
**Brain Processes, Theories, and Models**

edited by Roberto Moreno-Díaz and José Mira-Mira

The MIT Press
Cambridge, Massachusetts
London, England
INFORMATION REPRESENTATION IN A RANDOMLY AND
SYMMETRICALLY CONNECTED RECURRENT NEURAL NETWORK

Akira Date*  Koji Kurata**  Shun-ichi Amari***

* Department of Computer Science, Graduate School of Technology
  Tokyo University of Agriculture and Technology, 2-24-16 Nakamachi, Koganei, Tokyo 184, Japan
** Department of Biophysical Engineering, Faculty of Engineering Science
  Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560, Japan
*** Department of Mathematical Engineering and Information Physics, Faculty of Engineering
  University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113, Japan
**** RIKEN, Frontier Research System on Brain Information Processing
  Laboratory for Information Representation, Wako, Saitama 351-01, Japan

Abstract. A class of recurrent neural networks is considered in which \( w_{ij} \), the connection weight from the \( j \)th to the \( i \)th element, is randomly generated under the symmetricity condition \( w_{ij} = w_{ji} \). The expected number of equilibrium states in the network consisting of two-state threshold elements having outputs of \((\pm 1, 1)\) or \((0, 1)\) with a variable (but uniform throughout the network) threshold is derived by a method of statistical neurodynamics. It is shown that the expected number of equilibrium states is uniquely determined by the threshold value and that the equilibrium states are concentrated on the states having a specific activity level, i.e. the rate of excited neurons. For the \((0,1)\) model, a network which has a set of equilibrium states concentrated on the activity level 32% has the maximum number of equilibrium states of all networks. Applications of this network as a module for memory systems are discussed.

1 Introduction

Specific computation in the brain is not substantially affected by damage to a specific neural component. Since the brain is a large-scale system composed of neurons, in order to analyze the macroscopic properties of neural networks, random neural networks include associative memory in which the topology of the network is generated by a controlled probability model have been introduced by many researchers (see references in [2,5,7,11,12,14]).

So far in neural modelling of associative memory, it seems that the problem of how a representation to be memorized in the network should be created have not been discussed except primary feature extractions based on a set of input signals itself [9]. Usually, in order to study the capacity and the dynamics of associative memory, randomly generated binary strings have been used for a set of patterns to be memorized in the network [3,7,14]. In particular, a set of input patterns to the auto-associative network have been designed to be identical to a set of equilibrium states or attractors in the network.

Neurophysiological studies showed that neuronal short-term representations of long-term memory for a set of pictorial stimuli in the monkey inferior temporal cortex did not depend solely on
geometric similarity among visual stimuli but depend on temporal sequences of a set of stimuli exposed to the monkeys in the training session to some extent [17, 18]. The results suggest that the strategies used by single neurons to categorize objects utilize not only visual shapes themselves but also temporal relationship among objects presented to the monkeys [6, 13, 15, 17].

Morita (1988) proposed an associative memory system in which input patterns were associated with the attractors of a random symmetric network, a class of recurrent neural networks in which \( w_{ij} \), the connection weight from the \( j \)th to the \( i \)th element, is randomly generated under the symmetry condition \( w_{ij} = w_{ji} \). The Hamming distance between a pair of input patterns to the network does not affect the distance between the corresponding pair of attractors in the network [19]. Controlling the threshold value \( \theta \), that is equivalent to set the external signal \(-\theta\) to all elements, plays a central role in encoding input patterns to stable states in the random symmetric network. Under the condition, however, how the properties of the network vary, such as the number of equilibrium states, the distribution of those in the state space, have not been known yet, except that the network has a large number of equilibrium states [4, 22] and a set of the equilibrium states has a ultrametric structure [20].

In this article, with the method of statistical neurodynamics, we analyze a set of random symmetric neural networks consisting of two-state threshold elements having outputs of \((-1, 1)\) or \((0, 1)\) with variable thresholds, but uniform throughout the network. The main result of this paper is, in common for the \((-1, 1)\) and the \((0, 1)\) model, that 1) the expected number of equilibrium states in the network is in the exponential order of number of neurons (The result has already known for \( \theta = 0 \) by the method of statistical mechanics [22], but we obtained the result without spin glass analogy [7, 10]), 2) the equilibrium states are concentrated on the state having a specific activity for a fixed threshold parameter. For the \((-1,1)\) model, the network with a threshold value adjusted to make equilibrium states in which half components are firing, is best in a sense of maximizing the number of equilibrium states. For the \((0,1)\) model, on the contrary to the \((-1,1)\) model, a network consisting of the neurons with a threshold value adjusted to make equilibrium states in which 32% of neurons are firing has the maximum number of equilibrium states, although in the set of \( 2^n \) states the combinatorial number of the states is maximum at the activity level \( p = 0.5 \). Furthermore, for the \((-1,1)\) model, continuous change of the threshold parameter leads a continuous change of the activity \( p \) of equilibrium states at \( 0 < p < 1 \). On the contrary, the property is lost for the \((0,1)\) model, i.e., a \( p = 0 \) state is an only equilibrium one when the threshold is set larger than a specific value.

How neuronal representation for a novel visual stimulus is maintained during a short-term memory task [17] and how that for familiar stimulus is acquired by learning [15, 17, 18] are basic questions that remain to be examined. Primate memory system must at least consists of subsystems which keep any signals temporarily as steady states in order to transform the neuronal excitation patterns into patterns of synaptic connection. Since a random symmetric network has a large number of equilibrium states, we think that the network is potentially capable of memorizing a large number of items [5,6]. Finally we remark an applications of the random symmetric network as a module for memory systems.

2 Information representation in a random symmetric network

2.1 Model of a neuron and a neural network

Here we consider a fully interconnected network consisting of \( n \) neural-like elements in which the \( i \)th output \( x_i, i = 1, \ldots, n \) takes on binary states \((x_i = -1, 1 \text{ or } 0, 1)\) and \( w_{ij}, 1 \leq i, j \leq n \), the
connection weight from the \( j \)th to the \( i \)th neuron, is independent and identically distributed (i.i.d.) under the symmetricity condition \( w_{ij} = w_{ji} \). Note that \( \{w_{ij}\} \), \( i \geq j \), are chosen independently from a normal distribution with mean \( \mu_w = 0 \) and variance \( \sigma_w^2 = 1 \). We use \( N(0, 1) \) to denote this distribution. As we will mention later, the restriction is not essential to the analysis to follow. The \((-1, 1)\) model is identical to the SK spin glass model [21].

The state of each neuron is updated based on the sign of a linear form computed by the intersection weights and the current state of the system. When each neuron works synchronously at discrete times \( t = 0, 1, 2, \ldots \), and takes binary value of -1 or 1, the dynamics of the network is described as

\[
x_i^{t+1} = \operatorname{sgn} \left( \sum_{j=1}^{n} w_{ij} x_j^t - \theta \right),
\]

where \( x_i^t \) is the state or the output of the \( i \)th neuron at time \( t \), and \( \theta \) is a variable threshold value which is uniform throughout the network, and \( \operatorname{sgn} \) is the sign function, \( \operatorname{sgn}(u) = 1 \) when \( u > 0 \), and is otherwise \(-1\). By using the nonlinear operator \( T_w \), equation (2.1) is represented as

\[
T_w x = \operatorname{sgn}(W x - \theta),
\]

where \( x \) is an \( n \)-dimensional vector of the current state of the network (neuronal firing pattern), \( x = (x_1, \ldots, x_n)^T \), \( (T \) denotes transpose) and \( W = \{w_{ij}\} \) is the \( n \times n \) matrix whose \((i, j)\) component is \( w_{ij} \), and \( \operatorname{sgn} \) is operated componentwise.

When each neuron takes a binary value of 0 or 1, the dynamics of the network is described as

\[
T_w x = I(W x - \theta)
\]

where \( I \) is the unit step function, \( I(u) = 1 \) when \( u > 0 \), and is otherwise 0.

The network has additional input lines from the outside, and these inputs are used to assign an initial state \( x_0 \) to the network. The dynamics of state transition then follows, subject to equation (2.2) or (2.3).

The network as a whole evolves in the state space \( \{-1, 1\}^n \) or \( \{0, 1\}^n \). Therefore, the network has \( 2^n \) states, in which the state \( x \) is said to be an equilibrium state or a fixed point when \( x = T_w x \) holds. Once the network falls into an equilibrium state, it remains in that state until a cancellation signal comes from the outside. The basin of an equilibrium state \( x \) is a set of the states each of which falls in the state \( x \) after a finite number of state transitions. An equilibrium state which has large basin of the attraction is said to be a stable state.

2.2 Number of equilibrium states in the \((-1,1)\) representation network

First, we calculate the probability \( P \) that a state \( x \) in which \( m \) components are firing or activity \( p = m/n, 0 \leq p \leq 1 \) state, is an equilibrium state. It is easy to show that \( P \) does not depend on specific \( x \) but the same for any \( x \) having activity \( p \). Therefore we calculate the probability \( P \) that a state \( x \) whose first \( m \) components are firing

\[
x = \left( 1, \ldots, 1, -1, \ldots, -1 \right)
\]

is an equilibrium state. For the above \( x \), we put

\[
u_i = \sum_{j=1}^{n} w_{ij} x_j = \sum_{j=1}^{m} w_{ij} - \sum_{j=m+1}^{n} w_{ij}
\]
where $u_i$ is the $i$th component of $Wx$. All the $u_i$ are normally distributed with mean 0 and variance $n \sigma^2_x = n$. Notice that $u_i$ and $u_j$ are not independent because of $w_{ij} = w_{ji}$, and their covariance is given by

$$\text{Cov} (u_i, u_j) = \begin{cases} 
1, & i, j \leq m \text{ or } i, j > m \\
-1, & \text{otherwise.} 
\end{cases}$$

These correlated $u_i, i, \ldots, n$ can be represented by using $n + 1$ mutually independent normal random variables $s_i, i = 1, \ldots, n$ and $r$ subject to $N(0, 1)$ as

$$u_i = \begin{cases} 
\frac{\sqrt{n - 1} s_i + r}{\sqrt{n - 1}}, & i \leq m \\
\frac{\sqrt{n - 1} s_i - r}{\sqrt{n - 1}}, & i > m.
\end{cases}$$

The probability $P$ that the state $x$ is an equilibrium one then becomes

$$P = \text{Prob} \left\{ u_1 > \theta, \ldots, u_m > \theta, u_{m+1} \leq \theta, \ldots, u_n \leq \theta \right\}$$

In order to calculate (2.8), we first fix $r$, and calculate the conditional probability $P_r$.

$$P_r = \text{Prob} \left\{ s_i > \frac{\theta - r}{\sqrt{n - 1}}, i = 1, \ldots, m \right\} \times \text{Prob} \left\{ s_i \leq \frac{\theta + r}{\sqrt{n - 1}}, i = m + 1, \ldots, n \right\}$$

$$= \left[ \text{Prob} \left\{ s_i > \frac{\theta - r}{\sqrt{n - 1}} \right\} \right]^m \times \left[ \text{Prob} \left\{ s_i \leq \frac{\theta + r}{\sqrt{n - 1}} \right\} \right]^{n-m}$$

We define the macroscopic parameter $\Theta = \theta / \sqrt{n - 1}$. When $r$ is fixed, the events $s_i$ are mutually independent in the sense of the conditional probability. Then, we have

$$P_r = \left( 1 - \Phi \left( \Theta - \frac{r}{\sqrt{n - 1}} \right) \right)^n \times \left( \Phi \left( \Theta + \frac{r}{\sqrt{n - 1}} \right) \right)^{n(1-p)}$$

where $\Phi$ is cumulative distribution function of the standard normal distribution

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left\{ -\frac{t^2}{2} \right\} \, dt.$$ 

We then take the expectation with respect to $r$, and we obtain

$$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left\{ -\frac{r^2}{2} + n p \left[ 1 - \Phi \left( \Theta - \frac{r}{\sqrt{n - 1}} \right) \right] + n(1-p) \left[ \Phi \left( \Theta + \frac{r}{\sqrt{n - 1}} \right) \right] \right\} \, dr.$$ 

Let $\frac{r}{\sqrt{n - 1}} = y$, then the equation can be written as

$$P = \sqrt{\frac{n - 1}{2\pi}} \int_{-\infty}^{\infty} \exp \left\{ -n \left[ 1 - \left( \frac{1}{n} \right) \frac{y^2}{2} - p \log \left[ 1 - \Phi(\Theta - y) \right] - (1-p) \log \left[ \Phi(\Theta + y) \right] \right] \right\} \, dy.$$ 

When $n$ is sufficiently large, by using the saddle-point approximation method, we obtain

$$P = \exp \left[ -n G(y_0) \right]$$

where $G(y_0)$ is the saddle-point argument.
where $y_0$ is the value of $y$ minimizing

\[(2.15) \quad G(y) = \frac{y^2}{2} - p \log\left[1 - \Phi(\Theta - y)\right] - (1 - p) \log\left[\Phi(\Theta + y)\right].\]

Since there are $nC_{np}$ (binomial coefficient) states whose activity is $p$, the expected number of equilibrium states $N(p; \Theta)$ is equal to $P \times nC_{np}$. By using Stirling's formula

\[(2.16) \quad nC_{np} \approx \left[2\pi np(1 - p)\right]^{-\frac{1}{2}} \times (1 - p)^{-n(1 - p)} \times p^{-np},\]

for fixed $p$ and $\Theta$,

\[(2.17) \quad \frac{\log N(p; \Theta)}{n} = -G(p, y_0) - (1 - p) \log(1 - p) - p \log p\]

can be solved numerically.

**Fig.1.** Properties of the equilibrium states in random symmetric networks consisting of two-state threshold elements having outputs of $(-1, 1)$. a. The expected number of equilibrium states $N(p; \Theta)$ as a function of the activity level $p$, i.e., the rate of excited neurons, in the networks for threshold parameters $\Theta = -0.8, 0.0, 0.5, 1.2$. b. A network of $\Theta = 0.0$ has the maximum number of equilibrium states in which $p = 0.5$. c. Continuous change of $\Theta$ leads a continuous change of the activity level $p$ in equilibrium states.

We then calculated $N(p; \Theta)$, the expected number of equilibrium states having activity $p$, $0 < p < 1$, for fixed threshold parameters $\Theta = \theta/\sqrt{n - 1} = -0.8, 0.0, 0.5, 1.2$. In the numerical calculation $p$ was varied every 0.01 and we use 0 log 0 = 0. Fig.1(a) represents the relationship between the activity $p$ and $(\log N(p; \Theta))/n$. In each network for $\Theta = -0.8, 0.0, 0.5, 1.2$, the expected number of equilibrium states $N(p; \Theta)$ is maximal at the activity $p = 0.71, 0.50, 0.37, 0.17$ respectively, and is equal to $e^{0.127n}$, $e^{0.1892n}$, $e^{0.0688n}$, $e^{0.0667n}$ respectively. Notice that the ordinate of Fig.1(a) is logarithmically scaled. Therefore, the results in Fig.1(a) imply that most of the equilibrium states in the network for a fixed $\Theta$ are the states having activity $p_0$ which maximize $\log N(p; \Theta)$. In other words, the number of equilibrium states having activity $p'(\neq p_0)$ is negligibly smaller than those having activity $p_0$. In this sense, we can say that the equilibrium states in the network for a fixed threshold value are concentrated on the states having a specific activity.
Then, in a network for a fixed threshold value, the relationship between $\Theta$ and the expected number of equilibrium states, and the relationship between $\Theta$ and the activity in the equilibrium states are numerically calculated and shown in Fig.1 (b) and (c) with a solid line. In the numerical calculation, $p$, $\Theta$ were varied in $0 \leq p \leq 1$ and $-3 \leq \Theta \leq 3$ with every 0.01. As $\max N(p, \Theta) = N(0.5, 0.0) = e^{0.1990n}$, the network of $\Theta = 0.0$ is best in a sense of maximizing the number of equilibrium states. This value is identical to the result derived by the method of statistical physics [22].

The result of the numerical calculation depicted with the solid line in Fig.1(c) can be used to design a network having equilibrium states maximally at a specific activity $p'$ by adjusting the threshold parameter $\Theta$.

Under a assumption that $u_i$, $i = 1, \cdots, n$ are mutually independent, i.e. $r = 0$, to maximize the probability that the activity $p'$ states are equilibrium ones, from equation (2.10), one might choose the threshold parameter so that

$$p' = 1 - \Phi(\Theta)$$

holds. The dotted line in Fig.1(c) shows the relationship between $\Theta$ and $p'$. As the dotted line and the solid one in Fig.1(c) are not identical, the network with threshold parameter $\Theta$ chosen by equation (2.18) does not have the equilibrium states maximally at activity $p'$. That is because of the symmetric property of connection weights.

### 2.3 Number of equilibrium states in the (0,1) representation network

Here we consider a random symmetric network consisting of threshold elements that have states of 0 or 1. As we did for the (-1,1) model, first we calculate the probability $P$ that a activity $p = m/n$, $0 \leq p \leq 1$, state $z$ whose first $m$ components are firing is an equilibrium state. For the $x$, we put

$$u_i = \sum_{j=1}^{n} w_{ij} x_j = \sum_{j=1}^{m} w_{ij},$$

All the $u_i$ are normally distributed with mean 0 and variance $m \sigma^2 = m = np$. Notice that the variance depend on the activity level $p$, and that $u_i$ and $u_j$ are not independent because of $w_{ij} = w_{ji}$, and their covariance is given by

$$\text{Cov}(u_i, u_j) = \begin{cases} 1, & i, j \leq m \\ 0, & \text{otherwise}. \end{cases}$$

These correlated $u_i, i, \cdots, n$ can be represented by using $n + 1$ mutually independent normal random variables $s_i, i = 1, \cdots, n$ and $r$ subject to $N(0, 1)$ as

$$u_i = \begin{cases} \sqrt{m-1} s_i - r, & i \leq m \\ \sqrt{m} s_i, & i > m. \end{cases}$$

For the (0,1) model, $u_i, i = m+1, \cdots, n$ are mutually independent. We proceed along the lines explained in previous section and obtain the following results (The detail derivation is described elsewhere [8]). For fixed values $p$ and $\Theta$,

$$\frac{\log N(p; \Theta)}{n} = -F(p, y_0) + (1 - p) \log \left( \frac{\Theta}{\sqrt{p}} \right) - (1 - p) \log(1 - p) - p \log p.$$
Fig. 2. Properties of the equilibrium states in random symmetric networks consisting of two-state threshold elements having outputs of \((0,1)\). 

**a.** The expected number of equilibrium states \(N(p; \Theta)\) as a function of the activity level \(p\) in the networks for threshold parameters \(\Theta = -0.8, 0.0, 0.3, 0.5, 0.96\). A network of \(\Theta = 0.45\) has the maximum number of equilibrium states at \(p = 0.32\), although the combinatorial number of the states is maximum at \(p = 0.5\). 

**b.** Continuous change of \(\Theta\) leads to a continuous change of \(p\) in the equilibrium states at \(\Theta \leq 0.96\), but any states except a \(p = 0\) state cannot be an equilibrium state at \(\Theta > 0.96\).

It can be solved numerically. We should separately calculate the probability that a \(p = 0\) state is an equilibrium state. The \(p = 0\) state is an equilibrium state when \(\Theta \geq 0\).

We then calculated \(N(p; \Theta)\), the expected number of equilibrium states having activity \(p\), \(0 \leq p \leq 1\), for fixed threshold parameters \(\Theta = \Theta / \sqrt{n} = -0.80, 0.00, 0.30, 0.50, 0.96\). In the numerical calculation \(p\) was varied every 0.01 and we use \(\log \Theta = 0\). Fig. 2(a) represents the relationship between the activity \(p\) and \(\log N(p; \Theta)\). In each network with \(\Theta = -0.80, 0.00, 0.30, 0.50, 0.96\), the expected number of equilibrium states \(N(p; \Theta)\) is maximal at activity \(p = 0.82, 0.55, 0.40, 0.29, 0.25\) respectively, and is equal to \(e^{0.0327n}\), \(e^{0.1964n}\), \(e^{0.1844n}\), \(e^{0.1979n}\), \(e^{0.0011n}\) respectively.

For the networks exemplified in Fig. 2(a) except for a network of \(\Theta = 0.30\), the \(N(p; \Theta)\) have only one local maximum which is a global maximum, and the result has a same property as for the \((-1,1)\) case. For a network of \(\Theta = 0.30\), \(N(p; \Theta)\) has two local maximums at \(p = 0.03, 0.40\). Numerical calculation show \(N(p; \Theta)\) has two local maximums at \(0.0 < \Theta \leq 0.31\). The expected numbers of equilibrium states at the local maximums are \(N(0.03) = e^{0.0335n}\), \(N(0.40) = e^{0.1344n}\). Depending on the initial state, the network state might fall in the \(p = 0.03\) state. However, the number of equilibrium states having activity \(p = 0.03\) is negligibly smaller than that of the states having activity \(p = 0.40\). In this sense, we can say that the equilibrium states in the network for a fixed \(\Theta\) are concentrated on the state having a specific activity \(p_0\) which maximize \(N(p; \Theta)\).

Then, in the network for a fixed threshold value, the relationship between \(\Theta\) and the expected number of equilibrium states, and the relationship between \(\Theta\) and the activity in equilibrium states are numerically calculated and represented in Fig. 2(b) and (c) with a solid line. In the numerical calculation, \(p, \Theta\) were varied in \(0 \leq p \leq 1\) and \(-3 \leq \Theta \leq 3\) with every 0.01.

Although in a set of \(2^n\) network states the combinatorial number of states having activity \(p\) is maximal at \(p = 0.5\), as \(\max N(p; \Theta) = N(0.32, 0.45) = e^{0.1468n}\), the network of \(\Theta = 0.45\) is best in a sense of maximizing the number of equilibrium states in which \(p = 0.32\).
Continuous change of the threshold parameter $\Theta$ leads a continuous change of the activity in equilibrium states at $\Theta \leq 0.96$. As shown in Fig.2(c), when $\Theta$ is larger than 0.96, any states except a $p = 0$ state cannot be an equilibrium state.

Here we intuitively introduce a reason of why the $p = 0$ state is the only equilibrium one for the (0,1) network, when a threshold value is set larger than a specific value. For the (-1,1) case, the variance of $u_i$, $i = 1, \ldots, n$ is independent of the activity of the current state, while for the (0,1) case, the variance of $u_i$ is $np$ which depends on the activity. Here we assume that $u_i$, $i = 1, \ldots, n$ are mutually independent. In order to maximize the probability that the activity $p$ states are equilibrium states, the threshold parameter $\Theta$ should satisfy

$$p' = 1 - \Phi \left( \frac{\Theta}{\sqrt{p'}} \right),$$

since $u_i \sim \mathcal{N}(0, np)$. Equation (2.23) should hold also in a sense of keeping the total activity in the equilibrium state. The dotted line in Fig.2(c) shows the relationship between $\Theta$ and $p'$, and indicates that the network state falls into a $p = 0$ state when a threshold value is set larger than a specific value.

As one can also easily imagine to see the network of $\Theta = 0.96$ depicted in Fig.2(a), the maximum value of $N(p; \Theta)$ is negative at $0 < p \leq 1$ for the networks of $\Theta > 0.96$. Then, in such networks, the expected number of equilibrium states is smaller than 1. Since a $p = 0$ state is an equilibrium state when $\Theta \geq 0$, the expected number of equilibrium states is 1 or more at $\Theta \geq 0$. Therefore, at $\Theta > 0.96$, the $p = 0$ state becomes to be an only equilibrium (stable) state.

The result of the numerical calculation depicted with the solid line in Fig.2(c) can be used to design a network which has equilibrium states maximally at a specific activity $p'$ by adjusting the threshold parameter $\Theta$ at $p' \geq 0.25$. In a network designed based on equation (2.23) depicted with the dotted line in Fig.2(c), the equilibrium states are not concentrated on the activity $p'$ states, because $u_i$, $i = 1, \ldots, m$ are not mutually independent by the effect of the symmetricity of connection weights.

When a network is designed to have equilibrium states at activity $p' < 0.25$, the equilibrium states at activity $p' > 0.25$ are also generated. $\sum_{i=0}^{p'}$, the combinatorial number of states whose activity is $p' < 0.25$, is negligibly small compared to that of states whose activity is $0.25 < p' < 0.75$. Therefore one cannot design a network which has a set of equilibrium states concentrated on a specific activity $p' < 0.25$. That becomes to be possible, when the mean of connection weights is negative and smaller than a specific value. (Date, Kurata and Amari to appear).

3 Discussion

We have presented here the statistical properties of equilibrium states in random symmetric networks. In the foregoing analysis, $\{w_{ij}\}$, $i \geq j$, are assumed to be normally distributed i.i.d. random variables with mean 0 and variance 1. Actually, the results described in previous section holds when $\{w_{ij}\}$, $i \geq j$, are subject to any i.i.d. random variables with mean 0 and variance 1. If $\{w_{ij}\}$, $i \geq j$, are i.i.d. random variables, from the central limit theorem, equation (2.5) asymptotically becomes identical to equation (2.7) for the (-1,1) case, and equation (2.19) asymptotically becomes identical to equation (2.21) for the (0,1) case for a sufficiently large number $m$ (number of excited neurons).

When the variance of connection weights takes an arbitrary value $\sigma_w^2$, the results described in this article hold exactly by replacing $\Theta = \theta/((\sqrt{\sigma_w})).$ For an arbitrary network in which $w_{ij}$
are i.i.d. subject to $\mu_0 \neq 0$, the expected number of equilibrium states can be analyzed by using the same method. Since one significant property of random symmetric networks is that one can vary the distribution or activity level of equilibrium states to control the threshold value of the network, simple networks of $\mu_0 = 0$ were analyzed in this paper.

The brain is a self-organizing or adaptive system which can automatically capture the structure of external complex environments and represent it internally. In particular, a computational mechanism of obtaining novel information and arranging it without any inconsistency to the prior knowledge is what we are eager to know. Although a fundamental principle for self-organizing long-term memory structure of real-world environments is still unknown, it is suggested that the hippocampus could act as a temporary memory for the instantaneous storage of new data, while the neocortex served as a permanent memory (see references in [1, 16, 23]). Here we remark on a applications of the random symmetric network as a module for memory systems.

When a novel or unlearned information is presented, human can learn it even if the source signal is presented briefly. Then, it is likely that primate memory systems consists of some modules which maintain the novel signal using a network attractor till the network finish learning it by changing the connection weights. A conventional model of associative network [3, 14] solely cannot keep a novel signal. If we have a random symmetric network as a subsystem, the network can maintain the novel information using a network attractor, which depend on the input signal. In other words, the random symmetric network has ability to create a representation for novel information using a network attractor.

Furthermore, by using the property that controlling the threshold value vary the distribution of attractors, a random symmetric network might represent information independent of the similarity between patterns to be memorized. If the memory system has also a representation of information which depends on the similarity between patterns to be memorized, the associative process will be effective [12, 19].

A mechanism of resetting or changing a mode whether it is time to memorize or not can be realized by using the property that a $p = 0$ state is an only equilibrium one when threshold is set larger than a specific value, i.e. inhibiting signals are driven to all neurons, for the (0,1) model. Indeed, neurophysiological experiment showed the activities of hippocampal inhibitory neurons were suppressed when rats were walking in the novel environment [23]. The implementation of a random symmetric network as a functional module for memory systems will be described in our future papers.

Neural networks have ability to represent information in various manner. In particular, they can create the representations by learning or self-organization. The important point is to elucidate the characteristics, i.e. capabilities and limitations, of various types of neural information representation and to show their merits and demerits. We continue to construct theoretical basis of memory functioning to answer the question of how external information is represented in the brain for recognition, memory, and control.

Acknowledgements We thank Klaus-Robert Müller for reading earlier version of the manuscript. A.D. wishes to thank Takashi Onori for continuous support. The work was supported in part by Grant-in-Aid #06260228 for Scientific Research on Priority Areas on "Higher-Order Brain Information Processing" from the Ministry of Education, Science and Culture of Japan. A.D. is currently supported by a Research Fellowship from Support Center for Advanced Telecommunications Technology (SCAT, 4B-06).
References


