Retrieval Properties of Hopfield and Correlated attoractors in an Associative Memory Model

Tatsuya Uezu, Aya Hirano and Masato Okada[1,2]

Graduate School of Humanities and Sciences, Nara Women's University, Nara 630-8506, Japan
[1] *Laboratory for Mathematical Neuroscience,*
RIKEN Brain Science Institute, Saitama 351-0198, Japan
[2] *"Intelligent Cooperation and Control", PRESTO, JST,*
c/o RIKEN BSI, Saitama 351-0198, Japan

(Received February 6, 2004)

We examine a previouly introduced attractor neural network model that explains the persistent activities of neurons in the anterior ventral temporal cortex of the brain. In this model, the coexistence of several attractors including correlated attractors was reported in the cases of finite and infinite loading. In this paper, by means of a statistical mechanical method, we study the statics and dynamics of the model in both finite and extensive loading, mainly focusing on the retrieval properties of the Hopfield and correlated attractors. In the extensive loading case, we derive the evolution equations by the dynamical replica theory. We found several characteristic temporal behaviours, both in the finite and extensive loading cases. The theoretical results were confirmed by numerical simulations.

KEYWORDS: attractor neural network model, hopfield attractor, correlated attractor, dynamical replica theory

§1. Introducion

The attractor neural network model has been considered a plausible model for memory processing in the brain. Many experimental findings implying its existence have been reported. Among them, Miyashita and his group reported the persistent activities of neurons in the anterior ventral temporal (AVT) cortex.[1,2] They trained monkeys to recognize and match a set of 100 colored fractal patterns in the Delayed Match to Sample (DMS) paradigm. Miyashita and Chang reported that the neurons in the AVT are highly selective toward a few of the 100 colored fractal patterns, and that the activities of these neurons persist for 16 seconds after the removal of the presented stimulus.[2] Miyashita studied development of this selectivity, and found that serial positions number of stimuli during learning are converted into the spatial correlation of the neural activities.[1]

To explain these experiments, Griniasty et al. proposed a model for an attractor neural network.[3] By analysing equilibrium states of this model in the finite loading case, they found not only stored patterns (Hopfield attractors) but also patterns which have overlap with stored patterns (correlated attractors). For example, when the number of patterns is 13, correlated attractors are expressed by $\boldsymbol{m} = \frac{1}{2^7}{}^t(77, 51, 13, 3, 1, 0, 0, 0, 0, 1, 3, 13, 51)$ and its cyclic rotations, whereas Hopfield attractors are expressed by $\boldsymbol{m} =^t (1, 0, \cdots, 0)$ and its cyclic rotations. Here, $\boldsymbol{m}$ is the overlap vector $\boldsymbol{m} =^t (m_1, m_2, \cdots, m_c)$, of which component m_μ is the overlap between the state of neurons and the μ-th pattern. t denotes transposition. See eq. (6). Griniasty et al. found that the correlation between correlated attractors dicreases as the separation in the sequence of the patterns to which the attractors belong increases. These results support the work done by Miyashita.

In recent years, we have studied the statics and dynamics of the model introduced by Griniasty et al. in the finite loading case. As for the statics, we investigated equilibrium states and found the coexistence of several attractors such as Hopfield attractors, correlated attractors, and mixed states of several patterns in several parameter regions. The coexistence of attractors was also reported by Shiino et al.[4,5] and by Fukai et al.[6] As for the dynamics, we investigated the transient behaviour by performing numerical calculations of the overlap dynamics derived theoretically and by numerical simulations. In particular, we studied what happens when a Hopfield attractor and a correlated attractor coexist and when a Hopfield attractor does not exist and a correlated attractor does exist. In the former case, we studied the basin of attraction and in the latter case, we found that the trajectory initially approaches the state where the Hopfield attractor existed, but finally it tends toward the correlated attractor. We found that the results by numerical calculations of the overlap dynamics and the results by numerical simulations agree qualitatively.

The coexistence of multiple attractors and the dynamical behaviour obtained in the finite loading case are significant. However, the situation of the finite loading seems to be unrealistic. Thus, it is necessary to investigate whether the results in the finite loading case are obtained when applied practically, i.e. in the extensive loading case.

In this paper, we study the model introduced by Griniasty et al. by the statistical mechanical method, mainly focusing on the extensive loading case. In the extensive loading case, as to the study of the equilibrium states, Cugliandolo et al. derived the saddle point equations(S.P.E.) for order parameters using the replica method.[7] They identified not only the Hopfield attractors but also the correlated attractors. We have solved the saddle point equations numerically by scanning parameters and found that there are parameter regions where several attractors coexist. The coexistence of multiple attractors was found by Shiino et al. in the present model as well.[4,5] As for the dynamical behaviour, we have derived evolution equations for order parameters by using the dynamical replica theory by Laughton and Coolen.[8] As in the finite loading case, we have studied the dynamical behaviour in a coexistence region of a Hopfield attractor and a correlated attractor, and also in a region where a Hopfield attractor does not exist but a correlated attractor does exist. From these studies, we have obtained qualitatively similar results to those in the finite loading case.

In the next section, we explain the present model. Then, we summarize the theory and results for the finite loading case in §3. In §4, the extensive loading case is examined. The summary and discussion are given in §5.

§2. Model

Let us explain the model. The instantaneous state of each neuron is expressed by s_i which takes ±1, where i labels the neuron($i = 1, \cdots, N$), and the time evolution is given by

$$s_i(t+1) = \mathrm{sign}(h_i(t)), \tag{1}$$

where

$$h_i(t) = \sum_{j(\neq i)} J_{ij} s_j(t), \tag{2}$$

and J_{ij} is the strength of the synaptic connection from the j-th neuron to the i-th neuron. The stochastic dynamics is also taken into consideration by introducing temperature T, that is, the probability that $s_i(t+1)$ takes ±1 is given by

$$\mathrm{Prob}[s_i(t+1)] = \frac{1 \pm \tanh(\beta h_i(t))}{2}, \tag{3}$$

where $\beta = 1/T$. In the present model, the synaptic weight J_{ij} is defined as

$$\begin{aligned} J_{ij} \quad &= \quad \frac{1}{N} \sum_{\mu=1}^{c} (\xi_i^\mu \xi_j^\mu + a\xi_i^\mu \xi_j^{\mu-1} + a\xi_i^\mu \xi_j^{\mu+1}) = \frac{1}{N} \sum_{\mu,\nu} \xi_i^\mu A_{\mu\nu} \xi_j^\nu \quad \text{for } i \neq j, \\ & \quad J_{ii} = 0, \;\; \xi_i^0 \equiv \xi_i^c, \;\; \xi_i^{c+1} \equiv \xi_i^1, \end{aligned} \tag{4}$$

where ξ_i^μ represents the value of i-th neuron for μ-th pattern $\boldsymbol{\xi}^\mu \equiv^t (\xi_1^\mu, \cdots, \xi_N^\mu)$ and it takes value +1 or -1 with a probability of 1/2. c is the total number of patterns. $\boldsymbol{A}$ is a $c \times c$ matrix defined as

$$\boldsymbol{A} \quad \equiv \quad \{A_{\mu\nu}\} = \begin{pmatrix} 1 & a & 0 & \cdots & 0 & a \\ a & 1 & a & 0 & \cdots & 0 \\ 0 & a & 1 & a & 0 & 0 \\ 0 & 0 & a & 1 & a & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ a & 0 & \cdots & 0 & a & 1 \end{pmatrix}. \tag{5}$$

The order parameter is the overlap vector $\boldsymbol{m} =^t (m_1, m_2, \cdots, m_c)$, of which component m_μ is the overlap between the state of neurons $\boldsymbol{s} =^t (s_1, \cdots, s_N)$ and the μ-th pattern $\boldsymbol{\xi}^\mu$, that is,

$$m_\mu = \frac{1}{N} \sum_{i=1}^{N} s_i \xi_i^\mu. \tag{6}$$

§3. Finite Loading Case

In this section, we study the finite loading case of the model described by eqs.(1)-(4). That is, we consider what happens when $N \gg 1$ and $\alpha = \frac{c}{N} \ll 1$.

3.1 Theory

To study the equilibrium state, we consider the following Hamiltonian H.

$$H = -\frac{1}{2}\sum_{i\neq j} J_{ij} s_i s_j \tag{7}$$

$$= -\frac{1}{2N}\sum_{i\neq j}\sum_{\mu,\nu} \xi_i^{\mu} A_{\mu\nu} \xi_j^{\nu} s_i s_j. \tag{8}$$

We calculate the free energy of the system

$$f = -\frac{1}{\beta N}\ln Z, \quad Z = \mathrm{Tr}_{\boldsymbol{s}} e^{-\beta H},$$

where $\mathrm{Tr}_{\boldsymbol{s}}$ denotes the summation with respect to $s_i (i = 1, \cdots, N)$. By using the saddle point method, we obtain[9)]

$$f = -\frac{\beta}{2}\sum_{\mu,\nu} m_\mu A_{\mu\nu} m_\nu - i\sum_{\mu=1}^{c} m_\mu \hat{m}_\mu - [\ln\{2\cosh(-i\sum_{\mu=1}^{c}\hat{m}_\mu \xi^\mu)\}]_{\boldsymbol{\xi}}, \tag{9}$$

where $\boldsymbol{\xi} = {}^t(\xi^1, \xi^2, \cdots, \xi^c)$ and $[A]_{\boldsymbol{\xi}}$ is the average over $\boldsymbol{\xi}$, that is $[A]_{\boldsymbol{\xi}} = \frac{1}{2^c}\sum_{\{\xi_\mu=\pm 1\}}$A. Then, the S.P.E. becomes

$$\hat{m}_\mu = i\beta\sum_{\nu=1}^{c} A_{\mu\nu} m_\nu, \tag{10}$$

$$m_\mu = [\xi^\mu \tanh(\beta \sum_{\nu,\nu'=1}^{c} \xi^\nu A_{\nu\nu'} m_{\nu'})]_{\boldsymbol{\xi}}. \tag{11}$$

The physical meaning of m_μ is the overlap between the equilibrium state and the μ -th pattern,

$$m_\mu = \frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu \langle s_i\rangle, \tag{12}$$

where $\langle s_i\rangle$ is the thermal average of the i-th neuron.

Next, we derive the evolution equation of the system. Let us consider the probability $\mathcal{P}_t(\boldsymbol{m})$ that at time t the state $\boldsymbol{s}$ has the overlap m_μ with μ-th pattern $\boldsymbol{\xi}^\mu$ for $\mu = 1, \cdots, c$

$$\mathcal{P}_t(\boldsymbol{m}) = \mathrm{Tr}_{\boldsymbol{s}} p_t(\boldsymbol{s}) \prod_{\mu=1}^{c} \delta(m_\mu - m_\mu(\boldsymbol{s})), \tag{13}$$

$$m_\mu(\boldsymbol{s}) = \frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu s_i, \tag{14}$$

where $p_t(\boldsymbol{s})$ is the probability that the system takes the state $\boldsymbol{s}$ at time t. We assume that the transition probability $w_i(\boldsymbol{s})$ from the state $\boldsymbol{s} = (s_1, \cdots, s_i, \cdots, s_N)$ to the state $F_i\boldsymbol{s} = (s_1, \cdots, -s_i, \cdots, s_N)$ takes the following form

$$w_i(\boldsymbol{s}) = \frac{1 - s_i\tanh\{\beta h_i(\boldsymbol{s})\}}{2}, \tag{15}$$

where F_i is the flip operator of the i-th neuron and $h_i(\boldsymbol{s})$ is expressed as

$$h_i(\boldsymbol{s}) = \sum_{j(\neq i)}^{N} J_{ij} s_j \simeq \sum_{\mu,\nu \leq c} \xi_i^{\mu} A_{\mu\nu} m_{\nu}(\boldsymbol{s}). \tag{16}$$

The master equation for $p_t(\boldsymbol{s})$ is given by

$$\frac{d}{dt} p_t(\boldsymbol{s}) = \sum_{i=1}^{N} \{ w_i(F_i \boldsymbol{s}) p_t(F_i \boldsymbol{s}) - w_i(\boldsymbol{s}) p_t(\boldsymbol{s}) \}. \tag{17}$$

Then, using asynchronous dynamics, we obtain the evolution equations for the overlap $\boldsymbol{m}$ as

$$\frac{d}{dt} \boldsymbol{m} = -\boldsymbol{m} + [\boldsymbol{\xi} \tanh(\beta^t \boldsymbol{\xi} \boldsymbol{A} \boldsymbol{m})]_{\boldsymbol{\xi}}. \tag{18}$$

The equation for the stationary state of eq. (18) agrees to the eq. (11) for the equilibrium state.

3.2 Results

c is finite and we put $c = 13$ as in ref.,[7] becasuse the correlated attractors obtained by Amit et al. have non-zero values up to their five nearest neighbours.

The equilibrium state

We solved the S.P.E.(11) numerically and found the coexistence of several attractors such as the Hopfield attractors, correlated attractors, and mixed states of several patterns in several parameter regions. For example, we show the result for $a = 0.4$ in fig. 1. The solid curve denotes the Hopfield attractor. It exists up to $T \simeq 0.1$. The dashed curves denote a correlated attractor. For this attractor, there is the following symmetry

$$m_2 = m_{13}, \quad m_3 = m_{12}, \cdots, m_7 = m_8.$$

This attractor exists up to $T \simeq 0.25$. It has been considered that correlated attractors exist only for $a > 0.5$.[7] However, we found that they exist even for $a < 0.5$. As is seen from the figure, up to $T \simeq 0.1$, both the Hopfield attractor and the correlated attractor exist. Further, we found that the mixed states of three patterns $\boldsymbol{\xi}^{\mu}$, $\boldsymbol{\xi}^{\mu+1}$ and $\boldsymbol{\xi}^{\mu+2}$ (which are not drawn in fig. 1) exist up to $T \simeq 0.05$ and the mixed state of 13 patterns exists up to $T \simeq 1.7$.

Dynamics

First, we investigate the case where several attractors coexist. For example, for $a = 0.4$ and $T = 0.04$, a Hopfield attractor, a correlated attractor, mixed states of three patterns and a mixed state of 13 patterns coexist. For this parameter, we investigate the basin of attraction of these attractors by using overlap dynamics (eq. (18)) and numerical simulations. In the numerical integrations of eq. (18), we set the initial condition as

$$m_{\mu}(t = 0) = m_0 \delta_{\mu,1}, \quad \mu = 1, \cdots, c. \tag{19}$$

On the other hand, in the numerical simulations, we generate an initial state $\boldsymbol{s}(0)$ according to the following probability,

$$\mathrm{Prob}[s_i = \pm 1] = \frac{1 \pm m_0 \xi_i^1}{2}. \tag{20}$$

When $N \to \infty$, the overlap m_μ between $\boldsymbol{s}(0)$ and $\boldsymbol{\xi}^\mu$ satisfies the relation (19).
In fig. 2, we show the results by numerical integrations and those by numerical simulations. In fig. 3, we compare the results by numerical integrations and those by numerical simulations in more detail. As is shown in fig. 3(a), in the numerical integrations, when the initial overlap with the pattern 1, m_0, is 0.15, the trajectory tends toward the correlated attractor and the trajectory with $m_0 = 0.16$ tends toward the Hopfield attractor. That is, the boundary between the basin of attraction for the Hopfield attractor and that for the correlated attractor, m_0^c, is between 0.15 and 0.16. We show the results of the simulations in fig. 3. When m_0 is 0.16, the trajectory tends toward the correlated attractor and when m_0 is 0.17, the trajectory tends toward the Hopfield attractor. From these results, we note that even if the trajectories finally tend toward the correlated attractor, they initially approach the Hopfield attractor.

Next, we investigate the case where a Hopfield attractor does not exist but a correlated attractor exists, which is shown in fig. 4. As the figures show, both in the numerical integrations of eq. (18) and numerical simulations, we found that for $m_0 < 1$, the trajectory initially approaches the state where the Hopfield attractor existed, but finally it tends toward the correlated attractor.

In the next section, we study the extensive loading case to see whether the characteristic feature in the finite loading case such as the coexistence of multiple attractors and characteristic temporal behaviour continues to exist or not.

§4. Extensive Loading Case

4.1 Theroy

As an example of an extensive loading case, we consider the following synaptic weight J_{ij},

$$\begin{aligned} J_{ij} &= \frac{1}{N}\{\sum_{\mu=1}^{c}(\xi_i^\mu \xi_j^\mu + a\xi_i^\mu \xi_j^{\mu-1} + a\xi_i^\mu \xi_j^{\mu+1}) + \sum_{\mu=c+1}^{p} \eta_i^\mu \eta_j^\mu\} \quad \text{for } i \neq j, \\ & J_{ii} = 0, \quad \xi_i^0 \equiv \xi_i^c, \quad \xi_i^{c+1} \equiv \xi_i^1. \end{aligned}$$

c is the number of condensed patterns and we set $c = 13$. p is the total number of patterns and we consider the case that $\alpha = \frac{p}{N}$ is finite. We assume that ξ_i^μ and η_i^μ take +1 or -1 with a probability of 1/2. The overlap $m_\mu(\boldsymbol{s})$ between the state of neurons $\boldsymbol{s}$ and the μ-th pattern $\boldsymbol{\xi}^\mu$ is defined as

$$\begin{aligned} m_\mu(\boldsymbol{s}) &= \frac{1}{N}\sum_{i=1}^{N} s_i \xi_i^\mu \quad \text{for } \mu = 1, \cdots, c \\ &= \frac{1}{N}\sum_{i=1}^{N} s_i \eta_i^\mu \quad \text{for } \mu = c+1, \cdots, p. \end{aligned}$$

The cross-talk noise $z_i(\boldsymbol{s})$ is defined as

$$z_i(\boldsymbol{s}) \equiv \sum\nolimits_{\mu>c} \eta_i^\mu m_\mu(\boldsymbol{s}).$$

As order parameters, we define $\boldsymbol{m} =^t (m_1, m_2, \cdots, m_c)$ and $r = \frac{1}{\alpha}\sum_{\mu=c+1}^{p}(m_\mu(\boldsymbol{s}))^2$ where αr is the variance of the cross-talk noise.

The equilibrium states

By using the replica method, the free energy f for the replica symmetric(RS) solution[7] is given as

$$\begin{aligned} f &= \frac{1}{2}{}^t\boldsymbol{mAm} + \frac{\alpha}{2\beta}\left(\ln(1-\beta+\beta q) - \frac{\beta q}{1-\beta+\beta q}\right) + \frac{\alpha\beta}{2}r(1-q) \\ &\quad -T[\int Dz \ln\{2\cosh\beta(\sqrt{\alpha r}z +^t \boldsymbol{\zeta Am})\}]_{\boldsymbol{\zeta}}, \end{aligned} \tag{21}$$

where $\boldsymbol{\zeta} =^t (\zeta^1, \cdots, \zeta^c)$ and $[A]_{\boldsymbol{\zeta}}$ is the average over $\boldsymbol{\zeta}$, that is , $\frac{1}{2^c}\Sigma_{\{\zeta^\mu=\pm1\}}A$. The S.P.E. for $\boldsymbol{m}$, q and r are given by

$$\boldsymbol{m} = [\boldsymbol{\zeta}\int Dz \tanh\beta({}^t\boldsymbol{\zeta Am} + \sqrt{\alpha r}z)]_{\boldsymbol{\zeta}}, \tag{22}$$

$$q = [\int Dz \tanh^2\beta({}^t\boldsymbol{\zeta Am} + \sqrt{\alpha r}z)]_{\boldsymbol{\zeta}}, \tag{23}$$

$$r = \frac{q}{(1-c)^2}, \tag{24}$$

where $c \equiv \beta(1-q)$ and $\boldsymbol{A}$ is the $c\times c$ matrix which is defined in eq. (5). For $T = 0$, these equations become

$$c = \sqrt{\frac{2}{\pi\alpha r}}[\exp\{-(\frac{{}^t\boldsymbol{\zeta Am}}{\sqrt{2\alpha r}})^2\}]_{\boldsymbol{\zeta}}, \tag{25}$$

$$\boldsymbol{m} = -2[\boldsymbol{\zeta}H(\frac{{}^t\boldsymbol{\zeta Am}}{\sqrt{\alpha r}})]_{\boldsymbol{\zeta}}, \tag{26}$$

$$r = \frac{1}{(1-c)^2}, \tag{27}$$

where $H(x) = \int_x^\infty \frac{dt}{\sqrt{2\pi}}e^{-t^2/2}$.

Dynamics

As for the dynamics, we derive the evolution equations for $(\boldsymbol{m}, r)$ using the dynamical replica theory(DRT).[8] We assume that the same transition probability $w_i(\boldsymbol{s})$ as in the finite loading case, but in the present case, $h_i(\boldsymbol{s})$ is expressed as

$$h_i(\boldsymbol{s}) = \sum_{\mu,\nu\le c} \xi_i^\mu A_{\mu\nu} m_\nu(\boldsymbol{s}) + z_i(\boldsymbol{s}). \tag{28}$$

The master equation for $p_t(\boldsymbol{s})$ is the same as before. Using asynchronous dynamics, we obtain the evolution equations for the overlap $\boldsymbol{m}$ and the variance of the cross-talk noise $\alpha r(\boldsymbol{s})$, assuming the self-averaging in the limit of $N \to \infty$. Then, we obtain

$$\frac{d}{dt}\boldsymbol{m} = -\boldsymbol{m} + \int dz[\mathcal{D}_{\boldsymbol{\zeta}}[z]\boldsymbol{\zeta} \tanh \beta({}^t\boldsymbol{\zeta}\boldsymbol{A}\boldsymbol{m} + z)]_{\boldsymbol{\zeta}}, \tag{29}$$

$$\frac{d}{dt}r = -r + 1 + \frac{1}{\alpha}\int dz[\mathcal{D}_{\boldsymbol{\zeta}}[z] z \tanh \beta({}^t\boldsymbol{\zeta}\boldsymbol{A}\boldsymbol{m} + z)]_{\boldsymbol{\zeta}}, \tag{30}$$

where $\mathcal{D}_{\boldsymbol{\zeta}}[z]$ is the probability density function of the cross-talk noise $z_i(\boldsymbol{s})$,

$$\mathcal{D}_{\boldsymbol{\zeta}}[z] \equiv 2^c \left\langle \frac{1}{N}\sum_i \delta(z - z_i(\boldsymbol{s}))\delta_{\boldsymbol{\zeta}\boldsymbol{\xi}_i} \right\rangle_{\boldsymbol{m},r;t},$$

$$\int dz[\mathcal{D}_{\boldsymbol{\zeta}}[z]]_{\boldsymbol{\zeta}} = 1,$$

where $\boldsymbol{\xi}_i = {}^t(\xi_i^1, \cdots, \xi_i^c)$. $\langle\Phi\rangle_{\boldsymbol{m},r;t}$ denotes the subshell average

$$\langle\Phi\rangle_{\boldsymbol{m},r;t} = \frac{\sum_{\boldsymbol{s}} p_t(\boldsymbol{s})\delta(\boldsymbol{m} - \boldsymbol{m}(\boldsymbol{s}))\delta(r - r(\boldsymbol{s}))\Phi(\boldsymbol{s})}{\sum_{\boldsymbol{s}} p_t(\boldsymbol{s})\delta(\boldsymbol{m} - \boldsymbol{m}(\boldsymbol{s}))\delta(r - r(\boldsymbol{s}))}.$$

Here, to obtain closed equations, we assume the following:

(1) The time evolutions of the order parameters $(\boldsymbol{m}, r)$ have the self-averaging property with respect to the average over the variables ξ_i^μ.

(2) The probability distribution of the micro variable $\boldsymbol{s}$ is independent of time and is uniform in the subshell where the macro variables $(\boldsymbol{m}, r)$ have the same values.

Then, we obtain the closed equations for $(\boldsymbol{m}, r)$ using the replica method under the RS ansatz,

$$\begin{aligned}\frac{d}{dt}\boldsymbol{m} &= -\boldsymbol{m} + [\boldsymbol{\zeta}\int Dx \int Dy\{1 - \tanh({}^t\boldsymbol{\zeta}\boldsymbol{\mu} + y\sqrt{\frac{\Delta}{2\epsilon\rho}}\lambda + \frac{\lambda^2}{\sqrt{2\epsilon\rho}}x)\} \\ &\quad \times \tanh\beta({}^t\boldsymbol{\zeta}\boldsymbol{A}\boldsymbol{m} + U^-)]_{\boldsymbol{\zeta}},\end{aligned} \tag{31}$$

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}r &= -r + 1 + [\int Dx \int Dy\frac{1}{\alpha}\{1 - \tanh({}^t\boldsymbol{\zeta}\boldsymbol{\mu} + y\sqrt{\frac{\Delta}{2\epsilon\rho}}\lambda + \frac{\lambda^2}{\sqrt{2\epsilon\rho}}x)\} \\ &\quad \times U^- \tanh\beta({}^t\boldsymbol{\zeta}\boldsymbol{A}\boldsymbol{m} + U^-)]_{\boldsymbol{\zeta}},\end{aligned} \tag{32}$$

where $\epsilon = \frac{\alpha r}{2}, \Delta = \frac{\alpha\rho(1-q)}{1-\rho(1-q)}$, $U^-(x) = \sqrt{2\epsilon}x - \Delta$ and $Dx = dx e^{-x^2/2}/\sqrt{2\pi}$. When $(\boldsymbol{m}, r)$ are given, the parameters λ, ρ, q and $\boldsymbol{\mu}$ are determined using the following equations.

$$\lambda = \frac{\rho\sqrt{\alpha q}}{1 - \rho(1-q)}, \tag{33}$$

$$r = \frac{1 - \rho(1-q)^2}{[1 - \rho(1-q)]^2}, \tag{34}$$

$$q = [\int Dw \tanh^2({}^t\boldsymbol{\zeta}\boldsymbol{\mu} + \lambda w)]_{\boldsymbol{\zeta}}, \tag{35}$$

$$\boldsymbol{m} = [\int Dw\boldsymbol{\zeta} \tanh({}^t\boldsymbol{\zeta}\boldsymbol{\mu} + \lambda w)]_{\boldsymbol{\zeta}}. \tag{36}$$

From eq. (34), ρ is given by

$$\rho = \frac{1}{2r(1-q)}[2r-1+q-\sqrt{(1-q)^2+4rq}] \geq 0. \tag{37}$$

In particular, the evolution equations for $T=0$ is given by

$$\begin{aligned}
\frac{d}{dt}\boldsymbol{m} &= -\boldsymbol{m} + [\boldsymbol{\zeta}\int Dy\{\int_0^\infty \frac{du}{\sqrt{2\pi}} e^{-\frac{1}{2}(u+x_-)^2}\,(1-\tanh(\Xi(u+x_-,y,\boldsymbol{\zeta})))\\
&\quad -\int_0^\infty \frac{du}{\sqrt{2\pi}} e^{-\frac{1}{2}(u-x_-)^2}\,(1-\tanh(\Xi(-(u-x_-),y,\boldsymbol{\zeta})))\}]_{\boldsymbol{\zeta}}, \qquad (38)\\
\frac{1}{2}\frac{d}{dt}r &= -r+1+\frac{1}{\alpha}[\int Dy\{\int_0^\infty \frac{du}{\sqrt{2\pi}} e^{-\frac{1}{2}(u+x_-)^2}\,(1-\tanh(\Xi(u+x_-,y,\boldsymbol{\zeta})))\,U^-(u+x_-)\\
&\quad -\int_0^\infty \frac{du}{\sqrt{2\pi}} e^{-\frac{1}{2}(u-x_-)^2}\,(1-\tanh(\Xi(-u+x_-,y,\boldsymbol{\zeta})))\,U^-(-u+x_-)\}]_{\boldsymbol{\zeta}}, \qquad (39)
\end{aligned}$$

where $\Xi(x,y,\boldsymbol{\zeta}) \equiv \boldsymbol{\zeta}\boldsymbol{\mu} + y\sqrt{\frac{\Delta}{2\epsilon\rho}}\lambda + \frac{\lambda^2}{\sqrt{2\epsilon\rho}}x$ and $x_- \equiv -\frac{1}{\sqrt{2\epsilon}}(\boldsymbol{\zeta}\boldsymbol{A}\boldsymbol{m}-\Delta)$.

In the next subsection, we give the results for $T=0$ by the numerical calculations of the S.P.E., and also the results of the evolution equations together with the results by numerical simultions.

4.2 Results

The equilibrium states

We solved the S.P.E. (25)-(27) numerically for several values of parameters a and α and found the coexistence region of correlated attractors, Hopfield attractors, and mixed states. In fig. 5, we show an example of the phase diagrams. When α is decreased, at $\alpha \simeq 0.0183$, a pair of correlated attractors appear. One of the correlated attractors continues to exist at least until $\alpha \simeq 0.0014$. This attractor is not shown in fig. 5 because it is unstable, that is, this is not an attractor but a repeller. Another correlated attractor, which is drawn in the figure, is stable and exists until $\alpha \simeq 0.0049$. At $\alpha \simeq 0.0049$, this correlated attractor disappears and then a mixed state with three patterns appears and continues to exist until $\alpha = 0$. Further, a mixed state with thirteen patterns exist for $\alpha \in (0, 0.3119)$. The Hopfield attractor exists only for $\alpha \in (0, 0.013)$. Thus, for $\alpha \in (0.0049, 0.013)$, the Hopfield attractor, a correlated attractor, and a mixed state with thirteen patterns coexist.

Dynamics

We studied the following two cases using deterministic dynamics.

Case 1 A Hopfield attractor and a correlated attractor coexist.

Case 2 A correlated attractor, but no Hopfield attractor.

In both cases, we performed numerical integrations of the evolution equations (38) and (39) and numerical simulations of eq. (1). In the numerical integration, we took $\boldsymbol{m}(0) = (m_0, 0, \cdots, 0)$ and $r(0) = 1$ as an initial condition. Since the integration of the evolution equations required a lot of computation time, we used the Euler method as the integration scheme.

In the numerical simulations, an initial value of s_i was determined according to the probability distribution (20). In both numerical integrations and numerical simulations, m_0 was taken as $m_0 = 0.1, 0.2, \cdots, 0.9$.

Case 1

As an example, we set $a = 0.35, \alpha = 0.01$. In fig. 6, we show the result of DRT and of numerical simulations. As the figure shows, the boundary between the basin of attraction for the Hopfiled attractor and that for the correlated attractor is around $m_1(0) = 0.5$. Except for the boundary of the basin of attraction, the results by two methods agree fairly well.

Case 2

As an example, we set $a = 0.35$ and $\alpha = 0.015$. As is seen from fig. 7, the trajectories initially approach the state where the Hopfield attractor existed but finally tend toward the correlated attractor in both the numerical integrations and numerical simulations except for $m_0 = 0.9$ in the numerical integration. In the case of $m_0 = 0.9$ in the numerical integration, our numerical scheme to solve the saddle point equations did not generate a solution at some time. We suggest that this is due to the precision of the numerical calculations.
The tendency was that trajectories obtained by numerical integrations tend toward the correlated attractor faster than those obtained by numerical simulations.

§5. Summary and Discussion

We studied the statics and dynamics of the model introduced by Griniasty et al. in the finite and extensive loading cases. In both cases, we qualitatively obtained similar results. First, we summarize the results common to both cases.

As for the statics, we solved the saddle point equations numerically and obtained equilibrium states. We found regions of parameters where the Hopfield attractor and a correlated attractor coexist. It had been previouly believed that correlated attractors exist only for $a > 0.5$. However, we found that the correlated attractors exist not only for $a > 0.5$ but also for $a < 0.5$.

As for the dynamics, we studied the following two cases by numerical integrations of the evolution equations and by the numerical simulations, deriving the evolution equations by dynamical replica theory in the extensive loading case.

First, we studied the transient behaviour when a Hopfield attractor and a correlated attractor coexist. We found that when we set initial conditions at $\boldsymbol{m}(0) = (m_0, 0, \cdots, 0)$ (and $r(0) = 1$ for the extensive loading case), there exists some critical initial overlap m_0^c, which is the boundary between the basin of attraction for the Hopfield attractor and that for the correlated attractor. Next, we studied the transient behaviour when a Hopfield attractor does not exist but a correlated attractor does exist. Then, we found that trajectories initially approach the state where the Hopfield attractor existed but finally they tend toward the correlated attractor.

Next, we compare the results obtained in the finite loading case and the results obtained in the extensive loading case.

Equilibrium state

In both cases, the Hopfield attractor, the mixed state, and the correlated attractor exist, and there are parameter regions where the Hopfield attractor, the mixed state, and the correlated attractors coexist.

Dynamics

1. When several attractors coexist

In the finite loading case ($a = 0.4, T = 0.04$), the boundary between the basin of attraction for the Hopfiled attractor and that for the correlated attractor is $0.15 < m_1^c < 0.16$ in the numerical integrations and $0.16 < m_1^c < 0.17$ in the numerical simulations. On the other hand, in the extensive loading case($a = 0.35, \alpha = 0.01, T = 0$), the boundary between the basin of attraction for the Hopfiled attractor and that for the correlated attractor is $0.4 < m_1^c < 0.5$ in both the numerical integrations and the numerical simulations. In the finite and extensive loading cases, the results of numerical integrations and the numerical simulations agree fairly well. One reason for the difference between the finite and extensive loading cases might be due to the difference in parameters.

2. When a Hopfield attractor does not exist but a correlated attractor does exist

In the finite loading case, we used $m_1(0) = 0.1, 0.2, \cdots, 1.0$ as the initial conditions. Then the trajectories finally tended toward the correlated attractor in both the numerical integrations and numerical simulations. On the other hand, for the extensive loading case ($a = 0.35, \alpha = 0.015, T = 0$), we used $m_1(0) = 0.1, 0.2, \cdots, 0.9$ and $r(0) = 1$ as initial conditions. Then, trajectories finally tended toward the correlated attractor in both the numerical integrations and numerical simulations, except for $m_1(0) = 0.9$ in the numerical integration for which it seems that the precision of the numerical calculation is not good enough. Except for the last case, the results of numerical integrations and numerical simulations agree fairly well in both the finite and extensive loading cases.

Therefore, we demonstrate that the results similar to those in the finite loading case have been obtained in the extensive loading case.

It might be considered that if we could construct a physiological experiment to observe the transient behaviour of a firing pattern of a monkey, we would be able to judge whether an attractor neural network exists in the brain of the monkey. As Miyashita and Chang reported, the neurons in the AVT are highly selective towards a few of the 100 colored fractal patterns. This implies that the firing rate of neurons in the AVT is very low, while the firing rate is set to 50% in the present model. Thus, in the future it will be necessary to develop a model with a low firing rate,[10] so that we can compare results of that model with physiological experimental results.

Acknowledgements

We are grateful to A. C. C. Coolen for valuable discussions and suggestions of this work. This work was partially supported by Grant-in-Aid for Scientific Research on Priority Areas No. 14084212 and Grant-in-Aid for Scientific Research (C) No. 14580438.

1) Y. Miyashita, Nature **335**(1988)817.
2) Y. Miyashita and H. S. Chang, Nature **331**(1988)68.
3) M. Griniasty, M. Tsodyks and D. J. Amit, Neural Comp.**5**(1993)1.
4) M. Shiino, NOLTA'97(Honolulu) 665.
5) M. Shiino and A. Yamano, NOLTA'98(Crans-Montana) 1027.
6) T. Fukai, T. Kimoto, M. Doi and M. Okada, J. Phys. A: Math. Gen. **32**(1999) 5551.
7) L. F. Cugliandolo and M. V. Tsodkys, J. Phys. A: Math. Gen. **27**(1994) 741.
8) S. N. Laughton and A. C. C. Coolen, Phys. Rev. E **51**(1995)2581.
9) J. Hertz, A. Krogh and R. G. Palmer: Introduction to the theory of neural computation(Addison-Wesley, Redwood, 1991) p.251.
10) M. Okada, Neural Networks, **9**(1996) 1429.

Captions

Fig.1: Equilibrium states in finite loading case. $a = 0.4$ and $c = 13$. The abscissa is T and the ordinate is $m_\mu(\mu = 1, \cdots, 7)$. Solid curve: Hopfield attractor. Dashed curves: correlated attractor. Dot-dash-curve: mixed state of thirteen patterns.

Fig.2: Results of numerical integrations and numerical simulations for the finite loading case. $a = 0.4, T = 0.04$ and $c = 13$. Initial condition is $\boldsymbol{m} = (m_0, 0, \cdots, 0)$ and $m_0 = 0.1, 0.2, \cdots, 1.0$. Dased curves: simulation for $N = 60,000$. Solid curves: theory. (a) Trajectory in space of (m_1, m_2). C: correlated attractor. H: Hopfield attractor. (b)Time evolution of m_1.

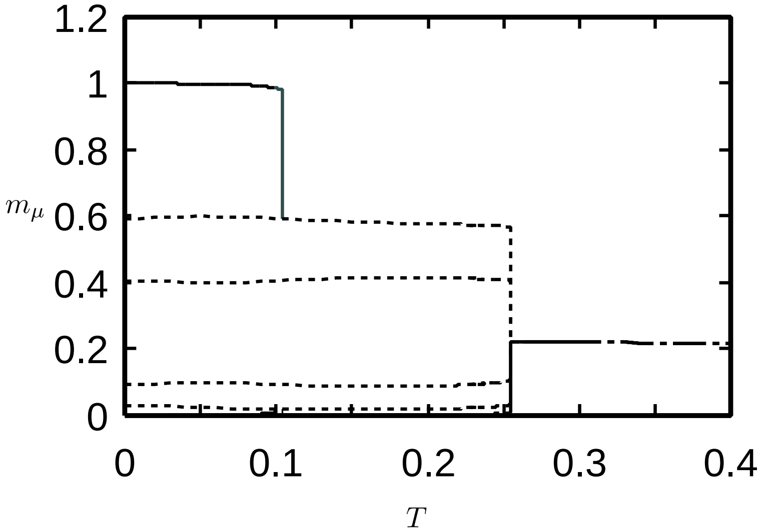
Fig.3: Results of numerical integrations (a) and numerical simulations (b) for finite loading case. $a = 0.4$, $T = 0.04$ and $c = 13$. Initial condition is $\boldsymbol{m} = (m_0, 0, \cdots, 0)$. (a) Solid curve: $m_0 = 0.15$. Trajectory tends toward correlated attractor. Dashed curve: $m_0 = 0.16$. Trajectory tends toward Hopfield attractor. (b) $N = 60,000$. Solid curve: $m_0 = 0.16$. Trajectory tends toward correlated attractor. Dashed curve: $m_0 = 0.17$. Trajectory tends toward Hopfield attractor.

Fig. 4: Results of numerical integrations and numerical simulations for finite loading case. $a = 0.4, T = 0.15$ and $c = 13$. Dashed curves: simulation for $N = 60,000$. Solid curves: theory. (a) Trajectory in space of (m_1, m_2). C: correlated attractor. H: point where Hopfield attractor existed. (b) Time evolution of m_1.

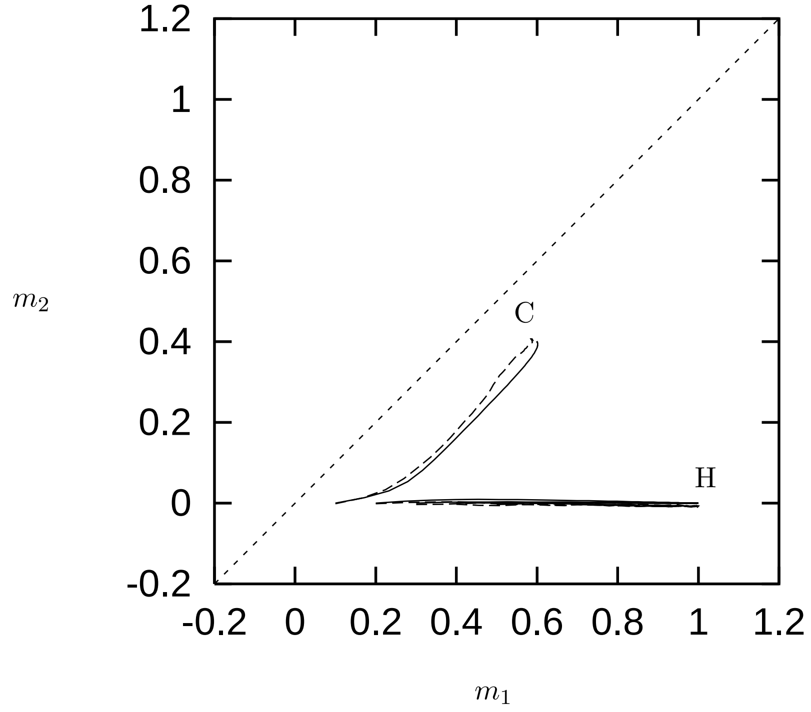
Fig.5: Equilibrium states for $a = 0.35$ and $T = 0$ for case of extensive loading. The abscissa is α and the ordinate is $m_\mu(\mu = 1, \cdots, 7)$. Solid curve: Hopfield attractor. Dashed curves: correlated attractor. Long dashed curves: mixed state of three patterns. Doted curve: mixed state of thirteen patterns.

Fig.6: Results of numerical integrations and numerical simulations for extensive loading case. $a = 0.35, \alpha = 0.01, T = 0$ and $c = 13$. Dashed curves: simulation for $N = 60,000$. Solid curves: DRT. (a) Trajectory in space of (m_1, m_2). C: correlated attractor. H: point where Hopfield attractor existed. (b)Time evolution of m_1.

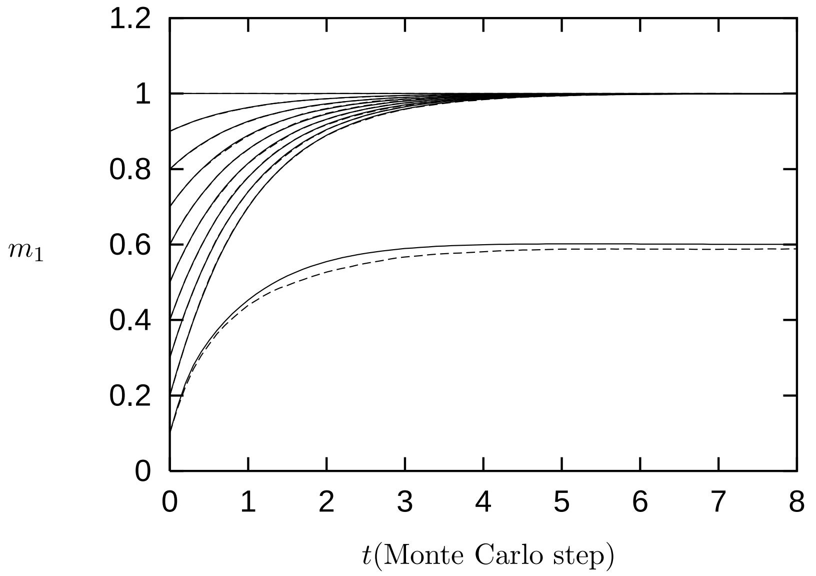
Fig.7: Results of numerical integrations and numerical simulations for the extensive loading case. $a = 0.35, \alpha = 0.015, T = 0$ and $c = 13$. Dashed curves: simulation for $N = 60,000$. Solid curves: DRT. (a) Trajectory in space of (m_1, m_2). C: correlated attractor. H: Hopfield attractor. (b)Time evolution of m_1.


1.2
1
0.8
0.6
0.4
0.2
0
m_μ
0
0.1
0.2
0.3
0.4
T

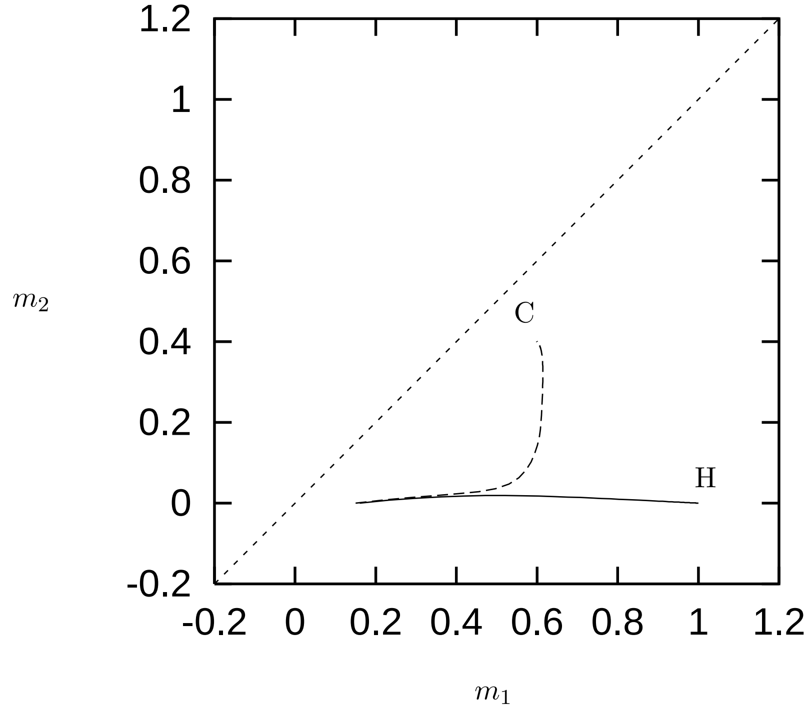
(a)


1.2
1
0.8
0.6
0.4
0.2
0
-0.2
m_2
C
H
-0.2 0 0.2 0.4 0.6 0.8 1 1.2
m_1

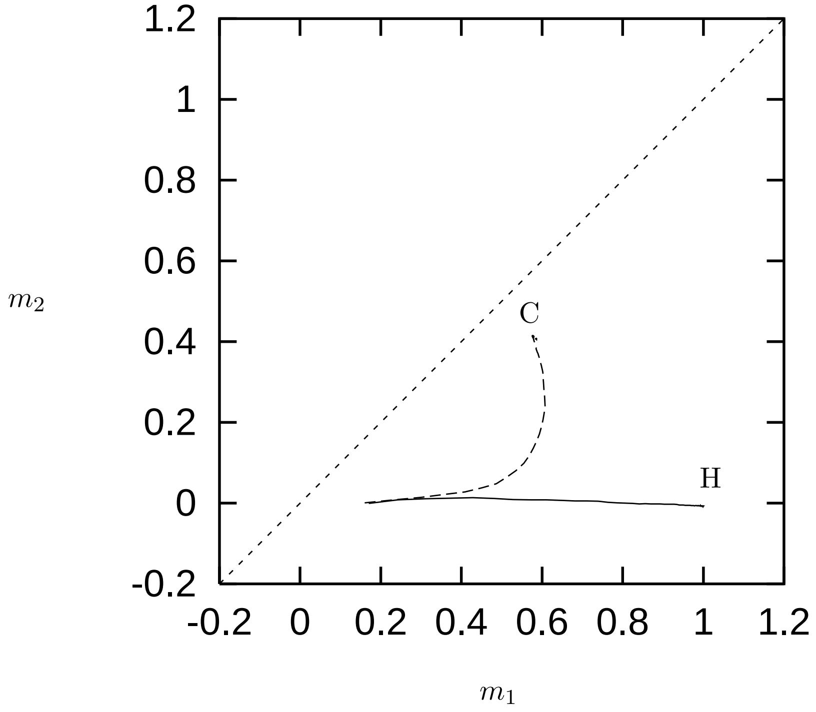
(b)


1.2
1
0.8
0.6
0.4
0.2
0
m_1
0 1 2 3 4 5 6 7 8
t(Monte Carlo step)

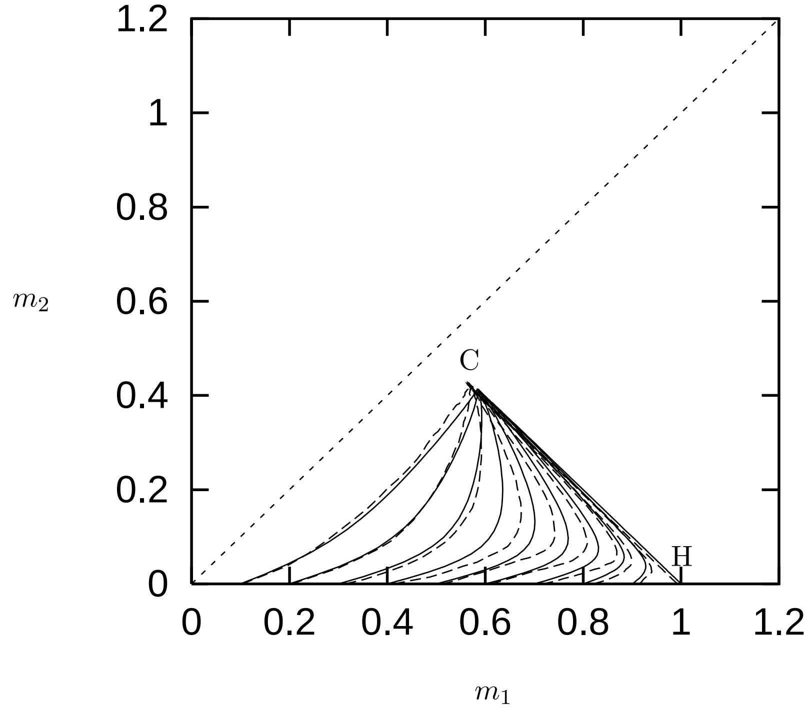
(a)


1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.2 0 0.2 0.4 0.6 0.8 1 1.2
m_2
m_1
C
H

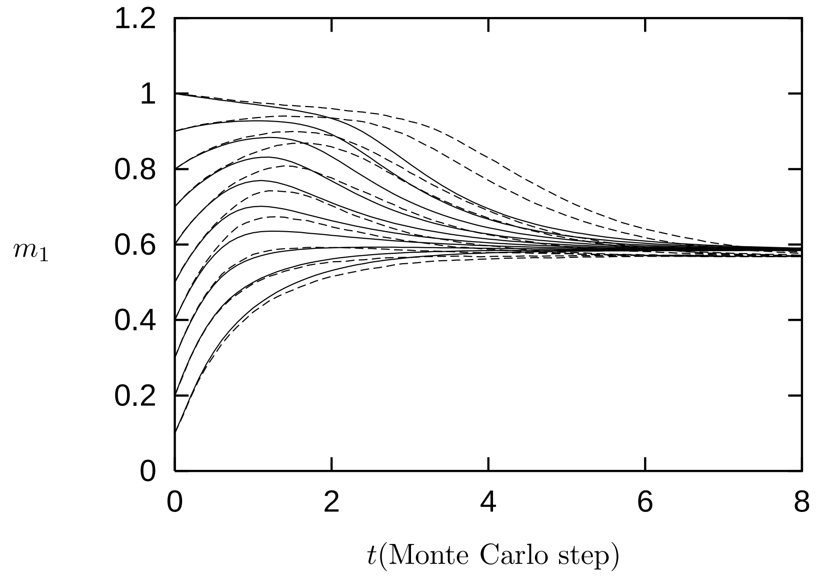
(b)

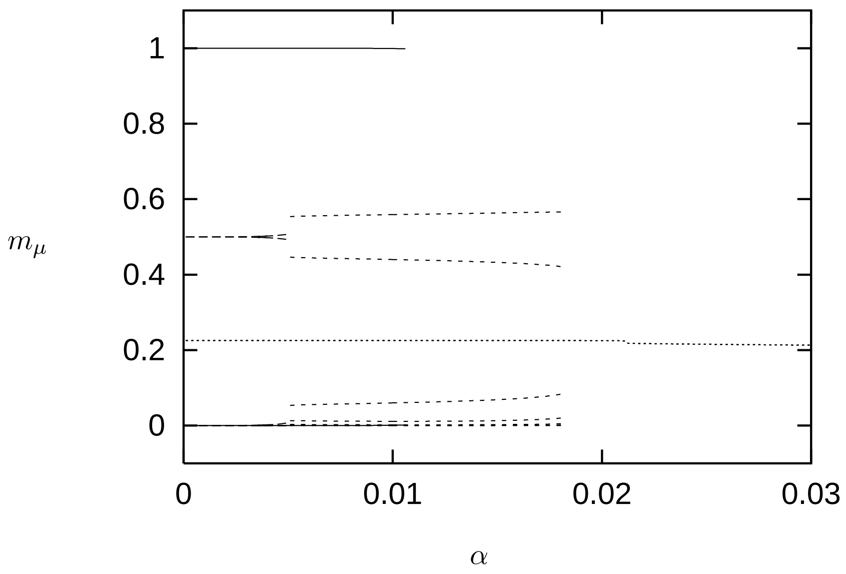

1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.2 0 0.2 0.4 0.6 0.8 1 1.2
m_2
m_1
C
H

(a)

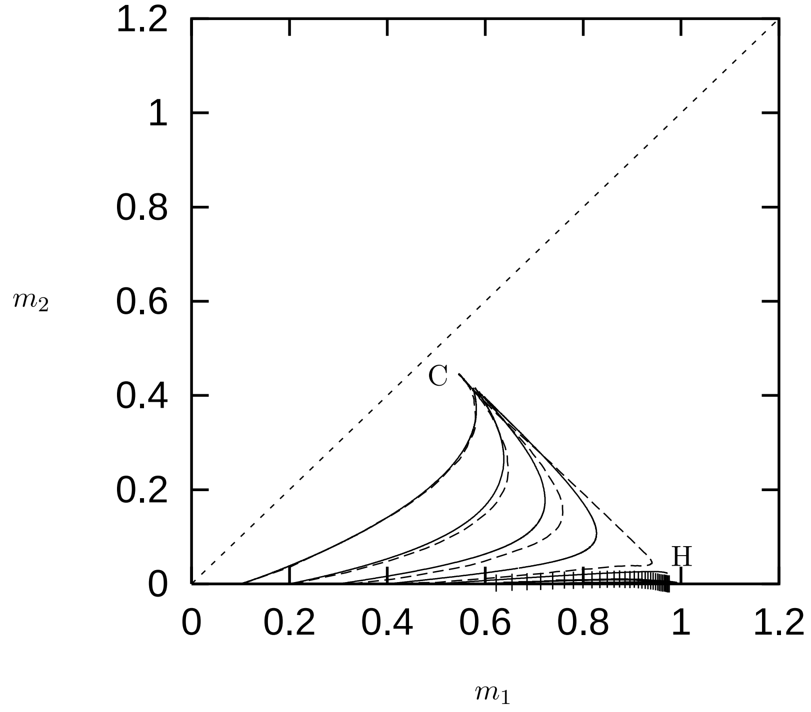

1.2
1
0.8
0.6
0.4
0.2
0
m_2
C
H
0
0.2
0.4
0.6
0.8
1
1.2
m_1

(b)

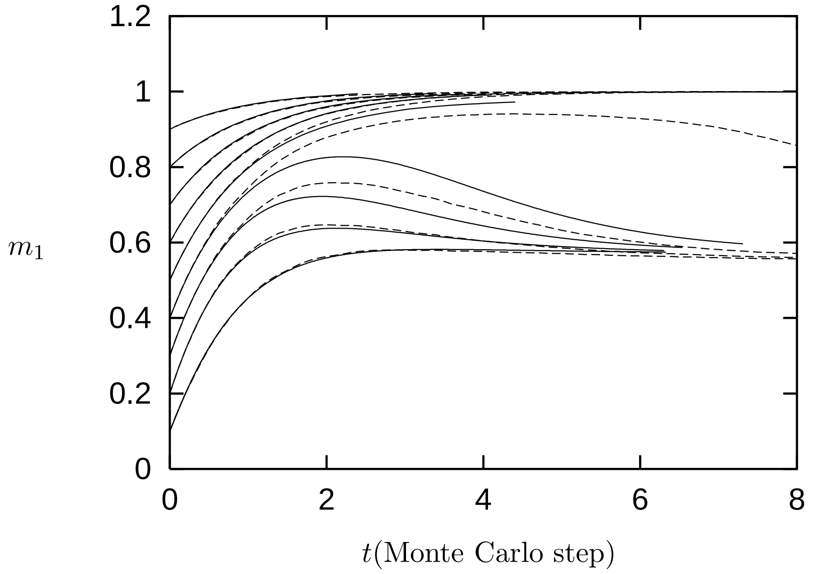

1.2
1
0.8
0.6
0.4
0.2
0
m_1
0
2
4
6
8
t(Monte Carlo step)


1
0.8
0.6
0.4
0.2
0
0
0.01
0.02
0.03
m_μ
α

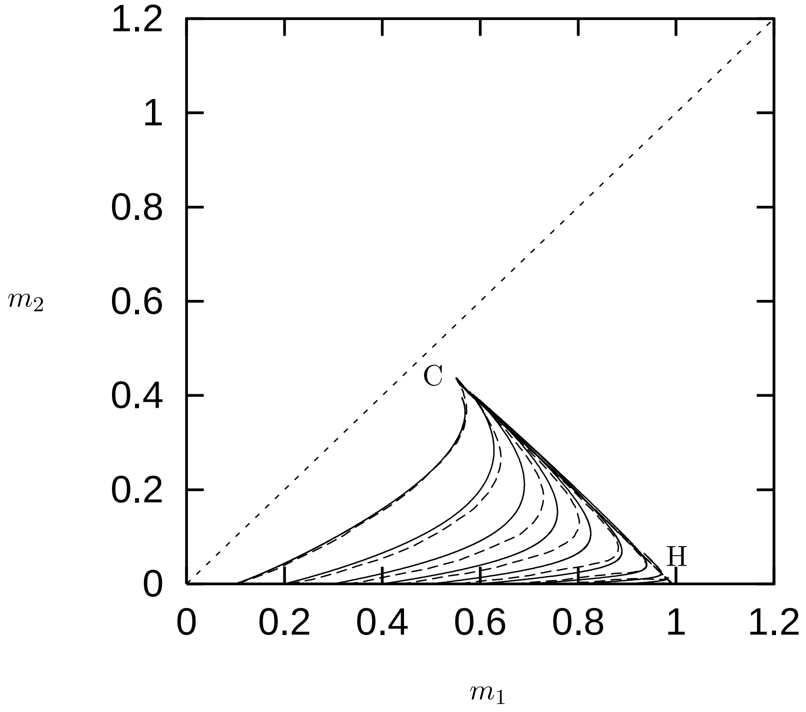
(a)


1.2
1
0.8
m_2
0.6
0.4
0.2
0
C
H
0
0.2
0.4
0.6
0.8
1
1.2
m_1

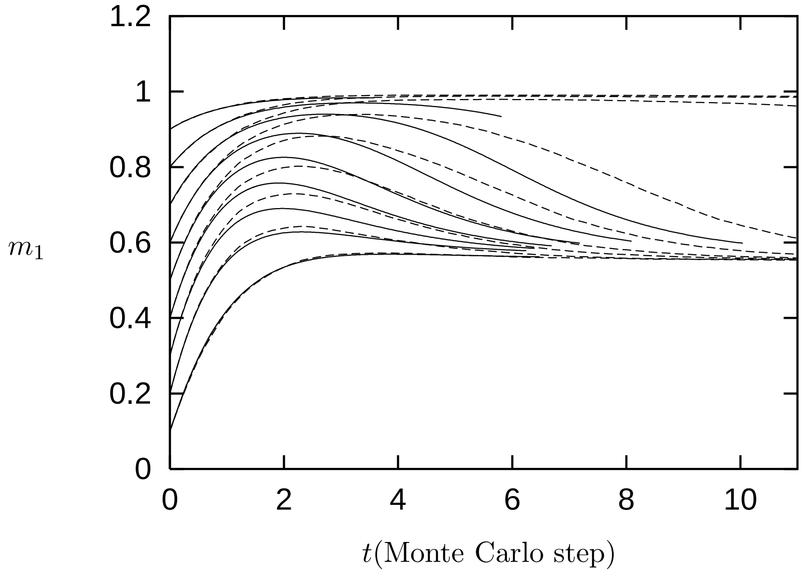
(b)


1.2
1
0.8
m_1
0.6
0.4
0.2
0
0
2
4
6
8
t(Monte Carlo step)

(a)


1.2
1
0.8
0.6
0.4
0.2
0
m_2
C
H
0
0.2
0.4
0.6
0.8
1
1.2
m_1

(b)


1.2
1
0.8
0.6
0.4
0.2
0
m_1
0
2
4
6
8
10
t(Monte Carlo step)