Journal of Statistical Physics, Vol. 70, Nos. 5/6, 1993

On the Parallel Dynamics of the Q-State Potts and Q-Ising Neural Networks

D. Bollé,[1] B. Vinck,[1] and V. A. Zagrebnov[1,2]

Received May 4, 1992; final September 15, 1992

Using a probabilistic approach, the parallel dynamics of the Q-state Potts and Q-Ising neural networks are studied at zero and at nonzero temperatures. Evolution equations are derived for the first time step and arbitrary Q. These formulas constitute recursion relations for the exact parallel dynamics of the extremely diluted asymmetric versions of these networks. An explicit analysis, including dynamical capacity–temperature diagrams and the temperature dependence of the overlap, is carried out for $Q=3$. Both types of models are compared.

KEY WORDS: Multistate networks; parallel dynamics; extreme dilution; probabilistic approach.

1. INTRODUCTION

During the last few years there has been considerable interest in neural networks with multistate neurons.[1–26] Models using Q-Ising neurons[1–10] as well as models using clock-type[11–14] and Potts-type[14–26] neuron states have been discussed. Networks of Q-Ising and clock-type neurons can function as associative memories for gray-toned or colored patterns, while Potts networks concern in fact two-state response functions allowing one of the states to have a more complicated internal structure (representing, e.g., a distinction between background and pattern).

In this paper we consider the parallel dynamics of Q-state Potts and Q-Ising neural networks, using a probabilistic approach (see, e.g., refs. 27 and 28). In more detail, employing a signal-to-noise ratio analysis based on the law of large numbers and the central limit theorem, we derive evolution

[1] Instituut voor Theoretische Fysica, K.U. Leuven, B-3001 Leuven, Belgium.

[2] On leave of absence from the Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, Dubna 141980, Russia.

822/70/5-6-2

equations for the first time step at zero and nonzero temperatures for arbitrary Q. For simplicity we take the patterns and the neurons out of the same set of variables. This is no essential restriction. In the case of extremely diluted asymmetric versions of these networks the one-step evolution equations describe the exact dynamics. This dynamics is solved explicitly for $Q=3$.

The rest of this paper is organized as follows. In Section 2 we introduce the Q-state Potts model and we present the general formula for one-step parallel dynamics at arbitrary temperature. Section 3 does the same for the Q-state Ising case. The fundamental difficulty to obtain more steps in these fully connected symmetric networks is the strong feedback (see, e.g., ref. 29). In Section 4 we consider the extremely diluted asymmetric versions of these models. In this case the feedback is suppressed (see refs. 30 and 31 and the references therein) such that the one-step dynamics accounts for the full dynamical evolution. We write down explicitly the corresponding fixed-point equations for $Q=3$. Comparing the two types of models, given a uniform distribution of the patterns, we find that in the case of the Potts model the fixed-point equation for the overlap involves three Gaussian integrations, while in the case of the Ising model the fixed-point equation contains only one Gaussian integration but we also need a fixed-point equation for the activity of the patterns to have a closed system of equations. Finally, in Section 5 we discuss and compare in detail the dynamical capacity–temperature diagrams and the temperature dependence of the relevant order parameters for these extremely diluted $Q=3$ models.

2. PARALLEL DYNAMICS OF THE Q-STATE POTTS NETWORK

2.1. The Model

Consider a network of N neurons, $\Lambda=\{1, 2,..., N\}$. We suppose that each neuron $i\in\Lambda$ can be described by a Potts spin $\sigma_i\in\mathcal{Q}=\{1, 2,..., Q\}$. The patterns to be stored in this network, $\{\xi_i^\mu\in\mathcal{Q}\}$, $i\in\Lambda$, $\mu\in\mathcal{P}=\{1, 2,..., p\}$, are taken to be independent identically distributed random variables (i.i.d.r.v.) with uniform distribution on $\mathcal{Q}$, i.e.,

$$\Pr\{\xi_i^\mu=k\in\mathcal{Q}\}=\frac{1}{Q} \tag{1}$$

for $i\in\Lambda$, $\mu\in\mathcal{P}$. The neurons are interconnected by a synaptic matrix $J=[J_{ij}^{kl}(\{\xi_s^\mu\}, s\in\Lambda, \mu\in\mathcal{P})]$ for $i, j\in\Lambda$, and $k, l\in\mathcal{Q}$. Given a network configuration $\boldsymbol{\sigma}_\Lambda=\{\sigma_j\}$, $j\in\Lambda$, the energy potential of neuron $i\in\Lambda$ is then

$$h_i(\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i})=-\sum_{j\in\Lambda\setminus i}\sum_{k,l\in\mathcal{Q}} J_{ij}^{kl}u(k, \sigma_i)\,u(l, \sigma_j) \tag{2}$$

with u the Potts spin operator given by

$$u(\xi_j^\mu, \sigma_j) = Q\delta(\xi_j^\mu, \sigma_j) - 1 \tag{3}$$

The nonzero-temperature ($T = \beta^{-1} > 0$) parallel dynamics $D_t^{(\beta)}$ of the system is defined by simultaneous updating of the neurons according to the conditional probabilities:

$$\Pr\{\sigma_i(t+1) = \sigma \in \mathcal{Q} \mid \boldsymbol{\sigma}_{\Lambda\setminus i}(t)\} = \frac{\exp[-\beta h_i(\sigma; \boldsymbol{\sigma}_{\Lambda\setminus i}(t))]}{\sum_{s\in\mathcal{Q}} \exp[-\beta h_i(s; \boldsymbol{\sigma}_{\Lambda\setminus i}(t))]} \tag{4}$$

To build in the capacity for learning and memory in this network, its stationary configurations representing the retrieved patterns, $D_t^{(\beta)}\boldsymbol{\sigma}^* = \boldsymbol{\sigma}^*$, i.e., the attractors of $D_t^{(\beta)}$, must be correlated with the stored patterns $\{\boldsymbol{\xi}^\mu\}$, $\mu \in \mathcal{P}$. This can be achieved[15] by choosing the learning rule

$$J_{ij}^{kl} = \frac{1}{Q^2 N} \sum_{\mu\in\mathcal{P}} u(\xi_i^\mu, k)\, u(\xi_j^\mu, l) \tag{5}$$

We remark that (5) tells us that each pair of neurons is connected.

2.2. One-Step Dynamics at Arbitrary Temperature

We first look at the Potts network introduced above at zero temperature ($\beta \to \infty$). Then the dynamics (4) at $i \in \Lambda$ gets the form

$$\sigma_i(t+1) = D_{t=1}^{(\infty)}\sigma_i(t): \quad h_i(\sigma_i(t+1); \boldsymbol{\sigma}_{\Lambda\setminus i}(t)) = \min_{s\in\mathcal{Q}} h_i(s_i; \boldsymbol{\sigma}_{\Lambda\setminus i}(t)) \tag{6}$$

Let the initial configuration $\boldsymbol{\sigma}_\Lambda(t=0)$ [in the sequel we denote this by $\boldsymbol{\sigma}_\Lambda(0)$] correspond to a realization of the i.i.d.r.v. $\{\sigma_i(0)\}$, $i \in \Lambda$, with uniform distribution on $\mathcal{Q}$ correlated with only one pattern, say $\boldsymbol{\xi}^\nu$, i.e.,

$$\Pr\{\sigma_i(0) = \xi_i^\mu\} = \frac{1}{Q}(1 + \delta_{\mu\nu} m_0^\nu) \tag{7}$$

where $m_0^\nu > 0$. By the law of large numbers (LLN), the overlap with pattern $\boldsymbol{\xi}^\nu$ at $t=0$ is defined as

$$m^\mu(0) \equiv \lim_{N\to\infty} m_\Lambda^\mu(0) = \lim_{N\to\infty} \frac{1}{N} \sum_{j\in\Lambda} u(\xi_j^\mu, \sigma_j(0))$$

$$\overset{\Pr}{=} E[u(\xi_j^\mu, \sigma_j(0))] = \delta_{\mu\nu} m_0^\nu \tag{8}$$

where the convergence is in probability.[32] According to ref. 27, the evolution of the main overlap (8) generated by $D_{t=1}^{(\infty)}$ can be calculated via an explicit computation of the internal network noise in the limit $N \to \infty$ at fixed $\alpha \equiv p/N$. (In the sequel the latter will be denoted as the α-limit.) Let $\mathscr{P}_{k,i}^{(\nu)} = \{\mu \in \mathscr{P} \setminus \nu : \xi_i^\mu = k\}$. Then the energy potential (2) for neuron i can be written as

$$-h_i(s_i; \boldsymbol{\sigma}_{\Lambda \setminus i}(0)) = u(\xi_i^\nu, s_i)\, m_{\Lambda \setminus i}^\nu(0) + \sum_{k \in \mathscr{Q}} u(k, s_i) \frac{1}{N} \sum_{\mu \in \mathscr{P}_{k,i}^{(\nu)}} \sum_{j \in \Lambda \setminus i} u(\xi_j^\mu, \sigma_j(0)) \tag{9}$$

The first term on the r.h.s. of (9) represents the contribution from the signal, the second term describes the noise contribution. Using (7), we see that for $\mu \neq \nu$

$$E[u(\xi_j^\mu, \sigma_j(0))] = 0 \tag{10}$$

$$\mathrm{Var}[u(\xi_j^\mu, \sigma_j(0))] \equiv a^\mu(0) = Q - 1 \tag{11}$$

such that $\{u(\xi_j^\mu, \sigma_j(0))\}$, $\mu \in \mathscr{P}$, $j \in \mathbb{N}$, is a sequence of i.i.d.r.v. Therefore, using the central limit theorem (CLT), we find that in the limit $N \to \infty$ ($\Lambda \uparrow \mathbb{N}$) the kth component of the noise term in (9) gets the form

$$\alpha\text{-lim} \frac{1}{N} \sum_{\mu \in \mathscr{P}_{k,i}^{(\nu)}} \sum_{j \in \Lambda \setminus i} u(\xi_j^\mu, \sigma_j(0)) \overset{\mathscr{D}}{=} \left[\frac{\alpha(Q-1)}{Q}\right]^{1/2} \mathscr{N}_{\nu,i}^{(k)}(0, 1) \tag{12}$$

where, as indicated, the convergence is in distribution (see, e.g., ref. 32). The quantity $\mathscr{N}(0, 1)$ is a Gaussian random variable with expectation 0 and variance 1. We have used the fact that in the α-limit, large-deviation arguments[32] tell us that

$$\Pr\left\{\left|\frac{|\mathscr{P}_{k,i}^{(\nu)}|}{p} - \frac{1}{Q}\right| \geqslant \delta\right\} \leqslant \exp(-\lambda_\delta p), \quad \lambda_\delta > 0 \quad \text{for any} \quad \delta > 0$$

Hence, in the α-lim the random energy potential (9) converges in distribution to the random variable

$$-h_i(s_i; \boldsymbol{\sigma}(0)) = u(\xi_i^\nu, s_i)\, m^\nu(0) + \sum_{k \in \mathscr{Q}} u(k, s_i) \left[\frac{\alpha(Q-1)}{Q}\right]^{1/2} \mathscr{N}_{\nu,i}^{(k)}(0, 1) \tag{13}$$

where, using the initial condition (7), we know that $\{\mathscr{N}_{\nu,i}^{(k)}(0, 1)\}$, $k \in \mathscr{Q}$, $i \in \mathbb{N}$, are independent Gaussian random variables.

To calculate the configuration at the next time step $t=1$, we have to consider the following two possibilities:

If $s_i=\xi_i^v$, then, recalling (3) and (13), the energy potential becomes

$$\begin{aligned}-h_i^{(a)} &\equiv -h_i(s_i=\xi_i^v;\boldsymbol{\sigma}(0))\\ &=(Q-1)\left(m^v(0)+\left[\frac{\alpha(Q-1)}{Q}\right]^{1/2}\mathcal{N}_{v,i}^{(k=\xi_i^v)}(0,1)\right)\\ &\quad+\sum_{k\in\mathcal{Q}\setminus\xi_i^v}(-1)\left[\frac{\alpha(Q-1)}{Q}\right]^{1/2}\mathcal{N}_{v,i}^{(k)}(0,1)\end{aligned}\tag{14}$$

else $s_i=\tilde{k}$ $(\neq\xi_i^v)$ and the energy potential is given by

$$\begin{aligned}-h_i^{(b)} &\equiv -h_i(s_i=\tilde{k};\boldsymbol{\sigma}(0))\\ &=(-1)\left(m^v(0)+\left[\frac{\alpha(Q-1)}{Q}\right]^{1/2}\sum_{k\in\mathcal{Q}\setminus\tilde{k}}\mathcal{N}_{v,i}^{(k)}(0,1)\right)\\ &\quad+(Q-1)\left[\frac{\alpha(Q-1)}{Q}\right]^{1/2}\mathcal{N}_{v,i}^{(\tilde{k})}(0,1)\end{aligned}\tag{15}$$

To solve the minimization problem in (6), we compute the difference

$$\Delta_i=h_i^{(b)}-h_i^{(a)}=Qm^v(0)+Q\left[\frac{\alpha(Q-1)}{Q}\right]^{1/2}[\mathcal{N}_{v,i}^{(k=\xi_i^v)}(0,1)-\mathcal{N}_{v,i}^{(\tilde{k})}(0,1)]\tag{16}$$

such that the configuration at $t=1$ is

$$\sigma_i(1)=\begin{cases}\xi_i^v & \text{if } \Delta_i>0\\ \tilde{k}=k_{\max} & \text{if } \Delta_i\leqslant 0\end{cases}\tag{17}$$

where $k_{\max}$ is defined by the relation

$$\mathcal{N}_{v,i}^{(k_{\max})}(0,1)=\max_{\tilde{k}\in\mathcal{Q}\setminus\xi_i^v}\mathcal{N}_{v,i}^{(\tilde{k})}(0,1)\tag{18}$$

Since, as we have noticed already, the $\{\mathcal{N}_{v,i}^{(k)}\}$, $k\in\mathcal{Q}$, are a collection of i.i.d.r.v., we can write

$$\Pr\{\mathcal{N}_{v,i}^{k_{\max}}\leqslant x\}=(\Phi(x))^{Q-1}\tag{19}$$

$$\Phi(x)\equiv\frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{x}dz\exp\left(-\frac{z^2}{2}\right)\tag{20}$$

Therefore we arrive at

$$\Pr\{\sigma_i(1)=\xi_i^\nu \mid \boldsymbol{\sigma}(0)\}=\int_{-\infty}^{+\infty} Dy\left[\Phi\left(\left[\frac{Q}{\alpha(Q-1)}\right]^{1/2} m^\nu(0)+y\right)\right]^{Q-1} \tag{21}$$

$$Dy\equiv\frac{dy}{(2\pi)^{1/2}}\exp\left(-\frac{y^2}{2}\right) \tag{22}$$

and finally [cf. (8)] we obtain

$$m^\nu(1)=Q\Pr\{\sigma_i(1)=\xi_i^\nu \mid \boldsymbol{\sigma}(0)\}-1 \tag{23}$$

$$m^\mu(1)=0, \qquad \mu\neq\nu \tag{24}$$

Furthermore, using (11), we calculate

$$a^\mu(1)=\mathrm{Var}[u(\xi_j^\mu, \sigma_j(1))]=Q-1=a^\mu(0) \tag{25}$$

At this point we remark that the key observation in the foregoing derivation is that the CLT can be used in (12) because $\{\sigma_i(0)\}$, $i\in\mathbb{N}$, are i.i.d.r.v. correlated with only one pattern $\boldsymbol{\xi}^\nu$.

At nonzero temperatures the dynamical evolution can be calculated using the method of auxiliary temperature noise (see, e.g., ref. 27).

For each $i\in\mathbb{N}$ let $\{\phi_i(l)\}$, $l\in\{1, 2,..., n\}$ be a collection of i.i.d.r.v. with joint distribution

$$F_\beta(x_1, x_2,..., x_n)\equiv\Pr\left\{\bigcap_{l=1}^{n}\{\phi_i(l)<x_l\}\right\}=\left(1+\sum_{l=1}^{n} e^{-\beta x_l}\right)^{-1} \tag{26}$$

The conditional probability (4) for $D_i^{(\beta)}$ then gets the form

$$\begin{aligned}&\Pr\{\sigma_i(t+1)=\sigma\in\mathcal{Q}\mid\boldsymbol{\sigma}_{\Lambda\setminus i}(t)\}\\&\quad=\Pr\left\{\bigcap_{s\in\mathcal{Q}\setminus\sigma}\{\phi_i(s)<h_i(s;\boldsymbol{\sigma}_{\Lambda\setminus i}(t))-h_i(\sigma;\boldsymbol{\sigma}_{\Lambda\setminus i}(t))\}\right\}\end{aligned} \tag{27}$$

This representation of the dynamics allows us to deal with the temperature and the internal noise *simultaneously*. Substituting (14)–(16) in (27) taken at the time step $t=1$, we arrive at the following result:

$$\begin{aligned}&\Pr\{\sigma_i(1)=\xi_i^\nu\mid\boldsymbol{\sigma}(0)\}\\&\quad=\Pr\left\{\bigcap_{s\in\mathcal{Q}\setminus\xi_i^\nu}\left\{\phi_i(s)<Qm^\nu(0)+Q\left[\frac{\alpha(Q-1)}{Q}\right]^{1/2}\right.\right.\\&\qquad\left.\left.\times[\mathcal{N}_{\nu,i}^{(\xi_i^\nu)}(0,1)-\mathcal{N}_{\nu,i}^{(s)}(0,1)]\right\}\right\}\end{aligned} \tag{28}$$

Finally, using (26), we obtain the explicit formula

$$\Pr\{\sigma_i(1)=\xi_i^\nu \mid \boldsymbol{\sigma}(0)\}$$
$$=\int_{R^1} Dy \int_{R^{Q-1}} Dx_1 \cdots Dx_{Q-1}$$
$$\times\left[1+\sum_{l\in\mathscr{Q}\setminus Q} \exp\left(-\beta\left\{Qm^\nu(0)+Q\left[\frac{\alpha(Q-1)}{Q}\right]^{1/2}(y-x_l)\right\}\right)\right]^{-1} \tag{29}$$

where Dx_i is defined in (22). The evolution of the main overlap reads [compare (23)–(24)]

$$m^\nu(1)=Q\Pr\{\sigma_i(1)=\xi_i^\nu \mid \boldsymbol{\sigma}(0)\}-1 \tag{30}$$

We remark that for $\beta\to\infty$, Eq. (29) reduces to (21).

3. PARALLEL DYNAMICS OF THE *Q*-ISING NETWORK

3.1. The Model

Consider a network Λ of N neurons which can take values in the set

$$\mathscr{S}=\{-1=s_1<s_2<\cdots<s_{Q-1}<s_Q=+1\}$$

The p patterns to be stored in this network, $\{\xi_i^\mu\in\mathscr{S}\}$, $i\in\Lambda=\{1, 2,..., N\}$, $\mu\in\mathscr{P}=\{1, 2,..., p\}$, are supposed to be i.i.d.r.v. with zero mean, $E[\xi_i^\mu]=0$, and variance $a=\mathrm{Var}[\xi_i^\mu]$. The latter is a measure for the activity of the patterns.

Given a configuration $\boldsymbol{\sigma}_\Lambda=\{\sigma_j\}$, $j\in\Lambda$, the local field h_i in neuron $i\in\Lambda$ is

$$h_i(\boldsymbol{\sigma}_\Lambda)=\sum_{j\in\Lambda\setminus i} J_{ij}\sigma_j \tag{31}$$

where the synaptic couplings are given by

$$J_{ij}=\frac{1}{Na}\sum_{\mu\in\mathscr{P}}\xi_i^\mu\xi_j^\mu \tag{32}$$

The parallel dynamics of this network is then defined by the transition probabilities

$$\Pr\{\sigma_i(t+1)=s_k\in\mathscr{S}\mid\boldsymbol{\sigma}_{\Lambda\setminus i}(t)\}=\frac{\exp[-\beta\varepsilon_i(s_k\mid\boldsymbol{\sigma}_{\Lambda\setminus i}(t))]}{\sum_{s\in\mathscr{S}}\exp[-\beta\varepsilon_i(s\mid\boldsymbol{\sigma}_{\Lambda\setminus i}(t))]} \tag{33}$$

Here the energy potential $\varepsilon_i(s|\boldsymbol{\sigma})$ of neuron i is taken to be[4]

$$\varepsilon_i(s|\boldsymbol{\sigma}_{\Lambda\setminus i}) = -\tfrac{1}{2}[h_i(\boldsymbol{\sigma}_{\Lambda\setminus i})\, s - bs^2] \qquad b>0 \tag{34}$$

3.2. One-Step Dynamics at Arbitrary Temperature

Considering the network first at zero temperature, the dynamics (33)–(34) at $i\in\Lambda$ takes the form

$$\sigma_i(t)\to\sigma_i(t+1)=s_k: \quad \min_{s\in\mathscr{S}} \varepsilon_i(s|\boldsymbol{\sigma}_{\Lambda\setminus i}(t)) = \varepsilon_i(s_k|\boldsymbol{\sigma}_{\Lambda\setminus i}(t)) \tag{35}$$

This rule is equivalent to using a gain function ($Q>2$)

$$\sigma_i(t+1) = g[h_i(\boldsymbol{\sigma}_{\Lambda\setminus i}(t))] \tag{36}$$

which has, since $\mathscr{S}$ is finite, a steplike shape. In the case of equidistant states

$$\mathscr{S}=\mathscr{S}_Q=\left\{s_k=-1+\frac{2(k-1)}{Q-1},\ k=1,2,\ldots,Q\right\}$$

it looks like

$$g(x)=\begin{cases}+1 & \text{if} \quad s_{Q-1}+(Q-1)^{-1}\leqslant x/2b\\ s_k & \text{if} \quad s_k-(Q-1)^{-1}<x/2b<s_k+(Q-1)^{-1}\\ -1 & \text{if} \quad x/2b\leqslant s_2-(Q-1)^{-1}\end{cases} \tag{37}$$

To measure the retrieval quality of the network, i.e., to estimate the evolution of an initial configuration $\boldsymbol{\sigma}_\Lambda(0)$ to a stored pattern $\boldsymbol{\xi}^\nu$, one can use the Hamming distance

$$\begin{aligned} d_{\mathrm{H}}(\boldsymbol{\sigma}(t),\boldsymbol{\xi}^\nu) &= \frac{1}{N}\sum_{i\in\Lambda}[\sigma_i(t)-\xi_i^\nu]^2 \\ &= \frac{1}{N}\sum_{i\in\Lambda}(\xi_i^\nu)^2+\frac{1}{N}\sum_{i\in\Lambda}[\sigma_i(t)]^2-\frac{2}{N}\sum_{i\in\Lambda}\xi_i^\nu\sigma_i(t) \end{aligned} \tag{38}$$

From this relation it is clear that the dynamics can be completely described in terms of the main overlap

$$m_\Lambda^\nu(t)=\frac{1}{Na}\sum_{i\in\Lambda}\xi_i^\nu\sigma_i(t) \tag{39}$$

and the arithmetic mean of the neuron activities

$$a_{\Lambda}(t)=\frac{1}{N}\sum_{i\in\Lambda}[\sigma_i(t)]^2 \tag{40}$$

Suppose that the initial configuration $\{\sigma_i(0)\}$, $i\in\Lambda$, is a collection of i.i.d.r.v. with mean $E[\sigma_i(0)]=0$, variance $\mathrm{Var}[\sigma_i(0)]=a(0)$, and correlated with only one stored pattern:

$$E[\xi_i^{\mu}\sigma_i(0)]=\delta_{\mu\nu}m_0^{\nu}a, \qquad m_0^{\nu}>0 \tag{41}$$

This equation corresponds to (8) for the Potts model. This implies that by the LLN one gets for the activity and the main overlap at $t=0$

$$a(0)\equiv\lim_{N\to\infty}\frac{1}{N}\sum_{i\in\Lambda}[\sigma_i(0)]^2\overset{\Pr}{=}E[\sigma_i^2(0)] \tag{42}$$

$$m^{\nu}(0)\equiv\lim_{N\to\infty}\frac{1}{Na}\sum_{i\in\Lambda}\xi_i^{\nu}\sigma_i(0)\overset{\Pr}{=}\mathrm{m}_0^{\nu} \tag{43}$$

To obtain the configuration at $t=1$ we first have to calculate the local field (31) at $t=0$. Recalling the learning rule (32), we obtain

$$h_i(\boldsymbol{\sigma}_{\Lambda}(0))=\xi_i^{\nu}\frac{1}{Na}\sum_{j\in\Lambda\setminus i}\xi_j^{\nu}\sigma_j(0)+\sqrt{\alpha}\frac{1}{(pa)^{1/2}}\sum_{\mu\in\mathscr{P}\setminus\nu}\xi_i^{\mu}\frac{1}{(Na)^{1/2}}\sum_{j\in\Lambda\setminus i}\xi_j^{\mu}\sigma_j(0) \tag{44}$$

where $\alpha=p/N$. The properties of the initial configurations (41)–(43) assure us that the first term on the r.h.s of (44) converges, at least in distribution, to the random variable

$$\lim_{N\to\infty}\xi_i^{\nu}\frac{1}{Na}\sum_{j\in\Lambda\setminus i}\xi_j^{\nu}\sigma_j(0)\overset{\mathscr{D}}{=}\xi_i^{\nu}m^{\nu}(0) \tag{45}$$

This random variable is independent of the second term on the r.h.s. of (44). In this second term we apply the CLT to find

$$\lim_{N\to\infty}\frac{1}{(pa)^{1/2}}\sum_{\mu\in\mathscr{P}\setminus\nu}\xi_i^{\mu}\frac{1}{(Na)^{1/2}}\sum_{j\in\Lambda\setminus i}\xi_j^{\mu}\sigma_j(0)\overset{\mathscr{D}}{=}[a(0)]^{1/2}\mathscr{N}(0,1) \tag{46}$$

Therefore, in the limit $N\to\infty$ the local field is the sum of two independent random variables, i.e.,

$$h_i(\boldsymbol{\sigma}(0))\equiv\lim_{N\to\infty}h_i(\boldsymbol{\sigma}_{\Lambda}(0))\overset{\mathscr{D}}{=}\xi_i^{\nu}m^{\nu}(0)+[\alpha a(0)]^{1/2}\mathscr{N}(0,1) \tag{47}$$

This allows us to compute the probability distribution of the local field at $t=0$,

$$F_{h(0)}(x) \equiv \Pr\{h_i(\boldsymbol{\sigma}(0)) \leqslant x\}$$
$$= \sum_{s_k \in \mathscr{S}} \Pr\{h_i(\boldsymbol{\sigma}(0)) \leqslant x \mid \xi_i^v\} \Pr\{\xi_i^v = s_k\} \tag{48}$$

as a function of the probabilities $\Pr\{\xi_i^v = s_k \in \mathscr{S}\}$. For the latter one could take, e.g., in analogy with (1), a uniform distribution. Recalling (36), (42), and (43), the activity and the main overlap for the configuration $\boldsymbol{\sigma}(1)$ can then be obtained using

$$a(1) = \sum_{s_k \in \mathscr{S}} (s_k)^2 \Pr\{\sigma_i(1) = s_k\} \tag{49}$$

$$m^v(1) = \frac{1}{a} \sum_{s_k, s_{k'} \in \mathscr{S}} s_k s_{k'} \Pr\{\sigma_i(1) = s_{k'} \mid \xi_i^v = s_k\} \Pr\{\xi_i^v = s_k\} \tag{50}$$

For the gain function (37), e.g., we arrive at the following results for the probabilities needed in (49), (50):

$$\Pr\{\sigma_i(1) = +1\} = \int_{2b(s_{Q-1} + (Q-1)^{-1})}^{+\infty} dF_{h(0)}(x)$$
$$\Pr\{\sigma_i(1) = s_k\} = \int_{2b(s_k - (Q-1)^{-1})}^{2b(s_k + (Q-1)^{-1})} dF_{h(0)}(x)$$
$$\Pr\{\sigma_i(1) = -1\} = \int_{-\infty}^{2b(s_2 - (Q-1)^{-1})} dF_{h(0)}(x) \tag{51}$$

and

$$\Pr\{\sigma_i(1) = +1 \mid \xi_i^v = s_k\}$$
$$= \Pr\left\{1 - (Q-1)^{-1} \leqslant \frac{1}{2b} \{s_k m^v(0) + [\alpha a(0)]^{1/2} \mathscr{N}(0,1)\}\right\} \tag{52}$$

$$\Pr\{\sigma_i(1) = s_{k'} \neq \pm 1 \mid \xi_i^v = s_k\}$$
$$= \Pr\left\{s_{k'} - (Q-1)^{-1} < \frac{1}{2b} \{s_k m^v(0) + [\alpha a(0)]^{1/2} \mathscr{N}(0,1)\}\right.$$
$$\left. < s_{k'} + (Q-1)^{-1}\right\} \tag{53}$$

$$\Pr\{\sigma_i(1) = -1 \mid \xi_i^v = s_k\}$$
$$= \Pr\left\{\frac{1}{2b} \{s_k m^v(0) + [\alpha a(0)]^{1/2} \mathscr{N}(0,1)\} < -1 + (Q-1)^{-1}\right\} \tag{54}$$

The evolution equations at nonzero temperature can be derived by making use of the same method as in the Q-state Potts case in Section 2.2. Then the analogue of (27) reads

$$\Pr\{\sigma_i(1)=s_k\,|\,\boldsymbol{\sigma}(0)\} = \Pr\left\{\bigcap_{s_l\in\mathscr{S}\setminus s_k}\{\phi_i(l)<[\varepsilon_i(s_l\,|\,\boldsymbol{\sigma}(0))-\varepsilon_i(s_k\,|\,\boldsymbol{\sigma}(0))]\}\right\} \tag{55}$$

where $\varepsilon_i(s\,|\,\boldsymbol{\sigma}(0))$ is defined by (34) and the ϕ are distributed according to (26).

At this point it is important to remark that in order to calculate the distribution of the configuration $\{\sigma_i(1)\}$, $i\in\Lambda$, it is sufficient to know the main overlap $m^{\nu}(0)$, the activity $a(0)$, and the distribution $\Pr\{\xi_i^{\nu}\}$ of the stored patterns.

Although the formulas (49)–(55) are completely determined, they are in a less explicit form than those for the Potts model [recall (29)–(30)] because here the distribution of the patterns is not supposed to be uniform, but it is only required to satisfy $E[\xi_i^{\mu}]=0$. Therefore, any calculation heavily depends on a full *a priori* distribution of the patterns. However, for the nontrivial case $Q=3$, $E[\xi_i^{\mu}]=0$, the parameter $\mathrm{Var}[\xi_i^{\mu}]=a$ uniquely determines the distribution of the patterns: for $\{\xi_i^{\mu}\}$, $i\in\Lambda$, $\mu\in\mathscr{P}$ i.i.d.r.v., $\Pr\{\xi_i^{\mu}=\pm1\}=a/2$ and $\Pr\{\xi_i^{\mu}=0\}=1-a$.

In the next section we explicitly study and compare the three-state Ising network and the three-state Potts network in the limit of extreme asymmetric dilution. Indeed, as we will argue, in this limit the one-step parallel dynamics is the exact dynamical evolution.

4. EXTREMELY DILUTED ASYMMETRIC THREE-STATE NETWORKS

The networks we have considered in the foregoing sections are fully connected and symmetric. The fundamental difficulty of the treatment of parallel dynamics in such systems is the strong feedback.[29]

To avoid this difficulty, we consider highly diluted asymmetric versions of the models above. This class of neural networks was introduced in connection with $Q=2$ Ising models.[30]

Let $\{c_{ij}=0,1\}$, $i,j\neq i\in\Lambda$, be i.i.d.r.v. with distribution $F_c^{(\Lambda)}\equiv\Pr\{c_{ij}=1\}=c/N$, $c>0$. The diluted Potts glass neural network is then defined via the random synaptic coupling matrix

$$J_{ij}^{kl}(c)\equiv\frac{c_{ij}}{c}NJ_{ij}^{kl},\qquad i,j\neq i\in\Lambda;\quad k,l\in\mathscr{Q} \tag{56}$$

In the limit $N \to \infty$ one obtains an infinite extremely diluted network. Two important properties of this network are essential for a further analysis.[30,31]

The first property is the high asymmetry of the connections,

$$\Pr\{c_{ij} = c_{ji}\} = \left(\frac{c}{N}\right)^2$$

$$\Pr\{c_{ij} = 1 \wedge c_{ji} = 0\} = \frac{c}{N}\left(1 - \frac{c}{N}\right) \tag{57}$$

Therefore, the number of symmetric connections in the infinite configuration $\mathbf{c} = \{c_{ij}\}$, $i, j \neq i \in \mathbb{N}$, is finite with probability one, i.e., almost all connections of the graph $G_{\mathbb{N}}(\mathbf{c}) = \{(i, j): c_{ij} = 1, i, j \neq i \in \mathbb{N}\}$ are directed: $c_{ij} \neq c_{ji}$.

The second property in the limit of extreme dilution is a directed local tree structure of the graph $G_{\mathbb{N}}(\mathbf{c})$. By the arguments above the probability $F_k^{(\Lambda)}(c)$ that k connections are directed toward a given site $i \in \Lambda$ is

$$F_k^{(\Lambda)}(c) \equiv \Pr\{k = |T_i^{(\text{in})}|\} = \frac{N!}{k!\,(N-k)!}\left(\frac{c}{N}\right)^k \left(1 - \frac{c}{N}\right)^{N-k} \tag{58}$$

where $T_i^{(\text{in})} = \{c_{ji} = 1, j \in \Lambda \backslash i\}$ is the in-tree for i and $|T_i^{(\text{in})}|$ its cardinality. This probability is equal to $\Pr\{k = |T_i^{(\text{out})}| = |\{c_{ij} = 1, j \in \Lambda \backslash i\}|\}$ for connections directed outward at a given site $i \in \Lambda$. In the limit of extreme dilution we get a Poisson distribution:

$$\lim_{N \to \infty} F_k^{(\Lambda)}(c) = \frac{c^k}{k!} e^{-c} \tag{59}$$

Hence, the mean value of the number of in (out) connections for any site $i \in \Lambda$ is $E[|T_i^{(\text{in})(\text{out})}|] = c$. The probability that two sites i and i' have site j as a common ancestor is obviously equal to c/N. From $E[|T_i^{(\text{in})}|] = c$ it follows that after t time steps the cardinality of the cluster of ancestors for site i will be of the order of c^t. The same is valid for site i'. Therefore, the probability that the sites i and i' have disjoint clusters of ancestors approaches $(1 - c^t/N)^{c^t} \simeq \exp(-c^{2t}/N)$ for $N \gg 1$.

Summarizing, we find that in the limit of extreme dilution:

(i) Almost all (i.e., with probability 1) feedback loops in $G_{\mathbb{N}}(\mathbf{c})$ are eliminated.

(ii) With probability one, any finite number of neurons have disjoint clusters of ancestors.

It would be interesting to study less severe types of dilution, e.g., Hamiltonian cycles,[33] in which case the feedback is probably not suppressed completely.

We now return to the parallel dynamics (4) for the extremely diluted network $G_{\mathbb{N}}(\mathbf{c})$ in the limit $c \to \infty$. This means that we have infinite average connectivity, allowing us to store infinitely many patterns $p = \alpha c$.

Since, as argued above, there is no feedback in this system, the states of the neurons from different $T_i^{(\text{in})}$ will remain independent in time: they receive their input from disjoint trees. So the configuration $\{\sigma_i(1)\}$, $i \in \mathbb{N}$, is again a collection of i.i.d.r.v. correlated with only one pattern ξ^ν. Hence Eqs. (29)–(30) for the Potts neurons and Eqs. (49)–(55) for the Q-Ising neurons remain valid at any moment t.

Restricting ourselves to $Q = 3$ neurons (29), we get in the Potts case

$$\Pr\{\sigma_i(t+1) = \xi_i^\nu \mid \boldsymbol{\sigma}(t)\}$$

$$= \int_{R^1} Dy \int_{R^2} Dx_1\, Dx_2 \times \left(1 + \sum_{l=1}^{2} \exp\{-\beta[3m^\nu(t) + (6\alpha)^{1/2}(y - x_l)]\}\right)^{-1} \tag{60}$$

$$m^\nu(t+1) = 3 \Pr\{\sigma_i(t+1) = \xi_i^\nu \mid \boldsymbol{\sigma}(t)\} - 1 \tag{61}$$

$$m^\mu(t+1) = 0, \qquad \mu \neq \nu \tag{62}$$

such that the corresponding fixed-point equation has the form

$$m^* = 3\left\{\int_{R^1} Dy \int_{R^2} Dx_1\, Dx_2 \times \left(1 + \sum_{l=1}^{2} \exp\{-\beta[3m^* + (6\alpha)^{1/2}(y - x_l)]\}\right)^{-1} - \frac{1}{3}\right\} \tag{63}$$

where $m^* = \lim_{t \to \infty} m^\nu(t)$. We also recall [see (25)] that $a^\mu(1) = a^\mu(0) = Q - 1$, implying that $a^\mu(t) = Q - 1$ is constant. The zero-temperature version of these equations can easily be written down recalling (19)–(24). For the fixed-point equation we obtain

$$m^* = 3\left\{\int_{-\infty}^{+\infty} Dy \left[\Phi\left(\frac{3}{2\alpha}\right)^{1/2} m^* + y\right)\right]^2 - \frac{1}{3}\right\} \tag{64}$$

For the $Q = 3$ Ising network with a gain function of the type (37) with $s_1 = -1$, $s_2 = 0$, $s_3 = +1$, we arrive, after some straightforward calculations, at the following results for the activity and the overlap:

$$a(t+1)=\int_{-\infty}^{+\infty} Dz\,[aU_\beta(m^\nu(t)-[\alpha a(t)]^{1/2}z)$$
$$+(1-a)\,U_\beta([\alpha a(t)]^{1/2}z)] \tag{65}$$

$$m^\nu(t+1)=\int_{-\infty}^{+\infty} Dz\,V_\beta(m^\nu(t)-[\alpha a(t)]^{1/2}z) \tag{66}$$

$$m^\mu(t)=0,\ \mu\neq\nu \tag{67}$$

where

$$U_\beta(x)\equiv\frac{\cosh(\beta x/2)}{\gamma+\cosh(\beta x/2)} \tag{68}$$

$$V_\beta(x)\equiv\frac{\sinh(\beta x/2)}{\gamma+\cosh(\beta x/2)} \tag{69}$$

$$\gamma\equiv\frac{1}{2}e^{\beta b/2} \tag{70}$$

The fixed-point equations can be read off from (65)–(67) by using again $m^*=\lim_{t\to\infty}m^\nu(t)$ and $a^*=\lim_{t\to\infty}a(t)$. In the zero-temperature case Eqs. (65)–(67) reduce to

$$a(t+1)=a-\frac{a}{2}\left[\operatorname{erf}\left(\frac{b+m^\nu(t)}{[\alpha a(t)]^{1/2}}\right)+\operatorname{erf}\left(\frac{b-m^\nu(t)}{[\alpha a(t)]^{1/2}}\right)\right]$$
$$+(1-a)\left[1-\operatorname{erf}\left(\frac{b}{[\alpha a(t)]^{1/2}}\right)\right] \tag{71}$$

$$m^\nu(t+1)=\frac{1}{2}\left[\operatorname{erf}\left(\frac{b+m^\nu(t)}{[\alpha a(t)]^{1/2}}\right)-\operatorname{erf}\left(\frac{b-m^\nu(t)}{[\alpha a(t)]^{1/2}}\right)\right] \tag{72}$$

These zero-temperature equations have also been studied in ref. 2. A fully connected version of this model has been simulated extensively in ref. 1.

To end this section, we remark that the derivation of the $Q=3$ results (65)–(70) and a similar calculation for analogous $Q=4$ models suggest the following general formulas for the Q-Ising case:

$$a(t+1)$$
$$=\int_{-\infty}^{+\infty} Dz\left\langle\!\!\left\langle\frac{\sum_{s\in\mathscr{S}}s^2\exp[-\frac{1}{2}\beta s(\xi^\nu m^\nu(t)+[\alpha a(t)]^{1/2}z-bs)]}{\sum_{s\in\mathscr{S}}\exp[-\frac{1}{2}\beta s(\xi^\nu m^\nu(t)+[\alpha a(t)]^{1/2}z-bs)]}\right\rangle\!\!\right\rangle \tag{73}$$

$$m^\nu(t+1)$$
$$=\frac{1}{a}\int_{-\infty}^{+\infty} Dz\left\langle\!\!\left\langle\frac{\sum_{s\in\mathscr{S}}\xi^\nu s\exp[-\frac{1}{2}\beta s(\xi^\nu m^\nu(t)+[\alpha a(t)]^{1/2}z-bs)]}{\sum_{s\in\mathscr{S}}\exp[-\frac{1}{2}\beta s(\xi^\nu m^\nu(t)+[\alpha a(t)]^{1/2}z-bs)]}\right\rangle\!\!\right\rangle \tag{74}$$

where $\langle\langle \cdots \rangle\rangle$ denotes the average over the distribution of the patterns. For zero temperature these formulas reduce to

$$a(t+1) = \int_{-\infty}^{+\infty} Dz \left\langle\!\!\left\langle g^2(\xi^\nu m^\nu(t) + [\alpha a(t)]^{1/2} z) \right\rangle\!\!\right\rangle \tag{75}$$

$$m^\nu(t+1) = \int_{-\infty}^{+\infty} Dz \frac{1}{a} \left\langle\!\!\left\langle \xi^\nu g(\xi^\nu m^\nu(t) + [\alpha a(t)]^{1/2} z) \right\rangle\!\!\right\rangle \tag{76}$$

We will come back to these formulas at a future occasion.

5. DISCUSSION OF THE NUMERICAL RESULTS

We now turn to an explicit discussion of the extremely diluted $Q = 3$ networks. We start with the Potts model described by the evolution equations (63). The properties of the solutions of these equations are best summarized in a temperature–capacity diagram (Fig. 1) and an overlap bifurcation diagram (Fig. 2). In Fig. 1 the line T_M indicates the temperature at which the retrieval states disappear discontinuously as a function of the loading capacity α. This happens above the line T_0, the temperature at which the zero solution becomes stable. This behavior is reminiscent of a first-order transition, as can be read off from Fig. 2. Before studying the latter, we still remark that the maximal loading capacity is $\alpha_c = 1.226$ in comparison with $\alpha_c = 2/\pi = 0.637$ for the extremely diluted Hopfield model.[30] For the fully connected $Q = 3$ Potts model with sequential dynamics a mean-field theory approach leads to $\alpha_c = 0.414$[15] versus $\alpha_c = 0.138$ for the Hopfield model.[34]

Figure 2 presents the overlap as a function of the temperature. First, we see that the $\alpha = 0$ (finite number of patterns) behavior completely coin-

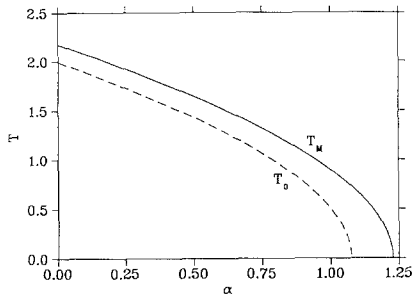

Fig. 1. The temperatures T_M (full line) for the retrieval state and T_0 (dashed line) for the zero state as a function of the loading α for the $Q = 3$ Potts network.

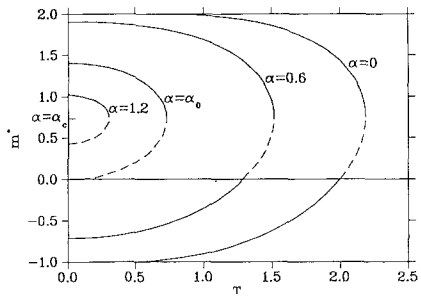

Fig. 2. The overlap m^* as a function of the temperature T for different values of the loading α for the $Q=3$ Potts network. The full line denotes a stable solution. $\alpha_0=1.073$, $\alpha_c=1.226$, $m^*=0.733$ at α_c.

cides with that found (again by a mean-field theory approach) for the fully connected $Q=3$ network with sequential dynamics.[35] There are three solutions of the fixed-point equations (63): a retrieval state, a zero state, and a state with negative overlap. The retrieval state is stable from $T=0$ up to $T=2.185$, the negative overlap state is stable from $T=0$ up to $T=T_0=2$, and the zero state is stable for $2\leqslant T<\infty$. For this case nearly maximal overlap $m^*=2$ is attained over some temperature region.

When increasing α the values of the overlaps become smaller. For $0<\alpha<\alpha_0=1.074$ the situation is analogous to the $\alpha=0$ case: there is a critical temperature $T_0(\alpha)$ under which the negative overlap state is stable and the zero state unstable. Above $T_0(\alpha)$, besides the retrieval state, also the zero state is stable. (This is in contrast to the fully connected sequential case, where for $\alpha\neq 0$ there are five solutions of the fixed-point equations but only the retrieval state is stable.[35]) For $\alpha_c>\alpha>\alpha_0$ only the retrieval state and the zero state are stable. At critical capacity α_c the value of the

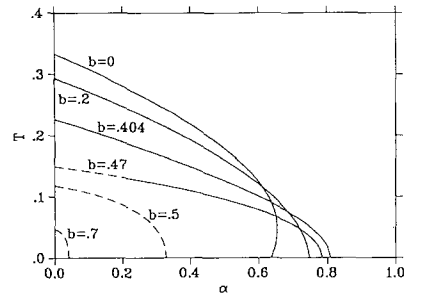

Fig. 3. A (T,α) diagram for different values of the parameter b for the $Q=3$ Ising network with $a=2/3$. The full line denotes a second-order transition, the dashed line a first-order one.

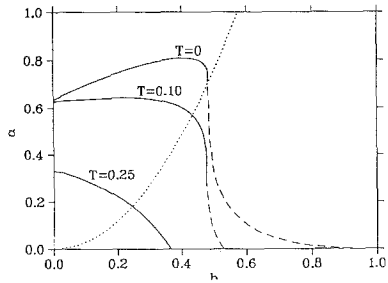

Fig. 4. The loading α as a function of the parameter b for different values of the temperature T for the $Q=3$ Ising network with $\alpha=2/3$. Full and dashed lines are as in Fig. 3. Below the dotted line there are no stable ($m^*=0$, $a^*>0$) solutions for $T=0$.

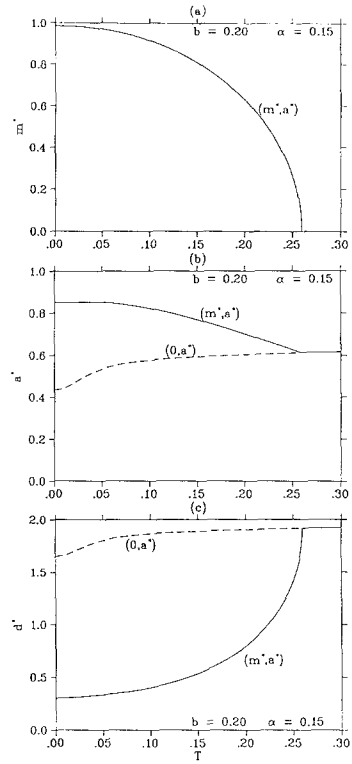

Fig. 5. The retrieval properties for the $Q=3$ Ising network with $a=2/3$ as a function of the temperature T for $b=0.2$ and $\alpha=0.15$: (a) the overlap m^*, (b) the activity a^*, and (c) the Hamming distance d^* for both the retrieval and the sustained activity solutions. The full line denotes a stable solution.

overlap at $T=0$ is $m^*=0.733$ versus $m^*=1.906$ for the fully connected sequential model.

Next we consider the $Q=3$ Ising model. Solving the evolution equations (65)–(70) is now more involved because the time evolution of the overlap and the activity are coupled. We find three different types of fixed points, depending on the parameters of the model, i.e., the pattern activity a and the parameter b controlling the relative importance of the states which are low in absolute value (i.e., the zero state in the $Q=3$ case): a retrieval fixed point ($m^*>0$, $a^*>0$), a sustained activity fixed point ($m^*=0$, $a^*>0$), and the zero fixed point ($m^*=0$, $a^*=0$).

Again the overall situation is best summarized in a temperature–capacity diagram (Fig. 3). The specific model that we consider has bias zero and pattern activity $a=2/3$ such that a pattern contains exactly as many zeros as plus or minus ones. As before the curves represent

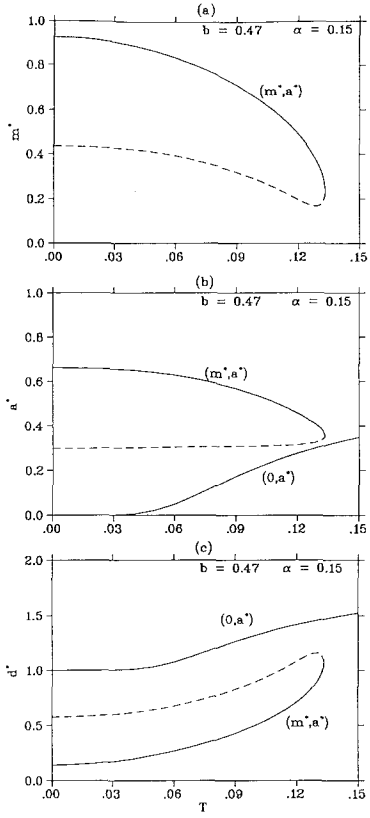

Fig. 6. The same as Fig. 5 for $b=0.47$ and $\alpha=0.15$.

the temperatures at which the retrieval solution(s) disappear(s). For $b=0$, implying that the neurons do not take on the zero state, we find the maximal capacity $\alpha_c = 2/\pi$ at $T=0$ in agreement with the extremely diluted Hopfield model.[30] For increasing b, α_c increases up to 0.809 for $b=0.404$. For greater values of b, α_c decreases again. This behavior at $T=0$ (see also ref. 2) is quite analogous to that for the fully connected sequential $Q=3$ model, derived using a mean-field theory approach.[4] Here the maximal capacity is an order of magnitude bigger (e.g., for $b=1/2$ we find $\alpha_c=0.329$ versus $\alpha_c=0.047$ for the connected model). In Fig. 4 we indicate how this behavior changes as a function of the temperature. For $T=0$ we can draw a curve below which no sustained activity solutions can be found. For $T\neq 0$ this type of solution can be found everywhere. We also remark that these solutions are independent of the pattern activity a. We clearly see in

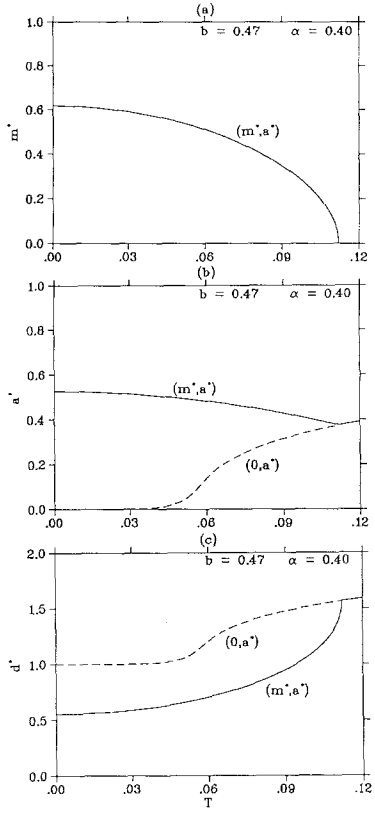

Fig. 7. The same as Fig. 5 for $b=0.47$ and $\alpha=0.40$.

this figure in which (b, α) region the network functions as an associative memory, i.e., where dynamically stable states $m^* > 0$ exist.

In Figs. 3 and 4 we have also indicated the nature of the disappearance of the retrieval solutions. For $0 < b < 0.462$ the transition is second order, for $0.462 \leqslant b \leqslant 0.483$ the order of this transition depends on α, and for $0.483 \leqslant b < 1$ the transition is completely first order.

These results are further illustrated in Figs. 5–7. There the overlap m, the activity a, and the Hamming distance d are depicted as a function of T for representative values of b and α. From these figures we also conclude that in order to evaluate the retrieval quality of the network, we have to take into account all three quantities m, a, and d. To fix our ideas, let us consider, e.g., the following example at $T = 0$. For fixed $\alpha = 0.15$, Figs. 5a and 6a show that m decreases as a function of a, whereas Figs. 5c and 6c indicate that also d decreases. This at first sight conflicting behavior is explained by the results on the activity a: in Fig. 6b this activity is closer to the pattern activity $a = 2/3$ than in Fig. 5b.

Hence, to determine the optimal retrieval quality of the network we have to take into account the values of the parameters m, a, and d together with the size of the basin of attraction. This is a different problem requiring further investigation.

ACKNOWLEDGMENTS

We thank E. Vanleeuw for discussions on the $Q = 4$ Ising results. This work has been supported in part by the Research Fund of the K.U. Leuven (grant OT/91/13). One of us (D.B.) thanks the Belgian National Fund for Scientific Research for support as a Research Director.

REFERENCES

1. C. Meunier, D. Hansel, and A. Verga, *J. Stat. Phys.* **55**:859 (1989).
2. J. S. Yedidia, *J. Phys. A* **22**:2265 (1989).
3. J. Stark and P. Bressloff, *J. Phys. A* **23**:1633 (1990).
4. H. Rieger, *J. Phys. A* **23**:L1273 (1990).
5. H. Rieger, in *Statistical Mechanics of Neural Networks*, L. Garrido, ed. (Springer, Berlin, 1990), p. 33.
6. D. Bollé, P. Dupont, and J. van Mourik, *Europhys. Lett.* **15**:893 (1991); *Physica A* **185**:357 (1992).
7. S. Mertens, H. M. Köhler, and S. Bös, *J. Phys. A* **24**:4941 (1991).
8. G. A. Kohring, *J. Stat. Phys.* **62**:563 (1991).
9. D. Bollé, P. Dupont, and B. Vinck, *J. Phys. A* **25**:2859 (1992).
10. M. Bouten and A. Engel, Basin of attraction in networks of multi-state neurons (1992).
11. A. J. Noest, *Phys. Rev. A* **38**:2196 (1988).
12. J. Cook, *J. Phys. A* **22**:2057 (1989).

13. F. Gerl, K. Bauer, and U. Krey, Learning wth Q-state clock neurons: Optimal storage capacity and Ada Tron-algorithm (1992).
14. A. E. Patrick, P. Picco, J. Ruiz, and V. A. Zagrebnov, *J. Phys. A* **24**:L637 (1991).
15. I. Kanter, *Phys. Rev. A* **37**:2739 (1988).
16. D. Bollé and F. Mallezie, *J. Phys. A* **22**:4409 (1989).
17. D. Bollé, P. Dupont, and F. Mallezie, in *Neural Networks and Spin Glasses*, W. K. Theumann and R. Köberle, eds. (World Scientific, Singapore, 1989), p. 151.
18. D. Bollé and P. Dupont, in *Statistical Mechanics of Neural Networks*, L. Garrido, ed. (Springer, Berlin, 1990), p. 365.
19. D. Bollé, P. Dupont, and J. van Mourik, *J. Phys. A* **24**:1065 (1991).
20. J. P. Nadal and A. Rau, *J. Phys. I* (Paris) **1**:1109 (1991).
21. T. Watkin, A. Rau, D. Bollé, and J. van Mourik, *J. Phys. I* (Paris) **2**:167 (1992).
22. G. M. Shim, D. Kim, and Y. M. Choi, *Phys. Rev. A* **45**:1238 (1992).
23. D. Bollé, P. Dupont, and J. Huyghebaert, *Phys. Rev. A* **46**:4194 (1992).
24. P. A. Ferrari, S. Martinez, and P. Picco, *J. Stat. Phys.* **66**:1643 (1992).
25. V. Gayrard, *J. Stat. Phys.* **68**:977 (1992).
26. H. Vogt and A. Zippelius, *J. Phys. A* **25**:2209 (1992).
27. A. E. Patrick and V. A. Zagrebnov, *J. Stat. Phys.* **63**:59 (1991).
28. V. A. Zagrebnov and A. S. Chvyrov, *Sov. Phys.-JETP* **68**:153 (1989).
29. E. Barkai, I. Kanter, and H. Sompolinsky, *Phys. Rev. A* **41**:590 (1990).
30. B. Derrida, E. Gardner, and A. Zippelius, *Europhys. Lett.* **4**:167 (1987).
31. R. Kree and A. Zippelius, in *Models of Neural Networks*, E. Domany, J. L. van Hemmen, and K. Schulten, eds. (Springer-Verlag, Berlin, 1991), p. 193.
32. A. N. Shiryayev, *Probability* (Springer-Verlag, New York, 1984).
33. B. Bollobas, *Graph Theory: An Introductory Course* (Springer-Verlag, New York, 1979).
34. D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Phys. Rev. Lett.* **55**:1530 (1985); *Ann. Phys.* (N.Y.) **173**:30 (1987).
35. D. Bollé, P. Dupont, and J. Huyghebaert, *Physica A* **185**:363 (1992).