The mean-field theory of a Q-state neural network model

1989 J. Phys. A: Math. Gen. 22 2057

(http://iopscience.iop.org/0305-4470/22/12/011)

J. Phys. A: Math. Gen. **22** (1989) 2057-2067. Printed in the UK

The mean-field theory of a Q-state neural network model

J Cook

Department of Physics, University of Edinburgh, Edinburgh EH9 3JZ, UK

Received 26 October 1988

Abstract. The mean-field equations of a Q-state clock neural network model are derived in the replica-symmetric approximation using the replica method. These equations are studied for the cases $Q=2, 3, 4$ and $Q\to\infty$. It is shown that an infinite number of patterns can be stored in the network even in the limit $Q\to\infty$. The phase diagram and storage capacity of the network are calculated and the information content considered. Although the overlap achieved with a nominated pattern is shown to decrease as Q increases, the information stored in the network is shown to increase.

1. Introduction

In the field of neural networks there has been a great deal of research on the Hopfield model and related models with two-state variables (e.g. Amit *et al* 1985, 1987a, Lautrup 1988, Bruce *et al* 1987, Canning and Gardner 1988). Recently, consideration of neural networks with variables with more than two states has begun (Kanter 1988, Noest 1988). The aim of this work is to explore how the properties of a neural network vary when the variables have an increasing number of states.

One Hamiltonian with which this can be examined is an extension of the Q-state clock model and clock spin glass (Nobre and Sherrington 1986):

$$H=-\frac{1}{2N}\sum_{i\neq j}\sum_{\mu}\cos\frac{2\pi}{Q}[(n_i-\xi_i^\mu)-(n_j-\xi_j^\mu)] \qquad (1)$$

where N is the number of sites, n_i is the value of the Q-state variable at site i ($n_i=0$, $1, 2, \ldots, Q-1$), and $\{\xi_i^\mu\}$ is the set of nominated configurations, or patterns, to be stored in the network.

Each pattern is labelled by a superscript, $\mu=1, 2, \ldots, p$, where p is the total number of nominated patterns to be stored. The nominated patterns are chosen to be random (i.e. ξ_i^μ takes any value in the allowed set $\{0, 1, 2, \ldots, Q-1\}$ with equal probability, $1/Q$) and uncorrelated (in the sense that $\Sigma_i \cos(\xi_i^\mu-\xi_i^\lambda)\sim O(\sqrt{N})$ for all $\mu\neq\lambda$). It will be found useful to define a parameter α such that $\alpha=p/N$.

The Hamiltonian (1) has the attractive feature that for $Q=2$ it reduces to the Hopfield model. Also for $Q\to\infty$, where the variable becomes continuous, it is an extension of the x-y model. The Hamiltonian is invariant under the transformation $n_i\to n_i+k$, where k is an integer. Hence, if a configuration, η say, is stored, the $Q-1$ related configurations, η_k, are also stored, where

$$\eta_{k_i}=(\eta_i+k)_{\mathrm{mod}\, Q} \qquad k=1, 2, \ldots, Q-1. \qquad (2)$$

In § 2 the mean-field theory will be presented and the mean-field equations derived for general $Q(>2)$. In § 3 these equations will be solved for various values of Q for

0305-4470/89/122057+11\$02.50

the case in which an intensive number of nominated patterns is stored. The following sections deal in detail with the specific cases $Q=3$, 4 and $Q\to\infty$. Phase diagrams of the same general struture are obtained in each case. The information content of the networks is also considered. Finally § 7 contains some general comments and conclusions.

2. Mean-field theory

The mean-field theory is performed using the replica trick (Sherrington and Kirkpatrick 1975), i.e. the free energy per spin, f, is written as

$$f=\lim_{n\to 0}\lim_{N\to\infty} -\frac{1}{nN\beta}(\langle\langle Z^n-1\rangle\rangle) \tag{3}$$

where $\langle\langle\ldots\rangle\rangle$ denotes a quenched average over the nominated patterns, ξ^μ, and $\beta=1/T$. The method used follows closely that of Amit *et al* (1987a), to be referred to as (I). Familiarity with the ideas presented in (I) is assumed. For ease of notation only the case $Q>2$ will be considered. The remaining case, $Q=2$, has been studied in (I).

Using γ as a replica index:

$$Z^n=\mathrm{Tr}_{\{n_i^\gamma\}}\exp(-\beta H^\gamma) \tag{4}$$

where

$$H^\gamma=\frac{p}{2}-\frac{N}{2}\sum_{\mu,\gamma}\left[\left(\frac{1}{N}\sum_i\cos\frac{2\pi}{Q}(n_i^\gamma-\xi_i^\mu)\right)^2+\left(\frac{1}{N}\sum_i\sin\frac{2\pi}{Q}(n_i^\gamma-\xi_i^\mu)\right)^2\right] \tag{5}$$

where the first term in (5) subtracts off self-interaction terms.

Now, using the Gaussian integral formula

$$\exp\left(\frac{wx^2}{2}\right)=\int_{-\infty}^{+\infty}\exp(-\tfrac{1}{2}z^2)\exp(\sqrt{w}zx)\frac{\mathrm{d}z}{\sqrt{2\pi}}$$

$\langle\langle Z^n\rangle\rangle$ can be written as

$$\begin{aligned}\langle\langle Z^n\rangle\rangle=\exp(\beta np/2)(N\beta)^{np}\Bigg\langle\!\!\Bigg\langle \mathrm{Tr}_{\{n_i^\gamma\}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\prod_{\gamma,\mu}\left(\frac{\mathrm{d}a_\mu^\gamma\,\mathrm{d}b_\mu^\gamma}{2\pi}\right)\\ \times\exp\left(-\frac{\beta N}{2}\sum_{\mu,\gamma}(a_\mu^{\gamma^2}+b_\mu^{\gamma^2})\right)\\ \times\exp\left[\beta\sum_{\mu,\gamma}\left(a_\mu^\gamma\sum_i\cos\frac{2\pi}{Q}(n_i^\gamma-\xi_i^\mu)+b_\mu^\gamma\sum_i\sin\frac{2\pi}{Q}(n_i^\gamma-\xi_i^\mu)\right)\right]\Bigg\rangle\!\!\Bigg\rangle.\end{aligned} \tag{6}$$

It can be shown within the saddle point approximation, which is imposed later, that a_μ^γ and b_μ^γ can be identified as

$$a_\mu^\gamma=\frac{1}{N}\left\langle\!\!\left\langle\sum_i\left\langle\cos\frac{2\pi}{Q}(n_i^\gamma-\xi_i^\mu\right\rangle\right\rangle\!\!\right\rangle \tag{7}$$

$$b_\mu^\gamma=\frac{1}{N}\left\langle\!\!\left\langle\sum_i\left\langle\sin\frac{2\pi}{Q}(n_i^\gamma-\xi_i^\mu\right\rangle\right\rangle\!\!\right\rangle \tag{8}$$

where $\langle\ldots\rangle$ represents a thermal average.

These identifications allow the significance of these parameters to be determined. Consider the symmetry of the Hamiltonian commented upon in the introduction.

Suppose ξ^μ is one of the nominated patterns. This nominated pattern and the $Q-1$ shifted patterns, ξ_k^μ, defined by (2), will be stored with equal accuracy in the network. For a solution in which the system closely matches a nominated pattern, or a combination of nominated patterns, the parameters b_μ^γ will be of order $1/\sqrt{N}$. The parameters a_μ^γ then give a measure of the overlap of the system with the nominated patterns. These are the only category of solutions that will be considered in this paper. If the overlap of the system with a shifted pattern in sought then $(a_\mu^{\gamma^2}+b_\mu^{\gamma^2})$ will give a measure of the overlap and a_μ^γ and b_μ^γ must be considered to determine which of the shifted patterns is involved.

Assume that for a finite number of the nominated patterns, ν, a_μ^γ is of order 1. These a_μ^γ will be called macroscopic. There then remains an extensive number of microscopic overlaps, i.e. $a_\mu^\gamma \sim \mathrm{O}(1/\sqrt{N})$, for $\mu>\nu$. The nominated patterns with microscopic overlap will also have b_μ^γ microscopic. The mean-field theory then procedes in an entirely analogous manner to that contained in (I) to give

$$\begin{aligned}\langle\langle Z^n\rangle\rangle = (N\beta)^n\left(\frac{N\beta^2\alpha\mathrm{i}}{4}\right)^{n(n-1)/2} &\exp(-\beta np/2)\left\langle\!\!\left\langle \mathrm{Tr}_{\{n_i^\gamma\}}\int\int\int\int\prod_{\mu=1}^{\nu}\left(\frac{da_\mu^\gamma db_\mu^\gamma}{2\pi}\right)\right.\right.\\ &\times\prod_{\gamma<\varepsilon}\mathrm{d}r_{\gamma\varepsilon}\,\mathrm{d}q_{\gamma\varepsilon}\exp\left(-\frac{N\beta}{2}\sum_\gamma\sum_{\mu=1}^{\nu}(a_\mu^{\gamma^2}+b_\mu^{\gamma^2})\right)\\ &\times\exp\left[\beta\sum_{\gamma,i}\sum_{\mu=1}^{\nu}\left(a_\mu^\gamma\cos\frac{2\pi}{Q}(n_i^\gamma-\xi_i^\mu)+b_\mu^\gamma\sin\frac{2\pi}{Q}(n_i^\gamma-\xi_i^\mu)\right)\right]\\ &\times\exp\left(-\frac{N\beta^2\alpha}{4}\sum_{\gamma<\varepsilon}q_{\gamma\varepsilon}r_{\gamma\varepsilon}+\frac{\beta^2\alpha}{4}\sum_i\sum_{\gamma<\varepsilon}r_{\gamma\varepsilon}\cos\frac{2\pi}{Q}(n_i^\gamma-n_i^\varepsilon)\right)\\ &\left.\left.\times\exp\{-(p-\nu)\,\mathrm{Tr}\ln[(1-\tfrac{1}{2}\beta)\delta_{\alpha\beta}+\tfrac{1}{2}\beta q_{\alpha\beta}]\}\right\rangle\!\!\right\rangle_\nu\end{aligned} \tag{9}$$

where $\langle\langle\ldots\rangle\rangle_\nu$ denotes a quenched average over the remaining patterns, ξ^μ, $\mu<\nu$.

In this paper only replica-symmetric solutions will be considered, i.e.

$$a_\mu^\gamma = a_\mu \tag{10}$$

$$b_\mu^\gamma = b_\mu \tag{11}$$

$$r_{\gamma\varepsilon} = r_{\varepsilon\gamma} = r \tag{12}$$

$$q_{\gamma\varepsilon} = q_{\varepsilon\gamma} = q \tag{13}$$

where $\gamma\neq\varepsilon$.

Now the remaining trigonometric functions can be expanded and two Gaussian transformations made to allow the trace to be taken, giving

$$\begin{aligned}\langle\langle Z^n\rangle\rangle = (N\beta)^n\left(\frac{N\beta^2\alpha\mathrm{i}}{4}\right)^{n(n-1)/2} &\exp(-nN\beta^2\alpha r/8)\exp(-\beta np/2)\\ &\times\int\int\int\int\prod_{\mu=1}^{\nu}\left(\frac{\mathrm{d}a_\mu\,\mathrm{d}b_\mu}{2\pi}\right)\mathrm{d}r\,\mathrm{d}q\\ &\times\exp\left(-\frac{nN\beta}{2}\sum_{\mu=1}^{\nu}(a_\mu^2+b_\mu^2)-\frac{n(n-1)N\beta^2\alpha}{8}qr\right)\\ &\times\exp\{(p-\nu)\mathrm{Tr}\log[(1-\tfrac{1}{2}\beta)\delta_{\gamma\varepsilon}+\tfrac{1}{2}\beta q_{\gamma\varepsilon}]\}\\ &\times\exp nN\left\langle\!\!\left\langle\int\int\frac{\mathrm{d}z\,\mathrm{d}z'}{2\pi}\exp(-\tfrac{1}{2}z^2)\exp(-\tfrac{1}{2}z'^2)\ln X\right\rangle\!\!\right\rangle_\nu\end{aligned} \tag{14}$$

where

$$X = \mathrm{Tr}_n \exp\left(A \cos \frac{2\pi n}{Q} + B \sin \frac{2\pi n}{Q} \right) \tag{15}$$

$$A = \beta \left(\frac{\sqrt{\alpha r}}{2} z + \sum_{\mu=1}^{\nu} (a_\mu c^\mu - b_\mu s^\mu) \right) \tag{16}$$

$$B = \beta \left(\frac{\sqrt{\alpha r}}{2} z' + \sum_{\mu=1}^{\nu} (a_\mu s^\mu + b_\mu c^\mu) \right) \tag{17}$$

with

$$c^\mu = \cos 2\pi \xi^\mu / Q \qquad s^\mu = \sin 2\pi \xi^\mu / Q \qquad n = 0, 1, \ldots, Q-1.$$

Hence X completely contains the dependence upon Q (for $Q > 2$).

Evaluating the integrals by the method of steepest descents taking the thermodynamic limit and letting $n \to 0$ in the same way as in (I), the free energy per spin is found to be

$$f = \frac{1}{2} \sum_{\mu=1}^{\nu} (a_\mu^2 + b_\mu^2) + \frac{\alpha}{2} + \frac{\beta \alpha r}{8} (1-q) + \frac{\alpha}{\beta} \left(\ln[1 - \tfrac{1}{2}\beta(1-q)] - \frac{\beta q}{2[1 - \frac{1}{2}\beta(1-q)]} \right)$$
$$- \frac{1}{\beta} \left\langle\!\!\left\langle \int \int \frac{\mathrm{d}z\, \mathrm{d}z'}{2\pi} \exp(-\tfrac{1}{2} z^2) \exp(-\tfrac{1}{2} z'^2) \ln X \right\rangle\!\!\right\rangle_\nu . \tag{18}$$

The mean-field equations then follow from the saddle point conditions

$$\frac{\partial f}{\partial a_\mu} = \frac{\partial f}{\partial b_\mu} = \frac{\partial f}{\partial r} = \frac{\partial f}{\partial q} = 0. \tag{19}$$

They are

$$a_\mu = \frac{1}{\beta} \left\langle\!\!\left\langle \int \int \frac{\mathrm{d}z\, \mathrm{d}z'}{2\pi} \exp(-\tfrac{1}{2} z^2) \exp(-\tfrac{1}{2} z'^2) \frac{\partial \ln X}{\partial a_\mu} \right\rangle\!\!\right\rangle_\nu \tag{20}$$

$$b_\mu = \frac{1}{\beta} \left\langle\!\!\left\langle \int \int \frac{\mathrm{d}z\, \mathrm{d}z'}{2\pi} \exp(-\tfrac{1}{2} z^2) \exp(-\tfrac{1}{2} z'^2) \frac{\partial \ln X}{\partial b_\mu} \right\rangle\!\!\right\rangle_\nu \tag{21}$$

$$\beta(1-q) = \frac{8}{\beta} \left\langle\!\!\left\langle \int \int \frac{\mathrm{d}z\, \mathrm{d}z'}{2\pi} \exp(-\tfrac{1}{2} z^2) \exp(-\tfrac{1}{2} z'^2) \frac{\partial \ln X}{\partial r} \right\rangle\!\!\right\rangle_\nu$$
$$= 2\sqrt{\frac{\alpha}{r}} \left\langle\!\!\left\langle \int \int \frac{\mathrm{d}z\, \mathrm{d}z'}{2\pi} \exp(-\tfrac{1}{2} z^2) \exp(-\tfrac{1}{2} z'^2) \left(z \frac{\partial \ln X}{\partial A} + z' \frac{\partial \ln X}{\partial B} \right) \right\rangle\!\!\right\rangle_\nu \tag{22}$$

$$r = \frac{2q}{[1 - \frac{1}{2}\beta(1-q)]^2}. \tag{23}$$

Applying the saddle point conditions before the replica-symmetric ansatz is taken, the order parameters can be identified, yielding in addition to (7) and (8),

$$r_{\gamma\varepsilon} = \sum_{\mu=\nu}^{p} (a_\mu^\gamma a_\mu^\varepsilon + b_\mu^\gamma b_\mu^\varepsilon) \tag{24}$$

$$q_{\gamma\varepsilon} = \frac{1}{N} \left\langle\!\!\left\langle \sum_i \left\langle \cos \frac{2\pi}{Q} (n_i^\gamma - n_i^\varepsilon) \right\rangle \right\rangle\!\!\right\rangle_\nu . \tag{25}$$

Having derived the general mean-field equations, (20)-(23), the remainder of the paper will consider their solution for various cases. First the limit $\alpha=0$ will be discussed. This is the case in which an intensive number of nominated patterns is stored. Then the storage of an extensive number of patterns will be considered for the cases $Q=3$, 4 and $Q\to\infty$.

3. Storage of an intensive number of patterns

In all that follows only solutions with $a_\mu=0$ for $\mu=2, 3, \ldots, \nu$ and $b_\mu=0$ for all μ will be considered. These include the true memory or retrieval states.

In the limit $\alpha\to 0$

$$A\to\beta\sum_{\mu=1}^{\nu}(a_\mu c^\mu - b_\mu s^\mu) \tag{26}$$

$$B\to\beta\sum_{\mu=1}^{\nu}(a_\mu s^\mu + b_\mu c^\mu). \tag{27}$$

In this limit the mean-field equations reduce to one equation for the order parameter $a_1\equiv a$. For example,

$$Q=2:\quad a=\tanh\beta a \tag{28}$$

$$Q=3:\quad a=\frac{\sqrt{3}}{4}\left\langle\!\!\left\langle s\tanh\frac{\sqrt{3}}{2}\beta as+\left(\sqrt{3}c-s\tanh\frac{\sqrt{3}}{2}\beta as\right)\times\tanh\left[\frac{3ac\beta}{4}-\tfrac{1}{2}\ln\left(2\cosh\frac{\sqrt{3}}{2}\beta as\right)\right]\right\rangle\!\!\right\rangle \tag{29}$$

$$Q=4:\quad a=\tanh\beta a/2 \tag{30}$$

$$Q\to\infty:\quad a=\frac{I_1(\beta a)}{I_0(\beta a)}. \tag{31}$$

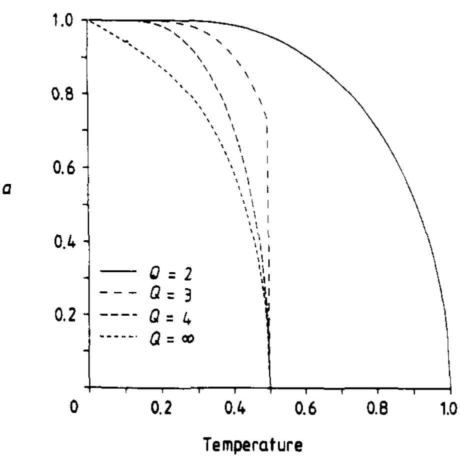

Figure 1. Plot of the overlap, a, as a function of temperature at $\alpha=0$ for $Q=2$, 3, 4 and $Q\to\infty$.

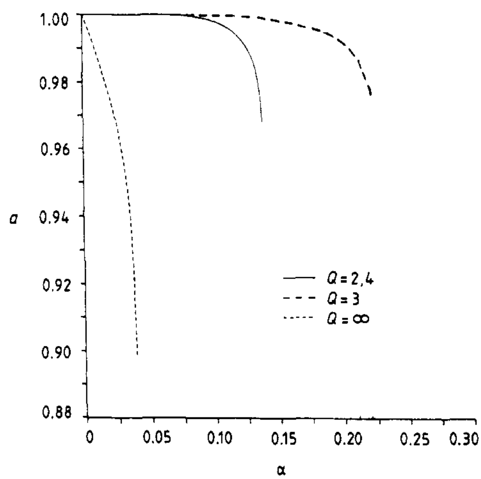

Figure 2. Plot of the overlap, a, as a function α at zero temperature for the cases $Q=2$, 3 and $Q\to\infty$.

The equations are solved to give the overlap, a, as a function of temperature, T (see figure 1). The curves are all of similar shape. In the limit $T\to 0$, the overlap tends to unity. For the case $Q=2$, the overlap is non-zero for $T<1$. For all other cases the overlap vanishes at $T=0.5$. This occurs because for $Q=2$ the mean value of $\cos 2\pi n/Q$ ($n=0,1,\ldots,Q-1$) is 1, whereas for all $Q>2$ it takes the value 0.5. At any given temperature the overlap is seen to decrease as Q increases.

4. Storage of an extensive number of patterns: $Q=3$

For the case $Q=3$ it can be shown that

$$X=\tfrac{1}{2}\ln\left(2\cosh\frac{\sqrt{3}}{2}B\right)+\ln\left\{2\cosh\left[\frac{3A}{4}-\tfrac{1}{2}\ln\left(2\cosh\frac{\sqrt{3}B}{2}\right)\right]\right\} \tag{32}$$

where the terms that average to zero when integrals over z and z' or when quenched averages are carried out have been ignored. This leads to the mean-field equations

$$a_\mu=\frac{\sqrt{3}}{4}\left\langle\!\!\left\langle\int\int\frac{\mathrm{d}z\,\mathrm{d}z'}{2\pi}\exp(-\tfrac{1}{2}z^2)\exp(-\tfrac{1}{2}z'^2)\left\{s\tanh\frac{\sqrt{3}B}{2}+\left(\sqrt{3}c-s\tanh\frac{\sqrt{3}B}{2}\right)\right.\right.\right.$$
$$\left.\left.\left.\times\tanh\left[\frac{3A}{4}-\tfrac{1}{2}\ln\left(2\cosh\frac{\sqrt{3}B}{2}\right)\right]\right\}\right\rangle\!\!\right\rangle \tag{33}$$

$$b_\mu=\frac{\sqrt{3}}{4}\left\langle\!\!\left\langle\int\int\frac{\mathrm{d}z\,\mathrm{d}z'}{2\pi}\exp(-\tfrac{1}{2}z^2)\exp(-\tfrac{1}{2}z'^2)\left\{c\tanh\frac{\sqrt{3}B}{2}-\left(\sqrt{3}s-c\tanh\frac{\sqrt{3}B}{2}\right)\right.\right.\right.$$
$$\left.\left.\left.\times\tanh\left[\frac{3A}{4}-\frac{1}{2}\ln\left(2\cosh\frac{\sqrt{3}B}{2}\right)\right]\right\}\right\rangle\!\!\right\rangle \tag{34}$$

$$\beta(1-q)=\frac{1}{2}\sqrt{\frac{3}{\alpha r}}\left\langle\!\!\left\langle\int\int\frac{\mathrm{d}z\mathrm{d}z'}{2\pi}\exp(-\tfrac{1}{2}z^2)\exp(-\tfrac{1}{2}z'^2)\right.\right.$$
$$\times\left\{z'\tanh\frac{\sqrt{3}B}{2}+\left(\sqrt{3}z-z'\tanh\frac{\sqrt{3}B}{2}\right)\right.$$
$$\left.\left.\left.\times\tanh\left[\frac{3A}{4}-\frac{1}{2}\ln\left(2\cosh\frac{\sqrt{3}B}{2}\right)\right]\right\}\right\rangle\!\!\right\rangle \tag{35}$$

$$r=\frac{2q}{[1-\tfrac{1}{2}\beta(1-q)]^2}. \tag{36}$$

The solution of the mean-field equations for the three-state model, (33)-(36), will now be discussed, first in the low-temperature limit and then for general temperature.

4.1. The low-temperature limit

In the limit, $T\to 0$, the mean-field equations simplify somewhat. q tends to 1 and the problem reduces to solving one non-linear equation in the variable y, where $y\equiv a/\sqrt{\alpha r}$. This is just as occurred in the same limit in the Hopfield model (see (I)) and will be the case for all Q.

Considering $Q=3$ the resulting equation is

$$yf_1(y)-3f_2(y)+8\sqrt{3\pi\alpha}\,y=0 \tag{37}$$

where

$$f_1(y)=6[1-\mathrm{erf}(\sqrt{2}y)]+12\exp(-\tfrac{3}{2}y^2)[1+\mathrm{erf}(y/\sqrt{2})] \tag{38}$$

$$f_2(y)=\left\langle\!\!\left\langle\int \mathrm{d}z'\exp(-\tfrac{1}{2}z'^2)\left\{s\,\mathrm{sgn}(B)[1+\mathrm{erf}(z_0/\sqrt{2})]-\sqrt{3}c\,\mathrm{erf}(z_0/\sqrt{2})\right\}\right\rangle\!\!\right\rangle \tag{39}$$

with

$$z_0=2\left(\frac{1}{\sqrt{3}}\left|ys+\frac{z'}{2}\right|-yc\right).$$

Non-zero solutions of equation (37) are found for α less some critical value, α_c, found to be 0.22. This compares with $\alpha_c=0.138$ in the case $Q=2$. The trivial solution $y=0$ exists for all α. The variation in overlap, with α for stable solutions of (37) is shown in figure 2. At the critical storage the overlap takes a value of 0.978, compared with 0.968 at critical storage in the Hopfield model.

4.2. *The phase diagram*

The mean-field equations (33)-(36) can be solved numerically for any value of α and temperature. The stability of solutions is determined by calculating the eigenvalues of the matrix (of order 0) of second derivatives of the free energy with respect to the order parameters. Consideration of the solutions and their stability with respect to replica-symmetric fluctuations allows the construction of the phase diagram, figure 3. The stability of the solutions with respect to replica-symmetry-breaking fluctuations has not been considered.

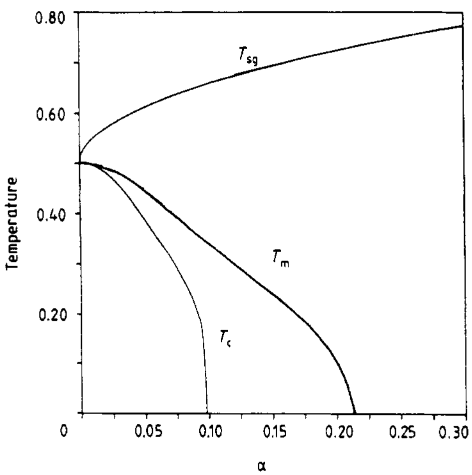

Figure 3. Phase diagram for $Q=3$.

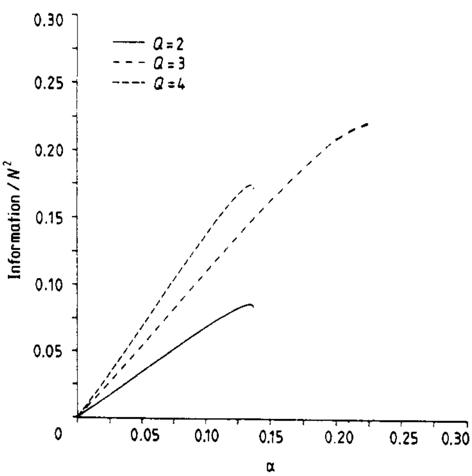

Figure 4. Plot of the information content of the networks with $Q=2$, 3, 4 against α at zero temperature.

Above the line T_{sg} the system is paramagnetic (i.e. $a=b=q=r=0$). Along T_{sg} there is a second-order transition to the spin-glass state in which $a=b=0$, $q\neq 0$, $r\neq 0$. T_{sg} is given by

$$T_{sg}=\tfrac{1}{2}+\sqrt{\alpha}/2. \tag{40}$$

Below the line T_m the memory states, ($a\neq 0$, $b=0$, $r\neq 0$, $q\neq 0$), become stable discontinuously. Below T_c the memory states have a lower free energy than the spin-glass states. The phase diagram is of very similar form to that of the Hopfield model, see (I).

4.3. *Information content*

To measure the amount of the information stored in the network, account must be taken not only of the overlaps achieved, but also of the number of states of the variables. Using the same procedure for evaluating information content as Amit *et al* (1987b), the information stored in the three-state network can be shown to be given by

$$\text{Information stored}=\alpha N^2\left[\frac{2a}{3}\ln\left(\frac{1+2a}{1-a}\right)+\tfrac{1}{2}\ln[(1-a^2)(1+2a)]\right]. \tag{41}$$

A plot of the information stored against α is shown in figure 4. The maximum information content for $Q=3$ is $0.22N^2$. This can be seen to be much higher than the maximal information content of the Hopfield model, which is only $0.09N^2$.

5. Storage of an extensive number of patterns: $Q=4$

$Q=4$ is a special case as the model decomposes into two $Q=2$ models. Hence, the storage capacity, $\alpha_c=0.138$, is the same as for the Hopfield model. The phase diagram is the same as that for the Hopfield model (see (I) or Lautrup 1988), but with the temperature rescaled by a factor of 0.5.

Although the mean-field equations at zero temperature are identical for $Q=2$ and $Q=4$, a four-state variable carries intrinsically more information than an Ising variable and so the information content, as shown in figure 4, is much higher for $Q=4$.

6. Storage of an extensive number of patterns: $Q\to\infty$

In the limit that $Q\to\infty$ the variables become continuous. To determine the mean-field equations for this case the quantity X given by (15) must first be calculated:

$$X=\mathrm{Tr}_n\exp\left(A\cos\frac{2\pi n}{Q}+B\sin\frac{2\pi n}{Q}\right).$$

In the limit that $Q\to\infty$:

$$X=\frac{Q}{2\pi}\int_0^{2\pi}\exp(A\cos y+B\sin y)\,\mathrm{d}y \tag{42}$$

i.e.

$$X=QI_0\sqrt{A^2+B^2}). \tag{43}$$

The mean-field equations can then be shown, for the category of solutions considered, to be

$$a=\left\langle\!\!\left\langle \int\int \frac{\mathrm{d}z\mathrm{d}z'}{2\pi}\exp(-\tfrac{1}{2}z^2)\exp(-\tfrac{1}{2}z'^2)\left(c\frac{\sqrt{\alpha r}}{2}z+s\frac{\sqrt{\alpha r}}{2}z'+a\right)\frac{I_1(D)}{D'I_0(D)}\right\rangle\!\!\right\rangle \tag{44}$$

$$\beta(1-q)=\frac{2}{\sqrt{\alpha r}}\left\langle\!\!\left\langle \int\int \frac{\mathrm{d}z\mathrm{d}z'}{2\pi}\exp(-\tfrac{1}{2}z^2)\exp(-\tfrac{1}{2}z'^2)\right.\right.$$

$$\left.\left.\times\left(\frac{\sqrt{\alpha r}}{2}(z^2+z'^2)+a(cz+sz')\right)\frac{I_1(D)}{D'I_0(D)}\right\rangle\!\!\right\rangle \tag{45}$$

$$r=\frac{2q}{[1-\frac{1}{2}\beta(1-q)]^2} \tag{46}$$

where $D=\sqrt{A^2+B^2}$ and $D'=D/\beta$.

The following subsections consider the solution of these mean-field equations.

6.1. *The low-temperature limit*

In the limit that $T\to 0$ the equations reduce to the non-linear equation

$$yf_2(y)+2\pi y\sqrt{2\alpha}-f_1(y)=0 \tag{47}$$

where

$$f_1(y)=\left\langle\!\!\left\langle \int\int \mathrm{d}z\,\mathrm{d}z'\exp(-\tfrac{1}{2}z^2)\exp(-\tfrac{1}{2}z'^2)\frac{(cz+sz'+2y)}{[(z+2yc)^2+(z'+2ys)^2]^{1/2}}\right\rangle\!\!\right\rangle \tag{48}$$

$$f_2(y)=\left\langle\!\!\left\langle \int\int \mathrm{d}z\,\mathrm{d}z'\exp(-\tfrac{1}{2}z^2)\exp(-\tfrac{1}{2}z'^2)\frac{(z^2+z'^2+2y(cz+sz'))}{[(z+2yc)^2+(z'+2ys)^2]^{1/2}}\right\rangle\!\!\right\rangle. \tag{49}$$

The stable solutions of equation (47) can be seen in figure 2. The critical value of α is 0.038 and at this point the overlap is 0.899.

The information stored in the network is given by

$$\text{Information stored}=\lim_{Q\to\infty}\alpha a N^2\ln Q. \tag{50}$$

So, below $\alpha_c=0.038$, the information content of the network is infinite.

6.2. *Phase diagram*

Figure 5 shows the phase diagram resulting from the mean field equations (44)-(46). The line T_c has not been drawn as it proved difficult to locate the spin-glass solutions at these small values of α with the numerical techniques used. The paramagnetic to spin-glass transition occurs along the same line as for $Q=3$, as noted in § 4. The general structure of the diagram is very similar to that of the Hopfield model and that derived in § 4.

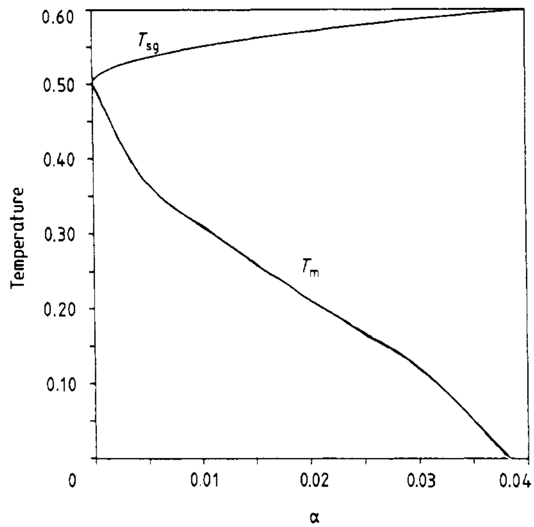

Figure 5. Phase diagram for $Q \to \infty$.

7. Concluding remarks

Using the replica trick the mean-field equations for the clock neural network have been derived in the replica-symmetric approximation. These have been solved for non-zero overlap with one nominated pattern only for the cases $Q = 2$, 3, 4 and for the continuous variable.

From figure 1 it seems that as the number of states of the variables increases at given α and temperature, the overlap decreases, so the retrieval of the nominated pattern is poorer. However, when information content is considered, it has been shown that the increase in the number of states of the variables more than outweighs any decrease in the overlap and the information content is seen to increase with Q for given α (less than $\alpha = 0.038$).

It has been shown that storage is still possible for the case $Q \to \infty$. However the storage capacity falls from 0.138 for $Q = 2$ and 0.266 for $Q = 3$ to a value of 0.038.

Mixture states (i.e. those with non-zero overlaps with more than one pattern) have not been considered. It may be interesting to find out whether any symmetric mixture states are stable for this model. Symmetric mixture states with non-zero overlaps with an odd number of nominated patterns have been shown to be stable for the Hopfield model (Amit *et al* 1985) and so should also be stable for the case $Q = 4$. However, for a Potts-type model Kanter (1988) found that symmetric mixture states were not stable.

Lastly, this paper has been confined to replica-symmetric solutions. Calculation of the entropy at zero temperature for the cases $Q = 2$, 3, 4 gives a negative entropy. This indicates that at least in these cases replica-symmetry breaking is occurring. In the cases considered the entropies calculated for the memory states were small, of the order of -0.001. This means that the effects of replica-symmetry breaking on the memory states will be small. A detailed consideration of replica-symmetry-broken solutions may be interesting, especially as there is unusual behaviour in the broken symmetry states of the three-state clock glass (Nobre and Sherrington 1987, Elderfield and Sherrington 1983, Gross *et al* 1985).

Acknowledgment

I would like to acknowledge the help of Dr E Gardner in suggesting this work and in providing advice during its development. Financial support for this work was from SERC.

References

Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev.* **32** 1007
—— 1987a *Ann. Phys., NY* **173** 30
—— 1987b *Phys. Rev.* A **35** 2293
Bruce A D, Gardner E and Wallace D J 1987 *J. Phys. A: Math. Gen.* **20** 2909
Canning A and Gardner E 1988 *J. Phys. A: Math. Gen.* **21** 3275
Elderfield D and Sherrington D 1983 *J. Phys. C: Solid State Phys.* **16** L497
Gross D J, Kanter I and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 304
Kanter L 1988 *Phys. Rev. A* **37** 7
Lautrup B 1988 *The Theory of the Hopfield Model* (Lecture notes, unpublished)
Nobre F D and Sherrington D 1986 *J. Phys. C: Solid State Phys.* **19** L181
Noest A J 1988 *Europhys. Lett.* **6** 469
Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **32** 1792