The Blume-Emery-Griffiths neural network with synchronous updating and variable dilution

D. Bollé[a] and J. Busquets Blanco[b]

Instituut voor Theoretische Fysica, Katholieke Universiteit Leuven, Celestijnenlaan 200 D, 3001 Leuven, Belgium

Abstract. The thermodynamic and retrieval properties of the Blume-Emery-Griffiths neural network with synchronous updating and variable dilution are studied using replica mean-field theory. Several forms of dilution are allowed by pruning the different types of couplings present in the Hamiltonian. The appearance and properties of two-cycles are discussed. Capacity-temperature phase diagrams are derived for several values of the pattern activity. The results are compared with those for sequential updating. The effect of self-coupling is studied. Furthermore, the optimal combination of dilution parameters giving the largest critical capacity is obtained.

PACS. 05.20.-y Classical statistical mechanics – 64.60.Cn Order-disorder transformations; statistical mechanics of model systems – 87.18.Sn Neural networks

1 Introduction

Recently, we have studied the possible different physics arising from sequential and synchronous updating of the spins in multi-state Ising-type ferromagnets [1]. In general, it is known that besides fixed-point attractors, synchronous updating in spin-models can give rise to cycles as stationary behaviour [2,3]. In [1] we have shown that for the $Q = 3$ Ising ferromagnet with synchronous updating the same stationary solutions appear as for sequential updating except for negative couplings. In that case, for synchronous updating, cycles of period two in the magnetization occur in the ferromagnetic phase. For the Blume-Emery-Griffiths (BEG) ferromagnet, however, we have found that synchronous updating allows for a much richer free energy landscape and phase diagram.

In view of these results it is of interest to extend such a study to multi-state neural networks in order to see whether the different types of updating give rise to different retrieval properties. For the Hopfield neural network model, e.g., involving binary neurons, it is known that in a replica symmetric treatment the retrieval region is slightly larger in the synchronous case [4]. A natural question is then whether a similar effect is present in the BEG neural network, which is known to optimize the mutual information content for three-state networks [5,6] and, hence, to guarantee the best retrieval properties in comparison with other three-state networks (see [7] for a review). For sequential updating, the thermodynamic and retrieval prop-

erties are studied in detail in the literature for fully connected architectures [8] and diluted architectures [9,10] (and references therein), for synchronous updating such a study has not yet appeared.

The purpose of this work is precisely to report on a study of the BEG neural network with synchronous updating and variable dilution. In particular, we discuss the thermodynamic and retrieval properties of this model using replica symmetric mean-field theory. Several forms of dilution are allowed by cutting some of the couplings appearing in the Hamiltonian: the couplings multiplying the term bilinear in the neurons and/or the couplings related to the term biquadratic in the neurons. Capacity-temperature phase diagrams are derived for several values of the pattern activity. In particular, the appearance and properties of cycles are discussed. Apart from these cycles it turns out that the asymptotic behaviour is almost identical to the one for sequential updating. For some values of the (fully-connected) network parameters the retrieval region is marginally enhanced. Only the addition of self-coupling, not allowed in the sequential model if one wants to satisfy detailed balance [3,11], can enlarge the retrieval region substantially, especially in the case of dilution. The optimal combination of the network dilution parameters giving the largest critical capacity is obtained.

The rest of the paper is organized as follows. In Section 2 the model with variable dilution is introduced from a dynamic point of view. Section 3 presents the replica-symmetric mean-field calculation of the free energy and obtains the fixed-point equations for the relevant order parameters. The possible appearance and properties of

[a] e-mail: desire.bolle@fys.kuleuven.be
[b] e-mail: Jordi.Busquets@fys.kuleuven.be

two-cycles are examined. In Section 4 these equations are studied in detail for arbitrary temperatures and different types of dilution. In particular, capacity-temperature phase diagrams are obtained and compared with the corresponding results for sequential updating. The optimal dilution parameters leading to the largest critical capacity are obtained. Also the specific thermodynamic properties are discussed. Furthermore, the effect of self-coupling on the retrieval properties and on the phase-diagrams is analyzed. Some concluding remarks are given in Section 5.

2 The model

Consider a network of N neurons, $\boldsymbol{\sigma} = \{\sigma_1, ..., \sigma_N\}$, which can take values from the set $\mathcal{S} = \{-1, 0, 1\}$. 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

$$\Pr(\xi_i^\mu) = a\,\delta(1 - (\xi_i^\mu)^2) + (1 - a)\,\delta(\xi_i^\mu), \qquad (1)$$

with a the pattern activity, namely

$$\mathrm{E}((\xi_i^\mu)^2) = a. \qquad (2)$$

The neurons are updated synchronously according to the transition probability

$$\Pr(\boldsymbol{\sigma}|\boldsymbol{\sigma}') = \prod_{i=1}^N \Pr(\sigma_i = s \in \mathcal{S}|\boldsymbol{\sigma}') \qquad (3)$$

$$\Pr(\sigma_i = s \in \mathcal{S}|\boldsymbol{\sigma}') = \frac{\exp[-\beta\epsilon_i(s|\boldsymbol{\sigma}')]}{\sum\limits_{s \in \mathcal{S}} \exp[-\beta\epsilon_i(s|\boldsymbol{\sigma}')]} \qquad (4)$$

with β the inverse temperature and $\epsilon_i(s|\boldsymbol{\sigma})$ an effective single site energy function given by

$$\epsilon_i(s|\boldsymbol{\sigma}) = -h_{1,i}(\boldsymbol{\sigma})s - h_{2,i}(\boldsymbol{\sigma})s^2, \quad s \in \mathcal{S}, \qquad (5)$$

where the random local fields are defined by

$$h_{k,i}(\boldsymbol{\sigma}) = \sum_{j=1}^N J_{k,ij}^c \sigma_j^k, \quad k = 1, 2. \qquad (6)$$

In the sequel the index k will always take the values $k = 1, 2$. The couplings $J_{k,ij}^c$ are taken to be of the form

$$J_{k,ij}^c = \frac{c_{k,ij}}{c_k} J_{k,ij} \qquad (7)$$

where the probability distribution of the $\{c_{k,ij}\}$ is given by

$$\Pr(c_{k,ij}) = c_k \delta(c_{k,ij} - 1) + (1 - c_k)\delta(c_{k,ij}). \qquad (8)$$

Hence, they allow for a diluted architecture. The $J_{k,ij}$ are determined via the Hebb rule

$$J_{k,ij} = \frac{1}{a_k^2 N} \sum_{\mu=1}^p \eta_{k,i}^\mu \eta_{k,j}^\mu \qquad (9)$$

where $\eta_{1,i}^\mu = \xi_i^\mu$ and $\eta_{2,i}^\mu = ((\xi_i^\mu)^2 - a)$. For convenience we define $a_1 \equiv \mathrm{E}((\eta_{1,i}^\mu)^2) = a$, $a_2 \equiv \mathrm{E}((\eta_{2,i}^\mu)^2) = a(1 - a)$.

As can be seen from (7)–(9), two types of dilution are possible. Some of the couplings related to the bilinear term in the neurons, $k = 1$, can be cut and/or some of the couplings related to the biquadratic term, $k = 2$, can be pruned, both simultaneously or independently. In the following we denote this as k-type dilution. We remark that in the literature $k = 2$-type dilution is frequently indicated as dilution of the active patterns, since only the fact whether the neuron is active or not (and not the sign of the neuron state) is important. In order for the model to satisfy detailed balance, the dilution has to be symmetric ($c_{k,ij} = c_{k,ji}$). In the case of extreme dilution, when $c_k = 0$, the average number of connections per neuron, $c_k N$, is still infinite

$$\lim_{N \to \infty} \frac{1}{c_k N} = 0, \, \forall c_k. \qquad (10)$$

Finally, we recall that the detailed balance property for synchronous updating is not destroyed by the presence of self-couplings, i.e., couplings of the form $J_{k,ii}^c$, $k = 1, 2$. Hence, we do allow for this type of couplings and write them as $J_{k,ii}^c = J_{k,0}\alpha_k/a_k$, with $J_{k,0}$ a parameter and $\alpha_k = p/c_k N$ the k-type capacities.

The long-time behaviour of this model is governed by the pseudo-Hamiltonian [1]

$$H(\boldsymbol{\sigma}) = -\frac{1}{\beta} \sum_{i=1}^N \ln\left\{2\exp\left(\beta h_{2,i}(\boldsymbol{\sigma})\right)\cosh\left(\beta h_{1,i}(\boldsymbol{\sigma})\right) + 1\right\} \qquad (11)$$

with as duplicate-neuron (spin) representation

$$H(\boldsymbol{\sigma}, \boldsymbol{\tau}) = -\sum_k \sum_{i,j} J_{k,ij}^c \sigma_i^k \tau_j^k \qquad (12)$$

such that $\operatorname*{Tr}_{\boldsymbol{\sigma}}[\exp(-\beta H(\boldsymbol{\sigma}))] = \operatorname*{Tr}_{\boldsymbol{\sigma}}\operatorname*{Tr}_{\boldsymbol{\tau}}[\exp(-\beta H(\boldsymbol{\sigma}, \boldsymbol{\tau}))]$.

In the next sections we study the thermodynamic and retrieval properties of this model starting from the free energy.

3 Replica mean-field theory

3.1 Replicated free energy

We adapt the standard replica technique as applied to dilute models [12]–[16] to synchronous updating in order to calculate the free energy of the model. Since the method is rather standard by now, although the specific details are rather involved, we only indicate the main steps. For the completely detailed calculation we refer to [17].

We start from the replicated partition function based upon the two-spin Hamiltonian (12)

$$\langle \mathcal{Z}^n \rangle_{c,\xi} = \prod_\rho \left(\operatorname*{Tr}_{\boldsymbol{\sigma}_\rho}\operatorname*{Tr}_{\boldsymbol{\tau}_\rho}\right) \prod_{k,i} \exp\left(\beta\frac{\alpha_k}{a_k}J_{k,0}\sum_\rho \sigma_{i,\rho}^k \tau_{i,\rho}^k\right)$$
$$\left\langle \prod_k \prod_{i,j \neq i} \exp\left(\beta J_{k,ij}^c \sum_\rho \sigma_{i,\rho}^k \tau_{i,\rho}^k\right)\right\rangle_{c,\xi} \qquad (13)$$

where the average is over the patterns and the dilution and where we have explicitly separated off the self-interactions.

We first average over the dilution by expanding the exponential with respect to $J_{k,ij}^c$ up to order $\mathcal{O}((c_k N)^{-3/2})$. After using that the couplings $J_{k,ij} \sim \mathcal{O}((\alpha_k c_k/a_k^2 N)^{1/2})$ and assuming a macroscopic overlap with only one pattern we arrive at

$$\langle \mathcal{Z}^n \rangle_{c,\xi} = \prod_\rho \left(\underset{\boldsymbol{\sigma}_\rho \boldsymbol{\tau}_\rho}{\mathrm{Tr}\,\mathrm{Tr}} \right) \mathcal{Z}^c \langle \mathcal{Z}^{nc} \rangle_\xi$$

$$\times \prod_k \left\{ \exp \left(\frac{\beta \alpha_k}{a_k}(J_{k,0} - c_k) \sum_{i,\rho} \sigma_{i,\rho}^k \tau_{i,\rho}^k \right) \right.$$

$$\left. \times \exp \left(\frac{\beta^2 \alpha_k (1 - c_k)}{4 a_k^2 N} \sum_{i,j \neq i} \left[\sum_\rho \left(\sigma_{i,\rho}^k \tau_{j,\rho}^k + \sigma_{j,\rho}^k \tau_{i,\rho}^k \right) \right]^2 \right) \right\}$$

(14)

where $\mathcal{Z}^c$ denotes the condensed part (e.g., pattern $\mu = 1$) and $\mathcal{Z}^{nc}$ the non-condensed part (patterns $\mu = 2, ..., p$)

$$\mathcal{Z}^c \langle \mathcal{Z}^{nc} \rangle_\xi = \left\langle \prod_k \left(\prod_{i,j} \exp \left(\beta \sum_\rho J_{k,ij} \sigma_{i,\rho}^k \tau_{j,\rho}^k \right) \right) \right\rangle_\xi .$$

(15)

Next, we consider the average over the disorder. A Hubbard-Stratonovich transformation allows us to write the condensed part, introducing $\sqrt{N\beta} m_{kk',\rho}^1$ with $k, k' = 1, 2$ (we forget about the upper-index 1 in the sequel), as

$$\mathcal{Z}^c \propto \int \prod_{k,k'} \left(d\mathbf{m}_{kk'} \exp \left(-\frac{\beta N}{2} \sum_\rho m_{kk',\rho}^2 \right) \right)$$

$$\times \prod_k \exp \left(\frac{\beta}{a_k} \sum_\rho \sum_i \left[\eta_{k,i}(\sigma_{i,\rho}^k + \tau_{i,\rho}^k) m_{k1,\rho} \right. \right.$$

$$\left. \left. + i\eta_{k,i}(\sigma_{i,\rho}^k - \tau_{i,\rho}^k) m_{k2,\rho} \right] \right)$$

(16)

with $d\mathbf{m}_{kk'} = \prod_\rho m_{kk',\rho}$. For the non-condensed part we first do a Hubbard-Stratonovich transformation introducing the variables $m_{kk',\rho}^\mu$ for $\mu > 1$, expand then the exponential up to order $\mathcal{O}(N^{-3/2})$, and finally use $\cosh(x) \simeq \exp(x^2/2)$ to obtain

$$\langle \mathcal{Z}^{nc} \rangle_\xi^{1/(1-p)} \propto \int d\boldsymbol{\Phi} \exp \left(-\frac{1}{2}\boldsymbol{\Phi}^2 + \frac{\beta}{2}\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\Phi}^\dagger \right),$$

(17)

where $\boldsymbol{\Phi}$ is a $4n$-dimensional vector of the form $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, ..., \boldsymbol{\Phi}_n)$ with $\boldsymbol{\Phi}_\rho = (m_{11,\rho}, m_{12,\rho}, m_{21,\rho}, m_{22,\rho})$. In the same way, $\boldsymbol{\Psi}$ is a $4n \times 4n$ matrix of the form

$$\boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Psi}_{11} & i\boldsymbol{\Psi}_{12} & 0 & 0 \\ i\boldsymbol{\Psi}_{13} & -\boldsymbol{\Psi}_{14} & 0 & 0 \\ 0 & 0 & \boldsymbol{\Psi}_{21} & i\boldsymbol{\Psi}_{22} \\ 0 & 0 & i\boldsymbol{\Psi}_{23} & -\boldsymbol{\Psi}_{24} \end{pmatrix} .$$

Each element in itself is an $n \times n$ matrix with elements defined in terms of the neuron variables as

$$\psi_{k1,\rho\gamma} = \frac{1}{2Na_k} \sum_{i=1}^N (\sigma_{i,\rho}^k + \tau_{i,\rho}^k)(\sigma_{i,\gamma}^k + \tau_{i,\gamma}^k),$$

$$\psi_{k2,\rho\gamma} = \frac{1}{2Na_k} \sum_{i=1}^N (\sigma_{i,\rho}^k + \tau_{i,\rho}^k)(\sigma_{i,\gamma}^k - \tau_{i,\gamma}^k),$$

$$\psi_{k3,\rho\gamma} = \frac{1}{2Na_k} \sum_{i=1}^N (\sigma_{i,\rho}^k - \tau_{i,\rho}^k)(\sigma_{i,\gamma}^k + \tau_{i,\gamma}^k),$$

$$\psi_{k4,\rho\gamma} = \frac{1}{2Na_k} \sum_{i=1}^N (\sigma_{i,\rho}^k - \tau_{i,\rho}^k)(\sigma_{i,\gamma}^k - \tau_{i,\gamma}^k).$$

The integral over the $\boldsymbol{\Phi}$ in (17) can be done, yielding

$$\langle \mathcal{Z}^{nc} \rangle_\xi \propto \exp \left(-\frac{N}{2} \sum_k \alpha_k c_k \mathrm{Tr}\left(\ln \left[\mathbf{I} - \beta \boldsymbol{\Psi}_{k1} \right] \right) \right)$$

$$\times \exp \left(-\frac{N}{2} \sum_k \alpha_k c_k \mathrm{Tr}\left(\ln \left[\mathbf{I} + \beta \boldsymbol{\Psi}_{k4} \right. \right. \right.$$

$$\left. \left. \left. + \beta^2 \boldsymbol{\Psi}_{k3}(\mathbf{I} - \beta \boldsymbol{\Psi}_{k1})^{-1} \boldsymbol{\Psi}_{k2} \right] \right) \right).$$

(18)

At this point we introduce the relevant Edwards-Anderson (EA) order parameters

$$q_{k1,\rho\gamma} = \frac{1}{N} \sum_{i=1}^N \sigma_{i,\rho}^k \sigma_{i,\gamma}^k$$

(19)

$$q_{k2,\rho\gamma} = \frac{1}{N} \sum_{i=1}^N \tau_{i,\rho}^k \tau_{i,\gamma}^k$$

(20)

$$r_{k,\rho\gamma} = \frac{1}{N} \sum_{i=1}^N \sigma_{i,\rho}^k \tau_{i,\gamma}^k$$

(21)

and we redefine the $m_{kk',\rho}$ for technical reasons (recall, e.g., Eq. (16)) such that the τ and the σ can be considered separately

$$m_{kk',\rho} = m_{k1,\rho} + (-1)^{k'-1} i m_{k2,\rho}.$$

(22)

Using its self-averaging property in the thermodynamic limit, the replicated free energy can then be written in terms of these redefined overlaps, the EA order parameters

and their conjugate variables as

$$
\begin{aligned}
f_n &= \lim_{N \to \infty} -\frac{1}{N\beta} \log\left(\langle \mathcal{Z}^n \rangle_{c,\xi}\right) \\
&= \sum_k \sum_\rho m_{k1,\rho} m_{k2,\rho} - \sum_k \frac{\alpha_k}{a_k}(J_{k,0} - c_k)\sum_\rho r_{k,\rho\rho} \\
&\quad + \frac{1}{2\beta}\sum_k \alpha_k c_k \mathrm{Tr}\left(\ln[\boldsymbol{\Sigma}_k]\right) \\
&\quad - \frac{\beta}{2}\sum_k \alpha_k \sum_{\rho,\gamma}\left(2 r_{k,\rho\gamma}\tilde{r}_{k,\rho\gamma} + \sum_{k'} q_{kk',\rho\gamma}\tilde{q}_{kk',\rho\gamma}\right) \\
&\quad - \frac{\beta}{2}\sum_k \frac{\alpha_k(1-c_k)}{a_k^2}\sum_{\rho,\gamma}\left(q_{k1,\rho\gamma}q_{k2,\rho\gamma} + r_{k,\rho\gamma}r_{k,\gamma\rho}\right) \\
&\quad - \frac{1}{\beta}\left\langle \log\left[\prod_\rho\left(\mathop{\mathrm{Tr}}_{\sigma_\rho}\mathop{\mathrm{Tr}}_{\tau_\rho}\right)\exp\left(-\beta\tilde{H}(\bar\sigma,\bar\tau)\right)\right]\right\rangle_{\xi^c} \quad (23)
\end{aligned}
$$

with $\bar\sigma = \{\sigma_1,\ldots,\sigma_n\}$ denoting a vector in replica space and

$$
\begin{aligned}
\boldsymbol{\Sigma}_k &= (\mathbf{I} - \beta\boldsymbol{\Psi}_{k1})(\mathbf{I} + \beta\boldsymbol{\Psi}_{k4}) \\
&\quad + \beta^2(\mathbf{I} - \beta\boldsymbol{\Psi}_{k1})\boldsymbol{\Psi}_{k3}(\mathbf{I} - \beta\boldsymbol{\Psi}_{k1})^{-1}\boldsymbol{\Psi}_{k2}. \quad (24)
\end{aligned}
$$

Here $\tilde{H}(\bar\sigma,\bar\tau)$ is the effective single-site replicated Hamiltonian

$$
\begin{aligned}
\tilde{H}(\bar\sigma,\bar\tau) &= -\sum_k\left(\frac{1}{a_k}\eta_k\sum_\rho(\sigma_\rho^k m_{k1,\rho} + \tau_\rho^k m_{k2,\rho})\right. \\
&\quad \left. + \frac{\beta\alpha_k}{2}\sum_{\rho,\gamma}\left(2\tilde{r}_{k,\rho\gamma}\sigma_\rho^k\tau_\gamma^k + \tilde{q}_{k1,\rho\gamma}\sigma_\rho^k\sigma_\gamma^k + \tilde{q}_{k2,\rho\gamma}\tau_\rho^k\tau_\gamma^k\right)\right). \\
&\tag{25}
\end{aligned}
$$

We remark that the $\boldsymbol{\Psi}_{kl}$ can be expressed in terms of the order parameters defined in (19)–(21)

$$
\boldsymbol{\Psi}_{k1} = \frac{1}{2a_k}\left(\mathbf{q}_{k1} + \mathbf{r}_k + \mathbf{r}_k^\dagger + \mathbf{q}_{k2}\right) \quad (26)
$$

$$
\boldsymbol{\Psi}_{k2} = \frac{1}{2a_k}\left(\mathbf{q}_{k1} - \mathbf{r}_k + \mathbf{r}_k^\dagger - \mathbf{q}_{k2}\right) \quad (27)
$$

$$
\boldsymbol{\Psi}_{k3} = \frac{1}{2a_k}\left(\mathbf{q}_{k1} + \mathbf{r}_k - \mathbf{r}_k^\dagger - \mathbf{q}_{k2}\right) \quad (28)
$$

$$
\boldsymbol{\Psi}_{k4} = \frac{1}{2a_k}\left(\mathbf{q}_{k1} - \mathbf{r}_k - \mathbf{r}_k^\dagger + \mathbf{q}_{k2}\right). \quad (29)
$$

The physical meaning of the parameters becomes clear when writing

$$
m_{k1,\rho} = \frac{1}{a_k}\left\langle \eta_k\langle\sigma_\rho^k\rangle_\beta\right\rangle_{\xi^c}, \; q_{k1,\rho\gamma} = \left\langle\langle\sigma_\rho^k\sigma_\gamma^k\rangle_\beta\right\rangle_{\xi^c} (30)
$$

$$
m_{k2,\rho} = \frac{1}{a_k}\left\langle \eta_k\langle\tau_\rho^k\rangle_\beta\right\rangle_{\xi^c}, \; q_{k2,\rho\gamma} = \left\langle\langle\tau_\rho^k\tau_\gamma^k\rangle_\beta\right\rangle_{\xi^c} (31)
$$

$$
r_{k,\rho\gamma} = \left\langle\langle\sigma_\rho^k\tau_\gamma^k\rangle_\beta\right\rangle_{\xi^c}, \quad (32)
$$

where $\langle\cdot\rangle_{\xi^c}$ denotes the average over the condensed pattern and initial conditions and where the thermal average $\langle\cdot\rangle_\beta$

is defined as

$$
\begin{aligned}
&\langle A(\bar\sigma,\bar\tau)\rangle_\beta = \\
&\frac{\prod_\rho\left(\mathop{\mathrm{Tr}}_{\sigma_\rho}\mathop{\mathrm{Tr}}_{\tau_\rho}\right) A(\bar\sigma,\bar\tau)\exp\left(-\beta\tilde{H}(\bar\sigma,\bar\tau)\right)}{\prod_\rho\left(\mathop{\mathrm{Tr}}_{\sigma_\rho}\mathop{\mathrm{Tr}}_{\tau_\rho}\right)\exp\left(-\beta\tilde{H}(\bar\sigma,\bar\tau)\right)}. \quad (33)
\end{aligned}
$$

Next, we want to study the replicated free energy and its minima within, as usual, the replica-symmetric approximation in order to obtain the phase diagrams for the model. To make this picture complete we also have to address the occurrence of two-cycles as possible stationary solutions of the relevant saddle-point equations.

3.2 Replica symmetry and non-cycle ansatz

In [4] it has been stated, within a replica-symmetric treatment, that in the Little-Hopfield model ([18,19]) cycles do not occur in the replica-symmetric retrieval phase. In that work the authors assume, however, that the order parameter $r_{k,\rho\gamma}$ (recall (21)) is symmetric, implying already absence of two-cycles in the retrieval phase right from the start. Extensive simulations for the Little-Hopfield model have shown [10] that, at least in the retrieval phase, cycles do not occur.

We have performed extensive simulations for the fully connected BEG neural network with equal dilution parameters (implying $\alpha_1 = \alpha_2 \equiv \tilde\alpha$) for uniform patterns ($a = 2/3$) with synchronous updating at zero temperature. The basic quantity we have studied is the *cycle parameter* $\Delta(t)$, defined as

$$
\Delta(t) = \frac{1}{N}\sum_{i=1}^N (\sigma_i(t) - \sigma_i(t-1))^2. \quad (34)
$$

The results are shown in Figure 1. In particular, we have run simulations for several values of the capacity $\tilde\alpha$ at zero temperature during 1000 time steps of synchronous updating, which in all cases reached stationarity (i.e. a constant value $\Delta(t) \geq 0$). We have determined how many of the runs (200 in total) end up in a cycle and how many in a fixed-point. Moreover, we have computed the average value of $\Delta(t)$ for the cycles. In addition, we have checked what is the average number of neurons involved in the cycles. To have an idea of the finite size effects, the simulations are performed for $N = 4000$ up to $N = 7000$ neurons. From the results it seems natural to conjecture that in the limit $N \to \infty$ no cycles will exist in the retrieval region ($\tilde\alpha \leqslant 0.091$ for $T = 0$ [8]). In Figure 1 we show results for the initial conditions $m_k(0) = 0.6$, $q_0(0) = 0.7$ but simulations for several other initial conditions (even for $m_k(0) = 0$) lead to the same conclusion.

The first order transition from retrieval to spin-glass behaviour includes as well a transition from a region where two-period cycles are not occurring to a region in which they basically dominate the space of solutions, as can be
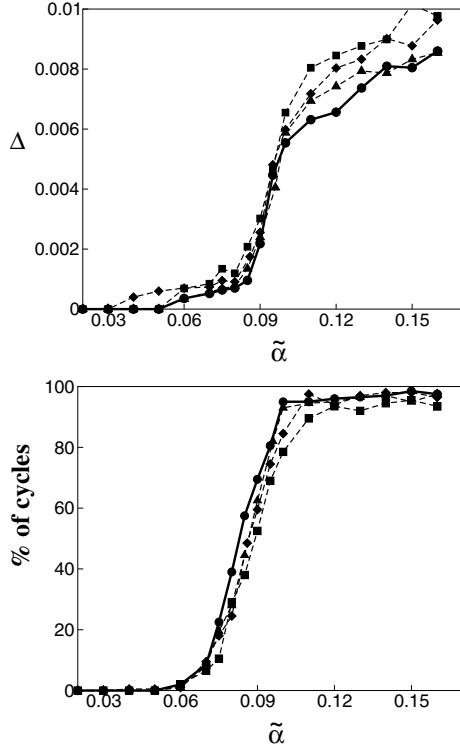
Fig. 1. Simulations with $N = 4000$ (squares), $N = 5000$ (diamonds), $N = 6000$ (triangles) and $N = 7000$ (circles) for $T = 0$ and $m_k(0) = 0.6, q_0(0) = 0.7$. We show the average value of $\Delta(t)$ for the runs that reached a cycle (top) and the percentage of cycles among the 200 runs performed (bottom). Only the results for the largest system size have been plotted with a full line.

concluded from the figure showing the percentage of cycles.

Furthermore, we see that the cycles never involve a large number of neurons (always less than 0.5% of the total for the simulations shown in the figures). The average number of flipping neurons increases monotonically with $\tilde{\alpha}$ but does seem to saturate with increasing system size. This is in agreement with the results obtained in a finite-size scaling study of the capacity problem for the Hopfield model [20] where cycles also involve only a tiny fraction of all neurons. Moreover, these authors find that there is always a fixed-point solution of the sequential updating nearby. In this sense, cycles seem to be stationary solutions in which the system is unable to reach an accessible spin-glass minimum but jumps around it.

For $T > 0$, one does not expect new cycles to appear, in analogy with the fact that one does not find new fixed-point attractors different from those at $T = 0$.

Given these findings, we neglect cycles in the sequel of the replica calculation since we are interested mainly in the retrieval phase. In order to implement this non-cycle ansatz, we eliminate any distinction between σ and

τ. Therefore, we introduce the following assumptions

$$m_{k1,\rho} = m_{k2,\rho} = m_{k,\rho} \tag{35}$$

$$q_{k1,\rho\gamma} = q_{k2,\rho\gamma} = q_{k,\rho\gamma}, \quad \tilde{q}_{k1,\gamma} = \tilde{q}_{k2,\rho\gamma} = \tilde{q}_{k,\rho\gamma} \tag{36}$$

$$r_{k,\rho\gamma} = q_{k,\rho\gamma}, \quad \text{only for } \rho \neq \gamma. \tag{37}$$

We explicitly remark that introducing $r_{k,\rho\rho} = q_{k,\rho\rho}$ would not be correct since both quantities imply products of spins from the same replica, but σ and τ are affected by different thermal fluctuations.

Finally, we impose symmetry in the replica space. Before doing so we solve for the conjugate variables in order to further simplify the free energy. This allows us to replace $\tilde{r}_{k,\rho\gamma}$ by $\tilde{q}_{k,\rho\gamma}$ for $\rho \neq \gamma$. The replica symmetry (RS) ansatz then implies

$$m_{k,\rho} = m_k \tag{38}$$

$$q_{k,\rho\gamma} = q_k, \quad q_{k,\rho\rho} = q_0 \tag{39}$$

$$\tilde{q}_{k,\rho\gamma} = \tilde{q}_k, \quad \tilde{q}_{k,\rho\rho} = \tilde{q}_{k,0} \tag{40}$$

$$r_{k,\rho\rho} = r_{k,0}, \quad \tilde{r}_{k,\rho\rho} = \tilde{r}_{k,0}. \tag{41}$$

As a result, the free energy per spin in the limit $n \to 0$ can be written as

$$\lim_{n \to 0} \frac{f_n}{n} = \sum_k m_k^2$$

$$+ \frac{1}{2\beta} \sum_k \alpha_k c_k \left(\ln \left[(1 - \chi_{kr})^2 - \chi_k^2 \right] - \frac{2\beta q_k}{a_k(1 - \chi_{kr} - \chi_k)} \right)$$

$$- \sum_k \frac{\alpha_k}{a_k} (J_{k,0} - c_k) r_{k,0}$$

$$- \frac{\beta}{2} \sum_k \frac{\alpha_k(1 - c_k)}{a_k^2} \left(q_0^2 - 2q_k^2 + r_{k,0}^2 \right)$$

$$+ \frac{\beta}{2} \sum_k \alpha_k (2q_0 \tilde{q}_{k,0} - 4q_k \tilde{q}_k + 2r_{k,0} \tilde{r}_{k,0})$$

$$- \frac{1}{\beta} \left\langle \int \mathcal{D}z_1 \mathcal{D}z_2 \log \left[\underset{\sigma}{\text{Tr}} \underset{\tau}{\text{Tr}} \exp \left(-\beta \tilde{H}(\sigma, \tau; z_k) \right) \right] \right\rangle_{\xi^c} \tag{42}$$

with z_1 and z_2 Gaussian variables with measure $\mathcal{D}z = dz(2\pi)^{-1/2} \exp(-z^2/2)$. The effective Hamiltonian now reads

$$\tilde{H}(\sigma, \tau; z_k) = - \sum_k \left(\frac{1}{a_k} \eta_k m_k(\sigma^k + \tau^k) + \alpha_k a_k \tilde{\chi}_{kr} \sigma^k \tau^k \right.$$

$$\left. + \frac{1}{2}(\sigma^2 + \tau^2)\alpha_k a_k \tilde{\chi}_k + \sqrt{\alpha_k \tilde{q}_k}(\sigma^k + \tau^k)z_k \right) \tag{43}$$

with the implicit definitions for the susceptibilities

$$\chi_k = \frac{\beta}{a_k}(q_0 - q_k), \chi_{kr} = \frac{\beta}{a_k}(r_{k,0} - q_k) \tag{44}$$

$$\tilde{\chi}_k = \frac{\beta}{a_k}(\tilde{q}_{k,0} - \tilde{q}_k), \tilde{\chi}_{kr} = \frac{\beta}{a_k}(\tilde{r}_{k,0} - \tilde{q}_k). \tag{45}$$

We remark that by setting $a = 1$, as well as $J_{2,0} = r_{2,0} = q_2 = \tilde{r}_{2,0} = \tilde{q}_2 = 0$, $\tilde{q}_{2,0} = \tilde{q}_{1,0} = 0$ and $q_0 = 1$ we exactly recover the free energy for the Hopfield model with synchronous updating and variable dilution [10].

3.3 Saddle-point equations

The extremization of the free energy in (42) yields

$$m_k = \frac{1}{a_k} \left\langle \eta_k \left\langle \sigma^k \right\rangle_\beta \right\rangle_{\xi^c, z}, \tag{46}$$

$$q_0 = \left\langle \left\langle \sigma^2 \right\rangle_\beta \right\rangle_{\xi^c, z}, \quad q_k = \left\langle \left\langle \sigma^k \right\rangle_\beta^2 \right\rangle_{\xi^c, z}, \tag{47}$$

$$r_{k,0} = \left\langle \left\langle \sigma^k \tau^k \right\rangle_\beta \right\rangle_{\xi^c, z}, \tag{48}$$

with $\langle \cdot \rangle_z$ the average over the Gaussian noise (recall the last term in (42)) and $\langle \cdot \rangle_\beta$ the thermal average performed with respect to (43). The conjugate variables satisfy

$$\tilde{q}_k = \frac{q_k}{a_k^2} \left(1 - c_k + \frac{c_k}{(1 - \chi_{kr} - \chi_k)^2} \right), \tag{49}$$

$$\tilde{\chi}_k = \frac{\chi_k}{a_k^2} \left(1 - c_k + \frac{c_k}{(1 - \chi_{kr})^2 - \chi_k^2} \right), \tag{50}$$

$$\tilde{\chi}_{kr} = \frac{1}{a_k^2} \left((1 - c_k)\chi_{kr} + \frac{c_k(1 - \chi_{kr})}{(1 - \chi_{kr})^2 - \chi_k^2} + J_{k,0} - c_k \right). \tag{51}$$

In the absence of thermal noise ($T = 0$), the system reaches a minimum of the free energy such that $\sigma_i = \tau_i$, $\forall i$. Therefore, it is interesting to consider the limit $\sigma = \tau$ of the free energy and the saddle-point equations. The trace of the effective partition function (recall (43)) has to be evaluated again and, in addition, the following equivalences hold

$$r_{k,0} = q_0, \quad \chi_{kr} = \chi_k, \tag{52}$$

$$\tilde{\chi}_{kr} = \tilde{\chi}_k, \quad \tilde{r}_{k,0} = \tilde{q}_{k,0} + \frac{1}{a_k \beta} J_{k,0}. \tag{53}$$

For zero self-couplings ($J_{k,0} = 0$), this leads to the relation

$$\lim_{\beta \to \infty} \beta f_{syn}(\beta) = \lim_{\beta \to \infty} 2\beta f_{seq}(2\beta), \tag{54}$$

as is also the case for the $Q = 3$ Ising network [10]. This is easy to understand when noticing that for $\tau_j^k = \sigma_j^k$ the Hamiltonian (12) for synchronous updating becomes simply twice the one for sequential updating. An immediate consequence is that the saddle-point equations are identical to those for sequential updating [8,10].

For the saddle-point equations at $T = 0$ we then obtain, with obvious notation

$$m_k = \frac{1}{a_k} \left\langle \eta_k \text{sign}^k(\tilde{h}_1(z_1)) \, \Theta(|\tilde{h}_1(z_1)| + \tilde{h}_2(z_2)) \right\rangle_{\xi^c, z} \tag{55}$$

$$q_0 = \left\langle \Theta(|\tilde{h}_1(z_1)| + \tilde{h}_2(z_2)) \right\rangle_{\xi^c, z} \tag{56}$$

$$\chi_k = \frac{1}{a_k \sqrt{\alpha_k \tilde{q}_k}}$$
$$\times \left\langle z_k \, \text{sign}^k(\tilde{h}_1(z_1)) \, \Theta(|\tilde{h}_1(z_1)| + \tilde{h}_2(z_2)) \right\rangle_{\xi^c, z} \tag{57}$$

where the effective fields are given by

$$\tilde{h}_k(z_k) = \frac{1}{a_k}\eta_k m_k + \delta_{k,2} \sum_{k'=1}^{2} \alpha_{k'} \left(a_{k'} \tilde{\chi}_{k'} + \frac{J_{k,0}}{2a_{k'}} \right) + \sqrt{\alpha_k \tilde{q}_k} z_k. \tag{58}$$

The conjugated variables read

$$\tilde{q}_k = \frac{q_k}{a_k^2} \left(1 - c_k + \frac{c_k}{(1 - \chi_k)^2} \right), \tag{59}$$

$$\tilde{\chi}_k = \frac{\chi_k}{a_k^2} \left(1 - c_k + \frac{c_k}{1 - \chi_k} \right). \tag{60}$$

4 Thermodynamic and retrieval properties

In this section we study the thermodynamic and retrieval properties of the BEG network with variable dilution and synchronous updating by numerically solving the saddle-point equations (46)–(51) and compare the results with those for sequential updating. As in the model with sequential updating [8], there are four phases: the retrieval phase R ($m_k > 0$, $q_1 > 0$), the spin-glass phase SG ($m_k = 0$, $q_1 > 0$), the paramagnetic phase P ($m_k = 0$, $q_1 = 0$) and the quadrupolar phase Q ($m_1 = 0$, $m_2 > 0$, $q_1 > 0$). In the sequel we discuss the phase diagrams for low loading ($\alpha_k = 0$), finite loading ($\alpha_k > 0$), k-type dilution and the effect of the self-couplings $J_{k,0}$.

4.1 Low loading

For $\alpha_k = 0$, the saddle-point equations can be derived without the use of replicas (e.g., [21,22] for the Hopfield model) and we find

$$m_{k,\sigma} = \frac{1}{a_k} \left\langle \eta_k \frac{\text{Tr}_\sigma \sigma^k \exp\left(\beta \sum_{k'} \frac{1}{a_k} \eta_{k'} m_{k',\tau} \sigma^{k'} \right)}{\text{Tr}_\sigma \exp\left(\beta \sum_{k'} \frac{1}{a_k} \eta_{k'} m_{k',\tau} \sigma^{k'} \right)} \right\rangle_{\xi^c} \tag{61}$$

and similarly for $m_{k,\tau}$. One can show that the only minimum of the free energy is given by $m_{k,\sigma} = m_{k,\tau}$, such that the solution coincides which the one for sequential updating. Physically, this means that in equilibrium there can be no cycles: the two neuron variables have decoupled becoming independent and leading to $\beta f_{syn}(\beta) = 2\beta f_{seq}(\beta)$, as seen in the other models mentioned above [1,10,21,22]. Hence, the dynamics does not influence the equilibrium solutions and the $a - T$ phase diagram is identical to the one for sequential updating. Furthermore, we know [8] that for low loading the form of the architecture (fully connected, extremely diluted) is not important. Hence we do not treat this case in more detail here but we refer to, e.g., [8] (Fig. 1) and the discussion therein.
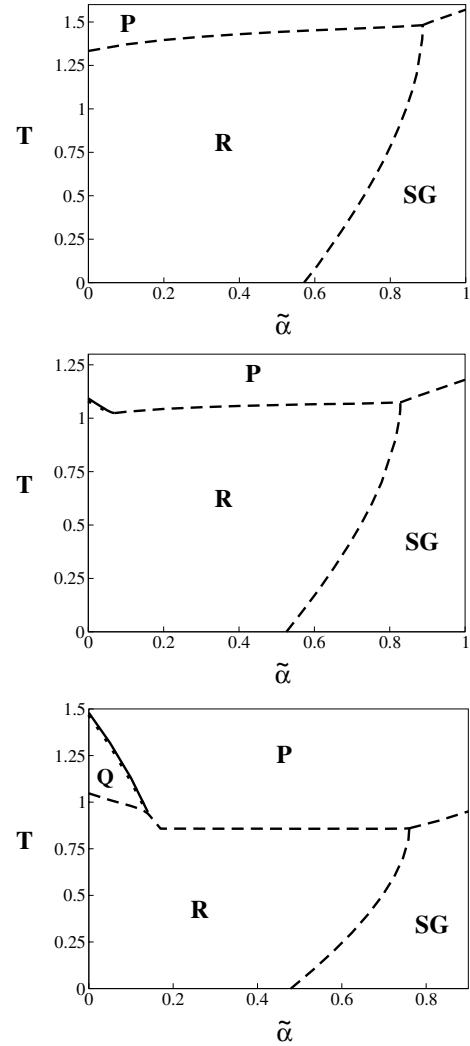
4.2 Finite loading

For the fully connected architecture, $c_k = 1$, Figure 2 presents the capacity-temperature phase diagrams for
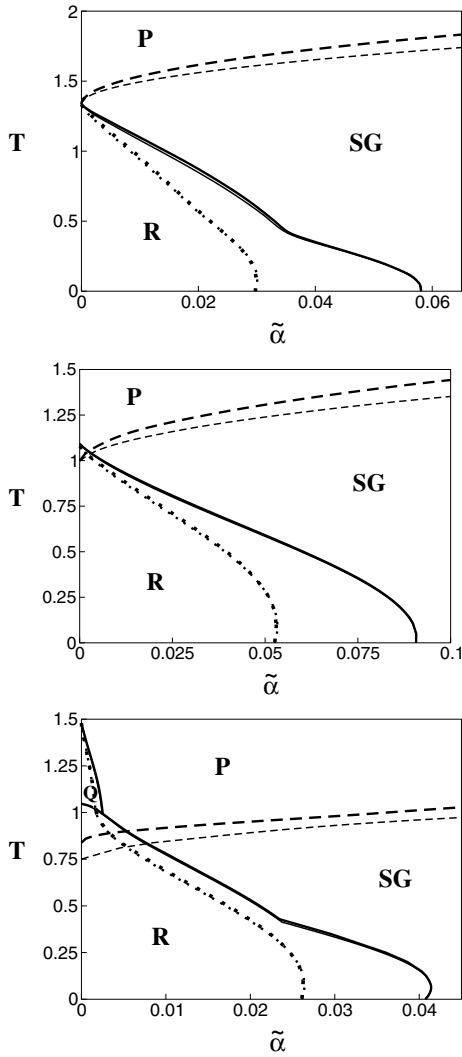
Fig. 2. The RS $\tilde{\alpha} - T$ phase diagram for the fully connected BEG neural network with activities, from top to bottom, $a = 0.5$, $2/3$ and 0.8 for synchronous and sequential updating (thick respectively thin lines).

Fig. 3. The RS $\tilde{\alpha} - T$ phase diagram for the extremely diluted BEG neural network with activities, from top to bottom, $a = 0.5$, $2/3$ and 0.8.

three typical values of the activity $a = 0.5$, $2/3$ (uniformly distributed patterns) and 0.8. For the moment, the self-couplings are set to zero and the dilution parameters are equal, i.e., $c_1 = c_2$ (recall that in that case $\alpha_k = \tilde{\alpha}$). Dashed lines correspond to second order transitions, full lines to first order transitions and dotted lines to thermodynamic transitions. The thick lines indicate the results for synchronous updating; for comparison we have indicated those for sequential updating [8] with thin lines.

As can be seen the main features for both types of updating are very similar. For synchronous updating the spin glass region increases visibly while the retrieval region increases marginally. A similar behaviour is observed for the Hopfield model [4,10]. We note the well-known weak re-entrant behaviour in the retrieval region at low temperatures signaling the breaking of the RS approximation. A calculation of the entropy at zero temperature for the model with asynchronous updating [8] suggests that

this breaking is comparable to that of the Hopfield model. Finally, for $a = 0.8$, the quadrupolar region is slightly enlarged as well.

For extreme symmetric dilution ($c_k = 0$) the phase diagrams nearly coincide with the ones for sequential updating [10]. Since both sets of results did not yet appear in the literature we show them in Figure 3. Compared with the fully connected architecture, most of the first order transitions become second order. Only the quadrupolar-paramagnet first order transition remains. Furthermore, there is strong re-entrant behaviour in the retrieval region. This is related to the failure of the RS approximation at low temperatures. One expects that in a full replica calculation the transition between the retrieval and the spin glass region is a vertical line starting from the tricritical point. This has also been observed in the Hopfield model [12,13] and in the $Q = 3$-Ising model [23]. We remark that, at first instance, the general behaviour of this model under RS seems to be very similar to the known results for other three state models such that an explicit

4.3 k-type dilution

Next, we look at the effect of k-type dilution on the system (recently introduced as asynchronous dilution in a Gardner optimal capacity calculation [25]). Indeed, the BEG model allows one to prune independently the two types of couplings appearing in the Hamiltonian (12).

The results can be summarized as follows. The performance of the network strongly depends on the activity. In any case, the ratio p/N decreases when the pruning becomes stronger, as one expects, while the standard capacity $\alpha = p/cN$ shows a maximum.

Hence, given the number of connections per neuron, $c = c_1 + c_2$ (with $0 \leq c \leq 2$, recall Eq. (8)) there has to be an optimal combination of both dilution variables c_1 and c_2. In order to be able to predict this combination once the other parameters of the system are known we have performed some numerical experiments. We present in Figure 4 the critical capacity α_c for the activities $a = 0.5$, $2/3$ and 0.8 as a function of c_1. In each figure the values of c are fixed at $c = 0.5$, 0.75, 1.0, 1.25 and 1.5, from top to bottom, and the temperature is fixed at $T = 0.5$.

We find a maximum in all figures, corresponding to the optimal combination of the dilution parameters $c = c_1^* + c_2^*$. Looking, e.g., at the optimal value c_1^* as a function of the number of connections per neuron c for $T = 0.5$ for several values of the activity, $a = 0.9$, 0.8, $2/3$, 0.5 and 0.4 we find that these data points can be fitted by

$$c_1^* = (1-a)(c_1 + c_2) = (1-a)c. \tag{62}$$

We conjecture that the temperature does not modify the position of the maximum but only its height, since c_1^* is determined by the role of both terms in the Hamiltonian and therefore by the activity.

Furthermore, we notice a second maximum for $a = 0.8$. This is due to the existence of a second retrieval solution occurring for large activities [8]. This solution only appears for $0 < c \lesssim 1.4$ (for $T = 0.5$). We can fit this second peak by

$$c_1^* = a(c_1 + c_2) = ac. \tag{63}$$

Since for larger activities this relation implies that there are fewer $k = 2$-type connections at the maximum c_1^*, the second retrieval solution has a smaller overlap m_2.

Furthermore, when c_1 is kept finite, both critical capacities for $c_2 = 0$ are always finite. However, when c_2 is kept finite, setting $c_1 = 0$ causes m_1 to be zero. This suggests that it is actually the first term of the Hamiltonian that mainly ensures the process of retrieval. Keeping only the second term of the Hamiltonian, the system still recognizes correctly the zero and the active states, but does not discern between the $+1$ and -1 states. It still implies though that the information content of the system is nonzero. A practical realization of this behaviour may be in pattern recognition where, looking at black and white
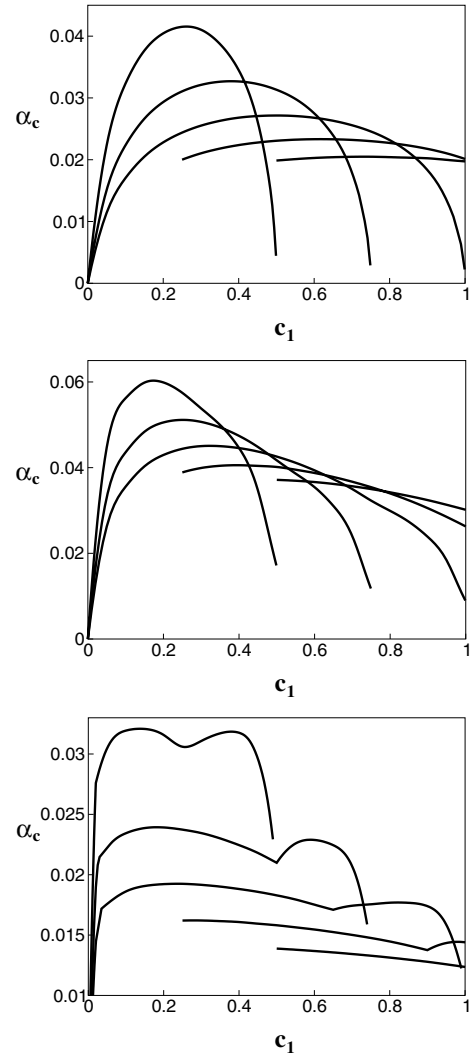
Fig. 4. The RS critical capacity at $T = 0.5$ as a function of c_1. From top to bottom, $a = 0.5$, $2/3$ and 0.8. In each figure, from top to bottom, $c = 0.5$, 0.75, 1.0, 1.25 and 1.5.

pictures on a grey background, the system would tell us the exact location of the picture with respect to the background without finding the details of the picture itself. These findings seem to be in agreement with the results in [25] for the Gardner optimal capacity in the presence of dilution, where it is concluded that the BEG network is indeed more robust against $k = 2$-type dilution of the couplings.

4.4 The self-couplings effect

Finally, we turn to the study of the effect of self-coupling on the phase diagrams of the model. In the case of very large but finite self-couplings $J_{1,0}$ and/or $J_{2,0}$, the σ and τ tend to stay in the same state. Indeed, when $J_{1,0}$ respectively $J_{2,0}$ are very large but finite, the two spins tend to fulfill $\sigma_i = \tau_i$ respectively $\sigma_i^2 = \tau_i^2$, and the parallel

Hamiltonian can be written as

$$\lim_{J_{1,0}\gg 1} H_{syn}(\boldsymbol{\sigma}) = 2H_{seq}(\boldsymbol{\sigma}) - \sum_{k=1}^{2}\frac{\alpha_k}{a_k}J_{k,0}\sum_{i=1}^{N}\delta_{|\sigma_i|,1}, \quad (64)$$

$$\lim_{J_{2,0}\gg 1} H_{syn}(\boldsymbol{\sigma}) = 2H_{seq}(\boldsymbol{\sigma}) + \sum_{i,j\neq i}J_{1,ij}^{c}\sigma_i(\sigma_j - \tau_j)$$
$$-\frac{\alpha_1}{a_1}J_{1,0}\sum_{i=1}^{N}\sigma_i\tau_i - \frac{\alpha_2}{a_2}J_{2,0}\sum_{i=1}^{N}\delta_{|\sigma_i|,1}. \quad (65)$$

In both cases, the last term favours the elimination of the zero states from the system as can also be seen from equations (4)–(5) in these limits. The order parameters of the system evolve to $m_2 = 0$, $q_0 = 1$, $q_2 = 1$, $r_{2,0} = 1$ and

(i) for $J_{1,0} \to \infty$ or $J_{1,0}$ and $J_{2,0} \to \infty$

$$m_1 = \frac{1}{a_1}\left\langle \eta_1 \tanh\left(2\beta\tilde{h}_1(z_1)\right)\right\rangle_{\xi^c, z_1} \quad (66)$$

$$q_1 = \left\langle \tanh^2\left(2\beta\tilde{h}_1(z_1)\right)\right\rangle_{\xi^c, z_1} \quad (67)$$

$$r_{1,0} = 1. \quad (68)$$

(ii) for $J_{2,0} \to \infty$

$$m_1 = \frac{1}{a_1}\left\langle \eta_1 \frac{\sinh\left(2\beta\tilde{h}_1(z_1)\right)}{\cosh\left(2\beta\tilde{h}_1(z_1)\right) + e^{-2\beta\alpha_1 a_1\tilde{\chi}_{1r}}}\right\rangle_{\xi^c, z_1} \quad (69)$$

$$q_1 = \left\langle \frac{\sinh^2\left(2\beta\tilde{h}_1(z_1)\right)}{\cosh^2\left(2\beta\tilde{h}_1(z_1)\right) + e^{-4\beta\alpha_1 a_1\tilde{\chi}_{1r}}}\right\rangle_{\xi^c, z_1} \quad (70)$$

$$r_{1,0} = \left\langle \frac{\cosh\left(2\beta\tilde{h}_1(z_1)\right) - e^{-2\beta\alpha_1 a_1\tilde{\chi}_{1r}}}{\cosh\left(2\beta\tilde{h}_1(z_1)\right) + e^{-2\beta\alpha_1 a_1\tilde{\chi}_{1r}}}\right\rangle_{\xi^c, z_1} \quad (71)$$

with $\tilde{h}_1(z_1)$ given by (58) ($k = 1$). As can be seen from these equations, in the first limiting case we arrive at a Hopfield model with sequential updating when $a \to 1$, with the usual rescaling in β ($\beta \to \beta/2$). For the second limiting case, we only evolve to the Hopfield model with sequential updating in the limit $\beta \to \infty$, $a \to 1$. Indeed, in the first case, the second local field h_2 becomes irrelevant when $a = 1$, and the Hamiltonian reduces to a Hopfield-like one with a rescaled β and $\sigma_i = \tau_i$ for any temperature. In the second case $\sigma_i^2 = \tau_i^2$ for any temperature, but only for $T = 0$ we are ensured that $\sigma_i = \tau_i$ as well.

In Figure 5 some phase diagrams in the presence of self-couplings $J_{k,0} = a_k$ are shown in comparison with the results without self-couplings.

For the fully connected architecture, $c_k = 1$, there is little difference between synchronous and sequential updating. While the spin glass region is visibly enlarged, the retrieval region has increased only marginally.

The extremely diluted network is much more sensitive to the presence of self-couplings. The retrieval region is
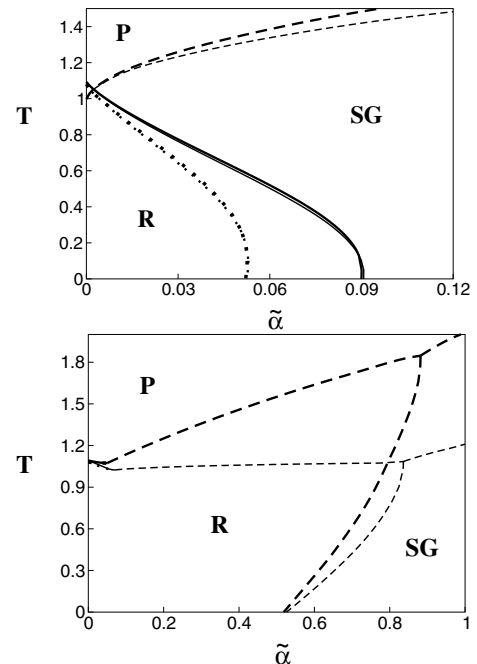

Fig. 5. The RS $\tilde{\alpha}-T$ phase diagrams for $a = 2/3$ with $J_{k,0} = 0$ (thin lines) and $J_{k,0} = a_k$ (thick lines) for the fully connected (top) and the extremely diluted (bottom) BEG networks.

substantially enlarged, as well as the spin glass region. A new feature is the different re-entrance point for different values of the self-coupling. This is not observed in the Hopfield model [10] because the limit of large self-couplings in that case is the Hopfield model for sequential updating, which is known to have the same RS critical capacity. Here the phase diagrams must evolve from the BEG phase diagram with synchronous updating to the Hopfield one with sequential updating, which has a different critical capacity. Therefore, the re-entrance point also evolves as we increase $J_{k,0}$. A similar effect has been observed for the Q-Ising model with finite dilution ($c = 0.01$, [10]).

Hence, the self-couplings in three-state neural networks with variable dilution are very relevant. Different values can produce families of phase diagrams that evolve towards the Hopfield phase diagram for sequential updating with the rescaling in β mentioned before. An illustration of this effect is presented, for $T = 0$, in Figure 6, where the critical capacity is shown for different values of the self-couplings as a function of the activity. We remark that for $a = 1$ we recover the value for the Hopfield model, $\alpha_c \simeq 0.139$. A maximum in the critical capacity appears unless the self-couplings are large enough so that the three-state nature of the model is breaking down since all zero states are completely eliminated.

5 Concluding remarks

We have considered the thermodynamic and retrieval properties of BEG networks with synchronous updating and variable dilution. Saddle-point equations for the relevant order parameters have been derived for arbitrary
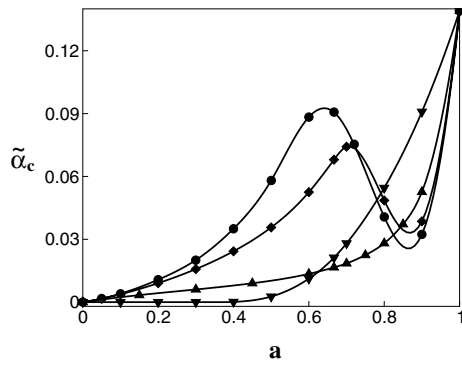
Fig. 6. The RS critical capacity as a function of the activity a for $T = 0$ and different values of the self-couplings $J_{k,0} = 0$ (circles), 10 (diamonds), 100 (triangles up) and ∞ (triangles down).

temperature using a replica mean-field treatment. In the stationary limit, the system allows for fixed-points as well as two-cycles. It is found numerically that two-cycles only occur in the spin-glass phase of the system, never involve a large number of neurons and are always close to a fixed-point spin glass solution.

Making a non-cycle and replica symmetry ansatz, the phase diagrams of the model are obtained. Two types of dilution involving the two types of couplings are discussed. The optimal combination of the network dilution parameters giving the largest critical capacity is obtained. It is also found that the model is more robust against dilution of the $k = 2$-type couplings.

The results for different pattern activities and variable dilution are compared with those for sequential updating. It is observed that the asymptotic behaviour of the network with synchronous and sequential updating is almost identical. For some values of the network parameters the retrieval region is enhanced marginally. Only the presence of self-couplings can enlarge the retrieval region substantially, especially in the case of extreme dilution.

References

1. D. Bollé, J. Busquets Blanco, Eur. Phys. J. B **42**, 397 (2004)
2. P. Peretto, *An introduction to the modeling of neural networks* (Cambridge University Press, Cambridge, 1992), Chap. 3
3. A.C.C Coolen, in *Handbook of Biological Physics*, Vol 4, edited by F. Moss, S. Gielen (Elsevier Science B.V., 2001), p. 531
4. J.F. Fontanari, R. Köberle, J. Phys. France **49**, 13 (1988); J.F. Fontanari, R. Köberle, Phys. Rev. A **36**, 2475 (1987)
5. D.R.C. Dominguez, E. Korutcheva, Phys. Rev. E **62**, 2620 (2000)
6. D. Bollé, T. Verbeiren, Phys. Lett. A **297**, 156 (2002)
7. D. Bollé, in *Advances in condensed matter and statistical mechanics*, edited by E. Korutcheva, R. Cuerno (Nova Science Publishers, New York, 2004), 319
8. D. Bollé, T. Verbeiren, J. Phys. A **36**, 295 (2003)
9. D. Bollé, D.R.C. Dominguez, R. Erichsen Jr., E. Korutcheva, W.K. Theumann, Phys. Rev. E **68**, 062901 (2003)
10. T. Verbeiren, Ph.D. thesis, K.U. Leuven, Leuven, Belgium (2003)
11. P. Peretto, Biol. Cybern. **50**, 51 (1984)
12. T.L.H. Watkin, D. Sherrington, Europhys. Lett. **14**, 791 (1991)
13. A. Canning, J.-P. Naef, J. Phys. I France **2**, 1791 (1992)
14. L. Viana, A.J. Bray, J. Phys. C: Solid State Phys. **18**, 3037 (1985)
15. R. Bidaux, J.P. Carton, G. Sarma, J. Phys. A **9**, L87 (1976)
16. W.K. Theumann, R. Erichsen Jr., Phys. Rev. E **64**, 061902 (2001)
17. J. Busquets Blanco, Ph.D. thesis, K.U. Leuven, Leuven, Belgium (2005)
18. W.A. Little, Math. Biosci. **19**, 101 (1974)
19. J.J. Hopfield, Proc. Nat. Acad. Sci. USA **79**, 2554 (1982)
20. T. Stiefvater, K.-R. Müller, R. Kühn, Physica A **232**, 61 (1996)
21. D.J. Amit, H. Gutfreund, H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985)
22. J.L. van Hemmen, Phys. Rev. A **34**, 3435 (1986)
23. D. Bollé, D.M. Carlucci, G.M. Shim, J. Phys. A **33**, 6481 (2000)
24. J.R.L. de Almeida, D.J. Thouless, J. Phys. A **11**, 983 (1987)
25. D. Bollé, I. Pérez Castillo, Physica A **349**, 548 (2005)