(2)

Static and dynamic properties of collective phenomena in systems composed of formal neurons

Yasuyuki NAKAMURA

December 1994

Abstract

Static and dynamic properties of collective phenomena of many-body systems, mainly composed of formal neurons, are studied numerically and theoretically. For theoretical analyses, we use methods of equilibrium and nonequilibrium statistical mechanics. The model systems which we adopted are neural networks and coupled oscillator systems.

First, the Glauber dynamics of the Hopfield neural network model is studied with use of the Mori-Zwanzig projection operator formalism for irreversible processes and exact evolution equations for overlaps are derived.

Secondly, effects of correlations among the embedded patterns on dynamics in a neural network are studied. Combined with delay of signal transmission, the correlations among the embedded patterns are shown to give rise to a time sequence of patterns. The theoretical analysis on the basis of notion of sublattice magnetization are performed and evolution equations which govern the overlap dynamics are derived. Numerical solutions of the equations support the results of computer simulations.

Thirdly, as an extension of the Hopfield model, we propose a neural network composed of D-dimensional spin neurons ($D \geq 1$). Our model is equivalent to the Hopfield model in the case of $D = 1$ and is related to the clock neural network in the case of $D = 2$. We derive the free-energy of our model using the replica symmetric theory. When a finite number of patterns are embedded, they are found to be retrievable if the temperature T is lower than $1/D$. The phase diagram and the storage capacity of the network are also obtained with the storage capacity $\alpha_c = 0.0743 (D = 2)$ and $\alpha_c = 0.0432 (D = 3)$ for $T = 0$.

Finally, we analyze both numerically and theoretically dynamical behaviors of systems composed of limit-cycle oscillators, which are coupled by time-delayed nearest-neighbor interaction. In this system we found cluster states which are characterized by a phase

difference of neighboring oscillators and its collective frequency. Time-delay turns out to have an important effects on the stability of various cluster states. The linear and nonlinear stability of cluster states are analyzed theoretically.

Acknowledgment

The author would like to express his sincere gratitude to Professor Toyonori Munakata, Department of Applied Mathematics and Physics, Kyoto University, for his guidance, discussions, encouragements and critically reading the manuscript. He is also grateful to Assistant Professor Akito Igarashi, Dr. Yutaka Kaneko and Dr. Toshio Aoyagi, Department of Applied Mathematics and Physics, Kyoto University, for valuable discussions and encouragements.

He wishes to express his gratitude to Mr. Futoshi Tominaga, Kawasaki Steel Corporation, and Mr. Kan Torii, Department of Applied Mathematics and Physics, Kyoto University, for useful discussions.

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

Contents

Chapter 1

Introduction

1.1 Overview

The human brain which consists of roughly 3×10^{10} neurons and 10^{14} synapses [1] has an excellent ability of information processing; especially learning, pattern recognition and associative memory. The motivation of studying neural networks lies in understanding the principles and mechanisms of the information processings in the brain. The study of neural networks also aims at the application of the principle of information processings of the brain to the engineering; it is expected to construct a new concept of information processings by which hard tasks for existing computers, e.g. a pattern recognition, are well dealt with.

For the above purposes, a number of studies have been made with two main approaches: One is the physiological approach, in which experiments have revealed behavior of a single neuron, the structure of the brain, and so on. The other approach is modeling the brain function with artificial neural networks composed of model neurons, on the basis of the results of neurophysiology, and analyzing the model, which is our standpoint throughout this article.

Each neuron receives input signals through synapses from many other neurons. If the sum of these signals exceeds the threshold value, the neuron becomes active. Otherwise it is inactive. If the neuron is active, it gives an output signal to other neurons and the same

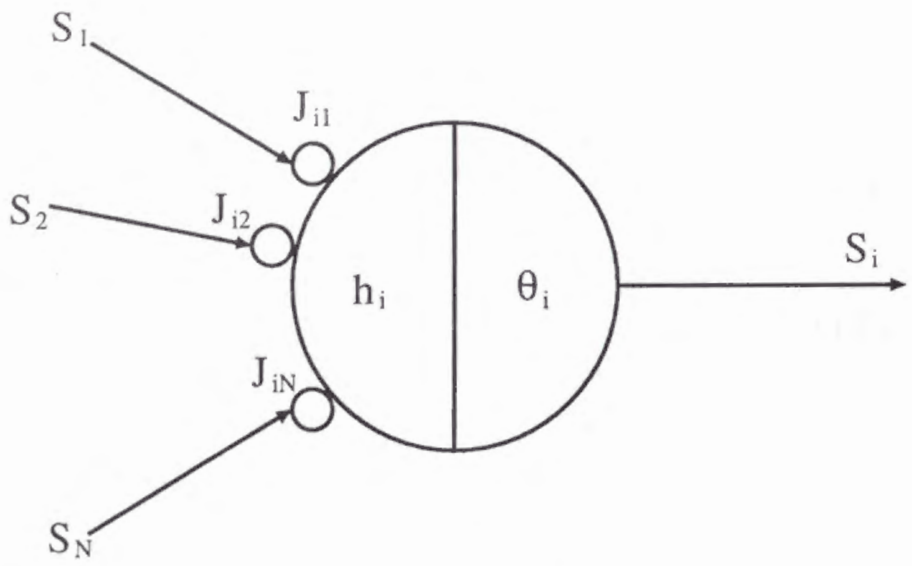

Figure 1.1: The formal neuron

process is repeated successively. Considering this mechanism of a neuron, W.S.McCulloch and W.Pitts proposed the first model neuron, so called a formal neuron, in 1943 [2]. The formal neuron is illustrated in Fig.1.1 and works as follows: The number of neurons is N, and each neuron is assumed to take the state $+1(-1)$ when it is active (inactive) respectively. (Note that the states were represented by 1(active) and 0(inactive) instead of 1 and -1 in the original paper, but for the mathematical convenience we adopt the representation $+1, -1$ throughout this article.) Let us denote the state of ith neuron at time t by $S_i(t)$ which takes $+1$ or -1. J_{ij} are the synaptic efficacies from jth neuron to ith neuron. The sum of all signals weighted by J_{ij} at ith neuron is called the local field at neuron i. If it is expressed by $h_i(t)$, we can write as

$$h_i(t) = \sum_{j=1}^{N} J_{ij} S_j(t). \tag{1.1}$$

If the threshold value is θ_i, the state of ith neuron at the next time step $t + \Delta t$ becomes

$$S_i(t + \Delta t) = \mathrm{sgn}(h_i - \theta_i) \tag{1.2}$$

where

$$\mathrm{sgn}(x) = \begin{cases} +1, & (x \geq 0) \\ -1, & (x < 0). \end{cases} \tag{1.3}$$

In this way a neuron updates its state at each time step deterministically. As an extension of this model, a stochastic version is also considered. In this model the neuron will be in the state ± 1 at the next time step with probability,

$$\begin{aligned} \mathrm{Prob}\{S_i(t + \Delta t) = \pm 1\} &= \frac{\exp(\pm h_i/T)}{\exp(h_i/T) + \exp(-h_i/T)} \\ &= \frac{1}{2}\left\{1 \pm \tanh\left(\frac{h_i(t)}{T}\right)\right\}. \end{aligned} \tag{1.4}$$

Here T is the parameter which determines the strength of fluctuation. T does not have a meaning of physical temperature in neural networks, but it is often called temperature. In the case of $T \to 0$ the dynamics (1.4) reduces to the one (1.2) with $\theta_i = 0$.

The way of updating can be asynchronous or synchronous. Asynchronous updating means a dynamical process in which at each time step only one neuron is updated. On the other hand synchronous updating means the one in which at each time step every neuron is selected and updated.

Though very simple, the formal neuron is attractive because it is easy to construct a network with such neurons and moreover the network shows various collective properties. On the other hand A.L.Hodgkin and A.F.Huxley described the dynamics of a single neuron with four variables nonlinear differential equations, so called Hodgkin-Huxley equations [3]. While this model is very excellent to explain the mechanism of a single neuron, it is difficult to construct a network. Since our interest is centered around collective properties of networks composed of neurons, we prefer to the formal neuron for modeling a neural network.

With use of formal neurons, many kinds of neural network model have been proposed. They are classified into two main types as are illustrated in Fig.1.2: One is a layered feed-forward network which consists of an input layer, some hidden layers and an output

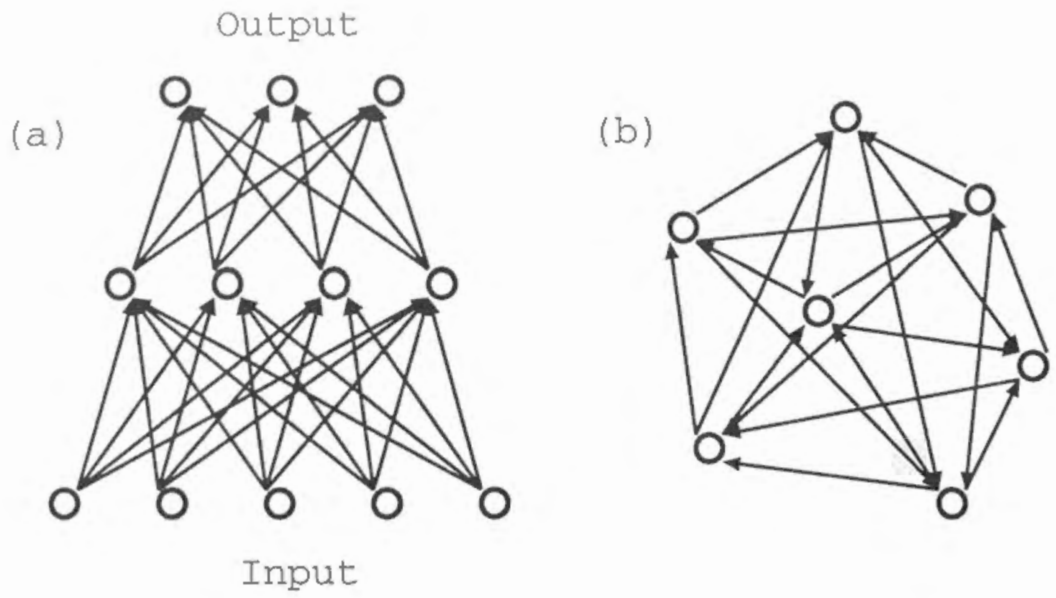

Figure 1.2: (a)Layered feed-forward network, (b)Recurrent network

layers and has only feed-forward connections between one layer and the next. Rosenblatt's perceptron [4] is the most simple example of this kind of networks. The other is a recurrent network in which neurons are interconnected. In this thesis we mainly investigate the latter type.

Among many models which have the property of associative memory, the model proposed by J.J.Hopfield [5] seems to be one of the most successful ones. Here we say that a system possesses the property of associative memory if it memorizes several prototype patterns and, when a distorted version of one of these patterns is presented, it is able to retrieve the corresponding prototype automatically. Hopfield's main contribution is introducing the idea of an energy function which plays a role of a Lyapunov function. If it exists, the system flows toward the local or global minima of the energy, which are called attractors, and settles in one of them. He related this dynamical process to the retrieval process of associative memory as stated below: Suppose that minima of the energy function correspond to the memorized pattern. Then starting from the point inside

the potential well around a minimum (the initial input pattern which is slightly different from the memorized one), the state of the network flows to the attractor (the memorized pattern). Thus the relaxation toward the minimum corresponds to retrieving the stored pattern from its partial or distorted one. He embedded p patterns represented by $\{\xi_i^\mu\}(\mu = 1,\dots,p;\ i = 1,\dots,N)$ into the neural network using Hebb's learning rule [6],

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu(1-\delta_{ij}). \tag{1.5}$$

Here ξ_i^μ takes the state ± 1. Note that these p patterns which he treated are uncorrelated, that is, the correlation $C_{\mu\nu}$ defined by

$$C_{\mu\nu} = \frac{1}{N}\sum_{i=1}^{N}\langle\xi_i^\mu\xi_i^\nu\rangle \tag{1.6}$$

satisfies the condition

$$C_{\mu\nu} = \delta_{\mu\nu}. \tag{1.7}$$

With synaptic couplings specified by Eq.(1.5), each pattern corresponds to the minimum of the energy

$$E = -\frac{1}{2}\sum_{i\neq j}^{N}J_{ij}S_iS_j, \tag{1.8}$$

under the condition, $p \ll N$. In the case of asynchronous dynamics, when the state of the network changes, the energy always decreases.

Stimulated by the Hopfield's work, many physicists have become interested in the study of neural networks. When a network has an energy function, methods of equilibrium statistical mechanics are powerful to analyze a performance of a network quantitatively. Amit et.al. succeeded in deriving a free-energy of the Hopfield model composed of stochastic neuron specified by Eq.(1.4) using the replica method. With use of this result, some properties in the equilibrium state have been understood: For example, they estimated the storage capacity as 0.138 under the assumption of replica symmetry. This result is nearly coincides with the value 0.15 which is the result obtained by Hopfield with the computer simulation. Here the storage capacity α_c is defined as $\alpha_c = p_c/N$, where p_c is the maximum number of patterns which can be embedded and N is the number of neurons.

1.2 Role of statistical mechanics

As is stated in the previous subsection, in order to know the principle of information processings in the human brain, artificial neural network models have been studied. On the other hand, neural networks composed of McCulloch-Pitts' formal neurons are attractive subjects as many-body problems in physics. There exists many analogies between neural networks and physical systems; most significant parallel is that both are made of many strongly interacting components. Physics, more so than any other discipline, has devoted itself to the systematic study of multicomponent systems, and some of the most important concepts developed in physics have emerged from such studies, in particular, the ideas connected with the phenomena of pattern formations. Here pattern formation refers to collective behavior in which systems comprising many simple elements spontaneously forms ordered structures. One of the most fruitful models in modern physics is an extremely simplified description of magnetism, the Ising model [7]. Despite the simplifications and perhaps due to them the model has provided deep insight into the properties of magnetic systems. Thus in order to know properties of matter, it is useful to examine the properties as collective phenomena of the system composed of simple elements. Statistical mechanics have offered mathematical frames for such kinds of studies. Therefore it was rather natural that physicists begin to study neural networks after Hopfield model, in which the similarity between neural networks composed of formal neurons and spin glass was pointed out. Amit el.al. examined static properties of the Hopfield model with use of the method of equilibrium statistical physics and made great success [29, 30].

Systems composed of coupled oscillators are also interesting systems for the study of

collective phenomena. Such systems can model many systems not only in physics but also in chemistry and biology. Dynamics of each oscillator is not particularly interesting indeed, but when oscillators interact to form a network, many interesting phenomena can be observed; for example, they tend to oscillate with phase or frequency locked, and this phenomenon is called synchronization or entrainment. Being different from spin systems, coupled oscillator systems are dissipative systems, and thus the partition function can not be defined. Therefore we can not apply the method of traditional equilibrium statistical mechanics for the study of oscillator systems. Kuramoto have developed the theory of oscillator systems with the method of contraction of evolution equations [59]. He described the motion of oscillators with phase equations under the condition of weak couplings. This way of description of oscillator systems is now widely adopted. We also use this method for the study of oscillator systems in this thesis.

Recently, there is an interesting topic in connection with physiology; synchronization may play an important role for information processings in the cat visual cortex [18, 19, 20]. When a stimulus is given, a group of neurons which correspond to the stimulus show an oscillatory firing; each neuron in the group does not oscillate individually but oscillate in synchrony. Synchronization of oscillatory responses may provide a suitable mechanism for the binding of distributed features of a patterns and thus may contribute to the segmentation of visual scenes. In order to study such phenomena, oscillator neural networks [21, 22, 23, 24] seem to be more appropriate than the conventional ones. Therefore the study of oscillator systems is rather closely related to that of neural networks.

1.3 Outline of the thesis

In chapter 2, with the interest in dynamical behavior of a neural network, for example an associative process, we investigate overlap dynamics in a Hopfield model in the thermodynamic limit, $N \to \infty$, under the condition that p, the number of embedded patterns, remains finite ($p/N \to 0$). For this kind of problem, the method of equilibrium statistical mechanics can not be applied, but the notion of sublattice and sublattice magnetization proposed by van Hemmen et.al. [8] is one of the powerful method. Although exact evolution equations for overlap dynamics are already derived [9, 10], we show another approach with use of the Mori-Zwanzig projection-operator formalism for irreversible processes to derive the same equations [25]. Here the overlap is a parameter which measures the nearness between the state of the system and one of the memorized patterns, and is defined as

$$m^{\mu}(t) = \frac{1}{N}\sum_{i=1}^{N} \xi_i^{\mu} S_i(t). \tag{1.9}$$

The method of projection-operator is powerful approach in the field of nonequilibrium statistical mechanics and we are first to apply this method in the field of neural networks. It is expected that a further progress might be made in the study of dynamical behavior of neural networks with our method.

We have so far mentioned about associative memory for static patterns. However, the problem of temporal association, which refers to the storage and retrieval of time sequence of patterns, is also important. Typical example is the recall of a tune, which can be provoked by a very simple stimulus. Another example is the system of neural group, called a central pattern generators, which controls a rhythm of the muscles or a rhythmic function, such as breathing, chewing, swimming, scratching , etc. A number of neural network models for temporal association have been proposed. In such models, it is expected that the system is quasi-stationary in one pattern before it performs a transition to the next. Then the main difficulty comes from the competition between two forces: one for stabilizing a pattern and the other for causing transition to the next pattern [11]. There are some approaches to overcome the difficulty: introducing short-term modifiable synapses [12, 13]; assuming a complicated synaptic structure [14, 15]; introducing a delay into the transition force, in other words, to use additional asymmetric

couplings with delay [16, 17]. Most of the models mentioned above treated uncorrelated patterns specified by Eq.(1.7) and sequences are given from outside. However it seems to be natural that one learns a pattern by correlating it with other similar ones. It also seems to be natural that when we recalls a main target pattern, being given an initial pattern we associate another similar pattern one after another and finally reach the main target one. Therefore in chapter 3, a neural network which stores correlated patterns are treated. First we propose a new model for associative memory of correlated patterns [26]. In our model, being stimulated by an initial pattern, the network gives rise to spontaneous transitions with use of correlations among embedded patterns. A time delay in transition term is assumed in our model. Secondly, we make some computer simulations, in order to ascertain whether or not our model behave as we expect. Finally, we consider overlap dynamics theoretically in the case of $p/N \to 0$ as in the previous chapter. In our analysis, we use the notion of sublattice and sublattice magnetization which are introduced by van Hemmen and co-workers [8]. We derive coupled differential equations which govern overlap dynamics in our model and compare the numerical results obtained from these equations with simulation results.

In chapter 4, we study the network composed of multi-state neurons. A model neuron which can take more than two states, so called a multi-state neuron, has been proposed as an extension of a formal Ising spin neuron. From a biological point of view, the multi-state neuron is plausible in the sense that a real neuron possesses a richer structure than the formal two-state neuron. Furthermore, from the standpoint of application of neural networks to information processings, the multi-state neuron is useful, because a colored or gray-toned pixel can be represented by one multi-state neuron. We have also a theoretical interest in the properties of network composed of multi-state neurons. As a new model of multi-state neuron, we propose a D-dimensional spin neuron, which is represented by D-dimensional unit vectors. Our model neuron is equivalent to the formal neuron in the case of $D = 1$ and is related to the clock-type neuron proposed by Cook [27] and Noest [28] in the case of $D = 2$. We simulate our model in the case of $D = 2, 3$ and confirm that our model works as an associative memory. We derive a free-energy of the network composed of our model neurons in the case of general dimension D and of $p/N \neq 0$, with use of the replica symmetric theory following the work of Amit et.al. [29, 30]. We will see that the free-energy coincides with that of the Hopfield model for $D = 1$. On the basis of this results, we study some properties of our model, e.g. the storage capacity, the effect of noise, etc. These properties will be summarized in the $T - \alpha$ phase diagram which is to be calculated as our main task in this chapter.

In chapter 5, coupled oscillator systems are investigated. One of the remarkable phenomenon observed in a system of coupled oscillators is a clustering behavior, which is our main concern in this chapter. Globally coupled oscillator systems have been intensively investigated , but there are still some kinds of systems of coupled oscillators which have so far been paid little attention to; for example, the systems composed of short-range coupled oscillators and those composed of time-delayed coupled oscillators. Thus the model system which we treat is that of oscillators which are coupled by time-delayed nearest-neighbor interaction [31]. The time-delay originates from the fact that it takes a finite time for information transmission between two elements in biological networks. The nearest-neighbor interaction makes it possible for the network to yield spatial structures which are absent in the global interaction. We analyze dynamical behaviors of the system both numerically and theoretically. First, we find the clustering behavior in our model and examine the linear and nonlinear stability of cluster states by means of computer simulations. Here the nonlinear stability means the relative stability among cluster states, which is examined under the fluctuation induced by white noise. Next, we explain some of the results theoretically with the energy analysis for the system without a time-delay and with the perturbation method for the system with a finite time-delay. Finally, we

give two comments; one is on an extension of our model to a higher dimensional system and the other is on some computer experiments for the system with global coupling.

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

Chapter 2

Projection-operator approach to overlap dynamics in a Hopfield network

The Hopfield model [5] for neural networks with symmetric connection J_{ij} between two neurons i and j has been extensively studied from the standpoint of (equilibrium) statistical mechanics of phase transitions in disordered spin systems [32]. Typical information which can be afforded by the theory is the number of the patterns which can be stored and retrieved in the network and the rate of errors in the retrieved pattern. Recently dynamical behavior of the model gathers considerable interests in connection with the speed of the association process and the size and structure of the basins of attraction [10, 33, 46]. These questions can only be answered within a nonequilibrium statistical dynamic treatment of the system, which is the main concern in this paper.

Basic evolution of the Hopfield model, consisting of N formal neurons(or Ising spins) is usually governed by the Glauber dynamics [30]

$$\begin{aligned}\frac{\partial p(\vec{S},t)}{\partial t} &= -\frac{1}{\Gamma}\sum_{i=1}^{N}\sum_{s_i'=\pm 1} S_iS_i'f(-S_i'|h_i)p(\vec{S}(S_i'),t) \\ &= L_G p(\vec{S},t) \qquad (2.1)\end{aligned}$$

where $p(\vec{S},t)$ denotes the N-body probability distribution at time t with $\vec{S}=(S_1,\cdots,S_N)$ and $\vec{S}(S_i')=(S_1,\cdots,S_{i-1},S_i',S_{i+1},\cdots,S_N)$. The Γ is the cycle time in which each of the

N neurons updates once on average. The function $f(S_i|h_i)$ represents the probability for the i-th spin to flip into the state S_i in the field h_i, which is given at temperature T by

$$f(S_i|h_i) = \frac{\exp(h_iS_i/T)}{\exp(h_i/T)+\exp(-h_i/T)}, \qquad (2.2)$$

and the field h_i is determined by $N-1$ spins as

$$h_i = \sum_{j(\neq i)} J_{ij}S_i, \qquad (2.3)$$

when there is no external field. The connection J_{ij} is expressed in terms of the p embedded patterns $\vec{\xi}^{\mu} = \{\xi_i^{\mu}\}(i=1,\cdots,N)$ by the Hebbian rule,

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p}\xi_i^{\mu}\xi_j^{\mu}(1-\delta_{ij}). \qquad (2.4)$$

With regard to a retrieval process the most important quantity is the overlap, defined by [1]

$$\hat{m}_{\mu} \equiv \frac{1}{N}\sum_{i=1}^{N}\xi_i^{\mu}S_i \equiv \frac{\hat{M}_{\mu}}{N}, \qquad (2.5)$$

which quantitatively measures similarity between the state $\vec{S}$ of the system and the pattern μ. The retrieval or association process is the one in which one of the $\hat{m}_{\mu}(t)$ approaches one with the other $p-1$ overlaps remaining small(typically of order $N^{-1/2}$). The equilibrium distribution $p_{eq}(\vec{S})$, which satisfies $L_G p_{eq}=0$ is given by

$$p_{eq}(\vec{S}) = \frac{1}{Z}\exp\left[(2NT)^{-1}\sum_{\mu}\hat{M}_{\mu}^2\right], \qquad (2.6)$$

where Z is the partition function of the system. When the number of the patterns p is much smaller than that of neurons N, as is often the case, mapping the Glauber dynamics to the overlap dynamics belongs to a typical problem of deriving a reduced description and this is most effectively performed by utilizing some kind of projection operator method , which has been playing an important role in irreversible statistical mechanics [35, 36, 37].

[1] Later we will introduce $m_{\mu}(M_{\mu})$, which is not a dynamical variable(a function of $\vec{S}$) like $\hat{m}_{\mu}(\hat{M}_{\mu})$

In this Brief Report we apply the Mori-Zwanzig projection operator formalism to the Glauber dynamics (2.1) in order to investigate dynamics of the retrieval process. Before proceeding to the details of the calculation we briefly present the main framework of the formalism in a form suitable for later application. Let us assume that the probability distribution function $p(\vec{x},t)$ is governed by

$$\frac{\partial p(\vec{x},t)}{\partial t} = Lp(\vec{x},t) \tag{2.7}$$

where L is a general time-independent linear operator. Although original formulation is given for the Hamiltonian system [36] and a stochastic system described by a Fokker-Planck operator [37], the L in Eq.(2.7) can be the Glauber operator L_G (2.1) where $\vec{x}$ denotes $\vec{S}$ and integration over $\vec{x}$ in Eq.(2.8) below should be interpreted to be a summation over $\vec{S}$. The operator adjoint to L is defined by

$$\int d\vec{x} f(\vec{x}) Lg(\vec{x}) = \int d\vec{x} g(\vec{x}) \Lambda f(\vec{x}) \tag{2.8}$$

for arbitrary functions $f(\vec{x})$ and $g(\vec{x})$. The operator Λ is seen to control time evolution of dynamical variables, i.e.,

$$\frac{dA(t)}{dt} = \Lambda A(t) \quad \text{or} \quad A(t) = \exp[\Lambda t]A. \tag{2.9}$$

To see the implication of Eq.(2.9) we consider that at time $t=0$ we know that the system is located in the state space at $\vec{x}_0$, that is, $p(\vec{x},t=0) = \delta(\vec{x}-\vec{x}_0)$. Then the expectation of an arbitrary variable $A(\vec{x})$ at time t is calculated as

$$\begin{aligned} \langle A(t)\rangle_{x_0} &= \int d\vec{x} A(\vec{x}) \exp[Lt] \delta(\vec{x}-\vec{x}_0) \\ &= \exp[\Lambda t] A(\vec{x}_0), \end{aligned} \tag{2.10}$$

where now the Λ is an operator working in the space $\vec{x}_0$. Let us consider collective dynamics of a set of dynamical variables $A_i(\vec{x})(i=1,\cdots,I)$. Mutual correlation among

the variables $\{A_i(\vec{x})\}$ is best extracted by introducing a projection operator P onto a space spanned by $\{A_i(\vec{x})\}$, defined by [36, 37]

$$PB(\vec{x}) = \sum_{i,j} (B(\vec{x}), A_i(\vec{x}))(\Xi^{-1})_{ij} A_j(\vec{x}), \tag{2.11}$$

where the innerproduct $(f(\vec{x}), g(\vec{x}))$ of two dynamical variables f and g denotes the equilibrium average $\int d\vec{x} f(\vec{x}) g(\vec{x}) p_{eq}(\vec{x})$ and $\Xi_{ij} \equiv (A_i(\vec{x}), A_j(\vec{x}))$. Applying the operator P on Eq.(2.9) one obtains, what is called, a generalized Langevin(GL) equation [36, 37] ,

$$\frac{dA_i(t)}{dt} = \sum_j \omega_{ij} A_j(t) - \int_0^t ds \Psi_{ij}(t-s) A_j(s) + f_i(t), \tag{2.12}$$

where the frequency and damping matrices are given by

$$\omega_{ij} = \sum_k (\Lambda A_i, A_k)(\Xi^{-1})_{kj}, \tag{2.13}$$

$$\Psi_{ij}(t) = -\sum_k (\Lambda f_i(t), A_k)(\Xi^{-1})_{kj}, \tag{2.14}$$

and

$$f_i(t) = \exp[t(1-P)\Lambda]\{(1-P)\Lambda A_i\} \equiv \exp[t(1-P)\Lambda] f_i. \tag{2.15}$$

Usually calculations of the damping kernel $\Psi(t)$ is prohibitively difficult. However by including all the relevant variables in the set $\{A_i(\vec{x})\}$ we may have rather good description of collective dynamics of the system even if we neglect the damping kernel entirely. For dense gases and liquids [38], for example, by choosing as $\{A_i(\vec{x})\}$ the density in a μ-space, $f_{rp}(\vec{x}) = \sum_n \delta(\vec{r}-\vec{r}_n)\delta(\vec{p}-\vec{p}_n)$, where the suffix i of $A_i(\vec{x})$ now becomes continuous parameters $(\vec{r},\vec{p})$ and $\vec{x} = (\vec{r}_1,\cdots,\vec{r}_N,\vec{p}_1,\cdots,\vec{p}_N)$, we have a kinetic equation for $f_{rp}(t)$ in which the matrix ω gives a Vlasov(mean-field) term and the kernel Ψ represents effects of collisions. From many examples of applications of Eq.(2.12) to various systems [39, 36], it is reasonable to call the GL equation (2.12) without the kernel as the mean-field approximation. Effects of damping could be approximately taken into account as $\Psi(t) \cong \Psi(0)\psi(t)$ where $\psi(t)$ is chosen to be a simple function with some parameter(s) determined by sum-rule arguments. Below we mainly consider the mean-field approximation with a brief comment on the damping kernel.

Now we turn to the problem of overlap dynamics. From Eqs.(2.7) and (2.1) the operator adjoint to L_G, Eq.(2.1), is given by

$$\Lambda_G g(\vec{S}) = \frac{1}{\Gamma}\sum_i f(-S_i|h_i)[g(\vec{S}(-S_i)) - g(\vec{S})]. \tag{2.16}$$

Since we are interested in dynamics of retrieve we take as the collective variables $\{A_i(\vec{S})\}$ the quantity

$$g_{\vec{M}}(\vec{S}) = \prod_{\mu=1}^{p} \Delta(\hat{M}_\mu - M_\mu) \equiv \Delta(\hat{\vec{M}} - \vec{M}), \tag{2.17}$$

where $\Delta(n)$ is defined for integer n with $\Delta(n) = 1$ for $n = 0$ and 0 otherwise. $g_{\vec{M}}(\vec{S})$ represents the event in which the overlap $\hat{M}_\mu(\mu = 1, \cdots, p)$ takes the value M_μ for each μ. The ensemble average $\langle g_{\vec{M}}(\vec{S})\rangle$, which expresses the probability of observing the overlap $\vec{M} = (M_1, \cdots, M_p)$ in the equilibrium state, the correlation $\Xi(\vec{M}, \vec{M}') \equiv \langle g_{\vec{M}}(\vec{S}) g_{\vec{M}'}(\vec{S})\rangle$ and its inverse are given by

$$\langle g_{\vec{M}}(\vec{S})\rangle = \frac{1}{Z}\exp\left[\sum_\mu \frac{M_\mu^2}{2NT}\right]\Omega_N(\vec{M}), \tag{2.18}$$

$$\Xi(\vec{M}, \vec{M}') = \Delta(\vec{M} - \vec{M}')\langle g_{\vec{M}}(\vec{S})\rangle, \tag{2.19}$$

$$\Xi^{-1}(\vec{M}, \vec{M}') = \frac{\Delta(\vec{M} - \vec{M}')}{\langle g_{\vec{M}}(\vec{S})\rangle} \tag{2.20}$$

where $\Omega_N(\vec{M}) = \sum_{\vec{S}} g_{\vec{M}}(\vec{S}) = T_r g_{\vec{M}}(\vec{S})$ is the number of the events, among the 2^N possible microscopic spin configurations, in which $g_{\vec{M}}(\vec{S}) = 1$. From Eqs.(2.16) and (2.17)

$$\begin{aligned} \Lambda_G g_{\vec{M}}(\vec{S}) &= \frac{1}{\Gamma}\sum_i f(-S_i|h_i)\prod_\mu[\Delta(\hat{M}_\mu(i) - \xi_i^\mu S_i - M_\mu) \\ &\quad -\Delta(\hat{M}_\mu(i) + \xi_i^\mu S_i - M_\mu)], \end{aligned} \tag{2.21}$$

where we have introduced $\hat{M}_\mu(i)$ by

$$\hat{M}_\mu(i) = \sum_{j(\neq i)} \xi_j^\mu S_j, \tag{2.22}$$

and the field h_i, Eq.(2.3), is expressed in terms of $\hat{M}_\mu(i)$ as

$$h_i = \frac{1}{N}\sum_\mu \xi_i^{(\mu)}\hat{M}_\mu(i) \equiv \frac{1}{N}\vec{\xi}_i \cdot \hat{\vec{M}}(i), \tag{2.23}$$

where $\vec{\xi}_i$ and $\hat{\vec{M}}(i)$ are p-dimensional vectors $(\xi_i^1, \cdots, \xi_i^p)$ and $(\hat{M}^1(i), \cdots, \hat{M}^p(i))$. The first step to obtain $\omega_{\vec{M},\vec{M}'} \equiv \sum_{\vec{M}''}(\Lambda_G g_{\vec{M}}, g_{\vec{M}''})\ \Xi^{-1}(\vec{M}'', \vec{M}') = (\Lambda g_{\vec{M}}, g_{\vec{M}'})/ < g_{\vec{M}'} >$ (see Eqs.(2.13) and (2.18)) consists in calculating $D_{\vec{M},\vec{M}'} \equiv (\Lambda g_{\vec{M}}, g_{\vec{M}'})$. From Eqs.(2.21) and (2.6) we see that, putting $\gamma \equiv \exp[\sum_\mu M_\mu'^2/2NT]$

$$\begin{aligned} \frac{D_{\vec{M},\vec{M}'}}{\gamma} &= \sum_i \mathrm{Tr}\Big\{\prod_\mu \Delta(M_\mu' - M_\mu - 2\xi_i^\mu S_i) f(-S_i|h_i) \\ &\quad \times\Delta(\hat{M}_\mu(i) - \xi_i^\mu S_i - M_\mu) \\ &\quad -\prod_\mu \Delta(M_\mu' - M_\mu) f(-S_i|h_i)\Delta(\hat{M}_\mu(i) + \xi_i^\mu S_i - M_\mu)\Big\}. \end{aligned} \tag{2.24}$$

We take the trace over $S_i = \pm 1$ and then over the remaining spin variables to obtain

$$\begin{aligned} \frac{D_{\vec{M},\vec{M}'}}{\gamma} &= \sum_i \Big[\Delta(\vec{M}' - \vec{M} - 2\vec{\xi}_i) f\left(-1|h_i = \frac{1}{N}\sum_\mu \xi_i^\mu(M_\mu + \xi_i^{(\mu)})\right)\Omega_{N-1}(\vec{M} + \vec{\xi}_i) \\ &\quad +\Delta(\vec{M}' - \vec{M} + 2\vec{\xi}_i) f\left(1|h_i = \frac{1}{N}\sum_\mu \xi_i^{(\mu)}(M_\mu - \xi_i^{(\mu)})\right)\Omega_{N-1}(\vec{M} - \vec{\xi}_i) \\ &\quad -\Delta(\vec{M}' - \vec{M}) f\left(-1|h_i = \frac{1}{N}\sum_\mu \xi_i^\mu(M_\mu - \xi_i^{(\mu)})\right)\Omega_{N-1}(\vec{M} - \vec{\xi}_i) \\ &\quad -\Delta(\vec{M}' - \vec{M}) f\left(1|h_i = \frac{1}{N}\sum_\mu \xi_i^{(\mu)}(M_\mu + \xi_i^{(\mu)})\right)\Omega_{N-1}(\vec{M} + \vec{\xi}_i)\Big]. \end{aligned} \tag{2.25}$$

From Eqs.(2.19),(2.21) and (2.25) we see that our next task is to calculate

$$p(\vec{\xi}_i; \vec{M}) \equiv \frac{\Omega_{N-1}(\vec{M} - \vec{\xi}_i)}{\Omega_N(\vec{M})}, \tag{2.26}$$

which appears in

$$\omega_{\vec{M},\vec{M}'} = \frac{D_{\vec{M},\vec{M}'}}{\Gamma\gamma\Omega_N(\vec{M}')}. \tag{2.27}$$

From the definition (2.26) $p(\vec{\xi}_i; \vec{M})$ stands for the probability of $S_i = 1$ under the condition $\sum_j S_j \vec{\xi}_j = \vec{M}$. Without loss of generality we take $i = 1$. First we consider the case

$p = 1$, *i.e.*, there is only one pattern. From the condition $\Sigma_j S_j \xi_j^1 = M_1$ we have $S_1 + \Sigma_{j(\neq 1)} \xi_1^1 \xi_j^1 S_j = M_1 \xi_1^1$. From the law of large number, half of $\xi_1^1 \xi_j^1$ take the value 1 and we relabel these j to yield $\sum_{i=1}^{N/2} S_i - \sum_{i=N/2+1}^{N} S_i = M_1 \xi_1^1$ or

$$\sum_{i=1}^{N/2} S_i = M_1 \xi_1^1 / 2. \tag{2.28}$$

Since the probabilistic variables $S_1, \cdots, S_{N/2}$ are equivalent each other we have from Eq.(2.28) $p(\xi_1^1; M_1)N/2 - [1 - p(\xi_1^1; M_1)]N/2 = M_1\xi_1^1/2$, leading to

$$p(\xi_1^{(1)}; M_1) = \frac{1}{2}\left(\frac{1 + M_1\xi_1^1}{N}\right) = \frac{1}{2}(1 + m_1\xi_1^1). \tag{2.29}$$

If there are two patterns, $p = 2$, we have an additional condition $S_1 + \Sigma_{j\neq 1} \xi_1^2 \xi_j^2 S_j = M_2 \xi_1^2$. Using the law of large number again we note that half of the coefficients $\xi_1^2 \xi_j^2 (j \leq N/2)$ take the value 1, which are relabelled again to have $\sum_{i=1}^{N/4} S_i - \sum_{i=N/4+1}^{N/2} S_i = M_2\xi_1^2/2$. From Eq.(2.28) and the above we have

$$\sum_{i=1}^{N/4} S_i = \frac{1}{4}\left(M_1\xi_1^{(1)} + M_2\xi_1^{(2)}\right), \tag{2.30}$$

which gives $p(\vec{\xi}_1; \vec{M}) \times N/4 - [1 - p(\vec{\xi}_1; \vec{M})] \times N/4 = [M_1\xi_1^1 + M_2\xi_1^2]/4$, thus

$$p(\vec{\xi}_1; \vec{M}) = \frac{1}{2}\left(1 + \sum_{i=1}^{2} m_i \xi_1^i\right). \tag{2.31}$$

So long as p is kept finite as N becomes large, we can continue the arguments above to generally obtain

$$p(\vec{\xi}_i; \vec{M}) = \frac{1}{2}\left(1 + \sum_{\mu=1}^{p} m_\mu \xi_i^{(\mu)}\right). \tag{2.32}$$

The mean-field term $\Sigma_j \omega_{ij} A_j$ in Eq.(2.12) is now ready to calculate, from Eqs.(2.27) and (2.32) leading to

$$\begin{aligned}
\sum_{\vec{M}'} \omega_{\vec{M},\vec{M}'} g_{\vec{M}'} = & \frac{1}{\Gamma}\sum_i \left[g_{\vec{M}+2\vec{\xi}_i} f\left(-1 | h_i = \frac{1}{N}\sum_\mu \xi_i^{(\mu)}(M_\mu + \xi_i^{(\mu)})\right)\right. \\
& \left. \times p(\vec{\xi}_i; \vec{M} + 2\vec{\xi}_i) - (\vec{M} \to \vec{M} - 2\vec{\xi}_i)\right] \\
& - \frac{1}{\Gamma}\sum_i \left[g_{\vec{M}} f\left(1 | h_i = \frac{1}{N}\sum_\mu \xi_i^\mu(M_\mu + \xi_i^{(\mu)})\right)\right. \\
& \left. \times [1 - p(\vec{\xi}_i; \vec{M})] - (\vec{M} \to \vec{M} - 2\vec{\xi}_i)\right],
\end{aligned} \tag{2.33}$$

where $(\vec{M} \to \vec{M} - 2\vec{\xi}_i)$ means that we replace all the $\vec{M}$ appearing on the left hand side by $\vec{M} - 2\vec{\xi}_i$. The remaining task is to represent the result in terms of the overlap $\{m_\mu\}$ rather than $\{M_\mu\}$. Since $\sum_{\vec{M}} g_{\vec{M}} = 1$ we define $g(\vec{m}) \equiv N^p g_{\vec{M}}$ which satisfies $\int d\vec{m} g(\vec{m}) = 1$. After multiplying N^p on both sides of Eq.(2.33) we use the relation

$$v(\vec{m}) - v\left(\vec{m} - 2\frac{\vec{\xi}_i}{N}\right) = \sum_\mu \frac{2\xi_i^\mu}{N}\frac{\partial v(\vec{m})}{\partial m_\mu} - \sum_{\mu\nu} \frac{1}{2}\frac{2\xi_i^\mu}{N}\frac{2\xi_i^\nu}{N}\frac{\partial^2 v(\vec{m})}{\partial m_\mu \partial m_\nu} + O(N^{-3}) \tag{2.34}$$

to derive finally

$$\frac{\partial g(\vec{m}, t)}{\partial t} = \frac{1}{N\Gamma}\left[\sum_{i,\mu} \xi_i^\mu \frac{\partial}{\partial m_\mu}\left\{\vec{m}\cdot\vec{\xi}_i - \tanh\left(\frac{\vec{m}\cdot\vec{\xi}_i}{T}\right)\right\} g(\vec{m}, t) + 0(N^{-1})\right], \tag{2.35}$$

where the order N^{-1} correction is given by

$$\frac{1}{N}\sum_{i,\mu,\nu} 2\xi_i^\mu \xi_i^\nu \frac{\partial}{\partial m_\mu}\left[(1 - \tanh(\vec{m}\cdot\vec{\xi}_i))\frac{\partial}{\partial m_\nu}\frac{(1 + \vec{m}\cdot\vec{\xi}_i)g(\vec{m}, t)}{4}\right]. \tag{2.36}$$

Here we give two comments on the GL Eq.(12) with regard to its application to overlap dynamics. The first one is concerned with the damping kernel. For the one-pattern case($p = 1$), which is equivalent to infinite range ferromagnitic system [30], we have calculated the kernel $\Psi(t = 0)$ to the lowest order in N^{-1} to find that it exactly vanishes. This strongly suggests that $\Psi(t)$ vanishes for $t \geq 0$ and $p \geq 1$ to the lowest order in N^{-1}. Secondly, since the random force $f_{\vec{M}}(t)$, which is given as $f_i(t)$ in Eq.(12), is orthogonal to $\{g_{\vec{M}}\}$, we can safely neglect it in discussing the probability distribution function based on the GL Eq.(12) [37].

Now from Eq.(2.35) it is seen that the average $\langle m_\mu(t)\rangle \equiv \int d\vec{m} m_\mu g(\vec{m}, t)$ follows the equation

$$\frac{d\langle m_\mu(t)\rangle}{dt} = \frac{1}{\Gamma}\left[-\langle m_\mu(t)\rangle + \left\langle\!\!\left\langle \xi^\mu \tanh\left(\frac{\langle\vec{m}(t)\rangle\cdot\vec{\xi}}{T}\right)\right\rangle\!\!\right\rangle\right], \tag{2.37}$$

where $\langle\langle\ \rangle\rangle$ denotes the average over the patterns according to the probability law $p(\xi^{(\mu)} = 1) = 1/2$. In deriving Eq.(2.37) we have replaced $\langle\tanh(\vec{m}\cdot\vec{\xi}/T)\rangle$ by $\tanh(\langle\vec{m}\rangle\cdot\vec{\xi}/T)$, which is consistent with our mean-field approximation in the limit $N \to \infty$ because in this limit we have only convection of probability and no conduction (no fluctuations), Eq.(2.35).

Equation (2.37) has been derived, so long as the authors are aware, by two methods. One is based on the notion of sublattice magnetization [10, 40, 41] and the other is based on a path-integral formulation of spin dynamics [46]. The sublattice idea [8] is elegant enough to be applied to more general Hopfield model. The merit of our mean-field approach is that it is besed on a general method of irreversible statistical mechanics and thus sheds some light on the implication of the overlap-dynamics equations (2.35) and (2.37). The remaining important problem is the case of extensively many patterns, that is, finite $\alpha \equiv p/N$. Overlap dynamics in this case is at present mainly investigated by computer simulations, except for a few theoretical works [33, 46, 9, 42]. With inclusion of some additional variables in the set of collective variables $\vec{A}$, it is hoped that overlap dynamics for a finite α case could be handled within the mean-field approximation developed in this paper.

Chapter 3

Correlated-data-driven dynamics in a neural network

A number of neural network models have been proposed to explain some properties of the nervous systems in terms of the formal two-state neurons [43]. In the Hopfield model [5] with N neurons, the symmetric connection J_{ij} between the neurons i and j is expressed as

$$J_{ij} = \frac{1}{N} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu}, \tag{3.1}$$

where $\xi_i^{\mu}(= \pm 1)(\mu = 1, \ldots, p;\ i = 1, \ldots, N)$ denotes the state of the neuron i in the μ-th pattern. The postsynaptic potential (PSP) or the field strength $h_i(t)$ at the neuron i is expressed as

$$h_i(t) = \sum_j J_{ij} S_j(t) = \sum_{\mu} \xi_i^{\mu} m_{\mu}(t), \tag{3.2}$$

where the overlap $m_{\mu}(t)$ with the μ-th pattern is defined by

$$m_{\mu}(t) = \frac{1}{N} \sum_i \xi_i^{\mu} S_i(t). \tag{3.3}$$

The collective behavior of the network with a symmetric connection, such as Eq.(3.1) is represented as a relaxation process toward the local minimum of the (free-) energy and the model works as content-addressable or associative memories [5].

Recently issues of temporal association in neural networks have gathered considerable attention [17, 16, 10, 9]. To effect transitions between patterns, a certain amount of asymmetry of the connection J_{ij} is required. After the suggestive proposal by Hopfield [5], some models have been put forth in which delay in signal transmission is introduced to yield, e.g.,

$$h_i(t) = \sum_\mu \xi_i^\mu m_\mu(t) + \varepsilon \sum_\mu \xi_i^{\mu+1} m_\mu(t-\tau). \tag{3.4}$$

For later convenience we briefly consider the transition mechanism within the framework of asynchronous dynamics [9]. We suppose that the system has been in the pattern ν for $0 \le t < \tau$. As for $m_\mu(t-\tau)$ in Eq.(3.4), we can put $m_\mu(t-\tau) = C_{\mu\nu}$ for $\tau \le t < 2\tau$ where the correlation $C_{\mu\mu'}$ between the patterns μ and μ' are defined by

$$C_{\mu\mu'} = \frac{1}{N}\sum_i \xi_i^\mu \xi_i^{\mu'}. \tag{3.5}$$

If we consider a network in which the correlations vanish on average, that is,

$$\overline{C}_{\mu\mu'} = \frac{1}{N}\sum_i \langle \xi_i^\mu \xi_i^{\mu'} \rangle = \delta_{\mu\mu'}, \tag{3.6}$$

and neglect fluctuations in $C_{\mu\mu'}$ of order $1/\sqrt{N}$, we have $m_\mu(t-\tau) = \delta_{\mu\nu}$ for $\tau \le t < 2\tau$. If ε is chosen to be larger than one and τ sufficiently longer than one cycle time (or one Monte Carlo time (MCT) in which N neurons are updated), we see that $m_{\nu+1}(t)$ increases until it becomes one at about $t \simeq \tau + \mathrm{MCT}$. During the time other overlaps $m_\mu(t)(\mu \ne \nu+1)$ decrease to zero. Thus after the pattern ν, the pattern $\nu+1$ is retrieved and then the pattern $\nu+2$ and so on. It should be noted that once $m_{\nu+1}(t)$ becomes large the first term on the right hand side(rhs) of Eq.(3.4) contributes to stabilization of the pattern $\nu+1$.

The model (3.4) and the modifications thereof have been successful in producing a pattern sequence [17, 16, 10]. We note however that the correlations among patterns $\overline{C}_{\mu\mu'}(\mu \ne \mu')$ are assumed to be zero, Eq.(3.6), in order to enable the network to retrieve the prescribed sequence $\nu, \nu+1, \ldots$. Putting the model (3.4) aside for a while, we consider what happens when there are strong correlations among patterns in a network(brain). If

the network is in a state(pattern) A and the pattern B alone is strongly correlated with the pattern A, it is expected that a spontaneous transition, if it occurs, is to the state B. That is, one can conceive that the internal correlations among the embedded patterns(data) can give rise to spontaneous or data-driven transitions. The purpose of this letter is to present and investigate a model which materialize the notion above.

Our model consists in the PSP given by

$$h_i(t) = \sum_\mu \varepsilon_\mu \xi_i^\mu m_\mu(t) + \sum_{\substack{\mu,\nu \\ \mu\ne\nu}} \varepsilon_\mu^T \xi_i^\mu m_\mu(t) m_\nu(t-\tau). \tag{3.7}$$

The correlations $\overline{C}_{\mu\mu'}$ defined by Eq.(3.6) now take non-zero values. Qualitatively the second(transition) term on the rhs of Eq.(3.7) are interpreted as follows: When the network is in the pattern ν_1 for $0 \le t < \tau$, $m_\nu(t-\tau)$ in Eq.(3.7) is $C_{\nu\nu_1}$ for $\tau \le t < 2\tau$. In the summation over ν in Eq.(3.7) the largest contribution comes from the term $\nu = \nu_1$ since $C_{\nu\nu_1} = 1$ for $\nu = \nu'$. By $[\nu_1]$ let us denote the pattern($\ne \nu_1$) which is most strongly correlated with the pattern ν_1. Now for t slightly larger than $\tau(t \ge \tau)$, $m_\mu(t)$ in Eq.(3.7) is nearly equal to $C_{\mu\nu_1}$. Since μ must be different from ν_1, the largest contribution in the summation over μ comes from the term $\mu = [\nu_1]$ and the PSP tends to drive the system toward the state $[\nu_1]$. Once this transition is initiated, the first term also begins supporting the transition. Of course this is a drastically simplified scenario and more quantitative argument is necessary to understand the transition mechanism in Eq.(3.7).

For the purpose we confine ourselves to the case where only three patterns A, B, and C are embedded in the network. The PSP is explicitly written down as

$$\begin{aligned} h_i(t) &= [\varepsilon_A m_A + \varepsilon_A^T m_A(\overline{m}_B + \overline{m}_C)]\xi_i^A \\ &\quad + [\varepsilon_B m_B + \varepsilon_B^T m_B(\overline{m}_A + \overline{m}_C)]\xi_i^B \\ &\quad + [\varepsilon_C m_C + \varepsilon_C^T m_C(\overline{m}_A + \overline{m}_B)]\xi_i^C, \end{aligned} \tag{3.8}$$

where ε_μ and $\varepsilon_\mu^T(\mu = A, B, C)$ are assumed to be positive and m_A and $\overline{m}_A$ denote $m_A(t)$ and $m_A(t-\tau)$, respectively. Suppose that at $t = 0$ the system is in a state A and the

PSP at the neuron i becomes

$$h_i(0) = \varepsilon_A \xi_i^A + \varepsilon_B C_{AB} \xi_i^B + \varepsilon_C C_{AC} \xi_i^C, \tag{3.9}$$

where the transition term in Eq.(3.7) is regarded to be effective after $t = \tau$. If

$$\varepsilon_A \geq \varepsilon_B C_{AB} + \varepsilon_C C_{AC}, \tag{3.10}$$

the system remains in the state A until $t = \tau$. At $t = \tau$, the PSP is given from Eq.(3.8) by putting $m_\mu = C_{\mu A} = \overline{m}_\mu$. Explicitly it reads as

$$\begin{aligned} h_i(t) &= [\varepsilon_A + \varepsilon_A^T (C_{AB} + C_{AC})]\xi_i^A \\ &\quad + [\varepsilon_B C_{AB} + \varepsilon_B^T C_{AB}(1 + C_{AC})]\xi_i^B \\ &\quad + [\varepsilon_C C_{AC} + \varepsilon_C^T C_{AC}(1 + C_{AB})]\xi_i^C \\ &\equiv h_A \xi_i^A + h_B \xi_i^B + h_C \xi_i^C. \end{aligned} \tag{3.11}$$

We now require that

$$h_B > h_A \quad \text{and} \quad h_B > h_C. \tag{3.12}$$

Under the condition (3.12), if $\xi_i^A = \xi_i^B$, $S_i (= \xi_i^A)$ is unchanged and if $\xi_i^B = \xi_i^C \neq \xi_i^A$, S_i flips from ξ_i^A to $-\xi_i^A = \xi_i^B$. Since we do not require $h_B > h_A + h_C$, which is too severe a condition to be satisfied, when $\xi_i^A = \xi_i^C \neq \xi_i^B$, it seems that the transition from $S_i = \xi_i^A$ to ξ_i^B is not assured. However as explained above, in the process of asynchronous updating, $S_i(t)$ takes the value ξ_i^B except for the neuron site, where $\xi_i^A = \xi_i^C \neq \xi_i^B$ and the coalition of A and C patterns resists the transition to the state B. At this point let us express the fraction Z of the coalition sites in terms of the correlations. If X, Y, Z and $(1 - X - Y - Z)$ denote the fractions of neuron sites at which $\xi_i^A = \xi_i^B = \xi_i^C$, $\xi_i^A = \xi_i^B \neq \xi_i^C$, $\xi_i^A = \xi_i^C \neq \xi_i^B$ and $\xi_i^B = \xi_i^C \neq \xi_i^A$, respectively, we have

$$C_{AB} = 2(X + Y) - 1, \quad C_{AC} = 2(X + Z) - 1, \quad C_{BC} = 1 - 2(Y + Z). \tag{3.13}$$

These equations lead to

$$Z = (1 + C_{AC} - C_{BC} - C_{AB})/4. \tag{3.14}$$

From the discussions above we see that the overlap m_B increases from C_{AB} at time $t = \tau$ to $m_B = 1 - 2Z = (1 + C_{BC} + C_{AB} - C_{AC})/2$ at $t = \tau + \text{MCT}$ and the system approaches the state B considerably as exemplified later by our numerical simulation. We now write down the conditions to be satisfied in order to keep the state B after $t = \tau + \text{MCT}$, Eq.(3.15), to effect the transition to the state C after $t = 2\tau$, Eq.(3.16), to keep the state C after $t = 2\tau + \text{MCT}$, Eq.(3.17) and finally to keep the state C forever($t > 3\tau$), Eq.(3.18).

$$\varepsilon_B + \varepsilon_B^T + \varepsilon_B^T C_{AC} \geq \varepsilon_A C_{AB} + \varepsilon_A^T (C_{AB})^2 + \varepsilon_A^T C_{AB} C_{AC} + \varepsilon_C C_{BC} + \varepsilon_C^T C_{BC} + \varepsilon_C^T C_{BC} C_{AB} \tag{3.15}$$

$$\varepsilon_C C_{BC} + \varepsilon_C^T C_{AB} C_{BC} + \varepsilon_C^T C_{BC} > \varepsilon_A C_{AB} + \varepsilon_A^T C_{AB} + \varepsilon_A^T C_{AB} C_{BC}, \quad \varepsilon_B + \varepsilon_B^T C_{AB} + \varepsilon_B^T C_{BC} \tag{3.16}$$

$$\varepsilon_C + \varepsilon_C^T C_{AB} + \varepsilon_C^T \geq \varepsilon_A C_{AC} + \varepsilon_A^T C_{AC} + \varepsilon_A^T C_{AC} C_{BC} + \varepsilon_B C_{BC} + \varepsilon_B^T C_{BC} C_{AB} + \varepsilon_B^T (C_{BC})^2 \tag{3.17}$$

$$\varepsilon_C \xi_i^C + \varepsilon_C^T C_{AC} + \varepsilon_C^T C_{BC} \geq \varepsilon_A C_{AC} + \varepsilon_A^T C_{AC} C_{BC} + \varepsilon_A^T C_{AC} + \varepsilon_B C_{BC} + \varepsilon_B^T C_{BC} C_{AC} + \varepsilon_B^T C_{BC} \tag{3.18}$$

If the conditions (3.10), (3.12) and (3.15)-(3.18) are satisfied we would have the sequence $A \to B \to C$, although the transition may be not perfect as noted in connection with the requirement Eq.(3.12)(see Eq.(3.16) also).

We now turn to computer simulations of the model (3.7). The stochastic Glauber dynamics is employed, in which the probability for the neuron i to take ± 1 at time $t + \Delta t$ is given by [30, 1]

$$\text{Prob}\{S_i(t + \Delta t) = \pm 1\} = \{1 \pm \tanh(h_i(t)/T)\}/2. \tag{3.19}$$

When the temperature of the system T goes to zero, the Glauber dynamics is reduced to the threshold dynamics

$$S_i(t + \Delta t) = \text{sgn}\{h_i(t)\}, \tag{3.20}$$

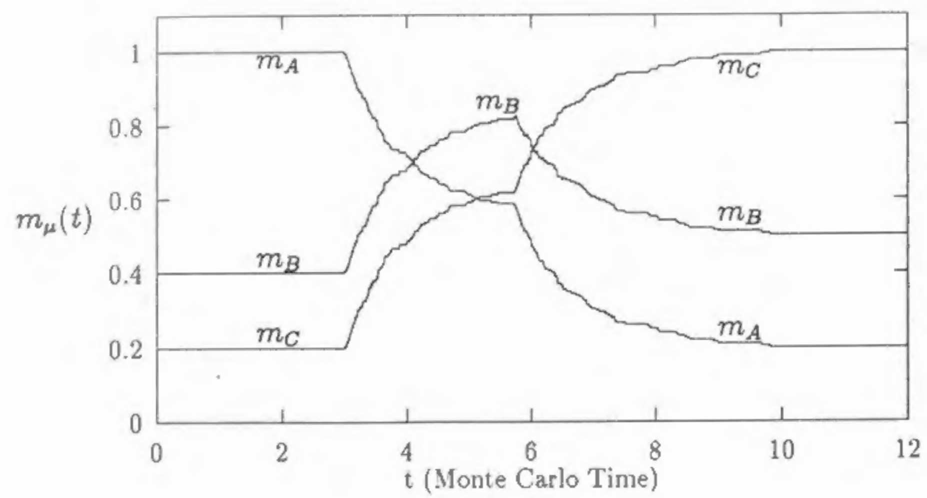

Figure 3.1: Simulation results for the overlap dynamics. $T = 0$, $C_{AB} = 0.4$, $C_{BC} = 0.5$, $C_{AC} = 0.2$, $\varepsilon_A = 0.6$, $\varepsilon_B = 0.8$, $\varepsilon_C = 1.0$, $\varepsilon_A^T = 1.0$, $\varepsilon_B^T = 1.9$ and $\varepsilon_C^T = 3.0$.

with $\mathrm{sgn}(x) = 1$ for $x \geq 0$ and $\mathrm{sgn}(x) = -1$ for $x < 0$. Three patterns A, B and C are embedded in the system consisting of $N = 400$ neurons and the delay τ is taken to be 3 MCT. When three correlations C_{AB}, C_{BC} and C_{AC} are given, we solve Eq.(3.13) to obtain X, Y and Z. We first produce the pattern, say A, by using N random numbers. Then the pattern B is determined by either taking ξ_i^B equal to ξ_i^A for the X and Y regions and equal to $-\xi_i^A$ for the remaining region. The pattern C is made in the same way.

Figure 1 Figure 3.1 shows the overlap dynamics for $T = 0$ and the parameters $\varepsilon_A = 0.6$, $\varepsilon_B = 0.8$, $\varepsilon_C = 1.0$, $\varepsilon_A^T = 1.0$, $\varepsilon_B^T = 1.9$, $\varepsilon_C^T = 3.0$, $C_{AB} = 0.4$, $C_{BC} = 0.5$ and $C_{AC} = 0.2$. Although the pattern B is not fully retrieved in the intermediate time region, we see that a transition sequence $A \to B \to C$ is realized with the parameters given above, which satisfy all the requirements listed before. The asymptotic ($t = \infty$) value of each overlap is seen to be consistent with the correlations in our data. We performed a simulation on a system with three uncorrelated patterns (other parameters are the same with the above)

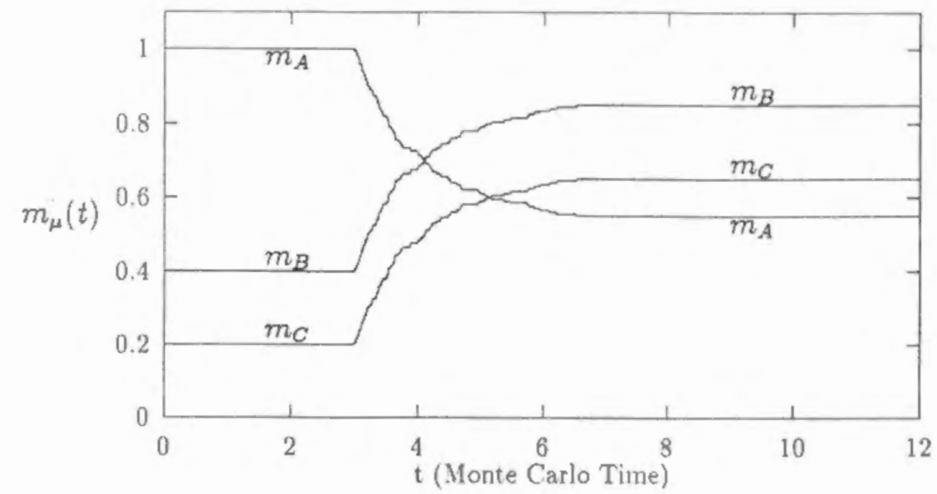

Figure 3.2: Simulation results for the overlap dynamics. ε_B^T is 2.0 and other conditions are the same with those for Fig.3.1.

and observed that all the overlaps remain constant, showing no sign of transitions. To see effects of finite temperature, we first change ε_B^T from 1.9 to 2.0, keeping other parameters as in Fig.3.1. In this case the requirement Eq.(3.16) is not satisfied and reflecting this, we notice in Fig.3.2 that the sequence retrieval is incomplete. Now we change T from zero to 0.06 and the result is shown in Fig.3.3, which reveals the familiar fact that fluctuations induced by temperature work positively for retrieval of patterns so long as it is not too high [30].

Here we give two comments on our model. The first one is concerned with the robustness of our scheme to the variation of the parameter values. To be concrete we change ε_B^T and ε_C^T with all the other parameters fixed as above. First we consider the case $T = 0$. If we put $\varepsilon_C^T = 3.0$ (the case of our simulation), ε_B^T should be in the range $1.9 \leq \varepsilon_B^T < 2.0$ in order to satisfy the conditions (3.10), (3.12) and (3.15) - (3.18). On the other hand, when we put $\varepsilon_C^T = 4.0$, ε_B^T is allowed to be in a little wider range $2.483 \leq \varepsilon_B^T < 2.778$.

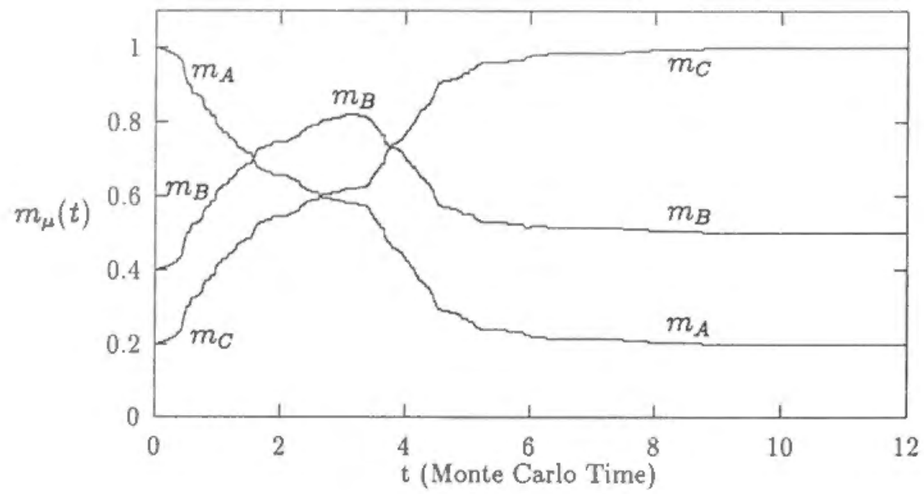

Figure 3.3: Simulation results for the overlap dynamics. $T = 0.06$ and other conditions are the same with those for Fig.3.2.

When the temperature is nonzero, the range of ε_B^T for successful retrieve of the state C becomes large. For example, a retrieval process $A \to B \to C$ is confirmed for the case $\varepsilon_B^T = 2.5$ and $\varepsilon_C^T = 3.0$. However it is remarked that when T is very small (e.g. 0.01) one must wait a longer time for the retrieve of the state C and that when T is too large (e.g. 0.3) the state B is skipped, resulting in the retrieval process $A \to C$.

The second comment is on the number of patterns embedded in our system. When a group of correlated patterns in a sequence increases its members (in our simulations only three members A, B, C), the tuning of the parameters, including the correlations $\{C_{\mu\nu}\}$, naturally becomes severe. However we can increase the number of patterns in our system by embedding many groups, each containing three patterns which are uncorrelated with patterns belonging to all the other groups. For example, in a system consisting of 60 neurons, we could embed 3 groups of correlated patterns ($(A_i, B_i, C_i), i = 1, 2, 3$).

Finally we consider overlap dynamics theoretically in the thermodynamic limit $N \to$

∞ with the number of the patter p kept finite $(p/N \to 0)$ [8, 25]. Our analysis is based on the notion of sublattice $I(\vec{x})$ and sublattice magnetization $m(\vec{x}, t)$, which are due to van Hemmen and his coworkers [8, 9]. Given p binary patterns, we have N vectors $\vec{\xi}_i = (\xi_i^1, \ldots, \xi_i^p)(i = 1, \ldots, N)$. Introducing the sublattice by

$$I(\vec{x}) = \{i; \vec{\xi}_i = \vec{x}\}, \quad \vec{x} \in \{-1, 1\}^p, \tag{3.21}$$

and the sublattice magnetization by

$$m(\vec{x}, t) = \frac{1}{|I(\vec{x})|} \sum_{i \in I(\vec{x})} S_i(t), \tag{3.22}$$

with $|I(\vec{x})|$ denoting the size of $I(\vec{x})$, one can easily verify that the PSP, Eq.(3.7), is expressed as

$$h_i(t) = \sum_{\mu, \vec{x}} \xi_i^\mu p_N(\vec{x}) x_\mu m(\vec{x}, t) \{\varepsilon_\mu + \varepsilon_\mu^T \sum_{\substack{\nu, \vec{y} \\ \nu \neq \mu}} p_N(\vec{y}) y_\nu m(\vec{y}, t - \tau)\}. \tag{3.23}$$

Here $p_N(\vec{x}) = |I(\vec{x})|/N$ denotes the weight of the sublattice $I(\vec{x})$, which depends on the correlations among the patterns. Since $\vec{\xi}_i = \vec{x}$ for all i in $I(\vec{x})$, the PSP depends on i only through the label $\vec{x}$ of the sublattice to which i belongs. Thus Eq.(3.23) can be rewritten as

$$h(\vec{x}, t) = \sum_{\mu, \vec{y}} x_\mu p_N(\vec{y}) y_\mu m(\vec{y}, t) \{\varepsilon_\mu + \varepsilon_\mu^T \sum_{\substack{\nu, \vec{z} \\ \nu \neq \mu}} p_N(\vec{z}) z_\nu m(\vec{z}, t - \tau)\}. \tag{3.24}$$

It is well-known that in the thermodynamic limit the sublattice magnetization $m(\vec{x}, t)$ is governed by the following set of 2^p coupled differential equations [9]:

$$\frac{dm(\vec{x}, t)}{dt} = -\{m(\vec{x}, t) - \tanh[h(\vec{x}, t)/T]\}. \tag{3.25}$$

Since the overlap $m_\mu(t)$ is expressed in terms of $m(\vec{x}, t)$ as $m_\mu(t) = \sum_{\vec{x}} p(\vec{x}) x_\mu m(\vec{x}, t) = \langle x_\mu m(\vec{x}, t) \rangle$, we finally obtain a set of p coupled equations for the overlaps,

$$\frac{dm_\mu(t)}{dt} = -m_\mu(t) + \left\langle x_\mu \tanh\left[\frac{1}{T} \sum_\nu x_\nu m_\nu(t) \left\{\varepsilon_\nu + \varepsilon_\nu^T \sum_{\nu'(\neq \nu)} m_{\nu'}(t - \tau)\right\}\right]\right\rangle. \tag{3.26}$$

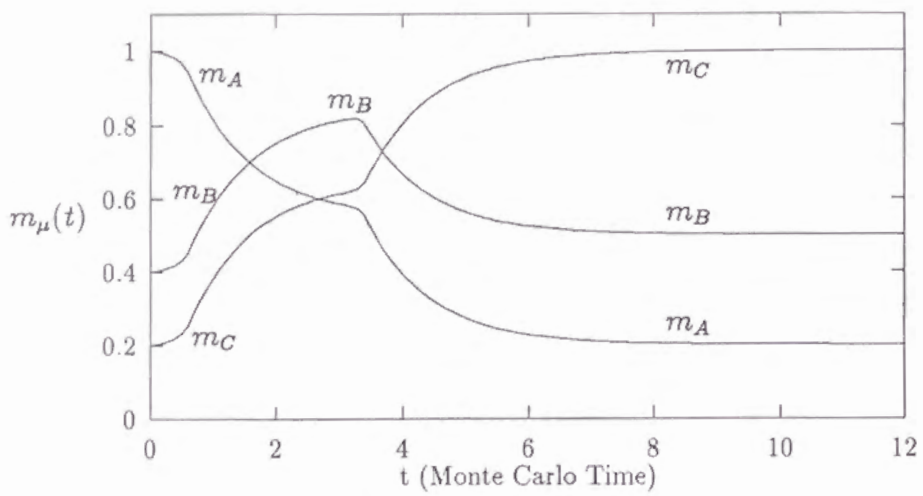

Figure 3.4: Numerical solutions of Eq.(3.26). The conditions are the same with those for Fig.3.3.

The second term on the rhs of Eq.(3.26) consists of 2^p terms, since there are 2^p possibilities for $\vec{x}$. We solve Eq.(3.26) with several values of T and the solutions to Eq.(3.26) are confirmed to be similar to our simulations. In Fig.3.4 we show the numerical results for the overlaps obtained from Eq.(3.26). The parameters are the same with those for Fig.3.3. From Figs.3.3 and 3.4 it is seen that Eq.(3.26) reproduces our experimental results for $N = 400$ and $p = 3$ excellently.

In this letter we investigated neural networks with correlated patterns. For the dynamics we proposed Eq.(3.7) as a model for sequential association. Concerning this model we obtained the following results:

a)With the appropriate values for the parameters $\varepsilon_\mu, \varepsilon_\mu^T$ the sequence was generated and finally the system settled in a target pattern. Effects of temperature turned out to be supportive for the retrieval of the sequence.

b)Our model, Eq.(3.7), itself does not induce transitions as stressed in connection with

the discussions on Fig.3.1. The model together with the correlations among the patterns make the sequence retrieval possible.

c)Coupled nonlinear differential equations (3.26) were derived for the model (3.7). Here also, the correlations among the patterns is implicit. Numerical solution to these equations are well-correlated with the simulation results.

Chapter 4

A neural network model composed of multi-dimensional spin neurons

4.1 Introduction

In the last decade, there has been a great deal of research on a neural network as an associative memory. Various properties, for example a storage capacity of a network, were studied with use of techniques of statistical mechanics of spin systems [29, 32, 45, 30, 1]. Most of these studies, however, treated Ising type neurons which represent only two states, firing and resting ones.

Recently there is a growing interest in neural networks with multi-state neurons. The merit of the model is that one neuron can express a complex state such as a color or shade of grey of each pixel in the pattern which can only be expressed with some Ising type neurons

Rieger [46] and Bollé et al. [47, 48] used neurons represented by spin variables $S_i(i = 1, \ldots, N)$ which can take Q values,

$$S_i = -1 + \frac{2(k-1)}{Q-1}, \qquad (i = 1, \ldots, N; k = 1, \ldots, Q). \tag{4.1}$$

Kanter proposed the model composed of Potts-neurons with q possible discrete states [49]. The dynamics of the Potts neural network is very complicated (see refs [49, 50]).

Other possibility for a multi-state neuron is the so-called circular representation in

which the state of the neuron is represented by points on the circle. The property of this model is that the state of each neuron can be expressed by the phase variable. We show here two examples of such models; one is the phasor model proposed by Noest [28, 51] and the other is the clock model proposed by Cook [27] and will discuss in Sec.2 how these models are similar to and different from our model.

Noest discussed a phasor network composed of unit-length 2-dimensional vectors (phasors) as neurons. The network has N phasors $S_i(i = 1, \ldots, N)$ which are complex numbers with $|S_i| = 1$. When the i-th phasor in the μ-th pattern is described as $\xi_i^\mu(i = 1, \ldots, N; \mu = 1, \ldots, p)$, synaptic couplings C_{ij} are defined as

$$C_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \bar{\xi}_j^\mu (1 - \delta_{ij}) \tag{4.2}$$

in order to store patterns. Note that δ_{ij} is Kronecker's delta function and $\bar{x}$ represents the complex conjugate of a complex x. Since ξ_j^μ can be expressed as $e^{i\theta_j^\mu}$, storing patterns is equivalent to storing phases. The dynamics of each phasors depends on the local (complex) field

$$h_i = \sum_j C_{ij} S_j. \tag{4.3}$$

For a discrete-time updating the state of i-th phasor at next step is defined as

$$S_i(t + \delta t) = \frac{h_i}{|h_i|}, \tag{4.4}$$

and for a continuous-time updating the i-th phasor evolves as

$$\frac{dS_i}{dt} = h_i - \bar{h}_i S_i^2. \tag{4.5}$$

In the network with an asynchronous discrete-time or a continuous-time updating, an energy can be defined as

$$E = -\frac{1}{2} \sum_{i \neq j} \bar{S}_i C_{ij} S_j, \tag{4.6}$$

which plays a role a Lyapunov-function. Noest studied this network with diluted synaptic connections. The other case of full connection was also studied by Gerl et al. [52].

Cook investigated a Q-state clock neural network model. The Hamiltonian of the system is

$$H = -\frac{1}{2N}\sum_{i\neq j}\sum_{\mu}\cos\frac{2\pi}{Q}\left\{(n_i-\xi_i^{\mu})-(n_j-\xi_j^{\mu})\right\}, \tag{4.7}$$

where N is the number of neurons, $n_i(=0,1,\ldots,Q-1)(i=1,\ldots,N)$ the state of i-th neuron and $\xi_i^{\mu}(=0,1,\ldots,Q-1)(i=1,\ldots,N;\mu=1,\ldots,p)$ the state of i-th neuron in the μ-th pattern stored in the network. This system is reduced to the Hopfield model [5] for $Q=2$ and to x-y spin systems for $Q=\infty$, where the state becomes continuous. It is to be noted that the Hamiltonian is invariant under the transformation $\{n_i\}\to\{n_i+k\ (\mathrm{mod}Q)\}$, where k is an integer. Hence, if a pattern, say $\{\eta_i\}$, is stored, the $Q-1$ related configurations, $\{\eta_{k_i}\}$, are also stored, where

$$\eta_{k_i} = \eta_i + k \ \ \mathrm{mod}Q \quad (i=1,\ldots,N;k=1,\ldots,Q-1). \tag{4.8}$$

Therefore notice that only the set of phase differences is meaningful as the information to be stored. Cook analyzed this model with the replica symmetric theory and estimated the storage capacity as $\alpha_c=0.038$ in the limit $Q=\infty$.

Noest's and Cook's model can be regarded as an extension of the Hopfield model to the clock type one. In this paper we want to consider a more general extension of the Hopfield model using D-dimensional spins (i.e. D-dimensional unit vectors) as neurons. This model is expected to realize a new type of neural network with multi-state neurons.

This paper is organized as follows: In section 2 we propose our neural network model composed of multi-dimensional spin neurons. Simulation results are shown in section 3. In section 4 we analyze our model theoretically using a replica symmetric theory. We derive the free-energy of our model near saturation. On the basis of this results, we calculated and discussed the phase diagram and the storage capacity. In the last section we summarize our results.

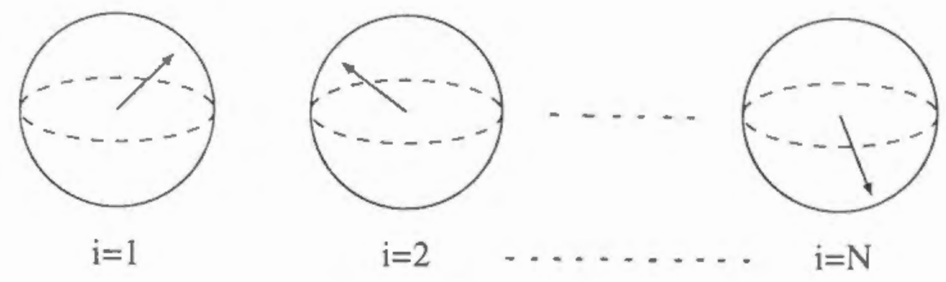

Figure 4.1: An illustration of neurons in the case of $D=3$.

4.2 Model

We consider a network composed of N neurons which are described by D-dimensional unit vectors $\boldsymbol{x}_i = {}^t(x_{i(1)},x_{i(2)},\ldots,x_{i(D)})$ $(i=1,\ldots,N)$, where ${}^t\boldsymbol{z}$ denotes the transposed vector of $\boldsymbol{z}$. Each neuron represents an arbitrary point on a surface of D-dimensional unit sphere. Neurons with $D=3$ are illustrated in Fig.4.1. Especially in the case of $D=2$, this model neuron is similar to that of Cook ($Q=\infty$) and of Noest. The neurons are interconnected with all the others through synaptic couplings. In this paper, we define synaptic couplings by an extended Hebb's rule as

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{p}\xi_i^{\mu}\,{}^t\xi_j^{\mu} \quad (i\neq j), \qquad J_{ii}=O, \tag{4.9}$$

where ξ_i^{μ} denotes the state of i-th neuron in the μ-th memorized pattern $(\mu=1,\ldots,p)$ and is also a D-dimensional unit vector. Note that a synaptic strength J_{ij} is here a $D\times D$ matrix.

There are two kinds of dynamics, one is under the zero temperature $(T=0)$ and the other under the finite temperature$(T\neq 0)$. First we consider the case $T=0$. We calculate the local field $\boldsymbol{h}_i(t)$ on the i-th neuron at time t as

$$\boldsymbol{h}_i(t) = \sum_{j=1}^{N} J_{ij}\boldsymbol{x}_j(t). \tag{4.10}$$

$\boldsymbol{h}_i(t)$ is a D-dimensional and in general non-unit vector. The state of the neuron i at the next time step is determined by normalizing $h_i(t)$ as

$$x_i(t+\Delta t) = \frac{h_i(t)}{|h_i(t)|}. \tag{4.11}$$

As to the way of updating, here we adopt an asynchronous one, for which we can define an energy function

$$E = -\frac{1}{2}\sum_{i\neq j} {}^t x_i J_{ij} x_j. \tag{4.12}$$

This energy is a Lyapunov function for the dynamics. This is seen as follows: Suppose one of the neuron, say i-th neuron, updates as $x_i(t) \to x_i(t+\Delta t)$, then the variation of energy ΔE is calculated as

$$\Delta E = -|h_i(t)|\{1 - {}^t x_i(t) x_i(t+\Delta t)\} \leq 0. \tag{4.13}$$

Next, in the case of $T \neq 0$, a random contribution is added to Eq.(4.10) so that the equilibrium state of the system realize the canonical ensemble [1, 28].

We now discuss the differences between our model and those of Noest and Cook. For the comparison we restrict our model to 2-dimensional case (clock type). In the Noest's model, when S_j is described as $e^{i\varphi_j}$ and ξ_j^μ as $e^{i\theta_j^\mu}$, the energy (4.6) is rewritten as

$$E = -\frac{1}{2N}\sum_{i\neq j}\sum_{\mu}\cos\{(\varphi_i - \theta_i^\mu) - (\varphi_j - \theta_j^\mu)\} \tag{4.14}$$

In the Cook's model ($Q = \infty$), with new notations $(2\pi/Q)n_i = \varphi_i$ and $(2\pi/Q)\xi_i^\mu = \theta_i^\mu$, the expression (4.7) coincides with Eq.(4.14). When we describe $x_i = {}^t(\cos\varphi_i, \ \sin\varphi_i)$ and $\xi_i^\mu = {}^t(\cos\theta_i^\mu, \ \sin\theta_i^\mu)$, the energy of our model (4.12) is rewritten as

$$E = -\frac{1}{2N}\sum_{i\neq j}\sum_{\mu}\cos(\varphi_i - \theta_i^\mu)\cos(\varphi_j - \theta_j^\mu). \tag{4.15}$$

Therefore the translational symmetry in Noest's and Cook's model is not existent in our model, thus the translated phase pattern $\{\theta_i + \bar{\theta}\}$ is quite different from the pattern $\{\theta_i\}$ in our model. However our model has an inversive symmetry in which the energy (4.12) is invariant under the transformation $\{x_i\} \to \{-x_i\}$ or $\{\xi_i^\mu\} \to \{-\xi_i^\mu\}$. This fact means that inverted patterns can also be memorized as in the Hopfield model.

4.3 Simulation

In this section we simulate our model according to the dynamics (4.10) and (4.11) with synaptic couplings (4.9). We adopt an asynchronous updating in which at each time step one neuron is selected randomly to be updated. The time step is set as $\Delta t = 1/N$. Simulation results are shown in the following.

First we examine the case $D = 2$. Patterns are written as $\xi_i^\mu = {}^t(\cos\theta_i^\mu, \ \sin\theta_i^\mu)$, with θ_i^μ distributed uniformly in $[0, 2\pi]$. These patterns are uncorrelated in the sense of $\sum_{i=1}^N \langle \xi_i^\mu \xi_i^\nu \rangle = \delta_{\mu\nu}$. Fig.4.2 shows the time evolutions of overlap $m^1(t)$ defined as

$$m^1(t) = \frac{1}{N}\sum_{j=1}^{N} {}^t\xi_i^1 x_i \tag{4.16}$$

which start from several initial value of $m^1(0)$. From Fig.4.2(a), which is for the case of $\alpha \equiv p/N = 0.05 (p = 20, N = 400)$, we see that the network can retrieve a memorized pattern. In Fig.4.3 the states of 10 neurons chosen randomly at time $t = 20$ are shown. Solid arrows represents x_i and dashed arrows ξ_i^1. With this configuration, the overlap value $m^1(20)$ is about 0.97. On the other hand in the case of $\alpha = 0.1 (p = 40, N = 400)$ the network can not retrieve a memorized pattern (Fig.4.2(b)). These results tell us that the storage capacity α_c is $0.05 < \alpha_c < 0.1$. Note that this storage capacity is larger than that of Cook ($\alpha_c = 0.038$ for $Q = \infty$) [27]. This will be discussed theoretically in the following section.

Next we make simulations in the case of $D = 3$. In this case p patterns are expressed as $\xi_i^\mu = {}^t(\sin\psi_i^\mu\cos\theta_i^\mu, \ \sin\psi_i^\mu\sin\theta_i^\mu, \ \cos\psi_i^\mu)$, with uniformly distributed ψ_i^μ in $[0, \pi]$ and θ_i^μ in $[0, 2\pi]$. These patterns are also uncorrelated. The result is that in the case of $\alpha = 0.025 (p = 10, N = 400)$ the network can retrieve a memorized pattern but in the case of $\alpha = 0.075 (p = 30, N = 400)$ the network can not retrieve a memorized pattern. Therefore the storage capacity α_c is $0.025 < \alpha_c < 0.075$. This result will also be considered theoretically in the following section.

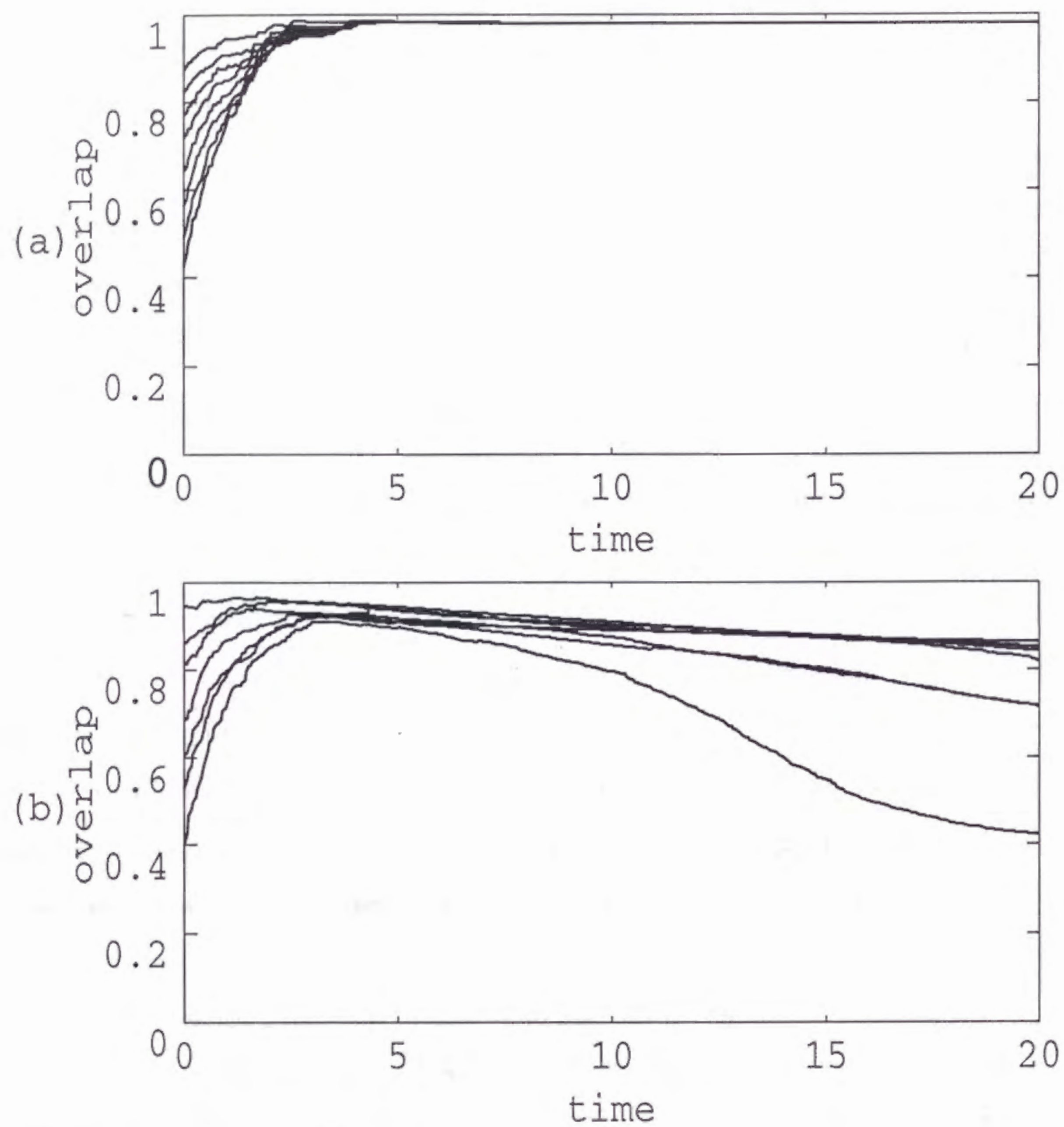

Figure 4.2: Time evolutions of overlaps which start from several initial conditions. (a)$N = 400, p = 20, \alpha = 0.05$; (b)$N = 400, p = 40, \alpha = 0.1$.

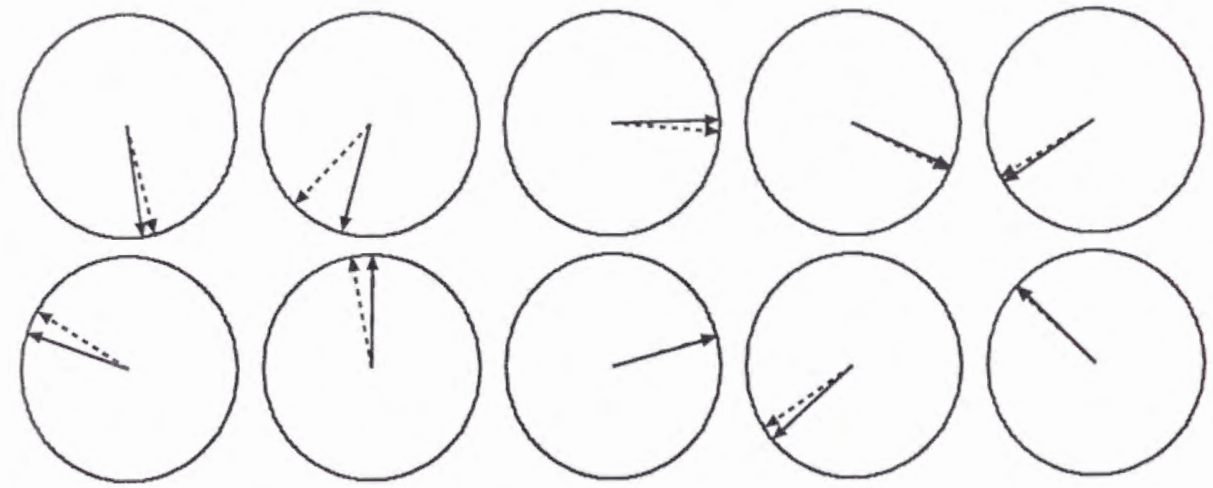

Figure 4.3: States of 10 neurons chosen randomly at time t=20 in the case of $p = 20$ and $D = 2$. Solid arrows represent $S_i(20)$ and dashed arrows ξ_i^1.

4.4 Mean-field theory

The mean-field theory is performed to calculate the free energy of the system with use of the replica method, i.e. the free energy per neuron, f, is written as

$$f = \lim_{n\to 0}\lim_{N\to\infty} -\frac{\langle\langle Z^n\rangle\rangle - 1}{nN\beta}, \tag{4.17}$$

where $\langle\langle\cdots\rangle\rangle$ denotes a quenched average over the memorized patterns, $\{\xi_i^\mu\}$, $\beta = 1/T$ an inverse temperature of the system, Z the partition function of the system, and n the number of replicas. In the following calculation we use the framework of Amit, Gutfreund and Sompolinsky [29, 30] (hereafter referred to as AGS), which is briefly reviewed in appendix. We adopt

$$H^\rho = -\frac{1}{2}\sum_{i\neq j} {}^t x_i^\rho J_{ij} x_j^\rho - \sum_{\nu=1}^{s} h^\nu \sum_{i=1}^{N} {}^t\xi_i^\nu x_i^\rho \tag{4.18}$$

as Hamiltonian of the ρ-th replica of the system. We assume that there are a finite number $s(\ll p)$ of fields $h^\nu(\nu = 1, \ldots, s)$ each of which is coupled to ν th condensed pattern. Then

$\langle\langle Z^n \rangle\rangle$ is calculated as

$$\langle\langle Z^n\rangle\rangle = \left\langle\!\left\langle \mathrm{Tr}_{\boldsymbol{x}^\rho} \exp(-\beta \sum_{\rho=1}^{n} H^\rho) \right\rangle\!\right\rangle \tag{4.19}$$

$$= \left\langle\!\left\langle \mathrm{Tr}_{\boldsymbol{x}^\rho} \exp\left[\frac{\beta}{2N}\sum_{ij\mu\rho}({}^t\xi_i^\mu \boldsymbol{x}_i^\rho)({}^t\xi_j^\mu \boldsymbol{x}_j^\rho) - \frac{\beta}{2N}\sum_{i\mu\rho}({}^t\xi_i^\mu \boldsymbol{x}_i^\rho)^2 + \beta\sum_{\nu\rho} h^\nu \sum_i {}^t\xi_i^\nu \boldsymbol{x}_i^\rho\right]\right\rangle\!\right\rangle. \tag{4.20}$$

Using Stratonovich-Hubbard transformation

$$\exp\left(\frac{A}{2a^2}\right) = \frac{1}{\sqrt{2\pi a A}}\int_{-\infty}^{\infty} dx \exp\left(-\frac{x^2}{2A} + ax\right), \tag{4.21}$$

we can proceed to calculate Eq.(4.20) as

$$\begin{aligned}\langle\langle Z^n\rangle\rangle &= (\beta N)^{\frac{sn}{2}} \left\langle\!\left\langle \mathrm{Tr}_{\boldsymbol{x}^\rho} \int \prod_{\mu\rho} \frac{dm_\rho^\mu}{\sqrt{2\pi}} \exp\left[\beta N\left\{-\frac{1}{2}\sum_{\mu\rho}(m_\rho^\mu)^2 + \sum_{\mu\rho} m_\rho^\mu \left(\frac{1}{N}\sum_i {}^t\xi_i^\mu \boldsymbol{x}_i^\rho\right) - \sum_{\mu\rho}\frac{1}{2N^2}\sum_i ({}^t\xi_i^\mu \boldsymbol{x}_i^\rho)^2\right\}\right]\right.\right. \\ &\times \int \prod_{\nu\rho} \frac{dm_\rho^\nu}{\sqrt{2\pi}} \exp\left[\beta N\left\{-\frac{1}{2}\sum_{\nu\rho}(m_\rho^\nu)^2 \left.\left. + \sum_{\nu\rho}(m_\rho^\nu + h^\nu)\left(\frac{1}{N}\sum_i {}^t\xi_i^\nu \boldsymbol{x}_i^\rho\right) - \sum_{\nu\rho}\frac{1}{2N^2}\sum_i ({}^t\xi_i^\nu \boldsymbol{x}_i^\rho)^2\right\}\right]\right\rangle\!\right\rangle. \end{aligned} \tag{4.22}$$

The sums $\sum_\nu$ and $\sum_\mu$ are over the first s patterns and over the remaining $p-s$ patterns, respectively. The products $\prod_\nu$ and $\prod_\mu$are also considered in the same way. In order to average the first exponential in Eq.(4.22) over the $p-s$ uncondensed patterns $\{\xi_i^\mu\}$, we make variable transformations

$$\boldsymbol{x}_i^\rho = \begin{pmatrix} x_{i(1)}^\rho \\ x_{i(2)}^\rho \\ x_{i(3)}^\rho \\ \vdots \\ x_{i(D)}^\rho \end{pmatrix} = \begin{pmatrix} \sin\zeta_{i(1)}^\rho \sin\zeta_{i(2)}^\rho \cdots \sin\zeta_{i(D-2)}^\rho \cos\varphi_i^\rho \\ \sin\zeta_{i(1)}^\rho \sin\zeta_{i(2)}^\rho \cdots \sin\zeta_{i(D-2)}^\rho \sin\varphi_i^\rho \\ \sin\zeta_{i(1)}^\rho \cdots\cdots\cdots \cos\zeta_{i(D-2)}^\rho \\ \vdots \\ \cos\zeta_{i(1)}^\rho \end{pmatrix} \tag{4.23}$$

and

$$\xi_i^\mu = \begin{pmatrix} \xi_{i(1)}^\mu \\ \xi_{i(2)}^\mu \\ \xi_{i(3)}^\mu \\ \vdots \\ \xi_{i(D)}^\mu \end{pmatrix} = \begin{pmatrix} \sin\psi_{i(1)}^\mu \sin\psi_{i(2)}^\mu \cdots \sin\psi_{i(D-2)}^\mu \cos\theta_i^\mu \\ \sin\psi_{i(1)}^\mu \sin\psi_{i(2)}^\mu \cdots \sin\psi_{i(D-2)}^\mu \sin\theta_i^\mu \\ \sin\psi_{i(1)}^\mu \cdots\cdots\cdots \cos\psi_{i(D-2)}^\mu \\ \vdots \\ \cos\psi_{i(1)}^\mu \end{pmatrix} \tag{4.24}$$

using polar coordinates in a D-dimensional space. $\zeta_{i(k)}^\rho, \psi_{i(k)}^\mu (k = 1, \ldots, D-2)$ vary from 0 to π and $\varphi_i^\rho, \theta_i^\mu$ from 0 to 2π. With use of these transformation, we can calculate an average over $\{\xi_i^\mu\}$ as

$$\langle\langle f(\xi_i^\mu)\rangle\rangle_{\xi_i^\mu} = \frac{1}{\mathcal{J}}\int_0^\pi d\psi_{i(1)}^\mu \int_0^\pi d\psi_{i(2)}^\mu \cdots \int_0^\pi d\psi_{i(D-2)}^\mu \int_0^{2\pi} d\theta_i^\mu \sin^{(D-2)}\psi_{i(1)}^\mu \sin^{(D-3)}\psi_{i(2)}^\mu \cdots \sin\psi_{i(D-2)}^\mu f(\psi_{i(1)}^\mu, \cdots, \psi_{i(D-2)}^\mu, \theta_i^\mu) \tag{4.25}$$

where

$$\mathcal{J} = \int_0^\pi d\psi_{i(1)}^\mu \int_0^\pi d\psi_{i(2)}^\mu \cdots \int_0^\pi d\psi_{i(D-2)}^\mu \int_0^{2\pi} d\theta_i^\mu. \tag{4.26}$$

Therefore the average is calculated as

$$\begin{aligned} &\left\langle\!\left\langle \exp\left[\beta N\left\{\sum_{\mu\rho} m_\rho^\mu \left(\frac{1}{N}\sum_i {}^t\xi_i^\mu \boldsymbol{x}_i^\rho\right) - \sum_{\mu\rho}\frac{1}{2N^2}\sum_i ({}^t\xi_i^\mu \boldsymbol{x}_i^\rho)^2\right\}\right]\right\rangle\!\right\rangle_{\xi_i^\mu} \\ &= \prod_{\mu i}\left\langle\!\left\langle \exp\left[\beta\sum_\rho\left\{m_\rho^{\mu t}\xi_i^\mu \boldsymbol{x}_i^\rho - \frac{1}{2N}({}^t\xi_i^\mu \boldsymbol{x}_i^\rho)^2\right\}\right]\right\rangle\!\right\rangle_{\xi_i^\mu} \\ &= \exp\left(-\frac{\beta p n}{2D}\right)\exp\left[\sum_{\mu i}\frac{\beta^2}{2D}\sum_{\rho\sigma} m_\rho^\mu m_\sigma^{\mu t} \boldsymbol{x}_i^\rho \boldsymbol{x}_i^\sigma\right], \end{aligned} \tag{4.27}$$

(see 4.6 Appendix). Inserting Eq.(4.27) into Eq.(4.22) and introducing new order parameters $r_{\rho\sigma}$ and $q_{\rho\sigma}$, we integrate over m_ρ^μ and make a saddle-point approximation for the integration over $m_\rho^\nu, r_{\rho\sigma}andq_{\rho\sigma}$ following the AGS. Thus we get the expression of $\langle\langle Z^n\rangle\rangle$ as

$$\begin{aligned}\langle\langle Z^n\rangle\rangle &= \exp\left(-\frac{\beta p n}{2D}\right)\left(\frac{\alpha\beta^2}{2D}\right)^{\frac{n(n-1)}{2}} \\ &\times \exp\left[-\frac{p}{2}\mathrm{Tr}\ln\left\{\left(1-\frac{\beta}{D}\right)I - \frac{\beta}{D}Q\right\} - \frac{N\alpha\beta^2}{2D}\sum_{\rho\neq\sigma} r_{\rho\sigma}q_{\rho\sigma}\right. \\ &-\frac{\beta N}{2}\sum_{\nu\rho}(m_\rho^\nu)^2 + \left\langle\!\left\langle N\ln \mathrm{Tr}_{\boldsymbol{x}^\rho}\exp\left\{\frac{\alpha\beta^2}{2D}\sum_{\rho\neq\sigma} r_{\rho\sigma}{}^t\boldsymbol{x}^\rho \boldsymbol{x}^\sigma\right.\right.\right. \\ &\left.\left.\left.\beta\sum_{\nu\rho}(m_\rho^\nu + h^\nu){}^t\xi^\nu \boldsymbol{x}^\rho - \frac{\beta}{2N}({}^t\xi^\nu \boldsymbol{x}^\rho)^2\right\}\right\rangle\!\right\rangle_{\xi^\nu}\right] \end{aligned} \tag{4.28}$$

where I is a unit matrix with $n \times n$ elements and Q is a matrix $\{q_{\rho\sigma}\}$ with zero diagonal elements. The parameters $m_\rho^\nu, q_{\rho\sigma},$ and$r_{\rho\sigma}$ are determined from the saddle point equations to be given as follows:

$$m_\rho^\nu = \left\langle\!\left\langle \frac{1}{N}\sum_i {}^t\xi_i^\nu \langle x_i^\rho \rangle \right\rangle\!\right\rangle, \qquad \nu = 1, \ldots, s, \tag{4.29}$$

$$q_{\rho\sigma} = \left\langle\!\left\langle \frac{1}{N}\sum_i {}^t\langle x_i^\rho \rangle \langle x_i^\sigma \rangle \right\rangle\!\right\rangle, \tag{4.30}$$

$$r_{\rho\sigma} = \frac{1}{\alpha}\sum_{\mu>s} \left\langle\!\left\langle m_\rho^\mu m_\sigma^\mu \right\rangle\!\right\rangle, \tag{4.31}$$

where $\langle\cdots\rangle$ denotes a thermal average. Eventually we get the free energy per neuron as

$$\begin{aligned} f = \lim_{n\to 0}\Bigg[& \frac{\alpha}{2D} + \frac{\alpha}{2\beta n}\mathrm{Tr}\ln\left\{\left(1-\frac{\beta}{D}\right)I - \frac{\beta}{D}Q\right\} \\ & + \frac{\alpha\beta}{2Dn}\sum_{\rho\neq\sigma} r_{\rho\sigma}q_{\rho\sigma} + \frac{1}{2n}\sum_{\nu\rho}(m_\rho^\nu)^2 - \frac{1}{n\beta}\left\langle\!\left\langle \ln \mathrm{Tr}_{x^\rho} \exp\left(\beta H_\xi\right)\right\rangle\!\right\rangle_{\xi^\nu}\Bigg], \end{aligned} \tag{4.32}$$

$$H_\xi = \frac{\alpha\beta}{2D}\sum_{\rho\neq\sigma} r_{\rho\sigma}{}^t x^\rho x^\sigma + \sum_{\nu\rho}(m_\rho^\nu + h^\nu){}^t\xi^\nu x^\rho. \tag{4.33}$$

In the case of $D = 1$ this result exactly reduces to that of AGS for the Hopfield model [29, 30]. When we take the replica symmetric assumption in which $m_\rho^\nu = m^\nu, q_{\rho\sigma} = q$, and $r_{\rho\sigma} = r$, Eq.(4.32) becomes

$$\begin{aligned} f = & \frac{\alpha}{2D} + \frac{\alpha}{2\beta}\ln\left\{1 - \frac{\beta}{D}(1-q)\right\} - \frac{\alpha}{2}\frac{q}{D-\beta(1-q)} \\ & + \frac{\alpha\beta r(1-q)}{2D} + \frac{1}{2}\sum_\nu (m^\nu)^2 - \frac{1}{\beta}\left\langle\!\left\langle \ln \mathrm{Tr}_{x} \exp\left(\sum_k A_k x_{(k)}\right)\right\rangle\!\right\rangle \end{aligned} \tag{4.34}$$

$$A_k = \beta\left\{\sqrt{\frac{\alpha r}{D}} z_k + \sum_\nu (m^\nu + h^\nu)\xi_{(k)}^\nu\right\}, \quad (k = 1, \ldots, D) \tag{4.35}$$

where $\langle\!\langle\cdots\rangle\!\rangle$ denotes the combined average over condensed patterns $\{\xi^\nu\}$ (Eq.(4.25)) and over the Gaussian noises $z_1, \ldots, z_k$ defined as

$$\langle\!\langle f(\xi^\nu)\rangle\!\rangle = \left\langle\!\left\langle \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{dz_1\cdots dz_D}{(2\pi)^{D/2}}\exp\left(-\frac{z_1^2+\cdots+z_D^2}{2}\right) f(\xi^\nu)\right\rangle\!\right\rangle_{\xi^\nu}. \tag{4.36}$$

Mean-field equations are determined as follows:

$$\frac{\partial f}{\partial m^\nu} = 0 \ : \ m^\nu = \frac{1}{\beta}\left\langle\!\left\langle \frac{\partial}{\partial m^\nu}\left[\ln \mathrm{Tr}_x \exp\left(\sum_k A_k x_{(k)}\right)\right]\right\rangle\!\right\rangle \tag{4.37}$$

$$\frac{\partial f}{\partial r} = 0 \ : \ \beta(1-q) = \frac{2D}{\alpha\beta}\left\langle\!\left\langle \frac{\partial}{\partial r}\left[\ln \mathrm{Tr}_x \exp\left(\sum_k A_k x_{(k)}\right)\right]\right\rangle\!\right\rangle \tag{4.38}$$

$$\frac{\partial f}{\partial q} = 0 \ : \ r = \frac{Dq}{\{D-\beta(1-q)\}^2} \tag{4.39}$$

We note that when we solve these equations, a set of solutions $m^\mu = 0, q = 0$ represents the paramagnetic state, $m^\mu = 0, q \neq 0$ the spin-glass state and $m^\mu \neq 0, q \neq 0$ the retrieval state. For the case of $D = 2$, Tr_x in Eq.(4.34) is calculated explicitly to give

$$\begin{aligned} \mathrm{Tr}_x \exp\left(\sum_k A_k x_{(k)}\right) &= \int_0^{2\pi} d\varphi \exp(A_1\cos\varphi + A_2\sin\varphi) \\ &= 2\pi I_0\left(\sqrt{A_1^2 + A_2^2}\right) \end{aligned} \tag{4.40}$$

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

$$I_k(z) = \frac{1}{2\pi}\int_0^{2\pi} d\phi e^{z\cos\phi}\cos k\phi. \tag{4.41}$$

In the following subsections we examine the properties of our model using the free energy (4.34) under the conditions

$$m^1 = m, \quad m^\nu = 0 \ (\nu \geq 2), \quad h^\nu = 0. \tag{4.42}$$

4.4.1 $\alpha = 0$

In this subsection we examine the case $\alpha = 0$ in which an intensive number of patterns are embedded in the network ($N = \infty$). We put $\alpha = 0$ in Eq.(4.34) and get

$$f = \frac{1}{2}m^2 - \frac{1}{\beta}\ln C_D\frac{I_{\frac{D}{2}-1}(\beta m)}{(\frac{1}{2}\beta m)^{\frac{D}{2}-1}}, \tag{4.43}$$

where

$$\begin{aligned} C_1 &= 2\sqrt{\pi}, \quad C_2 = 2\pi, \quad C_3 = 2\pi\sqrt{\pi}, \\ C_D &= \sqrt{\pi}\Gamma\left(\frac{D-1}{2}\right)\int_0^\pi d\zeta_2 \sin^{D-3}\zeta_2\cdots\int_0^\pi d\zeta_{D-2}\sin\zeta_{D-2}\int_0^{2\pi}d\varphi \quad (D \geq 4). \end{aligned}$$

The mean-field equation becomes

$$m = \frac{I_{\frac{D}{2}}(\beta m)}{I_{\frac{D}{2}-1}(\beta m)} \equiv g(\beta; m). \tag{4.44}$$

According to the property of the modified Bessel function

$$\frac{d}{dz}\frac{I_\nu(z)}{I_{\nu-1}(z)} > 0, \quad \frac{d^2}{dz^2}\frac{I_\nu(z)}{I_{\nu-1}(z)} \left\{ \begin{array}{ll} < 0 & (z > 0) \\ > 0 & (z < 0) \end{array} \right. \tag{4.45}$$

the function $g(\beta; m)$ is a sigmoid type one. Therefore under the condition

$$\left.\frac{d}{dm} g(\beta; m)\right|_{m=0} \leq 1, \tag{4.46}$$

which yields $\beta \leq D$ or $T \geq 1/D$, Eq.(4.44) has no solution except for $m = 0$. On the other hand as the noise T decreases from $1/D$ or β increases from D, Eq.(4.44) becomes to have non-zero positive solution which is stable. Fig.4.4 shows positive solutions of Eq.(4.44) as a function of temperature $T = 1/\beta$ for various case of dimension D.

4.4.2 $T = 0$

Next we consider the case of a low temperature limit $T = 0$ or $\beta = \infty$, where the mean-field equations (4.37), (4.38), and (4.39) become as follows:

$$m = \left\langle\!\!\left\langle \frac{\sum_k z_k \xi_{(k)} + y}{\sqrt{\sum_k z_k^2 + 2y \sum_k z_k \xi_{(k)} + y^2}} \right\rangle\!\!\right\rangle \equiv f_1(y), \tag{4.47}$$

$$\beta(1-q) = \left\langle\!\!\left\langle \frac{\sum_k z_k^2 + y \sum_k z_k \xi_{(k)}}{\sqrt{\sum_k z_k^2 + 2y \sum_k z_k \xi_{(k)} + y^2}} \right\rangle\!\!\right\rangle \equiv f_2(y), \tag{4.48}$$

$$r = \frac{1}{D\left\{1 - \frac{1}{D}\sqrt{\frac{D}{\alpha r}} f_2(y)\right\}^2}, \tag{4.49}$$

$$y \equiv \sqrt{\frac{D}{\alpha r}} m. \tag{4.50}$$

These equations are reduced to a single equation for the variable y,

$$y = \frac{D f_1(y)}{\sqrt{\alpha} + f_2(y)}. \tag{4.51}$$

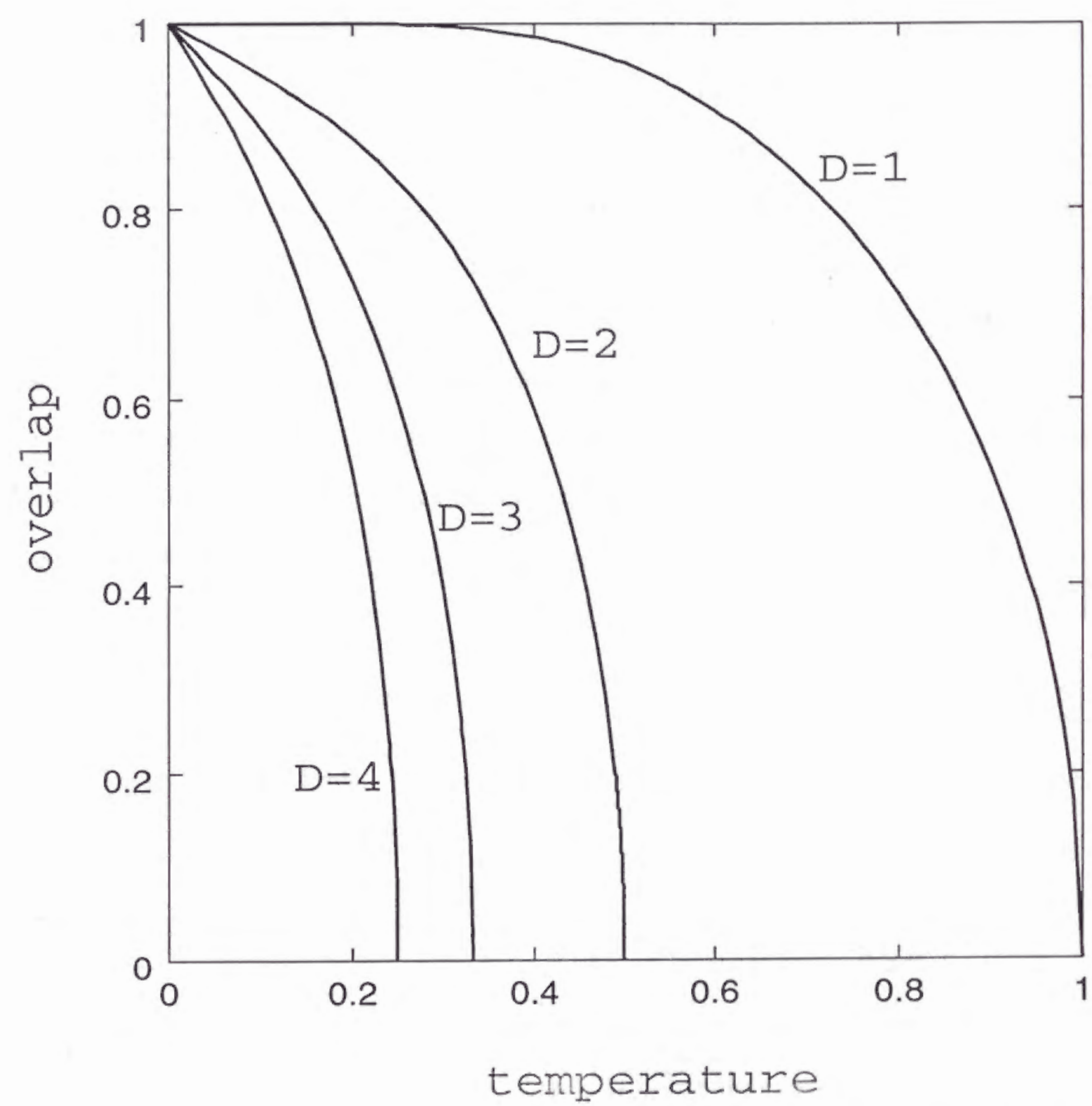

Figure 4.4: Nonzero solutions of Eq.(4.44) as a function of temperature $T = 1/\beta$ for $D = 1, 2, 3, \text{and} 4$.

This is a relation between the retrieval quality m and the storage level α. The storage capacity α_c is the value of α above which the Eq.(4.51) has no solution except for $y = 0$. The graphical solution of Eq.(4.51) is shown in Fig.4.5 for $m \geq 0$. The straight line represents the left hand side(l.h.s.). The dashed curves represents the right hand side (r.h.s.) plotted for two values of α, one below and one above α_c. For $\alpha < \alpha_c$ we have three non-negative solutions, $m_1 = 0, 0 < m_2 < m_3$. m_1, m_3 are stable and m_2 is unstable. The solution m_1 represents the spin-glass state because $m_1 = 0$ and $q \neq 0$ and m_3 the retrieval state because $m_1 \neq 0$ and $q \neq 0$. For $\alpha > \alpha_c$ there exists only one solution $m = 0$ which is the spin-glass state because of $q \neq 0$. Fig.4.5 tells us that the retrieval solution disappear abruptly. Fig.4.6 shows the solution m as a function of α in three cases $D = 1, 2$ and 3. In the case $D = 2$, retrieval solution disappears abruptly at $\alpha = 0.0743$ which is the storage capacity α_c. Note that this storage capacity is larger than that of the Cook's model [27] and is consistent with the simulation results in Sec.2. In the case $D = 3$, the storage capacity is calculated to be $\alpha_c = 0.0432$.

4.4.3 $T - \alpha$ phase diagram

We now turn to the full mean-field equations (4.37), (4.38), and (4.39), keeping T and α finite. In the case of $D = 2$, mean-field equations can be written explicitly as

$$m = \left\langle\!\!\left\langle \frac{\beta}{\sqrt{A_1^2 + A_2^2}} \frac{I_1\left(\sqrt{A_1^2 + A_2^2}\right)}{I_0\left(\sqrt{A_1^2 + A_2^2}\right)} \left\{\sqrt{\frac{\alpha r}{2}}(z_1\xi_{(1)} + z_2\xi_{(2)}) + m\right\} \right\rangle\!\!\right\rangle, \tag{4.52}$$

$$\beta(1-q) = \sqrt{\frac{2}{\alpha r}} \left\langle\!\!\left\langle \frac{\beta}{\sqrt{A_1^2 + A_2^2}} \frac{I_1\left(\sqrt{A_1^2 + A_2^2}\right)}{I_0\left(\sqrt{A_1^2 + A_2^2}\right)} \left\{\sqrt{\frac{\alpha r}{2}}(z_1^2 + z_2^2) + m(z_1\xi_{(1)} + z_2\xi_{(2)})\right\} \right\rangle\!\!\right\rangle, \tag{4.53}$$

$$r = \frac{2q}{\{2 - \beta(1-q)\}^2}. \tag{4.54}$$

The $T - \alpha$ phase diagram for $D = 2$ can be obtained by solving Eqs.(4.52), (4.53), and (4.54) numerically. But because of the difficulty of the numerical calculation, we obtained

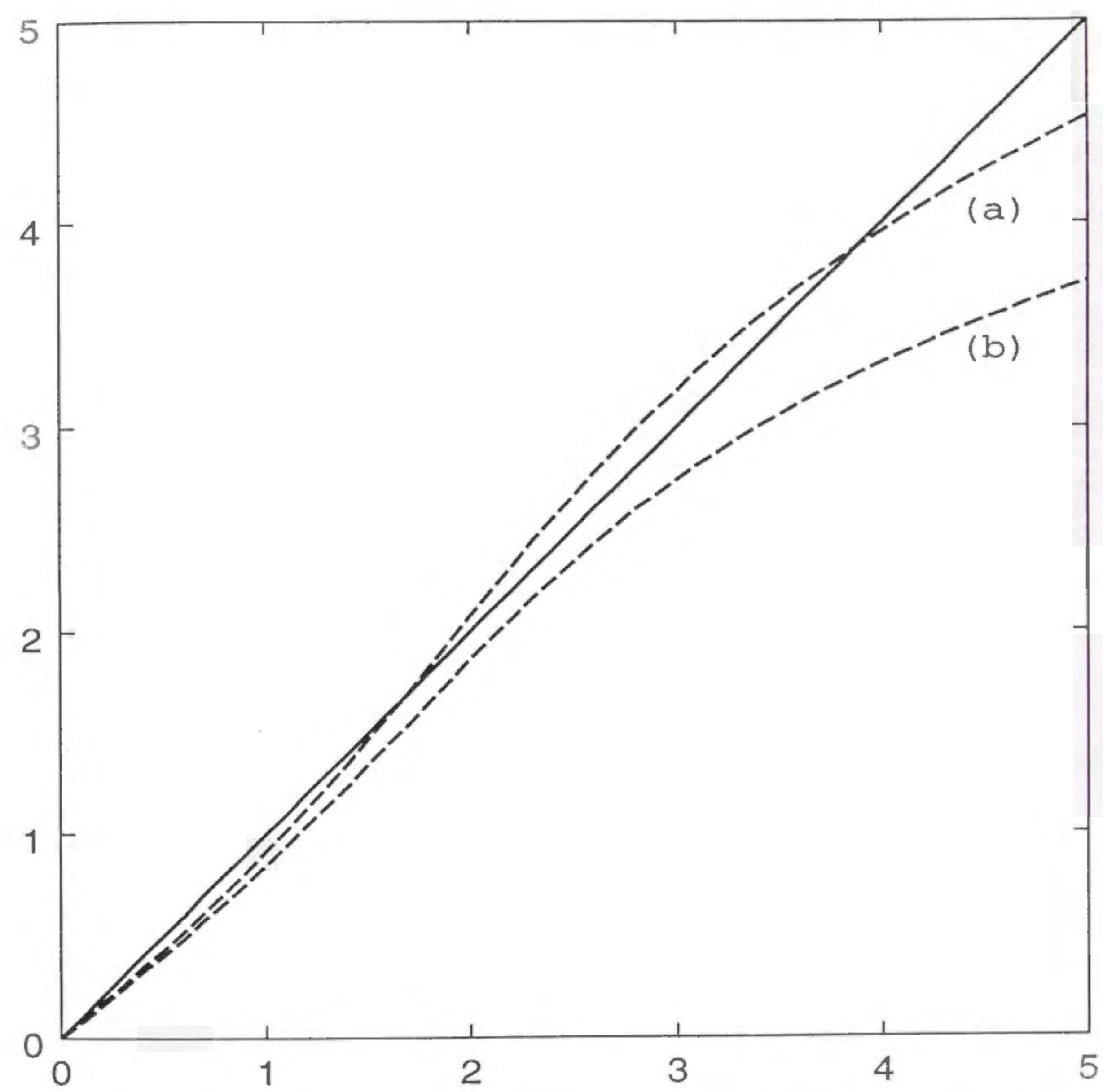

Figure 4.5: The graphical representation of the solutions of Eq.(4.51) in the case $D = 2$. (a)$\alpha = 0.05$, (b)$\alpha = 0.1$.

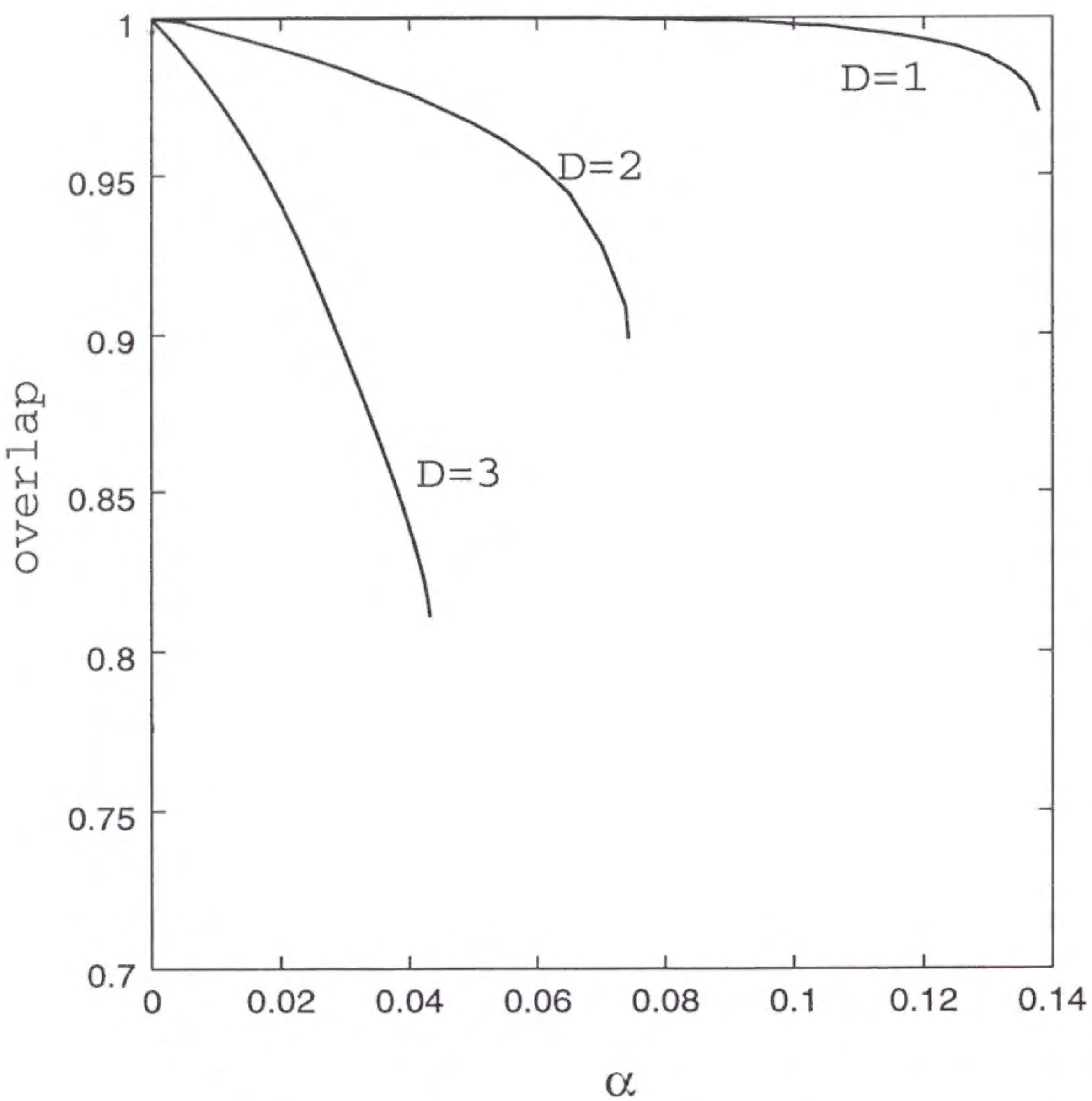

Figure 4.6: Solutions m of Eqs.(50)(4.51) as a function of α in the case of $D = 1, 2,$ and3.

only a part of the phase diagram, i.e., the boundary between the paramagnetic phase and the spin-glass one. In order to gain the information of the phase diagram as much as possible under this circumstances, we try to obtain the qualitative structure of the phase diagram analytically.

At high temperature T or low β, only the set of solutions $m = 0, q = 0$ (the paramagnetic state) is possible. Decreasing the temperature for fixed α, we cross the transition temperature $T_g(\alpha)$ below which a set of solutions $m = 0, q \neq 0$ (the spin-glass state) appears. With the anticipation that q will develop continuously from zero, the r.h.s. of Eq.(4.54) is expanded in powers of q to give in the lowest order

$$r \approx \frac{2q}{(2-\beta)^2}. \tag{4.55}$$

Setting $m = 0$ in Eq.(4.53) and using Eq.(4.55), we get

$$q = \frac{\beta^2 \alpha r}{4} = \frac{\beta^2 \alpha q}{2(2-\beta)^2}. \tag{4.56}$$

From this equation we find the transition temperature $T_g(\alpha)$ as

$$T_g(\alpha) = \frac{1}{2} + \frac{\sqrt{\alpha}}{2\sqrt{2}}. \tag{4.57}$$

When we furthermore decrease the temperature from $T_g(\alpha)$ for fixed $\alpha < \alpha_c = 0.0743$, we cross the new transition temperature $T_M(\alpha)$. Below $T_M(\alpha)$, a set of solutions $m \neq 0, q \neq 0$ (the retrieval state) is possible. In this case the transition is first order type, so there can be no expansion in m generally. But we can make an analytic calculation in the corner of the phase diagram near $\alpha = 0$ and $T = \frac{1}{2}$, because we expect both q and the discontinuity in m to be small there. There are three small parameters $m, q,$ and$t = \frac{1}{2} - T$. The equations (4.52), (4.53), and (4.54) are expanded in powers of these parameters to give

$$t = \frac{1}{2}m^2 + \alpha r, \tag{4.58}$$

$$q = 2m^2 + 2\alpha r, \tag{4.59}$$

$$r = \frac{q}{2(q-2t)^2}. \tag{4.60}$$

From these equations we get

$$g(y) \equiv y^3 - 2\tau y^2 + y + 2\tau = 0 \tag{4.61}$$

where

$$\tau \equiv \frac{t}{\sqrt{\alpha}}, \quad y \equiv \frac{m^2}{\sqrt{\alpha}}. \tag{4.62}$$

This equation has either two positive solutions or none at all. $T_M(\alpha)$ is determined by the disappearance of the two solutions. The value of τ at which the two solutions just disappear is calculated to be $\tau = 1.67$. Hence $T_M(\alpha)$ near $\alpha = 0$ and $T = 1/2$ is found to be

$$T_M(\alpha) = \frac{1}{2} - 1.67\sqrt{\alpha}. \tag{4.63}$$

In the case of $D = 3$, we can calculate $T_g(\alpha)$and$T_M(\alpha)$ in the same way as above to give

$$T_g(\alpha) = \frac{1}{3} + \frac{1}{3}\sqrt{\frac{5}{6}\alpha}, \tag{4.64}$$

$$T_M(\alpha) = \frac{1}{3} - 1.42\sqrt{\alpha}. \tag{4.65}$$

Summarizing the above results and the storage capacity which is discussed in the previous subsection, we get a phase diagram as is shown in Fig.4.7.

We note that the agreement between the line T_g obtained theoretically and that obtained numerically is excellent. Dashed lines are expected ones which are depicted for the guide of eyes. This figure tells us that the retrieval regime in the phase diagram becomes smaller as the dimension D increases, as expected.

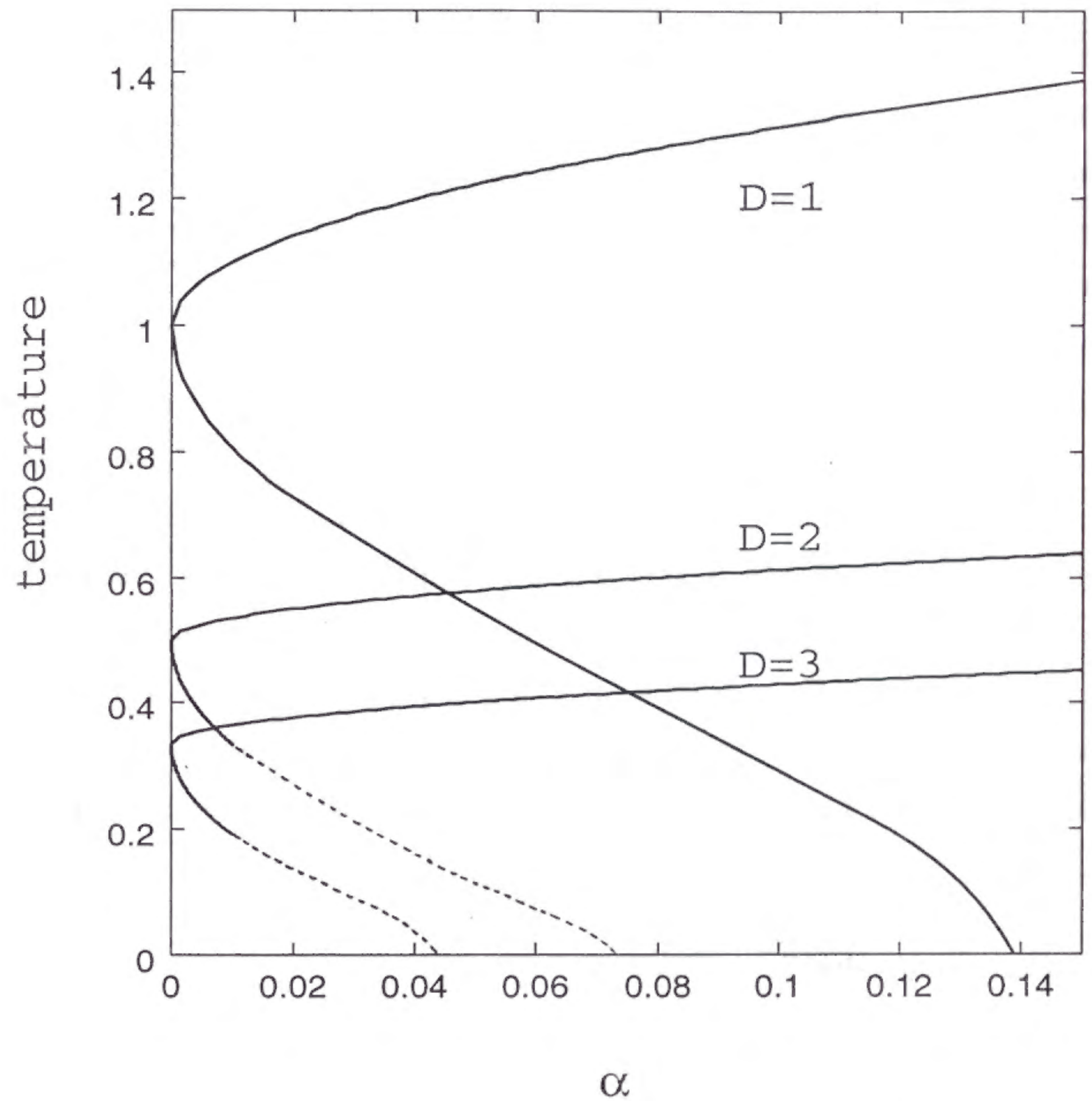

Figure 4.7: $T - \alpha$ phase diagram of a network in the case of $D = 1, 2,$ and3.

4.5 Summary

In this paper, we studied a neural network composed of D-dimensional spin neurons as an extension of the Hopfield model to a multi-dimensional one. In the case $D=2$, we found that a stored phase pattern $\{\theta_i\}$ itself, not a translated one $\{\theta_i+\bar{\theta}\}$, is retrieved. We analyzed the network by means of the replica symmetric theory and got the free-energy of the network in the case of general dimension D. For $D=1$, the free-energy was confirmed to coincide that of the Hopfield model. With use of this free-energy, first, the case of $\alpha=0$ was studied. This is the case in which a finite number of patterns are embedded in the network with an infinite number of neurons. We found that the patterns can be retrieved if the temperature T is lower than $1/D$. Next we calculated the storage capacity in the case of $D=2$ and $D=3$ and obtained $\alpha_c=0.0743(D=2)$ and $\alpha_c=0.0432(D=3)$ for $T=0$. In the case of $D=2$, our model is similar to the Cook's one, but the storage capacity of our model is larger than that of Cook's model. One of the reason is that the synaptic coupling J_{ij} of our model is expressed by matrix which needs more information. We also calculated the phase diagram and found that the retrieval regime shrank as dimension D increased. It was difficult to calculate the storage capacity for $D\geq 4$, but the dependence of the storage capacity on dimension D is interesting open problem.

4.6 Appendix

In this Appendix we show how to calculate $\langle\!\langle\cdots\rangle\!\rangle_{\xi_i^\mu}$ in eq.(4.27). First we estimate the exponential in second line of eq.(4.27).

$$\begin{aligned}
&\exp\left[\beta\sum_\rho\left\{m_\rho^{\mu t}\xi_i^\mu x_i^\rho-\frac{1}{2N}({}^t\xi_i^\mu x_i^\rho)^2\right\}\right]\\
=&\exp\left[\sum_{j=1}^D X_j\xi_{i(j)}^\mu-\sum_{j=1}^D Y_j(\xi_{i(j)}^\mu)^2-\sum_{j\neq k}^D Z_{jk}\xi_{i(j)}^\mu\xi_{i(k)}^\mu\right]
\end{aligned}$$

$$\approx 1+\sum_{j=1}^D X_j\xi_{i(j)}^\mu+\frac{1}{2}\left(\sum_{j=1}^D X_j\xi_{i(j)}^\mu\right)^2-\sum_{j=1}^D Y_j(\xi_{i(j)}^\mu)^2\sum_{j\neq k}^D Z_{jk}\xi_{i(j)}^\mu\xi_{i(k)}^\mu, \tag{4.66}$$

where

$$X_j \equiv \beta\sum_\rho m_\rho^\mu x_{i(j)}^\rho\sim O\left(\frac{1}{\sqrt{N}}\right), \tag{4.67}$$

$$Y_j \equiv \frac{\beta}{2N}\sum_\rho(x_{i(j)}^\rho)^2\sim O\left(\frac{1}{N}\right), \tag{4.68}$$

$$Z_{jk} \equiv \frac{\beta}{2N}\sum_\rho x_{i(j)}^\rho x_{i(k)}^\rho\sim O\left(\frac{1}{N}\right), \tag{4.69}$$

and we neglected terms which are smaller than $O(1/N)$ in the Taylor expansion. Using eq.(4.25), we easily find followings:

$$\langle\!\langle\xi_{i(j)}^\mu\rangle\!\rangle_{\xi_i^\mu} = 0 \quad (j=1,\ldots,D), \tag{4.70}$$

$$\langle\!\langle\xi_{i(j)}^\mu\xi_{i(k)}^\mu\rangle\!\rangle_{\xi_i^\mu} = 0 \quad (j,k=1,\ldots,D;j\neq k), \tag{4.71}$$

$$\langle\!\langle(\xi_{i(j)}^\mu)^2\rangle\!\rangle_{\xi_i^\mu} = \frac{1}{D} \quad (j=1,\ldots,D). \tag{4.72}$$

Therefore we can calculate as

$$\begin{aligned}
&\left\langle\!\!\left\langle\exp\left[\beta\sum_\rho\left\{m_\rho^{\mu t}\xi_i^\mu x_i^\rho-\frac{1}{2N}({}^t\xi_i^\mu x_i^\rho)^2\right\}\right]\right\rangle\!\!\right\rangle_{\xi_i^\mu}\\
=&\left\langle\!\!\left\langle 1+\sum_{j=1}^D X_j\xi_{i(j)}^\mu+\frac{1}{2}\left(\sum_{j=1}^D X_j\xi_{i(j)}^\mu\right)^2-\sum_{j=1}^D Y_j(\xi_{i(j)}^\mu)^2\sum_{j\neq k}^D Z_{jk}\xi_{i(j)}^\mu\xi_{i(k)}^\mu\right\rangle\!\!\right\rangle_{\xi_i^\mu}\\
=&1+\frac{1}{2D}\sum_{j=1}^D X_j^2-\frac{1}{D}\sum_{j=1}^D Y_j\\
=&1+\frac{\beta^2}{2D}\sum_{\rho\sigma}m_\rho^\mu m_\sigma^{\mu t}x_i^\rho x_i^\sigma-\frac{\beta p n}{2D}\\
=&\exp\left[\ln\left\{1+\frac{\beta^2}{2D}\sum_{\rho\sigma}m_\rho^\mu m_\sigma^{\mu t}x_i^\rho x_i^\sigma-\frac{\beta p n}{2D}\right\}\right]\\
=&\exp\left(-\frac{\beta p n}{2D}\right)\exp\left[\sum_{\mu i}\frac{\beta^2}{2D}\sum_{\rho\sigma}m_\rho^\mu m_\sigma^{\mu t}x_i^\rho x_i^\sigma\right],
\end{aligned} \tag{4.73}$$

where we used an approximation $\ln(1+\varepsilon)\approx\varepsilon$ for sufficiently small ε. eq.(4.73) is the final form of eq.(4.27).

Chapter 5

Clustering behavior of time-delayed nearest-neighbor coupled oscillators

5.1 Introduction

Quite a few papers have been devoted to the dynamics of populations of coupled oscillators. Such populations are known to model many systems of physics [53, 54], biology [55, 56, 57, 58] and chemistry [59]. It has been shown that coupled limit-cycle oscillators with weak coupling can be described by a system of phase coupled oscillators where each individual oscillator is described by a single variable, its phase [59]. Therefore, if we can show some features for a system of phase coupled oscillators, we may say that the same is true for more general oscillator systems.

One of the remarkable phenomena exhibited by a system of coupled oscillators is synchronization or entrainment; when oscillators, not necessarily identical each other, are allowed to mutually interact, it sometimes happens that dynamics of oscillators are entrained to an attractor in which every oscillator performs a synchronized (frequency- or phase-locked) oscillation. It is reported that when the natural frequency of each oscillator is distributed around an average one, there exists a critical strength of coupling, above which oscillators can be synchronized [60, 61, 62].

Furthermore an attractor, which represents a so-called cluster state, has been found for an oscillator system with mean-field coupling, i.e., a globally coupled oscillator system [63, 64]. This state is characterized by the coexistence of subgroups each of which consists of fully phase-synchronized oscillators. This clustering phenomenon is interesting from the stand point of a symmetry breaking because a collection of dynamical elements are spontaneously divided into some groups. In a sense the entrained or phase-locked state, mentioned before, may be called a 1-cluster state. The opposite extremity is the anti-phase state [72, 73], in which all the elements of the system (the number is N) oscillate with the same frequency and with equal spacing $2\pi/N$ in phase (a N-cluster state). Thus a system can have various cluster states.

In this paper we mainly investigate clustering behaviors in a phase coupled oscillator system. The model systems may be variously classified even if we confine our discussion to phase-coupled systems. First, if we pay attention to the range of interaction among oscillators, the models may belong to either the mean-field (MF) [60, 63, 64] or short-range (SR) [61, 62, 65, 66, 67, 68, 70] interaction model. Due to the simplification of analysis in the case of MF coupling, the MF model has attracted a lot of attention. However, spatial structures, which are absent in the MF model, are expected to yield variety of self-organized phenomena in a system with SR coupling. From this point of view we are mainly concerned with SR model. Secondly, we now turn to the time-delay of the interaction [69, 70, 71]. In models which are related to physical and/or chemical phenomena, time delay in the interaction among oscillators dose not play an important role. However in models which describe interaction among neurons or some biological elements, it is sometimes important to take into account effects of finite time τ required for information transmission between two elements. Therefore we consider both cases $\tau = 0$ and $\tau \neq 0$.

Summarizing the above we study collective behavior of the time-delayed nearest-neighbor (NN) coupled oscillator systems. Such systems have been studied by some authors. It is reported that due to the time delay there exists a multitude of synchronized solutions [69] and that time-delay leads to a frequency depression [70]. However it has not been reported whether clustering behavior can be observed or not in the NN coupling systems and how is the stability of cluster states modified under the influence of a finite time-delay if such states exist. This the main topic of this paper.

This paper is organized as follows: In Sec.5.2 we present our model and in Sec.5.3 results of our computer experiments are reported with the main emphasis put on the clustering behavior which is found for the NN coupling systems. In Sec.5.4 some theoretical studies are performed with use of an energy principle valid for the case of no time-delay ($\tau = 0$) and linear stability analysis for the case of finite time-delay ($\tau \neq 0$). Final section contains some remarks and summary of the results obtained in this paper.

5.2 Model

In this section we consider a one-dimensional system composed of N phase-coupled oscillators whose dynamics are described as

$$\frac{d\phi_k(t)}{dt} = \omega_0 + \frac{K}{I}\sum_j \sin\{\phi_j(t-\tau) - \phi_k(t)\} \quad (k = 1, \ldots, N) \tag{5.1}$$

where $\phi_k(t)$ denotes the phase of kth oscillator at time t and ω_0 common natural frequency, i.e the frequency in the absence of interaction among oscillators. K is the coupling strength, I is the number of oscillators which one oscillator interacts and τ is a time-delay. In Sec.5.5 we will comment on the extension to a higher dimensional model. Since we are mainly concerned with the NN coupling, the summation variable j takes the value $k-1$ and $k+1$, and $I = 2$. We also consider, for the sake of comparison, the case of MF coupling for which j takes all the values except for k and $I = N - 1$. We note that ϕ_{N+1} is identical to ϕ_1 due to the periodic boundary condition. Furthermore by variable

transformations,

$$\omega_0 t = \tilde{t}, \quad \frac{K}{\omega_0} = \tilde{K}, \quad \omega_0\tau = \tilde{\tau}, \quad \phi_k(t) = \tilde{\phi}_k(\tilde{t}) \tag{5.2}$$

we can get dimensionless equation

$$\frac{d\tilde{\phi}_k(\tilde{t})}{d\tilde{t}} = 1 + \frac{\tilde{K}}{I}\sum_j \sin\{\tilde{\phi}_j(\tilde{t} - \tilde{\tau}) - \tilde{\phi}_k(\tilde{t})\} \quad (k = 1, \ldots, N). \tag{5.3}$$

In what follows, we consider Eq.(5.3) with tides omitted for simplicity.

Systems with MF coupling and no time-delay ($\tau = 0$) have been studied intensively. For $\tau = 0$ it is natural to term the coupling with $K > 0 (K < 0)$ as attractive(repulsive), because a small deviation $\Theta_{jk} = \phi_j - \phi_k$ tends to decrease(increase) under the interaction $K \sin\Theta_{jk}$. We use this terminology even in the case of finite τ. These points will be discussed in detail in Sec.4 when we study the stability of cluster solutions based on an energy analysis.

In anticipation of the various cluster state which are found in our computer simulation, and for convenience of the later discussion, we give here a systematical rule to produce cluster states. Before proceeding to the case of the NN coupling we first explain the (symmetric) cluster states for the case of the MF coupling which are discussed in detail by Okuda [64]. A symmetric N_c-cluster state, where N_c is one of the divisor of N, is defined as the state in which each cluster $l (l = 0, 1, \ldots, Nc - 1)$ consisting of N/N_c oscillators with same phase rotate with same frequency Ω. Thus, when the phase of N_c clusters are equally separated, the phase ψ_l of cluster l is described as

$$\psi_l = \Omega t + \frac{2\pi l}{N_c}, \quad (l = 0, \ldots, N_c - 1). \tag{5.4}$$

In the case of $\tau = 0$, Ω is determined to be

$$\Omega = 1 + \frac{K}{N_c}\sum_{l=0}^{n-1} \sin\left(\frac{2\pi l}{N_c}\right). \tag{5.5}$$

Since there are many ways of distributing N oscillators into N_c groups with equal number of elements ($= N/N_c$), the solution (5.4), (5.5) is highly degenerate. The case of $N_c = N$

has attracted special attention because of the $(N-1)!$ attractors (a so-called attractor crowding) [72, 73].

We now turn to the cluster state of the NN case. The model is explicitly written as

$$\frac{d\phi_k(t)}{dt} = 1 + \frac{K}{2}\sum_{j=\pm 1} \sin\{\phi_j(t-\tau) - \phi_k(t)\}. \tag{5.6}$$

In this case the phase difference

$$\Phi = \phi_{k+1}(t) - \phi_k(t), \tag{5.7}$$

which turns out to be independent on k and t, plays the fundamental role as N_c dose in the case of MF model. Eq.(5.7) means that, in constructing a cluster state characterized by Φ, we first specify the phase ϕ_1 of the first oscillator. Then we take $\phi_2 = \phi_1 + \Phi$ and so on. Since Φ denotes a phase, it is in the range $0 \leq \Phi < 2\pi$. However if Φ happens to be larger than π, $\Phi = \phi_k - \phi_{k+1}$ can be taken in the range $0 \leq \Phi \leq \pi$ and we follow the prescription $\phi_N = \phi_1 + \Phi, \phi_{N-1} = \phi_N + \Phi$ and so on. Thus hereafter we will restrict Φ in the range $0 \leq \Phi \leq \pi$. With use of phase unit

$$\theta = \frac{2\pi}{N}, \tag{5.8}$$

Φ is expressed, from the condition $N\Phi = 2\pi n$, as

$$\Phi = n\theta \quad \begin{cases} n = 0, \ldots, N/2 & \text{for } N \text{ ; even} \\ n = 0, \ldots, (N-1)/2 & \text{for } N \text{ ; odd.} \end{cases} \tag{5.9}$$

Thus the number of the cluster, N_c, is determined, when $\Phi \neq 0$, from the condition

$$n\theta N_c = 2\pi m \tag{5.10}$$

where m denotes the smallest nonnegative integer that satisfies the equation (5.10). For example, in the case of $N = 30$, $N_c = 10$ for $n = 3$, $N_c = 5$ for $n = 12$, and so on. For $\Phi = 0$, it is evident that $N_c = 1$.

As to the dynamics of the cluster states, the phase of kth oscillator evolves in time as

$$\phi_k(t) = \Omega t + k\Phi \quad (k = 1, \ldots, N) \tag{5.11}$$

where Ω is the common frequency of the oscillators. Inserting Eq.(5.11) into Eq.(5.6), we get

$$\Omega = 1 - K\sin(\Omega\tau)\cos\Phi. \tag{5.12}$$

Even if Φ (or n) is given, depending on the parameter K and τ, the transcendental equation (5.12) for Ω may have more than one solution, which we denote as $\Omega_b(b = 1, \ldots, B)$ (b = 1,...,B). Thus in order to unambiguously define a cluster state we must specify not only Φ, Eq.(5.7), but also Ω_b, and we denote the cluster state by (Φ, Ω_b).

5.3 Computer simulation for NN coupling

We performed three types of computer simulations to examine :(1) the structure of cluster states, (2) linear stability of cluster states and (3) relative nonlinear stability among the linearly stable cluster states. Computer experiments were performed by integrating Eq.(5.3) with the use of Runge-Kutta-Gill algorithm with the time step $\Delta t = 0.01$. We have checked that decreasing the time step does not affect the results. We also examined systems consisting of different number of oscillators. Since general features of clustering behavior does not depend on N, we show results for $N = 30$ in the following. The initial conditions of the oscillators were chosen randomly or or some cluster states according to the purpose of experiments.

5.3.1 Examples of cluster states

In this subsection we show some examples of clustering behavior of the system with $N = 30$, thus phase unit $\theta = \pi/15$ (see Eq.(5.8)). For the initial condition we set

$$\phi_k(t) = 1(t+\tau) + \alpha_k \quad (-\tau \leq t \leq 0), \tag{5.13}$$

where phases α_k $(k = 1, \ldots, N)$ are randomly distributed over $[0, 2\pi]$. For the attractive coupling $(K > 0)$ either a 1-cluster $(\Phi = 0)$ or 30-cluster state $(\Phi = \theta)$ is produced depending on the initial condition and for the repulsive coupling $(K < 0)$ cluster states with large Φ, either a 15-cluster state $(\Phi = 14\theta)$ or a 2-cluster state $(\Phi = 15\theta)$ is produced. Some examples of clustering behaviors are given in Fig.5.1, which depicts trajectories of time evolution of oscillators $\phi_k(t)(k = 1, \ldots, N)$. These cluster states are also produced in the case of finite time delay, as shown in Fig.5.2, $(K, \tau) = (1, 1)$. As mentioned in Sec.5.3.2, other linear stable cluster solutions can exist in the case studied above. However, we could not get such cluster states in our simulations. This suggests that the 1- or 30-cluster states in the case of attractive coupling and 2- or 15-cluster states in the case of repulsive coupling have the largest basin of attraction. When there is no time delay these results are physically explainable from an energy principle (Sec.5.4). We also examine the time evolution of the average frequency $\langle\dot{\phi}\rangle = (1/N)\sum_k \dot{\phi}_k(t)$ for various time delays τ when a 1-cluster state is produced under the initial condition (5.13). It can be seen from Fig.5.3 that the averaged frequency of clusters become small as the time delay τ becomes large. In the case of parameters studied in Fig.5.3, Eq.(5.12) has only one solution, which is indicated by symbols in Fig.5.3.

5.3.2 Linear stability of the cluster state

Next we study linear stability of various cluster states. For this purpose the initial condition is set as

$$\phi_k(t) = \Omega(t + \tau) + k\Phi + \varepsilon_k \quad (-\tau \le t \le 0), \tag{5.14}$$

where ε_k $(k = 1, \ldots, N)$ are random numbers in the range $[-\pi/100, \pi/100]$. This initial state is the one that is slightly perturbed from a cluster state (Φ). If this cluster state is restored in spite of a perturbation, then it is stable; otherwise this cluster state is unstable. Our simulation results are summarized in Table 1. In this table, "S" denotes

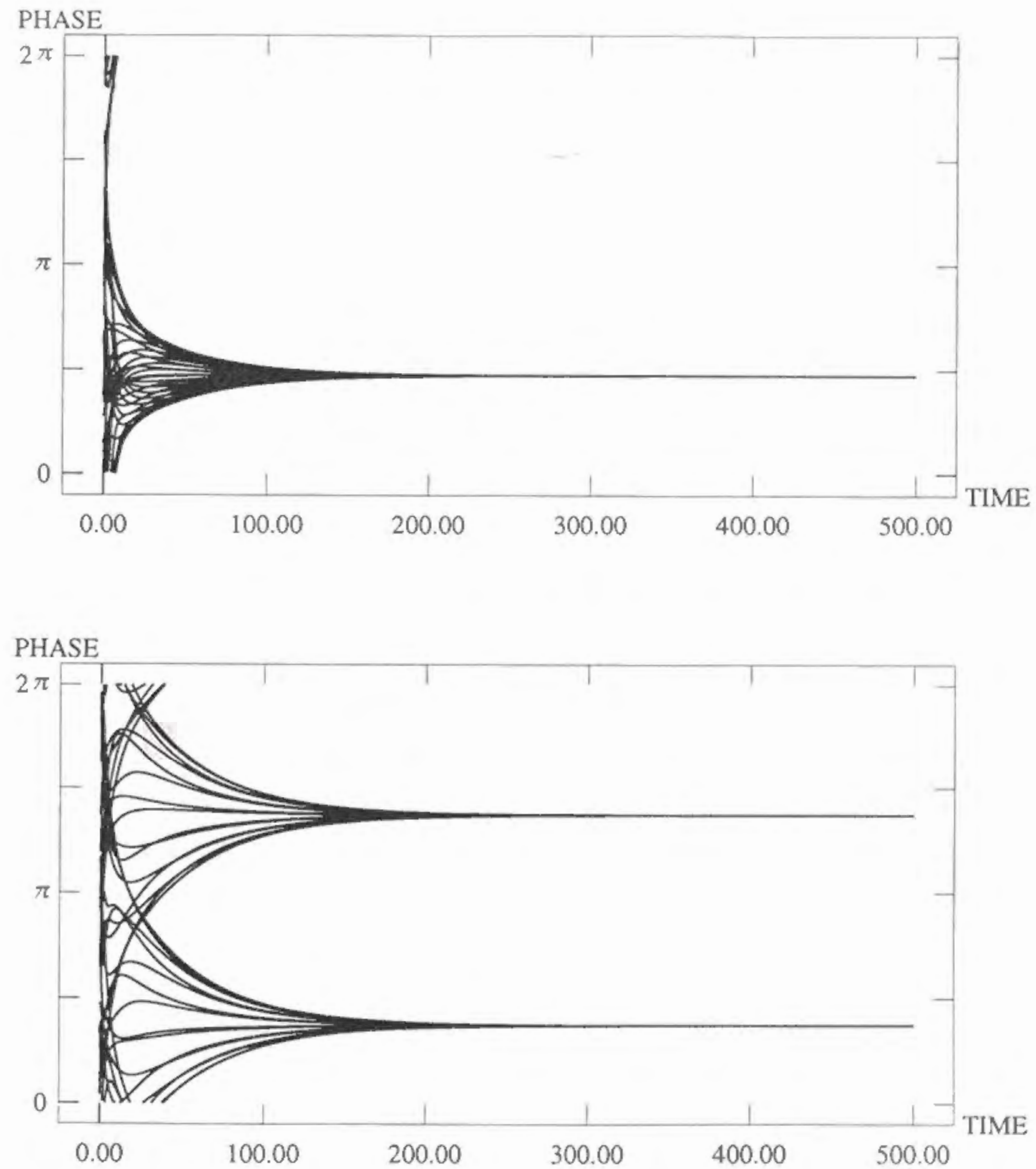

Figure 5.1: Time evolution of oscillators $\phi_k(t) - \Omega t$ $(k = 1, \ldots, 30)$ in the case of no time delay. (a) 1-cluster state $(\Phi = 0)$ for $(k, \tau) = (1, 0)$ with $\Omega = 1.0$ from Eq.(5.12). (b) 2-cluster state $(\Phi = \pi)$ for $(K, \tau) = (-1, 0)$ with $\Omega = 1.0)$ from Eq.(5.12).

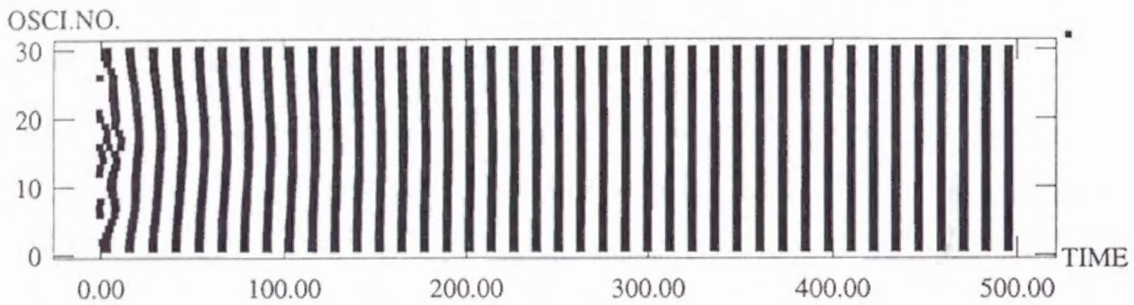

Figure 5.2: Time evolution of oscillators in the case of a finite time delay for $(k, \tau) = (1, 1)$ with $\Omega = 0.510973$ from Eq.(5.12). The 1-cluster state ($\Phi = 0$) is produced. The vertical axis denotes the number of oscillator. The dark region indicates that the phase takes a value between 0 and $\pi/4$.

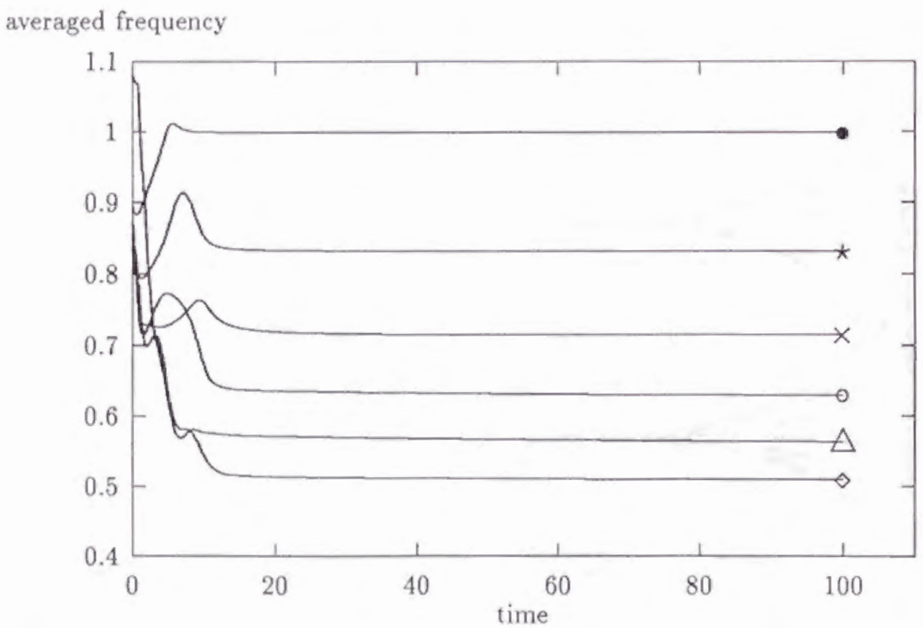

Figure 5.3: Time evolution of averaged frequencies $(1/N)\sum_k \dot{\phi}_k(t)$ for various time delays (from the top τ = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0)($K = 1$). Increasing a time delay induces frequency depression. The symbols indicate the solutions of Eq.(5.12) [$\bullet, 1.0(\tau = 0); \star, 0.833977(\tau = 0.2); \times, 0.717084(\tau = 0.4); \circ, 0.630602(\tau = 0.6); \triangle, 0.563973(\tau = 0.8); \diamond, 0.510973(\tau = 1.0)$], which precisely agree with the experiments.

Φ	N_c	$K=1$ $\tau=0$	$K=1$ $\tau=1$	$K=1$ $\tau=2$
0	1	S	S	S
θ	30	S	S	S
2θ	15	S	S	S
3θ	10	S	S	S
4θ	15	S	S	U
5θ	6	S	S	U
6θ	5	S	U	U
7θ	30	S	U	U
8θ	15	U	U	U
9θ	10	U	U	U
10θ	3	U	U	U
11θ	30	U	U	S
12θ	5	U	U	S
13θ	30	U	S	S
14θ	15	U	S	S
15θ	2	U	S	S

Table 5.1: Summary of computer experiments for linear stability with $K = 1$. Here the phase unit $\theta = \pi/15$ and N_c denotes the number of cluster. S means the cluster is stable and U unstable.

that the cluster state is stable and and "U" unstable. We show two examples of computer experiments in Fig.5.4.

It can be seen from this examples that the cluster state ($\Phi = 5\theta$) is stable (Fig.5.4(a)) and the cluster state ($\Phi = 6\theta$) is unstable (Fig.5.4(b)) in the case of $(K, \tau) = (1, 1)$. Table 1 shows that the stability of cluster states is modified by a time delay τ. These numerical results are verified theoretically in Sec.5.4. In passing we note that in the cases $(K, \tau) = (1, 0), (1, 1), (1, 2)$, Eq.(5.12) has only one solution Ω for any Φ. In the case of negative K, qualitatively similar results are obtained as to the linear stability of cluster states.

5.3.3 Relative stability among cluster states

In this section we study the relative stability among linearly stable cluster states. In the case $\tau \neq 0$, we have no theory at the moment to investigate this important nonlinear phenomenon and we rely on computer simulations with random noise.

We consider the case in which Eq.(5.12) has more than one solution. To be concrete We study the 1-cluster state ($\Phi = 0$), which is very stable for the attractive coupling (see subsection 5.3.1) for say, $(K, \tau) = (4.2)$. With this parameter Eq.(5.12) has five solutions $\Omega_1(= -2.58576), \Omega_2(= -1.99362), \Omega_3(= 0.111940), \Omega_4(= 1.65276)$ and $\Omega_5(= 2.89484)$ for $\Phi = 0$. In order to investigate the stability of these solutions, we employ the following initial conditions,

$$\phi_k(t) = \Omega_0(t + \tau) + \alpha \quad (-\tau \leq t \leq 0), \tag{5.15}$$

for all k, where α is an arbitrary constant in the range $[0, 2\pi]$ and Ω_0 is a parameter of our experiments. In Fig.5.5 we show one typical trajectory of an experiment with $\Omega_0 = -1.0$. In this case it is seen that the system is attracted to 1-cluster$state$(Ω_3). It is to be noted that all the oscillators have the same phase through the transition. In the experiment with $\Omega_0 = 2.3$ the final state was a 1-cluster state (Ω_5). From other similar experiments

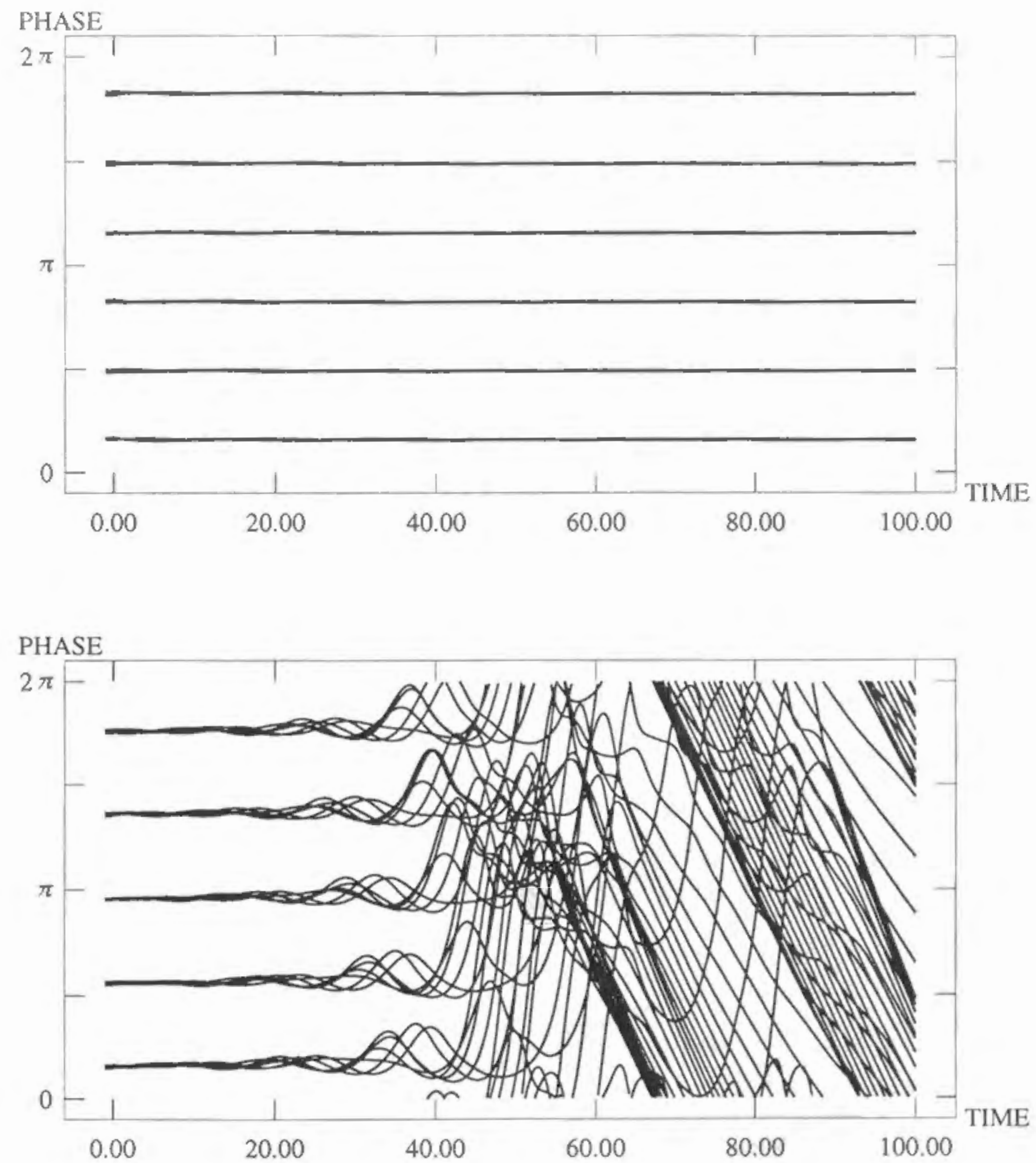

Figure 5.4: Time evolution of stable and unstable cluster states. (a)$(K, \tau) = (1, 1)$ and $\Phi = 5\theta$ with $\Omega = 0.684037$ from Eq.(5.12). The 6-cluster state ($\Phi = 5\theta$) is stable. (b) $(K, \tau) = (1, 1)$ and $\Phi = 6\theta$ with $\Omega = 0.782194$ from Eq.(5.12). The 5-cluster state ($\Phi = 6\theta$) is unstable and finally attracted to the antiphase (30-cluster) state.

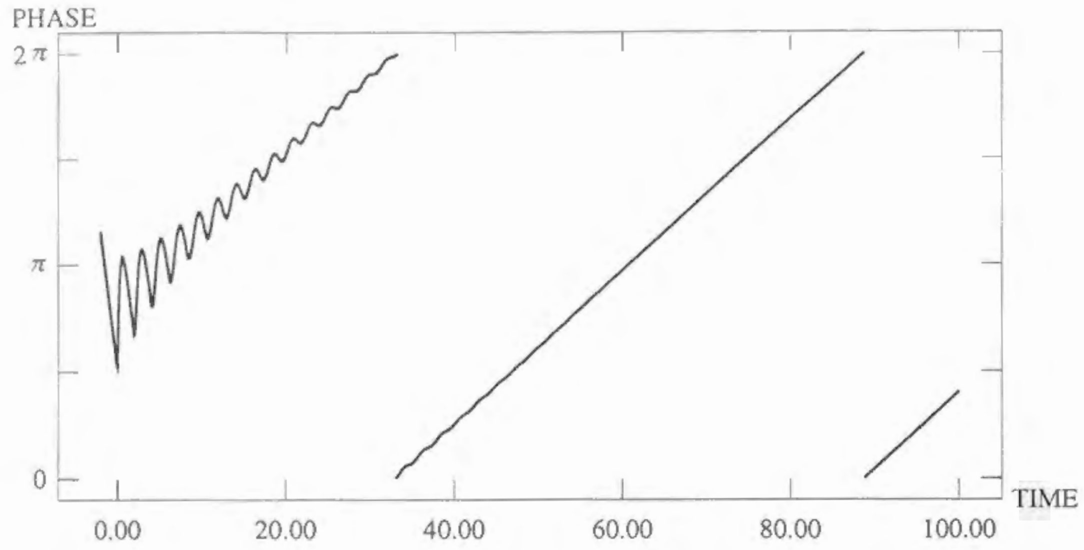

Figure 5.5: Time evolution of phases $\phi_k(t)$ $(k = 1, \ldots, 30)$ with initial condition (5.15) $(\Omega_0 = -1)$. The system is attracted to the 1-cluster state $(\Omega = \Omega_3)$.

and the theory given in Sec.5.4, we find that the 1-cluster state with $\Omega_1, \Omega_3 \text{and} \Omega_5$ are stable, and $\Omega_2 \text{and} \Omega_4$ are unstable.

Now the problem we pursue is the relative stability among these three stable 1-cluster states. In order to investigate this nonlinear problem we take the following Langevin equation.

$$\frac{d\phi_k(t)}{dt} = 1 + \frac{K}{2} \sum_{j=k\pm1} \sin\{\phi_j(t-\tau) - \phi_k(t)\} + \eta_k(t) \quad (k = 1, \ldots, N) \tag{5.16}$$

where the random force $\eta_k(t)$ denotes a Gaussian white noise

$$\begin{aligned} \langle \eta_k(t) \rangle &= 0 \\ \langle \eta_k(t)\eta_j(t') \rangle &= 2D\delta_{kj}\delta(t-t'). \end{aligned} \tag{5.17}$$

As stated in Sec.5.2 a cluster state is uniquely defined by the phase difference Φ and the frequency Ω_b of Eq.(5.12). The result of our experiment is shown in Fig.5.6. We set $D = 1.0 \times 10^{-7}K$ and take $\Omega_0 = \Omega_1$ in Eq.(5.15) for the initial condition. From Fig.5.6

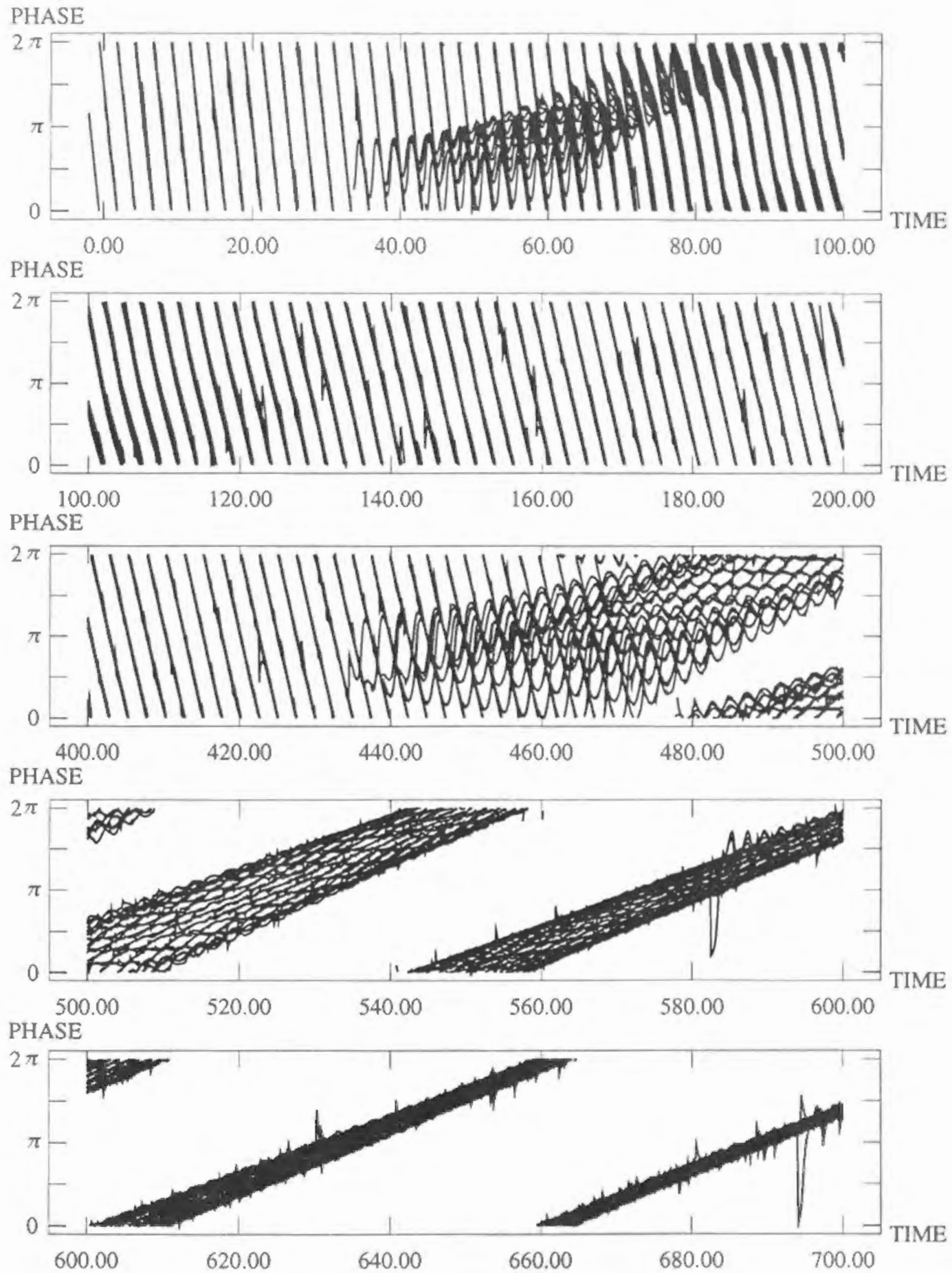

Figure 5.6: Time evolution of phases $\phi_k(t)(k = 1, \ldots, 30)$ under the noise for $(K, \tau) = (4, 2), D = 1.0 \times 10^{-7}K$. Initial state is the 1-cluster state $(\Phi = \tau)$ the system is finally attracted to the 1-cluster state. We note that trajectories from $t = 200$ to 400 are omitted. In this period the system was in the 2-cluster state. The system was in the 1-cluster state after $t = 700$ at least until $t = 1000$ when the experiment was stopped.

it can seen that 2-cluster states emerges as a transient state and then finally this state is attracted to the 1-cluster state (Ω_3). Similarly the 1-cluster state (Ω_5) is finally attracted to the 1-cluster state (Ω_3). On the contrary the 1-cluster state (Ω_3) remains stable under the noise disturbance. From these experiments we conclude that the 1-cluster state (Ω_3) is the most stable.

5.4 Theoretical analysis

5.4.1 Energy analysis for the system without delay

when there is no time delay, the energy of the system can be introduced to study properties of cluster states. In the case of $\tau = 0$ we modify Eq.(5.3) with use of

$$\tilde{\phi}_k(t) = 1 \cdot t + \hat{\phi}_k(t). \tag{5.18}$$

$\hat{\phi}_k(t)$ denotes the phase in coordinate rotating with frequency 1. Inserting Eq.(5.18) into Eq.(5.3) we obtain

$$\begin{aligned} \frac{d\phi_k}{dt} &= \frac{K}{2} \sum_{j=k\pm1} \sin(\phi_j - \phi_k) && (5.19) \\ &= -\frac{\partial V(\phi_1, \dots, \phi_N)}{\partial \phi_k}, && (5.20) \end{aligned}$$

where with the convention $\phi_{N+1} = \phi_1$,

$$V(\phi_1, \dots, \phi_N) = -\frac{K}{2} \sum_{j=1}^{N} \cos(\phi_{j+1} - \phi_j) \tag{5.21}$$

after omitting the tilde on $\phi_k(t)$. Therefor it follows that

$$\frac{dV}{dt} = \sum_k \frac{\partial V}{\partial \phi_k} \frac{d\phi_k}{dt} = -\sum_k \left(\frac{d\phi_k}{dt}\right)^2 \leq 0. \tag{5.22}$$

This means that energy $V(\phi_1, \dots, \phi_N)$ is the Lyapunov function for the system without a time delay. We also consider the energy of the cluster state which are introduced in

Sec5.2. Form the definition Φ by Eqs.(5.7) and (5.21), the energy per oscillator of the cluster state with phase Φ, E_{cs}, is

$$E_{cs} = \frac{V}{N} = -\frac{K}{2} \cos \Phi. \tag{5.23}$$

Based on the energy Eq.(5.23), we can examine the linear stability of cluster states. Let us consider the energy of the state which is slightly perturbed from a cluster state (Φ). If the phase difference is

$$\phi_{k+1} - \phi_k = \Phi + \delta\phi_k \tag{5.24}$$

where $\delta\phi_k$ represents a small perturbation, the energy per oscillator of this state is given as

$$\begin{aligned} E &= -\frac{K}{2N} \sum_{k=1}^{N} \cos(\Phi + \delta\phi_k) \\ &= E_{cs} + \frac{K}{2N} \left[\sum_{k=1}^{N} \delta\phi_k\right] \sin \Phi + \frac{K}{4N} \left[\sum_{k=1}^{N} (\delta\phi_k)^2\right] \cos \Phi. && (5.25) \end{aligned}$$

We note that due to the periodic boundary condition it holds that

$$\sum_{k=1}^{N} \delta\phi_k = 0 \tag{5.26}$$

Thus

$$E = E_{cs} + \frac{K}{4N} \left[\sum_{k=1}^{N} (\delta\phi_k)^2\right] \cos \Phi. \tag{5.27}$$

This equation tells us that $\pi/2 (= \Phi_c)$ is the critical point. For attractive coupling($K > 0$), if $\Phi < \Phi_c$ the perturbation makes the energy larger, which means that the cluster state(Φ) is linearly stable. The other case ($K < 0$) can be understood in the same way. This analysis explains the results of the case $(K, \tau) = (1, 0)$ of Table 5.1.

Next we consider the case in which the initial phases of oscillators are randomly distributed (see Sec.5.3.1). Equation (5.23) tells us that in the case of attractive (repulsive) coupling, the energy E_{cs} increases (decreases) as Φ becomes large in the range of $0 \leq \Phi\pi$.

The system starting from the random initial conditions is likely to fall into the attractor with lower energy. Therefore it easily happens that the system with attractive (repulsive) coupling is attracted to the cluster state with smallest (largest) Φ. This explains the results in Sec.5.3.1.

5.4.2 Linear stability analysis with a time-delay

In the cluster state the motion of oscillators is expressed by Eq.(5.11). We represent the phase of kth oscillator as

$$\phi_k(t) = \Omega t + k\Phi + \varepsilon a_k(t) \tag{5.28}$$

where ε is assumed to be very small. Inserting Eq.(5.28) into Eq.(5.6), we compare terms with the same order of ε. The equation we get for the order ε^0 is Eq.(5.12). For the order ε^1, we obtain an equation of motion for $a_k(t)$ as

$$\begin{aligned}\frac{da_k(t)}{dt} &= \frac{K}{2}\left[\{a_{k-1}(t-\tau) - a_k(t)\}\cos(\Omega\tau + \Phi)\right. \\ &\quad \left. + \{a_{k+1}(t-\tau) - a_k(t)\}\cos(\Omega\tau - \Phi)\right] \\ &\equiv \alpha\{a_{k-1}(t-\tau) - a_k(t)\} + \beta\{a_{k+1}(t-\tau) - a_k(t)\} \end{aligned} \tag{5.29}$$

where

$$a_1 = a_{N+1}, \quad a_0 = a_N \tag{5.30}$$

from the periodic boundary condition. We decompose $a_k(t)$ into Fourier modes as

$$a_k(t) = \sum_q f_q(t) e^{ikq}. \tag{5.31}$$

From Eq.(5.30) and Eq.(5.31) we can restrict q to

$$q = \frac{2\pi}{N} m \quad \left(-\frac{N}{2} < m \le \frac{N}{2}, m; \text{is an integer}\right). \tag{5.32}$$

From Eqs.(5.29) and (5.31) we obtain the equation for $f_q(t)$ as

$$\frac{df_q(t)}{dt} = (\alpha e^{-iq} + \beta e^{iq}) f_q(t-\tau) - (\alpha+\beta) f_q(t), \tag{5.33}$$

where

$$\alpha \equiv \frac{K}{2}\cos(\Omega\tau + \Phi) \quad \beta \equiv \frac{K}{2}\cos(\Omega\tau - \Phi). \tag{5.34}$$

Denoting the Laplace transform of $f_q(t)$ by $\hat{f}_q(s)$, we obtain from Eq.(33) the following equation:

$$\hat{f}_q(s) = \frac{C}{s + (\alpha+\beta) - (\alpha e^{-iq} + \beta e^{iq}) e^{-s\tau}}, \tag{5.35}$$

where C is the constant depending on the initial condition. Now we can examine the stability of the cluster state(Φ, Ω) as follows: First we calculate all the poles of $\hat{f}_q$(Eq.(5.35)) for each mode q. If none of real parts of the poles for each q exceed zero, this cluster state is stable. If there exists a mode q_0, for which a pole has a positive real part, this cluster state is unstable. It is to be noted that Eq.(5.35) has a large number of poles because of the term $e^{-s\tau}$

Examples of calculations are shown in Figs.5.7 and 5.8. In Fig.5.7 poles for $q = \frac{2\pi}{5}(m = 6)$ and $\Phi = 2\pi 5$ for $(K, \tau) = (1, 1)$ are plotted. Real part of one of the poles exceeds zero. Thus this cluster state is unstable. On the other hand, in Fig.5.8 poles for all q and $\Phi = \frac{\pi}{3}$ foe $(K, \tau) = (1, 1)$ are plotted. The largest value of real parts of pole is zero (Goldstone mode $q = 0$). Thus this cluster state is stable. These calculations can explain the results of Table 5.1 and are not contradictory to Fig.5.4.

5.5 Summary and remarks

In this paper we mainly studied the oscillator systems with NN coupling with a time delay. The cluster states is found to be a rather general pattern of motion in NN coupling systems and can be classified by the phase difference Φ between neighboring oscillators. The role of the time-delay consists in a modification of the collective frequency and of the linear stability of cluster states. In the case of no time delay, the results of the computer experiments on the linear stability of the cluster state and on the relative stability among cluster states are verified with theoretical analysis based on the energy of the system. In

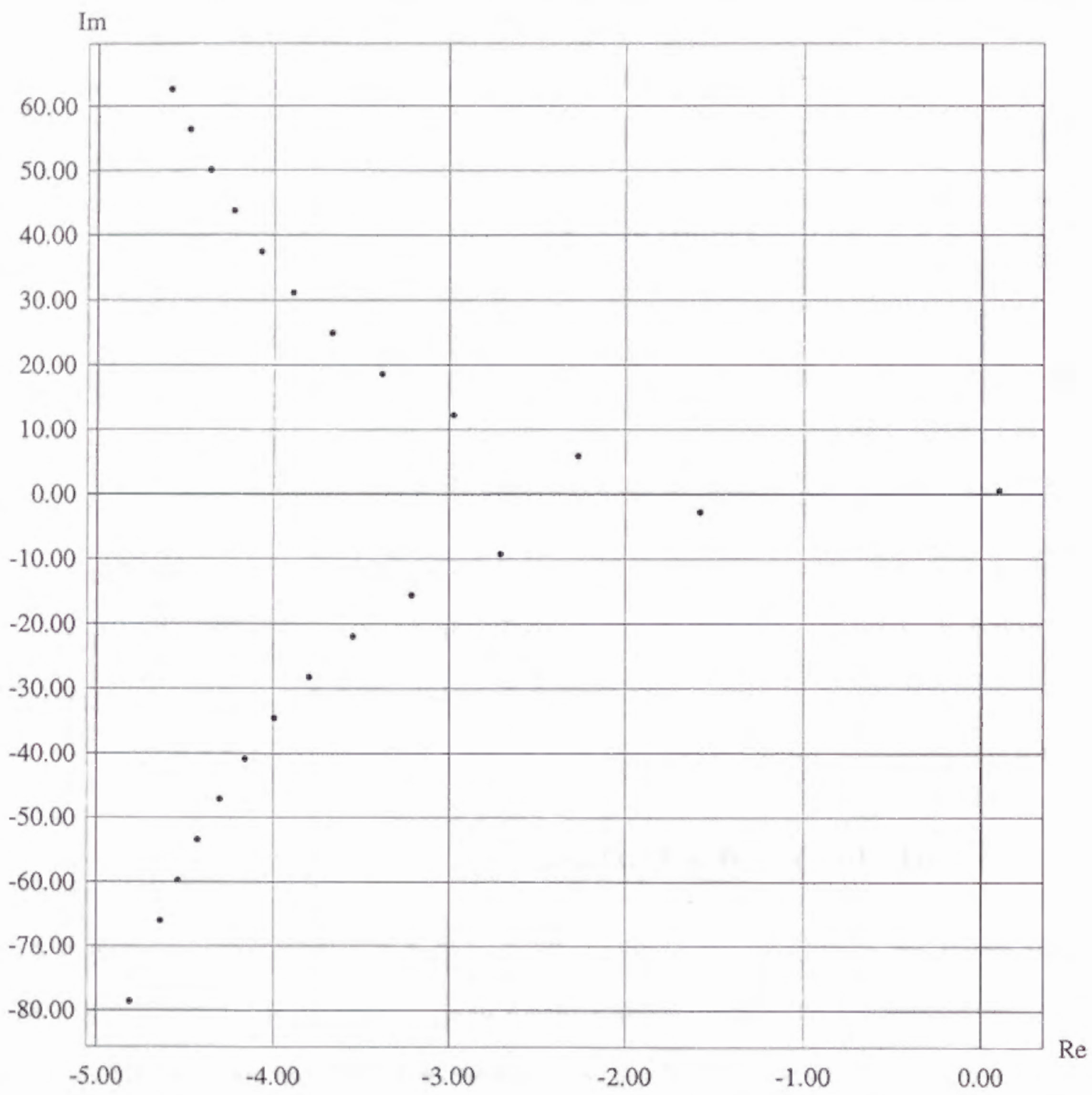

Figure 5.7: Poles for $q = 2\pi/5 (m = 6)$ and $\Phi = 2\pi/5$ in the case of $(K, \tau) = (1, 1)$.

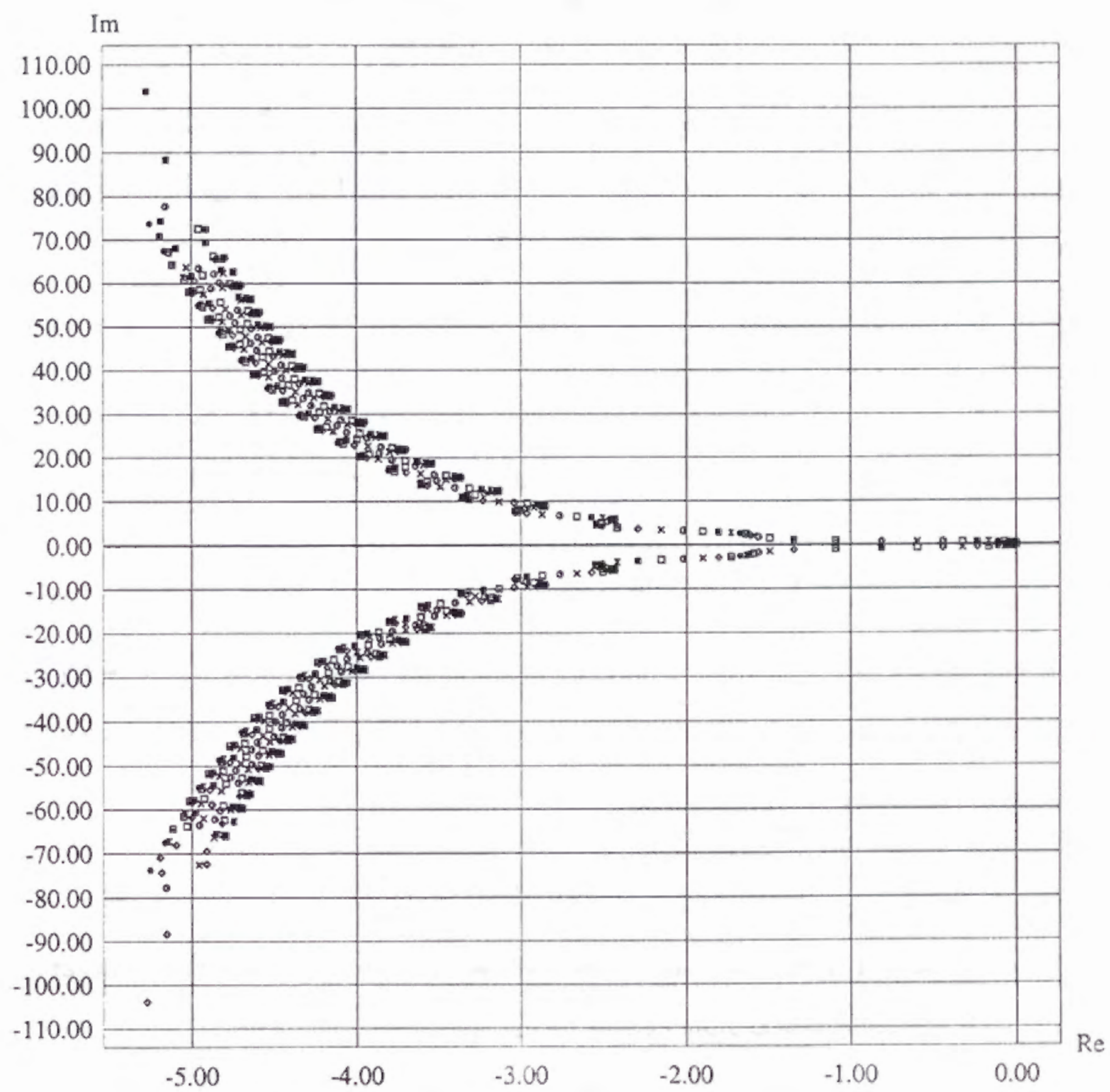

Figure 5.8: Poles for all q and $\Phi = \pi/3$ in the case of $(k, \tau) = (1, 1)$.

the case of a finite time delay, the linear stability can be analyzed with the technique of a Laplace transform, but we do not have a theory to discuss the relative (nonlinear) stability at present.

In the rest of this paper, we make two remarks; (1) on the extension to a higher dimensional coupled oscillator system and (2)on some computer experiments on the system with MF coupling

(1) We can extend our one-dimensional NN coupling model to high dimensional one. For example, in the case of the two-dimensional oscillator system with NN coupling the equations of motion becomes

$$\frac{d\phi_{k_1k_2}(t)}{dt} = 1 - \frac{K}{4}\sum_{(i,j)}{}' \sin\{\phi_{ij}(t-\tau) - \phi_{k_1k_2}(t)\} \quad (k_1 = 1,\ldots,N_1; k_2 = 1,\ldots,N_2), \quad (5.36)$$

where $\sum'_{(i,j)}$ means the summation over the nearest neighbors of the (k_1, k_2) oscillator and N_1, N_2 are numbers of oscillators in the 1-axis and 2-axis, respectively. The cluster state can be expressed as

$$\phi_{k_1k_2} = \Omega t + k_1\Phi_1 + k_2\Phi_2, \quad (5.37)$$

where Φ_1 and Φ_2 are phase differences between neighboring oscillators in 1-axis and 2-axis respectively and N_1 and N_2 are numbers of oscillators in the 1-axis and 2-axis, respectively. The collective frequency is calculated as

$$\Omega = 1 - \frac{K}{2}\sin(\Omega\tau)(\cos\Phi_1 + \cos\Phi_2). \quad (5.38)$$

Similarly in the case of n-dimensional systems, the equations of motion, the phase of the oscillators, and the collective frequency can be expressed as

$$\frac{d\phi_{k_1,\ldots,k_n}(t)}{dt} = 1-\frac{K}{2n}\sum_{(i_1,\ldots,i_n)}{}' \sin\{\phi_{i_1,\ldots,i_n}(t-\tau)-\phi_{k_1,\ldots,k_n}(t)\} \quad (k_1 = 1,\ldots,N_1;\ldots;k_n = 1,\ldots,N_n), \quad (5.39)$$

$$\phi_{k_1\cdots k_n} = \Omega t + k_1\Phi_1 + \cdots + k_n\Phi_n \quad (5.40)$$

$$\Omega = 1 - \frac{K}{n}\sin(\Omega\tau)(\cos\Phi_1 + \cdots + \cos\Phi_n). \quad (5.41)$$

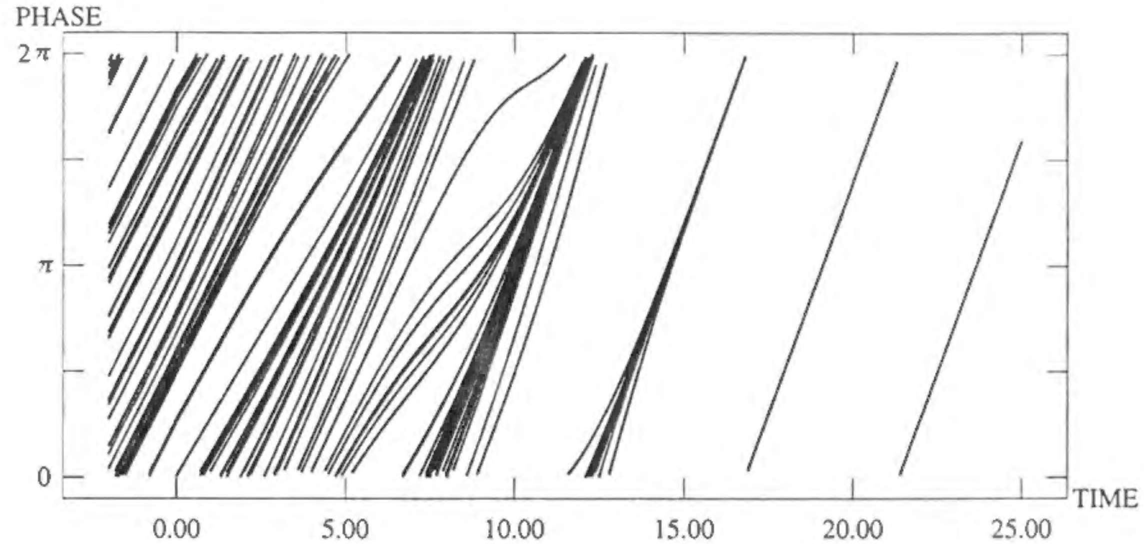

Figure 5.9: Time evolution of oscillators $\phi_k(t)$ in the case of MF coupling for $(K,\tau) = (-1,2)$.

(2) In Sec5.3.2 it was shown that the linear stability of the cluster state in the NN coupling system is modified by a time delay (see Table 5.1). We also made the experiments on the system with MF coupling for the comparison. Sin(x) coupling was chosen as in Eq.(5.3) and a similar tendency as to the stability of the cluster state is acquired. We show a result of the case for the 1-cluster state. According to the result of Okuda [64], the 1-cluster state was not produced in the case of MF repulsive coupling with $\tau = 0$. This means that the 1-cluster state is unstable. However, in the case $(K,\tau) = (-1,2)$ the 1-cluster state was produced as shown in Fig.5.9, indicating that the time delay has modified the stability of the 1-cluster state.

Chapter 6

Conclusions

In the preceding chapters, we investigated collective phenomena of neural networks and couple oscillator systems with use of equilibrium and nonequilibrium statistical mechanics. Here we summarize our results.

1. The dynamics of retrieval process in the Hopfield model was studied with use of the Mori-Zwanzig projection operator formalism for irreversible processes. Exact evolution equations for overlaps were derived under the condition that finite number of patterns are embedded, $\alpha = 0$. There has been only a few methods for the study of dynamical behavior of neural networks, therefore we expect that our approach will make a further progress in such a study. Especially, it is hoped that overlap dynamics for finite α case could be treated with our approach.

2. Neural networks with correlated patterns were investigated and we proposed a new model for associative memory of correlated patterns. In our model, driven by correlations among patterns, time sequence of patterns can be retrieved. Overlap dynamics were considered and exact evolution equations were derived. The agreement between numerical solution to these equations and the simulation results was excellent.

3. We proposed a neural network composed of D-dimensional spin neurons as an extension of the Hopfield model to a multi-dimensional one. It was confirmed with computer simulations ($D = 2, 3$) that our model works as an associative memory. A free-energy of our model was derived in the case of general dimension D with use of replica symmetric theory. In the case of $D = 1$, the free-energy was confirmed to coincide with that of the Hopfield model. Mean-field equations were derived from this free-energy, and $T - \alpha$ phase diagram was calculated with these equations. This phase diagram tells us fundamental properties of the network: patterns can be retrieved if the temperature T is lower than $1/D$; storage capacity was $\alpha_c = 0.0743(D = 2)$ and $\alpha_c = 0.0432(D = 3)$ for $T = 0$; the retrieval regime shrinks as dimension D increases as we expected. Some applications of our model is expected to appear in near future.

4. Clustering behavior in a time-delayed nearest-neighbor coupled oscillators was investigated numerically and theoretically. From computer simulations, we found that cluster states can be produced in the nearest-neighbor coupling systems and that the cluster states can be classified by the phase-difference between neighboring oscillators. The time-delay was found to modify the collective frequency and linear stability of cluster states. Linear and nonlinear stability was analyzed theoretically. When there is no time-delay, the results of computer simulation on the linear and nonlinear stability was explained theoretically, but in the case of finite time-delay only linear stability was explained. It is an open problem to explain the nonlinear stability of cluster states in the case of finite time-delay. The results obtained here may be applied to the modeling of neural networks, where clustering patterns could correspond to memories.

In this thesis, we discussed the properties of neural networks as collective phenomena. When we study neural networks from such a viewpoint, traditional methods of statistical mechanics can be naturally applied, and actually, the theory of neural networks has extensively developed with the methods. The theoretical approach with the methods of statistical mechanics in the field of neural networks will be more and more important to gain deeper insight into the modeling of neural networks. æ

Bibliography

[1] P. Peretto, *An Introduction to the Modeling of Neural Networks* (Cambridge University Press, Cambridge, 1989).

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

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

[4] F. Rosenblatt, *Principles of Neurodynamics* (Spartan Books, Washington DC, 1962).

[5] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79** (1982) 2554.

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

[7] E. Ising, Z. Phys, **31** (1925 253.

[8] J. L. van Hemmen and R. Kühn, Phys. Rev. Lett. **57** (1988) 913.

[9] R. Kühn and J. L. van Hemmen, *in Models of Neural Networks*, edited by E. Domany, J. L. van Hemmen and K. Schulten, (Springer-Verlag, Berlin, 1991).

[10] U. Riedel, R. Kühn and J. L. van Hemmen, Phys. Rev. A**38** (1988) 1105.

[11] S. Amari, IEEE trans. **C-21** (1972) 1197.

[12] P. Peretto and J. J. Niez, *in Disordered Systems and Biological Organization*, edited by E. Bienenstock, F. forgelman-Soulié and G. Weisbbuch, (Springer, Berlin, Heidelbeg, New York, 1986) 171.

[13] S. Dehaene, J. P. Changeux and J. P. Nadal, Proc. Natl. Acad. Sci. U.S.A. **84** (1987) 2727.

[14] J. Buhmann and K. Schulten, Europhys. Lett. **4** (1987) 1205.

[15] H. Nishimori, T. Nakamura and M. Shiino, Phys. Rev. A**41** (1988) 1105.

[16] D. kleinfeld, Proc. Natl. Acad. Sci. U.S.A. **83** (1986) 9469.

[17] H. Sompolinsky and I. Kanter, Phys. Rev. Lett. **57** (1986) 2861.

[18] C. M. Gray, P. König, A. E. Engel and W. Singer, Nature **338** (1989) 334.

[19] Edited by H. G. Schuster, *Nonlinear Dynamics and Neural Networks* (VCH, 1991).

[20] R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse and H. J. Reiboeck, Biol. Cybern. **60** (1988) 121.

[21] J. J. Collins and I. N. Stewart, J. Nonlin. Sci. (1992) 349.

[22] H. Sompolinsky, D. Golomb and D. Kleinfeld, Phys. Rev. A**43** (1991) 6990.

[23] H. G. Schuster and P. Wagner, Biol. Cybern. **64** (1990) 77, 83.

[24] L. F. Abbot, J. Phys. A**23** (1990) 3835.

[25] T. Munakata and Y. Nakamura, Phys. Rev. E**47** (1993) 3791.

[26] Y. Nakamura and T. Munakata, Phys. Rev. E**49** (1994) 1775.

[27] J. Cook, J. Phys. A**22** (1989) 2057.

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

[29] D. J. Amit, H. Gutfreund and H. Sompolinsky, Ann. Phys. **173** (1987) 30.

[30] D. J. Amit, *Modeling Brain Function* (Cambridge University Press, Cambridge, 1989).

[31] Y. Nakamura ,F. Tominaga and T. Munakata, Phys. Rev. E**49** (1994) 4849.

[32] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. **A32** (1985) 1007; **A35** (1987) 2293.

[33] H. Horner, D. Bormann, M. Frick, H. Kinzelbach and A. Schmidt, Z. Phys. **B76** (1989) 381.

[34] H. Rieger, M. Schreckenberg, and J. Zittartz, Z. Phys. **B72** (1988) 523.

[35] R. Zwanzig, Phys. Rev. **124** (1961) 983.

[36] H. Mori, Prog. Theor. Phys. **33** (1965) 423.

[37] H. Mori and H. Fujisaka, Prog. Theor. Phys. **49** (1973) 764.

[38] A. Z. Akcasu, and J. J. Duderstadt, Phys. Rev. **188** (1969) 479.

[39] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (W.A.Benjamin, London, 1975).

[40] A. C. C. Coolen, and Th. W. Ruijgrok, Phys. Rev. **A38** (1988) 4235.

[41] M. Shiino, H. Nishimori, and M. Ono, J. Phys. Soc. Jpn. **58** (1989) 763.

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

[43] W. A. Little, Math. Biosci. **6** (1974) 101.

[44] D. J. Amit, Proc. Natl. Acad. Sci. U.S.A. **85** (1988) 2141.

[45] D. J. Amit, H. Gutfreund and H. Sompolinsky, Phys. Rev. Lett. **55** (1985) 1530.

[46] H. Rieger, J. Phys. A**23** (1990) L1273.

[47] D. Bollé and J. von Mourik, J. Phys. A**27** (1994) 1151.

[48] D. Bollé, G. M. Shim, B. Vinck, and V. A. Zagrebnov, J. Stat. Phys. **74** (1987) 565.

[49] I. Kanter, Phys. Rev. A**37** (1988) 2739.

[50] D. Bollé, P. Dupont, and J. von Mourik, J. Phys. A**24** (1991) 1065.

[51] A. J. Noest, Phys. Rev. A**38** (1988) 2196.

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

[53] P. Hadray, M. R. Beasley and K. Wiesenfeld, Phys. Rev. B**38** (1988) 8712.

[54] D. Fisher, Phys. Rev. B**31** (1985) 1396.

[55] C. M. Gray, P. König, A. E. Engel and W. Singer, Nature **338** (1989) 334.

[56] A. T. Winfree, J. Theor. Biol. **16** (1967) 15.

[57] H. Sompolinsky, D. Golomb and D. Kleinfeld, Phys. Rev. A**43** (1991) 6990.

[58] H. G. Shuster and P. Wagner, Biol. Cybern. **64** (1990) 77,83.

[59] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, (Springer, Berlin, 1991)

[60] Y. Kuramoto and I. Nishikawa, J. Stat. Phys. **49** (1987) 569.

[61] S. Sakaguchi, S. Shinomoto and Y. Kuramoto, Prog. Theor. Phys. **79** (1988) 600.

[62] S. Shinomoto and Y. Kuramoto, Prog. Theor. Phys. **75** (1986) 1319.

[63] D. Golomb, D. Hansel, B. Shraiman and H. Sompolinsky, Phys. Rev. A**45** (1992) 3516.

[64] K. Okuda, Physica D**63** (1993) 424.

[65] H. Daido, Phys. Rev. Lett. **61** (1988) 231.

[66] S. H. Strogatz and R. E. Mirollo, Physica D**31** (1988) 143.

[67] E. Niebur, H. G. Shuster, D. M. Kammen and C. Koch, Phys. Rev. A**44** (1991) 6895.

[68] T. Aoyagi and Y. Kuramoto, Phys. Lett. A**155** (1991) 410.

[69] H. G. Shuster and P. Wagner, Prog. Theor. Phys. **81** (1989) 939.

[70] E. Niebur, H. G. Shuster and D. M. Kammen, Phys. Rev. Lett. **67** (1991) 2753.

[71] C. M. Marcus and R. M. Westervelt, Phys. Rev. A**39** (1989) 347.

[72] K. Wiesenfeld, C. Golubitsky, G. James and R. Roy, Phys. Rev. Lett. **65** (1990) 1749.

[73] K. Wiesenfeld and P. Hakley, Phys. Rev. Lett. **62** (1989) 1335.

Appendix A

Theory of sublattice for the overlap dynamics

In this appendix, we shall derive a set of coupled nonlinear differential equations which describe the time evolution of the overlaps of the Hopfield model in the thermodynamic limit, $N \to \infty$, under the condition that the number of patterns, p, remains finite ($p/N \to \infty$). A stochastic Glauber dynamics is used;

$$\text{Prob}\{S_i(t+\Delta t) = \pm 1\} = \frac{1}{2}\{1 \pm \tanh[\beta h_i(t)]\}. \tag{A.1}$$

The key idea of the following analysis is based on the notion of sublattice $I(\vec{x})$ and sublattice magnetization $m(\vec{x};t)$ proposed by Hemmen et.al. Given p binary patterns

$$\vec{\xi}^\mu = (\xi_1^\mu, \xi_2^\mu, \ldots, \xi_N^\mu), \quad (\mu = 1, \ldots, p), \tag{A.2}$$

at every site i there are p ± 1's. We express each p dimensional vector on the ith site as

$$\vec{\xi}_i = (\xi_i^1, \xi_i^2, \ldots, \xi_i^p), \quad (i = 1, \ldots, N). \tag{A.3}$$

There are altogether 2^p different possible combinations of the p bits at any site, namely the 2^p corners $\vec{x}$ of the hypercube $[-1,1]^p$. Introducing the corresponding sublattices

$$I(\vec{x}) = \{i; \vec{\xi}_i = \vec{x}\}, \quad \vec{x} \in C^p = \{-1,1\}^p, \tag{A.4}$$

and sublattice magnetizations

$$m(\vec{x};t) = \frac{1}{|I(\vec{x})|} \sum_{i \in I(\vec{x})} S_i(t), \tag{A.5}$$

one can easily calculate the local field on ith neuron as

$$h_i(t) = \sum_{\mu,\vec{x}} \xi_i^\mu p_N(\vec{x}) x^\mu m(\vec{x};t).. \tag{A.6}$$

Here $|I(\vec{x})|$ denotes the size of $I(\vec{x})$, $p_N(\vec{x}) = |I(\vec{x})|/N$. Since $\vec{\xi}_i = \vec{x}$ for all i in $I(\vec{x})$, the local field depends on i only through the label $\vec{x}$ of the sublattice to which i belongs. Thus (A.6) can be written as

$$h(\vec{x};t) = \sum_{\mu,\vec{y}} x^\mu p_N(\vec{y}) y^\mu m(\vec{y};t). \tag{A.7}$$

To derive a master equation for the probability $\rho(\vec{m};t)$ of finding the system at time t in a state with sublattice magnetizations $\vec{m} = (m(\vec{x}), \vec{x} \in C^p)$, we note that the $m(\vec{x})$ can only assume the values $-1, -1+\Delta(\vec{x}), -1+2\Delta(\vec{x}), \ldots, 1$ with $\Delta(\vec{x}) = 2/|I(\vec{x})|$, and that an asynchronous updating effects only single sublattice, changing the corresponding sublattice magnetization by at most $\pm\Delta(\vec{x})$. In order to find the system at time t in the state with sublattice magnetization $\vec{m}$,

1. the system at time t is in the state with sublattice magnetization $\vec{m} + \vec{\Delta}(\vec{x})$ and the site which is to update flips from $+1 \to -1$.

2. the system at time t is in the state with sublattice magnetization $\vec{m} - \vec{\Delta}(\vec{x})$ and the site which is to update flips from $-1 \to +1$.

3. the system at time t is in the state with sublattice magnetization $\vec{m}$ and the site which is to update doesn't flip $(+1 \to +1)$.

4. the system at time t is in the state with sublattice magnetization $\vec{m}$ and the site which is to update doesn't flip $(-1 \to -1)$.

Let us consider the first case. Suppose that among $i; i \in I(\vec{x})$, the number of spins which are $+1$ is X and the number of spins which are -1 is Y. According to (A.5),

$$\begin{aligned} X+Y &= |I(\vec{x})| \\ X-Y &= (m(\vec{x})+\Delta(\vec{x}))|I(\vec{x})|. \end{aligned}$$

Thus the probability $\rho(\vec{m}; t+\Delta t)$ of finding the system at time $t+\Delta t$ in the state with sublattice magnetization $\vec{m}$ is

$$\begin{aligned} \rho(\vec{m}; t+\Delta t) &= \rho(\vec{m}+\vec{\Delta}(\vec{x}); t)\frac{X}{X+Y}\mathrm{Prob}\{S_i(t+\Delta t)=+1\} \\ &= \rho(\vec{m}+\vec{\Delta}(\vec{x}); t)\frac{1+m(\vec{x})+\Delta(\vec{x})}{2}\frac{1-\tanh[\beta h(\vec{x},\vec{m}+\vec{\Delta}(\vec{x}))]}{2}. \quad (A.8) \end{aligned}$$

Considering the other three cases in the same way and taking into account that the probability of $\vec{\xi}_i = \vec{x}$ is $p_N(\vec{x})$,

$$\begin{aligned} &\rho(\vec{m}; t+\Delta t) \\ &= \sum_{\vec{x}\in C^q} p_N(\vec{x})\left\{\frac{1+m(\vec{x})+\Delta(\vec{x})}{4}(1-\tanh[\beta h(\vec{x},\vec{m}+\vec{\Delta}(\vec{x}))])\rho(\vec{m}+\vec{\Delta}(\vec{x});t)\right. \\ &+ \frac{1-m(\vec{x})+\Delta(\vec{x})}{4}(1+\tanh[\beta h(\vec{x},\vec{m}-\vec{\Delta}(\vec{x}))])\rho(\vec{m}-\vec{\Delta}(\vec{x});t) \\ &+ \frac{1+m(\vec{x})}{4}(1+\tanh[\beta h(\vec{x},\vec{m})])\rho(\vec{m};t) \\ &+ \left.\frac{1-m(\vec{x})}{4}(1-\tanh[\beta h(\vec{x},\vec{m})])\rho(\vec{m};t)\right\}. \quad (A.9) \end{aligned}$$

Subtracting $\rho(\vec{m}; t)$ from both sides, we obtain the master equation for $\rho(\vec{m}; t)$ as

$$\begin{aligned} &\rho(\vec{m}; t+\Delta t) - \rho(\vec{m}; t) \\ &= \sum_{\vec{x}\in C^q} p_N(\vec{x})\left\{\frac{1+m(\vec{x})+\Delta(\vec{x})}{4}(1-\tanh[\beta h(\vec{x},\vec{m}+\vec{\Delta}(\vec{x}))])\rho(\vec{m}+\vec{\Delta}(\vec{x});t)\right. \\ &+ \frac{1-m(\vec{x})+\Delta(\vec{x})}{4}(1+\tanh[\beta h(\vec{x},\vec{m}-\vec{\Delta}(\vec{x}))])\rho(\vec{m}-\vec{\Delta}(\vec{x});t) \\ &- \frac{1+m(\vec{x})}{4}(1-\tanh[\beta h(\vec{x},\vec{m})])\rho(\vec{m};t) \\ &- \left.\frac{1-m(\vec{x})}{4}(1+\tanh[\beta h(\vec{x},\vec{m})])\rho(\vec{m};t)\right\}. \quad (A.10) \end{aligned}$$

Here we have introduced the vector $\vec{\Delta}(\vec{x}) = (\Delta(\vec{x})\delta_{\vec{x}\vec{y}}; \vec{y} \in C^q)$. Equation (A.10) can be written as a differential equation

$$\begin{aligned} &\left[\exp\left(\Delta t\frac{\partial}{\partial t}\right)-1\right]\rho(\vec{x};t) \\ &= \sum_{\vec{x}\in C^q} p_N(\vec{x})\left\{-\frac{1}{2}\sinh\left(\Delta(\vec{x})\frac{\partial}{\partial m(\vec{x})}\right)(F(\vec{x},\vec{m})\rho(\vec{m};t))\right. \\ &+ \left.\frac{1}{2}\left[\cosh\left(\Delta(\vec{x})\frac{\partial}{\partial m(\vec{x})}\right)-1\right](D(\vec{x},\vec{m})\rho(\vec{m};t))\right\}, \quad (A.11) \end{aligned}$$

where

$$\begin{aligned} F(\vec{x},\vec{m}) &= -\{m(\vec{x}) - \tanh[\beta h(\vec{x},\vec{m})]\} \\ D(\vec{x},\vec{m}) &= 1 - m(\vec{x})\tanh[\beta h(\vec{x},\vec{m})]. \end{aligned}$$

For asynchronous dynamics the elementary time step Δt must be scaled with system size N according to $\Delta t = (\Gamma N)^{-1}$ to ensure that neuron is updated at a finite N-independent rate. Neglecting, as $N \to \infty$, terms of order N^{-1} and smaller in (A.11), we obtain differential equation

$$\frac{\partial\rho}{\partial t} = \Gamma\sum_{\vec{x}\in C^q}\left\{-\frac{\partial}{\partial m(\vec{x})}(F\rho)\right\}. \quad (A.12)$$

If the patterns are uncorrelated, $p(\vec{x})$ is simply given by $1/2^p$. With equation (A.12) one obtains the following set of 2^p coupled differential equations for the sublattice magnetizations $m(\vec{x};t) = \langle m(\vec{x})\rangle_{\rho(m;t)}$

$$\frac{dm(\vec{x};t)}{dt} = -\Gamma\{m(\vec{x};t) - \tanh[\beta h(\vec{x},t)]\}. \quad (A.13)$$

Furthermore one can reduce (A.13) to a set of p equations for the overlaps themselves

$$\frac{1}{\Gamma}\frac{dm^\mu(t)}{dt} = -m^\mu(t) + \left\langle x^\mu \tanh\left[\beta\sum_\nu x^\nu m^\nu(t)\right]\right\rangle \quad (A.14)$$

because the overlaps $m^\mu(t)$ can be expressed in terms of the sublattice magnetizations as

$$m^\mu(t) = \sum_{\vec{x}} p(\vec{x})x^\mu m(\vec{x};t) \equiv \langle x^\mu m(\vec{x};t)\rangle. \quad (A.15)$$

Appendix B

Review of AGS theory for the Hopfield model with $\alpha \neq 0$

As we have stressed in earlier chapters, neural networks are large interesting systems of simple units, like the physical systems studied in statistical mechanics. The methods and concepts of statistical physics are therefore natural tools for the study of neural networks. In this appendix, we briefly illustrate the theory developed by Amit, Gutfreund and Sompolinsky.

B.1 The replica method

In order to know the equilibrium properties of a system, we calculate the free-energy of the system. Since the free-energy generally depends on a set of embedded patterns, we should average it over the distribution of patterns. When the number of embedded patterns is kept finite as $N \to \infty$, the self-averaging produces an averaging over the distribution of patterns. However, when p is of order of N, this is no longer the case, and we are forced to do the averaging over the distribution of patterns more systematically. What is of interest to us is the average $\langle\langle \log Z \rangle\rangle$ over the distribution of all random binary patterns; this gives us the average free-energy whose derivatives give the average quantities we want to know,

such as the overlaps $m^{\mu}(\mu = 1, \ldots, p)$. Here Z is the partition function

$$Z = \mathrm{Tr}_S \exp(-\beta E\{S_i\}) \tag{B.1}$$

where the trace, Tr_S, means a sum over all possible states, $\{S_i = \pm 1\}$. Unfortunately this average is very hard to calculate directly. Luckily there is a technique, called the replica method, which lets us circumvent averaging $\log Z$. It is based on the identity

$$\log Z = \lim_{n \to 0} \frac{Z^n - 1}{n} \tag{B.2}$$

which allows us to compute $\langle\langle \log Z \rangle\rangle$ from knowledge of $\langle\langle Z^n \rangle\rangle$. Note that we need this for the parameter n close to 0, but we ignore that for a while and focus on $\langle\langle Z^n \rangle\rangle$ for integer n. In that case we can think of Z^n as the partition function of n copies, or replicas, of the original system, writing

$$Z^n = \mathrm{Tr}_{S^1} \mathrm{Tr}_{S^2} \cdots \mathrm{Tr}_{S^n} \exp[-\beta(E\{S_i^1\} + \cdots + E\{S_i^n\})]. \tag{B.3}$$

Each copy is labeled by a superscript on S_i, running from 1 to n.

B.2 The free-energy and the mean-field equations

We now proceed to employ the replica method to the computation of the free-energy of a system with a coupling matrix,

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^{\mu} \xi_j^{\mu} (1 - \delta_{ij}) \tag{B.4}$$

with $p = \alpha N$. The free-energy per spin averaged over the embedded patterns becomes

$$f = -\beta \lim_{n \to 0} \lim_{N \to \infty} \frac{1}{nN} (\langle\langle Z^n \rangle\rangle - 1). \tag{B.5}$$

We concentrate therefore on the computation of $\langle\langle Z^n \rangle\rangle$. It is written as

$$\begin{aligned} \langle\langle Z^n \rangle\rangle &= \langle\langle \mathrm{Tr}_S^{\rho} \exp(-\beta E\{S_i^{\rho}\}) \rangle\rangle \\ &= \left\langle\left\langle \mathrm{Tr}_S^{\rho} \exp\left(\frac{\beta}{2N} \sum_{ij\mu\rho} (\xi_i^{\mu} S_i^{\rho})(\xi_j^{\mu} S_j^{\rho}) - \frac{1}{2}\beta p n + \sum_{\nu=1}^{s} h^{\nu} \sum_{i\rho} \xi_i^{\nu} S_i^{\rho} \right) \right\rangle\right\rangle . \end{aligned} \tag{B.6}$$

Note that the trace is over the configurations of the n-fold replicated systems, labeled by $\rho(=1,\ldots,n)$. The second term, $\frac{1}{2}\beta pn$, in the exponent corrects for the presence of the self-interaction in the first, and the last term represents the effect of s external fields, each coupled to one of a finite number of patterns which are candidates for condensation. The difficulty in carrying our the trace is caused by the presence of products of S_i^ρ's with different i's in the quadratic term in the exponent. But it can be simplified by a Stratonovich-Hubbard transformation,

$$\exp\left(\frac{A}{2}a^2\right) = \frac{1}{\sqrt{2\pi A}}\int_{-\infty}^{\infty} dx \exp\left(-\frac{x^2}{2A}+ax\right), \tag{B.7}$$

and we get

$$\begin{aligned}\langle\langle Z^n\rangle\rangle &= e^{-\beta pn/2}(\beta N)^p n/2 \mathrm{Tr}_S^\rho \int_{-\infty}^{\infty}\prod_{\mu\nu\rho}\frac{dm_\rho^{\mu\nu}}{\sqrt{2\pi}} \\ &= \left\langle\!\!\left\langle \exp\left\{\beta N\left(-\frac{1}{2}\sum_{\mu\rho}(m_\rho^\mu)^2 + \sum_{\mu\rho} m_\rho^\mu \frac{1}{N}\sum_{i=1}^N \xi_i^\mu S_i^\rho\right)\right\}\right. \\ &= \left.\exp\left\{\beta N\left(-\frac{1}{2}\sum_{\nu\rho}(m_\rho^\mu)^2 + \sum_{\nu\rho}(m_\rho^\nu + h^\nu)\frac{1}{N}\sum_{i=1}^N \xi_i^\nu S_i^\rho\right)\right\}\right\rangle\!\!\right\rangle.\end{aligned} \tag{B.8}$$

Note that on the right hand side there are pn integrations but the sum in the exponent has been separated into two parts. One sum over μ includes the indices of the $p-s$ patterns which will not become extensive - they remain of order $O(N^{-1/2})$. The other sum is over the patterns $\nu(=1,\ldots,s)$. These are the candidates for condensation. The number of condensed patterns must remain finite as $N\to\infty$, hence, $p-s\approx p$.

Let us consider the first exponential in Eq.(two.exponential). We can calculate the averaging as

$$\begin{aligned}&\left\langle\!\!\left\langle \exp\left\{\beta N\left(\sum_{\mu\rho} m_\rho^\mu \frac{1}{N}\sum_{i=1}^N \xi_i^\mu S_i^\rho\right)\right\}\right\rangle\!\!\right\rangle_{\xi_i^\mu=\pm1} \\ &= \prod_{\mu i}\left\langle\!\!\left\langle \exp\left(\beta\xi_i^\mu\sum_\rho m_\rho^\mu S_i^\rho\right)\right\rangle\!\!\right\rangle_{\xi_i^\mu=\pm1} \\ &= \prod_{\mu i}\cosh\left(\beta\sum_\rho m_\rho^\mu S_i^\rho\right)\end{aligned}$$

$$\begin{aligned}&= \prod_{\mu i}\exp\left\{\log\cosh\left(\beta\sum_\rho m_\rho^\mu S_i^\rho\right)\right\} \\ &\approx \exp\left\{\frac{1}{2}\left(\beta\sum_\rho m_\rho^\mu S_i^\rho\right)^2\right\} \\ &= \exp\left(\frac{\beta^2}{2}\sum_{i\mu\rho\sigma} S_i^\rho S_i^\sigma m_\rho^\mu m_\sigma^\mu\right),\end{aligned} \tag{B.9}$$

where the approximation, $\log\cosh x \approx x^2/2$for small$x$, is used. The first exponential can therefore be written as

$$\begin{aligned}&\exp\left\{\beta N\left(-\frac{1}{2}\sum_{\mu\rho}(m_\rho^\mu)^2 + \frac{\beta}{2N}\sum_{i\mu\rho\sigma} S_i^\rho S_i^\sigma m_\rho^\mu m_\sigma^\mu\right)\right\} \\ &= \exp\left[-\frac{\beta N}{2}\left\{(1-\beta)\sum_{\mu\rho}(m_\rho^\mu)^2 - \beta\sum_{\mu\rho\sigma}\hat{q}_{\rho\sigma}\right\}\right] \\ &= \prod_{\rho>\sigma}\int dq_{\rho\sigma}\delta(q_{\rho\sigma}-\hat{q}_{\rho\sigma})\exp\left[-\frac{\beta N}{2}\left\{(1-\beta)\sum_{\mu\rho}(m_\rho^\mu)^2 - \beta\sum_{\mu\rho\sigma}q_{\rho\sigma}\right\}\right] \\ &= \prod_{\rho>\sigma}\int dq_{\rho\sigma}\int d(r_{\rho\sigma}\frac{\alpha\beta^2}{2})\exp\left\{-N\sum_{\rho\neq\sigma} r_{\rho\sigma}\frac{\alpha\beta^2}{2}(q_{\rho\sigma}-\hat{q}_{\rho\sigma})\right\} \\ &\quad\times\exp\left[-\frac{\beta N}{2}\left\{(1-\beta)\sum_{\mu\rho}(m_\rho^\mu)^2 - \beta\sum_{\mu\rho\sigma}q_{\rho\sigma}\right\}\right] \\ &= \left(\frac{\alpha\beta^2}{2}\right)^{n(n-1)}\prod_{\rho>\sigma}\int dq_{\rho\sigma}\int dr_{\rho\sigma}\exp\left\{-\frac{N}{2}\alpha\beta^2\sum_{\rho\neq\sigma} r_{\rho\sigma}(q_{\rho\sigma}-\hat{q}_{\rho\sigma})\right\} \\ &\quad\times\exp\left[-\frac{\beta N}{2}\left\{(1-\beta)\sum_{\mu\rho}(m_\rho^\mu)^2 - \beta\sum_{\mu\rho\sigma}q_{\rho\sigma}\right\}\right],\end{aligned} \tag{B.10}$$

where

$$\hat{q}_{\rho\sigma} \equiv \frac{1}{N}\sum_i S_i^\rho S_i^\sigma(1-\delta_{\rho\sigma}), \tag{B.11}$$

and we introduced Lagrange multipliers $r_{\rho\sigma}\alpha\beta^2/2$.

Now we obtain

$$\begin{aligned}\langle\langle Z^n\rangle\rangle &= \left(\frac{\alpha\beta^2}{2}\right)^{n(n-1)}(\beta N e^{-\beta})^{pn/2}\mathrm{Tr}_S^\rho\prod_{\mu\rho}\int\frac{dm_\rho^\mu}{\sqrt{2\pi}}\prod_{\nu\rho}\int\frac{dm_\rho^\nu}{\sqrt{2\pi}}\prod_{\rho>\sigma}\int dq_{\rho\sigma}dr_{\rho\sigma} \\ &\quad\times\exp\left[-\frac{\beta N}{2}\left\{(1-\beta)\sum_{\mu\rho}(m_\rho^\mu)^2 - \beta\sum_{\mu\rho\sigma}\hat{q}_{\rho\sigma}\right\}\right]\end{aligned}$$

$$\times \exp\left\{-\frac{N}{2}\alpha\beta^2 \sum_{\rho\neq\sigma} r_{\rho\sigma}(q_{\rho\sigma} - \frac{1}{N}\sum_i S_i^\rho S_i^\sigma)\right\}$$
$$\times \left\langle\!\!\left\langle \exp\left[\beta N\left\{-\frac{1}{2}\sum_{\nu\rho}(m_\rho^\nu)^2 + \sum_{\nu\rho}(m_\rho^\nu + h^\nu)\frac{1}{N}\sum_i \xi_i^\nu S_i^\rho\right\}\right]\right\rangle\!\!\right\rangle_{\xi_i^\nu} . \quad \text{(B.12)}$$

In all integrals except for that over m_ρ^μ, we take the saddle point value which identifies them as the order parameters. Saddle points are where the respective partial derivatives of the exponent vanish:

1. For the retrieved (condensed) patterns $\nu \leq s$, $\frac{\partial}{\partial m_\rho^\nu}$ gives
$$m_\rho^\nu = \left\langle\!\!\left\langle \frac{1}{N}\sum_i \xi_i^\nu \langle S_i^\rho\rangle \right\rangle\!\!\right\rangle, \quad \nu = 1,\ldots,s, \quad \text{(B.13)}$$
which is just the retrieval quality order parameter on pattern ν for the quenched thermal equilibrium of replica ρ.

2. From $\frac{\partial}{\partial r_{\rho\sigma}}$, one obtains the Edwards-Anderson order parameter,
$$q_{\rho\sigma} = \left\langle\!\!\left\langle \frac{1}{N}\sum_i \langle S_i^\rho\rangle\langle S_i^\sigma\rangle \right\rangle\!\!\right\rangle, \quad \text{(B.14)}$$
which, in the replica picture, characterizes the correlations caused by the equality of quenched patterns between two thermally independent replicas.

3. $\frac{\partial}{\partial q_{\rho\sigma}}$ gives
$$r_{\rho\sigma} = \frac{1}{\alpha}\sum_{\mu=s+1}^{p} \langle\!\langle m_\rho^\mu m_\sigma^\nu \rangle\!\rangle \quad \text{(B.15)}$$
as the order parameter of random overlaps with the non-retrieved patterns.

We calculate the integrals over m_ρ^μ. Because of the quadratic form of m_ρ^μ, the integrals are just multi-dimensional Gaussian by transforming to principal axes. The matrix of coefficients of $(m_\rho^\mu m_\sigma^\mu)$ is
$$(1-\beta)\mathcal{I} - \beta\mathcal{Q}, \quad \text{(B.16)}$$

where $\mathcal{I}$ is $n \times n$ unit matrix and $\mathcal{Q}$ is $n \times n$ matrix in which diagonal elements are 0 and off-diagonal $\rho\sigma$ elements are $q_{\rho\sigma}$. Assuming that the eigen-values of the matrix (B.16) are λ_ρ, we can calculate ¿the integrals over m_ρ^μ and obtain

$$\begin{aligned}
&\prod_{\mu\rho}\int \frac{dm_\rho^\mu}{\sqrt{2\pi}} \exp\left[-\frac{\beta N}{2}\left\{(1-\beta)\sum_{\mu\rho}(m_\rho^\mu)^2 - \beta\sum_{\mu\rho\neq\sigma} m_\rho^\mu m_\sigma^\mu q_{\rho\sigma}\right\}\right] \\
&= (\beta N)^{-pn/2}\prod_\rho \lambda_\rho^{-p/2} \\
&= (\beta N)^{-pn/2}\exp\left[-\frac{p}{2}\mathrm{Tr}\ln\{(1-\beta)\mathcal{I} - \beta\mathcal{Q}\}\right]. \qquad \text{(B.17)}
\end{aligned}$$

Thus we obtain the expression of $\langle\!\langle Z^n\rangle\!\rangle$ as

$$\begin{aligned}
\langle\!\langle Z^n\rangle\!\rangle = &\left(\frac{\alpha\beta^2}{2}\right)^{n(n-1)} e^{-\beta pn/2}\exp\left[-\frac{p}{2}\mathrm{Tr}\ln\{(1-\beta)\mathcal{I} - \beta\mathcal{Q}\}\right. \\
&\left. -\frac{N}{2}\alpha\beta^2\sum_{\rho\neq\sigma} r_{\rho\sigma}q_{\rho\sigma} - \frac{\beta N}{2}\sum_{\nu\rho}(m_\rho^\nu)^2\right] \\
&\times\left\langle\!\!\left\langle \mathrm{Tr}_{S^\rho}\exp\left[\frac{\alpha\beta^2}{2}\sum_{\rho\neq\sigma} r_{\rho\sigma}\sum_i S_i^\rho S_i^\sigma\right.\right.\right. \\
&\left.\left.\left. +\beta\sum_{\nu\rho}N(m_\rho^\nu + h^\nu)\frac{1}{N}\sum_i \xi_i^\nu S_i^\rho\right]\right\rangle\!\!\right\rangle_{\xi_i^\nu} . \qquad \text{(B.18)}
\end{aligned}$$

The average $\langle\!\langle\cdots\rangle\!\rangle$ in Eq.(B.18) is further calculated to be

$$\begin{aligned}
&\left\langle\!\!\left\langle \mathrm{Tr}_{S^\rho}\exp\left[\frac{\alpha\beta^2}{2}\sum_{i\rho\neq\sigma} r_{\rho\sigma}S_i^\rho S_i^\sigma + \beta\sum_{i\nu\rho}(m_\rho^\nu + h^\nu)\xi_i^\nu S_i^\rho\right]\right\rangle\!\!\right\rangle_{\xi_i^\nu} \\
&= \left\langle\!\!\left\langle \prod_i \mathrm{Tr}_{S^\rho}\exp\left[\frac{\alpha\beta^2}{2}\sum_{\rho\neq\sigma} r_{\rho\sigma}S^\rho S^\sigma + \beta\sum_{\nu\rho}(m_\rho^\nu + h^\nu)\xi_i^\nu S^\rho\right]\right\rangle\!\!\right\rangle_{\xi_i^\nu} \\
&= \left\langle\!\!\left\langle \exp N\left\{\frac{1}{N}\sum_i \ln \mathrm{Tr}_{S^\rho}\exp\left[\frac{\alpha\beta^2}{2}\sum_{\rho\neq\sigma} r_{\rho\sigma}S^\rho S^\sigma + \beta\sum_{\nu\rho}(m_\rho^\nu + h^\nu)\xi_i^\nu S^\rho\right]\right\}\right\rangle\!\!\right\rangle_{\xi_i^\nu} . \\
&\qquad \text{(B.19)}
\end{aligned}$$

The above transformations should be self-explanatory. Next we apply self-averaging to the sum in the exponent. The result is an average over the ξ's in the exponent. This is independent of the ξ's, and the external average, represented by the double brackets, becomes spurious. As a consequence, the double brackets move into the exponent. The next line in the above chain of equations is

$$\exp N \left\langle\!\!\left\langle \ln mboxTr_{S^\rho} \exp\left[\frac{\alpha\beta^2}{2}\sum_{\rho\neq\sigma} r_{\rho\sigma} S^\rho S^\sigma + \beta \sum_{\nu\rho}(m^\nu_\rho + h^\nu)\xi^\nu S^\rho\right]\right\rangle\!\!\right\rangle_{\xi^\nu_i} . \tag{B.20}$$

Now we obtain the final expression for the averaged partition function as

$$\begin{aligned}\langle\langle Z^n\rangle\rangle &= \left(\frac{\alpha\beta^2}{2}\right)^{n(n-1)} e^{-\beta pn/2}\exp\left[-\frac{p}{2}\mathrm{Tr}\ln\{(1-\beta)\mathcal{I}-\beta\mathcal{Q}\}\right. \\ &\quad \left. -\frac{N}{2}\alpha\beta^2\sum_{\rho\neq\sigma} r_{\rho\sigma}q_{\rho\sigma} - \frac{\beta N}{2}\sum_{\nu\rho}(m^\nu_\rho)^2\right] + \exp N\left\langle\!\!\left\langle \ln \mathrm{Tr}_{S^\rho} e^{\beta\mathcal{H}_\xi}\right\rangle\!\!\right\rangle_{\xi^\nu_i}, \end{aligned} \tag{B.21}$$

where

$$\mathcal{H}_\xi \equiv \frac{\alpha\beta}{2}\sum_{\rho\neq\sigma} r_{\rho\sigma} S^\rho S^\sigma + \sum_{\nu\rho}(m^\nu_\rho + h^\nu)\xi^\nu S^\rho \tag{B.22}$$

is a free-energy of a system in the replica space and the total number of variables is n.

The free-energy, Eq.(B.5), can be evaluated directly as

$$\begin{aligned} f &= \lim_{n\to 0}\left[\frac{1}{2}\alpha + \frac{1}{2n}\sum_{\nu\rho}(m^\nu_\rho)^2 + \frac{\alpha}{2\beta n}\mathrm{Tr}\ln[(1-\beta)\mathcal{I} - \beta\mathcal{Q}]\right. \\ &\quad \left. + \frac{\alpha\beta}{2n}\sum_{\rho\neq\sigma} r_{\rho\sigma}q_{\rho\sigma} - \frac{1}{\beta n}\left\langle\!\!\left\langle \ln\mathrm{Tr}_{S^\rho} e^{\beta\mathcal{H}_\xi}\right\rangle\!\!\right\rangle_{\xi^\nu_i}\right]. \end{aligned} \tag{B.23}$$

That the limit $n \to 0$ may exist can be made plausible by observing that:

- In the second term the summation over ρ could make the sum of order n;
- the third term, the Tr, is a trace of a matrix of order $n \times n$, which has n elements in its diagonal, and could be of order n;
- the fourth term has $n(n-1)$ terms in the sum, which in the limit could be proportional to $-n$;
- the last term is itself like a free-energy of a system with an energy $\mathcal{H}_\xi$. The free-energy should be 'extensive' and hence proportional to n.

B.2.1 Replica symmetry

To make further progress, we must make some simplifying assumptions about the nature of the matrices in replica space. the simplest such assumption is replica symmetry. It i defined to read

$$\begin{aligned} m^\nu_\rho &= m^\nu \\ q_{\rho\sigma} &= q \qquad \rho\neq\sigma \\ r_{\rho\sigma} &= r \qquad \rho\neq\sigma. \end{aligned} \tag{B.24}$$

With these matrices, the computation of the free-energy can be completed.

- The matrix $(1-\beta)\mathcal{I}-\beta\mathcal{Q}$ has one eigen-value $1-\beta-(n-1)\beta q$ and $n-1$ eigen-values $1-\beta(1-q)$. Hence,

$$\begin{aligned} &\lim_{n\to 0}\frac{1}{n}\mathrm{Tr}\ln[(1-\beta)\mathcal{I}-\beta\mathcal{Q}] \\ &= \lim_{n\to 0}\frac{1}{n}[(n-1)\ln[1-\beta(1-q)] + \ln[1-\beta-(n01)\beta q]] \\ &= \ln[1-\beta(1-q)] - \frac{\beta q}{1-\beta(1-q)}. \end{aligned} \tag{B.25}$$

- With Eqs.(B.25),

$$\lim_{n\to 0}\frac{1}{n}\sum_{\rho\neq\sigma} r_{\rho\sigma}q_{\rho\sigma} = -rq. \tag{B.26}$$

$$\lim_{n\to 0}\frac{1}{2n}\sum_{\nu\rho}(m^\nu_\rho)^2 = \frac{1}{2}\sum_\nu (m^\nu)^2 \tag{B.27}$$

- The 'free-energy' in the replica universe is,

$$\begin{aligned} &\frac{1}{n\beta}\langle\langle \ln\mathrm{Tr}_{S^\rho}\exp(\beta\mathcal{H}_\xi\rangle\rangle_{\xi^\nu} \\ &= \frac{1}{n\beta}\left\langle\!\!\left\langle \ln\mathrm{Tr}_{S^\rho}\exp\left[\frac{\alpha\beta^2}{2}r\left\{\left(\sum_\rho S^\rho\right)^2 - n\right\} + \beta\sum_{\nu\rho}(m^\nu+h^\nu)\xi^\nu S^\rho\right]\right\rangle\!\!\right\rangle_{\xi^\nu} \\ &= \frac{1}{n\beta}\left\langle\!\!\left\langle -\frac{\alpha\beta^2 rn}{2} + \ln\mathrm{Tr}_{S^\rho}\frac{1}{\sqrt{2\pi\alpha\beta^2 r}}\int_{-\infty}^{\infty} dx \right.\right. \end{aligned}$$

$$
\begin{aligned}
& \exp\left\{-\frac{x^2}{2\alpha\beta^2 r}+\sum_\rho S^\rho x+\beta\sum_{\nu\rho}(m^\nu+h^\nu)\xi^\nu S^\rho\right\}\Big\rangle\!\Big\rangle_{\xi^\nu} \\
= & -\frac{\alpha\beta r}{2}+\frac{1}{n\beta}\Big\langle\!\Big\langle \ln\int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi}}e^{-z^2/2}\mathrm{Tr}_{S^\rho}\exp\left\{\sum_\rho S^\rho\sqrt{\alpha\beta r}z+\beta\sum_{\nu\rho}(m^\nu+h^\nu)\xi^\nu S^\rho\right\}\Big\rangle\!\Big\rangle_{\xi^\nu} \\
= & -\frac{\alpha\beta r}{2}+\frac{1}{n\beta}\Big\langle\!\Big\langle \ln\int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi}}e^{-z^2/2}\left[2\cosh\left\{\sqrt{\alpha\beta^2 r}z+\beta\sum_\nu(m^\nu+h^\nu)\xi^\nu\right\}\right]^n\Big\rangle\!\Big\rangle_{\xi^\nu} \\
\xrightarrow{n\to 0} & -\frac{\alpha\beta r}{2}+\frac{1}{\beta}\Big\langle\!\Big\langle \int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi}}e^{-z^2/2}\ln\left[2\cosh\left\{\beta\sqrt{\alpha r}z+\beta\sum_\nu(m^\nu+h^\nu)\xi^\nu\right\}\right]\Big\rangle\!\Big\rangle_{\xi^\nu}.
\end{aligned} \tag{B.28}
$$

The result of the computation of the free-energy is

$$
\begin{aligned}
f = & \frac{\alpha}{2}+\frac{\alpha}{2\beta}\left(\ln[1-\beta(1-q)]-\frac{\beta q}{1-\beta(1-q)}\right)+\frac{\alpha\beta r}{2}(1-q)+\frac{1}{2}\sum_\nu(m^\nu)^2 \\
& -\frac{1}{\beta}\int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi}}e^{-z^2/2}\Big\langle\!\Big\langle \ln\left[2\cosh\beta\left\{\sqrt{\alpha r}z+\sum_\nu(m^\nu+h^\nu)\xi^\nu\right\}\right]\Big\rangle\!\Big\rangle_{\xi^\nu}.
\end{aligned} \tag{B.29}
$$

From equations

$$
\frac{\partial f}{\partial m^\nu}=0,\qquad \frac{\partial f}{\partial q}=0,\qquad \frac{\partial f}{\partial r}=0, \tag{B.30}
$$

we obtain mean-field equations

$$
m^\nu = \langle\!\langle \xi^\nu\tanh\beta[\sqrt{\alpha r}z+(\boldsymbol{m}+\boldsymbol{h})\cdot\boldsymbol{\xi}]\rangle\!\rangle \tag{B.31}
$$

$$
q = \langle\!\langle \tanh^2\beta[\sqrt{\alpha r}z+(\boldsymbol{m}+\boldsymbol{h})\cdot\boldsymbol{\xi}]\rangle\!\rangle \tag{B.32}
$$

$$
r = \frac{q}{[1-\beta(1-q)]^2}. \tag{B.33}
$$

The vectors $\boldsymbol{m}$,$\boldsymbol{h}$ and $\boldsymbol{\xi}$ have s components each, corresponding to the patterns that have condensed. The double angular brackets, the average over the random patterns, is composed of two parts. The first is an average over the distribution of discrete ξ^ν's. The second is an average over a continuous Gaussian distribution of the noise generated by the 'other' patterns. The combined average can be written as:

$$
\langle\!\langle O\rangle\!\rangle \equiv \int_{-\infty}^{\infty}dz\frac{e^{-z^2/2}}{\sqrt{2\pi}}\langle\!\langle O\{\xi\}\rangle\!\rangle_{\xi^\nu} \tag{B.34}
$$

Amit et.al. focused attention on attractors which retrieve at most one pattern and they derived the various phase diagrams - $(\alpha - T)$or$(h-\alpha)$ etc. - by numerical solution of the mean-field equations. They reduce to:

$$
m = \int\frac{dz}{\sqrt{2\pi}}e^{-z^2/2}\tanh[\beta(\sqrt{\alpha r}z+m+h)] \tag{B.35}
$$

$$
q = \int\frac{dz}{\sqrt{2\pi}}e^{-z^2/2}\tanh^2[\beta(\sqrt{\alpha r}z+m+h)] \tag{B.36}
$$

$$
r = \frac{q}{[1-\beta(1-q)]^2}. \tag{B.37}
$$

List of publication

1. T. Munakata and Y. Nakamura, Phys. Rev. E**47** (1993) 3791.

   'Projection-operator approach to overlap dynamics in a Hopfield network'

2. Y. Nakamura and T. Munakata, Phys. Rev. E**49** (1994) 1775.

   'Correlated-data-driven dynamics in a neural network'

3. Y. Nakamura, F. Tominaga and T. Munakata, Phys. Rev. E**49** (1994) 4849.

   'Clustering behavior of time-delayed nearest-neighbor coupled oscillators'

4. Y. Nakamura, K. Torii and T. Munakata, Phys. Rev. E (in press).

   'A neural network model composed of multi-dimensional spin neurons'