Thermodynamics of fully connected Blume–Emery–Griffiths neural networks

2003 J. Phys. A: Math. Gen. 36 295

(http://iopscience.iop.org/0305-4470/36/2/301)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) 295–305 PII: S0305-4470(03)54712-6

Thermodynamics of fully connected Blume–Emery–Griffiths neural networks

D Bollé and T Verbeiren

Instituut voor Theoretische Fysica, KU Leuven, B-3001 Leuven, Belgium

E-mail: desire.bolle@fys.kuleuven.ac.be and toni.verbeiren@fys.kuleuven.ac.be

Received 11 October 2002, in final form 11 November 2002
Published 17 December 2002
Online at stacks.iop.org/JPhysA/36/295

Abstract

The thermodynamic and retrieval properties of fully connected Blume–Emery–Griffiths networks are studied using replica mean-field theory. These networks can be considered as generalizations of the Hopfield model to the storage of ternary patterns. Capacity–temperature phase diagrams are derived for several values of the pattern activity. It is found that the retrieval phase is the largest in comparison with other three-state neuron models. Furthermore, the meaning and stability of the so-called quadrupolar phase is discussed as a function of both the temperature and the pattern activity. Where appropriate, the results are compared with the diluted version of the model.

PACS numbers: 05.20.−y, 64.60.Cn, 87.18.Sn

1. Introduction

It has been argued recently that an optimal Hamiltonian, guaranteeing the best retrieval properties for neural networks with multistate neurons can be found by maximizing the mutual information content of the network [1, 2]. In this way, for two-state neurons the well-known Hopfield model is recovered [3], for three-state neurons a Blume–Emery–Griffiths (BEG) type spin-glass model is obtained [4–7] (and references therein). In the discussion of the extremely diluted asymmetric version of this model [1], which has an exactly solvable dynamics because there are no feedback spin correlations [8], a new phase appeared, the so-called quadrupolar phase, that could yield new retrieval information. The time evolution of the order parameter characterizing this phase together with the stability properties of this phase in the extremely diluted model has been studied [9]. Furthermore, the zero-temperature parallel dynamics of the fully connected architecture taking into account all feedback correlations has been solved recently [10]. No quadrupolar phase has been found in this case.

A complete study of both the thermodynamic and retrieval properties of a fully connected architecture governed by such a BEG spin-glass Hamiltonian has not yet been given in the

literature. Although some preliminary results about the retrieval quality of this model for uniformly distributed patterns and nonzero temperatures have been presented in [2], one would like to have a complete temperature–capacity phase diagram as a function of the pattern activity. Furthermore, one would also like to solve the question posed in [2] about the stability of the quadrupolar phase. Finally, one would like to know in more detail how the retrieval quality of this model compares with other three-state neuron models. To fill these gaps is the purpose of this work.

The main results obtained are the following. Using a replica-symmetric mean-field theory analysis, following [11], it is shown that the retrieval phase is systematically larger than that of the three-state neuron models known in the literature, which have the same topology structure for the neurons. The critical capacity of the BEG neural network is about twice as large as that of the three-state neuron Ising model [12]. The region of thermodynamic stability of the retrieval states is much larger than that for the three-Ising model and, interestingly, even slightly larger than the corresponding region for the Hopfield model. Next, it is found that the quadrupolar phase is not a stable solution for low temperatures but can become stable at high temperatures for suitable choices of the network parameters. The physical meaning of this is discussed. Finally, by calculating the zero-temperature entropy we expect that, for uniformly distributed patterns, replica-symmetry breaking is of the same order as the breaking in the Hopfield model.

The rest of the paper is organized as follows. The model is introduced from a dynamic point of view in section 2. Section 3 presents the replica-symmetric mean-field approximation and obtains the fixed-point equations for the relevant order parameters. In section 4, these equations are studied in detail for arbitrary temperatures. In particular, a temperature–activity phase diagram for low loading and temperature–capacity phase diagrams for finite loading and several pattern activities are obtained. Also the specific thermodynamic properties are discussed. Where appropriate, the results are compared with the diluted version of the model. Concluding remarks are given in section 5. Finally, the appendix contains the explicit form of the fixed-point equations.

2. The model

Consider a network of N neurons, $\{\sigma_i\}, i = 1, \ldots, N$, which take values out of the set $\{-1, 0, 1\}$. In this network we want to store $p = \alpha N$ patterns, $\{\xi_i^\mu\}, i = 1, \ldots, N$ and $\mu = 1, \ldots, p$. They are supposed to be independent identically distributed random variables (iidrv) with respect to i and μ drawn from a probability distribution given by

$$\rho(\xi_i^\mu) = a\delta\left(1 - (\xi_i^\mu)^2\right) + (1-a)\delta(\xi_i^\mu) \tag{1}$$

with a the pattern activity, namely,

$$\lim_{N\to\infty} \frac{1}{N} \sum_i (\xi_i^\mu)^2 = a. \tag{2}$$

The neurons are updated asynchronously according to the transition probability

$$\Pr(\sigma_i' = s \in \{-1, 0, 1\} | \{\sigma_i\}) = \frac{\exp[-\beta\epsilon_i(s|\{\sigma_i\})]}{\sum_{s\in\{-1,0,1\}} \exp[-\beta\epsilon_i(s|\{\sigma_i\})]} \tag{3}$$

with β the inverse temperature and $\epsilon_i(s|\{\sigma_i\})$ an effective single site energy function given by [1]

$$\epsilon_i(s|\{\sigma_i\}) = -h_i(\{\sigma_i\})s - \theta_i(\{\sigma_i\})s^2 \qquad s \in \{-1, 0, 1\} \tag{4}$$

where the random local fields are defined by

$$h_i = \sum_{j=1}^{N} J_{ij}\sigma_j \qquad \theta_i = \sum_{j=1}^{N} K_{ij}\sigma_j^2. \tag{5}$$

The coefficients in these local fields are determined via the Hebb rule

$$J_{ij} = \frac{1}{a^2 N}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu \qquad K_{ij} = \frac{1}{a^2(1-a)^2 N}\sum_{\mu=1}^{p}\eta_i^\mu\eta_j^\mu \qquad \eta_i^\mu = \left(\left(\xi_i^\mu\right)^2 - a\right). \tag{6}$$

The rule for the couplings K_{ij} is consistent with the modified Hebb rule for the Hopfield model with biased patterns [13]. For zero temperature, the dynamical rule becomes

$$\sigma_i' = \operatorname{sign}(h_i(\{\sigma_i\}))\Theta(|h_i(\{\sigma_i\})| + \theta_i(\{\sigma_i\})). \tag{7}$$

The long-time behaviour of this network is governed by the Hamiltonian [1, 2]

$$H = -\frac{1}{2}\sum_{i\neq j} J_{ij}\sigma_i\sigma_j - \frac{1}{2}\sum_{i\neq j} K_{ij}\sigma_i^2\sigma_j^2. \tag{8}$$

Since we want to compare this model with the three-Ising model and we want to be able to change the relative importance of the two terms, we rewrite the Hamiltonian as

$$H = -\frac{A}{2}\sum_{i\neq j}\tilde{J}_{ij}\sigma_i\sigma_j - \frac{B}{2}\sum_{i\neq j}\tilde{K}_{ij}\sigma_i^2\sigma_j^2 \tag{9}$$

with

$$\tilde{J}_{ij} = aJ_{ij} \qquad \tilde{K}_{ij} = a(1-a)K_{ij}. \tag{10}$$

For

$$A = \frac{1}{a} \qquad B = \frac{1}{a(1-a)} \tag{11}$$

we trivially recover the above model. When we now take $K_{ij} = b\delta_{ij}$ and $A = B = 1$ we obtain the three-state Ising model

$$H = -\frac{1}{2}\sum_{i\neq j} J_{ij}\sigma_i\sigma_j - b\sum_i \sigma_i^2 \tag{12}$$

with b the gain parameter [12]. Finally, we find the Hopfield model back by first taking $B = 0$ and then $a = 1$, again with $A = 1$.

3. Replica-symmetric mean-field theory

We apply the standard replica technique [11, 14] in order to calculate the free energy of the model. Within the replica-symmetry approximation and for a finite number, s, of condensed patterns, we obtain

$$f(\beta) = \frac{1}{2}\sum_{\nu=1}^{s}\left(aAm_\nu^2 + a(1-a)Bl_\nu^2\right) + \frac{\alpha}{2\beta}\log(1-\chi) + \frac{\alpha}{2\beta}\log(1-\phi) + \frac{\alpha}{2\beta}\frac{\chi}{1-\chi} + \frac{\alpha}{2\beta}\frac{\phi}{1-\phi} + \frac{\alpha}{2}\frac{Aq_1\chi}{(1-\chi)^2} + \frac{\alpha}{2}\frac{Bp_1\phi}{(1-\phi)^2} - \frac{1}{\beta}\left\langle\int \mathrm{D}s\,\mathrm{D}t\,\ln \mathrm{Tr}_\sigma \exp(\beta\tilde{H})\right\rangle_{\{\xi^\nu\}} \tag{13}$$

with the effective Hamiltonian $\tilde{H}$ given by

$$\tilde{H} = A\sigma\left[\sum_{\nu} m_\nu \xi^\nu + \sqrt{\alpha r}s\right] + B\sigma^2\left[\sum_{\nu} l_\nu \eta^\nu + \sqrt{\alpha u}t\right] + \frac{\alpha}{2}\frac{A\chi}{1-\chi}\sigma^2 + \frac{\alpha}{2}\frac{B\phi}{1-\phi}\sigma^2 \tag{14}$$

and where $\mathrm{D}s$ and $\mathrm{D}t$ are Gaussian measures, $\mathrm{D}s = \mathrm{d}s(2\pi)^{-1/2}\exp(-s^2/2)$. Furthermore

$$\chi = A\beta(q_0 - q_1) \qquad \phi = B\beta(p_0 - p_1) \qquad r = \frac{q_1}{(1-\chi)^2} \qquad u = \frac{p_1}{(1-\phi)^2}. \tag{15}$$

In these expressions the relevant order parameters are

$$m_\nu = \frac{1}{a}\left\langle \xi^\nu \int \mathrm{D}s\,\mathrm{D}t\,\langle\sigma\rangle_\beta \right\rangle_{\{\xi^\nu\}} \tag{16}$$

$$l_\nu = \frac{1}{a(1-a)}\left\langle \eta^\nu \int \mathrm{D}s\,\mathrm{D}t\,\langle\sigma^2\rangle_\beta \right\rangle_{\{\xi^\nu\}} \tag{17}$$

$$q_0 = p_0 = \left\langle \int \mathrm{D}s\,\mathrm{D}t\,\langle\sigma^2\rangle_\beta \right\rangle_{\{\xi^\nu\}} \tag{18}$$

$$q_1 = \left\langle \int \mathrm{D}s\,\mathrm{D}t\,\langle\sigma\rangle_\beta^2 \right\rangle_{\{\xi^\nu\}} \tag{19}$$

$$p_1 = \left\langle \int \mathrm{D}s\,\mathrm{D}t\,\langle\sigma^2\rangle_\beta^2 \right\rangle_{\{\xi^\nu\}} \tag{20}$$

where $\langle\cdot\rangle_\beta$ represents the thermal average with respect to the effective Hamiltonian $\tilde{H}$. In the following, we take only one condensed pattern such that the index ν can be dropped.

The parameter m is the usual overlap between the condensed pattern and the network state, while l is related to the activity overlap, i.e., $\langle\sigma^2\xi^2\rangle$. In [9] it has been called the fluctuation overlap between the binary state variables σ_i^2 and η_i^2. Furthermore, q_0 is the activity of the neurons and q_1 and p_1 are the Edwards–Anderson order parameters with their conjugate variables r and u, respectively. Finally, χ and ϕ are the susceptibilities proportional to the fluctuation of the m and l overlaps, respectively.

We remark that these equations have been obtained in [2]. The trace over the neurons and the average over the patterns can be performed explicitly. The resulting expressions are given in the appendix.

4. Thermodynamic and retrieval properties

In this section, we study the thermodynamic and retrieval properties of the fully connected BEG network by numerically solving the fixed-point equations (16)–(20) for one condensed pattern. Depending on the temperature T and the system parameters α and a we recognize the following phases in terms of the order parameters.

There is a retrieval phase R ($m > 0,\ l > 0,\ q_1 > 0$) characterized by positive m and l, and a quadrupolar phase Q ($m = 0,\ l > 0,\ q_1 = 0$) where only l is nonzero, meaning that the active neurons (± 1) coincide with the active patterns but the signs are not correlated. This implies that this phase also carries some retrieval information. From the fact that $q_1 = 0$ in this phase we know that the spins are not frozen, so that we expect to find this phase only for high enough temperatures. Furthermore, there is the spin-glass phase SG ($m = 0,\ l = 0,\ q_1 > 0$) and the paramagnetic phase P ($m = 0,\ l = 0,\ q_1 = 0$). We first look at low loading $\alpha = 0$.

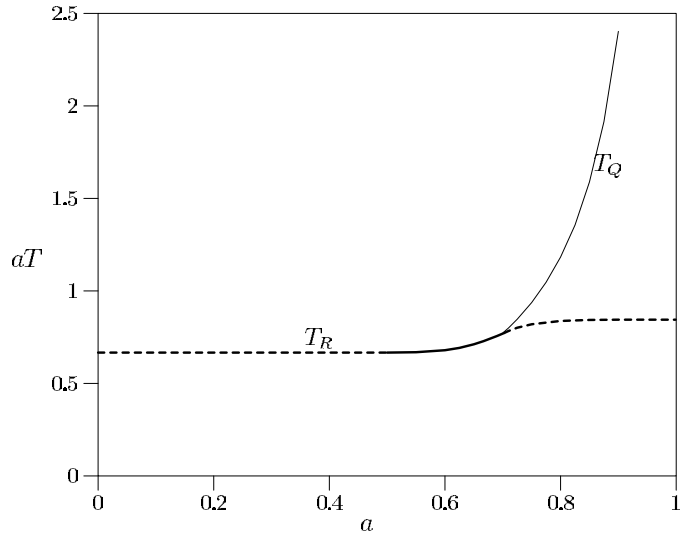

Figure 1. The BEG phase diagram for $\alpha = 0 : aT$ as a function of a. Dashed (full) line denotes continuous (discontinuous) transition.

4.1. Low loading

For $\alpha = 0$, the fixed-point equations simplify to a great extent because all the integrations (see the appendix) drop out. This allows us to carefully study the quadrupolar state as a function of the pattern activity a since it turns out that the effect of the quadrupolar phase is strongest for a small loading capacity and high temperatures. A temperature–activity phase is presented in figure 1. A dashed (full) line corresponds to a continuous (discontinuous) transition.

Below $a = 1/2$, there is a continuous transition at $aT_R = 2/3$ from the retrieval phase at low T to the paramagnetic phase at high T. This is similar to the three-Ising model where it occurs for all values of a [12]. At $a = 1/2$ the transition becomes discontinuous and, up to $a = 0.698$, the only phases present are R and P. The quadrupolar phase starts to appear at $a = 0.698$ and $aT = 0.767$, and beyond that point it keeps growing for increasing a. The transition R–Q remains discontinuous up to $a = 0.708$ and $aT = 0.78$. For larger values of a, it remains continuous and ends at $aT = (1 + (2\mathrm{e})^{-1})^{-1} = 0.84$ for $a = 1$. We remark that this phase diagram is the same as the corresponding one for the extremely diluted model [9], confirming the fact that for low loading the architecture is not important.

4.2. Finite loading

Next, in figures 2–4 we present the temperature–capacity phase diagrams for three typical values of the pattern activity, $a = 1/2$, $a = 2/3$ (uniformly distributed patterns) and $a = 0.8$, respectively. As before, a dashed line corresponds to a continuous transition, while a full line corresponds to a discontinuous transition (in all order parameters).

We start with some general observations. Below the line T_R retrieval states occur. The curve T_F represents the thermodynamic transition between retrieval states and spin-glass states. Hence, below T_F the retrieval states are global minima of the free energy while above this line the spin-glass states are. Thermodynamic transitions are shown as thick lines. The line T_{SG} denotes the transition from the spin-glass to the paramagnetic phase.

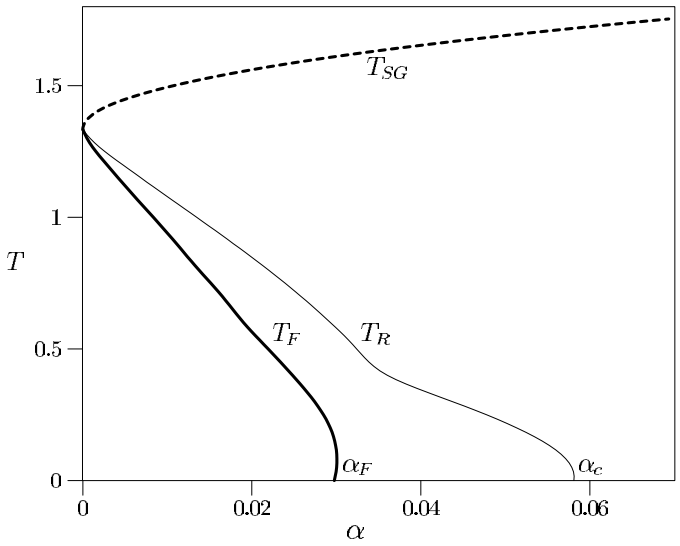

Figure 2. The BEG α–T phase diagram for $a = 1/2$. The meaning of the lines is explained in the text.

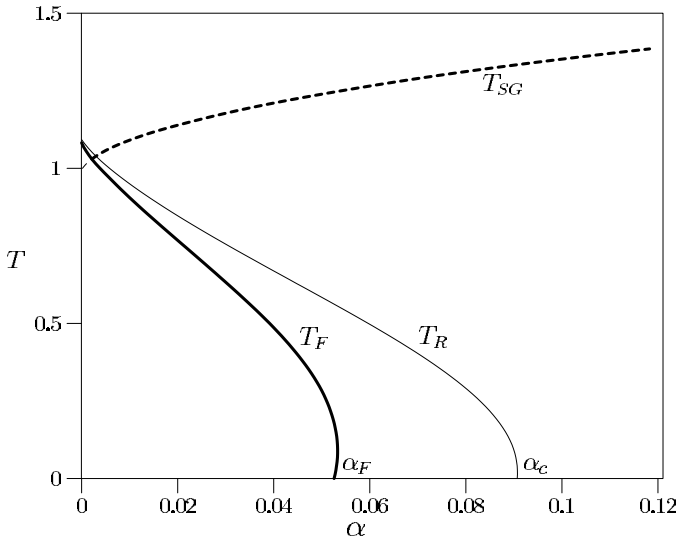

Figure 3. The BEG α–T phase diagram for $a = 2/3$. The meaning of the lines is explained in the text.

In all phase diagrams we note some reentrance, i.e., α_c at some finite T is larger than α_c at $T = 0$. This effect is well known from the literature (see, e.g., [15, 16]) and signals the breaking of replica symmetry. In order to have an idea about the size of this breaking compared with other models, we calculate the entropy at zero temperature [11]:

$$S = -\frac{\alpha}{2}\left[\ln(1-\chi) + \ln(1-\phi) + \frac{\chi}{1-\chi} + \frac{\phi}{1-\phi}\right]. \tag{21}$$

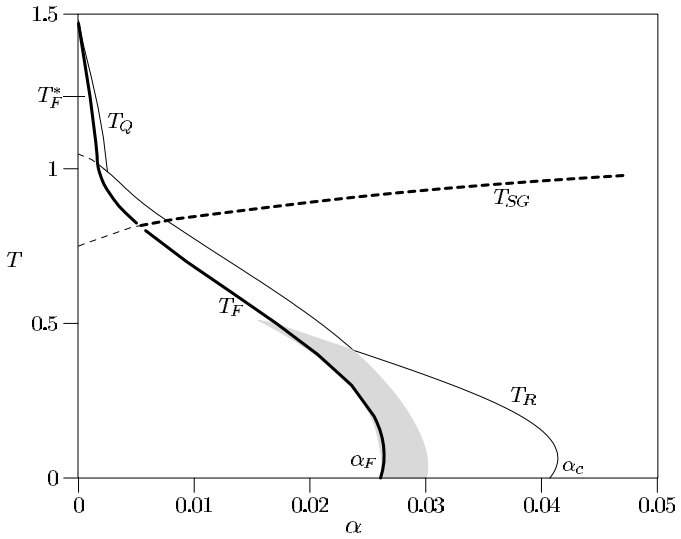

Figure 4. The α–T BEG phase diagram for $a = 0.8$. The meaning of the lines is explained in the text.

We find that this entropy is indeed negative but small. For uniform patterns, e.g., the entropy of the retrieval state at α_c is $S(\alpha_c) = -0.0017$, which is of the same order of magnitude as that for the Hopfield model, i.e., $S(\alpha_c) = -0.0014$, suggesting that the breaking is comparable. Moreover, the zero-temperature entropy becomes more negative for increasing a, e.g., $S(\alpha_c)$ is -0.0010, -0.0017 and -0.031 for $a = 1/2$, $a = 2/3$ and $a = 0.8$, respectively, which might suggest that the breaking becomes larger.

Next, we look at the different phase diagrams in more detail. In the case of $a = 1/2$ shown in figure 2, the diagram quantitatively resembles that for the three-Ising model [12]. At high temperatures there is the continuous transition from the disordered paramagnetic phase to the spin-glass phase. When crossing the curve T_R retrieval states show up as local minima of the free energy. At these points the overlap with the embedded patterns jumps from zero to a finite macroscopic value. In comparison with the three-Ising model there is a small kink in the line T_R. When lowering T further the retrieval states become global minima of the free energy. This happens along the curve T_F and this thermodynamic transition is first order. The value for the critical capacity at zero temperature is $\alpha_c = 0.058$, while the thermodynamic transition point is $\alpha_F = 0.030$.

For uniform patterns (see figure 3) several transition curves bordering the different phases discussed above show up. Here we note that the critical curves T_{SG} and T_R end in different temperature points at $\alpha = 0$ giving rise to a 'crossover' region for small α as it occurs in the Potts model [17]. This is related to the fact that for $\alpha = 0$ this model has a discontinuous transition at T_F. In this crossover region the retrieval states (global minima below T_F) and the paramagnetic states (local minima below T_F) coexist. We remark that a preliminary version of this figure has been discussed in [2].

Comparing these results with those found for the three-Ising model [12], we see that the $\alpha_c = 0.091$ found here is almost double the critical capacity of the latter, $\alpha_c = 0.046$. We recall that the latter number is obtained in the case of an optimal choice for the gain parameter b (see (12)), i.e. $b = 1/2$, which makes the Hamming distance minimal [12]. Furthermore, the region of thermodynamic stability for the retrieval states is about four times larger. Compared

with the Hopfield model, we note that α_c is smaller in the BEG model, 0.091 versus 0.13, but α_F is larger, 0.053 versus 0.051. So a large number of the retrieval states in the BEG network are global minima of the free energy.

Figure 4 shows the phase diagram for $a = 0.8$. We immediately remark that the structure of the phase diagram turns out to be more complicated, as expected. The major difference with the foregoing phase diagrams is the presence of the quadrupolar phase for high temperatures. Indeed, from the discussion in section 4.1 (see figure 1) we know that this quadrupolar phase shows up from an activity $a = 0.698$ onwards. As a consequence, the crossover or coexistence region is larger than that for the $a = 2/3$ phase diagram. The transition from the Q phase to the P phase is discontinuous and by comparing the relevant free energies of the different phases we find the thermodynamic transition line T_F^* below which the quadrupolar states are global minima. We see that in figure 4 T_F^* joins nicely with T_F, as it should. The transition from the R to the Q phase is continuous for small α up to $\alpha = 0.0023$ and discontinuous beyond that value. Finally, we note that there is a larger reentrance suggesting a stronger replica-symmetry breaking, which is consistent with the fact that the zero-temperature entropy is more negative in this case, as mentioned above.

The quadrupolar phase is situated in the high-temperature region and we can understand the physics behind it in the following way. The spin-glass order parameter q_1 is zero, meaning that the ± 1 spins are not frozen and as a consequence m can be zero. The fact that l is not zero practically means that the spins can flip freely between ± 1 but the probability that they jump to 0 or vice versa becomes very small. This effect arises from $a > 1/2$ onwards when the ratio between the second and the first terms in the Hamiltonian starts increasing as $(1 - a)^{-1}$. It implies that the information content of the system is nonzero in this phase. A practical example may be in pattern recognition where, looking at black and white pictures on a grey background, this phase would tell us the exact location of the picture with respect to the background without finding the details of the picture itself.

Returning to figure 4 we remark that in the shaded region two retrieval states coexist. A similar behaviour has also been seen in other multistate networks, e.g., the fully connected Potts model [17]. Their presence can be understood by studying the overlap order parameter m as a function of α. This is shown in figure 5 where the full line denotes stable solutions, while the dashed line corresponds to unstable solutions (saddle points). For small T two stable solutions occur. The one with the smallest overlap vanishes for large T.

We end this section with a technical remark. In studying the fixed-point equations of this model, it turns out that a large number of solutions are in fact saddle points and not minima of the free energy. This can, of course, be investigated by studying the local stability of the extrema of the free energy. For large a we find, e.g., that there may be more than four possible solutions involving some kind of quadrupolar character ($m = 0, l > 0$). Only the solution depicted in figure 4 is a stable one (a real attractor). This also answers the question posed in [2] about the stability of the quadrupolar state at low temperatures.

4.3. Coefficients A and B versus the critical capacity

In defining the BEG neural network model, we have included the possibility of varying the coefficients A and B in the Hamiltonian. The values of the coefficients A and B given by (11) stem from [1, 2]. They are obtained by optimizing the mutual information of the system. One could expect that this choice also improves other properties of the neural network, e.g., the basin of attraction, the critical capacity α_c.

Figure 6 shows α_c (upper line) and α_F (lower line) as a function of $\gamma = a(1 - a)B$ with A fixed at $A = 1/a$ for uniform patterns at $T = 0$. For $\gamma = 1$, we recover the BEG neural

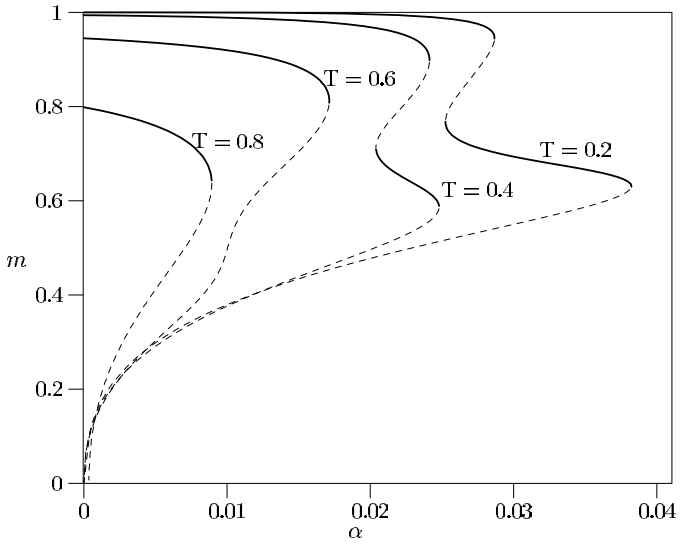

Figure 5. The overlap m as a function of α for $T = 0.2, 0.4, 0.6$ and 0.8. The full (dashed) line denotes the stable (unstable) solution.

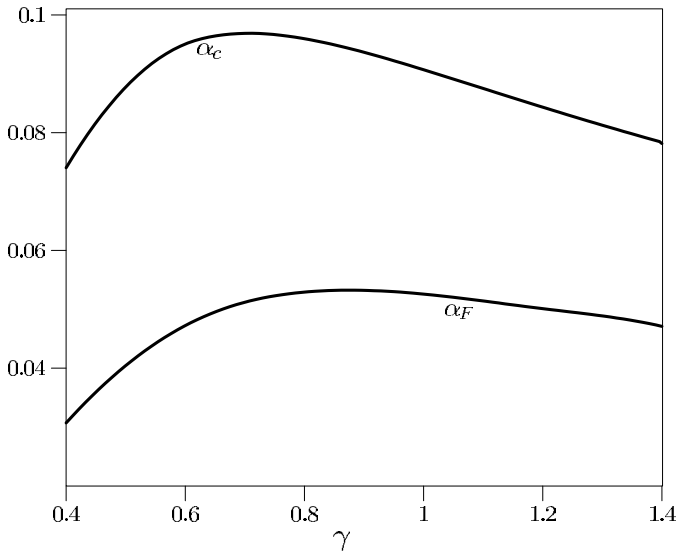

Figure 6. The capacities α_c and α_F as a function of $\gamma = a(1-a)B$ for uniform patterns at $T = 0$.

network as in (11) and studied above. It turns out that the maximum in the critical capacity is located at $\gamma = 0.712$ with a corresponding value of $\alpha_c = 0.096$. Also the maximum in the thermodynamic transition is located at a value smaller than 1. This does not agree with the expectation formulated above.

A reason for this is the approximation done in order to get the mean-field Hamiltonian in [1, 2]. At a certain point in its derivation the authors assume that $q_0 \sim a$. Consequently, the mutual information of the network is optimized under this assumption. Although this assumption is natural for having a complete match between the final state of the network and

the condensed pattern, in general, it may not be realized in a specific model. Furthermore, the fact that replica-symmetry breaking may be larger for larger α, as is also suggested by the zero-temperature entropy calculation, could be an extra reason why the maximum in α_c is further away from the value $\gamma = 1$ than is the maximum in α_F. A further investigation is non-trivial and beyond the scope of this work.

5. Concluding remarks

We have considered both the thermodynamic and retrieval properties of fully connected BEG networks. Fixed-point equations for the relevant order parameters have been derived for arbitrary temperatures in the replica-symmetric mean-field approximation. An activity–temperature phase diagram for low loading has been obtained. Near saturation, capacity–temperature phase diagrams have been discussed in detail for several values of the activity of the three-state patterns.

Compared with the existing three-state neuron models the retrieval region is larger and, e.g., for uniformly distributed patterns the critical capacity at zero temperature is almost twice as large. Also the region of thermodynamic stability of the retrieval states is much enlarged and even larger than that for the two-state Hopfield model. A Gardner type analysis of the corresponding perceptron model [18] confirms the better retrieval properties. A new information carrying phase, the quadrupolar phase, appears at larger values of the activity in the high-temperature region of the phase diagram and may extend the practical usefulness of this network, e.g., in pattern recognition.

Acknowledgments

We would like to thank J Busquets-Blanco, R Erichsen Jr, I Pérez-Castillo and W Theumann for fruitful discussions. One of the authors (DB) thanks W Theumann for hospitality and financial support during a stay at the Instituto de Física of the Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil. This work has been supported in part by the Fund of Scientific Research, Flanders-Belgium.

Appendix. Explicit expressions for the fixed-point equations

In this section, we write down explicitly the fixed-point equations for the order parameters of the BEG network. After performing the trace over the spins and the average over the condensed patterns in equations (16)–(20) we obtain

$$m = \int \mathrm{D}s\,\mathrm{D}t\,\mathbf{V}_\beta\left[(Am + A\sqrt{\alpha r}s), \left(Bl(1-a) + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right] \tag{A.1}$$

$$\begin{aligned} l = &\int \mathrm{D}s\,\mathrm{D}t\,\mathbf{W}_\beta\left[(Am + A\sqrt{\alpha r}s), \left(Bl(1-a) + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right] \\ &- \int \mathrm{D}s\,\mathrm{D}t\,\mathbf{W}_\beta\left[(A\sqrt{\alpha r}s), \left(-Bla + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right] \end{aligned} \tag{A.2}$$

$$\begin{aligned} q_0 = p_0 = a&\int \mathrm{D}s\,\mathrm{D}t\,\mathbf{W}_\beta\left[(Am + A\sqrt{\alpha r}s), \left(Bl(1-a) + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right] \\ &+ (1-a)\int \mathrm{D}s\,\mathrm{D}t\,\mathbf{W}_\beta\left[(A\sqrt{\alpha r}s), \left(-Bla + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right] \end{aligned} \tag{A.3}$$

$$q_1 = a \int \mathrm{D}s\,\mathrm{D}t \left(\mathbf{V}_\beta \left[(Am + A\sqrt{\alpha r}s), \left(Bl(1-a) + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right]\right)^2 + (1-a) \int \mathrm{D}s\,\mathrm{D}t \left(\mathbf{V}_\beta \left[(A\sqrt{\alpha r}s), \left(-Bla + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right]\right)^2 \tag{A.4}$$

$$p_1 = a \int \mathrm{D}s\,\mathrm{D}t \left(\mathbf{W}_\beta \left[(Am + A\sqrt{\alpha r}s), \left(Bl(1-a) + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right]\right)^2 + (1-a) \int \mathrm{D}s\,\mathrm{D}t \left(\mathbf{W}_\beta \left[(A\sqrt{\alpha r}s), \left(-Bla + B\sqrt{\alpha u}t + \frac{\alpha}{2}X\right)\right]\right)^2 \tag{A.5}$$

with

$$\chi = A\beta(q_0 - q_1) \qquad \phi = B\beta(p_0 - p_1) \qquad r = \frac{q_1}{(1-\chi)^2} \qquad u = \frac{p_1}{(1-\phi)^2} \tag{A.6}$$

and

$$X = A\frac{\chi}{1-\chi} + B\frac{\phi}{1-\phi}. \tag{A.7}$$

The functions $\mathbf{V}_\beta$ and $\mathbf{W}_\beta$ are defined by

$$\mathbf{V}_\beta(x, y) = \frac{\sinh(\beta x)}{\frac{1}{2}\exp(-\beta y) + \cosh(\beta x)} \tag{A.8}$$

$$\mathbf{W}_\beta(x, y) = \frac{\cosh(\beta x)}{\frac{1}{2}\exp(-\beta y) + \cosh(\beta x)} \tag{A.9}$$

and reduce, for zero temperature, to

$$\mathbf{V}_\infty(x, y) = \mathrm{sign}(x)\Theta(|x| + y) \tag{A.10}$$

$$\mathbf{W}_\infty(x, y) = \Theta(|x| + y). \tag{A.11}$$

References

[1] Carreta Dominguez D and Korutcheva E 2000 *Phys. Rev.* E **62** 2620
[2] Bollé D and Verbeiren T 2002 *Phys. Lett.* A **297** 156
[3] Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554
[4] Blume M, Emery V J and Griffiths R B 1971 *Phys. Rev.* A **4** 1071
[5] Blume M 1966 *Phys. Rev.* **141** 517
[6] Capel H W 1966 *Physica* **32** 966
[7] de Araújo J M, da Costa F A and Nobre F D 2000 *Eur. Phys. J.* B **14** 661
[8] Barkai E, Kanter I and Sompolinsky H 1990 *Phys. Rev.* A **41** 590
[9] Bollé D, Dominguez D R C, Erichsen E Jr, Korutcheva E and Theumann W K 2002 *Preprint* cond-mat/0208281
[10] Bollé D, Busquets-Blanco J and Shim G M 2002 *Preprint* cond-mat/0206569 (2002 *Physica* A at Press)
[11] Amit D J, Gutfreund H G and Sompolinsky H 1987 *Ann. Phys., NY* **173** 30
[12] Bollé D, Rieger H and Shim G M 1994 *J. Phys. A: Math. Gen.* **27** 3411
[13] Amit D J, Gutfreund H G and Sompolinsky H 1987 *Phys. Rev.* A **35** 2293
[14] Mézard M, Parisi G and Virasoro M A 1978 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[15] Watkin T L H and Sherrington D 1991 *Europhys. Lett.* **14** 791
[16] Naef J P and Canning A 1992 *J. Physique* I **2** 247
[17] Bollé D, Dupont P and Huyghebaert J 1992 *Phys. Rev.* A **45** 4194
[18] Bollé D, Pérez Castillo I and Shim G M 2002 *Preprint* cond-mat/0210256