Mixture states and storage of biased patterns in Potts-glass neural networks

D. Bollé* and J. Huyghebaert
Instituut voor Theoretische Fysica and Interdisciplinair Centrum voor Neurale Netwerken, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium
(Received 9 April 1993)

The presence and stability of mixture states in Q-state Potts neural networks are studied for different learning rules within the replica-symmetric mean-field-theory approach. The retrieval properties of the asymmetric mixture states are examined in the case of biased patterns. For the storage of a finite number of such patterns, these properties are compared for the usual Hebb learning rule and some variants obtained by subtracting, for a certain pattern, the average of the Potts neuron state over all the other patterns. The latter are introduced to suppress the symmetric mixture states. Furthermore, the embedding of an additional, infinite number of unbiased patterns stored with the Hebb rule is allowed. The storage capacity and the temperature-capacity phase diagram are discussed in these cases. A detailed analysis is made for the $Q=3$ model and two classes of representative bias parameters.

PACS number(s): 87.10.+e, 64.60.Cn, 75.10.Hk

I. INTRODUCTION

Neural networks with multistate neurons based upon the Q-state Potts-glass model have been analyzed in order to study, e.g., the storage and retrieval properties of gray-toned patterns. For an overview of the recent literature in this respect, we refer to Ref. [1]. In that paper a systematic treatment of the Mattis retrieval states and the symmetric mixture states for Q-state Potts neural networks with biased patterns and a Hebb learning rule, involving the bias parameters, has been given.

In general, for the storage and retrieval of (statistically) correlated patterns, the mixture states are important since the state of the network must have a nonzero overlap with all the patterns. As a first step towards this problem this type of states can be analyzed for biased (i.e., statistically independent but effectively correlated) patterns. The symmetric mixture states describe the inability of the network to perceive the details distinguishing the different patterns and hence should be suppressed. This can be done by different methods, e.g., a global dynamic constraint, an adjustable uniform field, a different learning rule (see, respectively, [2–4] for the Hopfield model). Some of the asymmetric mixture states, however, describe the retrieval behavior of the network. Hence a detailed study of these different mixture states for different learning rules appears to be interesting.

In the present work we consider symmetric and asymmetric mixture states and the storage of biased patterns for the Q-state Potts network with the usual Hebb rule and some variants obtained by subtracting, for a certain pattern, the average of the Potts neuron state over all the other patterns. These variants of the Hebb rule will have the property to suppress the unwanted symmetric mixture states. Since the treatment of biased patterns in Potts networks is rather involved and tedious (see, e.g., [1]), we restrict ourselves here to the storage of a finite number of biased patterns in combination with an infinite number of unbiased patterns. Analog problems have been studied for the Hopfield model: mixture states for low loading have been considered in [4] for the Hebb learning rule and some other local learning rules; in [5] mixture states for extensive loading have been discussed for the Hebb rule.

The rest of this paper is organized as follows. In Sec. II the model is introduced. Section III discusses, within mean-field theory, the low loading of biased patterns for the usual Hebb learning rule and its modifications introduced above. The possible types of solutions of the fixed-point equation for the overlap are determined and their stability is analyzed. The retrieval behavior as a function of the bias for the different learning rules is compared explicitly and temperature-bias stability diagrams are studied for some representative $Q=3$ models. In Sec. IV an extensive loading of patterns is considered whereby a finite number of biased patterns is learned with the different rules discussed before and the rest is learned with the usual Hebb rule. Replica-symmetric fixed-point equations for the order parameters are written down for general Q and arbitrary temperature T. The retrieval quality and the storage capacity α at zero temperature are discussed as functions of the bias and T-α phase diagrams are analyzed in detail for the $Q=3$ models introduced in Sec. III. Finally some conclusions comparing the different learning rules are drawn in Sec. V.

II. MODEL

Consider a system of N neurons. Each neuron can be described by a Potts spin $\sigma_i \in \{1,2,\ldots,Q\}$, $i=1,2,\ldots,N$. The neurons are interconnected with all the others by a synaptic matrix of strength J_{ij}^{kl} which determines the contributions of a signal fired by the jth presynaptic neuron in state l to the postsynaptic potential which acts on the ith neuron in state k. The energy potential h_{i,σ_i} of neuron i which is in a state σ_i is given by

$$h_{i,\sigma_i} = -\sum_{\substack{j=1 \\ j\neq i}}^{N} \sum_{k,l=1}^{Q} J_{ij}^{kl} u_{\sigma_i,k} u_{\sigma_j,l} , \tag{1}$$

with u the Potts spin operator defined as

$$u_{\sigma_i,k}=Q\delta_{\sigma_i,k}-1 \ . \tag{2}$$

The dynamics of the Q-state Potts model is defined as in [6]. At zero temperature the state of the neuron in the next time step is fixed to be the state which minimizes the induced local field (1). The stable states of the system are those configurations $\{\sigma_i\}$ where every neuron is in a state which gives a minimum value to $\{h_{i,\sigma_i}\}$. For symmetric couplings, i.e., $J_{ij}^{kl}=J_{ji}^{lk}$, this stability is equivalent to the requirement that the configurations $\{\sigma_i\}$ are the local minima of the Potts Hamiltonian

$$H=-\tfrac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N}\sum_{k,l=1}^{Q}J_{ij}^{kl}u_{\sigma_i,k}u_{\sigma_j,l} \ . \tag{3}$$

In the presence of noise there is a finite probability of having configurations other than the local minima. This can be taken into account by introducing an effective temperature $T=1/\beta$.

To build the capacity for learning and memory in this network, its stationary configurations representing the retrieved patterns must be correlated with the stored patterns $\{\xi^\mu\}$, $\mu=1,2\ldots,p$, fixed by the learning process. These patterns are chosen to be independent random variables. We allow $\tilde{p}$ of these patterns to be biased, i.e., those $\{\xi^\mu\}$ can take the values $1,2,\ldots,Q$ with probability

$$P(\xi_i^\mu=k)=\frac{1+B_k}{Q} \ , \qquad k=1,2,\ldots,Q, \quad \mu=1,2,\ldots,\tilde{p} \tag{4}$$

where the $\{B_k\}$ are the bias parameters. Since the $P(k)$ are probabilities, the $\{B_k\}$ satisfy

$$-1\le B_k\le Q-1, \quad \sum_{k=1}^{Q}B_k=0 \ . \tag{5}$$

The other $(p-\tilde{p})$ patterns are unbiased.

This model, with $\tilde{p}=p$, has been studied in [1,7] with the learning rule

$$J_{ij}^{kl}=\frac{1}{Q^2N}\sum_{\mu=1}^{p}(u_{\xi_i^\mu,k}-B_k)(u_{\xi_j^\mu,l}-B_l) \ . \tag{6}$$

In particular the Mattis retrieval states and the lowest symmetric states have been analyzed.

In the following we study the possible relevance of the mixture states in Potts-glass networks with different learning rules for the retrieval of correlated patterns in the sense described above.

III. LOW LOADING OF BIASED PATTERNS

A. Mean-field theory

We consider the storage of a finite number $\tilde{p}$ of biased patterns with the following learning rules:

$$J_{ij}^{kl}=\frac{1}{Q^2N}\sum_{\mu=1}^{\tilde{p}}\left[\left[u_{\xi_i^\mu,k}-\frac{v}{\tilde{p}-1}\sum_{\substack{\nu=1\\ \nu\neq\mu}}^{\tilde{p}}u_{\xi_i^\nu,k}\right]\times\left[u_{\xi_{j,l}^\mu}-\frac{w}{\tilde{p}-1}\sum_{\substack{\nu=1\\ \nu\neq\mu}}^{\tilde{p}}u_{\xi_{j,l}^\nu}\right]\right] . \tag{7}$$

For $v=w=0$ we find back the Hebb learning rule. For $v=0,w\neq0$ or $v\neq0,w=0$ the learning rule mixes in all the other patterns by subtracting in one of the terms the average of the Potts operator over these patterns. Both rules we get in this way are identical. For $v,w\neq0$ this average is subtracted in both terms.

First, we remark that in the limit $\tilde{p}\to\infty$, this subtraction becomes equal to the bias B_k and B_l such that the learning rule (7) for $v=w=1$ becomes the rule (6). Second, the following scaling relation is valid:

$$J_{ij}^{kl}(v=1,w=1)=\frac{\tilde{p}}{\tilde{p}-1}J_{ij}^{kl}(v=1,w=0) \ . \tag{8}$$

So in the following the rules for $v=w=1$ and $v=1,w=0$ can be treated together.

For these interactions we calculate the free energy using the method in [8]. The result is

$$f=\tfrac{1}{2}\sum_{\mu=1}^{\tilde{p}}\left[\left[m_\mu-\frac{v}{\tilde{p}-1}\sum_{\substack{\nu=1\\ \nu\neq\mu}}^{\tilde{p}}m_\nu\right]\left[m_\mu-\frac{w}{\tilde{p}-1}\sum_{\substack{\nu=1\\ \nu\neq\mu}}^{\tilde{p}}m_\nu\right]\right]-\frac{1}{\beta}\left\langle\!\left\langle \ln\frac{1}{Q}\sum_{\sigma=1}^{Q}\exp[\beta\mathcal{H}_\sigma(\xi)]\right\rangle\!\right\rangle \tag{9}$$

with $\mathcal{H}_\sigma(\xi)$ given by

$$\mathcal{H}_\sigma(\xi)=\sum_{\nu=1}^{\tilde{p}}u_{\xi^\nu,\sigma}\left[m_\nu-\frac{v+w}{\tilde{p}-1}\sum_{\substack{\gamma=1\\ \gamma\neq\nu}}^{\tilde{p}}m_\gamma+\frac{vw}{(\tilde{p}-1)^2}\sum_{\substack{\gamma=1\\ \gamma\neq\nu}}^{\tilde{p}}\sum_{\substack{\mu=1\\ \mu\neq\gamma}}^{\tilde{p}}m_\mu+h_\nu\right] \tag{10}$$

and h_ν the couplings of the external field terms. Here $\langle\!\langle\ \rangle\!\rangle$ denotes the average over the $\tilde{p}$ biased patterns.

The order parameter m_ν representing the overlap with the νth pattern, i.e.,

$$m_\nu=\frac{1}{N}\sum_{i=1}^{N}\langle\!\langle\langle u_{\xi_i^\nu,\sigma_i}\rangle\rangle\!\rangle \ , \tag{11}$$

where $\langle\ \rangle$ stands for the thermal average, satisfies the fixed-point equation

$$m_\nu=\left\langle\!\left\langle\frac{\sum_\sigma u_{\xi^\nu,\sigma}\exp[\beta\mathcal{H}_\sigma(\xi)]}{\sum_\sigma\exp[\beta\mathcal{H}_\sigma(\xi)]}\right\rangle\!\right\rangle \ . \tag{12}$$

These fixed-point equations allow for the following types of solution. In the case $v=w=0$ there exist a solution $\mathbf{m}=\mathbf{0}$ representing the disordered state. There are

no Mattis solutions $\mathbf{m}=(m,0,\ldots,0)$ and no symmetric solutions of the form $\mathbf{m}=(m_n,\ldots,m_n,0,\ldots,0)$, $n<\tilde{p}$. However, symmetric solution of the form $\mathbf{m}=(m_{\tilde{p}},\ldots,m_{\tilde{p}})$ do exist. They are unwanted because they represent confusion of the neural network. Finally there exist asymmetric solutions, some of which have retrieval properties, i.e., $\mathbf{m}=(m_1,m_{\tilde{p}-1},\ldots,m_{\tilde{p}-1})$ with $m_1 \gg m_{\tilde{p}-1}$. In the following we only discuss the retrieval-like asymmetric solutions.

For $v=w=1$, the only solutions that are present, compared with the Hebb rule, are the $\mathbf{m}=\mathbf{0}$ solution and the asymmetric retrieval-like solutions. This is also true for $v=1, w=0$ and $v=0, w=1$.

B. Results for the Hebb rule ($v=w=0$)

For general Q a signal-to-noise ratio analysis indicates that retrieval may be possible if

$$\frac{1}{Q}\sum_k B_k^2 \le \frac{Q-1}{\tilde{p}-1} . \tag{13}$$

The $\mathbf{m}=\mathbf{0}$ solution exists for all temperatures and it becomes stable above

$$T=Q-1+(\tilde{p}-1)\frac{1}{Q}\sum_{k=1}^{Q} B_k^2 . \tag{14}$$

For the symmetric solutions, an expansion of the fixed-point equation (12) in $m_{\tilde{p}}$ leads to

$$\begin{aligned} m_{\tilde{p}} = {} & \beta\left[Q-1+(\tilde{p}-1)\frac{1}{Q}\sum_{k=1}^{Q} B_k^2\right] m_{\tilde{p}} \\ & + \tfrac{1}{2}\beta^2\left[(Q-1)(Q-2)+3(\tilde{p}-1)(Q-2)\frac{1}{Q}\sum_{k=1}^{Q} B_k^2 \right. \\ & \left. + (\tilde{p}-1)(\tilde{p}-2)\frac{1}{Q}\sum_{k=1}^{Q} B_k^3\right] m_{\tilde{p}}^2 + O(m_{\tilde{p}}^3) . \end{aligned} \tag{15}$$

This tells us that the critical temperature is given by (14). Furthermore an inspection of the coefficients in m^2 (and m^3 for $Q=2$) shows that the transition to the $\mathbf{m}=\mathbf{0}$ solution is first order for $Q\ge 3$ and second order for $Q=2$. The latter is in agreement with the results of [4].

At $T=0$ the value of $m_{\tilde{p}}$ is given by

$$m_{\tilde{p}} = \frac{1}{\tilde{p}}\left\langle\!\left\langle \max_k \sum_{\nu=1}^{\tilde{p}} u_{\xi^\nu,k} \right\rangle\!\right\rangle . \tag{16}$$

This result is in agreement with [4] ($Q=2$) and [7] (zero bias).

The study of the stability properties requires a detailed investigation of the eigenvalues of the stability matrix. This has been done analogously to [7]. The analysis depends on the number of biased patterns and on the specific bias structure. Results will be described for the specific $Q=3$ models at the end of this subsection.

Concerning the asymmetric solutions, we have restricted ourselves to the form $\mathbf{m}=(m_1,m_{\tilde{p}-1},\ldots,m_{\tilde{p}-1})$. If we require retrieval-like properties, i.e., $m_1 \gg m_{\tilde{p}-1}$, we find that for $T=0$

$$m_1 = Q-1, \quad m_{\tilde{p}-1} = \frac{1}{Q}\sum_k B_k^2 , \tag{17}$$

provided the condition (13) is satisfied. This solution represents perfect overlap with one of the patterns. Also the stability properties of these asymmetric solutions have been studied. We present more detailed results for the specific $Q=3$ models introduced next.

As an illustrative example a $Q=3$ network is worked out for two representative classes of bias types, i.e., $\mathbf{B}_1=a(2,-1,-1)$ and $\mathbf{B}_2=a(1,0,-1)$ with $a\in[0,1]$. The first bias type indicates that one state is privileged and the other two states have equal probability to appear. In fact, for $a=1$ the probability distribution for the patterns is such that the lowest state has probability one. This means that there is no freedom left for the neuron. In the other case all three states have different probability.

In Figs. 1 and 2 we present a T-a diagram for $\tilde{p}=2,3,4$ biased patterns in the $\mathbf{B}_1$ and $\mathbf{B}_2$ model. For all T and a there exist a $\mathbf{m}=\mathbf{0}$ solution which is stable above the curve given by (14) for $Q=3$. Between the dashed and the dotted line the symmetric solutions exist and are stable. Below the dotted line they exist, but they are unstable. Below the solid line the retrieval-like asymmetric solutions exist and are stable. We remark that for $a=0$ the temperature below which these states are stable is $T=2.18$, exactly as for the Mattis retrieval states in the unbiased Potts model [9]. For increasing $\tilde{p}$, the stability region for the symmetric solutions becomes larger and the stability region for retrieval-like asymmetric solutions becomes smaller. We remark that at $T=0$ retrieval is only possible for certain values of the bias amplitude a in agreement with condition (13).

In view of these results the $\mathbf{B}_2$ model has better retrieval properties than the $\mathbf{B}_1$ model. Comparing with

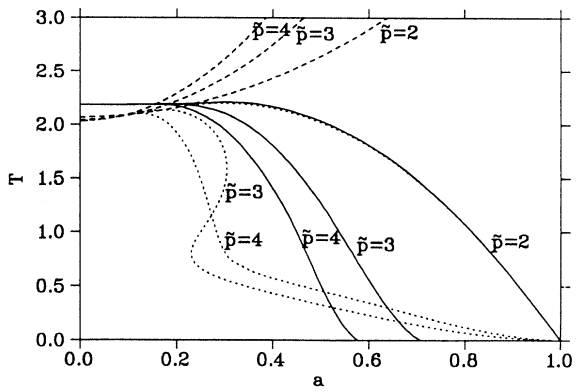

FIG. 1. The T-a diagram for the $Q=3$ $\mathbf{B}_1$ network with $v=w=0$ and $\tilde{p}=2,3,4$ biased patterns for $\alpha=0$. The dashed and the dotted (solid) curves concern the existence and stability of the symmetric (asymmetric) solutions.

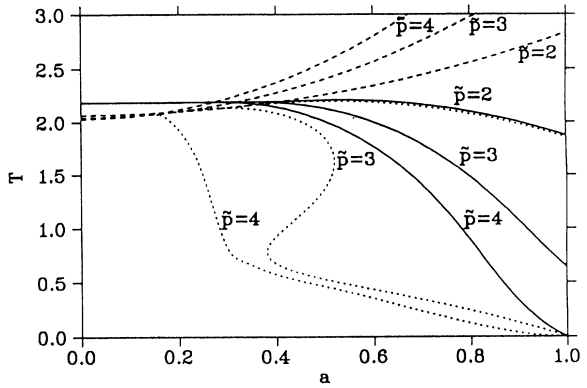

FIG. 2. Same as Fig. 1, but for the $Q=3$ $\mathbf{B}_2$ network.

the Hopfield model (see [4], Figs. 1 and 2) we find no difference between the stability properties of even and odd symmetric states: the symmetric states are never stable in the low-temperature regime.

C. Results for the generalized Hebb rule ($v=w=1$)

For general Q a signal-to-noise ratio analysis gives

$$\frac{1}{Q}\sum_k B_k^2 \leq Q-1 \tag{18}$$

and hence there is no extra condition on the bias parameters besides (5).

The $\mathbf{m}=0$ solution exists for all temperatures and it becomes stable above

$$T=\gamma\frac{\tilde{p}}{\tilde{p}-1}\left[(Q-1)-\frac{1}{Q}\sum_{k=1}^{Q}B_k^2\right]. \tag{19}$$

Here $\gamma=\tilde{p}/(\tilde{p}-1)$. For the $v=1, w=0$ model we have to take $\gamma=1$ according to the scaling property (8) of the learning rules.

As mentioned before there are no symmetric solutions. For the retrieval-like asymmetric solutions an expansion of the fixed-point equation (12) written in terms of $M=m_1-m_{\tilde{p}-1}$ gives

$$M=\beta\gamma\frac{\tilde{p}}{\tilde{p}-1}\left[Q-1-\frac{1}{Q}\sum_{k=1}^{Q}B_k^2\right]M$$
$$+\tfrac{1}{2}\beta^2\gamma^2\frac{\tilde{p}(\tilde{p}-2)}{(\tilde{p}-1)^2}\left[(Q-1)(Q-2)+3(Q-2)\frac{1}{Q}\sum_{k=1}^{Q}B_k^2+\frac{2}{(\tilde{p}-1)}\frac{1}{Q}\sum_{k=1}^{Q}B_k^3\right]M^2+O(M^3). \tag{20}$$

From (20) we find a critical temperature given by (19). Furthermore the order of the transition to the $\mathbf{m}=0$ solution not only depends on Q but also on $\tilde{p}$. For $\tilde{p}=2$ the coefficient of M^2 is zero such that we have to look at the coefficient of M^3. This coefficient is given by $\frac{1}{3}\beta^3\gamma^3Q^2(Q-6)(Q-1-\sum_k B_k^2/Q)$. Hence the transition is second order for $Q\leq 6$ (for $Q=6$ an inspection of the M^4 and M^5 terms has been done) and first order for $Q>6$. For $\tilde{p}>2$ this transition is second order for $Q=2$ and first order for $Q\geq 3$.

At $T=0$ the retrieval-like solution is explicitly given by

$$m_1=(Q-1)\left[1-\sum_k\left[\frac{1+B_k}{Q}\right]^{\tilde{p}}\right],$$
$$m_{\tilde{p}-1}=\frac{1}{Q}\sum_k B_k^2-(Q-1)\sum_k\left[\frac{1+B_k}{Q}\right]^{\tilde{p}}. \tag{21}$$

This does not represent perfect retrieval, but for growing $\tilde{p}$ the value of m_1 tends to $(Q-1)$ and the value of $m_{\tilde{p}-1}$ tends to $\sum_k B_k^2/Q$. We remark that these values do not depend on the scaling parameter γ.

We now turn to the specific $Q=3$ models discussed before. Figures 3 and 4 show a T-a diagram for $\tilde{p}=3,6,9$ patterns. Below the solid lines the retrieval-like asymmetric states exist and are stable. For increasing $\tilde{p}$ this curve tends to the dashed curve, i.e., the curve below which the Mattis states of the biased Potts model with learning rule (6) are stable [7]. We remark that there is an instability region for the latter (dotted curve in Fig. 3) as well as for the $\mathbf{B}_1$ model discussed here (dash-dotted curve in Fig. 3). This region for the $\mathbf{B}_1$ model grows with increasing $\tilde{p}$ and is only visible on the scale of Fig. 3 from

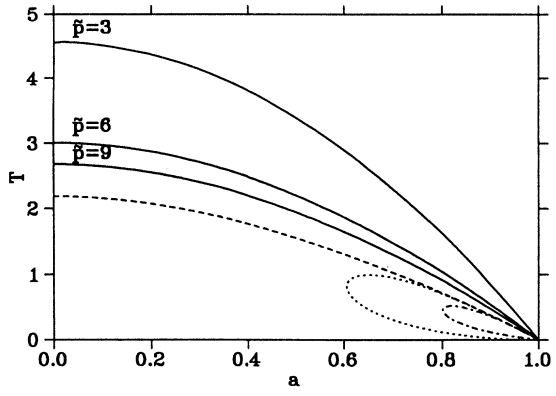

FIG. 3. The T-a diagram for the $Q=3$ $\mathbf{B}_1$ network with $v=w=1$ and $\tilde{p}=3,6,9$ biased patterns for $\alpha=0$. The solid and dash-dotted (dashed and dotted) lines represent the stability region for the asymmetric solutions (Mattis solutions for the Potts model [7]).

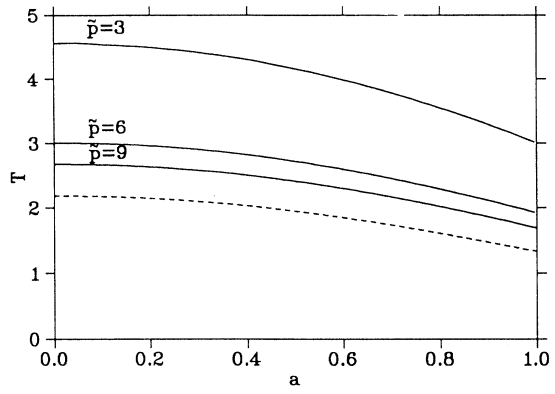

FIG. 4. Same as Fig. 3, but for the $Q=3$ $\mathbf{B}_2$ network.

$\tilde{p}=9$ onwards. Furthermore for $a=1$ the temperature stays different from zero for the $\mathbf{B}_2$ model. Finally the retrieval stability region for $\tilde{p}=2$ is bounded by (19), which leads to similar curves as shown in these figures, but starting at $T=4\gamma, a=0$ and ending at $T=0, a=1$ for the $\mathbf{B}_1$ model or $T=8\gamma/3, a=1$ for the $\mathbf{B}_2$ model.

In comparison with the Hebb learning rule ($v=w=0$) we not only have that there are no symmetric solutions but also the retrieval region is much larger. The latter is also true when comparing with the Hopfield model where the temperature lines are given by [see (19)]

$$T=\gamma\frac{\tilde{p}}{\tilde{p}-1}(1-a^2) . \tag{22}$$

IV. EXTENSIVE LOADING OF BIASED AND UNBIASED PATTERNS

A. Replica symmetric mean-field theory

In this section we consider an extensive loading of p patterns $\{\xi^\mu\}$, $\mu=1,\ldots,p$, of which a finite number $\tilde{p}$ are biased and the rest $p-\tilde{p}$ is unbiased. We are only interested in the condensation of the biased patterns.

The synaptic couplings are given by the learning rule

$$J_{ij}^{kl}=\frac{1}{Q^2N}\sum_{\mu=1}^{\tilde{p}}\left[\left[u_{\xi_i^\mu,k}-\frac{v}{\tilde{p}-1}\sum_{\substack{\nu=1\\ \nu\neq\mu}}^{\tilde{p}}u_{\xi_i^\nu,k}\right]\times\left[u_{\xi_j^\mu,l}-\frac{w}{\tilde{p}-1}\sum_{\substack{\nu=1\\ \nu\neq\mu}}^{\tilde{p}}u_{\xi_j^\nu,l}\right]\right]+\frac{1}{Q^2N}\sum_{\mu=\tilde{p}+1}^{p}u_{\xi_i^\mu,k}u_{\xi_j^\mu,l} . \tag{23}$$

In comparison with (7) the biased patterns are stored as before and the unbiased patterns are stored with the Hebb rule. We note that the remarks about the possible values of v and w after (7) remain valid. The scaling relation (9) no longer holds.

We introduce the following order parameters:

$$m_\nu=\frac{1}{N}\sum_{i=1}^{N}\langle\langle\langle u_{\xi_i^\nu,\sigma_i}\rangle\rangle\rangle , \tag{24}$$

which is the macroscopic overlap with a condensed pattern, and

$$q=\frac{1}{N}\sum_{i=1}^{N}\left\langle\left\langle\sum_{k=1}^{Q}\frac{1}{Q}\langle u_{k,\sigma_i}\rangle\langle u_{k,\sigma_i}\rangle\right\rangle\right\rangle , \tag{25}$$

which is the Edwards-Anderson order parameter. Using then mean-field theory and the replica-symmetric approximation one finds for the free energy, following [7] generalized to the Potts model,

$$f=\tfrac{1}{2}\sum_{\mu=1}^{\tilde{p}}\left[\left[m_\mu-\frac{v}{\tilde{p}-1}\sum_{\substack{\nu=1\\ \nu\neq\mu}}^{\tilde{p}}m_\nu\right]\left[m_\mu-\frac{w}{\tilde{p}-1}\sum_{\substack{\nu=1\\ \nu\neq\mu}}^{\tilde{p}}m_\nu\right]\right]+\tfrac{1}{2}\alpha(Q-1)-\tfrac{1}{2}\beta q^2 r-\frac{\alpha q}{2[1-\beta(Q-1-q)]}$$

$$+\frac{\alpha}{2\beta}\ln[1-\beta(Q-1-q)]+\tfrac{1}{2}\beta qr(Q-1)-\frac{1}{\beta}\left\langle\left\langle\int_{\mathbf{R}^Q}D\mathbf{z}\ln\frac{1}{Q}\sum_{\sigma=1}^{Q}\exp[\beta\mathcal{H}_\sigma(\xi,\mathbf{z})]\right\rangle\right\rangle , \tag{26}$$

with the Gaussian measure $D\mathbf{z}$ given by

$$D\mathbf{z}=\prod_{k=1}^{Q}dz_k(2\pi)^{-1/2}\exp(-z_k^2/2) . \tag{27}$$

Here $\mathcal{H}_\sigma(\xi,\mathbf{z})$ reads

$$\mathcal{H}_\sigma(\xi,\mathbf{z})=\sqrt{qr/Q}\sum_{l=1}^{Q}u_{l,\sigma}z_l+\sum_{\nu=1}^{\tilde{p}}u_{\xi^\nu,\sigma}\left[m_\nu-\frac{v+w}{\tilde{p}-1}\sum_{\substack{\gamma=1\\ \gamma\neq\nu}}^{\tilde{p}}m_\gamma+\frac{vw}{(\tilde{p}-1)^2}\sum_{\substack{\gamma=1\\ \gamma\neq\nu}}^{\tilde{p}}\sum_{\substack{\mu=1\\ \mu\neq\gamma}}^{\tilde{p}}m_\mu+h_\nu\right] . \tag{28}$$

Furthermore r is given by

$$r=\frac{\alpha}{[1-\beta(Q-1-q)]^2} . \tag{29}$$

It can be considered as the total mean-square random overlap with the noncondensed patterns, i.e.,

$$r=\frac{1}{\alpha}\sum_{\mu=\tilde{p}+1}^{p}\langle\langle m_\mu^2\rangle\rangle \tag{30}$$

with α the storage capacity, i.e., $\alpha=p/N$. Since we only consider condensation of the $\tilde{p}$ biased patterns the $\langle\langle\ \rangle\rangle$ in (26) denote the average over these biased patterns only.

The order parameters satisfy the following fixed-point equations:

$$m_\nu=\left\langle\!\!\left\langle \int_{\mathbf{R}^Q} D\mathbf{z}\frac{\sum_\sigma u_{\xi^\nu,\sigma}\exp[\beta\mathcal{H}_\sigma(\xi,\mathbf{z})]}{\sum_\sigma\exp[\beta\mathcal{H}_\sigma(\xi,\mathbf{z})]}\right\rangle\!\!\right\rangle , \tag{31}$$

$$q=\frac{1}{Q}\sum_{k=1}^{Q}\left\langle\!\!\left\langle \int_{\mathbf{R}^Q} D\mathbf{z}\left[\frac{\sum_\sigma u_{k,\sigma}\exp[\beta\mathcal{H}_\sigma(\xi,\mathbf{z})]}{\sum_\sigma\exp[\beta\mathcal{H}_\sigma(\xi,\mathbf{z})]}\right]^2\right\rangle\!\!\right\rangle , \tag{32}$$

together with (28) and (29). In the case $v=w=0$, these fixed-point equations have the following types of solution: asymmetric solutions $\mathbf{m}=(m_1,m_{\tilde{p}-1},\ldots,m_{\tilde{p}-1},0,\ldots,0)$, $q\neq 0$, some of which have retrieval properties $(m_1\gg m_{\tilde{p}-1})$; symmetric solutions $\mathbf{m}=(m_{\tilde{p}},\ldots,m_{\tilde{p}},0,\ldots,0)$, $q\neq 0$, representing confusion in the network; spin-glass solutions $\mathbf{m}=0, q\neq 0$ and a paramagnetic solution $\mathbf{m}=\mathbf{0}, q=0$. There are no Mattis solutions $\mathbf{m}=(m,0,\ldots,0)$ and no symmetric solutions of the form $\mathbf{m}=(m_n,\ldots,m_n,0,\ldots,0)$ with $n<\tilde{p}$.

For $v=w=1$, the only solutions that are present, compared with the Hebb rule, are the paramagnetic, the spin-glass, and the asymmetric retrieval-like solutions. This is also true for $v=1, w=0$ and $v=0, w=1$.

B. Results for the Hebb rule ($v=w=0$)

In this subsection we restrict ourselves to a study of a $Q=3$ network for the two classes of bias types $\mathbf{B}_1=a(2,-1,-1)$ and $\mathbf{B}_2=a(1,0,-1)$, introduced in Sec. III B.

Figures 5 and 6 show the storage capacity α, at $T=0$, for both models with $\tilde{p}=2,3,4$ biased and p-$\tilde{p}$ unbiased patterns as a function of the bias amplitude a. There always exist spin-glass solutions. Besides there exist symmetric solutions below the dashed line and there exist asymmetric retrieval-like solutions below the solid line. All transitions at these lines are of first order. We note that at $a=0$ the critical storage capacity for the asymmetric retrieval states is $\alpha_c=0.415$, such as for the Mattis retrieval states in the unbiased Potts model [6]. These figures show that the retrieval region becomes smaller and that the region where the symmetric solutions exist grows with increasing $\tilde{p}$. Furthermore at $\alpha=0$ retrieval is only possible for certain values of the bias amplitude a in agreement with condition (13).

Comparing both the $\mathbf{B}_1$ and the $\mathbf{B}_2$ case one sees that the storage capacity of the asymmetric states is always the largest for the $\mathbf{B}_2$ model ($a\neq 0$). Furthermore the storage capacity of the symmetric states grows slower, as a function of a, for the $\mathbf{B}_2$ than for the $\mathbf{B}_1$ model. To get a more specific idea we list these storage capacities at $a=1$: for the $\mathbf{B}_1$ model $\alpha=2.30$ ($\tilde{p}=2$), $\alpha=6.40$ ($\tilde{p}=3$), and $\alpha=12.91$ ($\tilde{p}=4$); for the $\mathbf{B}_2$ model $\alpha=0.38$ ($\tilde{p}=2$), $\alpha=0.80$ ($\tilde{p}=3$), and $\alpha=1.43$ ($\tilde{p}=4$).

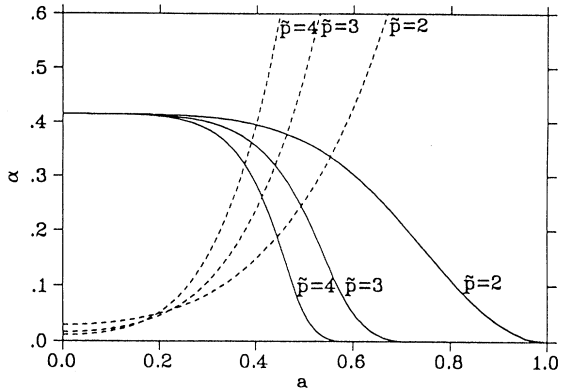

FIG. 5. The α-a diagram for the $Q=3$ $\mathbf{B}_1$ network with $v=w=0$ and $\tilde{p}=2,3,4$ biased patterns at $T=0$. The solid (dashed) curve concerns the existence of the asymmetric (symmetric) solutions.

For $T\neq 0$, the T-α phase diagrams for the $Q=3$ $\mathbf{B}_1$ and $\mathbf{B}_2$ models with $\tilde{p}=3$ at $a=0.4$ are shown in Figs. 7 and 8. The line T_g (dash-dotted line) indicates the transition from the spin-glass solution to the disordered paramagnetic state. This line can be calculated analytically for general Q: $T_g=Q-1+\sqrt{\alpha(Q-1)}$. For $Q\leq 6$ this transition is always second order, but for $Q>6$ we find that if $\alpha<\alpha_0=16(Q-1)(Q-6)^{-2}$ the transition is second order, while for $\alpha>\alpha_0$ the transition is first order [10]. Returning to $Q=3$ we find that below T_S (dotted line) symmetric solutions show up as local minima of the free energy and below T_A (solid line) the asymmetric retrieval solutions appear as local minima of the free energy. Below T_c (dashed line) the asymmetric retrieval solutions become global minima of the free energy. The transitions at the T_S, T_c, and T_A lines are first-order transitions

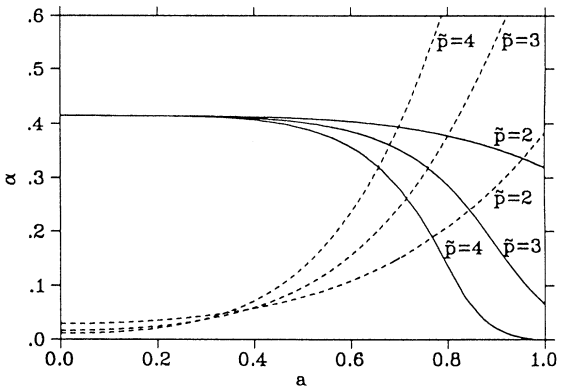

FIG. 6. Same as Fig. 5, but for the $Q=3$ $\mathbf{B}_2$ network.

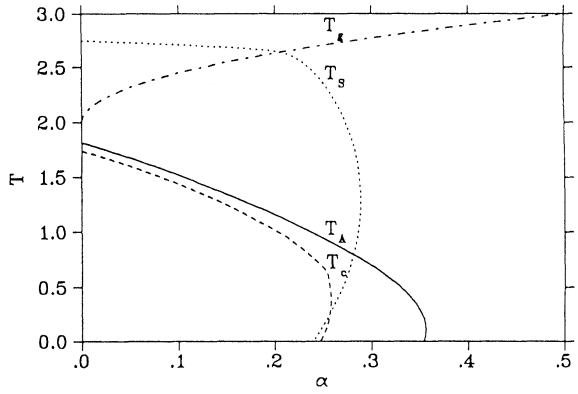

FIG. 7. The T-α diagram for the $Q=3$ $\mathbf{B}_1$ network with $v=w=0$ and $\tilde{p}=3$ biased patterns and p-$\tilde{p}$ unbiased patterns at $a=0.4$. The meaning of the curves is explained in the text.

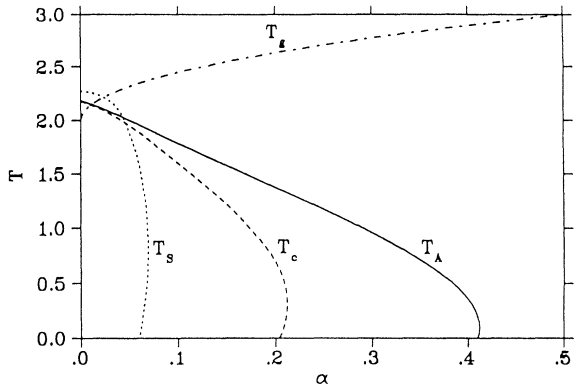

FIG. 8. Same as Fig. 7, but for the $Q=3$ $\mathbf{B}_2$ network.

while for the analog $Q=2$ problem the transitions from both the symmetric and the spin-glass solutions to the paramagnetic solution are second order [5]. For decreasing a the retrieval region becomes larger and the region for existence of the symmetric states shrinks. For $a=0$ we find back the phase diagram for the unbiased Potts model [10] with the line T_A now indicating the stability of the Mattis retrieval states.

Finally we remark that the T_S, T_c, and T_A transition curves have a negative slope at low temperatures suggesting the breaking of the replica symmetry approximation (see also [10] for the Potts model with learning rule (6) and [11] for the Hopfield model). The difference between the maximal value of α and its value at $T=0$ is the largest for the symmetric solutions. To get a more quantitative idea about this breaking we have calculated the entropy at $T=0$. For the $\mathbf{B}_1$ model we find at $\alpha=\alpha_c$ for the symmetric states $S_{\text{sym}}=-0.72\times10^{-1}$, while for the asymmetric states we have $S_{\text{asym}}=-0.46\times10^{-2}$. For the $\mathbf{B}_2$ model the corresponding values are $S_{\text{sym}}=-0.46\times10^{-1}, S_{\text{asym}}=-0.41\times10^{-2}$. This suggests a breaking that is stronger for the symmetric solutions.

C. Results for the generalized Hebb rule ($v=w=1$)

For general Q the fixed-point equations (31), and (32) for the asymmetric retrieval-like solutions can be written in terms of $M=m_1-m_{\tilde{p}-1}$ and q. At $T=0$, they can be further reduced to only one fixed-point equation in $y=\gamma M\sqrt{Q/qr}$. For $\tilde{p}=2$ biased patterns an expansion of this equation for small values of y learns that the transition from the asymmetric retrieval phase to the spin-glass phase is second order for $Q\leq 7$ and first order for $Q>7$. Furthermore for $Q\leq 7$, this transition occurs at

$$\alpha=\frac{Q^3}{2\pi(Q-1)}\left[2\gamma\left[Q-1-\frac{1}{Q}\sum_k B_k^2\right]\frac{1}{\sqrt{2\pi}}\times\int_{-\infty}^{\infty}dz\,e^{-z^2}E(z)^{Q-2}-\int_{-\infty}^{\infty}dz\,e^{-(1/2)z^2}zE(z)^{Q-1}\right]^2 \tag{33}$$

with $E(z)=\frac{1}{2}[1+\text{erf}(z/\sqrt{2})]$, provided that the term in between the square brackets is positive. Here we recall that $\gamma=\tilde{p}/(\tilde{p}-1)$, for the $v=1, w=0$ model $\gamma=1$. This implies the following conditions on the bias parameters

$$\frac{1}{Q}\sum_k B_k^2\leq Q-1-\sqrt{\pi/2}\,\frac{\int_{-\infty}^{\infty}dz\,e^{-(1/2)z^2}zE(z)^{Q-1}}{\gamma\int_{-\infty}^{\infty}dz\,e^{-z^2}E(z)^{Q-2}}. \tag{34}$$

Let us now turn to the specific $Q=3$ $\mathbf{B}_1$ and $\mathbf{B}_2$ model. At $T=0$, the storage capacity of the asymmetric retrieval states, for $\tilde{p}=2,3,4$ biased and $(p-\tilde{p})$ unbiased patterns in function of the bias amplitude a is shown in Fig. 9. We again remark that we have no symmetric states for this learning rule.

Clearly the capacity diminishes with growing $\tilde{p}$. At $\alpha=0$ we find that retrieval is only possible in a restricted region of the interval $a\in[0,1]$. This is not supported by the signal-to-noise ratio analysis (18) and the T-a diagrams for $\alpha=0$ (Figs. 3 and 4). Whether this effect is due

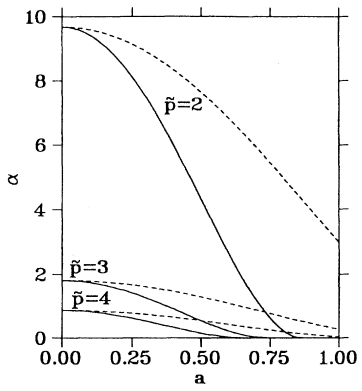

FIG. 9. The α-a diagram for the $Q=3$ $\mathbf{B}_1$ (solid curve) and $\mathbf{B}_2$ (dashed curve) network with $v=w=1$ and $\tilde{p}=2,3,4$ biased patterns at $T=0$.

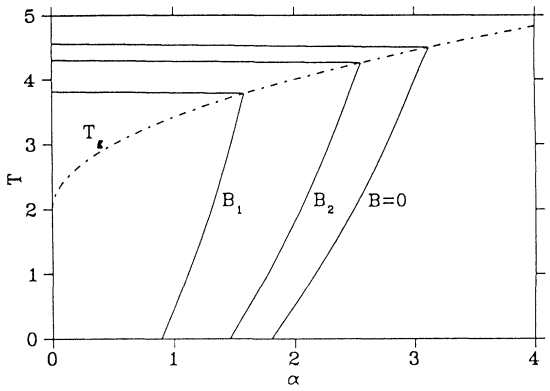

FIG. 10. Same as Figs. 7 and 8, but with $v=w=1$.

to replica symmetry breaking remains to be seen [12].

For finite T, the T-α phase diagram of the $v=w=1$ model with $\tilde{p}=3$ biased and $(p-\tilde{p})$ unbiased patterns is shown in Fig. 10. In the region bounded by the solid lines asymmetric retrieval-like states exist and they are the global minima of the free energy. Below the dash-dotted line spin-glass solutions appear. Remark that this spin-glass line is identical to the one in the $v=w=0$ model. Hence the transition properties along this line are the same as before. The transition from the asymmetric to the paramagnetic solutions at the solid line is first order.

From Fig. 10 we see that introducing bias leads to a smaller retrieval region. However, the shape of the critical lines remains nearly the same. The $\mathbf{B}_2$ model has better retrieval properties. Also here the transition lines have a negative slope at low temperatures and the maximal capacity is reached for $T\neq 0$. This suggests again a replica-symmetry breaking effect. This is substantiated by a calculation of the entropy at $T=0$ which gives at $\alpha=\alpha_c$: $S_{\text{asym}}=-0.1617$ for $a=0$, $S_{\text{asym}}=-0.1507$ for the $\mathbf{B}_1$ model with $a=0.4$, and $S_{\text{asym}}=-0.1604$ for the $\mathbf{B}_2$ model with $a=0.4$. Compared with the $v=w=0$ model we find that the entropy S_{asym} is more negative, suggesting a stronger breaking.

V. CONCLUDING REMARKS

In this paper we have studied symmetric and asymmetric mixture states and the storage of biased patterns in Potts-glass neural networks. In particular, we have discussed the retrieval of a finite number of these patterns for different learning rules in the possible presence of an extensive loading of unbiased patterns with the Hebb rule. Detailed results are presented for $Q=3$ models with two representative classes of bias parameters.

For $\alpha=0$ one concludes that the generalized Hebb learning rule $(v=w=1)$ leads to asymmetric retrieval solutions over the whole range of bias amplitudes a. Further, it leads to no symmetric solutions. Consequently its retrieval properties are better than the model with the usual Hebb rule $(v=w=0)$.

For $\alpha\neq 0$ one sees by comparing the $v=w=1$ model (Fig. 9) with the $v=w=0$ model (Figs. 5 and 6) that the storage capacity is always substantially larger for the network with the generalized Hebb rule $(v=w=1)$. Thus modifying the learning rule by subtracting in each term, for a certain pattern, the average of the Potts neuron state over all the other patterns enhances the storage capacity for the biased patterns. It is also interesting to remark at this point that the effective storage of these patterns with these learning rules (7) and (23) is possible without knowing explicitly the bias parameters B_k.

The learning rule $v=w=1$ leads always to a larger storage capacity than the rule $v=1, w=0$. The latter even gives rise to a smaller storage capacity than the Potts model with learning rule (6). For example, at $a=0$ and for $\tilde{p}=3$ one has $\alpha_c(v=w=1)=1.806$ and $\alpha_c(v=1,w=0)=0.355<\alpha_c(v=w=0)=0.415$. Similar results have been found for $Q=2$. So we conclude that, in contrast to the $v=w=1$ learning rule, the $v=1, w=0$ rule is not able to enhance the storage capacity of the biased patterns. Moreover the $v=w=1$ rule always leads to the largest retrieval region in the T-α plane (compare Figs. 7 and 8 with Fig. 10).

Finally we have checked that for all learning rules considered in this work the storage capacity of the unbiased patterns stays the same as in the Potts model without bias. So we have not damaged the ability of the network to retrieve the other patterns.

ACKNOWLEDGMENTS

This work has been supported in part by the Research Fund of the KU Leuven (Grant No. OT/91/13). The authors are indebted to R. Erichsen, Jr. and L. van Hemmen for stimulating discussions. They would like to thank the Belgian National Fund for Scientific Research and the Inter-University Institute for Nuclear Sciences for financial support.

*Electronic address: FGBDA18@cc1.kuleuven.ac.be

[1] D. Bollé, R. Cools, P. Dupont, and J. Huyghebaert, J. Phys. A **26**, 549 (1993).

[2] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **35**, 2293 (1987).

[3] C. J. Perez-Vicente and D. J. Amit, J. Phys. A **22**, 559 (1989).

[4] J. F. Fontanari and W. K Theumann, J. Phys. (Paris) **51**, 375 (1990).

[5] R. Erichsen, Jr. and W. K. Theumann, Int. J. Neural Syst. **1**, 347 (1991); in *Neural Networks and Spin Glasses,* edited by W. K. Theumann and R. Köberle (World Scientific, Singapore, 1990), p. 204.

[6] I. Kanter, Phys. Rev. A **37**, 2739 (1988).

[7] D. Bollé, P. Dupont, and J. van Mourik, J. Phys. A **24**, 1065 (1991).

[8] J. L. van Hemmen and R. Kühn, in *Models of Neural Networks,* edited by E. Domany, J. L. van Hemmen, and K. Schulten (Springer-Verlag, Berlin, 1991), p. 1.
[9] D. Bollé and F. Mallezie, J. Phys. A **22**, 4409 (1989).
[10] D. Bollé, P. Dupont, and J. Huyghebaert, Phys. Rev. A **45**, 4194 (1992).
[11] J.-P. Naef and A. Canning, J. Phys. (Paris) I **2**, 247 (1992).
[12] D. Bollé and J. Huyghebaert (unpublished).