Image evolution in Potts-glass neural networks

1989 J. Phys. A: Math. Gen. 22 4409

(http://iopscience.iop.org/0305-4470/22/20/017)

J. Phys. A: Math. Gen. 22 (1989) 4409-4418. Printed in the UK

Image evolution in Potts-glass neural networks

D Bollé† and F Mallezie

Instituut voor Theoretische Fysica, Universiteit Leuven, B-3030 Leuven, Belgium

Received 18 January 1989

Abstract. Neural networks of the Potts type, where neurons can occupy $q \geqslant 2$ discrete states, are considered for couplings which need not be symmetric. The time evolution of the macroscopic overlap between the instantaneous microscopic state of the system and the learned patterns is studied for small values of q and for a finite (small) number of patterns. Retrieval and limit-cycle behaviour in this type of model is discussed in some detail.

1. Introduction

Recently, the theory of neural networks has been extended to include neurons with more than two discrete states [1]. This has been done by viewing the neurons as Potts spins [2] having q possible discrete states. In [1] the capacity of storage and retrieval of information in this type of network has been discussed, mostly concentrating on the limit of zero temperature. In this discussion it was assumed that the synaptic connections J_{ij} are symmetric.

In the present paper we consider similar models extended to couplings which need not be symmetric. We study their general temporal development, thereby asking the specific question to what extent a built-in pattern can be retrieved at different noise levels in the system. In particular, we derive an evolution equation for the probability density of some macroscopic parameters measuring the overlap between the instantaneous microscopic state of the system with one of the built-in patterns. In this way we generalise the results obtained in [3] to these Potts networks.

We then discuss in detail the solutions of this differential equation for the overlap, for small values of q, for a few built-in patterns and for the number of neurons tending to infinity. Depending on the symmetry properties of the synaptic connections, we find the possibility of retrieval or limit-cycle behaviour.

The rest of this paper is organised as follows. In § 2 we describe the neural network models we study in the following. Section 3 derives the flow equation for the macroscopic overlap, starting from the master equation. In § 4 we study in detail the solution of this equation and its stability in the case of two, three and four Potts states with two and three built-in patterns for different symmetric and asymmetric synaptic connections. We compare our results with those obtained for the Hopfield model in [3, 4] and with some of the statements made for the Potts model with symmetric coupling and near zero temperature in [1].

† Onderzoeksdirecteur N.F.W.O., Belgium.

0305-4470/89/204409+10$02.50

2. The model

In this section we describe the model for the neural networks we study below. We follow and extend the recent work of Kanter [1].

We consider a network of N neurons. Each of the neurons can have more than two discrete states since it is viewed as a Potts spin having q possible values. The instantaneous configuration of all the spin variables at a given time describes then the state of such a network. The interactions among the neurons determine the dynamic evolution of the network in the phase space of q^N states.

The neurons are interconnected by a synaptic matrix of strength J_{ij} which determines the contributions of a signal fired by the jth presynaptic neuron to the postsynaptic potential which acts on the ith neuron. This contribution can either be positive (excitatory synapse) or negative (inhibitory synapse). The potential energy h_{σ_i} of neuron i which is in a Potts state σ_i is the sum of all postsynaptic potentials delivered to it in a time unit, i.e.

$$h_{\sigma_i} = -\sum_{j=1}^{N} \sum_{k,l=1}^{q} J_{ij}^{kl} m_{\sigma_i,k} m_{\sigma_j,l} \tag{2.1}$$

where $m_{\sigma_i,r}$ is an operator obeying Potts symmetry, given by

$$m_{\sigma_i,r} = q\delta_{\sigma_i,r} - 1. \tag{2.2}$$

Here q is the number of Potts states, σ_i a q-state Potts variable and δ the Kronecker delta. The potential energy (2.1) of the postsynaptic neuron i depends both on the state of neuron i and its neighbours and on their synaptic efficiency.

In this paper the J_{ij}^{kl} will not be assumed to be symmetric in (ij) and (kl). This means that the equilibrium properties of the system can no longer be described in terms of a Hamiltonian. At zero noise the stable states of the system are those configurations in which every spin σ_i is in a Potts state which gives a minimal value to h_{σ_i}. In the presence of noise there is a finite probability of having configurations other than these minima. This can be taken into account by introducing an effective temperature [4].

To build in the capacity for learning and memory in this network, its stable configurations must be correlated with the configurations determined by the learning process. To achieve this, we take the following form for the interaction, expressing the learning rule:

$$J_{ij}^{kl} = \frac{1}{q^2 N} \sum_{a,b=1}^{p} m_{k_i^a,k} A_{a,b} m_{k_j^b,l} \tag{2.3}$$

with the m given by (2.2). Here $A_{a,b}$ is a $p \times p$ matrix which need not be symmetric and the p sets $\{k_i^a\}$ are the patterns fixed by the learning process. The k_i^a are chosen as quenched random variables, assuming the values $1, 2, \ldots, q$ with equal probability. We remark that (2.3) tells us that every pair of neurons is connected. Furthermore the choice of random (anisotropic) interactions excludes the possibility that all spins occupy the same Potts state, i.e. there is no ferromagnetic order [5]. Using the learning rule (2.3), h_{σ_i} can be written as

$$h_{\sigma_i} = -\frac{1}{N} \sum_{j=1}^{N} \sum_{a,b=1}^{p} m_{k_i^a,\sigma_i} A_{a,b} m_{k_j^b,\sigma_j}. \tag{2.4}$$

This model will have the capacity of storage and retrieval of information if the emerging dynamically stable configurations $\{\sigma_i\}$ are correlated with the learned patterns

$\{k_i^a\}$. This is what we want to study in this paper by developing an equation specifying how an arbitrary initial state of the spins changes in time and to what extent one of the built-in patterns is approached. We will concentrate on the case of finite p and q for $N \to \infty$. In this study an important role is played by the order parameter Q^a, which is the macroscopic overlap with a pattern a:

$$Q^a = \frac{1}{N} \sum_i m_{\sigma_i, k_i^a}. \tag{2.5}$$

In a random spin configuration $Q^a = 0$. When $\sigma_i = k_i^a$, $i = 1, 2, \ldots, N$, which represents complete retrieval of pattern a, then $Q^a = q - 1$.

3. The flow equations for the overlap

In order to find out how an arbitrary initial state of the Potts spins $\{\sigma_i\}$, $i = 1, 2, \ldots, N$, of the network changes in time and to what extent, if at all, one of the built-in patterns is approached, we now derive a flow equation for the overlap.

We suppose that in our model sequential updating of the spins occurs, i.e. at each elementary time step only one spin is updated. The transition probability for this spin to go into another Potts state is given by the expressions

$$\begin{aligned} w(\sigma_i \to \sigma_i') &= \frac{1}{Z_i} \exp\left(-\frac{\beta}{2}(h_{\sigma_i'} - h_{\sigma_i})\right) \\ Z_i &= \sum_{\sigma_i = 1}^{q} \sum_{\sigma_i' \neq \sigma_i} \exp\left(-\frac{\beta}{2}(h_{\sigma_i'} - h_{\sigma_i})\right) \end{aligned} \tag{3.1}$$

which satisfy detailed balance. We then consider the probability $p(\boldsymbol{\sigma}, t)$ of observing the network in a state $\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_N\}$, $\sigma_i = 1, 2, \ldots, q$, at time t. The following master equation is valid:

$$\begin{aligned} \frac{\partial p(\boldsymbol{\sigma}, t)}{\partial t} = &\sum_i \sum_{\sigma_i'} w(\sigma_i' \to \sigma_i) p(\sigma_1, \ldots, \sigma_i', \ldots, \sigma_N; t) \\ &- \sum_i \sum_{\sigma_i'} w(\sigma_i \to \sigma_i') p(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N; t). \end{aligned} \tag{3.2}$$

From this basic equation we want to obtain an equation describing the evolution of the macroscopic overlap parameter. In order to do so we first note that the summation over the spin index i will be performed using the partition method introduced in [6]. The set of all indices is divided into subsets I_k which depend on the built-in patterns as follows:

$$\{i \leqslant N\} = \bigcup_{\boldsymbol{k}} I_{\boldsymbol{k}} \qquad I_{\boldsymbol{k}} = \{i \leqslant N \mid \boldsymbol{k} = \boldsymbol{k}_i\} \tag{3.3}$$

where $\boldsymbol{k}_i$ is the p-dimensional vector $\{k_i^1, k_i^2, \ldots, k_i^p\}$ for spin i. The number of p-dimensional vectors $\boldsymbol{k}$ is q^p, which is much smaller than the number N of vectors $\boldsymbol{k}_i$. Therefore the number of indices in each of the sets $I_{\boldsymbol{k}}$, denoted by $|I_{\boldsymbol{k}}|$, will almost be the same for all $\boldsymbol{k}$ if the k_i^a are chosen randomly. It is

$$|I_{\boldsymbol{k}}| = q^{-p} N + O(N^{1/2}). \tag{3.4}$$

Secondly, besides the (total) overlap Q^a, we introduce 'submagnetisations' for spins restricted to the set I_k, namely

$$\mu_k^\alpha(\boldsymbol{\sigma}) = \frac{1}{|I_k|} \sum_{i \in I_k} m_{\alpha,\sigma_i}. \tag{3.5}$$

We remark that $q^{-1}|I_k|(1+\mu_k^\alpha(\boldsymbol{\sigma}))$ denotes the number of spins in the set I_k for which the state $\sigma_i = \alpha$. The total overlap Q^a can then be written as

$$Q^a(\boldsymbol{\sigma}) = \frac{1}{q} \sum_k \frac{|I_k|}{N} \sum_\alpha \mu_k^\alpha(\boldsymbol{\sigma}) m_{k^a,\alpha}. \tag{3.6}$$

Thirdly, we define an operator F_i allowing us to change a Potts spin σ_i for all functions $\phi(\boldsymbol{\sigma})$:

$$\begin{aligned} &F_i : \phi(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N) \to F_i\phi(\boldsymbol{\sigma}) = \phi(\sigma_1, \ldots, F\sigma_i, \ldots, \sigma_N) \\ &F\sigma_i = \sigma_i \bmod q + 1. \end{aligned} \tag{3.7}$$

Consequently we have

$$\mu_k^\alpha(F_i^\nu \boldsymbol{\sigma}) = \begin{cases} \mu_k^\alpha(\boldsymbol{\sigma}) & \text{if } i \notin I_k \\ \mu_k^\alpha(\boldsymbol{\sigma}) + \dfrac{1}{|I_k|}(m_{\alpha,F^\nu\sigma_i} - m_{\alpha,\sigma_i}) & \text{if } i \in I_k \end{cases} \tag{3.8}$$

where ν indicates the power to which F_i is raised, i.e. $\nu = 1, 2, \ldots, q$ with $F^q = \mathbf{1}$.

In the rest of this section we work with the submagnetisations (3.5). We define the probability of finding the system at time t in a state with macroscopic parameters specified by the q^{p+1}-dimensional vector $\boldsymbol{\mu}(\boldsymbol{\sigma}) = \{\mu_k^\alpha(\boldsymbol{\sigma})\}$, $\alpha = 1, 2, \ldots, q$; $\boldsymbol{k} = \{k^1, k^2, \ldots, k^p\}$, $k^a = 1, 2, \ldots, q$ as

$$P(\boldsymbol{\mu}, t) = \sum_{\boldsymbol{\sigma}} p(\boldsymbol{\sigma}, t) \delta(\boldsymbol{\mu} - \boldsymbol{\mu}(\boldsymbol{\sigma})). \tag{3.9}$$

Using (3.2) and (3.7) we arrive at the following master equation for the probability (3.9):

$$\frac{\partial P(\boldsymbol{\mu}, t)}{\partial t} = \sum_{\boldsymbol{\sigma}} \sum_i \sum_{\nu=1}^{q-1} w(\sigma_i \to F^\nu \sigma_i) p(\boldsymbol{\sigma}, t)[\delta(\boldsymbol{\mu} - \boldsymbol{\mu}(F_i^\nu \boldsymbol{\sigma})) - \delta(\boldsymbol{\mu} - \boldsymbol{\mu}(\boldsymbol{\sigma}))]. \tag{3.10}$$

In the thermodynamic limit $N \to \infty$ the parameters $\boldsymbol{\mu}$ become continuous variables ($\mu_k^\alpha \in [-1, q-1]$) such that we can write

$$\begin{aligned} \frac{\partial P(\boldsymbol{\mu}, t)}{\partial t} = \sum_{\boldsymbol{\sigma}} \int \mathrm{d}\boldsymbol{\mu}'\, \delta(\boldsymbol{\mu}' - \boldsymbol{\mu}(\boldsymbol{\sigma})) p(\boldsymbol{\sigma}, t) \sum_i \sum_\nu \int \mathrm{d}\boldsymbol{x}\, \delta(-(\boldsymbol{\mu}(\boldsymbol{\sigma}) - \boldsymbol{\mu}(F_i^\nu \boldsymbol{\sigma}))) \\ \times w(\sigma_i \to F^\nu \sigma_i)[\delta(\boldsymbol{\mu} - \boldsymbol{\mu}' + \boldsymbol{x}) - \delta(\boldsymbol{\mu} - \boldsymbol{\mu}')]. \end{aligned} \tag{3.11}$$

To continue we need several observations. Firstly, recalling (3.1), (2.1) and (2.3) we infer that the transition rates w can be written as functions of the components of the q^{p+1}-dimensional vector $\boldsymbol{x}$ and the p-dimensional vector $\boldsymbol{Q}$ as follows:

$$w(\sigma_i \to F^\nu \sigma_i) = w((\boldsymbol{m}_{k_i,F^\nu\sigma_i} - \boldsymbol{m}_{k_i,\sigma_i}) \cdot A\boldsymbol{Q}) \tag{3.12}$$

with $\boldsymbol{m}_{k,\sigma}$ denoting the p-dimensional vector with components $m_{k^a,\sigma}$ and where, using (3.5)

$$\boldsymbol{m}_{k_i,F^\nu\sigma_i} - \boldsymbol{m}_{k_i,\sigma_i} = \frac{1}{q}|I_{k_i}| \sum_\alpha \boldsymbol{m}_{k_i,\alpha}(\mu_{k_i}^\alpha(F_i^\nu \boldsymbol{\sigma}) - \mu_{k_i}^\alpha(\boldsymbol{\sigma})). \tag{3.13}$$

Secondly, denoting the difference $[\boldsymbol{\mu}(\boldsymbol{\sigma})-\boldsymbol{\mu}(F_i^\nu\boldsymbol{\sigma})]$ for an arbitrary spin $i\in I_k$ in a Potts state σ_i by the q^{p+1}-dimensional vector $\boldsymbol{y}_k^{\alpha,\nu}$, it is straightforward to check that the $(\beta, \boldsymbol{k}')$ component of $\boldsymbol{y}_k^{\alpha,\nu}$ is only non-zero if $\boldsymbol{k}=\boldsymbol{k}'$.

Then the $\boldsymbol{x}$-dependent part of (3.11) can be written as

$$\sum_i\sum_\nu\int \mathrm{d}\boldsymbol{x}\,\delta(\boldsymbol{x}-(\boldsymbol{\mu}(\boldsymbol{\sigma})-\boldsymbol{\mu}(F_i^\nu\boldsymbol{\sigma})))G(\boldsymbol{x})=\sum_k\frac{|I_k|}{q}\sum_\alpha\sum_\nu(1+\mu_k^\alpha(\boldsymbol{\sigma}))G(\boldsymbol{y}_k^{\alpha,\nu}) \tag{3.14}$$

with

$$G(\boldsymbol{x})=w(\sigma_i\to F^\nu\sigma_i)[\delta(\boldsymbol{\mu}-\boldsymbol{\mu}'+\boldsymbol{x})-\delta(\boldsymbol{\mu}-\boldsymbol{\mu}')]. \tag{3.15}$$

This, together with the relation

$$\sum_{\boldsymbol{\sigma}}\delta(\boldsymbol{\mu}'-\boldsymbol{\mu}(\boldsymbol{\sigma}))p(\boldsymbol{\sigma},t)\mu_k^\alpha(\boldsymbol{\sigma})=\mu_k'^\alpha P(\boldsymbol{\mu}',t) \tag{3.16}$$

allows us to obtain for (3.11)

$$\frac{\partial P(\boldsymbol{\mu},t)}{\partial t}=\int \mathrm{d}\boldsymbol{\mu}'\,P(\boldsymbol{\mu}',t)\sum_k\frac{|I_k|}{q}\sum_\alpha(1+\mu_k'^\alpha)$$
$$\times\sum_\nu w_k(\alpha\to F^\nu\alpha)[\delta(\boldsymbol{\mu}-\boldsymbol{\mu}'+\boldsymbol{y}_k^{\alpha,\nu})-\delta(\boldsymbol{\mu}-\boldsymbol{\mu}')] \tag{3.17}$$

where $w_k(\alpha\to F^\nu\alpha)$ is the transition rate for $\alpha\to F^\nu\alpha$ for a spin in the set I_k. For every smooth function $\phi(\boldsymbol{\mu})$ we then get

$$\int \mathrm{d}\boldsymbol{\mu}\,\phi(\boldsymbol{\mu})\frac{\partial P(\boldsymbol{\mu},t)}{\partial t}$$
$$=\int \mathrm{d}\boldsymbol{\mu}'\,P(\boldsymbol{\mu}',t)\sum_k\frac{|I_k|}{q}\sum_\alpha(1+\mu_k'^\alpha)\sum_\nu w_k(\alpha\to F^\nu\alpha)[\phi(\boldsymbol{\mu}'-\boldsymbol{y}_k^{\alpha,\nu})-\phi(\boldsymbol{\mu}')]. \tag{3.18}$$

The last step in our derivation is an expansion in powers of $1/N$. Employing (3.4) and the observation about the vector $\boldsymbol{y}_k^{\alpha,\nu}$, made after (3.13), we find

$$\phi(\boldsymbol{\mu}'-\boldsymbol{y}_k^{\alpha,\nu})=\phi(\boldsymbol{\mu}')-\sum_\beta\left(\frac{\partial\phi(\boldsymbol{\mu})}{\partial\mu_k^\beta}\right)_{\boldsymbol{\mu}=\boldsymbol{\mu}'}\frac{1}{|I_k|}(m_{\beta,\alpha}-m_{\beta,F^\nu\alpha})+O(N^{-3/2}). \tag{3.19}$$

Inserting this result in (3.18) and doing a partial integration with respect to $\boldsymbol{\mu}'$, we arrive at

$$\int \mathrm{d}\boldsymbol{\mu}\,\phi(\boldsymbol{\mu})\frac{\partial P(\boldsymbol{\mu},t)}{\partial t}$$
$$=\int \mathrm{d}\boldsymbol{\mu}'\,\phi(\boldsymbol{\mu}')\sum_k\sum_\beta\frac{\partial}{\partial\mu_k'^\beta}\Bigg(P(\boldsymbol{\mu}',t)\frac{1}{q}\sum_\alpha(1+\mu_k'^\alpha)$$
$$\times\sum_\nu w_k(\alpha\to F^\nu\alpha)(m_{\beta,\alpha}-m_{\beta,F^\nu\alpha})\Bigg)+O(N^{-1/2}). \tag{3.20}$$

Since this equation holds for every smooth function ϕ it has to be valid for the integrands, leading to (after taking $N\to\infty$) a continuity equation for the density of a flow in the q^{p+1}-dimensional space. The corresponding equation for the flow itself is then

$$\frac{\mathrm{d}\mu_k^\beta}{\mathrm{d}t}=-\frac{1}{q}\sum_\alpha(1+\mu_k^\alpha)\sum_\nu w_k(\alpha\to F^\nu\alpha)(m_{\beta,\alpha}-m_{\beta,F^\nu\alpha}). \tag{3.21}$$

This equation is our final result. We remark that in the course of the derivation we have not used the specific form (3.1) for the transition probability w. Hence other expressions for w, satisfying detailed balance, could be employed.

Taking $q=2$, i.e. the Hopfield model, and using (3.1) and (3.6) it is straightforward to check that (3.21) leads to the flow equation for the total overlap Q^a written down in [3]. (See also [7].) For general q additional parameters have to be taken into account. For example, when $p=2$ it turns out that the q^3 equations (3.21) with w_k as in (3.1) can be reduced to the following flow equations in the three-dimensional space $(Q^1, Q^2, D \equiv \Sigma_k \mu^k_{\{k,k\}})$:

$$\begin{aligned}\frac{dQ^a}{dt} = \frac{1}{N_1}\Bigg\{ & 2q(q-1)\sinh\left(\frac{q}{2T}(R^a+R^b)\right) \\ & - D\left[(q-1)\exp\left(-\frac{q}{2T}(R^a+R^b)\right)+\exp\left(\frac{q}{2T}(R^a+R^b)\right)\right]\Bigg\} \\ & + \frac{1}{N_2}\Bigg\{ 2q(q-1)\left[\sinh\left(\frac{q}{2T}(R^a-R^b)\right)+(q-2)\sinh\left(\frac{q}{2T}R^a\right)\right] \\ & - q^2Q^a\left[\exp\left(-\frac{q}{2T}(R^a-R^b)\right)\right. \\ & \left.+(q-2)\exp\left(-\frac{q}{2T}R^a\right)+\exp\left(\frac{q}{2T}R^a\right)\right] \\ & - q^2Q^b\left[\exp\left(\frac{q}{2T}R^a\right)-\exp\left(\frac{q}{2T}(R^a-R^b)\right)\right] \\ & - D\left[2\sinh\left(\frac{q}{2T}(R^a-R^b)\right)-(q-2)\exp\left(-\frac{q}{2T}R^a\right)\right. \\ & \left.-2\exp\left(\frac{q}{2T}R^a\right)\right]\Bigg\} \end{aligned} \tag{3.22}$$

$$\begin{aligned}\frac{dD}{dt} = \frac{q^2}{N_1}\Bigg\{ & 2q(q-1)\sinh\left(\frac{q}{2T}(R^a+R^b)\right) \\ & - D\left[(q-1)\exp\left(-\frac{q}{2T}(R^a+R^b)\right)+\exp\left(\frac{q}{2T}(R^a+R^b)\right)\right]\Bigg\} \end{aligned} \tag{3.23}$$

where $a=1$, $b=2$ or $a=2$, $b=1$, $R^a=\sum_b A_{ab}Q^b$ and

$$N_1 = q^2(q-1)\left[2\cosh\left(\frac{q}{2T}(R^a+R^b)\right)+(q-2)\right]$$

$$\begin{aligned} N_2 = q^2\Bigg\{ & 2\cosh\left(\frac{q}{2T}(R^a-R^b)\right) \\ & + (q-2)\left[2\cosh\left(\frac{q}{2T}R^a\right)+2\cosh\left(\frac{q}{2T}R^b\right)+(q-3)\right]\Bigg\}. \end{aligned}$$

Finally we note that several generalisations of this derivation, similar to those in the Hopfield case [3] can be worked out. For example, a neuronal threshold can be included in the local field h_{σ_i}, a constant background coupling can be added to the connections J^{kl}_{ij}, biased patterns [8] can be considered by modifying the connections.

4. Retrieval and limit-cycle behaviour

In this section we discuss the numerical solutions of (3.21) and their stability as a function of the noise level (i.e. temperature) for the specific $q = 2, 3, 4$ state networks with two and three built-in patterns and three different forms of the matrix A appearing in the synaptic connection coefficients (2.3). From these solutions we calculate the overlap Q^a using (3.4) and (3.6) to study the retrieval properties of these networks. For a representative set of μ_k^α, our results are summarised in figures 1-5. In figures 1-3 the overlap flow diagrams projected on the (Q^1, Q^2) plane are shown.

For $p = 2$, $q = 2$ (see figure 1) we recover the results obtained in [3]. They are also a nice illustration of the extensive stability analysis given in [4] for the case of symmetric couplings. For completeness, and to be able to compare directly with results for higher p and q, we present a short discussion.

In the first row of diagrams we have taken A to be the unit matrix, corresponding to the standard Hebb connections. We can locate in the first diagram ($T = 0.1$) the $2p = 4$ Mattis states, thermodynamically equivalent to the ferromagnetic state, near the corners $(\pm 1, 0)$, $(0, \pm 1)$. They are local minima of the free energy for all $T < 1$. They move along the diagonals for increasing T to reach the origin $\boldsymbol{Q} = 0$ for $T = 1$. Furthermore, near the points $(\pm\frac{1}{2}, \pm\frac{1}{2})$ we locate the symmetric solutions discussed in [4] (corresponding to the stationary solutions of (3.21)) as unstable saddle points also going to the origin with increasing T. We see no asymmetric solutions.

In the second row of diagrams, corresponding to the second (symmetric) choice for A, we have an attractor near $(+\frac{1}{2}, -\frac{1}{2})$ and $(-\frac{1}{2}, +\frac{1}{2})$ also moving to the origin when

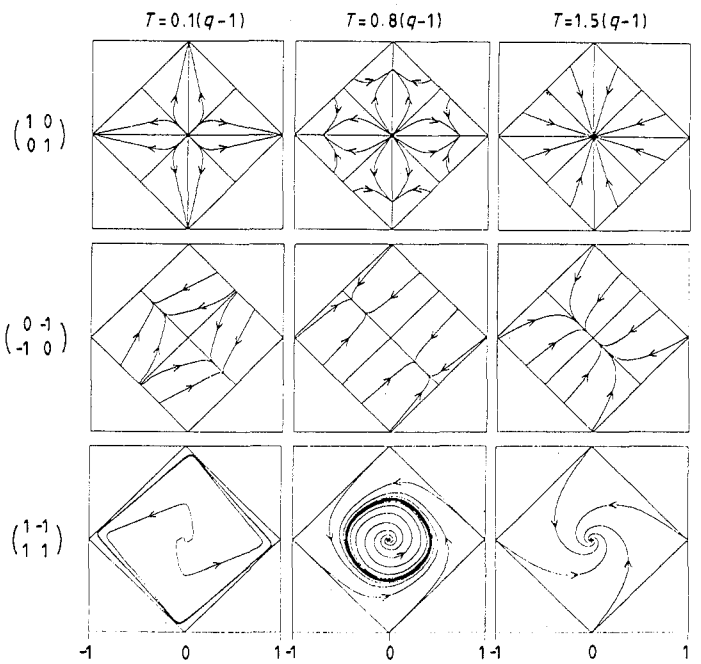

Figure 1 (see [3]). The flow lines in the (Q^1, Q^2) plane for $q = 2$, $p = 2$ for the choice of matrix A indicated in front of each row and at the temperature indicated above each column. In each diagram, the allowed values of Q are contained in the inner square with corners $(\pm 1, 0)$, $(0, \pm 1)$.

T increases. In the third row, where A is asymmetric, we find limit-cycle behaviour, with most of the time spent near the corners $(\pm 1, 0)$, $(0, \pm 1)$. Again the limit cycle contracts towards the origin with increasing temperature.

The conclusion is that for $T \geqslant 1$ the overlap vector $\boldsymbol{Q}$ approaches the origin in all cases, telling us that no correlation with the built-in patterns remains. The strongest correlation is found at low temperatures and in the third case the correlations switch back and forth between the learned patterns and their negatives.

For $p = 2$, $q = 3$ some results are shown in figure 2. At this point we have to explain the particular normalisation of the temperature, namely $T(q-1)$, introduced in order to get comparable temperatures in the different models. It turns out by looking at the first and second derivative of the free-energy density for the symmetric models (calculated, e.g., for the standard Hebb connections by using (7) of [1]) that the stationary solution $\boldsymbol{Q} = 0$ of (3.21) and (3.6) is an attractor for $T \geqslant q-1$. For the other models, the numerical solutions of (3.21) have the same property.

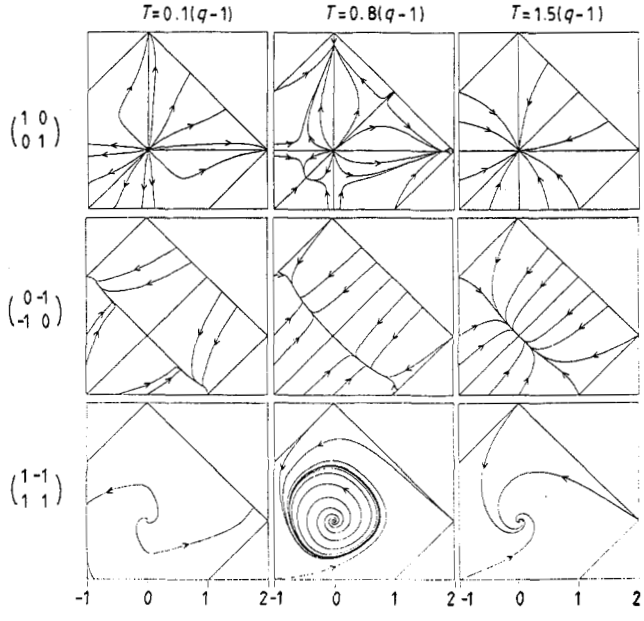

Figure 2. The flow lines in the (Q^1, Q^2) plane for $q = 3$, $p = 2$ for the choice of matrix A indicated in front of each row and at the temperature indicated above each column. In each diagram, the allowed values of $\boldsymbol{Q}$ are contained in the inner region with corners $(0, 2)$, $(2, 0)$, $(1, -1)$, $(-1, -1)$, $(-1, 1)$.

The general behaviour of the flow lines resembles the $q = 2$ case. However, some differences have to be noted, illustrating the remarks made by Kanter [1] in the case of symmetric couplings. In the first row of diagrams of figure 2, we observe the $p = 2$ Mattis states located near the corners $(2, 0)$ and $(0, 2)$ at low T and moving towards the origin with increasing T. From (3.22) and (3.23) it follows that the location of these Mattis states is given by

$$Q^a = 0 \qquad Q^b = \frac{(q-1)(\exp[(q/2T)Q^b] - \exp[-(q/2T)Q^b])}{(q-1)\exp[-(q/2T)Q^b] + \exp[(q/2T)Q^b]} \qquad D = qQ^b \quad (4.1)$$

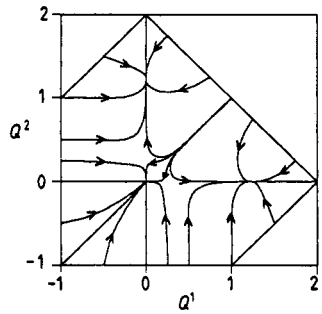

Figure 3. The flow lines in the (Q^1, Q^2) plane for $q=3$, $p=2$ with A the unit matrix and $T=2.1$.

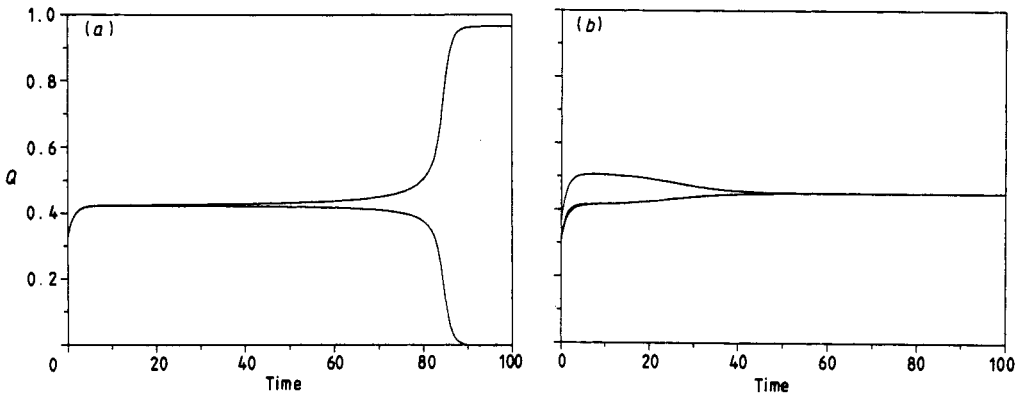

Figure 4. The overlap parameters Q^1, Q^2, Q^3 as a function of time for $q=2$, $p=3$ with A the unit matrix and at two different temperatures. In (a) $T=0.48$. The starting values are $Q^1=Q^2=0.333$, $Q^3=0.334$. The upper curve gives Q^3, the lower curve gives both Q^1 and Q^2. In (b) $T=0.4$. The initial values are $Q^1=0.31$, $Q^2=0.32$, $Q^3=0.37$.

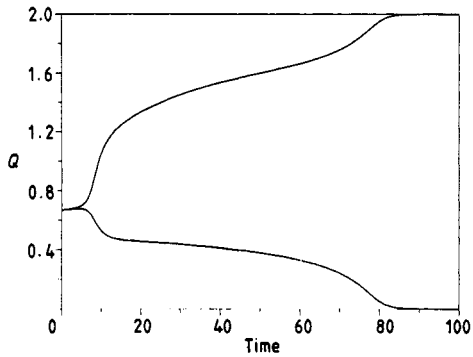

Figure 5. The overlap parameters Q^1, Q^2, Q^3 as a function of time for $q=3$, $p=3$ with A the unit matrix and $T=0.25$. The initial values are $Q^1=Q^2=0.6666$, $Q^3=0.6668$. The upper curve gives Q^3, the lower curve gives both Q^1 and Q^2.

for $a=1$, $b=2$ or $a=2$, $b=1$. For $q=3$, these equations have a solution $Q^b \neq 0$ for $T < T_3 = 2.18528(1)$. At T_3 the overlap jumps from $Q^b = 0.7545(5)$ to $Q^b = 0$. Furthermore, there is one symmetric saddle point solution at (1, 1) and an attractor at (−1, −1). The latter can be considered as the 'opposite' of the learned patterns (i.e. $\sigma_i \neq k_i^1$ and $\sigma_i \neq k_i^2$), so it does not represent recognition. They both move to the origin when T increases. Finally at $T = q-1 = 2$ this (−1, −1) attractor reaches the origin before the Mattis states do. So both the origin and the Mattis states are local minima of the free energy for $2 \leqslant T \leqslant T_3$ (see figure 3). This does not occur for $q=2$.

In the second row of diagrams of figure 2 we have again an attractor near (1, 0) and (0, 1) at low T, moving to the origin when T increases. In the third row we find limit-cycle behaviour analogous to the $q=2$ case. Again the conclusion is that in all cases of figure 2, the overlap becomes zero for $T \geqslant 2$. The strongest correlation is found at low T.

For $p=2$, $q=4$ we found qualitatively similar results, so we do not show any diagrams here. Again the overlap jumps from $Q^b = 1.519(3)$ to $Q^b = 0$ at $T_4 = 3.72816(1)$. Both the origin and the Mattis states are local minima of the free energy for $q-1=3 \leqslant T \leqslant T_4$.

For $p=3$ we have restricted ourselves here to models with the standard symmetric connections, illustrating in particular the behaviour of the symmetric solutions. For $q=2$ figures 4(a) and 4(b) show that for $T=0.40$ a nearly symmetric state is attracted towards a symmetric state while for $T=0.48$ such a state evolves to a Mattis state. This is in agreement with the observation in [3] that a symmetric state with three equal components becomes locally stable below $T=0.461$. For $q=3$ we see that at low T (figure 5, $T=0.25$) a nearly symmetric state becomes a Mattis state. This shows that the symmetric solutions are unstable for low T, as remarked also in [1]. So figures 4 and 5 illustrate that any stable solution near $T \to 0$ must have one dominant macroscopic overlap in the $q=3$ model, contrary to the $q=2$ Ising case.

Acknowledgments

We thank R Dekeyser and P Dupont for critical discussions and a careful reading of the manuscript. We are also grateful to A C C Coolen and Th W Ruijgrok for informing us about their work prior to publication.

References

[1] Kanter I 1988 *Phys. Rev.* A **37** 2739
[2] Wu F Y 1982 *Rev. Mod. Phys.* **54** 235
[3] Coolen A C C and Ruijgrok Th W 1988 *Phys. Rev.* A **38** 4253
[4] Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev.* A **32** 1007
[5] Gross D J, Kanter I and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 304
Elderfield D and Sherrington D 1983 *J. Phys. C: Solid State Phys.* **16** L497
[6] van Hemmen J L, Grensing D, Huber A and Kühn R 1986 *Z. Phys.* B **65** 53
[7] van Hemmen J L 1989 *Dynamical Networks* ed W Ebeling and M Peschel (Berlin, GDR: Akademie-Verlag) to appear
[8] Amit D J, Gutfreund H and Sompolinsky H 1987 *Phys. Rev.* A **35** 2293