\alpha}{2\pi}}\left\{\tilde{A}\exp(-x^2) + (1-\tilde{A})\exp(-(x/3)^2)\right\}. \tag{22}$$

In contrast to the $Q = 3$ case, equation (21) always has a solution for small enough $\alpha$. The maximum possible value of $\alpha$ is again $\alpha = 0.138$, which is attained for $\tilde{A} = 1$ (Ising $\pm 1$ patterns) and for $\tilde{A} = 0$ (Ising $\pm 1/3$ patterns) for any $b \leqslant 0.015$ as in the $Q = 3$ case. The

maximal storage capacity is smaller for uniformly distributed patterns ($\tilde{A}$=1/2) and increases for decreasing, as well as increasing $\tilde{A}$.

The complete $\alpha$–$b$ diagram is depicted in figure 2 for various parameters $\tilde{A}$ by solving equations (A.9)–(A.11) numerically. Considering equation (20), we expect essentially three different phases. For $b \to \infty$, we find, for all values of $\tilde{A}$, a retrieval phase $FM_\Delta$ with $m \approx \Delta$ and a maximum storage capacity $\alpha_c$ equal to the storage capacity at $b = 0$. In the limit $b \to 0$, one finds a retrieval phase $FM_{3\Delta}$ with $m \approx 3\Delta$ and a maximum storage capacity that is determined by equation (21). In between these limiting cases, one finds (for $\tilde{A} \neq 0, 1$) a phase with optimal retrieval $FM_1$, which means the phase with the highest possible storage capacity $\alpha$ and the smallest Hamming distance. However, these three different phases need not be separated by discontinuity lines for $\tilde{A} = 0.5$ (figure 2(*c*)) and $\tilde{A} = 0.25$ (figure 2(*d*)). For instance, at $\tilde{A} = 0.75$ the $FM_{3\Delta}$- and the $FM_1$-phase can be deformed into each other continuously, as can be seen from figure 2(b). Finally, for $\tilde{A} = 0$ ($\tilde{A} = 1$), the $FM_\Delta$- ($FM_{3\Delta}$-) and $FM_1$-phases are identical. At this point we remark that the $FM$-phases for uniform patterns found in figure 2(*c*) confirm the results of [16].

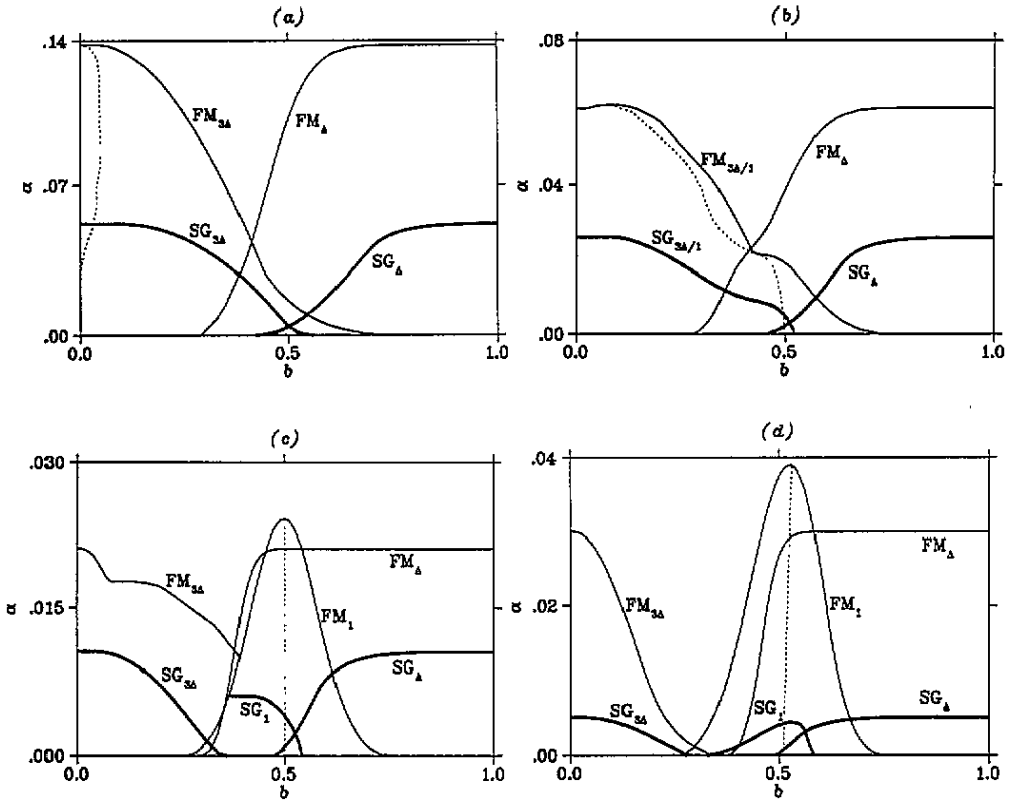


**Figure 2.** $Q = 4$ $\alpha$–$b$ retrieval phase diagrams at $T = 0$ for (*a*) $\tilde{A} = 1$, (*b*) $\tilde{A} = 0.75$, (*c*) $\tilde{A} = 0.5$ and (*d*) $\tilde{A} = 0.25$. The meaning of the lines is as in figure 1. The different phases according to (20) are indicated explicitly.

The largest possible storage capacity increases with $\tilde{A}$ and the value of $b$ that has to be chosen to achieve the maximal storage gets smaller, as expected. The location of the

characteristic bump in the phase diagram that indicates the $FM_1$-phase as well as the optimal retrieval line also shifts to smaller values of $b$ for increasing $\tilde{A}$.

In order to discuss the spin–glass (SG) phase ($m = 0$, $q > 0$) we have to solve the equations

$$q = 1 - \tfrac{8}{9}\,\mathrm{erf}\left(\frac{4\tilde{b}/3}{\sqrt{2\alpha r}}\right) \tag{23}$$

$$C = \frac{2}{3}\frac{1}{\sqrt{2\pi\alpha r}}\left\{1 + 2\exp\left(-\frac{(4\tilde{b}/3)^2}{2\alpha r}\right)\right\} \tag{24}$$

which are independent of $\tilde{A}$. In correspondence with the three different retrieval phases, we have to look for three different thermodynamic lines, $\alpha_{3\Delta}^{SG}(b, \tilde{A})$, $\alpha_1^{SG}(b, \tilde{A})$ and $\alpha_\Delta^{SG}(b, \tilde{A})$, where the SG (free) energy becomes smaller than that of the retrieval phases. For this reason we have to compare the energy per spin given by (note that $T = 0$)

$$E = -\tfrac{1}{2}m^2 A - \frac{\alpha r}{2} + q\left(b + \frac{\alpha}{2}\right). \tag{25}$$

In figures 2(*a*)–(*d*), we depict these thermodynamic lines for four different values of $\tilde{A}$. For $b \to \infty$, the value of $\alpha_\Delta^{SG}(b, \tilde{A})$ becomes $\alpha_{3\Delta}^{SG}(0, \tilde{A})$. As long as one can separate the $FM_1$-phase from the $FM_{3\Delta}$-phase, both lines, $\alpha_\Delta^{SG}(b, \tilde{A})$ and $\alpha_{3\Delta}^{SG}(b, \tilde{A})$, depend monotonically on $b$ and go to zero at some intermediate value of $b$. In this case (see figure 2(*c*) and (*d*)), one also has an $\alpha_1^{SG}(b, \tilde{A})$ line. One can also adopt the other point of view and ask for which value of $b$ at a given $\alpha$ does one obtain a thermodynamically stable retrieval state? This can also be read from figure 2.

### 4.3. $Q = \infty$

Finally, we turn to the case $Q = \infty$. Considering uniformly distributed patterns between $-1$ and 1 (and hence $A = \frac{1}{3}$), the fixed-point equations are still given by (13)–(15) with (16) replaced by

$$\langle\sigma\rangle = \frac{\int_{-1}^{1} \mathrm{d}\sigma\, \sigma \exp\left[\beta\sigma\left(\sum_\mu m_\mu \xi^\mu + \sqrt{\alpha r}\, z - \tilde{b}\sigma\right)\right]}{\int_{-1}^{1} \mathrm{d}s \exp\left[\beta s\left(\sum_\mu m_\mu \xi^\mu + \sqrt{\alpha r}\, z - \tilde{b}s\right)\right]}. \tag{26}$$

For the Mattis state, this leads to the explicit fixed-point equations presented in (A.12)–(A.14) of the appendix. These equations are written down for positive $\tilde{b}$. For $\tilde{b} \leqslant 0$ the limit $\tilde{b} \to 0$ has to be taken. (In this limit, $q = 1$.) In this case, one can further reduce the equations by introducing the variable $x = m/\sqrt{2\alpha r}$:

$$\sqrt{2\alpha} = \frac{1}{2}\left[\frac{\mathrm{erf}(x)}{x}\left(1 - \frac{3}{2x^2}\right) + \frac{3}{\sqrt{\pi}x^2}\exp(-x^2)\right] \tag{27}$$

together with, in view of (11), the following condition

$$b < b_0 = \sqrt{\frac{\alpha}{8}}\frac{\mathrm{erf}(x)}{x}. \tag{28}$$

The maximum possible value of $\alpha$ is $\alpha_0 \approx 0.0127$ for any $b \leqslant 0.0199$.

For $\tilde{b}$ positive, the transition line from the retrieval states to the spin–glass states, shown in the $\alpha$–$b$ phase diagram as a full curve (figure 3), has been obtained by solving (A.12)–(A.14) numerically. The transition is first order. At this point, it is interesting to compare this retrieval line with its analogues in the extremely diluted and layered feedforward $Q$-Ising neural networks [7, 8]. For all three models, the storage capacity is zero for $b \geqslant \frac{1}{2}$. For $b < \frac{1}{2}$, the storage capacity remains finite, though its value in the present model is relatively small compared with those mentioned above. This tells us that for increasing correlations among the neurons, the storage capacity decreases. As before, the retrieval state becomes the global minimum of the free energy under the thick full curve. The dotted curve indicates the values of the gain parameter $b$ needed to obtain the minimal Hamming distance.

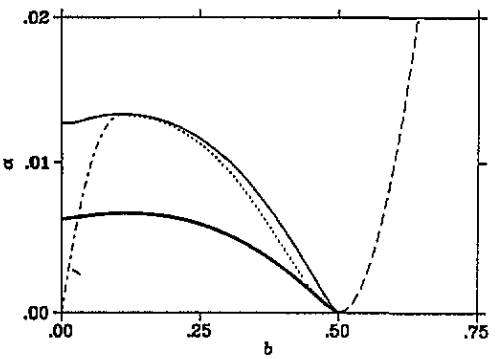


**Figure 3.** $Q = \infty$ $\alpha$–$b$ phase diagram at $T = 0$ for uniformly distributed patterns.

In contrast with the $Q = 3$ network, where the full memory phase at zero temperature is unstable against replica-symmetry breaking, the memory phase of the $Q = \infty$ model at zero temperature is only partially broken. The boundary below which breaking occurs is displayed as the chain curve in figure 3.

For the discussion of the spin–glass phase, one can follow the same argument as in the $Q = 3$ network, starting from one fixed-point equation analogous to (19). Here one finds that the minimum value $\alpha = 0$ occurs when $b \leqslant \frac{1}{2}$. For $b \geqslant \frac{1}{2}$, the region of existence of the spin–glass states is bounded by $\alpha > (b - \frac{1}{2})^2$, i.e. the broken curve in figure 3. The spin–glass transition is second order. These results are similar to the ones obtained in [6] for networks with graded-response neurons but *binary* stored patterns.

In the region where the spin–glass solutions appear, either the paramagnetic solutions do not exist or they lead to an ill-defined free energy. This can be seen by examining the solutions of the analogue of (30) together with the requirement $C \leqslant 1$.

We have also considered the effect of non-uniformity of the patterns by adding Ising patterns according to, for example, the output statistics $P(\xi^\mu) = \frac{1}{2}(1-a) + \frac{1}{2}a(\delta(\xi^\mu - 1) + \delta(\xi^\mu + 1))$ with $a$ between 0 and 1. The overall picture does not change. The actual values of the maximal capacity $\alpha$ in an $\alpha$–$b$ diagram decrease when decreasing $a$ from 1 to 0.

We end with the observation that the $\alpha$–$b$ diagrams for the fully connected and the layered feedforward $Q = \infty$ networks are similar to each other, while both are qualitatively different from that of the extremely diluted version of this network.

## 5. Phase diagrams for non-zero temperature

### *5.1.* $Q = 3$

We restrict ourselves to a discussion of the $\alpha$–$T$ phase diagrams for uniformly distributed patterns, but we consider different values of the gain parameter $b$.

Concerning the paramagnetic phase, we have found that it does not exist in a restricted region of the parameters of the network. Recalling the fixed-point equations (13)–(15) or (A.1)–(A.3), we see that the first two equations are trivially satisfied, such that the existence of this phase only depends on the equation for $C$. This leads to the following equation for the gain parameter $\tilde{b}$:

$$\tilde{b} = b - \frac{\alpha/2}{T(1 + \frac{1}{2}e^{\beta\tilde{b}}) - 1}. \tag{29}$$

For $\alpha = 0$ or $T = 0$, obviously, $\tilde{b} = b$. For finite $\alpha$, it is convenient to rewrite (29) in the form

$$\alpha = 2(b - \tilde{b})[T(1 + \tfrac{1}{2}e^{\beta\tilde{b}}) - 1]. \tag{30}$$

For $T > 1$, equation (30) always possesses a solution. For $0 < T \leqslant 1$, the situation is more complicated since the right-hand side is bounded. Therefore, given $b$ and $T$, no paramagnetic solution exists for $\alpha$ larger than this bound. At $T = 1$, this bound is $\exp(b-1)$; as $T$ approaches zero the bound goes to infinity.

If $b$ is smaller than 0.4631, then there is a temperature region $T_1 < T < T_2 < 1$ with $T_1$ and $T_2$ given as the solutions of the equation

$$\beta = 1 + \tfrac{1}{2}\exp(\beta b) \tag{31}$$

where $C > 1$, such that the free energy (9) of the paramagnetic state is not well defined. It is interesting to remark that the line given by equation (31) coincides with the boundary of the region where the zero state (at $\alpha = 0$) becomes unstable (see [9]).

All this is illustrated in the $\alpha$–$T$ phase diagrams of figure 4 for several values of $b$. In the region between the short-broken curve and (the far right-hand side of) the broken curve (starting at the merging point and ending at $\alpha = \exp(b - 1)$ for $T = 1$), no paramagnetic state exists. In fact, the broken curve represents the transition line between the spin–glass state and the paramagnetic state, and has been checked by comparing the corresponding free energies. This transition is first order. We further remark that the difference between these free energies diverges linearly as $\alpha$ goes to infinity.

The phase diagrams are completed by solving numerically the fixed-point equations (A.1)–(A.3) for the retrieval states, leading to the full curve. We also show the thermodynamic transition curve (thick full curve) below which the retrieval states become the global minima of the free energy. Finally, the chain curve denotes the Almeida–Thouless (AT) line, indicating the temperatures below which the replica-symmetric approximation is no longer valid. This curve is hardly seen on the scale of the figure.

The overall picture of the phase diagram for $b = \frac{1}{4}$ looks similar to that of the Hopfield model, while the phase diagram for $b = \frac{1}{2}$ is quite different in the sense that the paramagnetic phase exists between the retrieval phase and the spin–glass phase. (This has also been noted in [1], but the detailed phase diagram presented there is not complete.) This existence region of the paramagnetic phase extends in the direction of growing values of $\alpha$ for increasing $b$. This seems to be related to the fact that the larger $b$ is, the more the paramagnetic state becomes energetically favourable. In fact, one can easily check that, for $\alpha = 0$, the free energy of the paramagnetic state is the lowest for any $T$ if $b$ is greater than $\frac{1}{2}$.

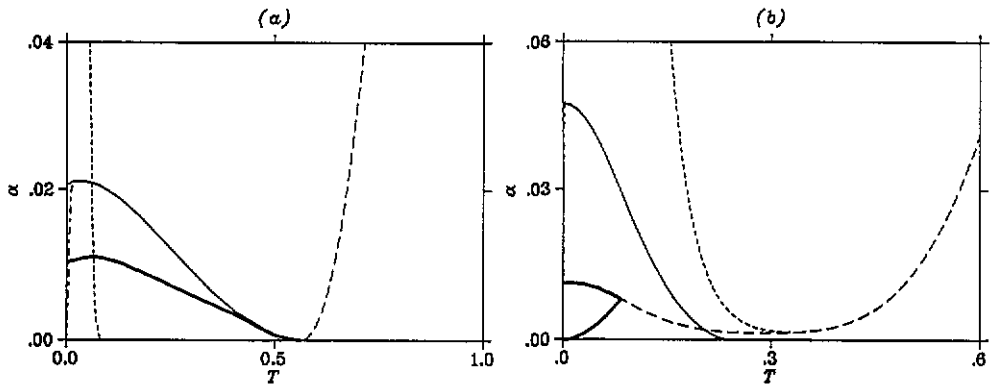


**Figure 4.** $Q = 3$ $\alpha$–$T$ phase diagrams for $b = \frac{1}{4}$ and $b = \frac{1}{2}$ and uniformly distributed patterns ((*a*) and (*b*), respectively). The meaning of the curves is as in figure 1. The chain curve is the AT line; the short broken curve indicates the border above which no paramagnetic state exists.

### 5.2. $Q = \infty$

In this case, the boundary of the spin–glass phase can be obtained by expanding the relevant equations (14)–(16) with respect to $r$. We arrive at

$$\alpha_{SG} = (C_0^{-1} - 1)^2 \tag{32}$$

with

$$C_0 = \beta \frac{\int_{-1}^{1} d\sigma\, \sigma^2 \exp(-\beta \tilde{b}_0 \sigma^2)}{\int_{-1}^{1} d\sigma\, \exp(-\beta \tilde{b}_0 \sigma^2)} \tag{33}$$

$$\tilde{b}_0 = b - \tfrac{1}{2}(C_0^{-1} - 1). \tag{34}$$

We expect that the spin–glass phase exists for all $\alpha$ if $T$ is not too high. Noting that $\alpha_{SG}$ becomes zero as $C_0 \to 1$, the temperature $T_{SG}$ below which spin–glass states exist for all $\alpha$ can be determined by the condition $C_0 = 1$ with $\tilde{b}_0 = b$. Such a solution is only possible for $b \leqslant \frac{1}{2}$.

To determine the thermodynamic transition line for this phase, we first need to investigate the properties of the paramagnetic phase. The existence of the latter depends on

$$\alpha = 2(b - \tilde{b})(C^{-1} - 1) \tag{35}$$

where $C$ is given by the right-hand side of (33) with $\tilde{b}_0$ replaced by $\tilde{b}$. By determining the limit of the right-hand side of (35) for $\tilde{b} \to \pm\infty$, we learn that the paramagnetic phase exists for all $\alpha$ if $T > 1$, while it does not exist for $\alpha > \alpha_{PM}$ if $T \leqslant 1$. Furthermore, the condition $C < 1$ implies that the derivative of this right-hand side at $\tilde{b} = b$ has to be negative, or $C(\tilde{b} = b) \leqslant 1$. Recalling that this is exactly the condition for obtaining $T_{SG}$, it means that there is no paramagnetic phase for $T < T_{SG}$ for all $\alpha$. Of course, in general, the spin–glass phase can exist below the boundary of the region for non-existence of the paramagnetic phase. Therefore, it is interesting to compare the free energies of both phases

for given $T > T_{SG}$ and $b$. Taking $\alpha = \alpha_{SG}(1+\epsilon)$, where $\epsilon$ is a small positive number, we get

$$\Delta f \approx \epsilon \frac{2T\sqrt{\alpha_{SG}}}{1+\sqrt{\alpha_{SG}}}\left(3 - \frac{\langle\sigma^4\rangle_0}{(\langle\sigma^2\rangle_0)^2}\right) \tag{36}$$

where

$$\langle\sigma^n\rangle_0 = \frac{\int_{-1}^{1} d\sigma\, \sigma^n \exp(-\beta\tilde{b}_0\sigma^2)}{\int_{-1}^{1} d\sigma\, \exp(-\beta\tilde{b}_0\sigma^2)} \tag{37}$$

*with $\tilde{b}_0$ being the solution of (34) at $\alpha = \alpha_{SG}$. We find that the energy difference $\Delta f$ is* always positive near $\alpha_{SG}$ and that it diverges linearly with $\alpha$ for $T > 1$ in the limit $\alpha \to \infty$. Therefore, the paramagnetic phase is always energetically favourable in comparison with the spin–glass phase. This means that the transition line between the spin–glass phase and the paramagnetic phase coincides with the boundary of existence of the paramagnetic states and that it is of first order.

To complete the $\alpha$–$T$ phase diagram we have calculated the temperature dependence of the storage capacity by solving numerically the fixed-point equations (A.12)–(A.14). The transition to the spin–glass states is first order.

Figure 5 shows the phase diagram for a typical value of the gain parameter, $b = \frac{1}{4}$. It is qualitatively similar to the $\alpha$–$T$ diagram for the Hopfield model. The full curve bounds the retrieval region. The retrieval states again become the global minima of the free energy below the thick full curve. The broken curve corresponds to the spin–glass transition discussed above.

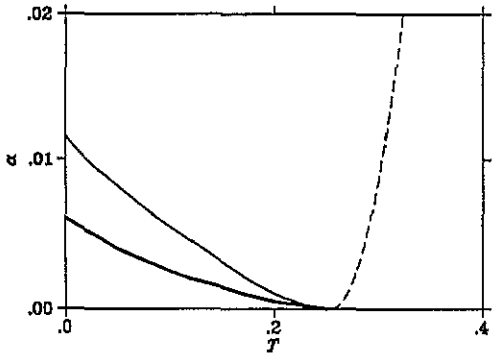


**Figure 5.** $Q = \infty$ $\alpha$–$T$ phase diagram and for $b = \frac{1}{4}$ and uniformly distributed patterns.

## 6. Concluding remarks

We have considered both the thermodynamic properties and retrieval properties of fully connected $Q$-Ising networks. Fixed-point equations have been derived for general temperature and arbitrary $Q$ in the replica-symmetric mean-field approximation. For $Q = 3, 4$ and $Q = \infty$, capacity–gain parameter diagrams and capacity–temperature phase

diagrams have been discussed in detail for different distributions of the stored patterns. The optimal gain function leading to the smallest Hamming distance has been determined.

Concerning the capacity–gain parameter diagrams, we find that the results are essentially different for odd and even $Q$ and that they are extremely dependent on the pattern activity. Furthermore, there are also interesting differences with the extremely diluted and layered feedforward versions of the models [7, 8]. Although the results for the latter are obtained using an exact dynamical approach, such a comparison makes sense since, in all cases we have studied here, replica-symmetry-breaking effects turn out not to be significant, as illustrated by the AT lines.

In the case of uniformly distributed patterns, we see that, for $Q = 3$, different retrieval regions show up for small $b$ and the capacity is reduced by a factor compared with the capacity for the extremely diluted and layered architectures. The line of optimal Hamming distance is given exactly by $b = \frac{1}{2}$; in the extremely diluted model it is located for the whole retrieval region in the interval $b \in [0.4, 0.5]$, while in the layered model it bends to smaller $b$ for growing $\alpha$.

For $Q = 4$, the storage capacity behaves entirely differently as a function of $b$. First of all, it does not decrease for $b > \frac{1}{2}$ and $\alpha(b = 0) = \alpha(b = \infty)$. Furthermore, there are three different retrieval phases. The optimal Hamming distance is again exactly $b = \frac{1}{2}$, while it is located in the interval $[0.4, 0.5]$ for the diluted model. Here, it is interesting to remark that for non-uniformly distributed patterns this optimal line is also always situated in the $FM_1$ retrieval phase (where $m = 1$).

For $Q = \infty$, the diagram for the layered and fully connected models are very similar in shape (but the capacity is reduced roughly by a factor of 10 in the latter). It is simple compared with the $Q = 4$ diagram. The line of optimal Hamming distance is close to $b = \frac{1}{2}$ in the whole retrieval region only for the diluted model. For the other architectures, it quickly bends towards smaller values.

From all this information, together with the detailed results on low loading ($\alpha = 0$) for the fully connected model [9, 16], we conclude that optimal retrieval with nearly vanishing Hamming distance becomes harder as $Q$ increases. To study optimal retrieval in networks with higher values of $Q$, we think it is sufficient to look at the retrieval phase with $m = 1$. For $Q = \infty$, the retrieval states are never perfectly correlated with the stored patterns.

Looking finally at the $\alpha$–$T$ diagrams, we immediately notice that for $Q = 3$ the situation is rather complicated, even for uniform patterns, and depends very much on the value of $b$. For $b$ close to, and greater than, the optimal $b = \frac{1}{2}$, the phase diagram is completely different from that of the Hopfield model, in the sense that the paramagnetic phase exists between the retrieval and the spin–glass phase. For $Q = 4$, we expect an even more complex behaviour, as exemplified already by the $\alpha$–$b$ diagrams, such that we defer a discussion of the temperature dependence to another occasion. For $Q = \infty$, the diagram is relatively simple again and qualitatively resembles that of the Hopfield model.

## Acknowledgments

This work has been supported in part by the Research Fund of the K U Leuven (grant OT/91/13). The authors are indebted to S Bös, J Huyghebaert, R Kühn and B Vinck for stimulating discussions. One of us (DB) would like to thank the Belgian National Fund for Scientific Research for financial support. The work of HR was performed within the Sonderforschungsbereich (SFB) 341 Köln-Aachen-Jülich.

## Appendix

In this appendix, we write down explicitly the fixed-point equations (13)–(16) for $Q = 3$, $Q = 4$ and $Q = \infty$. For a three-state network with patterns taking the values $\pm 1$ with probability $A/2$ and 0 with probability $(1 - A)$, we obtain for a Mattis state

$$m = \int \mathrm{D}z\, V_\beta(m + \sqrt{\alpha r}z, \tilde{b}) \tag{A.1}$$

$$q = \int \mathrm{D}z \left[ A V_\beta^2 \left( m + \sqrt{\alpha r}z, \tilde{b} \right) + (1 - A) V_\beta^2 \left( \sqrt{\alpha r}z, \tilde{b} \right) \right] \tag{A.2}$$

$$C = \frac{1}{\sqrt{\alpha r}} \int \mathrm{D}z\, z \left[ A V_\beta \left( m + \sqrt{\alpha r}z, \tilde{b} \right) + (1 - A) V_\beta \left( \sqrt{\alpha r}\, z, \tilde{b} \right) \right] \tag{A.3}$$

where

$$V_\beta(x, y) \equiv \frac{\sinh(\beta x)}{\frac{1}{2}\exp(\beta y) + \cosh(\beta x)}. \tag{A.4}$$

For zero temperature, the function $V_\beta(x, y)$ reduces to

$$V_\infty(x, y) = \mathrm{sign}(x)\theta(|x| - y) \tag{A.5}$$

and hence the Gaussian variable $z$ can be integrated out explicitly. For $\tilde{b} > 0$, one essentially obtains equations (15) of [1], taking into account the effective gain parameter (11).

For a $Q = 4$ model in which the patterns can take the value $\pm 1$ with probability $\tilde{A}/2$ and $\pm 1/3$ with probability $(1 - \tilde{A})/2$, where $\tilde{A} = (9A - 1)/8$, the fixed-point equations (13)–(16) for a Mattis state read

$$\begin{aligned} m &= \frac{1}{A} \int \mathrm{D}z \left[ \tilde{A} U_\beta \left( m + z\sqrt{\alpha r}, \tilde{b} \right) + (1 - \tilde{A}) U_\beta \left( m/3 + z\sqrt{\alpha r}, \tilde{b} \right) \right] \\ q &= \int \mathrm{D}z \left[ \tilde{A} U_\beta^2 \left( m + z\sqrt{\alpha r}, \tilde{b} \right) + (1 - \tilde{A}) U_\beta^2 \left( m/3 + z\sqrt{\alpha r}, \tilde{b} \right) \right] \\ C &= \frac{1}{\sqrt{\alpha r}} \int \mathrm{D}z\, z \left[ \tilde{A} U_\beta \left( m + z\sqrt{\alpha r}, \tilde{b} \right) + (1 - \tilde{A}) U_\beta \left( m/3 + z\sqrt{\alpha r}, \tilde{b} \right) \right] \end{aligned} \tag{A.6}$$

where

$$U_\beta(x, y) \equiv \frac{\sinh(\beta x) + \frac{1}{3}\exp(8\beta y/9)\sinh(\beta x/3)}{\cosh(\beta x) + \exp(8\beta y/9)\cosh(\beta x/3)}. \tag{A.7}$$

Again, for zero temperature, the function $U_\beta(x, y)$ reduces to

$$U_\infty(x, y) = \mathrm{sign}(x)[1 + 2\theta(|x| - 4y/3)]. \tag{A.8}$$

In that case, equation (A.6) implies for $\tilde{b} > 0$

$$m = \frac{1}{3A}\left\{\tilde{A}\,\mathrm{erf}\left(\frac{3m+4\tilde{b}}{3\sqrt{2\alpha r}}\right) + \tilde{A}\,\mathrm{erf}\left(\frac{3m-4\tilde{b}}{3\sqrt{2\alpha r}}\right) + \tilde{A}\,\mathrm{erf}\left(\frac{m}{\sqrt{2\alpha r}}\right) + \frac{1-\tilde{A}}{3}\,\mathrm{erf}\left(\frac{m+4\tilde{b}}{3\sqrt{2\alpha r}}\right) + \frac{1-\tilde{A}}{3}\,\mathrm{erf}\left(\frac{m-4\tilde{b}}{3\sqrt{2\alpha r}}\right) + \frac{1-\tilde{A}}{3}\,\mathrm{erf}\left(\frac{m}{3\sqrt{2\alpha r}}\right)\right\} \tag{A.9}$$

$$q = 1 - \frac{4}{9}\left\{\tilde{A}\,\mathrm{erf}\left(\frac{3m+4\tilde{b}}{3\sqrt{2\alpha r}}\right) - \tilde{A}\,\mathrm{erf}\left(\frac{3m-4\tilde{b}}{3\sqrt{2\alpha r}}\right) + (1-\tilde{A})\,\mathrm{erf}\left(\frac{m+4\tilde{b}}{3\sqrt{2\alpha r}}\right) - (1-\tilde{A})\,\mathrm{erf}\left(\frac{m-4\tilde{b}}{3\sqrt{2\alpha r}}\right)\right\} \tag{A.10}$$

$$C = \frac{2}{3}\frac{1}{\sqrt{2\pi\alpha r}}\left\{\tilde{A}\exp\left(-\frac{(3m+4\tilde{b})^2}{18\alpha r}\right) + (1-\tilde{A})\exp\left(-\frac{(m+4\tilde{b})^2}{18\alpha r}\right) + \tilde{A}\exp\left(-\frac{m^2}{2\alpha r}\right) + (1-\tilde{A})\exp\left(-\frac{m^2}{18\alpha r}\right) + \tilde{A}\exp\left(-\frac{(3m-4\tilde{b})^2}{18\alpha r}\right) + (1-\tilde{A})\exp\left(-\frac{(m-4\tilde{b})^2}{18\alpha r}\right)\right\}. \tag{A.11}$$

Finally, for $Q = \infty$ and uniformly distributed patterns ($A = \frac{1}{3}$), a retrieval state is given by the solution of the fixed-point equations (13)–(15) and (26). For the Mattis state at zero temperature, $\langle\sigma\rangle = \tilde{g}(m\xi + \sqrt{\alpha r}z)$, where the effective input–output function $\tilde{g}$ is given by (7) with $b$ replaced by the effective gain parameter $\tilde{b}$. This allows us to perform the Gaussian average in the fixed-point equations explicitly resulting in

$$m = \frac{3}{2}\int_{-1}^{1} \mathrm{d}\xi\,\xi\left[\left(1+\frac{m\xi}{2\tilde{b}}\right)\mathrm{erf}\left(\frac{2\tilde{b}+m\xi}{\sqrt{2\alpha r}}\right) + \frac{1}{\tilde{b}}\sqrt{\frac{\alpha r}{2\pi}}\exp\left(-\frac{(2\tilde{b}+m\xi)^2}{2\alpha r}\right)\right] \tag{A.12}$$

$$q = 1 + \frac{1}{2}\int_{-1}^{1}\mathrm{d}\xi\left[\left(\frac{\alpha r + (m\xi)^2}{(2\tilde{b})^2} - 1\right)\mathrm{erf}\left(\frac{2\tilde{b}+m\xi}{\sqrt{2\alpha r}}\right) + \frac{1}{\tilde{b}}\sqrt{\frac{\alpha r}{2\pi}}\left(\frac{m\xi}{2\tilde{b}} - 1\right)\exp\left(-\frac{(2\tilde{b}+m\xi)^2}{2\alpha r}\right)\right] \tag{A.13}$$

$$C = \frac{1}{4\tilde{b}}\int_{-1}^{1}\mathrm{d}\xi\ \mathrm{erf}\left(\frac{2\tilde{b}+m\xi}{\sqrt{2\alpha r}}\right) \tag{A.14}$$

for positive $\tilde{b}$. We remark that it is also straightforward to perform explicitly the integration associated with the random patterns but we do not need to write down the resulting algebraically complex expressions.

## References


[1] Rieger H 1990 *J. Phys. A: Math. Gen.* **23** L1273
[2] Meunier C, Hansel D and Verga A 1989 *J. Stat. Phys.* **55** 859
[3] Yedidia J S 1989 *J. Phys. A: Math. Gen.* **22** 2265
[4] Rieger H 1989 *PhD Thesis* University of Cologne, unpublished; 1990 *Statistical Mechanics of Neural Networks (Proc. XI. Sitges Conf.) (Springer Lecture Notes in Physics 368)* ed L Garrido (Berlin: Springer) p 33
[5] Bollé D, Vinck B and Zagrebnov V A 1993 *J. Stat. Phys.* **70** 1099
[6] Kühn R and Bös S 1993 *J. Phys. A: Math. Gen.* **26** 831
[7] Bollé D, Shim G M, Vinck B and Zagrebnov V A 1994 *J. Stat. Phys.* **74** to appear
[8] Bollé D, Shim G M and Vinck B 1994 *J. Stat. Phys.* **74** to appear
[9] Bollé D, Dupont P and Vinck B 1992 *J. Phys. A: Math. Gen.* **25** 2859
[10] Stiefvater T and Müller K R 1992 *J. Phys. A: Math. Gen.* **25** 5919
[11] Kohring G A 1992 *J. Physique* **2** 1549; 1993 *Neural Networks* **6** 573
[12] Mertens S, Köhler H M and Bös S 1991 *J. Phys. A: Math. Gen.* **24** 4941
[13] Bollé D, Dupont P and van Mourik J 1991 *Europhys. Lett.* **15** 893
[14] Bollé D, Kühn R and van Mourik J 1993 *J. Phys. A: Math. Gen.* **26** 3149
[15] Bös S, Kinzel W and Opper M 1993 *Phys. Rev.* E **47** 1384
[16] Bös S and Kühn R 1993 Multiplicity of metastable retrieval phases in networks of multi-state neurons *Preprint* Universität Heidelberg HD-TVP-93-19
[17] Amit D J, Gutfreund H and Sompolinky H 1987 *Ann. Phys.* **55** 1530