Properties of associative memory neural networks concerning biological information encoding

Katsunori KITANO

December 1999

Abstract

Associative abilities of the neural networks concerning the information coding in the real neural systems are studied. Two models which we adopted are the sparsely coded neural network and the oscillator neural network. We theoretically analyze such models with the replica theory and the theory of the statistical neurodynamics. These theories enable us to describe the states of the systems which consist of a number of units with a few macroscopic order parameters.

It is well known that a sparsely coded network in which the activity level is extremely low has intriguing equilibrium properties. Hence, first, we study the dynamical properties of a neural network designed to store sparsely coded sequential patterns rather than static ones. Applying the theory of statistical neurodynamics, we derive the dynamical equations governing the retrieval process which are described by some macroscopic order parameters such as the overlap. It is found that our theory provides good predictions for the storage capacity and the basin of attraction obtained through numerical simulations. The results indicate that the nature of the basin of attraction depends on the methods of activity control employed. Furthermore, it is found that robustness against random synaptic dilution slightly deteriorates with the degree of sparseness.

Second, we study the static and dynamical associative abilities of an oscillator neural network in which information is encoded by the relative timing of neuronal firing. In order to analyze such abilities, we apply the replica theory and the theory of statistical neurodynamics to the oscillator model. Using the theoretical results from these analyses, we can present the phase diagram showing both the basin of attraction and the equilibrium overlap in the retrieval state. Our results are supported by numerical simulation. Consequently, it is found that both the attractor and the basin are preserved even though dilution is promoted. Moreover, as compared with the basin of attraction

in the traditional binary model, it is suggested that the oscillator model is more robust against the synaptic dilution. Taking it into account the fact that oscillator networks contain more detailed information than binary networks, the obtained results constitute significant support for the plausibility of temporal coding.

Acknowledgement

The author would like to express his sincere gratitude to Professor Toyonori Munakata for his guidance, discussions, encouragements and critically reading the manuscript. He is also grateful to Assistant Professor Akito Igarashi, and Dr. Toshio Aoyagi for valuable discussions and encouragements.

He would like to express his deep gratitude to his parents and wife for their continual support and encouragement.

Contents

Chapter 1

Introduction

1.1 Background of the thesis

There is no doubt that the study of the information processing in the brain is one of the most important theme not only in recent years but in the future. It is expected that understanding of the mechanism of the information processing in the brain contributes to the advance in the various fields of study, for example, the medical technology of the brain, the construction of the computer based on the new architecture. However, it is true that it is hard to grasp such a mechanism only with the anatomical method which is effective for the other parts of the body. This is because, since it is considered that the function of the brain lies in the collective behavior of a number of nerve cells, we must comprehend this collective behavior as well as the behavior of a single nerve cell.

At this point, we mention the anatomical knowledge of the nerve system in brief. It is said that the brain is composed of about 10^{11} nerve cells, *neurons* [1]. A neuron consists of main four parts, a *soma*, *dendrites*, *axons*, and *synapses*. The schematic picture of a neuron is indicated in Figure 1.1. Dendrites are the parts to receive the signals from the other neurons and to transmit those signals to the cell body, *soma*. At a soma, the signals received at the dendrites converge, and the *action potential* is generated if the converged inputs exceed a *threshold* potential. Since an action potential exhibits sharp changes, it

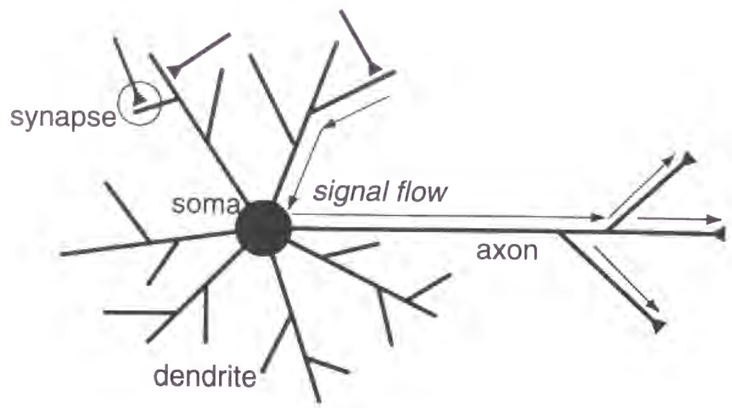


Figure 1.1: Schematic picture of a neuron

is also called a *spike*, or a *pulse*. Once an action potential is generated, such a potential is transmitted to the other neurons through *axons* as a output. An axon of a neuron contacts with a dendrite (or, sometimes, a soma) of another neuron through a *synapse*. The evoked action potential is transmitted through the synapse from the pre-synaptic neuron (the neuron sending the signal) to a post-synaptic neuron (the neuron receiving the signal). The synapse determine the efficiency of transmission of the signal from the pre-synaptic neuron to the post-synaptic neuron. Since changes of the synaptic efficiency depend on the context, it is said that such changes correspond to *learning*.

The study of the functions of the neural systems originated from the study by McCulloch and Pitts, in which a formal model of the real neuron mentioned above was proposed [2]. According to the simple formalization a neuron takes two states; one is the inactive state and the other is the active state. The former and the latter correspond to the state with few action potentials emitted and the state with action potentials emitted frequently, respectively. In the model, the inactive state and the active state are expressed as the value 0 and 1, respectively. Which states a neuron takes depends on the input to the neuron. The input to the ith neuron is determined by the summation of the outputs of the other neurons amplified by the synaptic efficacy between the other neurons and the

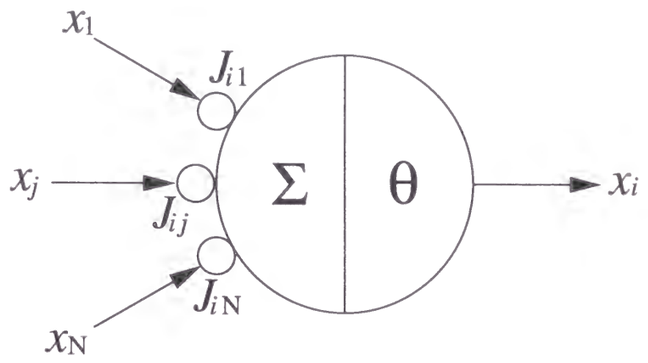


Figure 1.2: McCulloch-Pitts neuron

ith neuron. If the input exceed the threshold, the output of the neuron i takes the value 1, otherwise, 0 (see Figure 1.2). Namely, the state of the ith neuron, x_i, is determined by the equation

$$x_i = \Theta\left(\sum_j J_{ij}x_j - \theta\right), \tag{1.1}$$

where J_{ij} and θ are the synaptic efficacy between the jth neuron and the ith neuron and the threshold, respectively. The function $\Theta(u)$ is a step function; $\Theta(u) = 0$ $(u < 0)$ and 1 $(u > 0)$.

This formalization enabled us to make mathematical approaches to the study of the neural systems. And many kinds of mathematical models, *neural networks*, have been proposed and analyzed in order to understand how the functions are realized in the neural systems. Among these proposed models, *the associative memory model*, which is independently proposed by Nakano, Kohonen and Anderson, is one of the most successful ones [5, 6, 7]. The model possessing the following function is referred as the associative memory model: if a certain key as a input is given to the network, the network can retrieve the memory which is the most related to the input of the ones stored in the network. In particular, a lot of progress in the study of such a type of the models has been made since it was refined by Hopfield [8]. He discussed the analogy between his version of the associative memory model, *Hopfield model*, and the spin glass in the field of

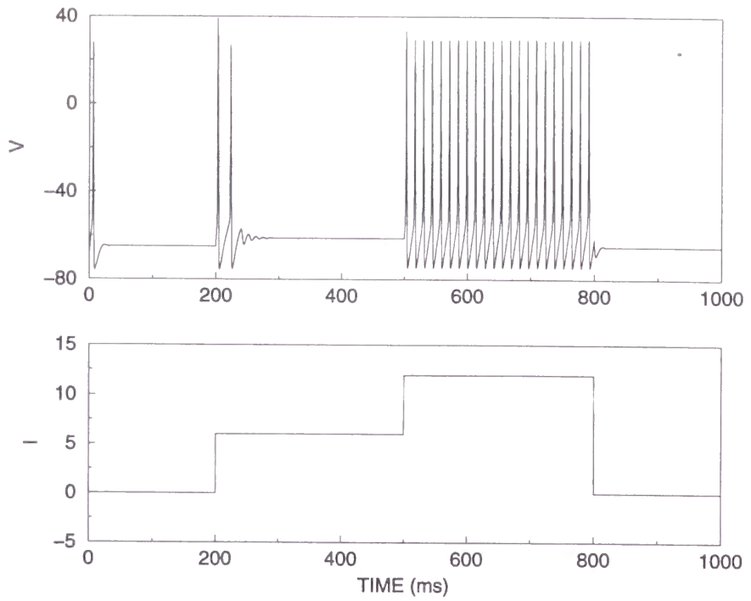


Figure 1.3: Typical behavior of Hodgkin-Huxley model

statistical physics, and proved the existence of the Lyapnov function (the energy function) of his model. In addition, the performance of the model such as the storage capacity was investigated. With this trend, using the replica theory which is a method of statistical physics, it was obtained by Amit et al that the storage capacity of the Hopfield model is 0.138 [10]. The equilibrium properties such as the critical overlap and spurious memory states were also obtained with the same method. In recent years, the studies for the retrieval process have been reported by several authors. Although attempts to analyze with path integral method were made, it is seemed that to solve this problem strictly is hard in the present circumstances [16, 17, 18]. However, the theory proposed by Amari and Maginu and developed by Okada is effective to obtain approximate solutions [20, 21]. Thus, the properties of the Hopfield type of the associative memory model are considerably understood.

On the other hand, the dynamics of a single neuron have been investigated by the

physiological studies. With experimental method, Hodgkin and Huxley derived the following equations describing the dynamics of the membrane potential V and the other parameters [3]

$$\begin{cases} C\dfrac{dV}{dt} &= -g_L(V-E_L) - g_{Na}m^3h(V-E_{Na}) - g_Kn^4(V-E_K) + I \\ \dfrac{dm}{dt} &= \alpha_m(V)(1-m) - \beta_m(V)m \\ \dfrac{dh}{dt} &= \alpha_h(V)(1-h) - \beta_h(V)h \\ \dfrac{dn}{dt} &= \alpha_n(V)(1-n) - \beta_n(V)n \end{cases} \tag{1.2}$$

Here, C, g_L, g_{Na}, g_K, E_L, E_{Na} and E_K are constants depending on the membrane properties. I is the external inputs. m, h and n are the variables representing dynamically changing conductances. As these equations indicating, even the dynamics of a single neuron is complicated. While the state of the MuCulloh-Pits neuron (1.1) is determined by only the external inputs in the passive way, the Hodgkin-Huxley equations (1.2) show the state of a neuron is determined by the internal dynamics as well as the external inputs. Taking account of physiological knowledge such as neuronal activities mentioned above, we cannot help but conclude that the picture of the information processing obtained through the study of the models on the traditional simple formulation is not necessarily plausible from biological points of view. In addition, it is expected that the model adopting the features observed in real biological systems can perform much more flexible information processing. Therefore, the studies on such models must be the next step.

1.2 Associative memory neural networks

In the studies of neural networks, the *association* is described in the following way. Let us consider that a network consists of N neurons and information is coded by the configurations of the network. When the state of the ith neuron at time t is denoted by the variable,

$S_i(t)$, the configuration is represented with the vector, $\mathbf{S}(t) = (S_1(t), S_2(t), \cdots, S_N(t))$. This N-dimensional vector is often called *pattern*, especially when the $S_i(t)$ s are binary variables taking 01 or ± 1. If the network is set to an initial input pattern $\mathbf{S}(0) = \mathbf{X}$, the network finally takes an output pattern $\mathbf{S}(\infty) = \mathbf{Y}$ after autonomously developing. The associative memory problem is to design the network capable of storing P input-output relations $(\mathbf{X}^1, \mathbf{Y}^1), (\mathbf{X}^2, \mathbf{Y}^2), \cdots, (\mathbf{X}^P, \mathbf{Y}^P)$ in such a way that when the input pattern $\mathbf{X}^k$ is presented to the network, the network produces the corresponding output pattern $\mathbf{Y}^k$. For this problem, two types of associative memory can be considered. One is the auto-associative memory in which the input pattern is the same as the output pattern, $\mathbf{X}^k = \mathbf{Y}^k$. The other is the hetero-associative memory, $\mathbf{X}^k \neq \mathbf{Y}^k$. The important characteristic of the associative memory is that even if the input pattern is not $\mathbf{X}^k$ but $\tilde{\mathbf{X}}^k$ which closely resembles $\mathbf{X}^k$, the network can retrieve the output $\mathbf{Y}^k$. Thus, the memory which is insensitive to small errors in the input patterns is referred to *content-addressable* memory.

Let us introduce the representative auto-associative memory model, the Hopfield model. The state of the ith neuron is determined by

$$S_i = \mathrm{sgn}\left(\sum_j J_{ij} S_j - \theta_i\right), \tag{1.3}$$

where

$$\mathrm{sgn}(x) = \begin{cases} 1 & (x \geq 0) \\ -1 & (x < 0) \end{cases}. \tag{1.4}$$

Here θ_i is the threshold, which is set to 0 in this model. There are at least two ways in which we might carry out the updating specified by (1.3). We could do it *synchronously*, updating all units simultaneously at each time step. Or we could do it *asynchronously*, updating them one at at time. While the dynamical properties extensively depend on which rules is adopted, there is few differences the in respect with equilibrium properties. The asynchronous updating is adopted in the Hopfield model.

The pattern to be memorized, $\xi^\mu = (\xi_1^\mu, \xi_2^\mu, \cdots, \xi_N^\mu)$, is embedded in the synaptic connections J_{ij} s. Although the form of J_{ij} is not unique, it is modified according to the Hebb rule, which is based on Hebb's postulate proposed from biological points of view [4],

$$\Delta J_{ij} \propto \xi_i^\mu \xi_j^\mu \tag{1.5}$$

in the Hopfield model. This indicates that if the state of the ith neuron is the same as that of the jth neuron, the coupling between them is strengthened, otherwise, weakened. In the case that the network stores P patterns, J_{ij} s turn to be

$$J_{ij} = \frac{1}{N} \sum_\mu^P \xi_i^\mu \xi_j^\mu. \tag{1.6}$$

In the case that P is small, even naive SN analysis enables us to proved that such a way to determine J_{ij} s is appropriate [1]. However, it is impossible using such an analysis in general.

Because the system consists of a number of units, it is useful to pursue the behavior of macroscopic order parameters rather than that of each neurons. As such an order parameter, the overlap between the configuration of the network $\mathbf{S}(t)$ and the target pattern ξ^μ can be defined:

$$m(t) = \frac{1}{N} \sum_j^N \xi_j^\mu S_j(t). \tag{1.7}$$

It is easy to find that $m(t)$ is the direction cosine of $\mathbf{S}(t)$ and ξ^μ and if $\mathbf{S}(t) \to \xi^\mu$, $m(t) \to 1$.

The properties of the network are described mainly by the behavior of the overlap $m(t)$. The important properties of the network are as follow.

- *storage capacity*

  When the number of memorized patterns P increases, the value for which the network comes to fails to retrieve *all* memorized patterns exists. This is caused by the fact that the memorized patterns are embedded into the synaptic connections in the

distributed way. Such a critical value of P is called the storage capacity P_C. Or $\alpha_C = P_C/N$, which is normalized by the number of neurons N, is often used.

- *critical overlap*

  $m(\infty)$ is the overlap when the network relax to the equilibrium state of the network and represents the accuracy of the retrieval. Since the non-target patterns play roles of noise in the retrieval processes, the increase in the number of memorized patterns stands for the increase in the variance of such a noise. Therefore, as P increases, $m(\infty)$ is gradually decreases. $m(\infty)$ at P_C, which is the accuracy in the worst case, is referred to the critical overlap. In the case of $P > P_C$, it is expected that $m(\infty)$ becomes 0.

- *basin of attraction*

  The associative memory neural networks are capable of the error correction. If an initial overlap $m(0)$ is larger than a certain value, the network succeeds in retrieval. Otherwise, the network fails to retrieve. The sets of the initial states which induce to succeed in retrieval is called the basin of attraction. It is easy to find that, as P increases, the basin of attraction is reduced.

The first and second properties of the Hopfield model have been solved by Amit et al [10]. Since Hopfield found the energy function of the system

$$H[\mathbf{S};\xi] = -\frac{1}{2}\sum_{i\neq j} J_{ij}S_iS_j, \tag{1.8}$$

they could apply the replica theory which is used for the study of spin glass. Introducing the system to thermal noise, we can grasp the equilibrium states through the free energy

$$F[\xi] = -\frac{1}{\beta}\log Z[\xi], \tag{1.9}$$

where $Z[\xi] = \sum_{\mathbf{S}} \exp(-\beta H[\mathbf{S};\xi])$ is the partition function. The replica theory is the method to average of the free energy $F[\xi]$ over the memorized patterns ξ. From such

an analysis, the averaged free energy is obtained with the coupled equations of the order parameters such as the overlap. As a result, it is found that the critical value of P for which non-zero solution of $m(\infty)$ exists is $0.138N$. In addition, the other properties of the equilibrium states also have been studied using the order parameters equations.

In order to investigate the third property, basin of attraction, it is necessary to analyze the dynamics in the retrieval process. In general, it is possible that such properties depend upon the update rule, synchronous update or asynchronous update. For the synchronous update model, the statistical neurodynamics theory proposed by Amari et al is the most practical one [20]. It is easy to find that the input to a neuron can be separated into the signal term which induces recollection and the crosstalk noise term caused by the non-target memorized patterns. Then, the essence of the theory is to assume such a crosstalk noise term as the Gaussian noise with the mean 0 and the temporally developing variance. However, the theory could describe the behavior of the system only qualitatively. Okada developed the theory taking into account of the temporal correlation of the crosstalk noise in different time steps [21]. Such analyses are in good agreement with the behavior of the system. On the other hand, the analysis with the dynamical replica theory by Coolen et al succeeds the behavior of the asynchronous update model.

1.3 Outline of the thesis

With the trend mentioned above, we studied two models taking account of the phenomena which are observed experimentally. These two model concern encoding of information in biological neural systems, one concerning *sparse coding* and the other concerning *temporal coding*.

In chapter 2, we investigate the properties of the neural network model processing the sparsely coded patterns. In real nerve systems, it is observed that the number of the active neurons is extremely small. According to a certain estimation, it was reported that

a fraction of active neurons is about 4 – 5 %. Based on this physiological knowledge, the model designed to process the patterns expressed with a small fraction of active bits was proposed by Willshaw et al [9]. These patterns, in which the active bits are sparsely distributed in a spatial sense, are called the sparsely coded patterns. From the viewpoint of information theory, the sparsely coded patterns is realistic rather than so-called random patterns. This is because, in general cases, the active bits and inactive bits corresponds to the foreground and the background, respectively, and the fraction of the background bits is larger that of the foreground bits.

By several authors, it has been reported that a sparsely coded network has intriguing equilibrium properties, for example, the storage capacity increases as the fraction of active neurons decreases [10, 11, 12, 13, 14]. We study the dynamical properties of a model designed to process sparsely coded *sequential* patterns rather than static ones. Applying the theory of statistical neurodynamics, we derive the dynamical equations governing the retrieval process which described by some macroscopic order parameters. The obtained theoretical results using the derived equations are compared with the result of numerical simulations.

Since the states of memorized patterns exhibit low activities, the activity level of the network is desirable to be kept low. This indicates that the mechanisms to stabilize the activity level significantly affect the retrieval abilities of the network. We compare the performances of the models with the three different types of the activity control mechanisms.

In the case of the sparsely coded network to process the static memories, as the activity level of memorized patterns a decreases, the storage capacity of memories diverge as the asymptotic form $1/a|\ln a|$ [11, 13]. This result is obtained expanding the equilibrium equations of the order parameters in term of a in the sparse coding limit $a \to 0$. We apply the same procedure to the present model for the *sequential* memories and discuss

the asymptotic property of the storage capacity.

Furthermore, the robustness against defects in some parts of the system is supposed to be an advantage which the distributed processing systems such as neural networks possess. The robustness against the random synaptic dilution is studied using both the theory and numerical simulations.

Chapter 3 is devoted to discuss the static and dynamical properties of oscillator neural networks. The collective oscillatory behavior of neuronal group has been observed in many biological neural systems – the primary visual cortex, olfactory bulb, hippocampus and so on. The experimental observation by Gray et al suggests that the temporal coherence of neuronal activity may contribute to information processing in real biological systems [30]. In addition, this observation makes the opportunity to reconsider traditional hypothesis that information in neural systems is encoded by firing rate, i.e. *rate coding*. From theoretical point of view, *temporal coding* is considered to be more suitable to dynamical information processing as real neural systems are. For instance, the utilization of temporal features of the neuronal activities enables us to solve so-called *binding problem* more neatly [31]. In anticipation of these points, as one of the models capable of describe the temporal aspects of real neuronal activities, the oscillator neural networks are worthy of being studied.

Many theoretical works concerning the associative memory type of oscillator neural networks have been reported by a number of authors [32, 33, 34, 35, 36, 37, 38, 39, 40, 41]. Among them, there are several theoretical works directly related to our model. In particular, Cook showed that the storage capacity of the oscillator model is 0.038, provided that the Hebbian rule with random phase patterns is used [40]. Hence, we discuss how the storage capacity is affected when the synapses are randomly diluted [27, 42, 43, 44]. If the diluted synapses are symmetric, we can apply the method used in statistical physics, the replica theory, to the model. Using such a theory, we derive the coupled equations

describing the equilibrium states. In addition, we compare the robustness against random synaptic dilution of the oscillator model with that of the traditional Hopfield model. It is interesting that how robust the oscillator model is, which is encoded by the more detailed information of the timings of spikes.

Next, we study the retrieval dynamics of oscillator neural networks. The ability of error correction which is one of the important abilities of associative memory models belongs to the dynamical properties. However, there is no dynamical theory with an extensive number of stored patterns in fully connected oscillator networks. Therefore, we apply the method of statistical neurodynamics to the oscillator model. The dynamical equations obtained with the theory enable us to predict initial permissible errors for successful recollection. We also investigate the ability of error correction in the case of random synaptic dilution.

Furthermore, we mention the relation between the replica theory and the statistical neurodynamics theory to show that the same result for equilibrium properties is obtained from both theories.

In the final chapter, we summarize main results and give some comments.

Chapter 2

Associative abilities of neural network for sparsely coded sequential patterns

2.1 Introduction

For the purpose of constructing more realistic mathematical neural network models (e.g, the Hopfield model [8]), so-called "random" patterns, which have been used for simple theoretical treatments, have been reconsidered. In a network capable of processing these random patterns, it is frequently supposed that statistically half of the neurons are allowed to be active. However, such a situation is not realistic for two reasons. First, according to the results of physiological studies, the activity level of real neural systems is thought to be low. Second, in a meaningful pattern, information is generally encoded by a small fraction of bits in a background which occupies most of the total area.

With these points in mind, neural networks loading sparsely coded patterns have been studied by many authors [9, 10, 11, 12, 13, 14]. These authors have reported that the maximal number of patterns stored in the network increases as the fraction of active neurons a decreases. Furthermore, the storage capacity in such a situation diverges as $-1/a \ln a$ which is the optimal asymptotic form obtained by Gardner [15]. However, considering the fact that the information content in a single pattern is reduced with the

degree of sparseness, we cannot immediately conclude that sparse coding enhances the associative ability. Rather, what we should note is that the optimal bound is obtained even for models with a relatively simple Hebbian learning rule.

While, owing to these studies, progress in the understanding of the equilibrium properties of sparsely coded networks has been made, many unsolved problems remain in regard to dynamical aspects. In particular for associative memory models, considering the associative ability for a noisy initial pattern to be dynamically corrected, it is necessary to consider the basin of attraction in order to grasp a network's characteristics properly. In recent years, several theories treating retrieval process have been proposed [16, 17, 18, 19, 20, 21]. Among these, we note that the method of statistical neurodynamics is practically useful, because it enables us to describe long-term behavior when a network succeeds in retrieval [20, 21]. However, for sparse coding, there is quantitative discrepancy between the results obtained from this theory and numerical simulation in the case of auto-association, which implies a difficulty in treating the strong feedback mechanism with this model [22]. On the other hand, sparse coding for sequential associative memory has not yet been studied in detail [24, 25]. In the present paper, we study this point by applying the method of statistical neurodynamics to a model for sequential associative memory.

2.2 Associative memory models for sparsely coded patterns

Let us consider the situation in which a neural network which consists of N McCulloch-Pitts neurons is designed to store sequential patterns rather than static ones. Each neuron obeys discrete *synchronous* dynamics described by

$$S_i(t+1) = F[h_i(t)], \tag{2.1}$$

$$h_i(t) = \sum_{j=1}^{N} J_{ij} S_j(t) - \theta, \tag{2.2}$$

where $S_i(t)$ and $h_i(t)$ are the state and the internal potential of the ith neuron at time t, respectively. Although we have written the transfer function in the general form $F(u)$, we consider the case $F(u) = \Theta(u)$; i.e. $F(u)$ is a step function. In this case, the state $S_i(t)$ takes only two values, 1 (firing state) and 0 (resting state). The quantities θ and J_{ij} represent the uniform threshold and the strength of the synaptic connection between the ith and jth neuron, respectively.

We assume that the stored patterns are generated with the probability

$$P(\xi_i^\mu) = a\delta(\xi_i^\mu - 1) + (1-a)\delta(\xi_i^\mu), \tag{2.3}$$

where ξ_i^μ is the state of the ith neuron in the μth pattern. Then, the activity for this network, $\frac{1}{N}\sum_i \xi_i^\mu$, assumes an average value of a. In particular, the case $a \to 0$ is referred to as "sparse coding". In order to make the network possess associative memory dealing with these patterns, the J_{ij}s must be designed appropriately. In the present paper, to construct a network capable of recalling a sequence of $P = \alpha N$ patterns in Figure 2.1, defined by such as $\xi^1 \to \xi^2 \to \cdots \to \xi^P \to \xi^1 \to \cdots$, we adopt covariance learning

$$J_{ij} = \frac{1}{a(1-a)N} \sum_{\mu}^{\alpha N} (\xi_i^{\mu+1} - a)(\xi_j^\mu - a), \tag{2.4}$$

which is usually adopted in the context of learning the sparsely coded patterns.

For such a network, the macroscopic state is found to be described by the following order parameters:

$$m^\mu(t) = \frac{1}{a(1-a)N} \sum_{j}^{N} (\xi_j^\mu - a) S_j(t) \tag{2.5}$$

$$x(t) = \frac{1}{aN} \sum_{j}^{N} S_j(t). \tag{2.6}$$

Here, $m^\mu(t)$ is the overlap with the target pattern ξ^μ. As the configuration of the network becomes close to the target pattern, this value approaches unity. The function $x(t)$

represents the activity of the network. On studying the retrieval processes, we mainly discuss the time evolution of these parameters.

2.3 Analysis with statistical neurodynamics theory

2.3.1 Derivation of equations describing retrieval dynamics

We consider the "condensed" situation in which only one overlap is sizable:

$$m^{\rho}(t) \sim O(1), \qquad m^{\mu}(t) \sim O\left(\frac{1}{\sqrt{N}}\right) \quad (\mu \neq \rho). \tag{2.7}$$

Here, ξ^{ρ} is the pattern to be retrieved at time t. Then, the internal potential $h_i(t)$ in Eq. (2.2) can be separated as

$$h_i(t) = \bar{\xi}_i^{\rho+1} m^{\rho}(t) - \theta + \sum_{\mu \neq \rho}^{\alpha N} \bar{\xi}_i^{\mu+1} m^{\mu}(t), \tag{2.8}$$

where we have written $\bar{\xi}_i^{\mu}$ as $\xi_i^{\mu} - a$. In this process, the first and the second terms in Eq. (2.8) are together regarded as the signal to induce recollection of the target pattern $\xi_i^{\rho+1}$ at the subsequent time step, $t+1$, while the remaining term is regarded as noise. For convenience, we define the noise term $z_i(t)$ as

$$z_i(t) = \sum_{\mu \neq \rho}^{\alpha N} \bar{\xi}_i^{\mu+1} m^{\mu}(t). \tag{2.9}$$

The quantity $z_i(t)$ is the crosstalk noise from the non-target patterns. The essence of the theory is to treat the crosstalk noise $z_i(t)$ as Gaussian noise with mean 0 and variance $\sigma(t)^2$ [25]. It has been confirmed numerically that this assumption is valid as long as the network succeeds in retrieval [26].

Now we derive the dynamical equations for the overlap $m(t)$ and the activity $x(t)$. The definition of the overlap leads to the equation

$$m^{\rho+1}(t+1) = \frac{1}{a(1-a)N} \sum_{i}^{N} \left\langle\!\!\left\langle \bar{\xi}_i^{\rho+1} F\left[\bar{\xi}_i^{\rho+1} m^{\rho}(t) - \theta + z_i(t)\right] \right\rangle\!\!\right\rangle_{\xi}, \tag{2.10}$$

where $\langle\!\langle \cdots \rangle\!\rangle_{\xi}$ denotes the average over the stored patterns. In the same way, we can write the equation for the activity $x(t)$,

$$x(t+1) = \frac{1}{aN} \sum_{i}^{N} \left\langle\!\!\left\langle F\left[\bar{\xi}_i^{\rho+1} m^{\rho}(t) - \theta + z_i(t)\right] \right\rangle\!\!\right\rangle_{\xi}. \tag{2.11}$$

Next, we examine the time development of the variance $\sigma(t)^2$. Expressing $z_i(t+1)$ as

$$z_i(t+1) = \frac{1}{a(1-a)N} \sum_{j \neq i}^{N} \sum_{\mu \neq \rho+1}^{\alpha N} \bar{\xi}_i^{\mu+1} \bar{\xi}_j^{\mu} F[h_j(t)], \tag{2.12}$$

we must consider the dependence of $h_j(t)$ on ξ_j^{μ} when summing over μ. In the internal potential $h_j(t)$, the term $\bar{\xi}_j^{\mu} m^{\mu-1}(t)$ is estimated to be $O(1/\sqrt{N})$. Therefore, we expand the function $F[h_j(t)]$ in terms of $\bar{\xi}_j^{\mu} m^{\mu-1}(t)$:

$$\begin{aligned} S_j(t+1) &= F[\hat{h}_j(t) + \bar{\xi}_j^{\mu} m^{\mu-1}(t)] \\ &= F[\hat{h}_j(t)] + F'[\hat{h}_j(t)] \bar{\xi}_j^{\mu} m^{\mu-1}(t) \\ &= \hat{S}_j(t+1) + \bar{\xi}_j^{\mu} F'[\hat{h}_j(t)] \frac{1}{AN} \sum_{k} \bar{\xi}_k^{\mu-1} S_k(t). \end{aligned} \tag{2.13}$$

Here $A = a(1-a)$ and $h_j(t) = \hat{h}_j(t) + \bar{\xi}_j^{\mu} m^{\mu-1}(t)$. We now assume that $\hat{S}_j(t+1)$ in Eq. (2.13) is independent of ξ_j^{μ}. As a result, we obtain the following equation for $z_i(t+1)$:

$$z_i(t+1) = \frac{1}{AN} \sum_{j} \sum_{\mu} \bar{\xi}_i^{\mu+1} \bar{\xi}_j^{\mu} \hat{S}_j(t+1) + U(t) \frac{1}{AN} \sum_{k} \sum_{\nu} \bar{\xi}_i^{\nu+1} \bar{\xi}_k^{\nu-1} S_k(t). \tag{2.14}$$

Squaring Eq.(2.14), we obtain

$$\begin{aligned} z_i(t+1)^2 &= \alpha a x(t+1) + U(t)^2 z_i(t)^2 \\ &\quad + U(t) \left(\frac{1}{AN}\right)^2 \sum_{j,k} \sum_{\mu,\nu} \bar{\xi}_i^{\mu+1} \bar{\xi}_i^{\nu+1} \bar{\xi}_j^{\mu} \hat{S}_j(t+1) \bar{\xi}_k^{\nu-1} S_k(t) \end{aligned} \tag{2.15}$$

Here the first term and the second term in Eq. (2.15) come from the square of the first term and the second term in Eq. (2.14), respectively. The last term in Eq. (2.15) arise from the product of the first term and the second term in Eq. (2.14). Applying the relation Eq. (2.13) to the term $S_k(t)$ in Eq. (2.15), (2.15) becomes

$$\begin{aligned} z_i(t+1)^2 &= \alpha a x(t+1) + U(t)^2 z_i(t)^2 + C(t+1,t) \\ &\quad + U(t)U(t-1) \left(\frac{1}{AN}\right)^2 \sum_{j,l} \sum_{\mu,\nu} \bar{\xi}_i^{\mu+1} \bar{\xi}_i^{\nu+1} \bar{\xi}_j^{\mu} \hat{S}_j(t+1) \bar{\xi}_l^{\nu-2} S_l(t-1) \end{aligned} \tag{2.16}$$

where we define

$$\begin{aligned} C(t+1,t) &= U(t)\frac{1}{AN}\sum_{j,k}\sum_{\mu,\nu}\bar{\xi}_i^{\mu+1}\bar{\xi}_i^{\nu+1}\bar{\xi}_j^{\mu}\hat{S}_j(t+1)\bar{\xi}_k^{\nu-1}\hat{S}_k(t) \\ &= U(t)\frac{1}{N}\sum_j \hat{S}_j(t+1)\hat{S}_j(t)\left[\frac{1}{A^2N}\sum_\mu(\bar{\xi}_i^{\mu+1})^2\bar{\xi}_j^{\mu}\bar{\xi}_j^{\mu-1}\right]. \end{aligned} \tag{2.17}$$

In the same way, substituting the relation Eq. (2.13) iteratively, we can take into account temporal correlations up to the initial time. Following this procedure, we obtain

$$\sigma(t+1)^2 = \alpha a x(t+1) + U(t)^2\sigma(t)^2 + \sum_{n=1}^{t+1} C(t+1,t+1-n) \tag{2.18}$$

with

$$C(t+1,t+1-n) = \prod_{\tau=1}^{n} U(t+1-\tau)\frac{1}{N}\sum_j \hat{S}_j(t+1)\hat{S}_j(t+1-n)\left[\frac{1}{A^2N}\sum_\mu(\bar{\xi}_i^{\mu+1})^2\bar{\xi}_j^{\mu}\bar{\xi}_j^{\mu-n}\right]. \tag{2.19}$$

In the case of auto-association, we cannot neglect temporal correlations of crosstalk noise $C(t+1,t+1-n)$, because this correlation plays a significant role in the retrieval process [21]. As for the present model, since ξ^μ and $\xi^{\mu-n}$ are independent of each other, except when $n = \alpha N, 2\alpha N, 3\alpha N, \cdots$, the last summation in Eq. (2.19) vanishes. Although the correlations for $n = \alpha N, 2\alpha N, 3\alpha N, \cdots$ remain, their effect can be regarded as negligible in the limit $N \to \infty$.

Consequently, the behavior of the network is described by the equations

$$m(t+1) = 1 - \frac{1}{2}[\mathrm{erfc}(\phi_1) + \mathrm{erfc}(\phi_0)] \tag{2.20}$$

$$x(t+1) = 1 - \frac{1}{2}\left[\mathrm{erfc}(\phi_1) - \frac{1-a}{a}\mathrm{erfc}(\phi_0)\right] \tag{2.21}$$

$$\sigma(t+1)^2 = \alpha a x(t+1) + U(t)^2\sigma(t)^2, \tag{2.22}$$

with

$$\phi_1 = \frac{(1-a)m(t) - \theta}{\sqrt{2}\sigma(t)} \tag{2.23}$$

$$\phi_0 = \frac{am(t) + \theta}{\sqrt{2}\sigma(t)} \tag{2.24}$$

$$U(t) = \frac{1}{\sqrt{2\pi}\sigma(t)}\left[ae^{-\phi_1^2} + (1-a)e^{-\phi_0^2}\right] \tag{2.25}$$

$$\mathrm{erfc}(t) = \frac{2}{\sqrt{\pi}}\int_{-t}^{\infty} e^{-u^2}du, \tag{2.26}$$

where we have set $F(u) = \Theta(u)$ and replaced the site average $\frac{1}{N}\sum_i^N \cdots$ with the average over the Gaussian noise $\langle\cdots\rangle_{z(t)}$ in the limit $N \to \infty$. For initial values, we can set $\sigma(0) = \sqrt{\alpha a x(0)}$ and choose arbitrary values for $m(0)$ and $x(0)$.

2.3.2 For several activity control mechanisms

Global inhibitory interaction

In a sparsely coded network, activity control is an important factor for good retrieval quality. Introducing the global inhibitory interaction such as

$$J_{ij}^{\mathrm{inh}} = J_{ij} - \frac{g}{aN}, \tag{2.27}$$

the activity can be dynamically controlled [10, 14]. The second term contributes as a global inhibitory interaction, and g represents its strength. In this case, the internal potential $h_i(t)$ is expressed as follows:

$$\begin{aligned} h_i(t) &= \sum_{j=1}^{N} J_{ij}^{inh}S_j(t) - \theta \\ &= \sum_{j=1}^{N} J_{ij}S_j(t) - gx(t) - \theta \end{aligned} \tag{2.28}$$

If the activity level of the network at time t, $x(t)$, greatly increases, each neuron receives a stronger inhibitory signal $-gx(t)$, so that $x(t+1)$ decreases. We can undertake a treatment of the retrieval process in this case in a manner similar to that undertaken above. We then find that equations (2.23) and (2.24) are modified as

$$\phi_1 = \frac{(1-a)m(t) - gx(t) - \theta}{\sqrt{2}\sigma(t)} \tag{2.29}$$

$$\phi_0 = \frac{am(t) + gx(t) + \theta}{\sqrt{2}\sigma(t)}. \tag{2.30}$$

Self-control threshold

Another model possessing an activity control mechanism is that with a time-dependent threshold which is calculated at each time step so that the activity of the network can be kept the same as that of the retrieved pattern [22]. Recently, as an improved model, a "self-control" model has been proposed [23]. In this model, the time-dependent threshold $\theta(t)$ adapts itself according to the activity a and the variance of crosstalk noise $\sigma(t)$. If a is sufficiently small, it takes the form

$$\theta(t) = \sigma(t)\sqrt{-2\ln a}. \tag{2.31}$$

However, from the biological point of view, it is not plausible that the network monitors the statistical quantity of the crosstalk noise. Hence, in the present paper, in place of $\sigma(t)$, we choose the leading term of $\sigma(t)$, $\sqrt{a\alpha x(t)}$. Then, we simply use

$$\theta(t) = \sqrt{-2x(t)\alpha a\ln a} \tag{2.32}$$

in place of the expression in Eq. (2.31).

2.4 The results

We carry out numerical simulation of the present model, and obtain Fig. 2.2 and 2.3 showing typical example in the case of successful association and in the case of failure association, respectively. In these figures, the top and the bottom corresponding to evolution of crosstalk noise and distribution of internal fields, respectively. Although the average of crosstalk noise $\langle z(t)\rangle$ become almost zero in both cases, temporal correlation $\langle z(t)z(t-1)\rangle$ takes finite value when the network fail to retrieve sequential patterns. From there results, the assumption of mean 0 of crosstalk noise is found to be valid, and it is numerically confirmed that temporal correlation of crosstalk noise vanish in the case of successful association. Furthermore, while the internal potentials are clearly separated

into two lumps, that is, one consists of active sites ($h > 0$) and the other of inactive sites ($h < 0$) in Fig. 2.2, the internal fields obey the single Gaussian like distribution with broad variance in Fig. 2.3.

We now compare our theoretical results with numerical simulations. Figures 2.4-2.6 display the results of the model using only a uniform threshold θ, a uniform threshold θ and the inhibitory interaction g, and a self-control threshold $\theta(t)$, respectively. In the first two cases, θ and g are optimized so as to maximize the storage capacity. The three curves displayed in the figures represent, from the top, the equilibrium activity, the equilibrium overlap, and the basin of attraction, respectively. The point where these three curves vanish indicates the storage capacity. From the results, it is found that the theoretical curves provide a good prediction of the retrieval properties in the network. Although with respect to storage capacity, these three cases differ very little, the differences among the activity control methods are reflected in the shapes of basin of attraction. While the basin becomes gradually narrow as α increases in the first case, the basin for $\alpha > 0$ is wider than that for $\alpha = 0$ in the second case. Furthermore, in the last case, the minimum initial overlap for which the network succeeds in retrieval becomes zero when $\alpha = 0$.

2.5 Storage capacity in the sparse coding limit

In the case of auto-associative memory model, it have been reported that the storage capacity diverges as the asymptotic form $-1/a\ln a$ in the sparse coding limit $a \to 0$. For the purpose of investigating the storage capacity of the present model in such a limit, we must describe the stationary state. Using the dynamical equations, we can describe the macroscopic order parameters in the stationary state as follows:

$$m = 1 - \frac{1}{2}\left[\mathrm{erfc}(\phi_1) + \mathrm{erfc}(\phi_0)\right] \tag{2.33}$$

$$x = 1 - \frac{1}{2}\left[\mathrm{erfc}(\phi_1) - \frac{1-a}{a}\mathrm{erfc}(\phi_0)\right] \tag{2.34}$$

$$\sigma = \sqrt{\frac{\alpha a x}{1-U^2}} \tag{2.35}$$

$$\phi_1 = \frac{(1-a)m-\theta}{\sqrt{2}\sigma} \tag{2.36}$$

$$\phi_0 = \frac{am+\theta}{\sqrt{2}\sigma} \tag{2.37}$$

In the limit $a \to 0$, we can approximate U to 0. Furthermore, to satisfy $x \sim 1$ which means that the network keeps the same activity as the target pattern, the order of error bits of misfiring $\mathrm{erfc}(\phi_1)$ must be the same as that of $\mathrm{erfc}(\phi_0)/a$. Namely, the relation

$$\mathrm{erfc}(\phi_0) \sim a\mathrm{erfc}(\phi_1) << \mathrm{erfc}(\phi_1) \tag{2.38}$$

is obtained. Therefore, the following approximations for the stationary state of order parameters in such a limit are obtained,

$$m \sim 1-\frac{1}{2}\mathrm{erfc}(\phi_1), \tag{2.39}$$

$$x \sim 1, \tag{2.40}$$

$$\sigma \sim \sqrt{a\alpha}, \tag{2.41}$$

$$\phi_1 \sim \frac{m-\theta}{\sqrt{2a\alpha}}, \tag{2.42}$$

$$\phi_0 \sim \frac{\theta}{\sqrt{2a\alpha}}. \tag{2.43}$$

The error bits $\mathrm{erfc}(\phi_0)/a$ must be less than 1, that is, $\mathrm{erfc}(\phi_0)/a < 1$. Because of $\phi_0 \to \infty$ when $a \to 0$, we can carry out the asymptotic expansion of $\mathrm{erfc}(\phi_0)$, obtaining

$$e^{-\phi_0^2} < a\phi_0. \tag{2.44}$$

Here, if $e^{-\phi_0^2} \sim a^n$ is assumed, ϕ_0 is found to be

$$\phi_0 \sim \sqrt{n|\ln a|}, \tag{2.45}$$

where $n \geq 1$ satisfies the inequality Eq. (2.44). From Eq. (2.43) and (2.45), the maximum storage capacity with $n = 1$ is described as follow:

$$\alpha \sim \frac{\theta^2}{2a|\ln a|}. \tag{2.46}$$

Using this result, it is found that $\phi_1 \sim (1-\theta)\sqrt{|\ln a|}$. Under the condition

$$1-\theta >> \frac{1}{\sqrt{|\ln a|}}, \tag{2.47}$$

the critical overlap is approximated to

$$m \sim 1-\frac{1}{2\sqrt{\pi}} \cdot \frac{\theta}{1-\theta} \cdot \frac{1}{\sqrt{|\ln a|}} \cdot a^{-(1/\theta-1)^2}. \tag{2.48}$$

This result is in agreement with the form obtained by Tsodyks et al [11]. In the limit $a \to 0$, keeping the condition (2.47), we can maximize θ to unity. Finally, we find the storage capacity of the present model in the sparse coding limit,

$$\alpha \sim \frac{1}{2a|\ln a|}. \tag{2.49}$$

To verify the crude calculation mentioned above, we numerically calculate the dynamical equations. The dependence of the storage capacity on the activity level a is shown in Fig. 2.7. The solid line for the reference expresses $\alpha a \sim -1/\ln a$, namely $\alpha \sim -1/a\ln a$. The data series obtained through numerical calculation approach the line as a decreases. In the present case, we have confirmed that the storage capacity diverges as $-1/a\ln a$ in the limit $a \to 0$, and it seems to approach such an asymptotic form quite slowly [17].

2.6 Random synaptic dilution

Next, we investigate robustness against random synaptic dilution. In this case, a randomly diluted synapse is represented by the random variable c_{ij}:

$$\tilde{J}_{ij} = \frac{c_{ij}}{c} J_{ij} \tag{2.50}$$

The variable c_{ij} takes the value 1 with probability c, and is 0 otherwise. In other words, c represents the ratio of connected synapses. It is known that random synaptic dilution can be statistically regarded as static noise in a synapse [27],

$$\tilde{J}_{ij} = J_{ij} + \eta_{ij}. \tag{2.51}$$

The synaptic noise η_{ij} is a Gaussian noise with mean 0 and variance η^2/N. The relation between the dilution ratio c and the variance parameter η^2 is found to be

$$\eta^2 = \frac{1-c}{c}\alpha. \tag{2.52}$$

In the model with random diluted synapses, noise component consists of two terms as follow:

$$\begin{aligned} \tilde{z}_i(t) &= z_i^c(t) + z_i^s(t) \\ &= \frac{1}{AN}\sum_{j\neq i}^{N}\sum_{\mu\neq\rho}^{\alpha N}\bar{\xi}_i^{\mu+1}\bar{\xi}_j^{\mu}S_j(t) + \sum_{j\neq i}^{N}\eta_{ij}S_j(t) \end{aligned} \tag{2.53}$$

In the above equation, the first term is the traditional crosstalk noise, and the second term arise from the random synaptic dilution. We can treat the crosstalk noise term in the same way as the discussion mentioned above. When we take the statistics of synaptic noise, we must take into account correlations between η_{ij} and η_{ji} in $S_j(t)$. Expanding $S_j(t)$ in terms of η_{ji} yields

$$z_i^s(t) = \sum_{j\neq i}^{N}\eta_{ij}\hat{S}_j(t) + S_i(t)\sum_{j\neq i}^{N}\eta_{ij}\eta_{ji}F'(h_j(t-1). \tag{2.54}$$

where $\hat{S}_j(t)$ is assumed to be independent of η_{ji}. In Eq. (2.54), the second term vanishes since η_{ij} and η_{ji} are independent. As a result, we obtain

$$\langle z_i^s(t)\rangle \equiv \sigma^s(t)^2 = \eta^2 a x(t). \tag{2.55}$$

Since $z_i^c(t)$ is assumed to be independent of $z_i^s(t)$, the resultant equation for the noise is modified as

$$\begin{aligned} \langle\tilde{z}_i(t)\rangle &\equiv \tilde{\sigma}(t+1)^2 \\ &= \sigma(t+1)^2 + a\alpha\frac{1-c}{c}x(t+1). \end{aligned} \tag{2.56}$$

In addition, ϕ_1 and ϕ_0 become

$$\phi_1 = \frac{(1-a)m(t) - gx(t) - \theta}{\sqrt{2}\tilde{\sigma}(t)} \tag{2.57}$$

and

$$\phi_0 = \frac{am(t) + gx(t) + \theta}{\sqrt{2}\tilde{\sigma}(t)}. \tag{2.58}$$

Clearly, when $c = 1$, we have $\tilde{\sigma} = \sigma$.

Figure 2.8-2.10 indicate results of theoretical analyses and numerical simulation for the three type of activity control mechanisms in the case of $c = 0.7$. All parameters except the ratio of connected synapses are the same as Fig. 2.4-2.6. In every cases, our theoretical analyses are in good agreement with the results of numerical simulations.

In Fig. 2.11-2.13, we display the dependence on the ratio of connected synapses c. From these results, it is seemed that the associative abilities are proportionate to the parameter c for all cases.

In order to examine the deterioration experienced with the decrease in the ratio of connection c at each activity a, we define the normalized storage capacity $\alpha_c^*(c) = \alpha_c(c)/\alpha_c(1)$, where $\alpha_c(c)$ is the storage capacity when the ratio of connection is c. Fig. (2.14) displays the normalized storage capacity $\alpha_c^*(c)$ as a function of c. As indicated by these results, the effect of dilution on the storage capacity is not significant. However, if the activity level a becomes small, the network gradually becomes sensitive as c decreases; the storage capacity decreases almost linearly decreases with the increase in the degree of dilution, $1-c$.

As for this point, we can make the similar discussion to that in the previous section. In this case, the variance of the crosstalk noise in the stationary states is found to be expressed as

$$\sigma \sim \sqrt{\frac{a\alpha}{c}}. \tag{2.59}$$

If $c >> a \quad (c \sim O(1))$, the storage capacity is approximated to

$$\alpha_c(c) \sim \frac{c}{2a|\ln a|}. \tag{2.60}$$

Therefore, the normalized storage capacity $\alpha_c^*(c)$ is obtained as follow:

$$\alpha_c^*(c) \sim c \tag{2.61}$$

With regard to the basin of attraction, the model with the optimized uniform threshold θ, which has the most narrow basin, is the most robust of the three.

2.7 Conclusion

Let us now summarize the results of the present study of retrieval dynamics in the neural network with sparsely coded sequential patterns.

- We derived dynamical equations for retrieval processes using the methods of statistical neurodynamics. Our theoretical results were found to be in good agreement with numerical simulations. This is due to the fact that temporal correlations of the crosstalk noise, which contribute significantly in the auto-associative memory model, vanish.

- Three types of the activity control mechanisms, the optimized uniform threshold, the global inhibitory interaction, and the self-control threshold, were investigated. Such mechanisms mainly affect the basins of attraction. Among the three models, in general the self-control model has the widest basin, and in addition, in this case there is no need to tune parameters by hand.

- For the present model, we confirmed that the storage capacity in the sparse coding limit $a \to 0$ diverges as the asymptotic form $-/a \ln a$ using both theoretical and numerical analyses.

- We applied the present theory to the model with random synaptic dilution. From the results, it is seemed that the associative abilities become reduced proportionately to the decrease of the ratio of connected synapses c. Furthermore, as the activity level becomes low, the robustness against random synaptic dilution deteriorates slightly. For low activity, the storage capacity decreases almost linearly with the ratio of connected synapses c.

Finally, let us comment on numerical simulations. When carrying out simulations, we can choose the two ways to generate patterns with a given activity a. One way is to draw from the distribution Eq. (2.3), and the other is to set Na neurons to unity and the remaining $N(1-a)$ neurons to zero. While the activities of the patterns produced in the former case are distributed about an average value of a, those produced in the latter case are exactly a. As reported by Nadal [29], we observed that there is a finite quantitative discrepancy between these two cases, and the present theory is applicable only to the former case.

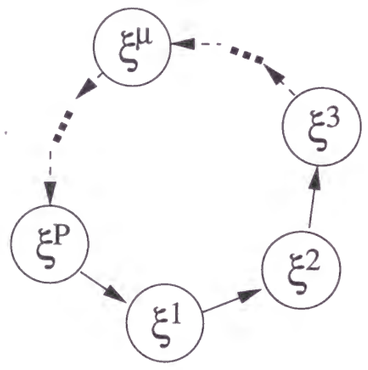
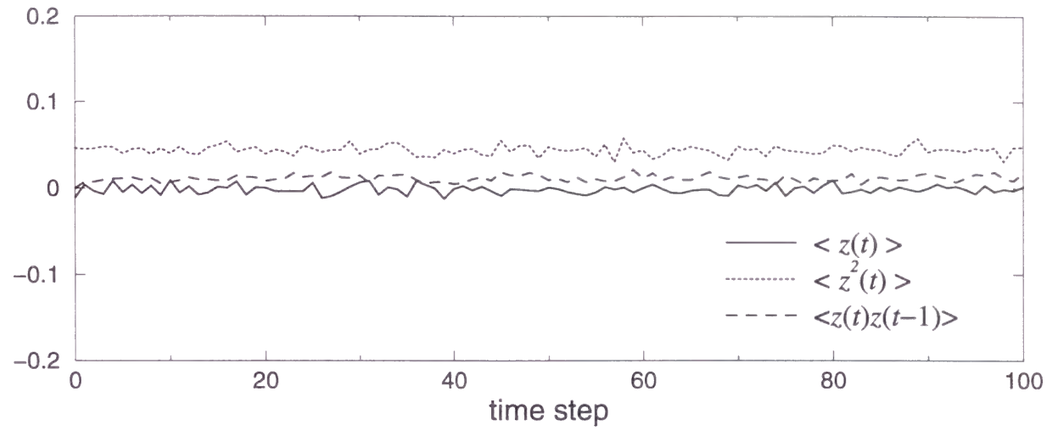


Figure 2.1: The sequential patterns treated by the present network model.



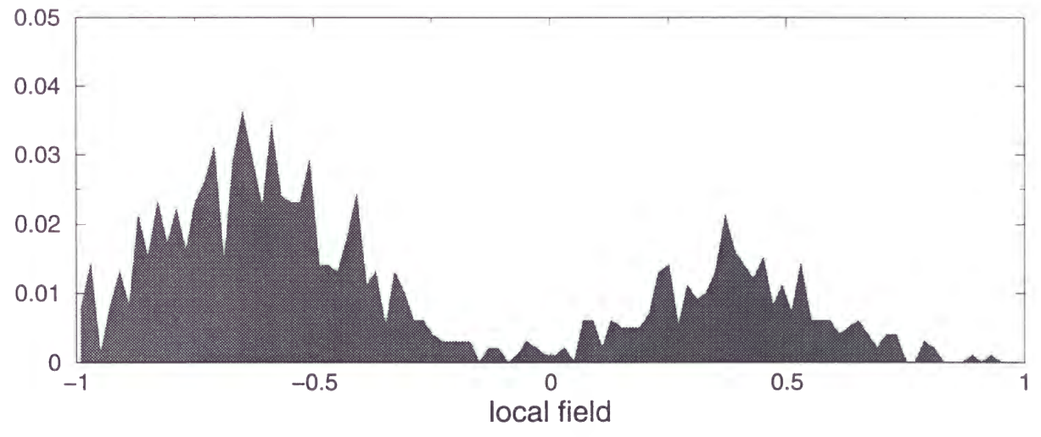
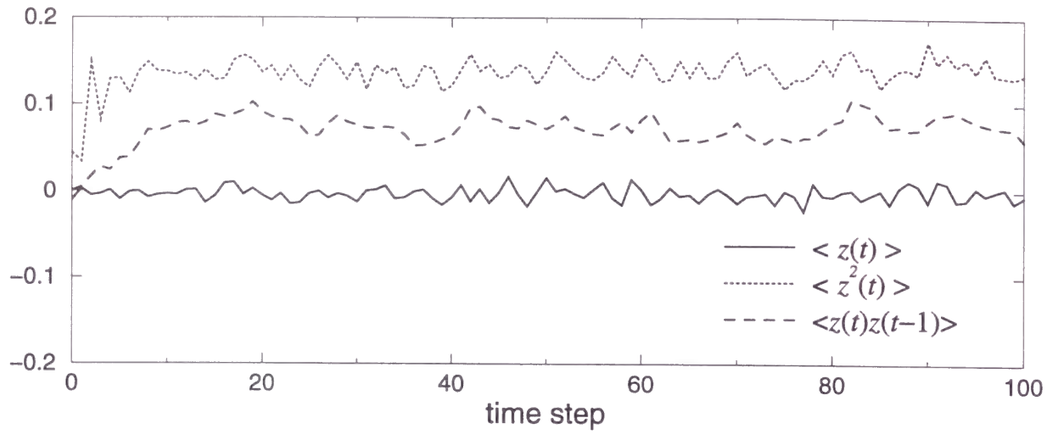


Figure 2.2: Top : Typical time evolution of statistics of crosstalk noise $z(t)$ for $a = 0.3$, $\alpha = 0.15$, $\theta = 0$, $g = 0.33$, and the initial overlap $m(0) = 0.6$. Bottom : Snap shot of distribution of local field $h(t)$.



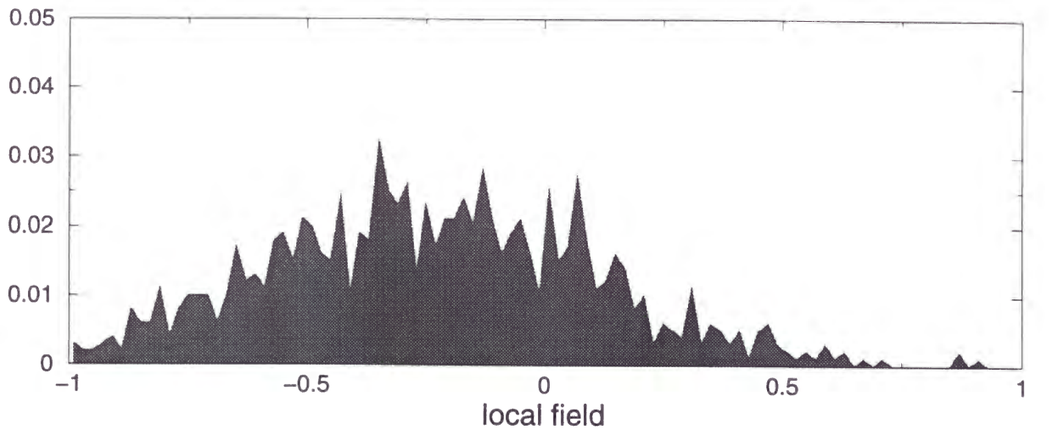


Figure 2.3: A plot similar to that in figure 2.2. The initial overlap $m(0) = 0.2$. The other parameters are the same as in figure 2.2.

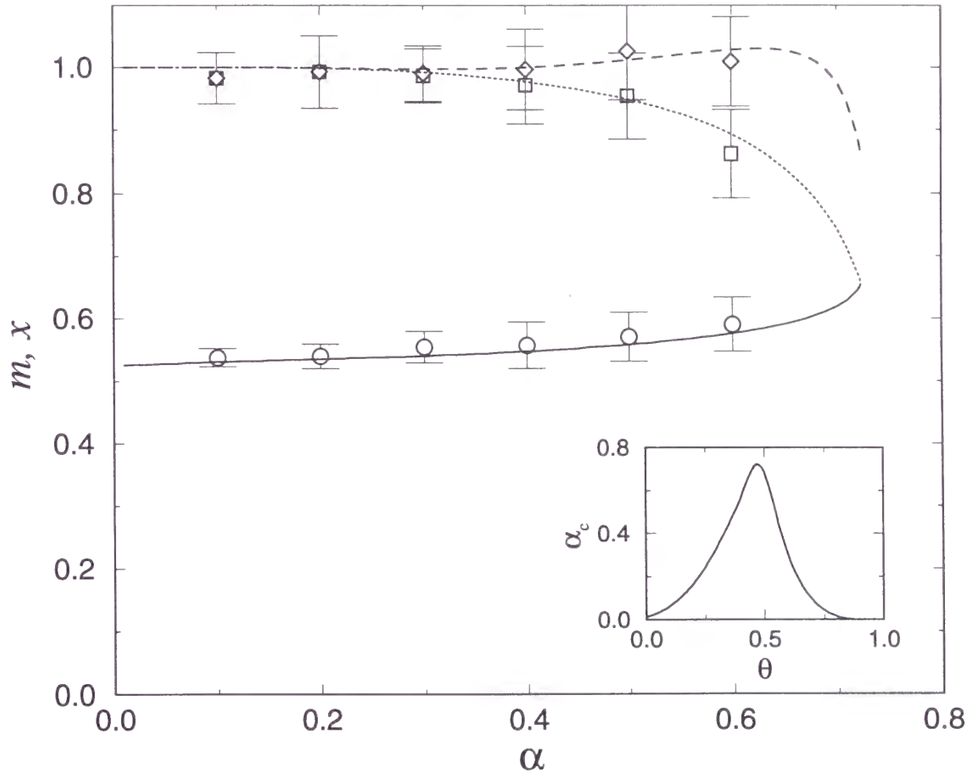


Figure 2.4: From the top, the equilibrium activity (dashed curve), equilibrium overlap (dotted curve), and basin of attraction (full curve) for $a = 0.1$ and $\theta = \theta_{opt}(= 0.47)$. The ordinate is the overlap m or the activity x, and the abscissa is the loading rate α. The data points indicate simulation results with $N = 2000$ for 20 trials. We take the initial activity as $x(0) = 1.0$. The inset shows the dependence of the storage capacity α_c on the uniform threshold θ. The value at the peak of the curve corresponds to θ_{opt}.

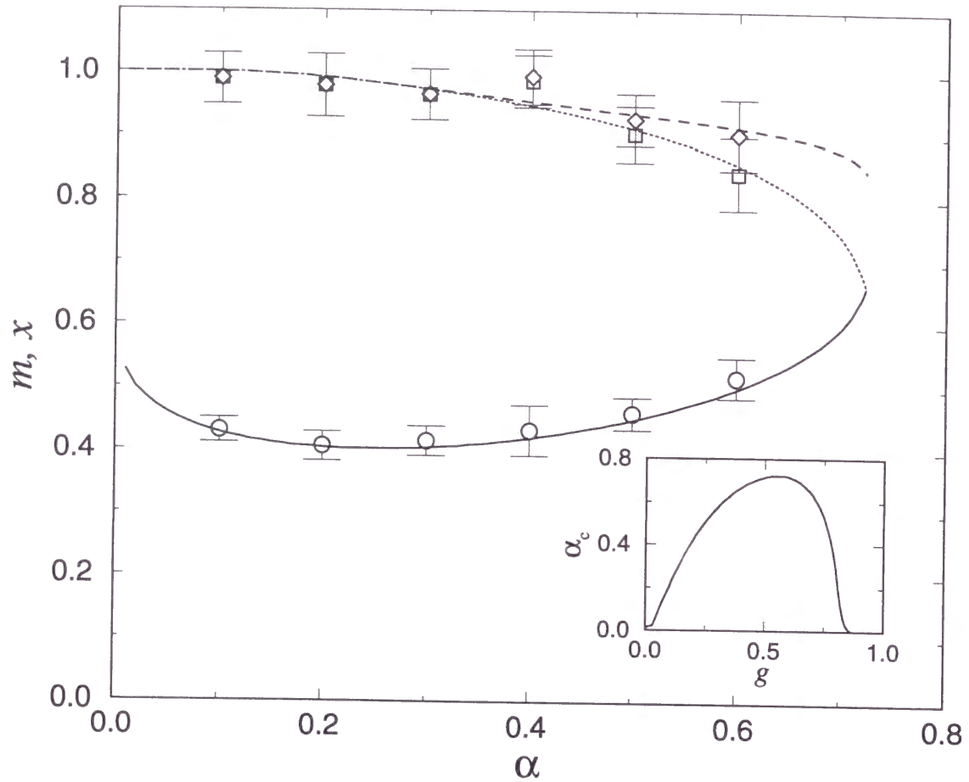


Figure 2.5: A plot similar to that in figure 2.4 for the case $\theta = 0$ and $g = g_{opt}(= 0.56)$. The other parameters are the same as in figure 2.4. The inset shows the dependence of the storage capacity α_c on the inhibitory interaction g when $\theta = 0$. The value at the peak of the curve corresponds g_{opt}.

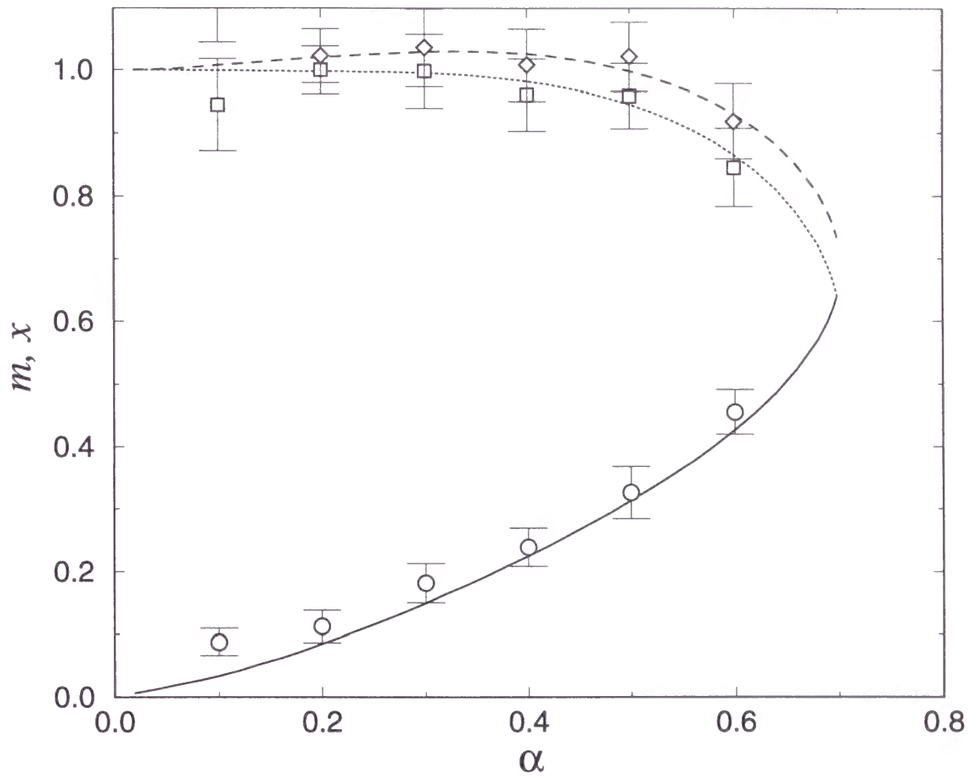


Figure 2.6: A plot similar to that in figure 2.4 for the self-control model. The other parameters are the same as in figure 2.4.

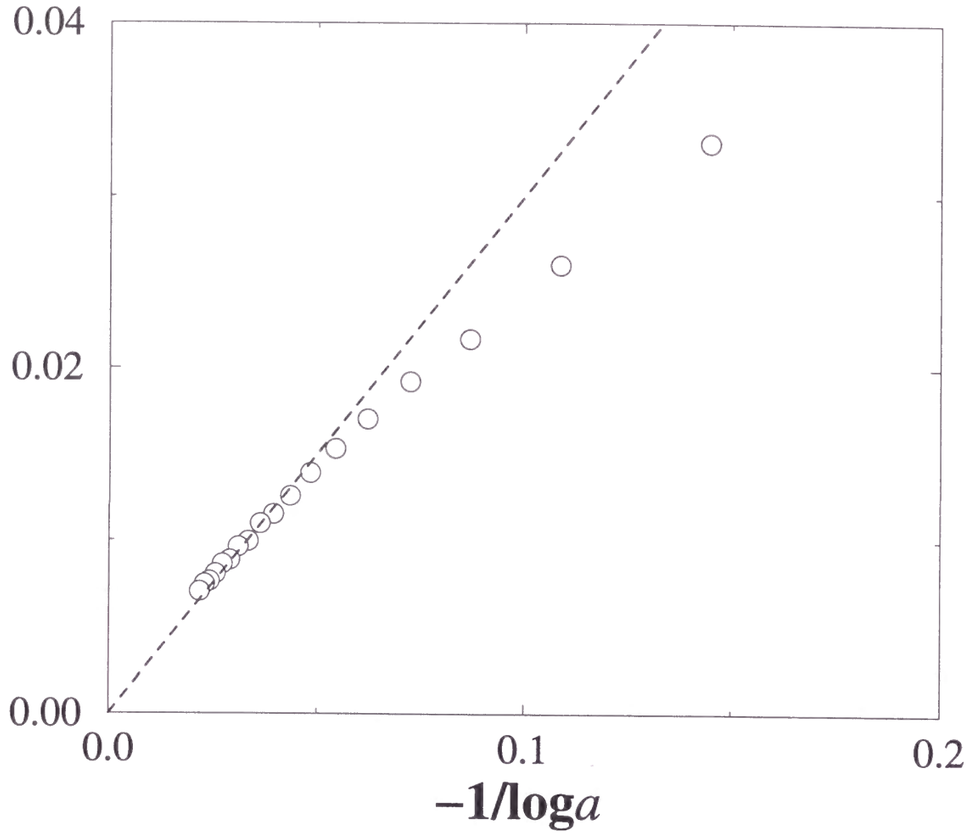


Figure 2.7: Dependence of the maximized storage capacity on the activity a. The maximized storage capacity can be obtained by optimally tuning θ.

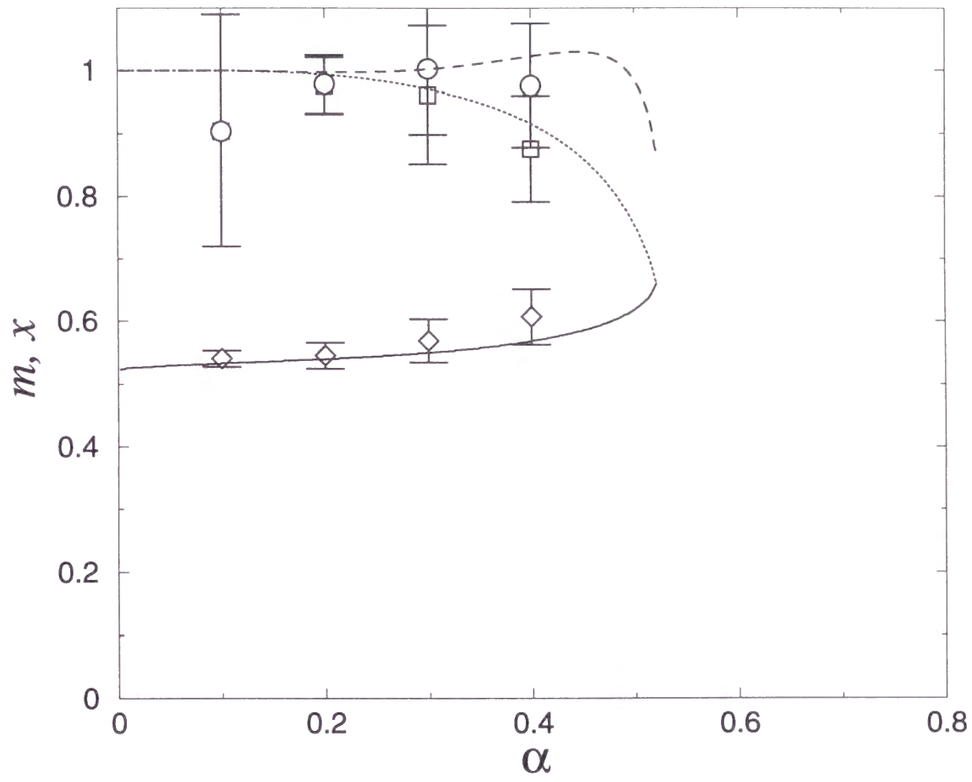


Figure 2.8: A plot similar to that in figure 2.4 for the case $c = 0.7$. The other parameters are the same as in figure 2.4.

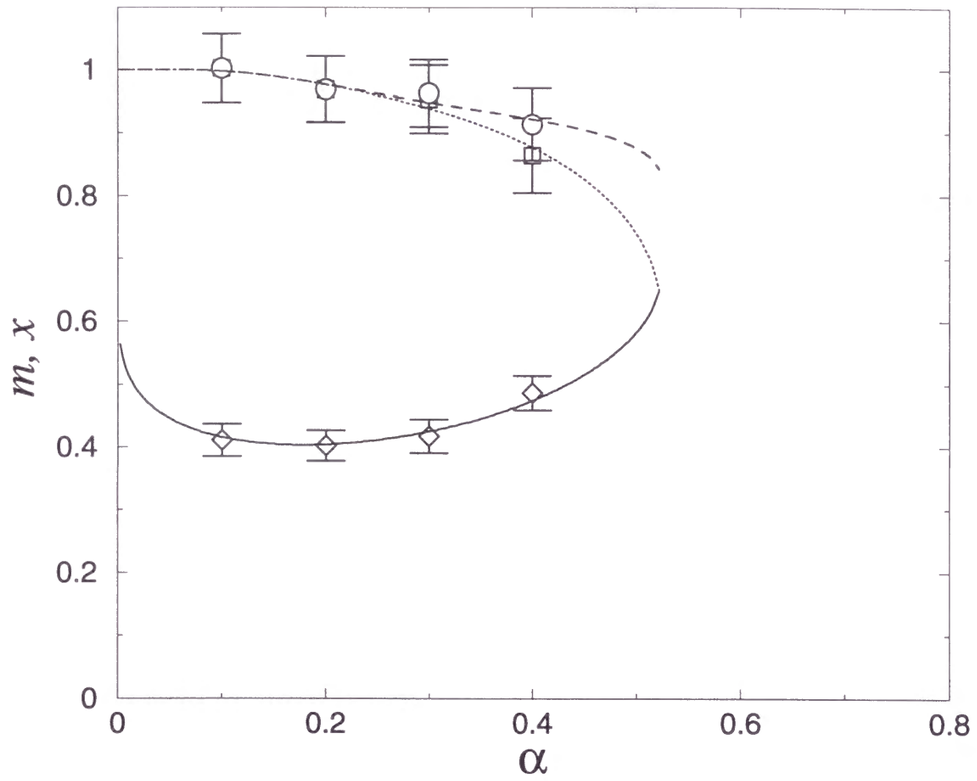


Figure 2.9: A plot similar to that in figure 2.4 for the case $c = 0.7$. The other parameters are the same as in figure 2.5.

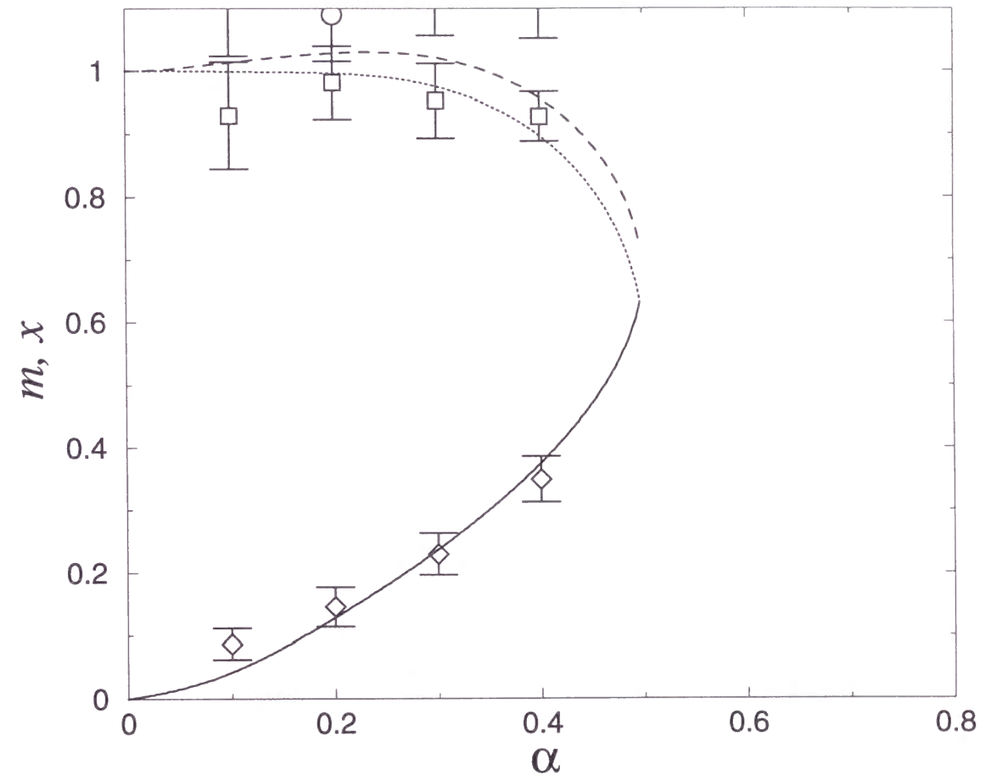


Figure 2.10: A plot similar to that in figure 2.4 for the case $c = 0.7$. The other parameters are the same as in figure 2.6.

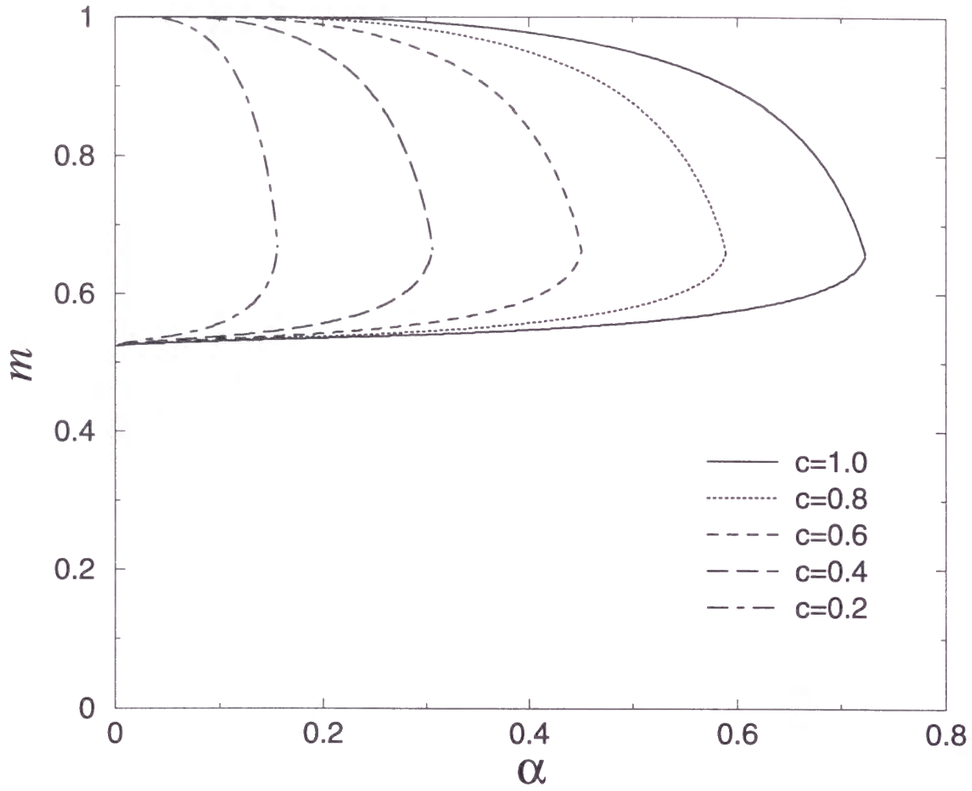


Figure 2.11: Dependence of the theoretical curve on the ratio of connected synapses c The other parameters are the same as in figure 2.4.

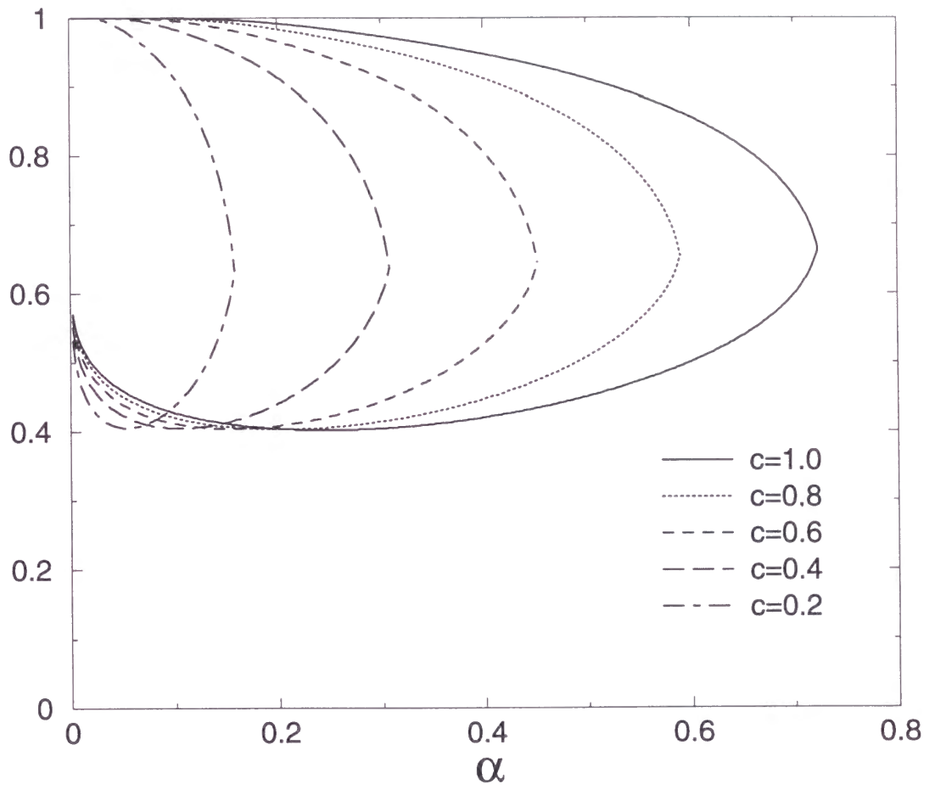


Figure 2.12: A plot similar to that in figure 2.11 The other parameters are the same as in figure 2.5.

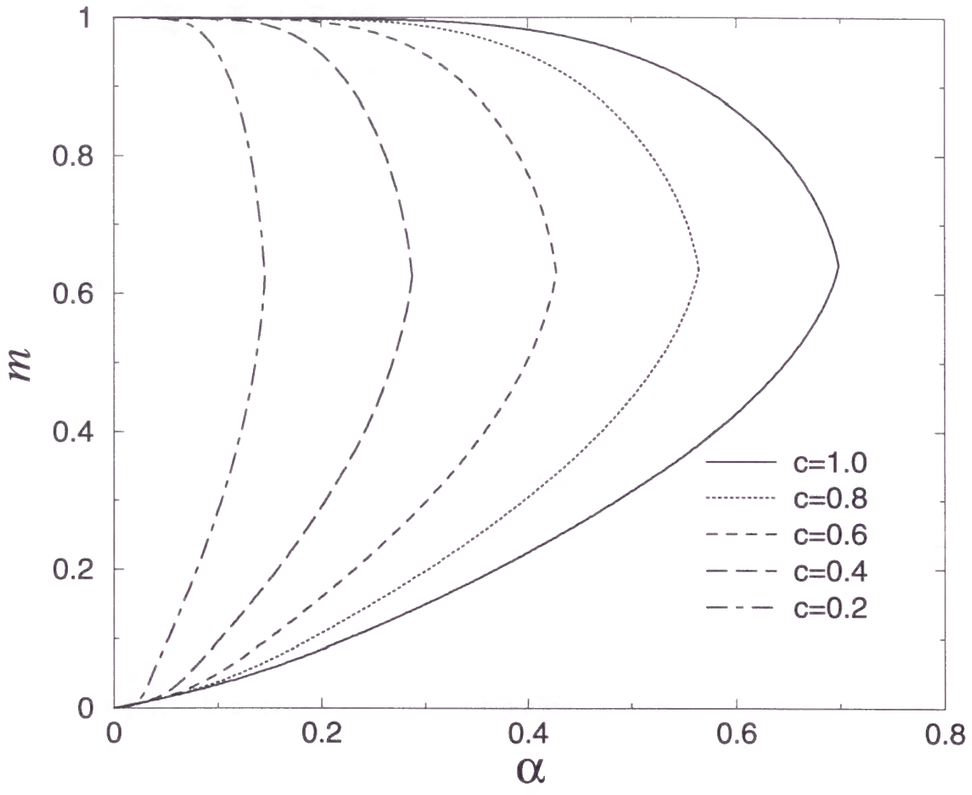


Figure 2.13: A plot similar to that in figure 2.11 The other parameters are the same as in figure 2.6.

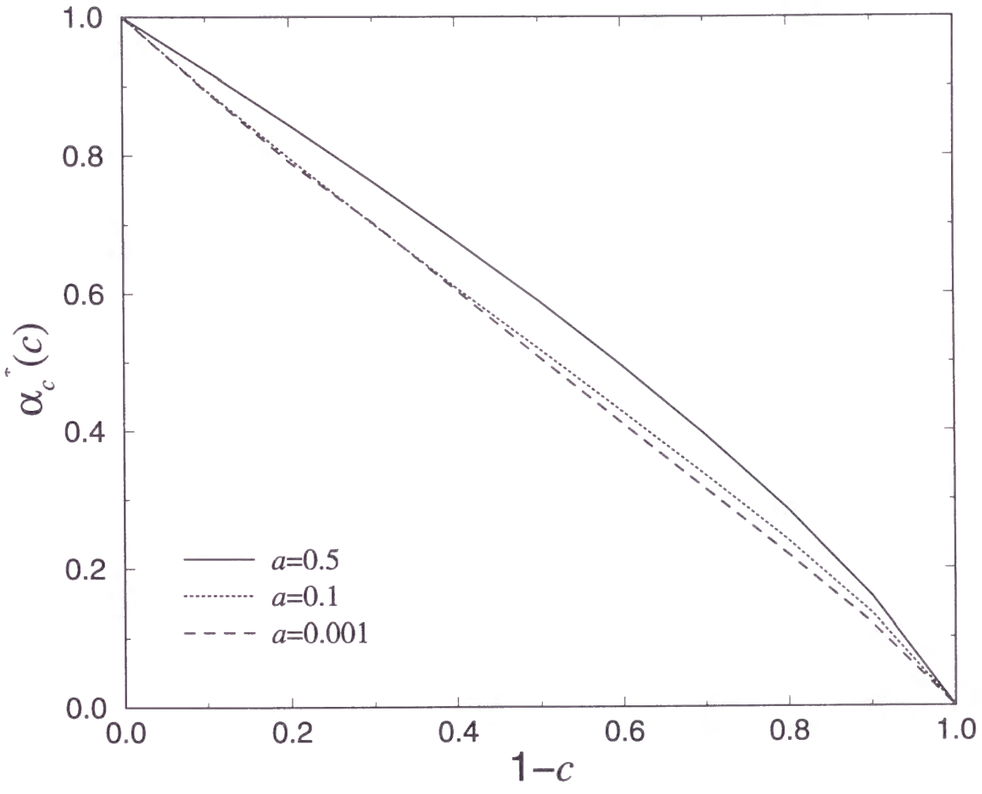


Figure 2.14: Dependence of the normalized storage capacity $\alpha_c^*(c)$ on the ratio of connected synapses c for $a = 0.5, 0.1, 0.001$.

Chapter 3

Associative abilities of oscillator neural networks

3.1 Introduction

In recent years, many attempts have been made to treat neural network models more realistic than traditional ones such as the Hopfield model [8]. Progress in the physiological understanding of real neural systems, for example, neuronal activity and morphology of synaptic connectivities, has led theoretical interests to those various models with biological validity. With this trend, many models whose purpose is to grasp the essence of more detail dynamics in a neuron have been proposed and analyzed. From the theoretical point of view, models capable of describing the continuous behavior of neuronal activities are expected to be superior in information processing. In particular, oscillator neural networks have come to be one of the most intriguing models in this context, since it was reported that collective oscillatory behavior may contribute to information processing in biological systems [30]. This is because such models are simple enough to allow for theoretical analysis, while they also contain the essence of the temporal features of neuronal activity. The results obtained through the analysis of such a simple model are sure to further our understanding not only of more complex models but of real neural systems.

Many interesting analyses concerning oscillator neural networks have been reported [32,

33, 34, 35, 36, 37, 38, 39]. As the representative study, in the case of auto-association of random phase patterns with Hebbian learning, the storage capacity $\alpha_c = 0.038$ is estimated using the replica theory [40]. However, despite these works, oscillator models remain poorly understood.

For the purpose of estimation of performance with regard to associative memory, it is necessary to consider two important aspects of association, one concerning the attractor and the other concerning the basin. However, most previous studies are restricted to properties of the attractor, such as equilibrium overlap and storage capacity. However, considering the associative ability for a noisy pattern to be dynamically corrected, in order to obtain a proper measure of a model's performance, it is necessary to study the basin of attraction also. In order to discuss the basin, we must treat the dynamics of recalling.

For traditional models, several theoretical analyses on dynamics of retrieval processes have been reported. Using the method of generating functionals and path integrals, a general theory can be formulated [16]. Although this method yields an exact description, a suitable approximation is required in order to obtain practical results [17]. In the case of parallel dynamics, the result turns out to be simple so that the exact prediction of the retrieval dynamics for the initial few steps is possible [18]. As for arbitrary finite time scale, the dynamical replica theory has been proposed recently [19]. On the other hand, as an approximation method, the statistical neurodynamics theory has been proposed [20, 21]. Although the approximation used in this approach is crude in a sense, it it practically useful to predict long term behavior when a network succeeds in retrieval [26]. Finally, we should note that, under suitable conditions, the theoretical result from the statistical neurodynamics can be obtained also by the path integral method.

Furthermore, one we cannot avoid is the problem of theoretically estimating robustness against damage of the synaptic structure. In the case of Hopfield model, this has already been studied by several authors [27, 42, 43, 44]. Oscillator neural networks can retrieve

more detailed information than can the Hopfield model, because the memorized patterns are described by continuous rather than binary variables. For this reason, one might guess that the retrieval ability of the oscillator neural network decreases faster with the ratio of the disconnected synapses than that of the Hopfield model. To clarify the above point, in this paper we wish to address the problem of how the oscillator neural network is affected by random synaptic dilution.

In the present paper, from the point of view mentioned above, we theoretically analyze and estimate the associative ability of the oscillator neural network. In the next section, we introduce an oscillator neural network model treated here. Section 3 contain a theoretical analysis of the equilibrium property in this network using the replica theory. It is investigated how the equilibrium states are affected by random synaptic dilution. In order to discuss the retrieval dynamics in the oscillator neural network, the statistical neurodynamics theory is applied to the model in Section 4. Using the derived dynamical equations describing the time development of some macroscopic parameters such as that representing overlap, we examine the shape of the basin of attraction. In Section 5, we show the relation between the replica theory and the statistical neurodynamics theory. In Section 6, we give a brief summary and conclusion.

3.2 Oscillator neural network model

Let us first consider the situation in which N periodic firing neuronal systems are coupled to each other. In general, such a system can be described by evolution equations involving a set of state variables, for instance, a membrane potential and several ionic leak currents. Under suitable conditions, it is well known that such a coupled system can be reduced to a system of simple coupled oscillators. Therefore, the state of the i-th system can be characterized by a single variable ϕ_i. This quantity ϕ_i is usually referred to as the phase, which represents the timing of the neuronal spikes. The reduced phase equations take the

general form [45]

$$\frac{d\phi_i}{dt} = \omega_i + \sum_{j=1}^{N} \Gamma_{ij}(\phi_j - \phi_i), \tag{3.1}$$

where ω_i is the frequency of the i-th neuron and Γ_{ij} represents the effect of the interaction between the i-th and j-th neurons. We should remark that the system (3.1) is invariant under uniform phase translation, $\phi_i \to \phi_i + \phi_0$, where ϕ_0 is an arbitrary real constant. It can be also shown that $\Gamma_{ij}(\phi)$ is a 2π-periodic function of ϕ. To be specific, we assume that all frequencies are equal to Ω and that $\Gamma_{ij}(\phi)$ is approximated by the lowest mode of the Fourier components. Eliminating Ω by applying the transformation $\phi_i \to \phi_i + \Omega t$, the model equations (3.1) become

$$\frac{d\phi_i}{dt} = \sum_{j=1}^{N} K_{ij} \sin(\phi_j - \phi_i + \beta_{ij}), \tag{3.2}$$

where J_{ij} and β_{ij} are parameters representing the effect of the interaction.

Using the complex representation $S_i = \exp(i\phi_i)$, we obtain the alternative form [41]

$$\begin{cases} \dfrac{dS_i}{dt} &= \dfrac{1}{2}(h_i - \tilde{h}_i S_i^2) \\ h_i &= \displaystyle\sum_{j=1}^{N} J_{ij} S_j \end{cases} \tag{3.3}$$

where $J_{ij} = K_{ij}\exp(i\beta_{ij})$ and $\tilde{h}_i$ denotes the complex conjugate of h_i. We now consider the case that the system is in a stable stationary state. Putting $dS_i/dt = 0$, it is found that such states satisfy the conditions

$$S_i = \frac{h_i}{|h_i|}. \tag{3.4}$$

Let us denote a set of patterns to be memorized as complex variables $\xi_i^\mu = \exp(i\theta_i^\mu)$ $(\mu = 1, \cdots, P)$, where P is the total number of patterns. For simplicity, we assume that the parameters θ_i^μ are chosen at random from a uniform distribution between 0 and 2π. As usual, the load parameter α is defined by $\alpha = P/N$. The overlap $M^\mu(t)$ between the state of the system and the pattern μ at time t is given by

$$\begin{aligned} M^\mu(t) &= m^\mu(t) e^{i\varphi^\mu(t)} \\ &= \frac{1}{N}\sum_{j=1}^{N} \tilde{\xi}_j^\mu S_j(t). \end{aligned} \tag{3.5}$$

In practice, the correlation of the system with the μ-th pattern at time t is measured by the amplitude component $m^\mu(t) = |M^\mu(t)|$. To realize the function of the auto-associative memory, we must define the synaptic efficacies appropriately. As a natural choice, we assume that the synaptic efficacies are given by the generalized Hebbian rule

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{P} \xi_i^\mu \tilde{\xi}_j^\mu. \tag{3.6}$$

Note that rotational symmetry is not broken by the above rule. Owing to this symmetry, all patterns generated by the uniform phase translation $\xi_i^\mu \exp(i\phi_0)$ represent the same pattern as ξ_i^μ. This nature stems from the fact that information is encoded not by the absolute time but by the relative timing of spikes. Using the above learning rule, it is expected that an initial noisy phase pattern can be corrected dynamically, as illustrated in Figure 3.1.

To discuss robustness against damage of synaptic connections, we define

$$\bar{J}_{ij} = \frac{c_{ij}}{c} J_{ij}, \tag{3.7}$$

as the formulation for randomly diluted synapses. Here, the c_{ij} are independent random variables, which assume the values 1 and 0 with probabilities c and $1-c$, respectively. Note that the dilution parameter c represents the ratio of connections. In the limit $N \to \infty$, the expression in Eq.(3.7) becomes equivalent to that of synaptic connections with static noise[27],

$$\begin{aligned} \bar{J}_{ij} &= J_{ij} + \eta_{ij} \\ &= \frac{1}{N}\sum_{\mu=1}^{P} \xi_i^\mu \tilde{\xi}_j^\mu + \eta_{ij} \end{aligned} \tag{3.8}$$

The synaptic noise η_{ij} is a complex Gaussian noise with mean 0 and variance η^2/N. It is easy to determine the relation between the dilution ratio c and the variance parameter

η^2 as

$$\eta^2 = \frac{1-c}{c}\alpha. \tag{3.9}$$

For the sake of simplicity in later theoretical analysis, we adopt the expression of Eq.(3.8).

In this model, the outputs of all neurons are continuously and simultaneously changing according to the equations (3.3). On the other hand, the various dynamical theories in the case of traditional neural networks have been greatly advanced in the last decade. However, when we apply such theories to oscillator neural networks and attempt an analytical treatment of the dynamics, we encounter difficulties owing to the updating rule. Therefore, to make the dynamics (3.3) more mathematically tractable, we discretize time and assume the synchronous updating rule. Furthermore, considering the fact that all neurons relax toward the state in which the relation (3.4) is satisfied, it is natural to adopt the following dynamics:

$$\begin{cases} S_i(t+1) &= \dfrac{h_i(t)}{|h_i(t)|} \\ h_i(t) &= \displaystyle\sum_{j=1}^{N} J_{ij} S_j(t) \end{cases} \tag{3.10}$$

Here, the quantity $X(t)$ represents the value of the X at time t. Throughout this paper, we treat the above model analytically. Since these dynamics are not equivalent to those of (3.2), the behavior they describe in general different. In fact, it is found that, for certain realizations of (3.2), the structure of phase space suffers a qualitative change under the simplification to (3.10). Fortunately, however, we have found that the solutions of (3.10) corresponding to such cases do not appear in the simulation of this equation under the conditions we presently consider. This point will be discussed again in the last section. We should also remark that if we use *asynchronous* updating the equilibrium states of (3.10) are equivalent to those of the original phase model. This can be easily shown by considering the noiseless limit in a statistical mechanics treatment. In this sense, (3.10) can be thought of as a synchronous update version of the oscillator neural network with

discrete time. In traditional neural networks, the analytical approach for synchronous updating has contributed to the understanding of retrieval dynamics in the more general cases. Therefore, we believe that this model serves as a convenient starting point for the theoretical study of retrieval dynamics in general oscillatory systems. In fact, we will find that the basins of attraction in the two models are qualitatively very similar.

In the next section, the equilibrium states for the oscillator neural network model (3.2), (3.8) are investigated. The storage capacity and the critical overlap is the properties of the equilibrium states. In the following section, we theoretically analyze the retrieval process of the present model. To analyze such a process, because of the reason mentioned above, we treat the model (3.10), (3.8) in practice and mainly discuss the time development of the overlap (3.5) along with certain other macroscopic parameters.

3.3 Theoretical analysis for equilibrium states

3.3.1 Replica theory

If the synaptic noise introduced in Eq. (3.8) is Hermitian, that is, $\eta_{ij} = \tilde{\eta}_{ji}$, then the synaptic matrix $\bar{J}_{ij}$ is also found to be Hermitian. In this case, the Hamiltonian, the energy function, can be defined for the present network model. Introducing an additional noise term to the present model in order to apply statistical mechanics for the analysis of the equilibrium states,

$$\begin{cases} \dfrac{d\phi_i}{dt} &= -\dfrac{\partial H}{\partial \phi_i} + \gamma_i(t) \\ H &= -\dfrac{1}{2}\displaystyle\sum_{i \neq j} \bar{J}_{ij} \tilde{S}_i S_j \end{cases} \tag{3.11}$$

where the last term $\gamma_i(t)$ is Gaussian white noise characterized by $\langle \gamma_i(t) \rangle = 0$ and $\langle \gamma_i(t)\gamma_j(t') \rangle = 2T\delta_{ij}\delta(t-t')$. The temperature T $(= \beta^{-1})$ gives a measure of the level of the stochastic noise in the dynamics. The introduction of this noise enables us to perform the standard mean-field analysis in terms of statistical mechanics. Therefore, the asymptotic behavior of the network at finite T is governed by the free-energy, and

the equilibrium probability density is given by the Gibbs distribution, $e^{-H/T}$. Using the synaptic matrix (3.8), the Hamiltonian H takes the form

$$\begin{aligned} H = & -\frac{1}{2N}\sum_{\mu}^{P}\left[\left(\sum_i \cos(\phi_i-\theta_i^\mu)\right)^2+\left(\sum_i \sin(\phi_i-\theta_i^\mu)\right)^2\right] \\ & +\frac{P}{2}-\sum_{i<j}\eta_{ij}^{Re}\cos(\phi_j-\phi_i)+\sum_{i<j}\eta_{ij}^{Im}\sin(\phi_j-\phi_i), \end{aligned} \tag{3.12}$$

where η_{ij}^{Re} and η_{ij}^{Im} are the real and the imaginary part of η_{ij}, respectively.

Let us consider the situation in which the network retrieved the pattern ξ^1, namely,

$$m_1 = m \sim O(1), \qquad m^\mu \sim O(\frac{1}{\sqrt{N}}) \quad (\mu > 1). \tag{3.13}$$

The overlap m^μ is defined as Eq. (3.5). Then, let us define the parameter α by $\alpha = P/N$. To proceed, we must first perform the quenched averaging of the free energy over the randomness. Using the replica method, the averaged free energy per neuron is computed from

$$\begin{aligned} f &\equiv \lim_{N\to\infty} -\frac{1}{N\beta}\langle\langle \ln Z\rangle\rangle \\ &= \lim_{N\to\infty}\lim_{n\to 0} -\frac{1}{N\beta n}\ln\langle\langle Z^n\rangle\rangle, \end{aligned} \tag{3.14}$$

where $\langle\langle\cdots\rangle\rangle$ indicates a quenched average over the patterns ξ_i^μ as well as over the synaptic noise η_{ij}. The partition function Z is defined by $Z = \mathrm{Tr}_{\{\phi_i\}} e^{-\beta H(\{\phi_i\})}$.

In the replica symmetry approximation, we find that the averaged free energy per neuron is given by

$$\begin{aligned} f = & \frac{1}{2}m^2+\frac{\alpha}{2}+\frac{\beta\alpha r}{8}(1-q) \\ & +\frac{\alpha}{\beta}\left\{\ln(1-\frac{1}{2}\beta(1-q))-\frac{\beta\eta^2}{8}(1-q)^2-\frac{\frac{\beta}{2}q}{1-\frac{\beta}{2}(1-q)}\right\} \\ & -\frac{1}{\beta}\left\langle\!\!\left\langle\int\int\frac{dz_1dz_2}{2\pi}\exp(-\frac{z_1^2+z_2^2}{2})\ln\int_0^{2\pi}d\phi\exp\beta\left[A\cos\phi+B\sin\phi\right]\right\rangle\!\!\right\rangle_\theta, \end{aligned} \tag{3.15}$$

with

$$\begin{aligned} A &= \frac{\sqrt{\alpha r+2\eta^2 q}}{2}z_1+m\cos\theta \\ B &= \frac{\sqrt{\alpha r+2\eta^2 q}}{2}z_2+m\sin\theta. \end{aligned} \tag{3.16}$$

The double angular brackets $\langle\langle\cdots\rangle\rangle_\theta$ denote an average over θ with the same distribution as θ_i^μ. The network is then characterized by order parameters m, r and q. The parameter m represents the overlap with the retrieved pattern ξ_i^1. The parameters q and r correspond to the Edwards-Anderson order parameter and the mean-square random overlap for the unretrieved patterns, respectively. The saddle-point equations for these order parameters are

$$\begin{cases} m = & \left\langle\!\!\left\langle\int\int\frac{dz_1dz_2}{2\pi}\exp(-\frac{z_1^2+z_2^2}{2})\frac{I_1(\beta\sqrt{A^2+B^2})}{I_0(\beta\sqrt{A^2+B^2})}\frac{A\cos\theta+B\sin\theta}{\sqrt{A^2+B^2}}\right\rangle\!\!\right\rangle_\theta \\ \beta(1-q) = & \frac{2}{\sqrt{\alpha r+2\eta q}}\left\langle\!\!\left\langle\int\int\frac{dz_1dz_2}{2\pi}\exp(-\frac{z_1^2+z_2^2}{2})\frac{I_1(\beta\sqrt{A^2+B^2})}{I_0(\beta\sqrt{A^2+B^2})}Az_1+Bz_2\sqrt{A^2+B^2}\right\rangle\!\!\right\rangle_\theta, \\ r = & \frac{2q}{(1-\frac{\beta}{2}+\frac{\beta}{2}q)^2} \end{cases} \tag{3.17}$$

where $I_k(z)$ is the k-th order modified Bessel function, defined by

$$I_k(z) = \frac{1}{2\pi}\int_0^{2\pi}d\phi e^{z\cos\phi}\cos k\phi \tag{3.18}$$

We are now ready to discuss the storage capacity α_c in the case of the random diluted synapses. In the limit of zero noise, $\beta^{-1}\to 0$, q tends to 1, and Eqs.(3.17) reduce to the equations

$$\begin{cases} m = & f_1\left(\frac{m}{\sqrt{\alpha r+2\eta^2}}\right) \\ r = & 2\left\{1-\frac{1}{\sqrt{\alpha r+2\eta^2}}f_2\left(\frac{m}{\sqrt{\alpha r+2\eta^2}}\right)\right\}^{-2} \end{cases}, \tag{3.19}$$

where f_1 and f_2 are defined by

$$\begin{aligned} f_1(y) &= \int_0^{2\pi}d\varphi\int_0^\infty dR\frac{(R\sin\varphi+2y)Re^{-\frac{R^2}{2}}}{2\pi(R^2+4Ry\sin\varphi+4y^2)^{\frac{1}{2}}} \\ f_2(y) &= \int_0^{2\pi}d\varphi\int_0^\infty dR\frac{(R^2+2Ry\sin\varphi)Re^{-\frac{R^2}{2}}}{2\pi(R^2+4Ry\sin\varphi+4y^2)^{\frac{1}{2}}} \end{aligned}. \tag{3.20}$$

Note that the relationship between the dilution parameter c and the mean square deviation of the synaptic noise η is given by Eq.(3.9). These equations always have a trivial solution $m = 0$, which corresponds to a spin glass state ($q \neq 0$). For $\alpha < \alpha_c$, there also exists

a solution, for which $m \neq 0$ corresponding to a retrieval state. This retrieval solution disappears discontinuously at α_c, where the overlap m jumps from the finite value m_c to zero, except for the case $c = 0$.

3.3.2 The results

In Figure 3.2, the storage capacity obtained from numerical solutions of Eqs.(3.19) is plotted as a function of the ratio of the disconnected synapses.

In the case of the fully connected network, that is, $c = 1(\eta = 0)$, we obtained that $\alpha_c = 0.038$ and $m_c = 0.90$. This result is essentially identical to those of the Q-state clock model in the limit $Q \to \infty$ estimated by Cook.

In general, it is expected that α_c falls monotonically from 0.038 to zero as η increases. In fact, analytical results show that the retrieval solution exists only in the case that $\eta < \eta_c$, $\eta_c = \sqrt{\pi}/2 \approx 0.886$, and α_c is a monotonically decreasing function of η.

On the other hand, in case of the Hopfield model, Sompolinsky has estimated that $\eta_c = \sqrt{2/\pi} \approx 0.797$. Using Eq.(3.9), we finally obtained a theoretical curve, as shown in Figure 3.2. We also carried out numerical simulations in which each value of α_c was averaged over 20 trials with $N = 1500$. As is clear from Figure 3.2, the simulation results are in reasonable agreement with analytical results.

Figure 3.3 shows the overlap of the retrieval state at α_c, where the solid and the dashed curves correspond to the oscillator model and the Hopfield model, respectively. In either case, the critical overlap is affected little by the synaptic dilution as long as $1-c$ is smaller than 0.8. Particularly in the oscillator neural network, the critical overlap m_c remains almost constant in the range 0 to 0.8.

As mentioned above, η_c is slightly larger in the present model than in the Hopfield model. This implies that our system is more robust against synaptic dilution than is the Hopfield model. Let us attempt to clarify this point quantitatively. Although α_c in

the oscillator network is generally smaller than in the Hopfield model[10], the oscillator network is able to retrieve phase patterns represented by continuous variables, not simple binary ones. Thus taking account of the information content in the retrieved patterns, it makes no sense to compare the storage capacities α_c of the two models. However, it is meaningful to estimate and compare how the random dilution of synapses in each model reduces its performance from the level without dilution ($c = 1$). For this purpose, we define the normalized maximum storage capacity as $\alpha_c^*(c) = \alpha_c(c)/\alpha_c(1)$, where $\alpha_c(1)$ is the maximum storage capacity at $c = 1$. Thus, $\alpha_c^0 = 0.038$ in the oscillator model, and $\alpha_c^0 = 0.138$ in the Hopfield model. The dependence of the normalized storage capacities $\alpha_c^*(c)$ on $1 - c$ is shown in Figure 3.4. It is obvious from this figure that the normalized capacity $\alpha_c^*(c)$ of the oscillator network is always larger than that of the Hopfield network. Nevertheless, for $c < 0.8$, the qualities of the retrieval patterns obtained in either model are largely independent of c, as seen in Figure 3.3. Therefore, we can conclude that the oscillator network is totally more robust against dilution than the Hopfield model.

3.4 Theoretical analysis for retrieval process

3.4.1 Statistical neurodynamics theory

Let us consider the situation in which the network is recalling the pattern ξ_i^1, namely,

$$m^1(t) \sim O(1), \qquad m^\mu(t) \sim O\left(\frac{1}{\sqrt{N}}\right) \quad (\mu \neq 1). \tag{3.21}$$

The internal potential $h_i(t)$ in Eq.(3.10) can be separated as

$$\sum_{j \neq i}^{N} \bar{J}_{ij} S_j(t) = \xi_i M(t) + \frac{1}{N} \sum_{j \neq i}^{N} \sum_{\mu=2}^{p} \xi_i^\mu \tilde{\xi}_j^\mu S_j(t) + \sum_{j \neq i}^{N} \eta_{ij} S_j(t), \tag{3.22}$$

where $\xi_i = \xi_i^1$ and $M(t) = M^1(t)$. From this point, for simplicity, we drop the index μ in the case of pattern 1. In this process, the first term on the r.h.s. of Eq.(3.22) is regarded as the signal to induce recollection of the target pattern ξ_i^1, while the remaining terms are

regarded as noise. For convenience, we define the noise terms $z_i(t)$ as

$$\begin{aligned} z_i(t) &= z_i^c(t) + z_i^s(t) \\ &= \frac{1}{N}\sum_{j\neq i}^{N}\sum_{\mu=2}^{p}\xi_i^\mu\tilde{\xi}_j^\mu S_j(t) + \sum_{j\neq i}^{N}\eta_{ij}S_j(t). \end{aligned} \tag{3.23}$$

In $z_i(t)$, $z_i^c(t)$ is the crosstalk noise from unretrieved patterns ($\mu \neq 1$), and $z_i^s(t)$ is caused by noise in the synapses. The essence of the theory is to treat the crosstalk noise $z_i^c(t)$ as complex Gaussian noise with mean 0 and variance $\sigma_c(t)^2$. It has been confirmed numerically that this assumption is valid as long as the network succeeds in retrieval [26]. In addition, the synaptic noise $z_i^s(t)$ is also assumed to be complex Gaussian with mean 0 and variance $\sigma_s(t)^2$ [27]. Therefore, $z_i(t)$ displays a complex Gaussian distribution with mean 0 and variance $2\sigma(t)^2 = \sigma_c(t)^2 + \sigma_s(t)^2$. Here, we also assume $z_i^c(t)$ and $z_i^s(t)$ to be independent. We note that $z_i(t)$ can be expressed with two independent Gaussian variables $x_i(t)$ and $y_i(t)$ satisfying

$$z_i(t) = x_i(t) + iy_i(t), \qquad x_i(t), y_i(t) \sim N(0, \sigma(t)^2), \qquad \langle x_i(t)y_i(t)\rangle = 0. \tag{3.24}$$

Now we derive a dynamical equation for the overlap with the recalled pattern. The definition of overlap (3.5) leads to the equation,

$$m(t+1)e^{i\varphi(t+1)} = \frac{1}{N}\sum_{j}^{N}\tilde{\xi}_j\frac{\xi_j m(t)e^{i\varphi(t)} + z_j(t)}{|\xi_j m(t)e^{i\varphi(t)} + z_j(t)|}. \tag{3.25}$$

The variable $z_j(t)$ represents Gaussian noise. Then, because of the symmetry of its distribution, we assume $z_j(t)$ produces no effect to change $\varphi(t)$. This assumption has been confirmed numerically. Using this assumption, i.e. setting $\varphi(t) = \varphi_0$, we obtain

$$\begin{aligned} m(t+1) &= \frac{1}{N}\sum_{j=1}^{N}\frac{m(t) + z_j(t)e^{i(\varphi_0+\theta_j)}}{|m(t) + z_j(t)e^{i(\varphi_0+\theta_j)}|} \\ &= \frac{1}{N}\sum_{j=1}^{N}\frac{m(t) + z_j(t)}{|m(t) + z_j(t)|}. \end{aligned} \tag{3.26}$$

Here, we use the fact that the distribution of $z_j(t)e^{i(\varphi_0+\theta_j)}$ can be obtained by simply rotating that of $z_j(t)$.

Next, we examine the time development of the variance $\langle|z_i(t)|^2\rangle = 2\sigma(t)^2$. First, we consider the synaptic noise $z_i^s(t+1) = \sum_{j\neq i}^{N}\eta_{ij}S_j(t+1)$. When we take the statistics of $z_i^s(t+1)$, we must take into account correlations between η_{ij} and η_{ji} in $S_j(t+1)$. Here, expanding $S_j(t+1) = h_j(t)/|h_j(t)|$ in terms of η_{ji} yields

$$z_i^s(t+1) = \sum_{j\neq i}^{N}\eta_{ij}\frac{h_j^0(t)}{|h_j^0(t)|} + S_i(t)\sum_{j\neq i}^{N}\frac{\eta_{ij}\eta_{ji}}{2|h_j^0(t)|}, \tag{3.27}$$

where $h_j^0(t)$ is assumed to be independent of η_{ji}. If the dilution is asymmetric, $\eta_{ji} \neq \tilde{\eta}_{ij}$ (or $c_{ij} \neq c_{ji}$), the second term vanishes. Even if it were symmetric, the assumption that the mean of the noise is 0 would lead us to neglect the second term proportional to $S_i(t)$, since it is related to the mean of $z_i^s(t+1)$. As a result, we obtain

$$\sigma_s(t+1)^2 = \eta^2. \tag{3.28}$$

Second, consider the crosstalk noise $z_i^c(t+1)$. We express $z_i^c(t+1)$ as

$$z_i^c(t+1) = \frac{1}{N}\sum_{j\neq i}^{N}\sum_{\mu=2}^{p}\xi_i^\mu\tilde{\xi}_j^\mu\frac{h_j(t)}{|h_j(t)|}. \tag{3.29}$$

When summing over μ, as in the case of (3.27), we must consider the dependence of $h_j(t)$ on ξ_j^μ. In the local field $h_j(t)$, the term

$$\frac{1}{N}\sum_{k\neq j}^{N}\xi_j^\mu\tilde{\xi}_k^\mu S_k(t) \sim \xi_j^\mu M^\mu(t), \tag{3.30}$$

which is caused by the non-target pattern μ, is estimated to be $O(1/\sqrt{N})$. Using this estimation, we expand the complex function $h_j(t)/|h_j(t)|$, obtaining

$$\begin{aligned} z_i^c(t+1) &= \frac{1}{N}\sum_{j\neq i}^{N}\sum_{\mu=2}^{p}\xi_i^\mu\tilde{\xi}_j^\mu\frac{h_j^\mu(t)}{|h_j^\mu(t)|} + \frac{1}{N}\sum_{j\neq i}^{N}\frac{1}{2|h_j^\mu(t)|}\cdot\frac{1}{N}\sum_{k\neq i}^{N}\sum_{\mu=2}^{p}\xi_i^\mu\tilde{\xi}_k^\mu S_k(t) \\ &\quad + O\left(\frac{1}{\sqrt{N}}\right), \end{aligned} \tag{3.31}$$

where $h_j^\mu(t) = \xi_j M(t) + \frac{1}{N}\sum_{k\neq j}^{N}\sum_{\nu\neq 1,\mu}^{p}\xi_j^\nu\tilde{\xi}_k^\nu S_k(t) + z_s(t)$ is assumed to be independent of ξ_j^μ. Accordingly, we find

$$z_i^c(t+1) = \frac{1}{N}\sum_{j\neq i}^{N}\sum_{\mu=2}^{p}\xi_i^\mu\tilde{\xi}_j^\mu S_j(t+1) + U(t)z_i^c(t), \tag{3.32}$$

and

$$U(t) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{2|\xi_j M(t) + z_j(t)|}. \tag{3.33}$$

where we have used the fact $h_j^\mu(t) \to h_j(t)$ in the limit $N \to \infty$ in Eq.(3.32). Squaring Eq.(3.32) and averaging in order to obtain $\sigma_c(t+1)$, we obtain

$$\sigma_c(t+1)^2 = \alpha + U(t)^2\sigma_c(t)^2 + 2U(t)\mathrm{Re}\left\langle \frac{1}{N}\sum_{j\neq i}^{N}\sum_{\mu=2}^{p}\xi_i^\mu\tilde{\xi}_j^\mu S_j(t+1)\tilde{z}_i^c(t)\right\rangle, \tag{3.34}$$

where $\alpha = p/N$.

We can calculate the last term in Eq.(3.34) by means of substituting Eq.(3.32) into Eq.(3.34) iteratively. Then, we need the following quantities:

$$\begin{aligned} X(t+1,t+1-\tau) &= \mathrm{Re}\left[\frac{1}{N}\sum_j S_j(t+1)\tilde{S}_j(t+1-\tau)\right] \\ &= \mathrm{Re}\left[\frac{1}{N}\sum_j \frac{\xi_j M(t)+z_j(t)}{|\xi_j M(t)+z_j(t)|}\cdot\frac{\tilde{\xi}_j\tilde{M}(t-\tau)+\tilde{z}_j(t-\tau)}{|\tilde{\xi}_j\tilde{M}(t-\tau)+\tilde{z}_j(t-\tau)|}\right]. \end{aligned} \tag{3.35}$$

To carry out the average in the above equation, we must generally take account of the correlation $2C(t,t-\tau) = \langle z(t)\tilde{z}(t-\tau)\rangle$. The estimation so obtained up to the n-th preceding time step is called the n-th order approximation [21]. In the n-th order approximation, we assume that the noise at each time, $z(t-1),\cdots,z(t-n+1)$, is correlated to $z(t)$, while $z(t-n)$ is independent of $z(t)$. Using Eq.(3.32) as eq.(3.34) was used above, we can obtain equations for $C(t,t-\tau)$.

Finally, the macro-dynamical equations for the n-th order approximation are given as follow:

$$m(t+1) = \left\langle \frac{m(t)+z(t)}{|m(t)+z(t)|}\right\rangle_{z(t)} \tag{3.36}$$

$$U(t) = \left\langle \frac{1}{2|m(t)+z(t)|}\right\rangle_{z(t)} \tag{3.37}$$

$$\begin{aligned} \sigma_c(t+1)^2 &= \alpha + U(t)^2\sigma_c(t)^2 \\ &\quad +2\alpha\sum_{\tau=1}^{n}X(t+1,t+1-\tau)\prod_{k=1}^{\tau}U(t+1-k) \end{aligned} \tag{3.38}$$

$$2\sigma(t+1)^2 = \sigma_c(t+1)^2 + \eta^2 \tag{3.39}$$

where

$$X(t+1,t+1-\tau) = \mathrm{Re}\left\langle \frac{m(t)+z(t)}{|m(t)+z(t)|}\cdot\frac{m(t-\tau)+z(t-\tau)}{|m(t-\tau)+z(t-\tau)|}\right\rangle_{z(t),z(t-\tau)}, \tag{3.40}$$

and

$$2C(t,t-\tau) = \begin{cases} \alpha X(t,t-\tau) + 2U(t-1)U(t-\tau-1)C(t-1,t-\tau-1) \\ \quad +\alpha\sum_{k=\tau+1}^{n-1}X(t,t-k)\prod_{\kappa=\tau+1}^{k}U(t-\kappa) \\ \quad +\alpha\sum_{k=1}^{n-1}X(t-k,t-\tau)\prod_{\kappa=1}^{k}U(t-\kappa)+\eta^2X(t,t-\tau) & (1\le\tau\le n-2) \\ \alpha X(t,t-\tau)+2U(t-1)C(t-1,t-\tau)+\eta^2X(t,t-\tau) & (\tau=n-1) \\ 0 \quad (\tau=n) \end{cases} \tag{3.41}$$

We have assumed here that the site average $\frac{1}{N}\sum_j\cdots$ does not depend on the memorized pattern and, for a given $m(t)$, $\sigma(t)$, does not depend on the initial pattern. In this case, this average is identical to $\langle\cdots\rangle_{z(t)}$, where $z(t)$ represents an arbitrary $z_i(t)$. Here, we should note that $U(t)$ is given by $m(t),\sigma(t)$ and $X(t+1,t+1-\tau)$ by $m(t),m(t-\tau),\sigma(t),\sigma(t-\tau)$. When calculating these coupled equations (3.36), (3.37), (3.38), (3.39), (3.40) and (3.41), it is necessary to give initial conditions $m(0)$, $\sigma(0)^2 = \alpha/c$ and $X(t,0)$. We use $X(t,0) = m(t)m(0)$ as an approximation.

3.4.2 The results

First, we compare the time evolution of the overlap predicted by our theory with numerical simulations for some choices of the initial overlap in Fig. 3.5. Figure 3.5 indicates that the predictions from the higher order approximation agree better with numerical results. We generally find two phases of the system behavior, depending on the storage level α and the initial overlap $m(0)$. One is a retrieval phase, in which $m(t) \to m(\infty) \sim 1$ for large t. The other is a non-retrieval phase, in which $m(t) \to 0$. When for a given α the retrieval phase exists, the network goes to the state$m(\infty) \sim 1$ provided that the initial overlap is larger than a certain critical value. The width of the basin to retrieve the pattern can be measured by this critical value.

Fig.3.6 indicates results of theoretical analysis and numerical simulation in the case of $\eta = 0$ (or $c = 1.0$). The upper part and the lower part of the theoretical curves represent the equilibrium overlap $m_\infty(\alpha)$ and the basin of attraction $m_0(\alpha)$, respectively. Both of these are obtained as functions of α. The vertical parts of the curves represent the storage capacity α_c. Of the four curves, the fourth order approximation is in best agreement with the numerical simulations. From this result, it is found that we must take account of higher order temporal correlations of noise $z(t)$ to predict the behavior of the present model.

For various values of the dilution parameter c, numerical simulation and theoretical analysis were carried out. These results are given in Fig.3.7. Here, the simulations were done in the case of symmetric dilution ($c_{ij} = c_{ji}$) and asymmetric dilution ($c_{ij} \neq c_{ji}$). We confirm that there is no discrepancy between in the two cases as long as we consider statistical properties. Moreover, it is observed that theoretical results are consistent with the simulations in this case as well.

In Fig.3.8, we display the dependence on the ratio of connected synapses c. If $c \geq 0.3$, although the vertical lines move to left, the upper and lower curves are affected only slightly. The two curves do not approach each other until c reaches to 0.1. In a previous section, we found that, for the case of symmetric dilution, equilibrium overlap remains comparatively large even if c is quite small. In the present study, this has been confirmed in the case of asymmetric dilution as well. Furthermore, we have found that the basin remains sufficiently wide even for small values of c.

We compare the width of basin of the present model with that of the traditional model in Fig.3.9. Since the network retrieves the target pattern when $m(0) > m_0(\alpha)$, we adopt $1 - m_0(\alpha)$ as the width of the basin. This figure contains a plot of $1 - m_0(0.8\alpha_c^{\rm 4th})$ for each model. Here, $\alpha_c^{\rm 4th}$ is the storage capacity obtained with the fourth order approximation. Making such a comparison, we see that the oscillator model has wider basin and is more

robust against synaptic dilution.

3.5 Relation between replica theory and statistical neurodynamics theory

In this section, we show, assuming that the system is in the equilibrium state, the result obtained by the replica method can be derived from the statistical neurodynamics theory. According to replica theory, using the replica symmetric ansatz, the equilibrium state in the oscillator neural network is characterized by two order parameters, m and r, corresponding to the overlap with the retrieval pattern and the mean-square random overlap for the unretrieved patterns, respectively. Using an another expression of Eqs. (3.20), we can write these order parameters in the equilibrium state satisfy the equations,

$$\begin{cases} m = \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} DxDy \frac{x + \frac{2m}{\sqrt{\alpha r + 2\eta^2}}}{\sqrt{(x + \frac{2m}{\sqrt{\alpha r + 2\eta^2}})^2 + y^2}} \\ r = \displaystyle 2\left\{1 - \frac{1}{\sqrt{\alpha r + 2\eta^2}} G_{rp}\left(\frac{m}{\sqrt{\alpha r + 2\eta^2}}\right)\right\}^{-2} \end{cases}, \tag{3.42}$$

where $DxDy$ denotes $\frac{1}{2\pi}\exp(-\frac{x^2+y^2}{2})dxdy$ and $G_{rp}(a)$ is defined by

$$G_{rp}(a) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} DxDy \frac{x^2 + y^2 + 2ax}{\sqrt{(x + 2a)^2 + y^2}}. \tag{3.43}$$

Introducing the variable $\sigma_{rp} = \sqrt{\frac{\alpha r}{2}}$, the above equations can be rewritten as

$$\begin{cases} m = \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} DxDy \frac{x + \frac{2m}{\sqrt{2\sigma_{rp}^2 + 2\eta^2}}}{\sqrt{(x + \frac{2m}{\sqrt{2\sigma_{rp}^2 + 2\eta^2}})^2 + y^2}} \\ \sigma_{rp}^2 = \displaystyle \alpha\left\{1 - \frac{1}{\sqrt{2\sigma_{rp}^2 + 2\eta^2}} G_{rp}\left(\frac{m}{\sqrt{2\sigma_{rp}^2 + 2\eta^2}}\right)\right\}^{-2} \end{cases} \tag{3.44}$$

We easily find that the equation here for m is equivalent to equation (3.36), provided that σ_{rp} coincides with the variance of the crosstalk noise z^c in equation (3.38), σ_c. Therefore, we need to prove that the variance of the crosstalk noise z^c in the equilibrium state

satisfies the condition represented by equation(3.44). The equilibrium state implies that $z_i^c(t+1) = z_i^c(t) = z_i^c$ in Eq. (3.32). Then, z_i^c is given by

$$z_i^c = \frac{1}{1-U}\frac{1}{N}\sum_{\nu\neq 1}^{P}\sum_{j=1}^{N}\xi_i^\nu\tilde{\xi}_j^\nu\frac{h_j^\nu}{|h_j^\nu|}. \tag{3.45}$$

Squaring and Averaging, we obtain

$$\sigma_c^2 = \frac{\alpha}{(1-U)^2}. \tag{3.46}$$

From the definition of U in equation (3.33), we find

$$\begin{aligned} U &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} DxDy\frac{1}{2\sqrt{(\sigma x+m)^2+(\sigma y)^2}} \\ &= \frac{1}{2\sigma}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} DxDy\frac{1}{\sqrt{(x+\frac{m}{\sigma})^2+y^2}} \end{aligned}. \tag{3.47}$$

Therefore, substituting the relation $\sigma^2 = \frac{\sigma_c^2+\eta^2}{2}$, the equation (3.46) becomes

$$\sigma_c^2 = \alpha\left\{1-\frac{1}{\sqrt{2\sigma_c^2+2\eta^2}}G_{sn}\left(\frac{m}{\sqrt{2\sigma_c^2+2\eta^2}}\right)\right\}^{-2}, \tag{3.48}$$

with

$$G_{sn}(a) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} DxDy\frac{1}{\sqrt{(x+2a)^2+y^2}}. \tag{3.49}$$

Now we can prove that $G_{sn}(a) = G_{rp}(a)$, because

$$\begin{aligned} G_{rp}(a) - G_{sn}(a) = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} DxDy\frac{x^2+y^2+2ax}{\sqrt{(x+2a)^2+y^2}} \\ &-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} DxDy\frac{1}{\sqrt{(x+2a)^2+y^2}} \\ = &\int_0^{2\pi}d\varphi\int_0^{\infty}dRR(R-2a\cos\varphi)\exp\left(-\frac{R^2-4aR\cos\varphi+4a^2}{2}\right) \\ &-\int_0^{2\pi}d\varphi\int_0^{\infty}dR\exp\left(-\frac{R^2-4aR\cos\varphi+4a^2}{2}\right) \\ = &\int_0^{2\pi}d\varphi\left\{\left[-R\exp\left(-\frac{R^2-4aR\cos\varphi+4a^2}{2}\right)\right]_{R=0}^{R=\infty}\right. \\ &\left.+\int_0^{\infty}dR\exp\left(-\frac{R^2-4aR\cos\varphi+4a^2}{2}\right)\right\} \\ &-\int_0^{2\pi}d\varphi\int_0^{\infty}dR\exp\left(-\frac{R^2-4aR\cos\varphi+4a^2}{2}\right) \\ = &\int_0^{2\pi}d\varphi\left[-R\exp\left(-\frac{R^2-4aR\cos\varphi+4a^2}{2}\right)\right]_{R=0}^{R=\infty} \\ = &\ 0. \end{aligned} \tag{3.50}$$

Hence, as for the equilibrium state, the result of the present theory coincides with that of the replica theory.

We should remark that this derivation is essentially equivalent to SCSNA, which has been proposed by Shiino and Fukai [46], Okuda [39] first applied this method to the analysis of the equilibrium state in oscillator neural networks. It is worth noting that equation (3.48) can be derived directly from equation (3.38) in the limit $n \to \infty$. Using the fact that, owing to $z_i(t+1) = z_i(t) = z_i$, $X(t,t') = 1$ in the equilibrium state, it can be shown that equation (3.38) reduces to the same form as equation (3.48).

3.6 Conclusion

From the theoretical analyses through the replica theory and the statistical neurodynamics theory, A good understanding of the properties of the oscillator neural network. We found the numerical simulations support our theoretical results. The main results obtained in

this study are as follows.

- Using the replica symmetric solution, we have estimated the influence of random synaptic dilution on the storage capacity and the critical overlap. Although the storage capacity of the oscillator neural network is less than that of the Hopfield model, the difference between the performance of the fully connected network and that for a diluted network is smaller in the case of the oscillator model than in the case of the Hopfield model.

- The statistical neurodynamics theory enable us to describe the retrieval process of the oscillator neural networks. Then, we must take into account the higher order temporal correlations of noise. The present study shows that it is necessary to consider at least the fourth order approximation.

- For all values of c, theoretical results are in good agreement with numerical simulation. These theoretical curves indicate equilibrium overlaps and basins change little even if c decreases to about 0.3. Furthermore, the basins remain sufficiently wide near saturation.

- The widths of basins in the oscillator model are wider than that in the binary model. Moreover, the oscillator model is found to be more robust against decrease of c.

In conclusion, we have found that the oscillator neural network exhibits good performance while processing detailed information such as the timing of neuronal firings. Our results support the plausibility of temporal coding.

Because our analysis has been simplified by discretizing time and assuming synchronous updating, in general, this simplification may results in behavior which differs from that of the original model possessing continuous time. In fact, it is found that, in some cases, the dynamics of (3.10) with synchronous updating possess period 2 solutions

which do not correspond to any solution of the original phase model (3.2). In this case, an attractor existing in the original model becomes an invariant torus constituted by such periodic solutions in our model, (3.10).

For example, this phenomenon is observed in certain situations in which the coupled phase oscillators have an in-phase synchronization attractor. However, under the conditions that the memorized phase patterns are uniformly random and the system size is large, we find that such irrelevant solutions are not observed in numerical simulations. Although we believe that the irrelevant solutions are not realized due to their high symmetry, at the present time we cannot confirm this belief theoretically. We think that this is an important problem to be considered further.

The question now arises whether the theoretical results given in this paper apply in the case of the original model or not. To clarify this point in the case of the auto-associative model, we compared the basins of attraction in the two models numerically, as shown in Figure 3.10. Although a slight difference between these models can be found, we can safely say that the basins of attraction are qualitatively very similar.

Finally, let us make a comment leading to future problems. When we carried out numerical simulation of the oscillator neural networks, it was observed that a network is rarely trapped in spurious states. This may be one of factor responsible for the ability of the oscillator network to recall from considerably noisy patterns. Though this point has not yet been investigated in the general case, the case of $\alpha = 0$ (that in which the number of patterns P remains finite in the limit $N \to \infty$) has been investigated, and it has been reported that the symmetric mixture states are all unstable [39]. Investigation of the case for finite α should provide a deeper understanding of characteristics of oscillator neural networks.

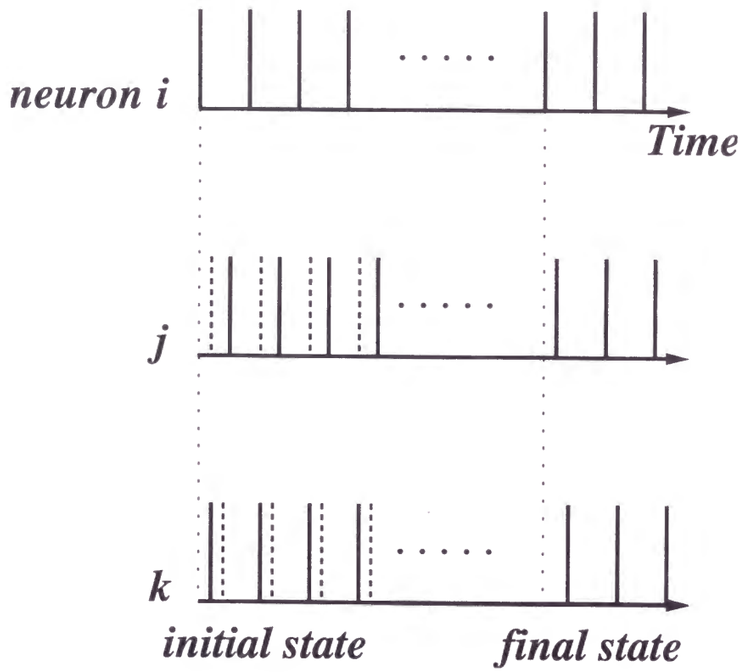


Figure 3.1: Retrieval process of a phase pattern in the oscillator neural network. In an initial noisy pattern, the relative timing of the spikes is disturbed from the memorized positions (dashed lines). These relative phases are corrected dynamically in the final state.

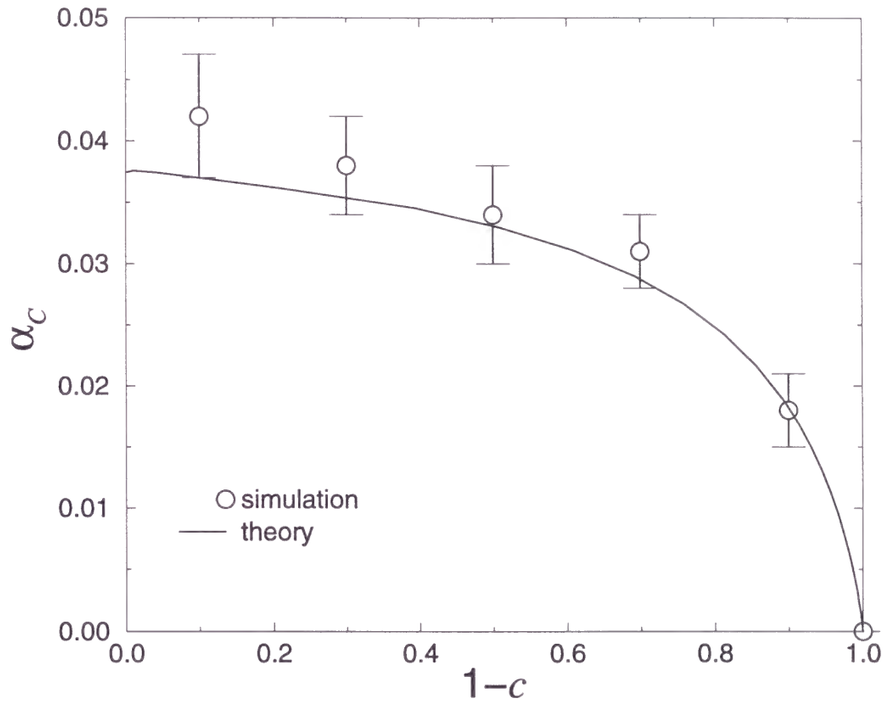


Figure 3.2: Dependence of the storage capacity α_C on the ratio of disconnected synapses $1 - c$. The solid curve represents theoretical results. The data points indicate simulation results with $N = 1500$ for 20 trials.

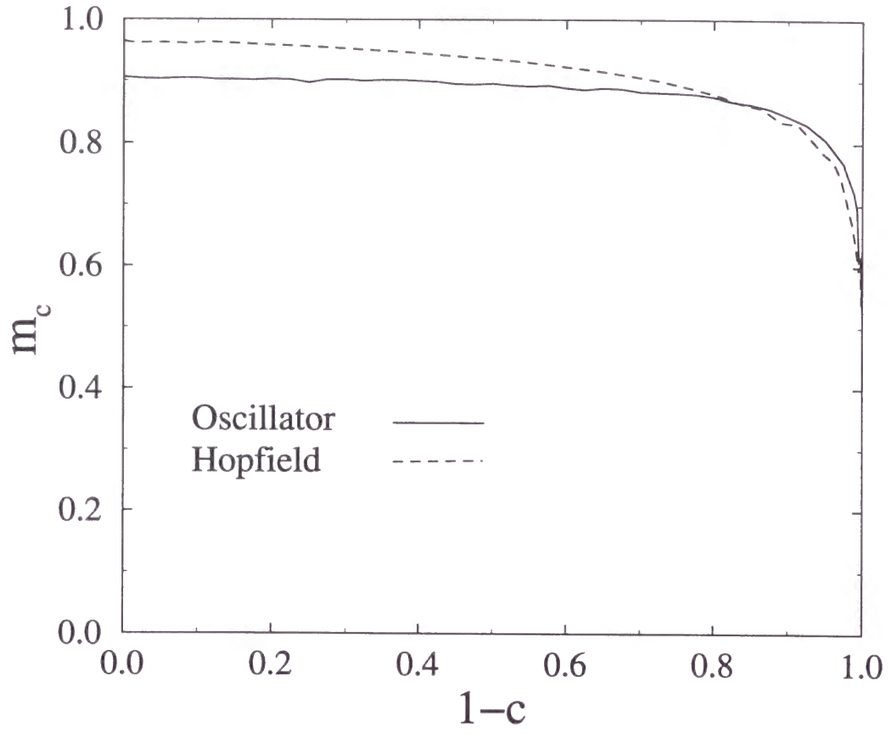


Figure 3.3: Dependence of the critical overlap m_C on the ratio of disconnected synapses $1-c$. The solid and dashed curves correspond to the oscillator and the Hopfield models, respectively.

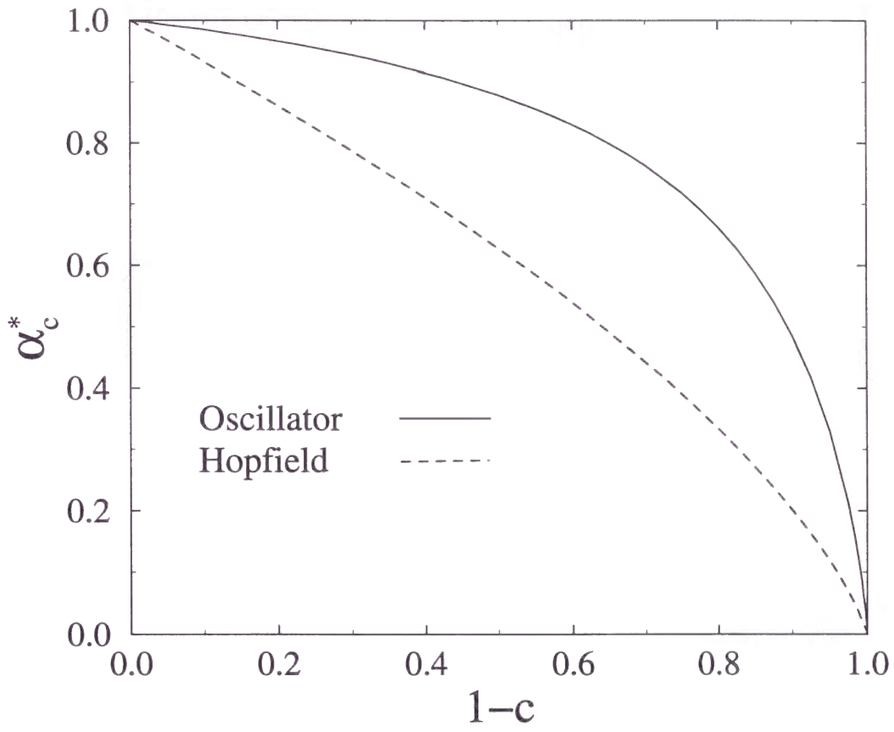


Figure 3.4: Comparison of the normalized storage capacities, α_C^*, between the oscillator model and the Hopfield model.

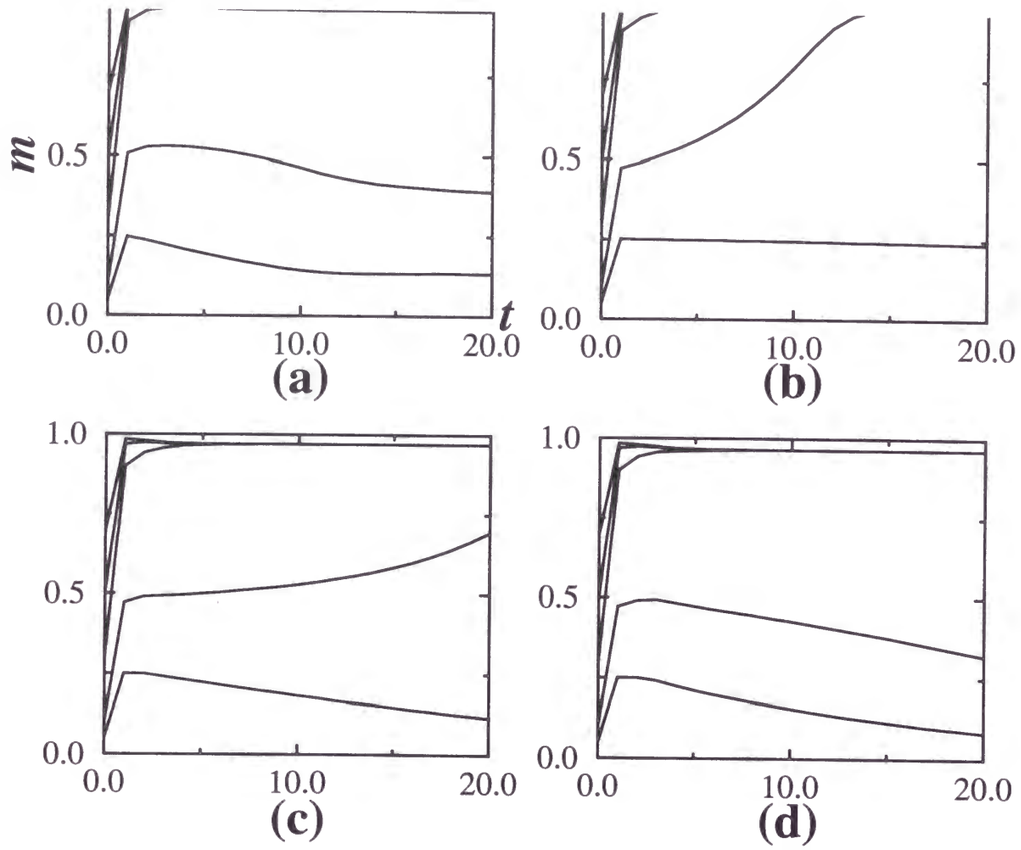


Figure 3.5: Typical time evolution of overlaps for $\alpha = 0.03$, and the initial overlaps $m = 0.05, 0.1, 0.3, 0.5$ and 0.7. (a)Numerical simulation with $N = 1000$ (b)Theoretical curves at first order approximation, (c) second order, and (d) third order.

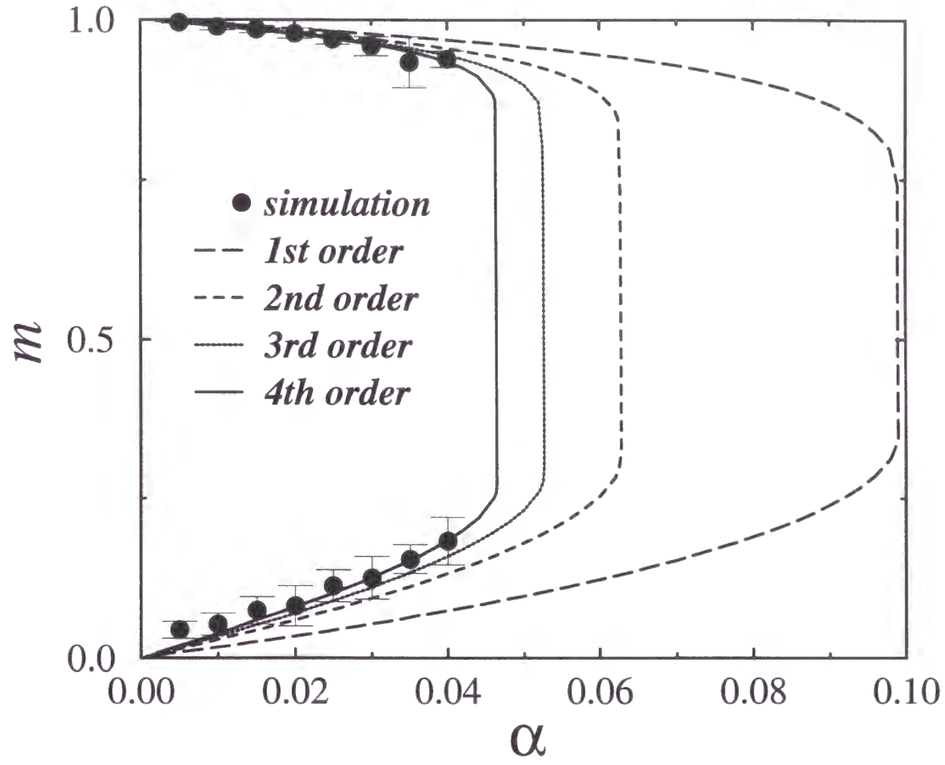


Figure 3.6: The equilibrium overlap and basin of attraction. The four curves represent the theoretical results for various order approximations. The ordinate is the overlap m and the abscissa is the storage ratio α. The data points indicate simulation results with $N = 1000$ for 20 trials. The upper part, the lower part, and the vertical part of the theoretical curves represent the equilibrium overlap, the basin of attraction, and the storage capacity, respectively.

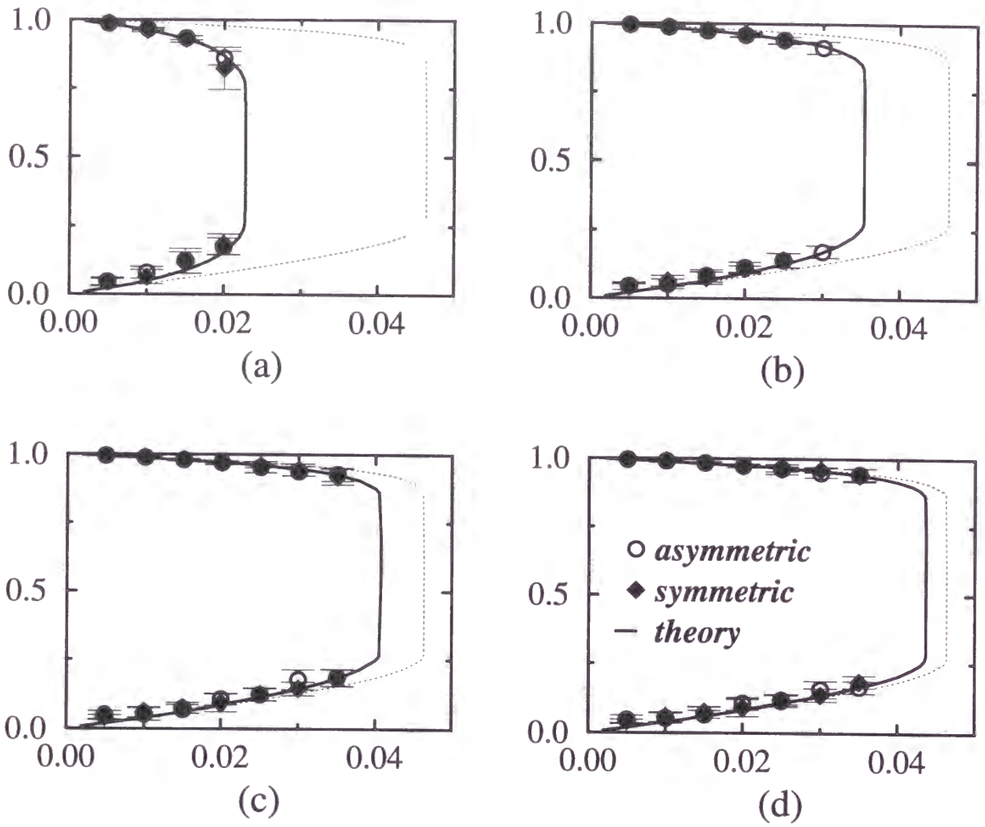


Figure 3.7: Comparison between the effects of symmetric and asymmetric dilution. The solid curves represent the theoretical results at fourth order. The ratios of connected synapses are (a)$c = 0.1$, (b)$c = 0.3$, (c)$c = 0.5$ and (d)$c = 0.7$. For reference, the result for the case $c = 1.0$ (i.e. the fully connected case) is indicated by the dashed curves.

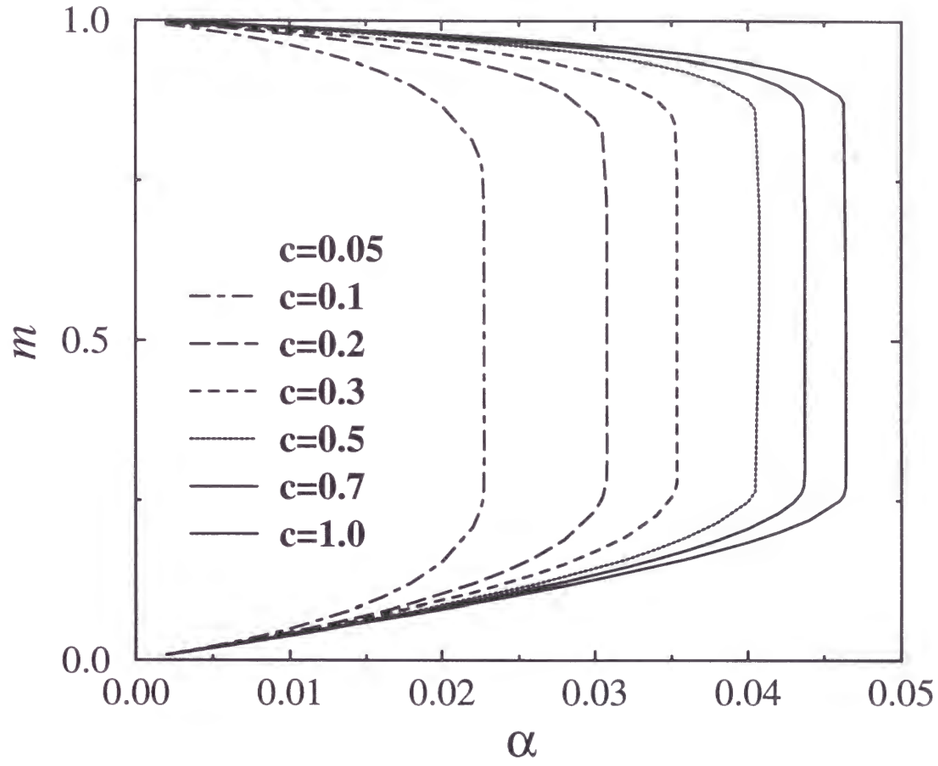


Figure 3.8: Dependence of the theoretical curve on the ratio of connected synapses c. These curves were obtained with the fourth order approximation.

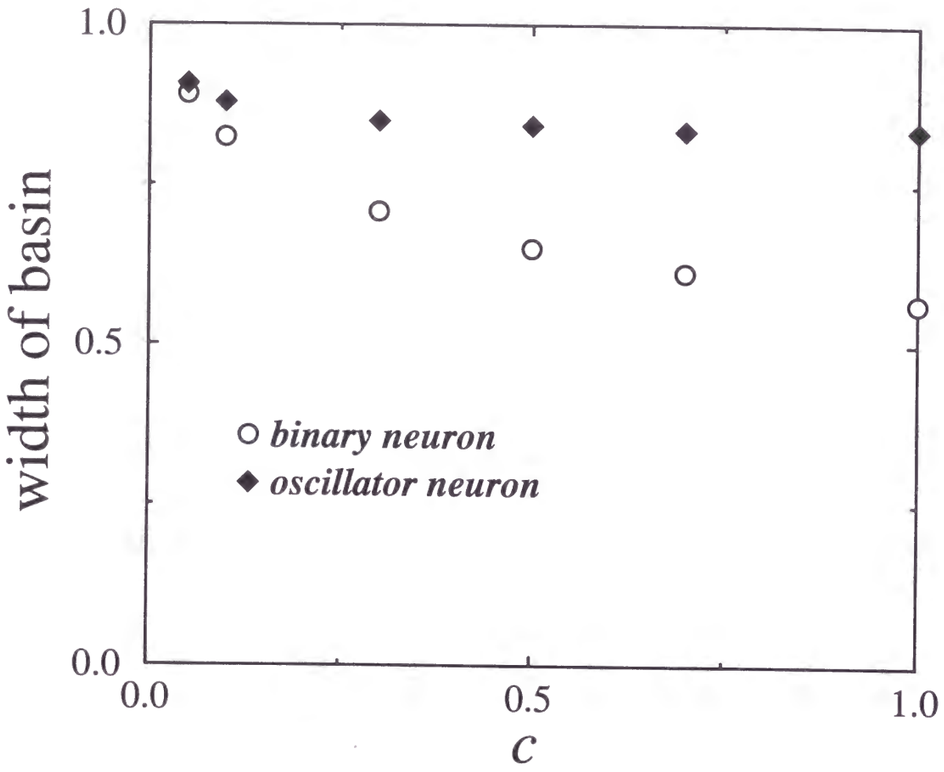


Figure 3.9: Comparison of the sizes of basins for the oscillator and binary models near saturation, α_c^{4th}.

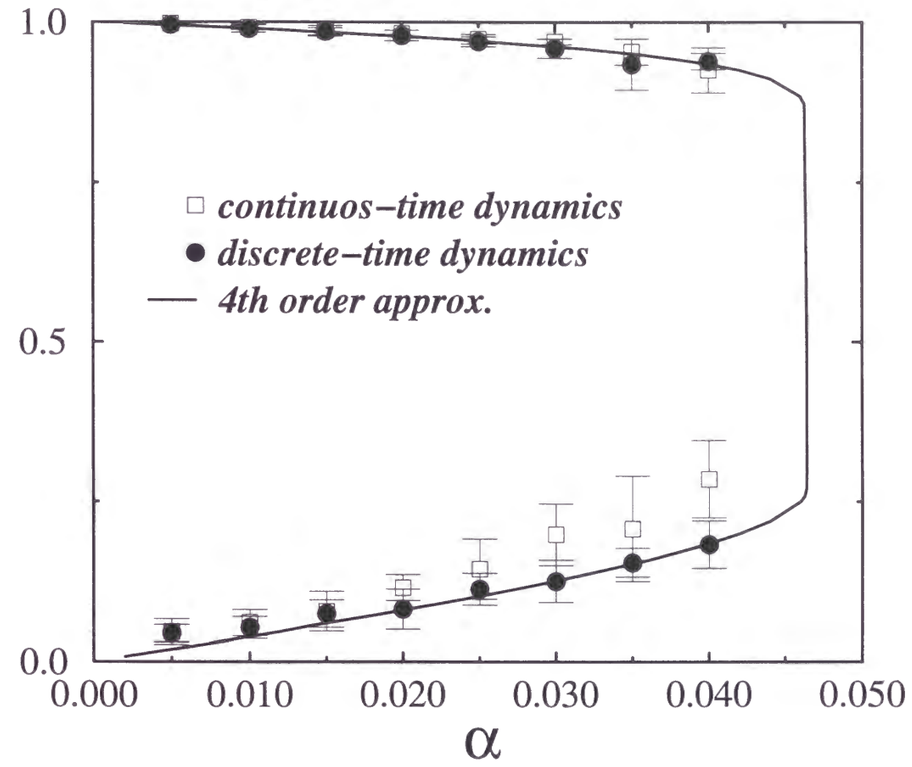


Figure 3.10: Comparison of basins of attraction. The circles represent numerical data in the discrete-time model associated with our theoretical results. The squares represent simulations in the original phase model, where the dynamics are govern by a set of ordinary differential equations. For reference, the theoretical curve (the solid line) obtained using the fourth order approximation is also shown.

Chapter 4

Conclusion

In the preceding chapters, we discussed the associative memory neural networks concerning biological information encoding. Here, we summarize our results.

- The properties of the neural networks for sparsely coded sequential patterns were studied. In order to analyze the retrieval processes of the model, we applied the method of statistical neurodynamics and obtained the coupled equations describing the dynamics of the order parameters such as the overlap and the activity. It is found that our theory provides good predictions for the storage capacity and the basin of attraction obtained through numerical simulations. The results indicates that the nature of the basin of attraction depends on the methods of activity control employed. In the case that the synapses are randomly diluted, the associative abilities deteriorate proportionately to the decrease of the ratio of connected synapses. Furthermore, it is found that robustness against random synaptic dilution slightly deteriorates with the degree of sparseness. Finally, it is confirmed that the storage capacity of our model also diverge as the same asymptotic form as the model for the sparsely coded static patterns.

- The properties of the oscillator neural networks were studied. At first, the equilibrium properties such as the storage capacity was investigated using the replica

theory. From the resultant saddle point equations for the equilibrium order parameters, it was showed that the oscillator model is more robust against random synaptic dilution than is the Hopfield model. Next, we applied the method of statistical neurodynamics to the oscillator model. Using the derived equations, we present the phase diagram showing both the basin of attraction and the equilibrium overlap in the retrieval state. From the results, we found that it is essential to take into account of the higher order temporal correlations of noise in order to predict the behavior of the retrieval processes of the oscillator model. In addition, it is turned out that both the attractor and the basin are preserved even though random synaptic dilution is promoted and that the oscillator model is more robust than the Hopfield model in respect with the basin of attraction. Taking into account of the fact that oscillator networks contain more detailed information than binary networks, the obtained results constitute significant support for the plausibility of temporal coding.

In conclusion, it is turned out that the models with biological information encoding exhibit different properties and performance from the Hopfield model.

Although the present study is restricted to the framework of the associative memory, it is important to consider such information encodings in various architectures of neural networks. On the other hand, it is true that the mathematically tractable models including our models are still long way from biological neural systems. However, the understandings of the basic models are expected to be useful on studying more realistic and more complicated models. At such opportunities, we hope that the present study serves as the valuable knowledge.

Bibliography

[1] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the theory of neural computation*, Addison-Wesley, (1991).

[2] W. S. McCulloch, and W. Pitts, *Bull. Math. Biophys.*, **5**, 115, (1943).

[3] A. L. Hodgkin, and A. F. Huxley, *J. Physiol.*, **117**, 500, (1952).

[4] D. O. Hebb, *The Organization of Behavior: A Neuropsycological Theory*, Wiley, New York, (1949).

[5] K. Nakano, *IEEE Transactions on System, Man and Cybernetics*, **26**, 859, (1972).

[6] T. Kohonen, and M. Ruohonen, *IEEE Transactions on Computers*, **22**, 701, (1973).

[7] J. A. Anderson, *Mathematical Biosciences*, **14**, 197, (1972).

[8] J. J. Hopfield, *Proc. Nat. Acad. Sci. USA*, **79**, 2554, (1982).

[9] D. J. Willshaw, O. P. Buneman, and H. C. Longuet-Higgins, *Nature*, **222**, 960, (1969).

[10] D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Phys. Rev. Lett.*, **55**, 1530, (1985). and, D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Phys. Rev. A*, **35**, 2293, (1987).

[11] M. V. Tsodyks, and M. V. Feigelman, *Europhys. Lett.*, **6**, 101, (1988).

[12] C. J. P. Vicente, and D. J. Amit, *J. Phys. A: Math. Gen.*, **22**, 559, (1989).

[13] H. Horner, *Z. Phys. B*, **75**, 133, (1989).

[14] J. Buhmann, R. Divko, and K. Schulten, *Phys. Rev. A*, **39**, 2689, (1989).

[15] E. Gardner, *J. Phys. A: Math. Gen.*, **21**, 257, (1989).

[16] H. Rieger H, M. Schreckenberg, and J. Zittartz, *Z. Phys. B*, **72**, 523, (1988).

[17] H. Horner, D. Bormann, M. Frick, H. Kinzelbach, and A. Schmidt, *Z. Phys. B*, **76**, 381, (1989).

[18] E. Gardner, B. Derrida, and P. Mottishaw, *J. Physique*, **48**, 741, (1987).

[19] A. C. C. Coolen, and D. Sherrington, *Phys. Rev. Lett.*, **71**, 3886, (1993).

[20] S. Amari, and K. Maginu, *Neural Networks*, **1**, 63, (1988).

[21] M. Okada, *Neural Networks*, **8**, 833, (1995).

[22] M. Okada, *Neural Networks*, **9**, 1429, (1996).

[23] D. R. C. Dominguez, and D. Bollé, *Phys. Rev. Lett.*, **80**, 2961, (1998).

[24] E. Domany, W. Kinzel, and R. Meir, *J. Phys. A: Math. Gen.*, **22**, 2081, (1989).

[25] S. Amari, *Neural and Synagetic Computers* Berlin, Springer, p85, (1988).

[26] H. Nishimori, and T. Ozeki, *J. Phys. A: Math. Gen.*, **26**, 859, (1993).

[27] H. Sompolinsky, *Phys. Rev. A*, **34**, 2571, (1986).

[28] K. Kitano, and T. Aoyagi, *Phys. Rev. E*, **57**, 5734, (1998).

[29] J. P. Nadal, *J. Phys. A: Math. Gen.*, **24**, 1093, (1991).

[30] C. M. Gray, P. Konig, A. K. Engel, and W. Singer, *Nature*(London), **338**, 334, (1989).

[31] C. V. Malsburg, and W. Schneider, *Biol. Cybern.*, **54**, 29, (1986).

[32] L. F. Abbot, *J. Phys.*, **A23**, 3835, (1990).

[33] D. L. Wang, J. Buhmann, and C. V. Malsburg, *Neural Comp.*, **2**, 95, (1990).

[34] F. Gerl, K. Bauer, and U. Krey, *Z. Phys. B*, **88**, 339, (1992).

[35] A. Arenas, and C. J. P. Vincente, *Europhys. Lett.*, **26**, 79, (1994).

[36] T. Fukai, and M. Shiino, *Neural Comp.*, **7**, 529, (1995).

[37] K. Park, and M. Y. Choi, *Phys. Rev. E*, **52**, 2907, (1995).

[38] T. Aoyagi, *Phys. Rev. Lett.*, **74**, 4075, (1995).

[39] K. Okuda, *unpublished.*

[40] J. Cook, *J. Phys.*, **A22**, 2057, (1989).

[41] A. J. Noest, *Europhys. Lett.*, **6**, 469, (1988).

[42] J. L. Van Hemmen, *Phys. Rev. A*, **36**, 1959, (1987).

[43] A. J. Noest, *Phys. Rev. Lett.*, **63**, 1739, (1989).

[44] M. Bouten, A. Komoda, and R. Serneels, *J. Phys. A*, **23**, 2605, (1990).

[45] Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence*, Springer, New York, (1984).

[46] M. Shiino, and T. Fukai, *J. Phys.*, **A25**, L375, (1992).

List of publication

1. Toshio Aoyagi, and Katsunori Kitano, Phys. Rev. E **55**, *7424*, (1997).
   'Effect of random synaptic dilution in oscillator neural networks'

2. Katsunori Kitano, and Toshio Aoyagi, Phys. Rev. E **57**, *5914*, (1998).
   'Effect of random synaptic dilution on recalling dynamics in an oscillator neural network'

3. Toshio Aoyagi, and Katsunori Kitano, Neural Comput., **10**, *1527*, (1998).
   'Retriaval Dynamics in Oscillator Neural Networks'

4. Katsunori Kitano, and Toshio Aoyagi, J. Phys. A: Math. and Gen., **31**, *L613*, (1998).
   'Retrieval dynamics of neural networks for sparsely coded sequential patterns'