Controlling the dynamics of multi-state neural networks

J. Stat. Mech. (2008) P06002

(http://iopscience.iop.org/1742-5468/2008/06/P06002)

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Controlling the dynamics of multi-state neural networks

Tao Jin and Hong Zhao[1]

Department of Physics, Xiamen University, Xiamen 361005, People's Republic of China
and
Institute of Theoretical Physics and Astrophysics, Xiamen University, Xiamen 361005, People's Republic of China
E-mail: jintao@xmu.edu.cn and zhaoh@xmu.edu.cn

Received 17 March 2008
Accepted 11 May 2008
Published 3 June 2008

Online at stacks.iop.org/JSTAT/2008/P06002
doi:10.1088/1742-5468/2008/06/P06002

Abstract. In this paper, we first analyze the distribution of local fields (DLF) which is induced by the memory patterns in the Q-Ising model. It is found that the structure of the DLF is closely correlated with the network dynamics and the system performance. However, the design rule adopted in the Q-Ising model, like the other rules adopted for multi-state neural networks with associative memories, cannot be applied to directly control the DLF for a given set of memory patterns, and thus cannot be applied to further study the relationships between the structure of the DLF and the dynamics of the network. We then extend a design rule, which was presented recently for designing binary-state neural networks, to make it suitable for designing general multi-state neural networks. This rule is able to control the structure of the DLF as expected. We show that controlling the DLF not only can affect the dynamic behaviors of the multi-state neural networks for a given set of memory patterns, but also can improve the storage capacity. With the change of the DLF, the network shows very rich dynamic behaviors, such as the 'chaos phase', the 'memory phase', and the 'mixture phase'. These dynamic behaviors are also observed in the binary-state neural networks; therefore, our results imply that they may be the universal behaviors of feedback neural networks.

Keywords: neuronal networks (theory), pattern formation (theory), network dynamics

[1] Author to whom any correspondence should be addressed.

1742-5468/08/P06002+11$30.00

Contents

1. Introduction

There are various models of multi-state neural networks. Among them, the Q-Ising model has been extensively studied in the past decades [1]–[8]. This model is significant not only in statistical physics as a generalization of the Ising model [9, 10], but also in network applications as a potential candidate for storing and retrieving colored patterns [11, 12]. Whether for theoretical studies or for practical applications, the dynamic behaviors of the network are crucial. It is known that the local fields play one basic role in the dynamics of neural networks. Recently, Bolle and Shim [13] have studied the time evolution of the local field in symmetric Q-Ising neural networks, and found that the equilibrium DLF differ in the retrieval phase and the spin-glass phase. Their work reveals that studying the local fields one can obtain useful information about the dynamic behaviors of neural networks.

However, there are two limitations in the investigation of Bolle and Shim. Firstly, they did not distinguish the roles of memory attractors and spurious attractors. The equilibrium DLF is actually composed of two parts: one part is induced by memory patterns which are stored as fixed-point attractors of the network; another part is induced by various spurious attractors. This two parts of the DLF play different roles in the dynamic behaviors of neural networks. Secondly, the equilibrium DLF is uncontrollable. The Hebb rule adopted in the Q-Ising model is a deterministic rule, and the neural networks designed using this rule are uniquely determined for a given set of memory patterns, so as to give the equilibrium DLF. This limitation prevents one from studying in detail the relationships between the DLF and the dynamic behaviors of neural networks.

In this paper, we first reveal the relationships between the DLF of memory patterns and the dynamic behaviors of multi-state neural networks with associative memories. For this purpose, in the next section, a fully connected Q-Ising neural network is considered. It will be noted that the DLF of memory patterns is quite different from the equilibrium DLF studied by Bolle and Shim, though it also changes continuously with increase of the memory patterns. Our studies show that it is the gap structure in the DLF of memory patterns that closely determines the dynamic behaviors of the network, such as the storage ratio of the network and the retrieval performance of memory patterns.

This fact motivates us to further investigate the dynamic behaviors of multi-state neural networks by controlling the DLF of memory patterns.

In section 3, we will extend the Monte Carlo adaptation (MCA) rule [14] to make it suitable for designing general multi-state neural networks with controllable DLF of memory patterns. The original MCA rule is proposed for designing binary-state neural networks with associative memories, and our extension is based on a general algorithm which is also applicable in the binary-state situation. Applying this extended algorithm, the gap structure in the DLF of memory patterns can be directly controlled by a single parameter. We will show that this parameter directly affects the retrieval properties of memory patterns, the amount of spurious attractors, and even the symmetry degree of the networks. More specifically, with increase of the parameter, we in turn observe the 'chaos phase' in which all random initial states are attracted to a chaotic orbit, the 'memory phase' in which all random initial states are attracted to the memory patterns, and the 'mixture phase' in which random initial states are attracted to either memory patterns or spurious attractors. These dynamic behaviors are also observed in the binary-state neural networks designed using the original algorithm of the MCA rule [14]; therefore, the findings in this paper imply that they may be the universal behaviors of the feedback neural networks.

We present the summary and discussion in section 4.

2. Role of the DLF of memory patterns in the Q-Ising model

A Q-Ising neural network is a dynamic system whose evolution is determined by

$$s_i(t+1) = \sigma(h_i(t)), \qquad h_i(t) = \sum_{j=1}^{N} J_{ij} s_j(t), \qquad i = 1, \ldots, N. \tag{1}$$

Here, N is the number of neurons, $s_i(t)$ is the state of ith neuron that can be any one of $\{\sigma_l = -1 + 2(l-1)/(Q-1), l = 1, \ldots, Q\}$, and $h_i(t)$ is the instantaneous local field of the ith neuron. For a fully connected Q-Ising neural network, the synaptic matrix J_{ij} is obtained using the Hebb rule

$$J_{ij} = \frac{1}{AN} \sum_{\mu=1}^{p} \xi_i^{\mu} \xi_j^{\mu}, \qquad i, j = 1, \ldots, N, \tag{2}$$

where $\xi_i^{\mu} \in \{\sigma_l\}$ is the ith bit of the μth memory pattern. A memory pattern is a network state denoted by $\{\xi_i^{\mu}\}$, and the parameter A measures its activity [7]. In this paper, we restrict to $Q = 4$ and take the memory patterns with uniform distribution. For this case, $A = \frac{5}{9}$. The storage ratio of network is defined as $\alpha = p/N$.

In the context of associative memory, the basic problem is storing memory patterns. A successfully stored memory pattern should be a fixed-point attractor of network. That is, it should satisfy

$$\xi_i^{\mu} = \sigma(h_i^{\mu}), \qquad h_i^{\mu} = \sum_{j=1}^{N} J_{ij} \xi_j^{\mu}, \qquad i = 1, \ldots, N, \tag{3}$$

where h_i^{μ} is the local field induced by μth memory pattern. We call (3) the fixed-point condition. It is known that the Q-Ising model works well on storing memory patterns as fixed-point attractors for small α, but not for large values. Low storage capacity is one

J. Stat. Mech. (2008) P06002

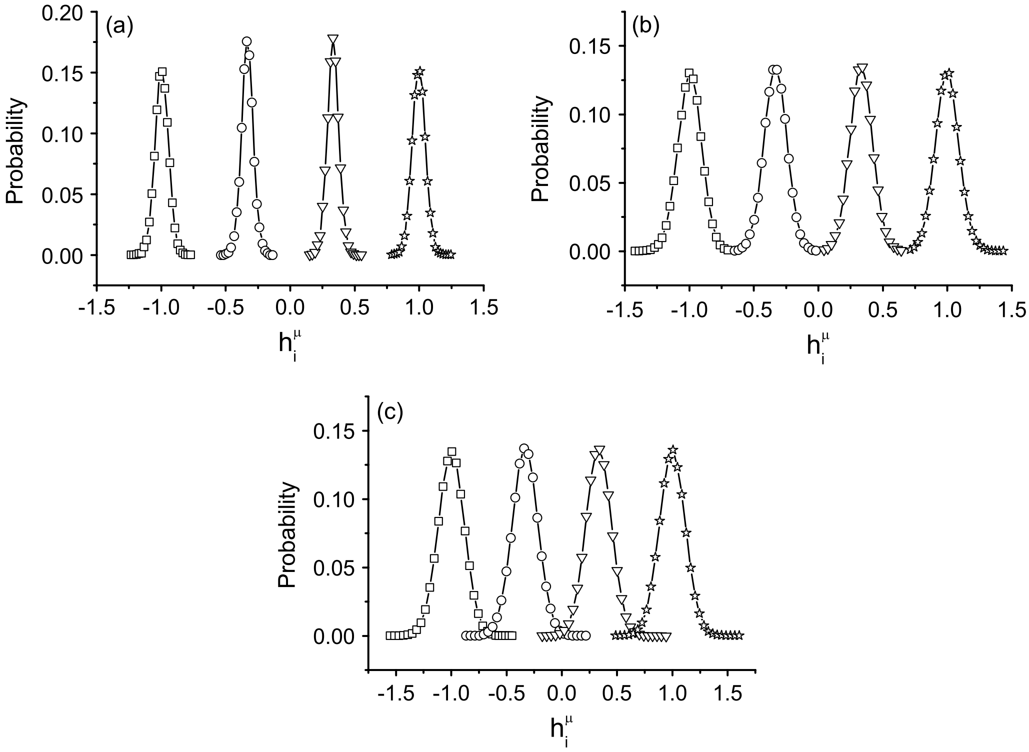

Figure 1. The DLF of memory patterns for the Q-Ising neural network with $Q = 4$ and $N = 1000$: (a) $\alpha = 0.005$, (b) $\alpha = 0.015$ and (c) $\alpha = 0.025$. The squares, circles, triangles, and stars represent the classes of Σ_1, Σ_2, Σ_3 and Σ_4, respectively.

reason that the Q-Ising model is unfavorable for practical application. To understand why the memory patterns may not be successfully stored as fixed-point attractors for large α, we analyze the probability distribution of h_i^μ, i.e., the DLF of memory patterns.

Our analysis starts with classifying h_i^μ into four classes Σ_l, i.e., $h_i^\mu \in \Sigma_l$ if $\xi_i^\mu = \sigma_l$, $l = 1, \ldots, 4$. Figure 1 shows the DLF of memory patterns for $\alpha = 0.005$, 0.015, and 0.025. In these plots, different classes Σ_l are marked by different symbols, and each plot is obtained by averaging over 20 sets of randomly selected memory patterns. Here and throughout this paper, we take $N = 1000$ in our numerical computation unless a special emphasis is applied.

The role of the evolution function $\sigma(x)$ is to divide the space of local fields into Q parts and for each part to assign a unique value which corresponds to the Q neuron states. Therefore, if different classes of h_i^μ are isolated by gaps in the space of local fields, as shown in figure 1(a), it is easy to find an evolution function $\sigma(x)$ ensuring that the fixed-point condition is satisfied for all of memory patterns. In this situation, there is some freedom to select the partition points in the corresponding gaps. For the case of figure 1(a), for example, the most favorable partition seems to be $h_i^\mu \in (-\infty, -\frac{2}{3})$ for Σ_1, $h_i^\mu \in (-\frac{2}{3}, 0)$ for Σ_2, $h_i^\mu \in (0, \frac{2}{3})$ for Σ_3, and $h_i^\mu \in (\frac{2}{3}, +\infty)$ for Σ_4. In fact, the evolution function

$$\sigma(h_i(t)) \equiv \sum_{l=1}^{Q} \sigma_l \{\Theta[B(\sigma_{l+1} + \sigma_l) - h_i(t)] - \Theta[B(\sigma_l + \sigma_{l-1}) - h_i(t)]\} \qquad (4)$$

with $B = 0.5$ satisfies the requirement of the above partition. Here, $\sigma_0 = -\infty$, $\sigma_{Q+1} = +\infty$, $\Theta(x) = 1$ for $x \geq 0$ and $\Theta(x) = 0$ for others. Parameter B decides the partition points, i.e., different B means different partition points. This evolution function is commonly adopted in the Q-Ising model [10, 13], and can be equivalently explained as: $s_i(t+1)$ takes the value of σ_l if $h_i(t)$ is bound by $\sigma_l + \sigma_{l-1} < h_i(t)/B < \sigma_l + \sigma_{l+1}$. This will be adopted in the present paper.

The variance of each class of h_i^μ increases with the increase of α. For the network considered above, the four classes of h_i^μ begin to overlap each other when $\alpha \approx 0.015$ as shown in figure 1(b). This crossover means that the four classes of h_i^μ combine into one inseparable set, and thus cannot be partitioned, no matter how one selects $\sigma(x)$. More importantly, it means that the fixed-point condition is no longer to be perfectly satisfied for all of the memory patterns, because part of h_i^μ surely extends out of the corresponding class boundaries, no matter how one selects the boundaries. For the case of $\alpha \approx 0.025$, one may notice the obvious crossover as shown in figure 1(c). In this situation, we have checked that only about 30% of memory patterns are successfully stored as fixed-point attractors. The value of α at which the gaps in the DLF of memory patterns disappear defines a critical storage ratio of the network below which all of memory patterns are surely stored as fixed-point attractors. In other words, it is the gap structure in the DLF of memory patterns that closely determines whether the memory patterns are surely stored as fixed-point attractors or not. Actually, under condition (3), the critical storage ratio of the Q-Ising model is expected to approach zero in the thermodynamics limit, i.e., $\alpha \to 0$ for $N \to \infty$, as in the Hopfield model [7, 15].

Being stored as fixed-point attractor is the basic requirement of a memory pattern. For associative memory, a memory pattern should be stored efficiently, which means that the attraction basin should be large enough. To measure the attraction basin size of the μth memory pattern, a direct way is to calculate the fraction of random initial states attracted to it, P_μ. Here, we say that an initial state is attracted to a memory pattern only if they have perfect overlap. Therefore, $P_{\text{total}} \equiv \sum_{\mu=1}^{M} P_\mu$ measures the relative size of the attraction basins of memory patterns, and $1 - P_{\text{total}}$ measures that of spurious attractors. This quantity is commonly employed to measure the retrieval performance of neural networks [14, 16, 17].

In figure 2(a), we give P_{total} versus B with $\alpha = 0.004$. One may note that the value of P_{total} depends upon B and reaches the maximum, i.e., $P_{\text{total}} \approx 0.3$, at $B \approx 0.38$. In fact, the value of B at which P_{total} reaches the maximum is independent of α. Fixing $B = 0.38$, we study the correlation between P_{total} and α, which is shown in figure 2(b). These results imply that, for the Q-Ising neural network considered in this paper, the maximum of P_{total} is about 0.3 which is obtained at $B \approx 0.38$ and $\alpha \approx 0.004$. In other words, no more than 30% of the random initial states can be attracted to the memory patterns. Poor retrieval performance is another reason that the Q-Ising model is unfavorable for practical application. Although the choice of partition points may improve the retrieval performance as shown in figure 2(a), the key factor is the width of the gaps in the DLF of memory patterns.

In section 3, we will show that by controlling the gap structure in the DLF of memory patterns, not only can the storage capacity be enlarged, but also the retrieval performance can be dramatically improved.

J. Stat. Mech. (2008) P06002

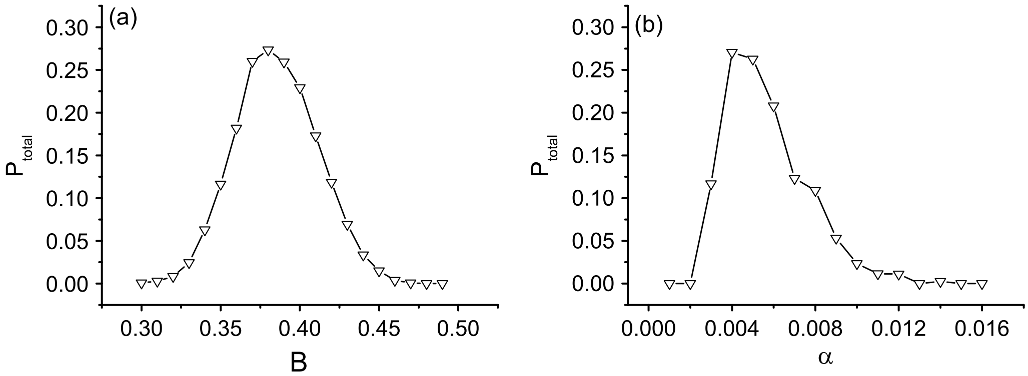

Figure 2. The Q-Ising neural network with $Q = 4$ and $N = 1000$: (a) P_{total} versus B for $\alpha = 0.004$; (b) P_{total} versus α for $B = 0.38$.

3. Control of the DLF of memory patterns using the extended MCA rule

From the previous section, we know that it is the gap structure in the DLF of memory patterns that directly affects the dynamic behaviors of multi-state neural networks with associative memories, e.g., the disappearance of the gap means that not all of the memory patterns are surely stored as fixed-point attractors. This implies that the dynamic behaviors of networks may be controlled if the gap structure in the DLF of memory patterns is controllable. In fact, the gap structure is based on a design rule. The Hebb rule adopted in the Q-Ising model is a deterministic rule. Applying this rule, the DLF of memory patterns is only influenced by the storage ratio of network α, and cannot be further controlled for a given set of memory patterns. In this section, we extend the MCA rule to design general multi-state neural networks with a controllable DLF of memory patterns.

Applying the MCA rule to design multi-state neural networks, the value of the neuron states σ_l can be selected quite flexibly. As an example, we consider a four-state neural network with neuron states $\sigma_1 = -2$, $\sigma_2 = -1$, $\sigma_3 = 1$, and $\sigma_4 = 2$. The corresponding evolution function is thus selected as

$$\sigma(h_i(t)) \equiv \begin{cases} -2 & h_i(t) \leq -U \\ -1 & -U < h_i(t) < 0 \\ +1 & 0 < h_i(t) < U \\ +2 & h_i(t) \geq U \end{cases} \tag{5}$$

and the local field $h_i(t)$ is still calculated as it is in (1). Here, U is a constant which determines the partition points of the space of local fields.

For designing multi-state neural networks, we should extend the original algorithm of the MCA rule. The purpose of this extended algorithm is to find one synaptic matrix J_{ij} not only ensuring that a given set of memory patterns are stored as fixed-point attractors, but also ensuring that all of the h_i^μ distribute over designated regions in the space of local fields. The value of J_{ij} is not limited in principle, but for the sake of simplicity, in this paper we restrict to $|J_{ij}| = 1$. To start the design procedure, we randomly assign binary numbers ± 1 to J_{ij}, and $J_{ii} = 0$ for all of the i as usual. A set of p independent and

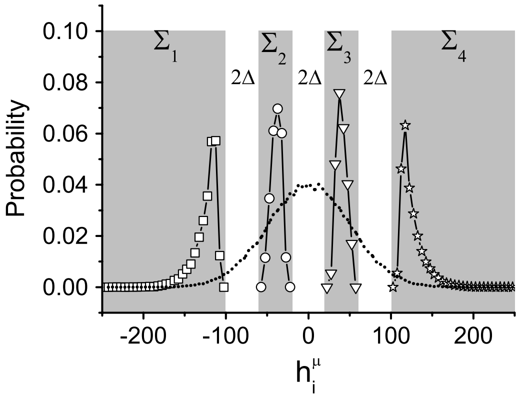

Figure 3. The DLF of memory patterns for the neural network designed using the MCA rule with $N = 1000$, $\alpha = 0.015$ and $\Delta = 50$. The squares, circles, triangles, and stars represent the classes of Σ_1, Σ_2, Σ_3 and Σ_4, respectively. The short-dot line represents the probability distribution of h_i^μ before proceeding with the design.

identically distributed random memory patterns $\{\xi_i^\mu\}$ $(\mu = 1, 2, \ldots, p)$ are selected to be stored as fixed-point attractors. Here, $\xi_i^\mu = \sigma_l$ with $\sigma_l \in \{-2, -1, 1, 2\}$. The class of h_i^μ is defined as $h_i^\mu \in \Sigma_l$ if $\xi_i^\mu = \sigma_l$, $l = 1, \ldots, 4$.

At the beginning, the J_{ij} and $\{\xi_i^\mu\}$ are selected randomly; therefore, h_i^μ should satisfy a Gaussian distribution with zero mean as shown in figure 3. For this situation, the purpose of the extended algorithm of the MCA rule is to 'drive' h_i^μ into the four corresponding intervals: $(-\infty, -U - \Delta]$ for Σ_1, $[-U + \Delta, -\Delta]$ for Σ_2, $[\Delta, U - \Delta]$ for Σ_3, and $[U + \Delta, +\infty)$ for Σ_4 as shown in figure 3 by the gray shadows. The width of the gap in the DLF of memory patterns is 2Δ, and $\Delta \geq 0$ means that all of the memory patterns are successfully stored as fixed-point attractors with non-vanishing attraction basin. In fact, the Δ determines the size of the attraction basin of memory patterns. The relationship between the width of the gap and the size of the attraction basin has been discussed in detail in [14] for binary-state neural networks, which can be directly extended to the multi-state situation.

For convenience in describing the extended algorithm of the MCA rule, we introduce a new parameter as follows:

$$b_i^\mu = \begin{cases} -\left|\bar{h}_i^\mu - \dfrac{U}{2}\right| + \left(\dfrac{U}{2} - \Delta\right) & |\xi_i^\mu| = 1 \\ \bar{h}_i^\mu - (U + \Delta) & |\xi_i^\mu| = 2 \end{cases} \tag{6}$$

with $\bar{h}_i^\mu = \mathrm{sgn}(\xi_i^\mu) h_i^\mu$. If $b_i^\mu \geq 0$ for $\mu = 1, \ldots, p$ and $i = 1, \ldots, N$, one may easily derive that $\Delta \leq \bar{h}_i^\mu \leq U - \Delta$ for $|\xi_i^\mu| = 1$ and $\bar{h}_i^\mu \geq U + \Delta$ for $|\xi_i^\mu| = 2$, which means that the four classes of h_i^μ have been correctly 'driven' into the desired regions. In other words, $b_i^\mu \geq 0$ implies perfect satisfaction of $\Sigma_1 \subseteq (-\infty, -U - \Delta]$, $\Sigma_2 \subseteq [-U + \Delta, -\Delta]$, $\Sigma_3 \subseteq [\Delta, U - \Delta]$, and $\Sigma_4 \subseteq [U + \Delta, +\infty)$. Therefore, $b_i^\mu \geq 0$ serves as the criterion for our design purpose being achieved. It is also employed as the stop condition of the design procedure, i.e., the design procedure is stopped when this criterion is satisfied.

J. Stat. Mech. (2008) P06002

The extended algorithm of the MCA rule is described in the following design procedure. Since $\bar{h}_i^\mu$ $(=\mathrm{sgn}(\xi_i^\mu)\sum_{j=1}^N J_{ij}\xi_j^\mu)$ is only correlated with the ith row of J_{ij}, the synaptic matrix can be designed row by row, independently. We apply three steps to finish the design of one row. In the first step, we calculate $\{b_i^\mu, \mu = 1,\ldots,p\}$ and find the minimum $b_i^{\min}$ of this set. There are usually many terms taking the same minimum $b_i^{\min}$, and the goal of this step is to find the set $\{b_i^{\mu_1},\ldots,b_i^{\mu_m}\}$ satisfying the condition $b_i^\mu = b_i^{\min}$ for $\mu \in \{\mu_1,\ldots,\mu_m\}$. In the second step, for each $J_{ij} \in \{J_{ij}, j = 1,\ldots,N, j \neq i\}$ we calculate $\{C_j^\mu\,\mathrm{sgn}(\xi_i^\mu)J_{ij}\xi_j^\mu, \mu = \mu_1,\ldots,\mu_m\}$ and count the number of negative terms m_i^j in each subset. Here, $C_j^\mu = -1$ for $|\xi_i^\mu| = 1$ and $\bar{h}_j^\mu > U/2$ or $C_j^\mu = +1$ for others. Let $m_i^{\max}$ represent the biggest value of m_i^j. Again, many terms may take the same value $m_i^{\max}$. We then record the indexes of $j \in \{j_1,\ldots,j_n\}$ satisfying $m_i^j = m_i^{\max}$. In the third step, we randomly pick an index j from the set $\{j_1,\ldots,j_n\}$ and make an adaptation $J_{ij} \to -J_{ij}$. This adaptation changes the sign of $\mathrm{sgn}(\xi_i^\mu)J_{ij}\xi_j^\mu$. As a result, there are $m_i^{\max}$ terms in $\{\bar{h}_i^{\mu_1},\ldots,\bar{h}_i^{\mu_m}\}$ that will be pushed towards the interval $[\Delta, U-\Delta]$ for $|\xi_i^\mu| = 1$ or the interval $[U+\Delta,+\infty)$ for $|\xi_i^\mu| = 2$. It is easy to find that $m_i^{\max} \geq m/2$ in general. Therefore, by continuously repeating these three steps, the set $\{\bar{h}_i^\mu\}$ will be pushed toward the desired intervals continuously until $b_i^\mu \geq 0$ is satisfied for $\mu = 1,\ldots,p$. Applying the same procedure to each row of J_{ij}, we obtain a network with $b_i^\mu \geq 0$ for $i = 1,\ldots,N$ and $\mu = 1,\ldots,p$. As an example, in figure 3 we plot the DLF of memory patterns with $\alpha = 0.014$ and $\Delta = 25$. It clearly shows that all of the h_i^μ have been correctly 'driven' into the desired intervals.

From the above description, we know that both U and Δ are control parameters in the extended algorithm of the MCA rule. They decide the partition points of the space of local fields and the width of the gaps in the DLF of the memory patterns, respectively. For the sake of simplicity, we fix $U - 2\Delta = N/100$ in the numerical computation of this paper. That is, we fix the widths of the two middle intervals of the DLF and make the width of the gap as big as possible.

The symmetry degree of the synaptic matrix J_{ij} can be defined as $\eta = 2\Gamma/N(N-1)$ [14], where Γ represents the number of symmetry elements. For a symmetric matrix, we have $\Gamma = N(N-1)/2$, and thus $\eta = 1$. Figure 4(a) shows η as a function of Δ for different storage ratios. One may note that η increases with the increase of Δ. In other words, the parameter Δ directly affects the symmetry degree of the synaptic matrix.

Compute P_{total} as a function of Δ; one may get the phase diagram of networks with different α as shown in figure 4(b). In this paper, each P_{total} is obtained by averaging over five sets of randomly selected memory patterns, while for each set of memory patterns 10 000 random initial states are checked. According to the value of P_{total}, for a certain α, the axis of Δ can be divided into three intervals by two points, named Δ_1 and Δ_2. In the range of $\Delta \leq \Delta_1$, we have $P_{\mathrm{total}} = 0$, which means almost no random initial state is attracted to the memory patterns. Detailed studies show that the random initial states are attracted to one chaotic attractor, and the attraction basins of the memory patterns are like some islands embedded in a 'chaos sea'. This range is called the 'chaos phase', and neural networks with this phase have particular superiority in pattern recognition [15]. In the range of $\Delta_1 < \Delta \leq \Delta_2$, we have $P_{\mathrm{total}} = 1$ which means that almost all random initial states are attracted to the memory patterns, that is, the attraction basins of the memory patterns almost fill up the whole state space. This range is called the 'memory phase', and

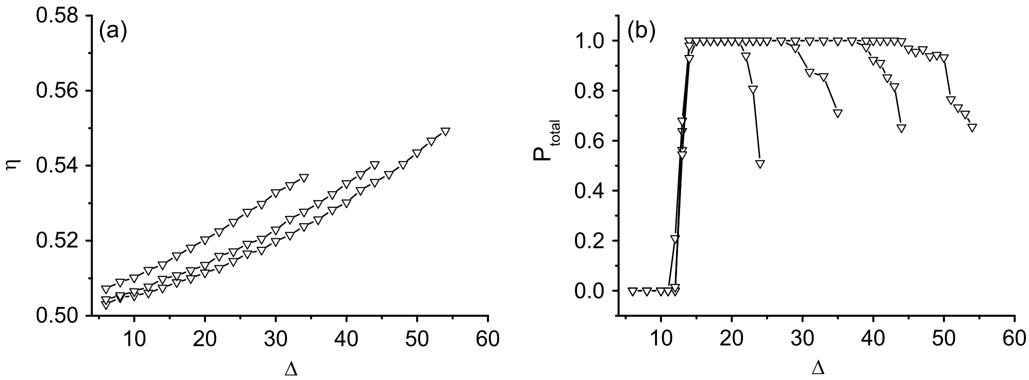

Figure 4. The neural network with $N = 1000$ designed using the MCA rule: (a) η versus Δ with, from top to bottom, $\alpha = 0.020, 0.016, 0.014$; (b) P_{total} versus Δ with, from left to right, $\alpha = 0.022, 0.020, 0.016, 0.014$ with $\Delta_2 \approx 22, 28, 38, 44$, respectively. For this situation, we have $\Delta_1 \approx 14$.

neural networks with this phase are more suitable for the purpose of associative memory than the Q-Ising model, because they have perfect retrieval performance. In the range of $\Delta > \Delta_2$, P_{total} decreases with increase of Δ, which means that the attraction basins of the spurious attractors expand with Δ. This range is called the 'mixture phase'. These dynamical phases are also observed in the binary-state neural networks designed using the original algorithm of the MCA rule; therefore, they may be universal behaviors of feedback neural networks.

To show the effects of system size, we plot the phase diagram in the α–Δ plane. The results for the four-state neural networks with $N = 800$, 1000, and 1500 are given in figures 5(a)–(c), respectively. For a certain N, it is found that Δ_1 is almost unchanged, while Δ_2 decreases with increase of α. As a result, the 'memory phase' disappears when α exceeds a critical value α_{c}. Although both Δ_1 and Δ_2 are different for different N, the critical storage ratio α_{c} seems independent of the network size. As can be found in the three plots, $\alpha_{\text{c}} \approx 0.024$ for any one of the three cases.

Figure 5(d) shows the phase diagram of a six-state neural network. For this network, the neuron states are taken as $\{\sigma_1 = -3, \sigma_2 = -2, \sigma_3 = -1, \sigma_4 = 1, \sigma_5 = 2, \sigma_6 = 3\}$, and the corresponding evolution function is

$$\sigma(h_i(t)) \equiv \begin{cases} -3 & h_i(t) \le -2U \\ -2 & -2U < h_i(t) \le -U \\ -1 & -U < h_i(t) < 0 \\ +1 & 0 < h_i(t) < U \\ +2 & U \le h_i(t) < 2U \\ +3 & h_i(t) \ge 2U. \end{cases} \qquad (7)$$

The local field $h_i(t)$ is also calculated as it is in (1), and parameter U still determines the partition points of the space of local fields. The purpose of this plot is to show that the extended algorithm of the MCA rule can be applied to design general multi-state neural networks with associative memories. Moreover, one may note that these neural networks may have similar dynamic behaviors.

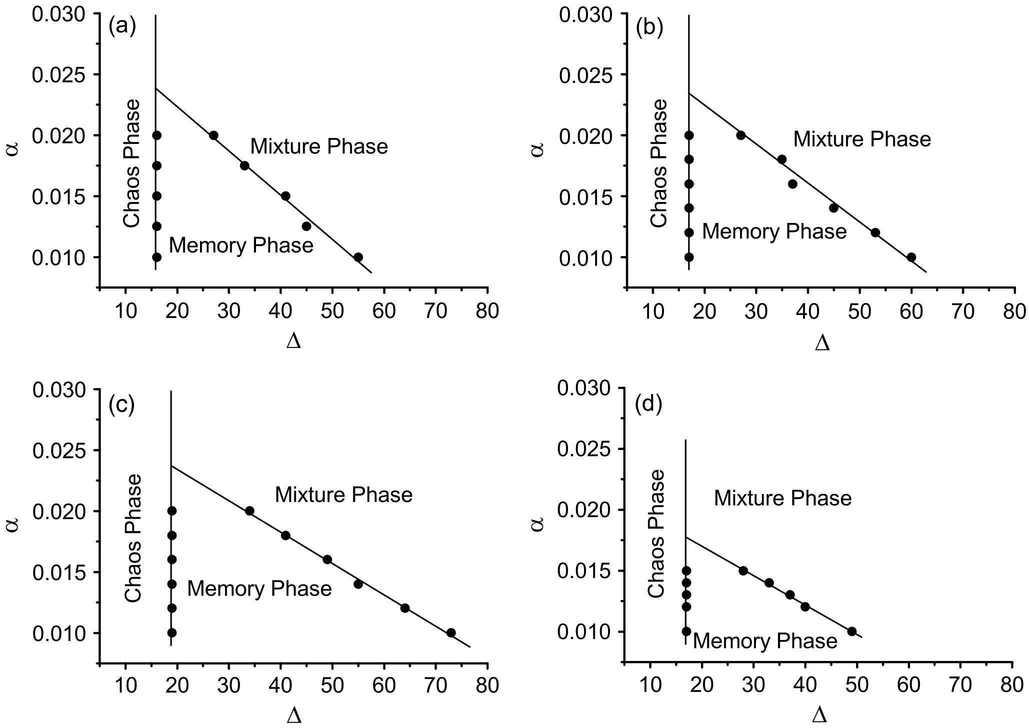

Figure 5. The phase diagrams of neural networks with (a) four-state neurons and $N = 800$, (b) four-state neurons and $N = 1000$, (c) four-state neurons and $N = 1500$, and (d) six-state neurons and $N = 1000$.

4. Summary and discussion

We have presented an extended algorithm for the MCA rule to design multi-state neural networks with associative memories. This algorithm is quite different from the Hebbian-like rules. When applying a Hebbian-like rule, the DLF of memory patterns is uniquely determined for a fixed set of memory patterns. This is because the Hebbian-like rules are described as deterministic functions of memory patterns. For different sets of memory patterns, we have found that in the Q-Ising model the DLF of memory patterns has different gap structures, which are closely related to the dynamic behaviors of neural networks. However, for a fixed set of memory patterns, these gap structures are not controllable by applying a Hebbian-like rule. In contrast, by applying the extended algorithm of the MCA rule, the gap structure in the DLF of memory patterns can be directly controlled. Controlling the gap structure, we have studied in detail the correlations between the DLF of memory patterns and the dynamic behaviors of multi-state neural networks. On the basis of these correlations, one not only may control the dynamic behaviors of the neural networks, but also can improve their storage capacity.

For a fixed set of memory patterns, we have shown that the neural networks may exhibit quite rich dynamic behaviors with change of the gap structure. As for the binary-state neural networks [14], we also observed a 'chaos phase', a 'memory phase', and a 'mixture phase' in the multi-state neural networks. This implies that these dynamic behaviors are universal in the feedback neural networks. However, as far as we know,

J. Stat. Mech. (2008) P06002

unlike 'mixture phases' which have been widely found in binary-state and multi-state neural networks with associated memories, both 'chaos phases' and 'memory phases' have not been observed in the neural networks designed using Hebbian-like rules. The reason is as follows: the 'chaos phases' and 'memory phases' are observed only when the DLF of memory patterns have certain special structures, as is shown in the previous section; these special structures can be easily obtained using the MCA rules, but are hardly encountered when using Hebbian-like rules. The 'mixture phase' actually corresponds to the 'retrieval phase' observed in the previous works. It is worth pointing out that, although the 'retrieval phase' is sometimes called a 'memory phase' in the previous studies, it is different from the 'memory phase' mentioned in this paper. In the 'retrieval phase', the network may have spurious attractors coexisting with the memory patterns, while in the 'memory phase' mentioned in this paper, the memory patterns are attractors of unique types.

The dynamical behaviors of multi-state neural networks designed using the extended algorithm of the MCA rule have also been studied as a function of α, and have been shown globally in the α–Δ parameter plane. An important finding is that the critical storage ratio α_c, below which the 'memory phase' exists, seems independent of the network size. Note that for the neural networks designed in the 'memory phase' all memory patterns are stored as fixed-point attractors with perfect retrieval performance. However, for Q-Ising neural networks, the critical storage ratio α, below which all memory patterns are only guaranteed to be stored as fixed-point attractors, is expected to approach zero in the thermodynamic limit, i.e., $\alpha \to 0$ for $N \to \infty$, as in the case of the Hopfield model [7, 15].

Finally, both the 'chaos phase' and the 'memory phase' are important not only for understanding the global properties of feedback neural networks but also for the practical applications [16]. Therefore, they merit further investigations in the future.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant No. 10475067, the National Basic Research Program of China (973 Program) (2007CB814800), and NCETXMU of Xiamen University.

References

[1] Noest A J, 1988 *Phys. Rev.* A **38** 2196
[2] Bolle D and Shim G M, 1994 *Phys. Rev.* E **50** 5043
[3] Meunier C, Hansel D and Verga A, 1989 *J. Stat. Phys.* **55** 859
[4] Bolle D, Dupont P and Vinch B, 1992 *J. Phys. A: Math. Gen.* **25** 2859
[5] Bolle D, Vinck B and Zagrebnov V A, 1993 *J. Stat. Phys.* **70** 1099
[6] Bolle D, Shim G M, Vinck B and Zagrebnov V A, 1994 *J. Stat. Phys.* **74** 565
[7] Bolle D, Rieger H and Shim G M, 1994 *J. Phys. A: Math. Gen.* **27** 3411
[8] Bolle D, Jongen G and Shim G M, 1998 *J. Stat. Phys.* **91** 125
[9] Bolle D, Jongen G and Shim G M, 1999 *J. Stat. Phys.* **96** 861
[10] Theumann W K and Erichsen R, 2001 *Phys. Rev.* E **64** 061902
[11] Rieger H, 1990 *J. Phys. A: Math. Gen.* **23** L1273
[12] Erichsen R and Theumann W K, 1999 *Phys. Rev.* E **59** 947
[13] Bolle D and Shim G M, 2002 *Phys. Rev.* E **65** 067101
[14] Zhao H, 2004 *Phys. Rev.* E **70** 066137
[15] Amit D J, Gutfreund H and Sompolinsky H, 1985 *Phys. Rev. Lett.* **55** 1530
[16] Jin T and Zhao H, 2005 *Phys. Rev.* E **72** 066111
[17] Hopfield J J, Feinstein D I and Palmer R G, 1983 *Nature* **304** 159