Main overlap dynamics for multistate neural networks

1991 J. Phys. A: Math. Gen. 24 L637

(http://iopscience.iop.org/0305-4470/24/11/011)

J. Phys. A: Math. Gen. 24 (1991) L637–L647. Printed in the UK

LETTER TO THE EDITOR

Main overlap dynamics for multistate neural networks

A E Patrick†, P Picco‡, J Ruiz‡ and V A Zagrebnov†

† Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, Dubna 141980, USSR

‡ Centre de Physique Théorique, CNRS Luminy-Case 907, F-13288 Marseille Cedex 9, France

Received 1 January 1991

Abstract. We present explicit formulae for one-step dynamics of retrieval-pattern errors (main overlap) for the q-state Potts and ($q \leqslant 4$) clock neural networks. Solutions of the fixed point equations and critical values of the saturation parameters $\alpha_c(q)$ in the one-step approximation are considered in the Potts case.

The understanding of neural networks exibiting learning and associative memory is based mainly on the Ising spin models (see, e.g., [1]) with energy function

$$H_\Lambda(\sigma_\Lambda) = -\sum_{i,j\in\Lambda} J_{ij}\sigma_i\sigma_j\,. \tag{1}$$

Here $\Lambda = \{1,2,\ldots N\}$ and $\boldsymbol{\sigma}_\Lambda = \{\sigma_i\}_{i\in\Lambda}$ is a configuration of neurons that have *two* main levels of activity, i.e. $\sigma_i \in Q_I = \{-1,1\}$. The bonds $\{J_{ij}\}_{i,j\in\Lambda}$ ($J_{ii} = 0$) are the synaptic efficacies which are assumed to be generated by learning in such a way that they ensure the dynamic stability of configurations $\{\boldsymbol{\sigma}^\mu_\Lambda\}_{\mu\in[M]}$ which are close to a stored patterns $\{\xi^\mu_i\}_{i\in\Lambda} \equiv \boldsymbol{\xi}^\mu$, $\mu \in [M]$; $[M] = \{1,2,\ldots,M\}$. According to the Hebb learning rule [2]

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{M}\xi^\mu_i\xi^\mu_j \qquad i \neq j \tag{2}$$

where the patterns $\{\boldsymbol{\xi}^\mu\}_{\mu\in[M]}$ are taken to be quenched random vectors with independent components $\xi^\mu_i \in \{\pm1\}; i \in [N]$, and $\Pr\{\xi^\mu_i = \pm1\} = \frac{1}{2}$.

A generalization of the Ising neural network (1) and (2) (the Little–Hopfield model [3]) to the case of neurons with more then two discrete states has been considered in recent papers [4,5]. In [4] each neuron $\{\sigma_i\}$ is viewed as a Potts spin with q possible states. Therefore, instead of binary encoded pattern configurations $\boldsymbol{\sigma}_\Lambda$ may represent coloured or q-grade shaded patterns. That paper was devoted to the thermodynamics of the Potts neural network and to a comparison of its properties with those of the Little–Hopfield model. In [5], the same approach is used for the clock model. In a recent paper [8], the authors introduced networks of three-state neurons, where the additional state embodies the absence of information.

0305-4470/91/110637+11$03.50

The aim of this letter is to consider *parallel dynamics* for two kinds of *multistate* networks: for q-state Potts neural networks and for their clock model counterparts. We obtain explicit formulae for the corresponding main overlap evolution after one step of the zero-temperature parallel dynamics. Note that in the Ising case, the state of each neuron after one step of the zero-temperature parallel dynamics is equal to the sign of the corresponding induced local field whereas for the multistate case each neuron makes a more complicated decision comparing induced local fields in all states.

We first present our scheme of calculation of the one-step evolution for the main overlap in the Ising case, which is also applicable in the general case of multistate neurons. Next, we derive the corresponding one-step formula for evolution of the main overlap for the q-state Potts neural network. Then we use the same method to derive an explicit one-step formula for evolution of the main overlap for the clock neural network for $q = 3, 4$. Finally we make some concluding remarks.

We first deal with the Little–Hopfield neural network. Recal that in this case the patterns $\{\xi_i^\mu\}_{i\in\Lambda}$, $\mu \in [M]$ are independent identically distributed random variables (IIDRV) with $\xi_i^\mu \in Q_I = \{-1, 1\}$. The Hamiltonian (1) with (2) takes the form

$$H_\Lambda(\sigma_\Lambda) = -\sum_{i\in\Lambda} \varepsilon_i^{(\Lambda)}[\sigma_i; \sigma_{\Lambda\setminus i}] \tag{3}$$

where we have introduced the local energy-function

$$\varepsilon_i^{(\Lambda)}[\sigma_i; \sigma_{\Lambda\setminus i}] = \sum_{\mu\in[M]} \xi_i^\mu \sigma_i \left(\frac{1}{N} \sum_{j\in\Lambda\setminus i} \xi_j^\mu \sigma_j \right) \equiv \sigma_i h_i^{(\Lambda)}(\sigma_{\Lambda\setminus i}) . \tag{4}$$

Let the configuration at time t be $\sigma_\Lambda(t)$. The zero-temperature parallel dynamics, see, e.g., [1]

$$\sigma_i(t+1) = \mathrm{sign}[h_i^{(\Lambda)}(\sigma_{\Lambda\setminus i}(t))]$$

is equivalent to the following rule:

$$\sigma_i(t) \to \sigma_i(t+1) : \ \varepsilon_i^{(\Lambda)}[\sigma_i(t+1); \sigma_{\Lambda\setminus i}(t)] = \max_{\sigma_i \in Q_I} \varepsilon_i^{(\Lambda)}[\sigma_i; \sigma_{\Lambda\setminus i}(t)] \qquad i \in \Lambda . \tag{5}$$

We introduce this—perhaps unusual—reformulation of the dynamics in terms of the local energies, because it allows an easy and transparent generalization to the zero-temperature dynamics in the Potts and clock models to be discussed below.

Let the initial condition $\{\sigma_i(t=0)\}_{i\in\Lambda}$ be IIDRV from Q_I correlated with *only one* pattern $\boldsymbol{\xi}^\nu$, i.e.

$$\Pr\{\sigma_i(t=0) = \xi_i^\mu\} = \tfrac{1}{2}(1 + \delta_{\mu\nu} m^\nu(t=0)) \qquad \mu \in [M] . \tag{6}$$

This means that for overlaps $m_\Lambda^\mu(t=0) = (1/N)\sum_{i\in\Lambda} \xi_i^\mu \sigma_i(t=0)$ by the strong law of large numbers (SLLN) one obtains

$$m^\mu(t=0) = \lim_{N\to\infty} m_\Lambda^\mu(t=0) = \delta_{\mu\nu} m^\nu(0) \qquad \mu \in [M] \tag{7}$$

where $m^\nu(0) > 0$ is an initial value of main overlap. For these initial *conditions*, the local energy-function (4) can be rewritten in the following form:

$$\varepsilon_i^{(\Lambda)}[\sigma_i; \sigma_{\Lambda\setminus i}(t=0)] = \sigma_i \xi_i^\nu \left[m_{\Lambda\setminus i}^\nu(t=0) + \frac{1}{N} \sum_{\mu\in[M]\setminus\nu} \xi_i^\mu \xi_i^\nu \sum_{j\in\Lambda\setminus i} \xi_j^\mu \sigma_j(t=0) \right] . \tag{8}$$

We now consider the so-called 'α'-limit, $N \to \infty$ at fixed $\alpha \equiv M/N$, in (8). By the initial conditions and the central limit theorem (CLT), the second term in brackets in (8) converges *in distribution* to $\sqrt{\alpha}\mathcal{N}(0,1)$, where $\mathcal{N}(a,b)$ denotes a Gaussian random variable with expectation a and variance b. Hence, the local energy functions (8) in the 'α'-lim are random variables themselves

$$\begin{aligned}\varepsilon_i[\sigma_i; \boldsymbol{\sigma}(t=0)] &\stackrel{d}{=} \text{'}\alpha\text{'-lim}\, \varepsilon_i[\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i}(t=0)] \\ &\stackrel{d}{=} \xi_i^\nu \sigma_i[m^\nu(t=0) + \sqrt{\alpha}\mathcal{N}(0,1)] \qquad i = 1, 2, \ldots \,. \end{aligned} \tag{9}$$

Now we can determine the configuration $\boldsymbol{\sigma}(t=1)$ for the neural network *after* the 'α'-lim (see (5) and (9))

$$\sigma_i(t=1): \ \varepsilon_i[\sigma_i(t=1); \boldsymbol{\sigma}(t=0)] = \max_{\sigma_i \in Q_I} \{\xi_i^\nu \sigma_i[m^\nu(t=0) + \sqrt{\alpha}\,\mathcal{N}(0,1)]\}\,. \tag{10}$$

According to (10) one obtains

$$\sigma_i(t=1) = \begin{cases} \xi_i^\nu & \text{if } m^\nu(t=0) + \sqrt{\alpha}\,\mathcal{N}(0,1) > 0 \\ -\xi_i^\nu & \text{if } m^\nu(t=0) + \sqrt{\alpha}\,\mathcal{N}(0,1) < 0 \end{cases} \tag{11}$$

i.e. the future state of σ_i is determined by the sign of the signal- and noise-contributions to the local energy. Hence, for the main overlap at $t = 0$ we obtain a well known formula, namely

$$\begin{aligned} m^\nu(t=1) &= \text{'}\alpha\text{'-lim}\, \frac{1}{N} \sum_{i\in\Lambda} \xi_i^\nu \sigma_i(t=1) \\ &= \boldsymbol{E}\{\mathrm{sign}[m^\nu(0) + \sqrt{\alpha}\,\mathcal{N}(0,1)]\} \\ &= \sqrt{\frac{2}{\pi}} \int_0^{m^\nu(0)} \mathrm{d}x\, \exp\left[-\frac{x^2}{2}\right]. \end{aligned} \tag{12}$$

The same arguments give $m^\mu(t=1) = 0$ for $\mu \neq \nu$ because $m^{\mu(\neq\nu)}(t=0) = 0$. To calculate the evolution of $m^\nu(t)$ for $t \geqslant 2$, one needs more complicated arguments [6, 7].

We deal next with the q-state Potts neural network. We suppose that each neuron has $q \geqslant 2$ levels of activity described by the Potts spin variables $\sigma_i \in \{1, 2, \ldots, q\} = Q_P$. The stored patterns $\{\xi_i^\mu \in Q_P\}_{i\in\Lambda}, \mu \in [M]$ are taken to be IIDRV with $\Pr\{\xi_i^\mu = k \in Q_P\} = 1/q$. The corresponding energy function (cf (1) and (2)) assumes the form

$$H_\Lambda(\boldsymbol{\sigma}_\Lambda) = -\sum_{\mu\in[M]} \frac{1}{N} \sum_{i\neq j}^{N} \left[\delta(\xi_i^\mu; \sigma_i) - \frac{1}{q}\right] \left[\delta(\xi_j^\mu; \sigma_j) - \frac{1}{q}\right] \equiv -\sum_{i\in\Lambda} \varepsilon_i^{(\Lambda)}[\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i}] \tag{13}$$

which for $q = 2$ reduces essentially to the Hamiltonian (1),(2). Here $\delta(\cdot;\cdot)$ is the Kronecker delta. Zero-temperature parallel dynamics is defined now by

$$\sigma_i(t) \to \sigma_i(t+1): \varepsilon_i^{(\Lambda)}[\sigma_i(t+1); \boldsymbol{\sigma}_{\Lambda\setminus i}(t)] = \max_{\sigma_i \in Q_P} \varepsilon_i^{(\Lambda)}[\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i}(t)] \qquad \forall i \in \Lambda\,. \tag{14}$$

We can now apply the above line of reasoning to derive the evolution of the main overlap for parallel dynamics in the case of a Potts neural network. Let the initial conditions $\{\sigma_i(t=0)\}_{i\in\Lambda}$ be IIDRV such that

$$\Pr\{\sigma_i(t=0)=\xi_i^\mu\}=\frac{1+(q-1)m^\nu(0)\delta_{\mu\nu}}{q} \qquad \mu\in[M] \tag{15}$$

where $m^\nu(0)>0$. For the Potts model, the overlaps (at finite Λ) are defined as

$$m_\Lambda^\mu(t=0)=\frac{q}{q-1}\frac{1}{N}\sum_{j\in\Lambda}\left[\delta(\xi_j^\mu;\sigma_j(t=0))-\frac{1}{q}\right] \qquad \mu\in[M] \tag{16}$$

where the prefactor is chosen so as to satisfy $m_\Lambda^\mu(t=0)=1$ for $\sigma_i(t=0)=\xi_i^\mu\ \forall i\in\Lambda$. By the initial condition (15) and the SSLN one obtains

$$m^\mu(t=0)=\lim_{N\to\infty}m_\Lambda^\mu(t=0)=\frac{q}{q-1}\left[\boldsymbol{E}\delta(\xi_i^\mu;\sigma_i(t=0))-\frac{1}{q}\right]=\delta_{\mu\nu}m^\nu(0). \tag{17}$$

For $q=2$, equations (15) and (16) coincide with corresponding definitions for the Ising case.

Let $\mathcal{M}_p^{(\nu)}=\{\mu\in[M]\setminus\nu:\ \xi_i^\mu=p,\forall i\in\Lambda\}$. Then by definition of the local energy function (13), one obtains

$$\frac{q}{q-1}\varepsilon_i^{(\Lambda)}[\sigma_i;\boldsymbol{\sigma}_{\Lambda\setminus i}(t=0)]=\left[\delta(\xi_i^\nu;\sigma_i)-\frac{1}{q}\right]m_{\Lambda\setminus i}^\nu(t=0)+\frac{q}{q-1}\sum_{p=1}^{q}\left[\delta(p;\sigma_i)-\frac{1}{q}\right]$$
$$\times\frac{1}{N}\sum_{\mu\in\mathcal{M}_\nu^{(p)}}\sum_{j\in[N]\setminus i}\left[\delta(\xi_j^\mu;\sigma_j(t=0))-\frac{1}{q}\right]. \tag{18}$$

Here we have again separated the signal from the noise contribution to the local energy, and we have split latter into q terms, each favouring one of the possible Potts states. The pth such contribution will henceforth be called the pth component of the noise for short.

By equation (15), we have

$$\boldsymbol{E}\left(\delta(\xi_j^\mu;\sigma_j(t=0))-\frac{1}{q}\right)=0 \qquad \mu\neq\nu$$
$$\mathrm{Var}\left(\delta(\xi_j^\mu;\sigma_j(t=0))-\frac{1}{q}\right)=\frac{q-1}{q^2} \qquad \mu\neq\nu$$

for the statistics of individual contributions to the noise term in (18). According to the CLT for the pth component of the noise in (18) one obtains

$$\text{‘}\alpha\text{’-lim}\,\frac{q}{q-1}\frac{1}{N}\sum_{\mu\in\mathcal{M}_\nu^{(p)}}\sum_{j\in[N]\setminus i}\left[\delta(\xi_j^\mu;\sigma_j(t=0))-\frac{1}{q}\right]\stackrel{d}{=}\sqrt{\frac{\alpha}{q(q-1)}}\mathcal{N}_\nu^{(p)}(0;1). \tag{19}$$

where we used the result that, by the SLLN, $|\mathcal{M}_\nu^{(p)}|/M \to 1/q$, when $M \to \infty$. Hence, at $t = 0$ the random local energy function (18) converges in distribution to the random variable

$$\varepsilon_i[\sigma_i; \boldsymbol{\sigma}(t=0)] \equiv \text{`}\alpha\text{'-}\lim \frac{q}{q-1}\varepsilon_i^{(\Lambda)}[\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i}(t=0)]$$
$$\stackrel{d}{=} \left[\delta(\xi_i^\nu; \sigma_i) - \frac{1}{q}\right] m^\nu(t=0) + \sum_{p=1}^{q}\left[\delta(p;\sigma_i) - \frac{1}{q}\right]\sqrt{\frac{\alpha}{q(q-1)}}\mathcal{N}_\nu^{(p)}(0;1). \tag{20}$$

By the initial conditions (15) $\{\mathcal{N}_\nu^{(p)}(0;1)\}_p$ are *independent* Gaussian variables.

The configuratuon $\{\sigma_i(t=1)\}_i$ is defined by relation (14) where the local energy function is given by (20). There are *two* different ways to achieve the max in (14), cf (20). Either (a) $\sigma_i = \xi_i^\nu$, then

$$\varepsilon_i^{(a)} = \left(1 - \frac{1}{q}\right)\left[m^\nu(0) + \sqrt{\frac{\alpha}{q(q-1)}}\mathcal{N}_{i;p=\xi_i^\nu}^\nu(0;1)\right] - \frac{1}{q}\sqrt{\frac{\alpha}{q(q-1)}}\sum_{p(\neq\xi_i^\nu)}^{q}\mathcal{N}_{i;p}^\nu(0;1)$$

or (b) $\sigma_i = p_{\max} \neq \xi_i^\nu$, then

$$\varepsilon_i^{(b)} = -\frac{1}{q}\left[m^\nu(t) - \sqrt{\frac{\alpha}{q(q-1)}}\sum_{p\in Q_P\setminus p^{\max}}\mathcal{N}_{i;p}^\nu(0;1)\right] + \left[1 - \frac{1}{q}\right]\sqrt{\frac{\alpha}{q(q-1)}}\mathcal{N}_{i;p^{\max}}^\nu(0;1).$$

In the latter case, $p^{\max}$ is defined through the following relation:

$$\mathcal{N}_{i;p^{\max}}^\nu(0;1) = \max_{p\in Q_P\setminus\xi_i^\nu}\mathcal{N}_{i;p}^\nu(0;1)$$

for *each fixed realization* of the independent Gaussian noises $\{\mathcal{N}_{i;p}^\nu(0;1)\}_p$. Which of the possible outcomes is chosen depends on the value of the difference

$$\Delta_i = \varepsilon_i^{(a)} - \varepsilon_i^{(b)} = m^\nu(t=0) + \sqrt{\frac{\alpha}{q(q-1)}}\,[\mathcal{N}_{i;p=\xi_i^\nu}^\nu(0;1) - \mathcal{N}_{i;p^{\max}}^\nu(0;1)]. \tag{21}$$

Equation (21) is an analogue of the standard *signal-noise relation* for the case of *Potts neurons*, cf (11). By (21) we obtain

$$\sigma_i(t=1) = \begin{cases} \xi_i^\nu & \text{if } \Delta_i > 0 \\ p_{\max} & \text{if } \Delta_i \leqslant 0 \end{cases} \qquad \text{i.e. } \delta(\sigma_i(t=1);\xi_i^\nu) = \theta(\Delta_i). \tag{22}$$

Recall that $\{\mathcal{N}_{i;p}^\nu\}_p$ are independent. Hence

$$\Pr\{\mathcal{N}_{i;p_{\max}}^\nu(0;1) \leqslant x\} = \left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{d}z\, \exp\left(-\frac{z^2}{2}\right)\right]^{q-1} \equiv [\Phi(x)]^{q-1}$$

and consequently by (21) and (22) we obtain

$$\Pr\{\Delta_i > 0\} = \Pr\left\{m^\nu(t=0)\sqrt{\frac{\alpha}{q(q-1)}} + \mathcal{N}_{i;p=\xi_i^\nu}^\nu(0;1) > \mathcal{N}_{i;p^{\max}}^\nu(0;1)\right\}$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}y\,\exp\left[-\frac{y^2}{2}\right]\left[\Phi\left(y + \sqrt{\frac{\alpha}{q(q-1)}}m^\nu(t=0)\right)\right]^{q-1} \tag{23}$$

As a result of (22) and (23) we obtain the following formula for the *main overlap* at $t = 1$ (cf (16)):

$$
\begin{aligned}
m^{\nu}(t=1) &= \text{`}\alpha\text{'-lim}\,\frac{q}{q-1}\frac{1}{N}\sum_{j\in\Lambda}\left[\delta(\sigma_j(t=1);\xi_j^{\nu}) - \frac{1}{q}\right] \\
&= \frac{q}{q-1}E\left\{\delta(\sigma_j(t=1);\xi_j^{\nu}) - \frac{1}{q}\right\} \\
&= \frac{q}{q-1}\left[\Pr\{\Delta_i > 0\} - \frac{1}{q}\right] \\
&= \frac{q}{q-1}\left\{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}y\,\exp\left(-\frac{y^2}{2}\right)\right. \\
&\quad \times \left.\left[\Phi\left(y + \sqrt{\frac{\alpha}{q(q-1)}}m^{\nu}(t=0)\right)\right]^{q-1} - \frac{1}{q}\right\}.
\end{aligned}
\tag{24}
$$

As above, one obtains the result that $m^{\mu}(t=1) = 0$, for $\mu \neq \nu$, because $m^{\mu(\neq\nu)}(t = 0) = 0$. One can also check that for $q = 2$ equation (24) coincides with (12).

We deal now with the clock neural network. As with the Potts neural network, each neuron has q levels of activity, i.e. $\sigma_i \in Q_P = \{1, 2, \ldots, q\}$, but the energy of configuration $\boldsymbol{\sigma}_\Lambda$ has a 'clock form' [5]

$$
H_\Lambda(\boldsymbol{\sigma}_\Lambda) = -\frac{1}{N}\sum_{i\neq j}^{N}\sum_{\mu\in[M]}\cos\left\{\frac{2\pi}{q}(\theta_i^{\mu} - \sigma_i)\right\}\cos\left\{\frac{2\pi}{q}(\theta_j^{\mu} - \sigma_j)\right\}. \tag{25}
$$

Here again M patterns $\{\boldsymbol{\theta}^{\mu}\}_{\mu\in[M]}$, $[M] = \{1, 2, \ldots, M\}$, are described by IIDRV $\{\theta_i^{\mu}\}_{i\in\Lambda}$, $\mu \in [M]$ with $\Pr\{\theta_i^{\mu} = p \in Q_P\} = 1/q$. The local energy function has the form, cf (25)

$$
\varepsilon_i^{(\Lambda)}[\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i}] = \frac{1}{N}\sum_{\mu\in[M]}\cos\left[\frac{2\pi}{q}(\theta_i^{\mu} - \sigma_i)\right]\sum_{j(\neq i)}^{N}\cos\left[\frac{2\pi}{q}(\theta_j^{\mu} - \sigma_j)\right] \tag{26}
$$

and the dynamics $\boldsymbol{\sigma}_\Lambda(t) \to \boldsymbol{\sigma}_\Lambda(t+1)$ for a finite Λ is defined by the relation

$$
\sigma_i(t) \to \sigma_i(t+1) :\ \varepsilon_i^{(\Lambda)}[\sigma_i(t+1); \boldsymbol{\sigma}_{\Lambda\setminus i}(t)]\ \max_{\sigma_i\in Q_P}\varepsilon_i^{(\Lambda)}[\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i}(t)] \qquad \forall i \in \Lambda. \tag{27}
$$

We define overlaps

$$
m_\Lambda^{\mu}(t) = \frac{1}{N}\sum_{i\in\Lambda}\cos\left[\frac{2\pi}{q}(\theta_i^{\mu} - \sigma_i(t))\right] \qquad \mu \in [M]. \tag{28}
$$

Let the initial conditions $\{\sigma_i(t=0)\}_{i\in\Lambda}$ be IIDRV correlated with only one pattern $\boldsymbol{\xi}^{\nu}$

$$
\Pr\{\sigma_i(t=0) = \theta_i^{\mu}\} = \frac{1}{q} + \delta_{\mu\nu}G_q(m^{\nu}(0)). \tag{29}
$$

Here we choose $G_q(\cdot)$ in such away that application of the SLLN for (28) with distribution (29) gives

$$m^{\mu}(t=0) = \lim_{N\to\infty} m^{\mu}_{\Lambda}(t=0) \simeq \delta_{\mu\nu} m^{\nu}(0) \tag{30}$$

where $m^{\nu}(0) \neq 0$ is a value of main overlap for $t = 0$.

To calculate the main overlap at $t = 1$, we use the same line of reasoning as before. But now the complexity of the max-problem (27) increases very fast with q. Therefore, we restrict ourselves to $q = 3$ and 4, which have a deep relation to the Potts and Ising cases, respectively, and which allow one to treat them explicitly.

For $q = 3$ and 4 one has $G_{q=3}(x) = \frac{2}{3}x$ and $G_{q=4}(x) = x$, correspondingly. As above, we define an appropriate (i-dependent) decomposition of the set $[M]$

$$\mathcal{M}_{i;p} = \{[M] \setminus \nu : \theta^{\mu}_i = \theta^{\nu}_i + p(\mathrm{mod}\, q-1),\ p = 0, 1, \ldots, q-1\} \tag{31}$$

Using this decomposition, we can write the local energy function (26) in the form

$$\begin{aligned}
\varepsilon^{(\Lambda)}_i[\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i}(t=0)] &= \cos\left[\frac{2\pi}{q}(\theta^{\nu}_i - \sigma_i)\right]\left\{ m^{\nu}_{\Lambda\setminus i}(t=0) \right. \\
&\quad \left. + \sum_{\mu\in\mathcal{M}_{i;p=0}} \sum_{j\in\Lambda\setminus i} \cos\left[\frac{2\pi}{q}(\theta^{\mu}_j - \sigma_j(t=0))\right]\right\} \\
&\quad + \sum_{p=1}^{q-1} \cos\left[\frac{2\pi}{q}(\theta^{\nu}_i - \sigma_i + p)\right] \\
&\quad \times \frac{1}{N} \sum_{\mu\in\mathcal{M}_{i;p}} \sum_{j\in\Lambda\setminus i} \cos\left[\frac{2\pi}{q}(\theta^{\mu}_j - \sigma_j(t=0))\right].
\end{aligned} \tag{32}$$

By the choice of the θ^{μ}_i, the decomposition (31) satisfies $|\mathcal{M}_{i;p}|/N \to \alpha/q$ for $N \to \infty$. Hence, by the SLLN and the CLT we obtain

$$\begin{aligned}
\varepsilon_i[\sigma_i; \boldsymbol{\sigma}(t=0)] &= \text{`}\alpha\text{'-}\lim \varepsilon^{(\Lambda)}_i[\sigma_i; \boldsymbol{\sigma}_{\Lambda\setminus i}(t=0)] \\
&\stackrel{d}{=} \cos\left[\frac{2\pi}{q}(\theta^{\nu}_i - \sigma_i)\right]\left[m^{\nu}(0) + \sqrt{\frac{\alpha}{q}}\mathcal{N}_0(0;\tfrac{1}{2})\right] \\
&\quad + \sum_{p=1}^{q-1} \sqrt{\frac{\alpha}{q}} \cos\left[\frac{2\pi}{q}(\theta^{\nu}_i - \sigma_i + p)\right] \mathcal{N}_p(0;\tfrac{1}{2})
\end{aligned} \tag{33}$$

where $\{\mathcal{N}_p(0;\frac{1}{2})\}_{p=1}^{q-1}$ are independent Gaussian random variables.

The max-problem (27) for the limiting local energy function (33) in the case of $q = 3$ can be resolved in the following ways (cf (21)). Either (a) $\sigma_i = \theta^{\nu}_i$, in which case

$$\varepsilon^{(a)}_i = \mathcal{E}^{\nu}_0 - \tfrac{1}{2}\sqrt{\alpha/3}\{\mathcal{N}_1(0;\tfrac{1}{2}) + \mathcal{N}_2(0;\tfrac{1}{2})\}$$

or (b)

$$\sigma_i = \begin{cases} \theta_i^\nu + 1(\mathrm{mod}\,3): & \varepsilon_i^{(b)} = -\frac{1}{2}\mathcal{E}_0^\nu + \sqrt{\alpha/3}\,[\mathcal{N}_1(0;\frac{1}{2}) - \frac{1}{2}\mathcal{N}_2(0;\frac{1}{2})] \\ & \text{if } \mathcal{N}_1(0;\frac{1}{2}) > \mathcal{N}_2(0;\frac{1}{2}) \\ \theta_i^\nu + 2(\mathrm{mod}\,3): & \varepsilon_i^{(b)} = -\frac{1}{2}\mathcal{E}_0^\nu + \sqrt{\alpha/3}\,[\mathcal{N}_2(0;\frac{1}{2}) - \frac{1}{2}\mathcal{N}_1(0;\frac{1}{2})] \\ & \text{if } \mathcal{N}_2(0;\frac{1}{2}) > \mathcal{N}_1(0;\frac{1}{2})\,. \end{cases}$$

Here $\mathcal{E}_0^\nu \equiv m^\nu(t=0) + \sqrt{\frac{1}{2}}\mathcal{N}_0(0;\frac{1}{2})$. Which of the two channels, (a) or (b) is actually chosen depends, again, on the sign of the difference

$$\Delta_i = \varepsilon_i^{(a)} - \varepsilon_i^{(b)} = \tfrac{3}{2}\mathcal{E}_0^\nu - \tfrac{3}{2}\sqrt{\alpha/3}\max\{\mathcal{N}_1(0;\tfrac{1}{2});\mathcal{N}_2(0;\tfrac{1}{2})\}\,. \tag{34}$$

As a consequence one obtains

$$\cos[(2\pi/3)(\theta_i^\nu - \sigma_i(t=1))] = \begin{cases} 1 & \text{with } \Pr\{\varepsilon_i^{(a)} - \varepsilon_i^{(b)} \geqslant 0\} \\ -\frac{1}{2} & \text{with } \Pr\{\varepsilon_i^{(a)} - \varepsilon_i^{(b)} > 0\}\,. \end{cases} \tag{35}$$

According to (34) the event

$$\{\varepsilon_i^{(a)} - \varepsilon_i^{(b)} \geqslant 0\} = \{\mathcal{E}_0^\nu \geqslant \sqrt{\alpha/3}\mathcal{N}_1(0;\tfrac{1}{2}); \mathcal{E}_0^\nu \geqslant \sqrt{\alpha/3}\mathcal{N}_2(0;\tfrac{1}{2})\}$$

i.e.

$$\begin{aligned} \Pr\{\varepsilon_i^{(a)} - \varepsilon_i^{(b)} \geqslant 0\} &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} \mathrm{d}x\, \exp(-x^2)\left[\frac{1}{\sqrt{\pi}}\int_{-\infty}^{x+\sqrt{3/\alpha}\,m^\nu(0)} \mathrm{d}y\, \exp(-y^2)\right]^2 \\ &\equiv \Phi_3(\alpha; m^\nu(0))\,. \end{aligned} \tag{36}$$

Using (35) and (36), we obtain for the main overlap at $t = 1$ (see (28)) the result that for $N \to \infty$

$$m_{q=3}^\nu(t=1) = \boldsymbol{E}\left\{\cos\left[\frac{2\pi}{3}(\theta_i^\nu - \sigma_i(t=1))\right]\right\} = \tfrac{3}{2}\Phi_3(\alpha; m^\nu(0)) - \tfrac{1}{2} \tag{37}$$

by the SLLN.

The same calculations for the overlaps with $\mu \neq \nu$ give $m^{\mu(\neq\nu)}(t=1) = 0$. Note, that (37) is very similar to the corresponding formula for the Potts model for $q = 3$, cf (24).

To resolve the max-problem (27) for the case $q = 4$ (see (33)), we have the following possibilities:

$$\begin{aligned} &\text{(a)} \qquad \sigma_i = \theta_i^\nu (= \theta_i^\nu + 2(\mathrm{mod}\,4)) : \ \varepsilon_i^{(a)} = \epsilon_0^\nu (= -\epsilon_0^\nu) \\ &\text{(b)} \qquad \sigma_i = \theta_i^\nu + 1(\mathrm{mod}\,4) (= \theta_i^\nu + 3(\mathrm{mod}\,4)) : \ \varepsilon_i^{(b)} = \mathcal{N} (= -\mathcal{N}) \end{aligned} \tag{38}$$

where $\epsilon_0^\nu \equiv m^\nu(t=0) + \sqrt{\alpha/4}\,[\mathcal{N}_0(0;\frac{1}{2}) - \mathcal{N}_2(0;\frac{1}{2})]$ and $\mathcal{N} \equiv \sqrt{\alpha/4}\,[\mathcal{N}_1(0;\frac{1}{2}) - \mathcal{N}_3(0;\frac{1}{2})]$. Then, using independence of the random variables $\{\mathcal{N}_p(0;\frac{1}{2})\}_{p=0}^{q-1}$, we obtain the result that the event

$$\begin{aligned}\{\sigma_i(t=1)=\theta_i^\nu\} &= \{\varepsilon_i^{(a)} \geqslant \varepsilon_i^{(b)}; \varepsilon_i^{(a)} \geqslant -\varepsilon_i^{(b)}; \varepsilon_i^{(a)} \geqslant -\varepsilon_i^{(a)}\} \\ &= \{m^\nu(t=0) + \sqrt{\alpha/2}\mathcal{N}_{02}(0;1) \geqslant \sqrt{\alpha/2}\mathcal{N}_{31}(0;1); m^\nu(t=0) \\ &\quad + \sqrt{\alpha/2}\mathcal{N}_{02}(0;1) \geqslant -\sqrt{\alpha/2}\mathcal{N}_{31}(0;1); m^\nu(t=0) + \sqrt{\alpha/2}\mathcal{N}_{02}(0;1) \geqslant 0\}\end{aligned}$$

where $\mathcal{N}_{02}(0;1) = \frac{1}{\sqrt{2}}[\mathcal{N}_0(0;\frac{1}{2}) - \mathcal{N}_2(0;\frac{1}{2})]$ and $\mathcal{N}_{31}(0;1) = \frac{1}{\sqrt{2}}[\mathcal{N}_3(0;\frac{1}{2}) - \mathcal{N}_1(0;\frac{1}{2})]$. Hence

$$\begin{aligned}\Pr\{\sigma_i(t=1)=\theta_i^\nu\} &= \frac{1}{2\pi}\int_{-\sqrt{2/\alpha}}^{\infty} \mathrm{d}x\, \exp\left(-\frac{x^2}{2}\right) \int_{-x-\sqrt{2/\alpha}\,m^\nu(0)}^{x+\sqrt{2/\alpha}\,m^\nu(0)} \mathrm{d}y\, \exp\left(-\frac{y^2}{2}\right) \\ &= \Phi_4(\alpha; m^\nu(0))\,.\end{aligned} \tag{39}$$

Similarly, one obtains that $\Pr\{\sigma_i(t=1) = \theta_i^\nu + 2(\mathrm{mod}\,4)\} = \Phi_4(\alpha; -m^\nu(0))$. If $\sigma_i(t=1) = \theta_i^\nu + 1(= \theta_i^\nu + 3)(\mathrm{mod}\,4)$, see (38), then $\cos[(2\pi/q)(\theta_i^\nu - \sigma_i(t=1))] = 0$ for $q=4$. Therefore, by the SLLN for (28), in the limit $N \to \infty$ and by (38) and (39), we obtain for the main overlap at $t=1$ the final result that

$$m_{q=4}^\nu(t=1) = E\left\{\cos\left[\frac{\pi}{2}(\theta_i^\nu - \sigma_i(t=1))\right]\right\} = \mathrm{erf}\left[\frac{m^\nu(t=0)}{\sqrt{\alpha}}\right]. \tag{40}$$

Similar calculations give $m^{\mu(\neq\nu)}(t=1) = 0$. It is interesting to note that (40) coincides with equation (12).

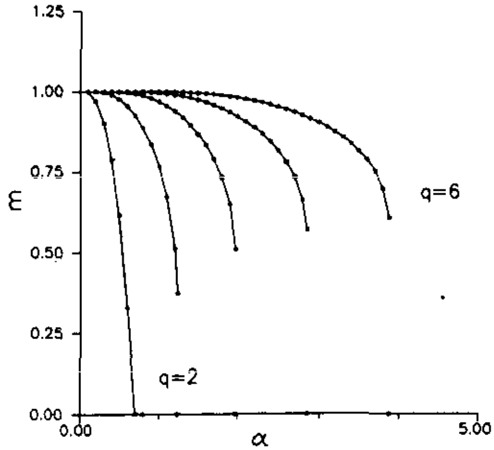

Figure 1. Solutions of the fixed-point equation (i.e. $m(t=1) = m(t=0) = m^*$ in (24)) for the q-state Potts model as functions of α for $q = 2,3,4,5,6$.

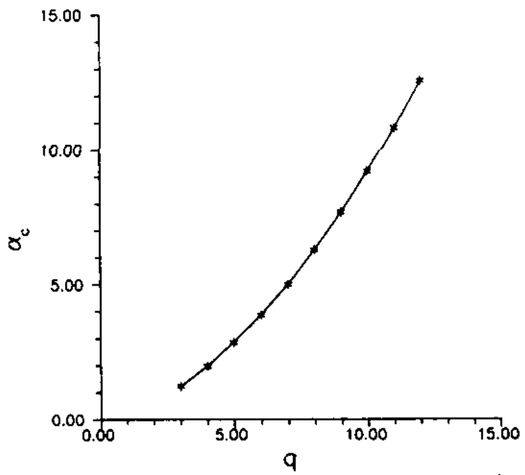

Figure 2. The critical value α_c as a function of q for the Potts model (one-step approximation).

We conclude this letter by making a few remarks. The explicit formulae derived here for the two types of multistate neural networks allow one to estimate the dependence of the network capacity as a function of the number q neuron activity levels, cf [4]. In particular, it is interesting to compare the two kinds of multistate networks corresponding to the Potts and clock neurons. Note that (37) coincides (up to variance) with the corresponding formula (24) for the $q = 3$ Potts neural network. This could have been anticipated because the $q = 3$ clock model is equivalent to the $q = 3$ Potts one. More striking (at first sight) is the full agreement between (40) and the corresponding formula (12) for the Ising neural network. On the other hand it is clear that for $q = 4$ and the special choice of the initial conditions (29), fluctuations in the transversal directions $\theta_i^\nu + 1(+3)(\mathrm{mod}\, 4)$ are irrelevant, because for them $\cos(\cdot) = 0$.

In figure 1 we present the fixed points of equation (24) for the q-state Potts model. In the one-step approximation the transition from continuous behaviour of $m^*(\alpha)$ at $\alpha_c(q)$ to discontinuous behaviour occurs at $q = 3$. In figure 2 we present the dependence of the critical parameter α_c on q for $q \geqslant 2$.

The non-zero temperature case can be easily taken into account by the introduction into the dynamics of an additional thermal noise, see e.g. [6]. To go beyond the one-step calculations seems difficult, however. We are faced with the fundamental difficulty that one has to take into account a feedback, creating strong correlations and a complicated perturbation of the noises [7] which are independent and Gaussian only for the one-step parallel dynamics.

We gratefully acknowledge Reimer Kühn for reading the manuscript and valuable remarks. The first draft of this letter was written when one of the authors (VAZ) visited the Centre de Physique Théorique, CNRS-Luminy, and SFB 123 of Heidelberg University. He thanks both organizations for the warm hospitality and support.

References

[1] Amit D G 1989 *Modeling Brain Function* (Cambridge: Cambridge University Press)

[2] Hebb D O 1949 *The Organization of Behavior* (New York: Wiley)
[3] Little W A 1974 *Math. Biosci.* **19** 101
Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554
[4] Kanter I 1988 *Phys. Rev.* A **37** 2739
[5] Cook J 1989 *J. Phys. A: Math. Gen.* **22** 2057
[6] Patrick A E and Zagrebnov V A 1989 *Preprint* CPT-89/P 2311, Marseille (1991 *J. Stat. Phys.* submitted)
[7] Patrick A E and Zagrebnov V A 1991 A probabilistic approach to parallel dynamics for the Little–Hopfield model *J. Phys. A: Math. Gen.* submitted
[8] Monier O, Hansel D and Verga A 1989 *J. Stat. Phys.* **55** 859