A trade-off formula in designing asymmetric neural networks

2013 J. Phys. A: Math. Theor. 46 435002

(http://iopscience.iop.org/1751-8121/46/43/435002)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **46** (2013) 435002 (10pp) doi:10.1088/1751-8113/46/43/435002

A trade-off formula in designing asymmetric neural networks

Daxing Xiong[1,3] **and Hong Zhao**[2]

[1] Department of Physics, Fuzhou University, Fuzhou 350108, People's Republic of China
[2] Department of Physics and Institute of Theoretical Physics and Astrophysics, Xiamen University, Xiamen 361005, People's Republic of China

E-mail: phyxiongdx@fzu.edu.cn and zhaoh@xmu.edu.cn

Received 22 May 2013, in final form 20 August 2013
Published 8 October 2013
Online at stacks.iop.org/JPhysA/46/435002

Abstract
We show that for asymmetric neural networks the symmetric degree η of the synaptic coupling can be related to the two main network parameters, the storage capacity α and another designing parameter κ by the formula $\eta = \alpha\kappa^2$. Such a relation has been well verified by the simulations of our neural network designing. The formula suggests that we cannot improve the network performances by tuning the parameters α and κ simultaneously. The result may provide useful information for optimizing the designing of asymmetric neural networks.

PACS numbers: 0705.Mh, 0510.Ln, 0570.Fh

(Some figures may appear in colour only in the online journal)

1. Introduction

In the past three decades, neural networks with associative memories have become a favored research topic, both for their theoretical significance as a generalization of the Ising-like model well known in statistical physics, and for their potential applications in many fields, such as pattern recognition, image manipulation, optimization problem and so on [1–4]. A neural network of this kind can generally be classified into two categories: symmetric ones [5–9] and asymmetric ones [6, 10–13], according to whether the synaptic couplings, $J_{i,j}$, between pairs of neurons, satisfy $J_{i,j} = J_{j,i}$ $(i \neq j)$. For symmetric neural networks, the storage capacity, defined by $\alpha = \frac{p}{N}$, where p is the number of memories to be stored and N is the size of neural networks, is a predominant parameter for evaluating the network performance. Over the past two decades, many studies have been devoted to estimating the storage capacity and so far some theoretical formulas have been proposed to perform such an estimate correctly (see [9] and references therein). However, symmetric neural networks have a serious limitation in their applications: a vast number of spurious memories exist in these networks causing them to fail

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

to recall the correct memories [13]. To suppress the spurious memories, asymmetric neural networks were then suggested and it has been found that by introducing asymmetric synaptic couplings, the spurious memories can indeed be diminished to some extent. To achieve this goal these asymmetric neural networks need to introduce an additional designing parameter (here we denote it as κ and will describe it in the next section). The parameter κ is correlated to the degree of symmetry of $J_{i,j}$, and more importantly, it is also closely related to the network performances: for example, it has been shown that by turning κ the spurious memories can be suppressed completely and the attraction basin of the memories can be made controllable [13]. Therefore compared with symmetric neural networks, there may be two main designing parameters, α and κ, that can be used to evaluate the network performances in asymmetric neural networks. Nevertheless, an explicit understanding of the relation between α and κ is still lacking. Thus, it would be desirable to gain more knowledge of this relation before designing an asymmetric neural network for applications.

Since the information of neural networks is primarily stored in their synaptic couplings, the investigation on the synaptic matrices' eigenvalue spectrum via the random matrix theory may provide useful clues to the above problem. It is well known that the statistical properties of the eigenvalue of random matrices have been proved as a powerful tool in the study of various problems throughout mathematics and physics [14]. Here we just recall some results relevant to our focused problem. For asymmetric random matrix theory, the first celebrated result is the circular law proposed by Girko [15], which states that the eigenvalues of an N-dimensional asymmetric random matrix whose elements are drawn from a Gaussian distribution with zero mean and variance $\frac{1}{N}$, uniformly lie within a circle in the complex plane. Such an investigation has been subsequently generalized to the matrix with a certain symmetry degree. It is found that the circle changes to an ellipse whose semi-axes depend on the symmetry degree [16]. However, these classic results cannot be directly applied to the realistic neural network models where neurons do not behave in the same way [17–20]. In view of this fact, Rajan and Abbott [19] present a first study of the spectrum of two-state (two-component) asymmetric neural networks. They assume that the neurons are either 'excitatory' or 'inhibitory', and the strengths of the synapses from different types of neurons have Gaussian distributions with different means and variances. Based on those assumptions they find that the eigenvalues of the synaptic matrices would lie within a circle if a certain balance condition is satisfied and the condition mainly depends on the variances of different distributions rather than the means. They also suggest that the appearance of eigenvalues at the edge of the circle implies the existence of slow-oscillating and long-lasting modes. Quite recently, this general conclusion was also extended to the case of multi-state neurons [20]. Another representative investigation on the eigenvalue spectrum of asymmetric neural networks is conducted by Zhou *et al* [21]. They find that two essentially different dynamical phases, the so-called 'chaos phase' and 'memory phase' that have already been found to exist in a two-state asymmetric neural network [13], can be correlated to the very distinct eigenvalue spectrum of the network synaptic couplings, namely, in the former phase all eigenvalues lie uniformly within a circle in the complex plane, however, in the latter phase the spectrum is split into two parts: one represents the noise part within the circle and the other corresponds to the information part outside the circle. Those results provide the first direct evidence of the correlation between the eigenvalue spectra and the dynamics of neural networks.

In this paper we shall perform a detailed investigation on the eigenvalue spectrum of two-state neural networks with asymmetric synaptic couplings. Through studying the dependence of the spectrum on the two main network parameters, α and κ, we are able to present a formula $\eta = \alpha\kappa^2$ that relates the symmetric degree η of the synaptic couplings with α and κ, which may enable us to optimize the designing of this kind of asymmetric neural network. The rest of this paper is organized as follows. The asymmetric neural network model to be studied will be described in the next section. The results on the spectrum of this model will be provided in section 3. The detailed derivations of the formula between η, α and κ will be presented in section 4, followed by a brief summary in the final section.

2. Model

We consider the two-state asymmetric neural network model, which is an attractor network usually composed of N neurons connected to each other through a synaptic coupling $J_{i,j}$. This kind of neural network is usually assumed to follow the following parallel dynamics [22] as

$$h_i(t) = \sum_{j=1}^{N} J_{i,j}s_j(t);\ s_i(t+1) = \mathrm{sgn}(h_i(t)),\ i, j = 1, \ldots, N, \tag{1}$$

where $h_i(t)$ is the instantaneous local field of the ith neuron at time t, s_i is the state of the ith neuron that may take the value 1 or -1, $s_i(t+1) = 1$ for $h_i(t) \geqslant 0$ and $s_i(t+1) = -1$ for $h_i(t) < 0$, respectively, after applying the function sgn.

For this kind of neural network, the central task is to find $J_{i,j}$, which guarantees a given set of p memories $\{\xi_i^\mu, i = 1, \ldots, N; \mu = 1, \ldots, p\}$ as the fixed point attractors of the networks. To be fixed point attractors, they should satisfy

$$h_{\xi_i^\mu} = \sum_{j=1}^{N} J_{i,j}\xi_j^\mu;\ \xi_i^\mu = \mathrm{sgn}(h_{\xi_i^\mu}), \tag{2}$$

or equivalently,

$$\bar{h}_{\xi_i^\mu} \geqslant \kappa, \tag{3}$$

where $\bar{h}_{\xi_i^\mu} = \xi_i^\mu h_{\xi_i^\mu}$ and the designing parameter κ must be positive. Thus, the problem of finding the solutions of $J_{i,j}$ is converted to a problem of solving a system of linear inequalities $\bar{h}_{\xi_i^\mu} \geqslant \kappa$, where each memory pattern gives rise to a system of N inequalities. The probability of solving this system of inequalities has first been carefully studied by Gardner and his collaborators [10, 11]. Several subsequent algorithms (the so-called learning rules), mainly classified as the methods of deterministic type [11, 12] and probabilistic type [13], have been proposed to derive the $J_{i,j}$. We have made a detailed comparison with these methods and found that the deterministic successive over-relaxation (SOR) rule [12] has a more rapid convergence over other methods. Therefore we shall adopt the SOR rule for our study here. In the following we shall only briefly describe the SOR rule but not present its mathematical derivation and proof of convergence, for the latter one can refer to [23].

The 'standard' SOR rule proposed in [12] is applicable for the condition $\bar{h}_{\xi_i^\mu} \geqslant \kappa$. Here, for the convenience of the later analysis, we shall mainly restrict our focus to the extreme case satisfying $\bar{h}_{\xi_i^\mu} = \kappa$. Then given the memory set $\{\xi_i^\mu, i = 1, \ldots, N; \mu = 1, \ldots, p\}$, one can realize the designing by performing the following adaptation rule:

$$J_{i,j}(t+1) = J_{i,j}(t) - \frac{1+\beta}{N}\left(h_{\xi_i^\mu} - \kappa\xi_i^\mu\right)\xi_j^\mu, \tag{4}$$

for each ξ_i^μ, where β is the over-relaxation factor and must be chosen between 0 and 1.

Note that for each randomly given memory set, the learning iterations (equation (4)) continue until all the p patterns in the set are memorized. Usually each learning iteration is called an epoch. During an epoch, each memory pattern is presented once. Then the learning speed is measured in the number of epochs. The synaptic matrices are adjusted only if a pattern has not already been memorized. Thus, during an epoch the maximum number of adjustments for a connection $J_{i,j}$ is equal to p. In our parallel dynamics here, all the adjustments for a single pattern can be done in one cycle. So each epoch requires p parallel updates of $J_{i,j}$.

Adopting the preceding adaptation rule, one is then able to design an asymmetric neural network with two-state neurons. Before going on to the next section, it is also worth noting that this kind of designed neural network has two main designing parameters: the storage capacity α and another designing parameter κ, which are crucial to their performances. The former determines how many patterns can be actually stored as memories; the latter is closely related to the symmetry degree of $J_{i,j}$ and also the attraction basin of memories. It has been found that by tuning κ to a proper range of values, the spurious memories that commonly exist in symmetric neural networks can be suppressed completely; in addition, the sizes of the attraction basins of memories can be controllable by endowing different values of κ to different memory patterns [13].

3. Eigenvalue spectra

We probe the eigenvalue spectrum of the designed asymmetric neural networks by studying the eigenvalue distribution of their synaptic couplings $J_{i,j}$. We are particularly interested in the dependence of these distributions on the two main network parameters: the storage capacity α and another designing parameter κ. For this purpose, we first fix the value of α and study how the eigenvalue spectrum depends on κ. To calculate the eigenvalue, a neural network with $N = 1000$ neurons (fixed throughout the paper) is prepared and p ($p = \alpha N$) system states are randomly selected as memories. Then the elements of the synaptic coupling $J_{i,j}$ are randomly initialized with values $+\frac{1}{\sqrt{N}}$ or $-\frac{1}{\sqrt{N}}$. For convenience of analysis and following [10, 11, 21], we restrict our discussion to the case in which the elements of $J_{i,j}$ should satisfy

$$\langle J_{i,j}\rangle = 0; \left\langle J_{i,j}^2\right\rangle = \frac{1}{N}, \tag{5}$$

during our designing process, where $\langle\cdot\rangle$ represents an average of the distribution. Thus, in the next step we iteratively apply equation (4) while keeping equation (5) until $|\bar{h}_{\xi_i^\mu} - \kappa| < 10^{-6}$ has been achieved for all the p memories. To achieve this goal one can actually perform a renormalization $J_{i,j} - \langle J_{i,j}\rangle$ and $\frac{J_{i,j}}{\sqrt{N\langle J_{i,j}^2\rangle}}$ to $J_{i,j}$ after one epoch in adopting the SOR rule, which will force the elements of $J_{i,j}$ to meet equation (5) during the whole designing process. Note that when applying equation (4) the choice of the over-relaxation factor β has been verified not to affect the final designing result [12], thus we fix $\beta = 0.6$ throughout our study. By continuously repeating the above procedure, we are able to derive a network satisfying both $\bar{h}_{\xi_i^\mu} = \kappa$ and $\langle J_{i,j}\rangle = 0$, $\langle J_{i,j}^2\rangle = \frac{1}{N}$, and finally calculate the eigenvalue spectrum of $J_{i,j}$.

Some typical results with $\alpha = 0.01$ for several κ are shown in figure 1. It can be seen that when κ is small, all eigenvalues uniformly lie within a circle of the complex plane (see figure 1(*a*)), which behaves much like those of asymmetric random matrices [15]; while as κ becomes larger, the eigenvalue spectrum is split into two parts—the left part still lies inside the circle but with a smaller radius, while the right part emerges outside the circle and lies in the real axis with values equal to κ. We also note that the number of eigenvalues in the right part is equal to p, the number of memory patterns, suggesting that the right part carries memory information. As an $N \times N$ matrix has N eigenvalues in total, then the number of eigenvalues in the left part is equal to $N - p$. In view of this fact, we call the left part the *noise part*, and the right part the *information part*. Those phenomena are in agreement with that presented in the asymmetric neural networks designed by the generalized perceptron rule [11, 21], where it has been found that the phenomenon of the eigenvalue spectrum can be correlated to the two essentially different dynamical phases [21]. Thus, regardless of the designing (learning) rules, our results here suggest that those phenomena of the spectrum may be a general feature for two-state asymmetric neural networks.

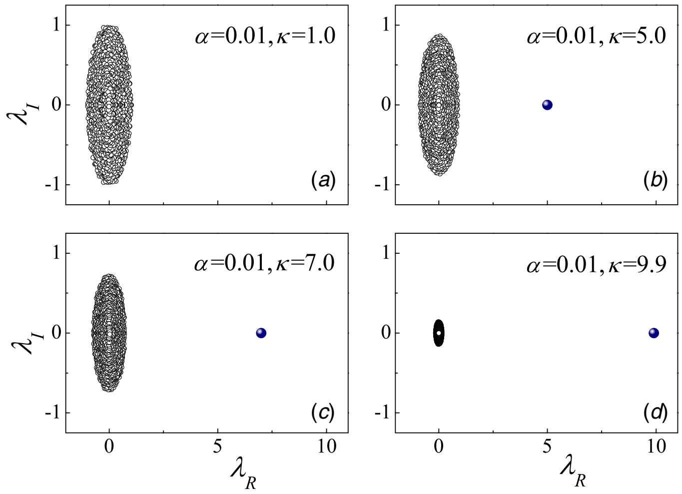

Figure 1. Eigenvalue spectrum of the asymmetric neural networks with $\alpha = 0.01$ for $\kappa = 1.0$ (a); $\kappa = 5.0$ (b); $\kappa = 7.0$ (c); $\kappa = 9.9$ (d), where λ is the eigenvalue of the synaptic matrix, and λ_I and λ_R represent the imaginary and real part of λ, respectively.

Another interesting piece of information from figure 1 which was not covered by [21] is that as κ continues to grow, the radius of the circle in the noise part becomes smaller. Then it is expected that when κ is large enough, the circle will become a dot in the real axis (see figure 1(d)). At this moment all eigenvalues become real indicating that the neural network becomes a symmetric one. Therefore, through studying the eigenvalue spectrum, we show here that the neural network will finally tend to symmetry by keeping on increasing κ. With the increase of κ we have indeed seen such a phenomenon for other several fixed α values as well. As shown in the example in figure 1 ($\alpha = 0.01$), it is found that when κ approaches the value of 10, the learning time (the time taken to successfully design the neural network by applying equations (4) and (5)) becomes especially long, indicating that there exists a critical value of κ above which one cannot succeed in designing the network.

In order to further check those phenomena, we next study the dependence of the eigenvalue spectrum on α with a fixed κ. The results are presented in figure 2. A picture similar to that of figure 1 can be clearly seen, i.e., for $\kappa = 5.0$, as α increases, all the spectra split into the noise part and the information part (see figures 2(a)–(c)), while when α is large enough, all eigenvalues tend to lie in the real axis (see figure 2(d)) suggesting that there is also a limiting behavior when we fix κ and increase α instead. Obviously a critical point of α exists for a given κ as well. For $\kappa = 5.0$ focused in figure 2, the critical value α_{cr} may be around $\alpha = 0.4$.

Combining the results of figures 1 and 2, one may find that whether by increasing κ or by increasing α alone, the spectrum shows a transition and the synaptic coupling matrices follow a similar behavior from being asymmetric to symmetric. Thus, deriving a general relation between κ and α based on such phenomena seems possible. In the following section we shall demonstrate that this is indeed the case.

4. Analysis

According to the preceding properties of the eigenvalue spectrum, we assume that the designed synaptic coupling $J_{i,j}$ can be written as a symmetric part $J^s_{i,j}$ with a weighting factor κ plus an asymmetric part $J^a_{i,j}$, i.e.,

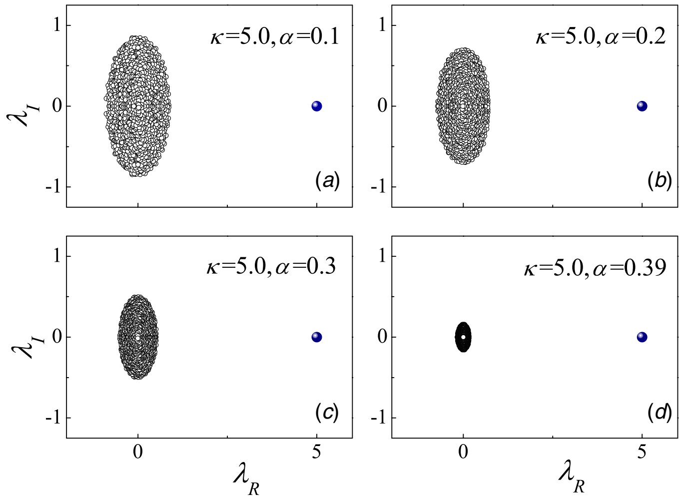

Figure 2. Eigenvalue spectrum of the asymmetric neural networks with $\kappa = 5.0$ for $\alpha = 0.1$ (*a*); $\alpha = 0.2$ (*b*); $\alpha = 0.3$ (*c*); $\alpha = 0.39$ (*d*), where λ is the eigenvalue of the synaptic matrix, and λ_I and λ_R represent the imaginary and real part of λ, respectively.

$$J_{i,j} = \kappa J^s_{i,j} + J^a_{i,j}. \tag{6}$$

For $J^s_{i,j}$, if $p \leqslant N$, it can be obtained by the pseudoinverse (PI) rule [24, 25], reading as $J^s_{i,j} = \frac{1}{N}\sum_{\mu,\nu=1}^{p} \xi_i^\mu \left(\mathbf{C}^{-1}\right)_{\mu\nu} \xi_j^\nu$, where $\mathbf{C}^{-1}$ is the inverse of the overlap matrix $\mathbf{C}$ defined by $C_{\mu\nu} = \frac{1}{N}\sum_{i=1}^{N} \xi_i^\mu \xi_i^\nu$. Obviously the $J^s_{i,j}$ designed by the PI rule is symmetric with $\langle J^s_{i,j}\rangle = 0$, $\langle (J^s_{i,j})^2\rangle = \frac{p}{N^2} = \frac{\alpha}{N}$. As $J_{i,j}$ is asymmetric and satisfies $\langle J_{i,j}\rangle = 0$, $\langle J^2_{i,j}\rangle = \frac{1}{N}$ according to equation (5), then we can easily find that $\langle J^a_{i,j}\rangle = 0$ and $\langle (J^a_{i,j})^2\rangle = \frac{1-\alpha\kappa^2}{N}$ provided that $\langle J^s_{i,j} J^a_{i,j}\rangle = 0$.

Such an assumption is based on the following facts. First, it is well known that the symmetric part $J^s_{i,j}$ designed by the PI rule only has two degenerate eigenvalues 0 and 1 [26]. Second, from [15, 19] one can learn that for an asymmetric matrix such as $J^a_{i,j}$ satisfying $\langle (J^a_{i,j})^2\rangle = \frac{1-\alpha\kappa^2}{N}$, all its eigenvalues uniformly lie within a circle of the complex plane with radius $R = \sqrt{1-\alpha\kappa^2}$. We have drawn the histograms of the absolute values of eigenvalues inside the circle in the complex plane for several κ and α. As examples, the results of $\alpha = 0.1$ for several κ are plotted in figure 3, from which one can see that these histograms verify that the distribution of the complex eigenvalues is indeed uniformly inside the circle. Then it can be expected that in the large-N limit, the synaptic coupling written as $J_{i,j} = \kappa J^s_{i,j} + J^a_{i,j}$ may have distributions of the eigenvalue spectrum like those in figure 1 or figure 2 for certain values of α and κ, which is in agreement with the analysis in [21].

Now we turn to explaining why a transition behavior of the eigenvalue spectrum from being asymmetric to symmetric can take place. First, according to equation (6) we can easily get $1-\alpha\kappa^2 \to 0$ if the asymmetric part of $J_{i,j}$ vanishes and only the deterministic symmetric part is left. Thus, α and κ should satisfy $\alpha\kappa^2 \leqslant 1$ regarding this limit, which provides a constrained relation of α and κ in a mathematical sense. However, the relation under this limiting case cannot characterize the intermediate transition process of the spectrum ($1-\alpha\kappa^2 \neq 0$). To derive a general relation that can describe the whole transition process of the eigenvalue spectrum, physically, it is natural to consider the symmetry degree of $J_{i,j}$ defined by

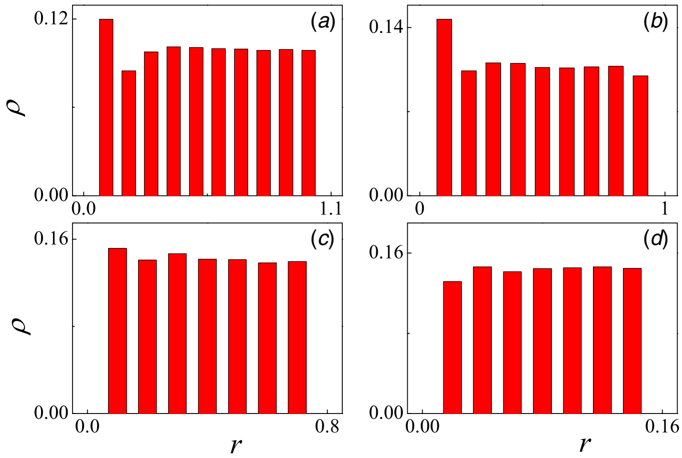

Figure 3. Histograms of the absolute values of eigenvalues inside the circle in the complex plane with $\alpha = 0.01$ for $\kappa = 1.0$ (a); $\kappa = 5.0$ (b); $\kappa = 7.0$ (c); $\kappa = 9.9$ (d), where r represents the radial distance from the origin to the focused position in the circle ($r < R$, the radius of the circle), and ρ is the density of the eigenvalue spectrum in the range $[r, r + \triangle r]$.

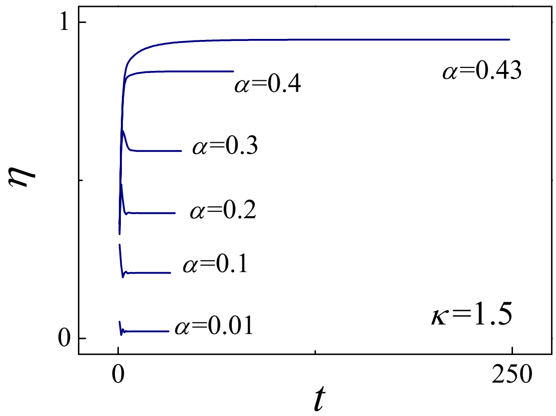

Figure 4. The symmetry degree η versus the time step t of the neural network designing (learning) with $\kappa = 1.5$ for several α.

$$\eta = \frac{\langle J_{i,j} J_{j,i} \rangle}{\langle J_{i,j}^2 \rangle}. \tag{7}$$

Under this definition, $\eta = 1$ indicates that we are considering a totally symmetric neural network, otherwise $\eta \neq 1$ corresponds to a general asymmetric network. For the asymmetric case, $\eta = 0$ indicates that there is no correlation between $J_{i,j}$ and $J_{j,i}$ on the average; $\eta = -1$ indicates that the synaptic matrix is fully antisymmetric with $J_{i,j} = -J_{j,i}$. The reason for considering η here is obvious: the aim is to characterize the preceding transition in the eigenvalue spectrum. While we also note that when an asymmetric neural network with given κ and α has been successfully designed, η is actually uniquely definite. In this sense the learning of asymmetric neural networks is essential for designing the synaptic coupling $J_{i,j}$ with a given symmetry degree, and thereby, η is expected to be as a function of the designing parameters, κ and α. Our results in figure 4 clearly confirm this conjecture, where with fixed

$\kappa = 1.5$ and for several α, η converges to a definite value lying between 0 and 1 after several time steps of designing. Therefore to derive the relation between κ and α in designing the neural networks, we begin by analyzing η in the following.

To derive η easily, we first perform a decomposition for the asymmetric part $J^a_{i,j}$. A general asymmetric matrix such as $J^a_{i,j}$ can usually be written as

$$J^a_{i,j} = T^s_{i,j} + \gamma T^{as}_{i,j}, \tag{8}$$

where $T^s_{i,j}$ ($T^{as}_{i,j}$) is symmetric (antisymmetric) with $T^s_{i,j} = T^s_{j,i}$, $T^{as}_{i,j} = -T^{as}_{j,i}$, and $\langle (T^s_{i,j})^2\rangle = \langle (T^{as}_{i,j})^2\rangle = \frac{1-\alpha\kappa^2}{N}\frac{1}{1+\gamma^2}$; γ denotes the ratio of $T^s_{i,j}$ to $T^{as}_{i,j}$ [11]. Then $\gamma = 0$ indicates that $J^a_{i,j}$ is symmetric; $\gamma = 1$ means that $J^a_{i,j}$ is a fully random matrix with equal weight for $T^s_{i,j}$ and $T^{as}_{i,j}$; $\gamma \gg 1$ suggests that $J^a_{i,j}$ is a totally antisymmetric matrix. We next substitute equations (6) and (8) into equation (7); η can be represented by

$$\begin{aligned}\eta &= \frac{\big(\big\langle \kappa J^s_{i,j} + T^s_{i,j} + \gamma T^{as}_{i,j}\big)\big(\kappa J^s_{i,j} + T^s_{i,j} - \gamma T^{as}_{i,j}\big)\big\rangle}{\frac{1}{N}} \\ &= \frac{\kappa^2\big\langle\big(J^s_{i,j}\big)^2\big\rangle + \big\langle\big(T^s_{i,j}\big)^2\big\rangle - \gamma^2\big\langle\big(T^{as}_{i,j}\big)^2\big\rangle}{\frac{1}{N}} \\ &= \alpha\kappa^2 + (1-\alpha\kappa^2)\left(\frac{1-\gamma^2}{1+\gamma^2}\right).\end{aligned} \tag{9}$$

Here it is assumed that $\langle J^s_{i,j}T^s_{i,j}\rangle = 0$, $\langle J^s_{i,j}T^{as}_{i,j}\rangle = 0$, and $\langle T^s_{i,j}T^{as}_{i,j}\rangle = 0$.

Now let us consider the value of η by taking into account the value of γ. First, if $J^a_{i,j}$ is antisymmetric, we have $\gamma \gg 1$. Thus, $\eta = 2\alpha\kappa^2 - 1$ according to equation (9). In this case the maximum storage capacity can be determined by $\alpha_{\max} = \frac{1+\eta}{2\kappa^2}$, which recovers Gardner's result for highly antisymmetric neural networks [11]. Next we consider the asymmetric neural network which is the focus of this paper. Since the eigenvalue spectrum of the noise part uniformly lies in a circle rather than an ellipse in the complex plane (see figures 1– 3), there is no correlation between $J^a_{i,j}$ and $J^a_{j,i}$ on the average, i.e., $\langle J^a_{i,j}J^a_{j,i}\rangle = 0$, suggesting that $J^a_{i,j}$ is a totally random matrix. Thus, we get $\gamma = 1$. Substituting $\gamma = 1$ into equation (9) we obtain

$$\eta = \alpha\kappa^2, \tag{10}$$

which relates η to α and κ. To verify such a relation, in figures 5(*a*) and (*b*) we plot η as a function of κ (α) for several given α (κ). η is actually calculated according to equation (7). To calculate $\langle J_{i,j}J_{j,i}\rangle$ and $\langle J^2_{i,j}\rangle$ for each α (κ), we take 100 instances of neural network designing where different random initial memory patterns are considered, for the average. From the figure one can see that the theoretical formula is in good agreement with the simulation results, strongly suggesting that $\eta = \alpha\kappa^2$ is a suitable formula for describing the relation. The reason for the slight deviation for $\kappa = 1.5$ in figure 5(*b*) may be that when α is a relatively large value between 0.2 and 0.4, the network size N is not large enough compared with the number p of memory patterns. Some correlations between $J^s_{i,j}$ and $J^a_{i,j}$ may then still exist, which results in the estimate errors by using the theoretical formula.

Interestingly, the formula $\eta = \alpha\kappa^2$ also shows a limiting case: as η should satisfy $\eta \leqslant 1$, then we get $\alpha\kappa^2 \leqslant 1$, which is consistent with the above mathematical analysis under the limiting case that $J^a_{i,j}$ vanishes and only the part of $J^s_{i,j}$ is left in equation (6). Therefore the fact that a critical point of κ_{cr} (α_{cr}) is shown in figure 1 (figure 2) is understandable because for a given κ or α, $\alpha\kappa^2 \leqslant 1$ provides an upper boundary of α or κ. In addition to this limiting case, it is worth mentioning that the proposed formula here essentially provides a trade-off relation

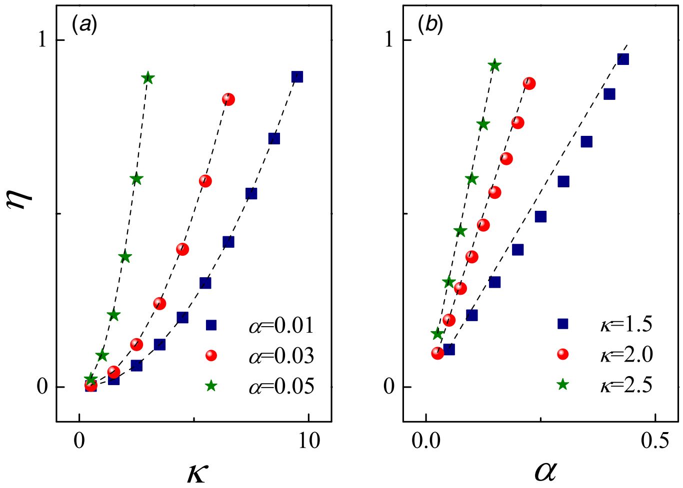

Figure 5. (*a*) The symmetry degree η versus κ for several α, where the squares, circles, and stars represent the case of $\alpha = 0.01$, 0.03, 0.05, respectively. (*b*) η versus α for several κ, where the squares, circles, and stars represent the case of $\kappa = 1.5$, 2.0, 2.5, respectively. The dashed lines are the theoretical results from the formula $\eta = \alpha\kappa^2$ proposed in this paper.

in designing asymmetric neural networks, i.e., given one of the designing parameters, κ or α, via measuring the symmetric degree η, another designing parameter is uniquely determined. This may enable us to optimize the designing of the neural networks.

Finally, we would like to point out that the proposed formula between η, α and κ is only applicable to the two-state asymmetric neural networks under the designing condition $\bar{h}_{\xi_i^\mu} = \kappa$ and with the restrictions $\langle J_{i,j}\rangle = 0$, $\langle J_{i,j}^2\rangle = \frac{1}{N}$ discussed here. For the networks under the condition $\bar{h}_{\xi_i^\mu} \geqslant \kappa$ or without the restrictions $\langle J_{i,j}\rangle = 0$, $\langle J_{i,j}^2\rangle = \frac{1}{N}$, although a constrained relation between α and κ may still exist, the relation between η, α and κ would doubtless be more complicated, leading us unable to provide an analysis here. The results of the latter case will be presented in our further studies [27].

5. Summary

To summarize, through studying the dependence of the eigenvalue spectrum of the designed two-state asymmetric neural networks on their two main network parameters, the storage capacity α and another designing parameter κ, we have found that by increasing one of the parameters and keeping another fixed, the spectra can show an interesting limiting behavior toward the spectra of a totally symmetric matrix. This phenomenon can be properly described by the increase of the symmetry degree η of their synaptic couplings. We have proposed a general formula $\eta = \alpha\kappa^2$ to characterize the phenomenon, which is in good agreement with the test of the simulations of our neural network designing. As η should be smaller than 1, the formula provides an explicit constrained relation between the two important designing parameters of asymmetric neural networks, suggesting that the performances determined by α and κ cannot be simultaneously improved. More importantly, the formula relates the symmetry degree η to the designing parameters α and κ in the neural network designing, which may provide useful information for optimizing the designing and further application of this kind of asymmetric neural network.

Acknowledgments

This work was supported by the NNSF (grant nos 11205032, 11147191 and 10925525) of China, the NSF (grants no 2013J05008) of Fujian province and the start-up fund (no 022390) from Fuzhou University.

References

[1] Hopfield J J and Tank D W 1985 'Neural' computation of decisions in optimization problems *Biol. Cybern.* **52** 141–52
[2] Tank D W and Hopfield J J 1986 Simple 'neural' optimization networks: an A/D converter, signal decision circuit, and a linear programming circuit *IEEE Trans. Circuits Syst.* **33** 533–41
[3] Jin T and Zhao H 2005 Pattern recognition using asymmetric attractor neural networks *Phys. Rev.* E **72** 066111
[4] Kryzhanovsky B V and Kryzhanovsky V M 2008 A binary pattern classification using potts model *Opt. Mem. Neural Netw.* **17** 308–16
[5] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities *Proc. Natl Acad. Sci. USA* **79** 2554–8
[6] Jin T and Zhao H 2008 Controlling the dynamics of multi-state neural networks *J. Stat. Mech.* P06002
[7] Kanter I 1988 Potts-glass models of neural networks *Phys. Rev.* A **37** 2739–42
[8] Xiong D X and Zhao H 2010 q-state Potts-glass neural network based on pseudoinverse rule *Phys. Rev.* E **82** 026114
[9] Xiong D X and Zhao H 2010 Estimates of storage capacity in the q-state Potts-glass neural network *J. Phys. A: Math. Theor.* **43** 445001
[10] Gardner E 1988 The space of interaction in neural network models *J. Phys. A: Math. Gen.* **21** 257–70
[11] Gardner E, Gutfreund H and Yekutieli I 1989 The phase space of interactions in neural networks with definite symmetry *J. Phys. A: Math. Gen.* **22** 1995–2008
[12] Heekuck Oh and Kothari S C 1991 A new learning approach to enhance the storage capacity of the Hopfield model *IJCNN'91: Proc. IEEE Int. Joint Conf. on Neural Networks* pp 2056–62
[13] Zhao H 2004 Designing asymmetric neural networks with associative memory *Phys. Rev.* E **70** 066137
[14] Metha M L 2006 *Random Matrices* (Amsterdam: Elsevier)
[15] Girko V L 1984 Circle law *Theory Probab. Appl.* **29** 694–706
[16] Sommers H J, Crisanti A, Sompolinsky H and Stein Y 1988 Spectrum of large random asymmetric matrices *Phys. Rev. Lett.* **60** 1895–8
[17] Shepherd G M 2004 *The Synaptic Organization of the Brain* (London: Oxford University Press)
[18] Holmgren C, Harkany T, Svennenfors B and Zilberter Y 2003 Pyramidal cell communication within local networks in layer 2/3 of rat neocortex *J. Physiol.* **551** 139–53
[19] Rajan K and Abbott L F 2006 Eigenvalue spectra of random matrices for neural networks *Phys. Rev. Lett.* **97** 188104
[20] Wei Y 2012 Eigenvalue spectra of asymmetric random matrices for multicomponent neural networks *Phys. Rev.* E **85** 066116
[21] Zhou Q, Jin T and Zhao H 2009 Correlation between eigenvalue spectra and dynamics of neural networks *Neural Comput.* **21** 2931–41
[22] Little W A 1974 The existence of persistent states in the brain *Math. Biosci.* **19** 101–20
[23] Young D M 1971 *Iterative Solution of Large Linear Systems* (New York: Academic)
[24] Personnaz L, Guyon I and Dreyfus G 1986 Collective computational properties of neural networks: new learning mechanisms *Phys. Rev.* A **34** 4217–28
[25] Kanter I and Sompolinsky H 1987 Associative recall of memory without errors *Phys. Rev.* A **35** 380–92
[26] Waugh F R, Marcus C M and Westervelt R M 1990 Fixed-point attractors in analog neural computation *Phys. Rev. Lett.* **64** 1986–9
[27] Xiong D X and Zhao H 2013 unpublished