Categorization in a Hopfield network trained with weighted examples: Extensive number of concepts

Rogério L. Costa and Alba Theumann
Instituto de Física, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves 9500, CP 15051, 91501-970 Porto Alegre, RS, Brazil

(Received 2 August 1999)

We consider the categorization problem in a Hopfield network with an extensive number of concepts $p=\alpha N$ and trained with s examples of weight λ_τ, $\tau=1,\ldots,s$ in the presence of synaptic noise represented by a dimensionless ''temperature'' T. We find that the retrieval capacity of an example with weight λ_1, and the corresponding categorization error, depend also on the arithmetic mean λ_m of the other weights. The categorization process is similar to that in a network trained with Hebb's rule, but for $\lambda_1/\lambda_m>1$ the retrieval phase is enhanced. We present the phase diagram in the T-α plane, together with the de Almeida–Thouless line of instability. The phase diagrams in the α-s plane are discussed in the absence of synaptic noise and several values of the correlation parameter b.

PACS number(s): 64.60.Cn, 75.10.Nr, 75.10.Hk

I. INTRODUCTION

In the seminal work by Amit, Gutfreund, and Sompolinsky [1] on the statistical equilibrium properties of the Hopfield model [2] for attractor neural networks, it was shown that, besides the desired retrieval states of single memories, there appeared undesirable mixture states where several patterns could be recalled. However, in later developments it was shown that mixture states could be very useful in the description of a different mental process, that of categorization. In the categorization process, the network learns several blurred examples of a concept, and there is a competition between the retrieval of an individual example and the state in which the network categorizes by retrieving the hidden concept common to the individual examples [3].

The problem of categorization has been widely investigated in a Hopfield network both in the absence [4] and in the presence [5] of synaptic noise, in layered neural networks both for binary [6] and multistate neurons [7], and analogical networks [8]. Dynamical studies were also performed in very diluted networks for binary [9] and multistate neurons [10]. The effect of gradual dilution and of synaptic noise in the categorization ability of an attractor neural network with hierarchically correlated patterns has been studied recently [11].

In a previous publication [12] we addressed ourselves to the problem of categorization in an attractor neural network trained with weighted examples in the presence of synaptic noise, for a finite numbers of concepts. We found that the retrieval capacity of an example in the competition with categorization depends on the ratio between the weight of the retrieved example and the mean value of the other weights. When this ratio is larger than unity we obtained that the line of first-order transitions between the retrieval and categorization phases ends at a critical point. Prior to us, a generalized Hebb's rule with weighted patterns has been used in several schemes of neural networks [13–17]. In [18] a mechanism was proposed to enhance the retrieval of a finite subset of ''marked'' patterns by increasing their relative weights. When the weight of the marked subset is sufficiently large, the discontinuous transition from retrieval to spin glass phase becomes continuous at a tricritical point. When the number of marked patterns is extensive, the transition is always discontinuous [19]. We refer the reader to our previous paper [12] for more details on the justification of the model, and we concentrate here on the extensions of the work to the case of an extensive number of concepts.

When we are in the presence of a number $p=\alpha N$ of concepts, they may interfere forming a spin glass phase that competes with the processes of retrieval and categorization, and it is the purpose of this paper to present our results and phase diagram when $\alpha\neq 0$. We obtained that the retrieval phase is never a global minimum of the free energy, although the region in which the retrieval of a particular example is a relative minimum increases with its weight.

The outline of the paper is as follows. We present the calculation of the free energy and saddle point equations within a replica symmetric calculation in Sec. II, together with the results for the Almeida-Thouless instability line. We leave Sec. III for conclusions, and the technical details of the calculation are presented in the Appendix.

II. MODEL

We study in this paper the categorization ability of a Hopfield network with N neurons $S_i=\pm 1$, trained with s weighted examples of a given concept, for an extensive number $p=\alpha N$ of concepts. By introducing the dimensionless temperature $T=1/\beta$ as a measure of synaptic noise, the categorization and retrieval overlaps can be obtained from the free energy associated to the Hamiltonian [12]

$$H=-\frac{1}{2}\sum_{i,j} J_{ij}S_iS_j+\sum_{\mu=1}^{p} h_\mu\sum_i S_i\xi_i^\mu, \tag{1}$$

where the synapses J_{ij} are given by the weighted Hebb's rule:

$$J_{ij}=\frac{1}{N}\sum_{\mu}^{p}\sum_{\tau}^{s}\lambda_{\tau}\xi_i^{\mu\tau}\xi_j^{\mu\tau} \tag{2}$$

and the h_μ are auxiliary fields. The examples $\xi_i^{\mu\tau}$ are binary random variables that may or may not be aligned to the concept $\{\xi_i^\mu\}$ with probability

$$P(\xi_i^{\mu\tau}|\xi_i^\mu)=\frac{1}{2}(1+b)\delta(\xi_i^{\mu\tau}-\xi_i^\mu)+\frac{1}{2}(1-b)\delta(\xi_i^{\mu\tau}+\xi_i^\mu). \tag{3}$$

The p concepts $\{\xi_i^\mu\}$ are uncorrelated and they may be equal to ± 1 with equal probability.

The quantities of interest are the categorization overlap

$$m_\mu=\frac{1}{N}\sum_i S_i\xi_i^\mu \tag{4}$$

and the related categorization error

$$\epsilon_\mu=\frac{1-m_\mu}{2}, \tag{5}$$

together with the retrieval overlap

$$m_{\mu\tau}=\frac{1}{N}\sum_i S_i\xi_i^{\mu\tau}. \tag{6}$$

As we are interested in the categorization process of a given concept, say $\mu=1$, we expect the retrieval and categorization overlaps m_1 and $m_{1\tau}$ in Eq. (4) and Eq. (6) to be $O(1)$, while all the other overlaps for $\mu>1$ would be $O(1/\sqrt{N})$ and contribute to the spin glass phase. Then we write the partition function

$$Z=\int\prod_\tau dm_{1\tau}e^{-(N\beta/2)\sum_\tau m_{1\tau}^2\lambda_\tau}\times\mathrm{Tr}\left\{\exp\left[\frac{\beta}{2N}\sum_{i,j}\sum_{\mu=2}^{p}\sum_{\tau=1}^{s}\lambda_\tau\xi_i^{\mu\tau}\xi_j^{\mu\tau}S_iS_j+\sum_i H_iS_i\right]\right\}, \tag{7}$$

where

$$H_i=h_1\xi_i^1+\sum_\tau m_{1\tau}\lambda_\tau\xi_i^1. \tag{8}$$

Using the replica method we write the free energy per site:

$$\beta F=-\lim_{n\to 0}\frac{1}{Nn}[Z_n-1], \tag{9}$$

where $Z_n=\langle Z^n\rangle$ and $\langle\cdots\rangle$ stands for an average over $\{\xi_i^{\mu\tau}\}$, $\{\xi_i^\mu\}$ in that order. We reserve the detailed calculations for the Appendix and present here only the relevant results. From Eq. (A2) we can write Z_n,

$$Z_n=e^{-(\beta N/2)\sum_{\alpha,\tau}(m_{1\tau}^\alpha)^2}\,\mathrm{Tr}\left\{\langle e^{\sum_{i,\alpha}H_i^\alpha S_i^\alpha}\rangle e^{pn\beta\hat{G}}\right\}, \tag{10}$$

where, from Eq. (A11),

$$n\beta\hat{G}=\frac{1}{2}\sum_{\kappa=2}\frac{1}{\kappa}\beta^\kappa\,\mathrm{Tr}(\underline{\underline{L}}^\kappa)\mathrm{Tr}(\underline{\underline{\hat{q}}}^\kappa) \tag{11}$$

and $\underline{\underline{L}}$ is an $s\times s$ matrix with elements

$$L_{\tau_1\tau_2}=\delta_{\tau_1\tau_2}\lambda_{\tau_2}+(1-\delta_{\tau_1\tau_2})b^2\lambda_{\tau_2} \tag{12}$$

while $\underline{\underline{\hat{q}}}$ is the $n\times n$ overlap matrix with elements

$$\hat{q}_{\alpha_1\alpha_2}=\frac{1}{N}\sum_i S_i^{\alpha_1}S_i^{\alpha_2}. \tag{13}$$

At this point we consider the simplified distribution of weights that was introduced in Ref. [12], $\lambda_\tau=\lambda_m\neq\lambda_1$, for $\tau\geqslant 2$ and we obtain from Eq. (11)

$$n\beta G=-\frac{1}{2}\mathrm{Tr}[1-\beta\omega_+\underline{\underline{\hat{q}}}]-\frac{1}{2}\mathrm{Tr}[1-\beta\omega_-\underline{\underline{\hat{q}}}]-\frac{1}{2}(s-2)\mathrm{Tr}\ln[1-\beta\omega_m\underline{\underline{\hat{q}}}]-\frac{1}{2}[\beta\omega_++\beta\omega_-+(s-2)\beta\omega_m]\mathrm{Tr}\,\underline{\underline{\hat{q}}} \tag{14}$$

with

$$\omega_m=\lambda_m(1-b^2),$$

$$\omega_\pm=\frac{1}{2}\{\lambda_1+\lambda_m[1+(s-2)b^2]\}\pm(\{\lambda_1-\lambda_m[1+(s-2)b^2]\}^2+4(s-1)b^4\lambda_1\lambda_m)^{1/2}. \tag{15}$$

We introduce the order parameters in Eq. (13) by means of the identities

$$1=\int_{-\infty}^{\infty}\prod_{\gamma\neq\delta}dq_{\gamma\delta}\int_{-\imath\infty}^{\imath\infty}\frac{\beta^2\alpha}{2}\frac{dr_{\gamma\delta}}{2\pi\imath}\times\exp\left\{\frac{\alpha\beta^2}{2}\sum_{\gamma\neq\delta}r_{\gamma\delta}\left(\sum_i S_i^\gamma S_i^\delta-Nq_{\gamma\delta}\right)\right\} \tag{16}$$

and we obtain for Z_n at the saddle point

$$\ln Z_n=-N\left\{\frac{\beta}{2}\sum_\gamma[\lambda_1(m_{11}^\gamma)^2+(s-1)\lambda_m(m_{s-1}^\gamma)^2]-\alpha\beta G(q_{\gamma\delta})+\frac{\alpha\beta^2}{2}\sum_{\gamma\neq\delta}r_{\gamma\delta}q_{\gamma\delta}\right\}-N\beta\Lambda, \tag{17}$$

where given the symmetry of the weights we look for a solution:

$$m_{1\tau}=m_{11}\delta_{\tau 1}+m_{s-1}(1-\delta_{\tau 1}) \tag{18}$$

and we call

$$e^{N\beta\Lambda}=\left\langle \mathrm{Tr}\left\{ e^{\sum_{\gamma}\beta H_j^{\gamma}S_j^{\gamma}} e^{(\alpha\beta^2/2)\sum_{\gamma\neq\delta} r_{\gamma\delta}\sum_j S_j^{\gamma}S_j^{\delta}}\right\}\right\rangle_{\{\xi_i^1\},\{\xi_i^{\tau}\}}. \tag{19}$$

Having in mind the calculation of the de Almeida–Thouless [20] instability line, we write the saddle-point solution:

$$\begin{aligned} q_{\gamma\delta}&=q+\eta_{\gamma\delta},\\ r_{\gamma\delta}&=r+\lambda_{\gamma\delta},\\ m_{11}^{\gamma}&=m_{11},\\ m_{s-1}^{\gamma}&=m_{s-1}, \end{aligned} \tag{20}$$

and we obtain to the second order in $\lambda_{\gamma\delta},\eta_{\gamma\delta}$, in the limit $n\to 0$,

$$-\frac{\ln Z_n}{N}=n\beta f_{\mathrm{RS}}+\beta\Delta f_{\mathrm{RSB}}, \tag{21}$$

where the replica symmetric free energy is given by

$$\begin{aligned}\beta f_{\mathrm{RS}}=&\beta\frac{1}{2}[\lambda_1 m_{11}^2+(s-1)\lambda_m m_{s-1}^2]\\ &-\int \mathcal{D}z\langle \ln 2\cosh[\beta h_1\xi^1+\beta m_{11}\lambda_1\xi^{11}+\beta\lambda_m m_{s-1}x\\ &+\beta\sqrt{\alpha r}z]\rangle+\alpha\frac{1}{2}\left\{\beta^2 rC+\sum_{i=+,-,m}a_i\left[\ln(1-\omega_i\beta C)\right.\right.\\ &\left.\left.-q\frac{\beta\omega_i}{1-\beta\omega_i C}+\beta\omega_i\right]\right\}\end{aligned} \tag{22}$$

with

$$\begin{aligned} a_+&=a_-=1,\\ a_m&=(s-2),\\ C&=(1-q). \end{aligned} \tag{23}$$

The average $\langle\ \rangle$ in Eq. (21) is over the examples and the concept, in that order. The variable

$$x=\sum_{\tau=2}^{s}\xi^{1\tau} \tag{24}$$

has a conditional binomial distribution [5,12]

$$\begin{aligned} P(x|\xi^1)&=\binom{s-1}{\kappa}\left[\frac{1}{2}(1+b)\right]^{\kappa}\left[\frac{1}{2}(1-b)\right]^{s-1-\kappa},\\ \kappa&=\frac{1}{2}[\xi^1x+(s-1)]. \end{aligned} \tag{25}$$

The correction Δf_{RSB} due to replica symmetry breaking is the quadratic form obtained in the Appendix, Eq. (A20)

A. Replica symmetric theory

The saddle-point equations for the replica symmetric order parameters q, r, m_{11}, and m_{s-1} are obtained by extremizing f_{RS} in Eq. (22) while m_1 in Eq. (4) is obtained by differentiating f_{RS} with respect to h_1. We obtain, by calling $n_\kappa=2\kappa-(s-1)$,

$$\begin{aligned} m_{11}=&\int_{-\infty}^{\infty}\mathcal{D}z\sum_{\kappa=0}^{s-1}\binom{s-1}{\kappa}\{P(\kappa+1)\tanh(\beta g_+)\\ &+P(\kappa)\tanh(\beta g_-)\}, \end{aligned} \tag{26}$$

$$\begin{aligned} m_{s-1}=&\frac{1}{s-1}\int_{-\infty}^{\infty}\mathcal{D}z\sum_{\kappa=0}^{s-1}\binom{s-1}{\kappa}n_\kappa\{P(\kappa+1)\tanh(\beta g_+)\\ &-P(\kappa)\tanh(\beta g_-)\}, \end{aligned} \tag{27}$$

$$\begin{aligned} m_1=&\int_{-\infty}^{\infty}\mathcal{D}z\sum_{\kappa=0}^{s-1}\binom{s-1}{\kappa}\{P(\kappa+1)\tanh(\beta g_+)\\ &-P(\kappa)\tanh(\beta g_-)\}, \end{aligned} \tag{28}$$

$$\begin{aligned} q=&\int_{-\infty}^{\infty}\mathcal{D}z\sum_{\kappa=0}^{s-1}\binom{s-1}{\kappa}\{P(\kappa+1)\tanh^2(\beta g_+)\\ &+P(\kappa)\tanh^2(\beta g_-)\}, \end{aligned} \tag{29}$$

where

$$g_{\pm}=m_{11}\lambda_1\xi^{11}+z\sqrt{\alpha r}\pm m_{s-1}\lambda_m n_{\kappa}. \tag{30}$$

The equation for r is

$$r=q\sum_{i=+,-,m}a_i\frac{\omega_i^2}{(1-\omega_i\beta C)^2} \tag{31}$$

with the coefficients a_i in Eq. (23)

B. Replica symmetry breaking

By solving for χ in Eq. (A22), we obtain that the lowest eigenvalue vanishes at the instability line,

$$\alpha\beta^2 rK=q, \tag{32}$$

where

$$\begin{aligned} K=&\int_{-\infty}^{\infty}Dz\sum_{\kappa=0}^{s-1}\binom{s-1}{\kappa}\{P(\kappa+1)\mathrm{sech}^4(\beta g_+)\\ &+P(\kappa)\mathrm{sech}^4(\beta g_-)\} \end{aligned} \tag{33}$$

and $g_{\pm}$ are given in Eq. (30).

III. NUMERICAL RESULTS AND DISCUSSIONS

We study in this paper the categorization ability of a Hopfield network trained with s weighted examples of each concept, for a macroscopic number $p=\alpha N$ of concepts, in the presence of synaptic noise represented by a dimensionless ''temperature'' T, thus extending our previous results for finite p [12]. Besides the categorization and retrieval overlap

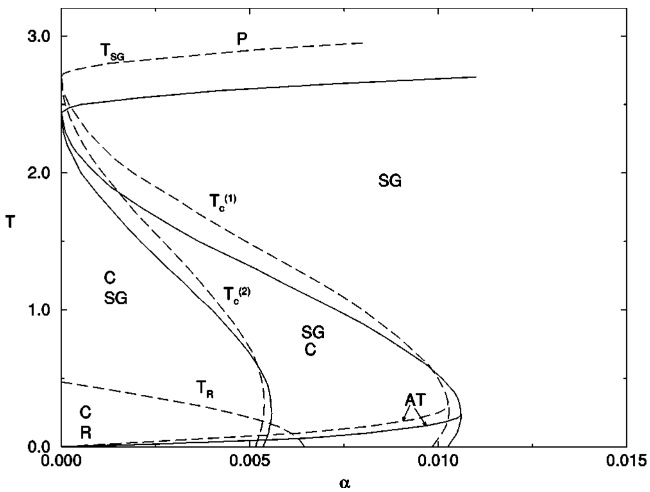

FIG. 1. Phase diagram in the T-α plane, for $b=0.4,s=10$, and values of $\lambda_1/\lambda_m=1.75$ (broken line) and $\lambda_1/\lambda_m=1.0$ (full line). At high temperatures $T>T_{SG}$ the system is in the paramagnetic phase, for $T_c(1)<T<T_{SG}$ there is a spin glass phase, for $T_c(2)<T<T_c(1)$ there appears a categorization solution that is only locally stable, and for $T<T_c(2)$ the categorization phase is globally stable. For $\lambda_1>\lambda_m$ there is also a locally stable retrieval phase when $T<T_R$. We indicate by AT the de Almeida–Thouless instability line. T and α are dimensionless parameters.

in Eq. (14) and Eq. (16), we have to consider now the spin glass overlap matrix in Eq. (13). Having in mind that the network has not been exposed to the underlying concepts but only to imperfect and differently weighted examples of them, we are interested in knowing how the relative weights influence the occurrence of the retrieval, categorization, and spin glass phases. As the concepts are statistically uncorrelated, it is sufficient to analyze the overlaps corresponding to one of

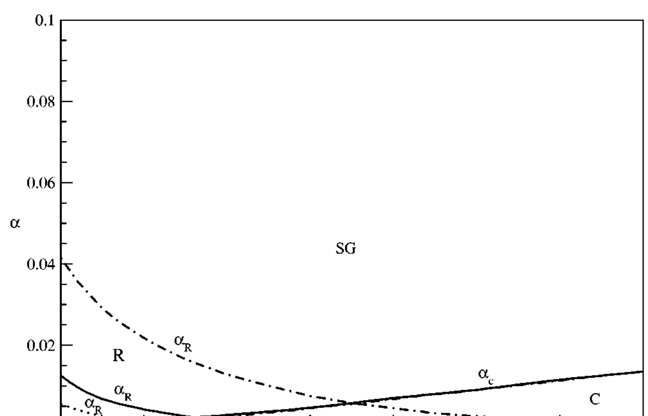

FIG. 2. Phase diagram in the α-s plane in the absence of synaptic noise $(T=0)$ and $b=0.2$ for values $\lambda_1/\lambda_m=0.75$ (pointed line), 1.0 (full line), and 1.75 (broken line). For high values of α the network is in the spin glass phase. For $\alpha<\alpha_R$ the network retrieves and for $\alpha_R<\alpha<\alpha_c$ the network categorizes. For small values of s, when $\alpha_c<\alpha_R$, the retrieval solution is globally stable. The three lines α_c for different λ_1/λ_m coincide for large s, but for small s the line for $\lambda_1/\lambda_m=1.75$ curves down, leaving only the retrieval solution for small values of s. T and α are dimensionless parameters.

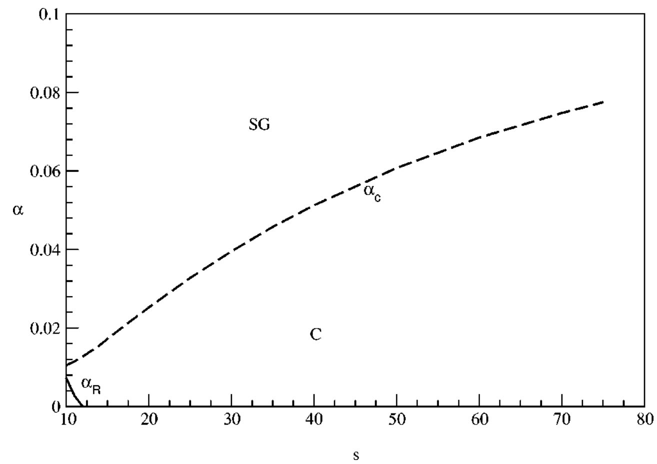

FIG. 3. Phase diagram in the α-s plane in the absence of synaptic noise for correlation parameter $b=0.4$, and for values $\lambda_1/\lambda_m=1.75$ (full line) and 0.75 (broken line). The meanings of lines α_R and α_c are as in Fig. 2. The two lines α_c coincide for both values of λ_1/λ_m. T and α are dimensionless parameters.

them, let us say $\mu=1$; then the retrieval phase of one particular example $\{\xi_i^{11}\}$ is characterized by the asymmetric solution $m_{11}>m_{s-1}$ in Eq. (18), while in the categorization phase $m_{11}=m_{s-1}$ and the categorization error ϵ_1 in Eq. (16) drops abruptly to small values. For finite values of $\alpha\neq 0$ there is also a spin glass phase with $q\neq 0$, $r\neq 0$ in Eq. (20), which interferes with the retrieval and categorization phases. We calculate the spin glass phase in the replica symmetric approximation, together with the de Almeida–Thouless instability line where the replicon eigenvalue vanishes.

In the present work we consider the simplest choice of differentiated weights in Eq. (2), namely $\lambda_\tau=\lambda_m,\tau\geqslant 2$, and $\lambda_m\neq\lambda_1$. We show in Fig. 1 the phase diagram in the T-α plane, obtained from the solution of the saddle point equations, Eqs. (26)–(31), for correlation parameter $b=0.4$, number of examples $s=10$, and for different values of the ratio λ_1/λ_m. For $\lambda_1=\lambda_m$ we recover the results of Ref. [15], which show only a spin glass and categorization ordered phases, while for $\lambda_1>\lambda_m$ we observe the occurrence of a locally stable retrieval phase at low temperatures. We also show the de Almeida–Thouless instability line, which limits the validity of the replica symmetric theory at very low temperatures. The effect of varying the number of examples s in the absence of synaptic noise is shown in Fig. 2, where we present the phase diagram in the α-s plane for correlation parameter $b=0.2$ for several values of λ_1/λ_m. A comparison with Fig. 1 shows that the relative enhancement of the retrieval versus categorization phases may be controlled by increasing λ_1/λ_m and decreasing the correlation parameter b. The phase diagrams in the α-s plane for $T=0$, several values of the correlation parameter b, and λ_1/λ_m are shown in Fig. 3.

To conclude, the present results show that weighted examples allow us to control the extent and importance of the retrieval versus categorization phases within a replica symmetric spin glass theory. The de Almeida–Thouless instability line limits our results in the lower part of phase diagram in the T-α plane, particularly in the phase boundary between the categorization phase and the spin glass phase. In contrast, the breakdown of the replica symmetric mean field equation is less important in the Hopfield model [1].

ACKNOWLEDGMENTS

This work was partially supported by Conselho Nacional de Pesquisa e Desenvolvimento (CNPq) and by Financiadora de Estudos e Projetos (FINEP).

APPENDIX

We present here the detailed derivation of the expressions used in the paper. From Eq. (17) we obtain that $pn\beta\hat{G}$ in Eq. (10) is given by

$$e^{pn\beta\hat{G}}=\left\langle\left\langle e^{(\beta/2N)\sum_{\alpha}\sum_{i\neq j}\sum_{\mu=2}^{p}\sum_{\tau=1}^{s}\lambda_{\tau}\xi_i^{\mu\tau}\xi_j^{\mu\tau}S_i^{\alpha}S_j^{\alpha}}\right\rangle\right\rangle, \tag{A1}$$

where the bracket indicates an average over $\{\xi_i^{\mu\tau}\}$, $\{\xi_i^{\mu}\}$ for $\mu\geqslant 2$, in that order. Taking advantage of the statistical independence of the $\{\xi_i^{\mu}\}$, we may write [21]

$$e^{pn\beta\hat{G}}=\prod_{\mu\geqslant 2}\left\langle\left\langle e^{(\beta/2N)\sum_{i,j}t_{ij}\sum_{\tau}\lambda\tau\xi_i^{\mu\tau}\xi_j^{\mu\tau}}\right\rangle_{\{\xi_i^{\mu\tau}\}}\right\rangle_{\{\xi_i^{\mu}\}}, \tag{A2}$$

where

$$t_{ij}=\sum_{\alpha}S_i^{\alpha}S_j^{\alpha}. \tag{A3}$$

From now on we will indicate by a double bracket an average as shown in Eq. (A2); then

$$\begin{aligned}\langle\langle\xi_i^{\mu\tau}\rangle\rangle&=\langle b\xi_i^{\mu}\rangle=0,\\ \langle\langle\xi_i^{\mu\tau}\xi_i^{\mu\tau'}\rangle\rangle&=b^2+(1-b^2)\delta_{\tau\tau'}.\end{aligned} \tag{A4}$$

By using the cumulant expansion in Eq. (A2), we may write

$$n\beta\hat{G}=\sum_{\kappa=2}^{\infty}\frac{1}{\kappa!}\left(\frac{\beta}{2}\right)^{\kappa}S_c(\kappa), \tag{A5}$$

$$S_c(\kappa)=\left\langle\left\langle\left[\sum_{i\neq j}\frac{t_{ij}}{N}\sum_{\tau}\lambda_{\tau}\xi_i^{\mu\tau}\xi_j^{\mu\tau}\right]^{\kappa}\right\rangle\right\rangle_{\text{cumulant}}. \tag{A6}$$

It has been shown elsewhere [21] that the only relevant contribution in the thermodynamic limit comes from cumulant averages with the maximum number of sums over sites, or ‘‘ring diagrams.’’ This gives from Eq. (A4) and Eq. (A6),

$$\begin{aligned}S_{\text{ring}}(\kappa)=&\frac{2^{(k-1)}(\kappa-1)!}{N^{\kappa}}\sum_{i_1\neq i_2\neq\cdots\neq i_{\kappa}}t_{i_1i_2}t_{i_2i_3}\cdots t_{i_{\kappa}i_1}\\ &\times\sum_{\tau_1\tau_2\cdots\tau_{\kappa}}L_{\tau_1\tau_2}L_{\tau_2\tau_3}\cdots L_{\tau_{\kappa}\tau_1},\end{aligned} \tag{A7}$$

where we introduced an $s\times s$ matrix $\underline{\underline{L}}$ with elements

$$L_{\tau_1\tau_2}=\delta_{\tau_1\tau_2}\lambda_{\tau_2}+(1-\delta_{\tau_1\tau_2})b^2\lambda_{\tau_2}. \tag{A8}$$

By performing first the sums over sites and later the sums over replica indices, we obtain in Eq. (A7)

$$S_{\text{ring}}(\kappa)=2^{(k-1)}(\kappa-1)!\,\text{Tr}(\underline{\underline{L}}^{\kappa})\text{Tr}\,(\underline{\underline{\hat{q}}}^{\kappa}), \tag{A9}$$

where $\underline{\underline{\hat{q}}}$ is the $n\times n$ overlap matrix with elements

$$\hat{q}_{\alpha_1\alpha_2}=\frac{1}{N}\sum_{i}S_i^{\alpha_1}S_i^{\alpha_2}. \tag{A10}$$

Introducing Eq. (A9) in Eq. (A5), we obtain

$$n\beta\hat{G}=\frac{1}{2}\sum_{\kappa=2}\frac{1}{\kappa}\beta^{\kappa}\,\text{Tr}(\underline{\underline{L}}^{\kappa})\text{Tr}(\underline{\underline{\hat{q}}}^{\kappa}). \tag{A11}$$

In order to evaluate the traces in Eq. (A11) we first replace the overlaps in Eq. (A10) by the order parameters $q_{\alpha_1\alpha_2}$, for $\alpha_1\neq\alpha_2$, by means of Eq. (16). We then find that the matrix $\underline{\underline{L}}$ has nondegenerate eigenvalues ω_+,ω_- in Eq. (16), and one eigenvalue $\omega_m=\lambda_m(1-b^2)$ that is $(s-2)$ degenerate; then

$$\text{Tr}\,\underline{\underline{L}}^{\kappa}=\omega_+^{\kappa}+\omega_-^{\kappa}+(s-2)\omega_m^{\kappa}. \tag{A12}$$

Introducing Eq. (A12) into Eq. (A11), we obtain

$$n\beta G=-\frac{1}{2}\text{Tr}\sum_{i=+,-,m}a_i\{\ln[1-\beta\omega_i\underline{\underline{q}}]+\beta\omega_i\underline{\underline{q}}\} \tag{A13}$$

with $a_+=a_-=1, a_m=(s-2)$.

To calculate the instability line, we parametrize $q_{\alpha_1\alpha_2}$ as in Eq. (20) and we evaluate the traces in Eq. (A13) by expanding the logarithms to second order in $\eta_{\alpha_1\alpha_2}$, with the result

$$n\beta G=n\beta G_{\text{RS}}+\beta\Delta G_{\text{RSB}}, \tag{A14}$$

where

$$\beta G_{\text{RS}}=-\frac{1}{2}\sum_{i=+,-,m}a_i\left\{\ln[1-\omega_i\beta C]-q\frac{\beta\omega_i}{1-\omega_i\beta C}+\beta\omega_i\right\} \tag{A15}$$

and $C=(1-q)$. For the replica symmetry breaking part, we obtain

$$\begin{aligned}\beta\Delta G_{\text{RSB}}=&-\frac{\alpha}{4}\sum_{\alpha_1\neq\alpha_2}\sum_{\alpha_3\neq\alpha_4}\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4}\\ &\times\sum_{i=+,-,m}a_i\beta^2\omega_i^2M^i_{\alpha_1\alpha_2}M^i_{\alpha_3\alpha_4},\end{aligned} \tag{A16}$$

where

$$\underline{\underline{M}}^i=[1-\beta\omega_i\underline{\underline{q}}]^{-1}. \tag{A17}$$

We also have to expand Λ in Eq. (19) to second order in $\lambda_{\gamma\delta}$ with the result

$$\beta\Lambda = n\left\{ \int_{-\infty}^{\infty} \mathcal{D}z \langle\langle \ln 2\cosh[\beta h_1 \xi^1 + \beta m_{11}\lambda_1 \xi^{11} + \beta m_{s-1}\lambda_m x + \beta\sqrt{\alpha r} z]\rangle\rangle - \frac{1}{2}\alpha\beta^2 r \right\} + \beta\Delta\Lambda_{\rm RSB}, \tag{A18}$$

$$\beta\Delta\Lambda_{\rm RSB} = \frac{1}{2}\left(\frac{\alpha\beta^2}{2}\right)^2 \sum_{\gamma\neq\delta}\sum_{\gamma'\delta'} \lambda_{\gamma\delta}\lambda_{\gamma'\delta'}\Gamma_{\gamma\delta;\gamma'\delta'}, \tag{A19}$$

where the correlation functions $\Gamma_{\gamma\delta;\gamma'\delta'} = \langle\langle S^\gamma S^\delta; S^{\gamma'} S^{\delta'}\rangle\rangle - q^2$ are calculated with the replica symmetric theory.

From Eq. (17), Eq. (A16), and Eq. (A19), we obtain for the replica symmetry breaking free energy

$$\beta\Delta f_{\rm RSB} = \beta\Delta G_{\rm RSB} + \beta\Delta\Lambda_{\rm RSB} - \frac{1}{2}\alpha\beta^2 \sum_{\gamma\neq\delta} \eta_{\gamma\delta}\lambda_{\gamma\delta}. \tag{A20}$$

We parametrize the replicon eigenvector [1,20] of the quadratic form in Eq. (A20),

$$\eta_{\theta\nu} = c_2, \eta_{\theta\alpha_1} = \eta_{\nu\alpha_1} = \frac{c_2}{(2-n)}, \quad \eta_{\alpha_1\alpha_2} = \frac{2c_2}{(2-n)(3-n)},$$

$$\lambda_{\theta\nu} = c_2', \lambda_{\theta\alpha_1} = \lambda_{\nu\alpha_1} = \frac{c_2'}{(2-n)}, \quad s\lambda_{\alpha_1\alpha_2} = \frac{2c_2'}{(2-n)(3-n)}, \tag{A21}$$

where $\alpha_1, \alpha_2 \neq \theta, \nu$. We obtain that the replicon eigenvalue χ is the smaller solution of the secular equation:

$$[K' - \chi][\alpha\beta^2 K - \chi] - 1 = 0, \tag{A22}$$

where

$$K' = \sum_{i=+,-,m} a_i \frac{\omega_i^2}{[1-\beta\omega_i C]^2}, \tag{A23}$$

$$K = \int_{-\infty}^{\infty} \mathcal{D}z \ \langle\langle \cosh^{-4}[\beta m_{11}\lambda_1 \xi^{11} + \beta m_{s-1}\lambda_m x_s + \beta\sqrt{\alpha r} z]\rangle\rangle, \tag{A24}$$

and the double bracket indicates the random average over ξ^{11}, x_s, and ξ^1, in that order.

[1] D.J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985); Ann. Phys. (N.Y.) **173**, 30 (1987).
[2] J.J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).
[3] J.F. Fontanari and R. Meir, Phys. Rev. A **40**, 2806 (1989).
[4] J.F. Fontanari, J. Phys. (France) **51** 2421 (1990).
[5] P.R. Krebs and W.K. Theumann, J. Phys. A **26**, 3983 (1993).
[6] J.A. Martins and W.K. Theumann, Physica A **253**, 38 (1998).
[7] D.C. Dominguez and W.K. Theumann, J. Phys. A **30**, 1403 (1997).
[8] D.A. Stariolo and F.A. Tamarit, Phys. Rev. A **46**, 5249 (1992).
[9] C.R. da Silva, F.A. Tamarit, Ney Lemke, J.J. Arenzon, and E.M.F. Curado, J. Phys. A **28**, 1593 (1995).
[10] D.C. Dominguez and W.K. Theumann, J. Phys. A **29**, 749 (1996).
[11] P.R. Krebs and W.K. Theumann, Phys. Rev. E **60**, 4580 (1999).
[12] Rogério L. Costa and Alba Theumann, Physica A **268**, 499 (1999).
[13] J.L. van Hemmen and V.A. Zagrebnov, J. Phys. A **20**, 3989 (1987).
[14] L. Viana, J. Phys. (France) **49**, 167 (1988).
[15] C. Marzban and R. Viswanathan, Phys. Lett. A **191**, 127 (1994).
[16] M. Mézard, J.P. Nadal, and G. Toulouse, J. Phys. (France) **47**, 1457 (1986).
[17] G. Parisi, J. Phys. A **19**, L617 (1986); **19**, L675 (1986).
[18] J.F. Fontanari and R. Köberle, J. Phys. A **21**, L253 (1988).
[19] S. Nicolis, Europhys. Lett. **12**, 583 (1990).
[20] J.R.L. de Almeida and D.J. Thouless, J. Phys. A **11**, 983 (1978).
[21] Alba Theumann, Phys. Rev. E **53**, 6361 (1996); M.A.P. Idiart and Alba Theumann, J. Phys. A **25**, 779 (1992).