Local minimum states of the binary multi-charge network model

2001 J. Phys. A: Math. Gen. 34 11241

(http://iopscience.iop.org/0305-4470/34/50/306)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 11241–11252 PII: S0305-4470(01)24926-9

Local minimum states of the binary multi-charge network model

Kazuo Nokura

Shonan Institute of Technology, Fujisawa 251-8511, Japan

E-mail: nokura@la.shonan-it.ac.jp

Received 8 May 2001, in final form 29 October 2001
Published 7 December 2001
Online at stacks.iop.org/JPhysA/34/11241

Abstract

We study the number of local minimum states of the multi-charge network spin glass model with binary spin variables $\rho = 0, 1$ and external fields. The energy function is defined by changing the sign of the Hopfield energy function. This model is motivated by considerations on the shape–complementary-shape interactions among protein molecules.

PACS numbers: 87.10.+e, 75.50.Lk, 64.60.Cn

1. Introduction

In recent years, studies of infinite range spin glass models [1, 2] have given much inspiration to studies of biological networks such as neural network models [3]. The energy structures of these models are complex, which is typically expressed by the enormous number of local minimum states.

A few years ago, we introduced a spin glass model which is defined by reversing a sign of the Hopfield energy function [4, 5]. With spin variables η_i $(i = 1, 2, \ldots, N)$ and external field h_0, this model has an energy function

$$H = -\frac{1}{2}\sum_{i\neq j} J_{ij}\eta_i\eta_j + h_0\sum_i \eta_i \tag{1}$$

where interactions J_{ij} are given by

$$J_{ij} = -\frac{1}{N}\sum_{\mu} \xi_i^{\mu}\xi_j^{\mu} \tag{2}$$

with $J_{ii} = 0$. The quenched variables $\xi_i^{\mu}(\mu = 1, 2, \ldots, P)$ are assumed to be ± 1 with probability $1/2$. Putting (2) into (1), we see that the spin variables interact as if they bear several charges given by ξ_i^{μ}. For this reason, we call this type of spin glass model the multi-charge network (MCN) model.

0305-4470/01/5011241+12$30.00

In this paper, with biological modelling in mind, we study the number of local minimum states of (1) for binary spin variables $\eta_i = \rho_i - r$, where $\rho_i = 0, 1$ and r is a parameter which controls the proximity to the Ising spin ($r = 1/2$) and pure binary spin ($r = 0$).

There are several reasons which make the MCN model interesting in the studies of biological systems. First, it reflects the effect of unlearning in random neural network models [4]. The idea of unlearning was introduced to discuss the function of REM sleep and to improve the neural network model [6, 7]. When this idea is applied to a random neural network model, interactions become correlated, which will be qualitatively expressed by (2). Secondly, this model is expected to simulate the protein molecule networks which are controlled by shape–complementary-shape interactions among protein molecules [5]. In this case, the MCN model will provide a simple spin-glass-like model of the immune system.

The function of the immune system has been an attractive subject in biology. The immune system is made of many different kinds of protein molecules, which respond to foreign materials and render them harmless. Further, information about foreign materials is maintained in the system to maintain immunity. More than two decades ago, the network model of the immune system was suggested to discuss these properties [8]. About a decade ago, a spin-glass-like model was suggested to discuss the capacity of the memory by assuming random interactions among protein molecule concentrations [9, 10]. The advantage of the spin-glass-like model is that we can use various results obtained for spin glass models. In addition, formulations by infinite range spin interactions make several analytic studies feasible.

Let us describe the statistical mechanics of the Ising MCN model with no external fields. By using the replica method, we found that there is a dynamical spin glass phase transition for $\alpha \equiv P/N < \alpha_c \sim 1.4$, while the phase transition is similar to the Sherrington–Kirkpatrick model for $\alpha > \alpha_c$ [4]. The MCN model for small α has properties very similar to the spin glass models which were introduced recently by several authors [11–13]. We also studied the number of local minimum states in the form of $\exp(Ng)$ and found that g increases to the possible maximum value $\ln 2$ as $\alpha \to 0$ [5]. Although there will be some corrections which disappear in the thermodynamic limit, this result implies that memory effects tend to the maximum for $\alpha \to 0$. Correspondingly, we found numerically that remanent magnetizations tend to 1 in this limit. These results imply that the memory effects are much stronger than the case of random interactions.

In this paper, we extend the studies of local minimum states of the MCN model to binary spin variables, which are more realistic for biological systems. In section 2, we describe the model in the context of protein molecule networks. Section 3 is devoted to the derivation of the mean-field equations for the number of local minimum states. Section 4 is devoted to studies of the saddle point equations for various parameters. In particular, we address the properties in the $\alpha \to 0$ limit. Section 5 is devoted to some discussions.

2. Model description

The MCN model is obtained simply by changing the sign of the Hopfield model. Originally, this model was motivated by the observations on unlearning in the SK model. In this section, we present another motivation for this model. Although the idea of multi-charge was described in the previous paper [5], we describe them in terms of protein molecule interactions.

It is known that the interactions among protein molecules are governed by dual shapes on the surfaces of the molecules. That is, molecules interact attractively via the complementarity of their shapes. The interactions are weak otherwise. To formulate this, we imagine formal protein molecules and introduce quenched variables ξ_i^μ to indicate that there is a shape μ if $\xi_i^\mu = 1$, its complementary shape if $\xi_i^\mu = -1$ on the surface of a kind of protein

molecule i. In this way, a kind of molecule i ($i = 1, 2, \ldots, N$) is characterized by shapes ξ_i^μ ($\mu = 1, 2, \ldots, P$). P represents the number of kinds of dual shapes of protein molecules and N is the number of different kinds of molecules which make the networks.

Now we discuss the effective energy function for protein molecule concentrations η_i, which will be given by the quadratic function of η_i. In terms of shape μ of molecules i and molecules j, shape–complementary-shape interactions will give a factor $\xi_i^\mu \xi_j^\mu - \left(\xi_i^\mu \xi_j^\mu\right)^2$ in the energy function. In this expression, the second term can be dropped since they are cancelled by terms which control protein concentrations. Further, there are no cross terms between different shapes since different shapes interact weakly. We can introduce self-interaction terms for binary spin variables. Such terms will be proportional to the squares of concentrations, which are expressed by external field terms for binary spin variables. In this way, we reach the energy function expressed by (1) with (2).

Since ξ_i^μ work like charges, the idea of neutralization naturally arises by writing the energy function in the form

$$H = \frac{1}{2N} \sum_\mu \left(\sum_i \xi_i^\mu \eta_i \right)^2 + h_0 \sum_i \eta_i - \frac{1}{2} \alpha \sum_i \{(1 - 2r)\eta_i + r(1 - r)\}. \quad (3)$$

This form implies that, for $h_0 = (1 - 2r)\,\alpha/2$, the minimum of the energy function is given by $\sum_i \xi_i^\mu \eta_i = 0$ for all μ. This requirement can be viewed as a neutralization of all kinds of charges. As we will see later, this value of h_0 roughly gives the maximum of the number of local minimum states for a given α. Although the neutralization is not satisfied exactly for discrete spin variables, this form of the energy function implies that, for $P < N$, there exists an $(N - P)$-dimensional configuration space which gives very low energy, while there is no such space when $P > N$. This suggests that there is a large number of local minimum states for $\alpha \sim 0$, which implies strong remanent properties. For these reasons, we are especially interested in the model with small α.

Let us give some remarks. Since N is the number of kinds of molecules, ξ_i^μ should be different vectors for different μ. Thus, if ξ_i^μ are assigned randomly, N should be much smaller than 2^P, which is the maximum number of kinds of protein molecules. This gives the relation $P \ln 2 \gg \ln N$ or $\alpha \gg \ln N/(N \ln 2)$. This is the lower bound of α, which is very small for large N, to simulate the protein molecule networks. In this paper, we assume that α is of order 1 to apply mean-field theory and concentrate on the small α.

These arguments suggest that the number of local minimum increases as the number of kinds of protein molecules increases when $h_0 = (1 - 2r)\,\alpha/2$. We will study this point in the next section.

3. Derivation of the saddle point equations

This section is devoted to the formulation of the mean-field method to study the number of local minimum states of the binary MCN model. We follow the calculations presented in [14].

When the temperature is zero, the energy function (1) defines a gradient dynamics which makes the energy lower and leads to the attractors of the dynamics. They are local minimum states of (1) which are defined by $\rho_i = 0$ if the local field is negative and $\rho_i = 1$ if the local field is positive. Thus, they should satisfy the equations

$$\rho_i = \theta\left(\sum_j J_{ij}(\rho_j - r) - h_0 \right) \quad (4)$$

for all i, where $\theta(x) = 0$ for $x < 0$ and $\theta(x) = 1$ for $x \geqslant 0$. Note that positive h_0 suppresses $\rho_i = 1$. For $r = 1/2$ and $h_0 = 0$, the equation (4) reduces to the Ising spin model without external

fields. In this case, $\eta_j = \pm 1/2$ will appear with probability $1/2$, implying $h_i = \sum_j J_{ij}\eta_j$ distribute evenly around 0 when all solutions are scanned. This is a kind of consistency between the right-hand side and the left-hand side of (4). When $0 \leqslant r < 1/2$, we expect that, by adjusting h_0, $\rho_i = 0, 1$ appear with the same probability.

Equation (4) implies that the local minimum states satisfy

$$\eta_i \left(\sum_j J_{ij}\eta_j - h_0 \right) > 0 \tag{5}$$

for all i, where $\eta_i = \rho_i - r = -r, 1 - r$. The number of local minimum states is given by

$$G = \sum_{\{\eta\}} \prod_i \int_0^\infty \delta \left(\eta_i \left(\sum_j J_{ij}\eta_j - h_0 \right) - h_i \right) \mathrm{d}h_i. \tag{6}$$

For $r = 0$, $\eta_i = -r$ are assumed to be infinitesimally small negative values.

ξ_i^μ averages of G, which are denoted by $\langle G \rangle$, are obtained after expressing the delta functions by integral representations. We describe the calculations in the appendix, where definitions of the saddle point variables are also given. Then the problem reduces to finding the extremum of

$$\langle G \rangle = \mathrm{Extr}\{\exp(Ng)\} \tag{7}$$

where

$$g = -\tfrac{1}{2}\alpha \ln\{(1+B)^2 - AC\} + \alpha B - \bar{A}A - \bar{B}B - \bar{C}C + \ln z \tag{8}$$

where

$$z = \sum_{\eta=-r,1-r} \Phi(w(\eta)) \exp(\eta^2 \bar{C})$$

with

$$\Phi(x) = \int_{-\infty}^{x} \exp\left(-\tfrac{1}{2}t^2\right) \frac{\mathrm{d}t}{\sqrt{2\pi}}$$

and

$$w(\eta) = \frac{\eta^2 \bar{B} - \eta h_0}{\sqrt{2\bar{A}\eta^2}}.$$

Note that, when $r = 0$, we should first take the limit $r \to 0$ in $w(-r)$.

Using g, we obtain the saddle point equations as follows:

$$\bar{A} = \frac{1}{2} \frac{\alpha C}{(1+B)^2 - AC}$$

$$\bar{B} = \alpha - \alpha \frac{1+B}{(1+B)^2 - AC}$$

$$\bar{C} = \frac{1}{2} \frac{\alpha A}{(1+B)^2 - AC}$$

$$A = -\frac{1}{z} \frac{1}{2\bar{A}} \sum_\eta w(\eta)\Phi'(w(\eta)) \exp(\eta^2 \bar{C})$$

$$B = \frac{1}{z} \frac{1}{\sqrt{2\bar{A}}} \sum_\eta |\eta| \Phi'(w(\eta)) \exp(\eta^2 \bar{C})$$

$$C = \frac{1}{z} \sum_\eta \eta^2 \Phi(w(\eta)) \exp(\eta^2 \bar{C})$$

where the η-sum is over $-r, 1 - r$.

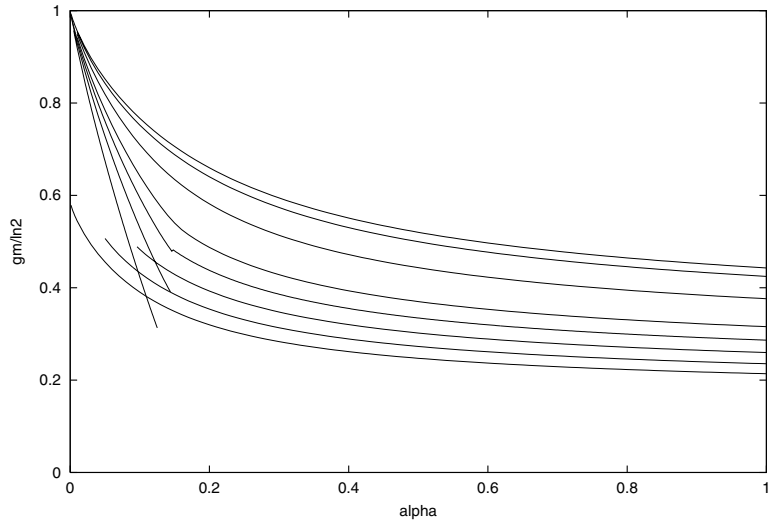

Figure 1. α-dependence of $g_m/\ln 2$ with $h_0 = 0$ and $r = 0.0, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5$ from the bottom.

Fortunately, analytic studies of the saddle point equations are feasible in the $\alpha \to 0$ limit, which is the most interesting situation. The details will be discussed in the next section.

4. Solutions of the saddle point equations

By using the saddle point equations presented in section 3, we can obtain the saddle point values of g, which is denoted by g_m, as a function of r, h_0 and α. The situation $r = 1/2$ and $h_0 = 0$, i.e. the Ising spin case was studied in [5]. Here, we first study the case $h_0 = 0$ with various r and then study the pure binary spin case $r = 0$ with varying h_0, which is more realistic biologically. Fortunately, analytic studies are feasible for $\alpha \sim 0$, since $w(\eta)$ with suitable parameter values tends to positive infinity in this limit.

Although we first focus on the case $h_0 = 0$, some expressions are not restricted to this situation as long as two $w(\eta)$ tend to positive infinity for $\alpha \to 0$.

4.1. Solutions with $h_0 = 0$

Let us first concentrate on the case $h_0 = 0$ with various r. Before discussing the analytic solutions, we describe the numerical results for $h_0 = 0$. Figure 1 shows the α-dependence of $g_m/\ln 2$ for various r. As expected, g_m increases as $\alpha \to 0$. On the other hand, at a given α, g_m monotonically decreases as r decreases from the Ising case $r = 1/2$. There are two peculiar aspects in figure 1. First, in the $\alpha \to 0$ limit, the solutions for $r \neq 0$ give the limit $g_m \to \ln 2$, while for $r = 0$, the limiting value of g_m is smaller than ln 2. Secondly, there is a crossover point around $\alpha \sim 0.1$ for small r. These results imply that $\alpha = 0$ is a singular point of g_m as a function of r.

Let us discuss the solutions of the saddle point equations for small α. When $h_0 = 0$, the equations for A and B in section 3 yield

$$A = -\frac{\bar{B}}{2\bar{A}}B. \tag{9}$$

On the other hand, numerical studies imply that $-A$ is positive and increases faster than $1/\alpha$, while other variables remain finite or tend to zero as $\alpha \to 0$. Then, the saddle point equations for $\bar{A}$, $\bar{B}$, $\bar{C}$ reduce to

$$\bar{A} \sim -\frac{\alpha}{2A} \qquad \bar{B} \sim \alpha \qquad \bar{C} \sim -\frac{\alpha}{2C}. \tag{10}$$

With these relations and the behaviour of $-A$, we assume that $\bar{B}/\sqrt{2\bar{A}} \to \infty$ and $\Phi(w(\eta)) \to 1$ as $\alpha \to 0$, except for the pure binary case $r = 0$, which will be discussed later. As we will see, $\Phi(w(\eta)) - 1 \to 0$ faster than α. Thus we simply assume that they can be set to 1 to discuss the first-order terms of α.

With these assumptions, we have the closed equations for C and $\bar{C}$, which give

$$C \sim \gamma - \frac{\alpha}{4\gamma}\{r^4 + (1-r)^4 - 2\gamma^2\} \tag{11}$$

to the first order of α, where $\gamma = (r^2 + (1-r)^2)/2$. Note that this expression holds for $h_0 \neq 0$ as long as $\Phi(w(\eta)) \sim 1$. On the other hand, using (9) and (10), we have $B \sim 1$, which reads

$$\sqrt{\frac{\alpha}{|A|}} \sim \frac{1}{2\sqrt{2\pi}} \sum_{\eta} |\eta| \exp\left(-\tfrac{1}{2}\eta^2 \alpha |A|\right) \tag{12}$$

where we used $z \sim 2$ and $\bar{C} \sim 0$. Since $r^2 < (1-r)^2$, the term with $\eta = -r$ contributes mainly on the right–hand side of this equation. Then we asymptotically obtain

$$r^2|A| \sim \tfrac{2}{\alpha} \ln \tfrac{1}{\alpha} \tag{13}$$

which really increases faster than $1/\alpha$. Summarizing these results, we obtain for $r \neq 0$,

$$g_m \sim \ln 2 - \frac{1}{2}\alpha \ln\left\{4 + C\frac{2}{r^2\alpha} \ln\frac{1}{\alpha}\right\} + \frac{1}{2}\alpha \tag{14}$$

where we used $z \sim 2 - \alpha$ in $\ln z$. This result gives $g_m/\ln 2 \to 1$ and the factor $1/r^2$ implies that the gradient for $\alpha \sim 0$ increases as $r \to 0$ in accordance with figure 1.

For $r = 0$, the contributing term in (12) changes from $\eta = -r$ to $\eta = 1 - r$. We also notice that $\Phi(w(0)) = 1/2$ and $\Phi(w(1)) \to 1$ as $\alpha \to 0$ by assuming $w(1) \to +\infty$. Repeating similar calculations, we have $C = C_0 \sim 2/3 - \alpha/6$, $z \sim 3/2(1-\alpha/2)$ and $|A| \sim (2/\alpha)\ln(1/\alpha)$, yielding

$$g_m \sim \ln\frac{3}{2} - \frac{1}{2}\alpha \ln\left\{4 + C_0\frac{2}{\alpha} \ln\frac{1}{\alpha}\right\} + \frac{1}{2}\alpha. \tag{15}$$

This result gives $g_m/\ln 2 \to \ln(3/2)/\ln 2 \sim 0.585$ as $\alpha \to 0$. By these calculations, we see that the difference of g_m is due to the difference of limiting values of $w(\eta)$.

Let us give some remarks. First, in the above calculations, we used $\Phi(w(\eta)) \sim 1$ repeatedly. This is true to the first order of α, since typically $1 - \Phi(\bar{B}/\sqrt{2\bar{A}}) \sim \exp(-\alpha|A|/2)/(\sqrt{2\pi\alpha|A|}) \sim \alpha/\sqrt{\ln(1/\alpha)}$. Secondly, we should note that the asymptotic behaviour of $|A|$ will be qualitatively correct also for $h_0 > 0$ as long as two $w(\eta)$ tend to positive infinity. This condition implies that h_0 should be scaled as $c\alpha$ with a suitable constant c. We discuss this point in the next subsection.

4.2. Two characteristic h_0 for $h_0 > 0$

To discuss non-zero h_0, it is helpful to identify two characteristic h_0 which characterize the behaviour of g_m: one is defined by $g_m \sim 0$ and the other is defined by the maximum condition on g_m. In the following arguments, we use $\bar{B} \sim \alpha$ to give some idea of them, especially their α-dependence. The picture obtained will be qualitatively correct for moderate α.

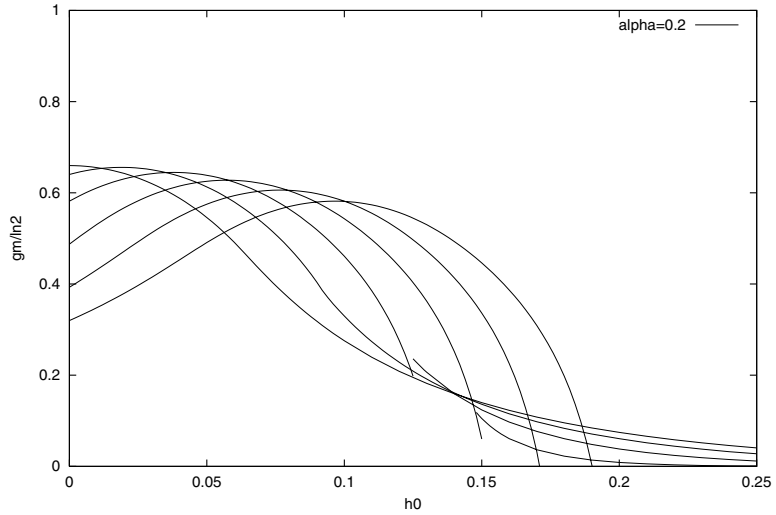

Figure 2. h_0-dependence of $g_m/\ln 2$ with $\alpha = 0.2$ and $r = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5$. The maxima are approximately located at $0.1(1 - 2r)$.

The first characteristic h_0 is denoted by h_b, beyond which g_m are practically zero. h_b is determined by $\Phi(w(1-r) = 1/2$ or $w(1-r) = 0$. For $h_0 > h_b$, g_m is very small especially for small α since $w(1-r) \to -\infty$. The condition $w(1-r) = 0$ gives

$$h_b = (1-r)\bar{B} \sim (1-r)\alpha. \tag{16}$$

Note that h_b scales as α in accordance with $J_{ij}^2 \sim \alpha/N$.

The second characteristic h_0, which is denoted by h_m, gives the maximum of g_m with other parameters fixed. This is defined by $\mathrm{d}g_m/\mathrm{d}h_0 = 0$, which reduces to $\partial g_m/\partial h_0 = 0$ at the saddle points, where the partial differential means differential with fixed saddle point variables. This gives

$$\Phi'(w(-r))\exp(r^2\bar{C}) = \Phi'(w(1-r))\exp((1-r)^2\bar{C}) \tag{17}$$

which can be solved in terms of h_m, giving

$$\begin{aligned} h_m &= \frac{(1-2r)(\bar{B}^2 - 4\bar{A}\bar{C})}{2\bar{B}} \\ &\sim \frac{(1-2r)\alpha}{2} \end{aligned} \tag{18}$$

where, for the second line, we used the asymptotic behaviour of the saddle point variables for $\alpha \to 0$.

For $\alpha = 0.2$, figure 2 shows $g_m/\ln 2$ as functions of h_0 for various r. As expected, g_m for the Ising spin case $r = 1/2$ takes a maximum value at $h_0 = 0$, whereas for $0 \leqslant r < 1/2$, g_m takes maxima at finite $h_0 \sim h_m$. For $r \sim 0$, the tails of g_m have break points, beyond which g_m are practically zero. These points are located at h_0 slightly smaller than h_b.

4.3. Solutions with $r = 0$ for various h_0

In this subsection, we concentrate on the situation $r = 0$, which is natural biologically since sites with $\eta_i = 0$ do not contribute to the local fields on other sites.

The behaviour of g_m with $r = 0$ was partly discussed in the previous subsections. Figure 1 shows that, with no external field, $g_m \to \ln 3/2$ instead of $\ln 2$ for $\alpha \to 0$. Figure 2 shows that,

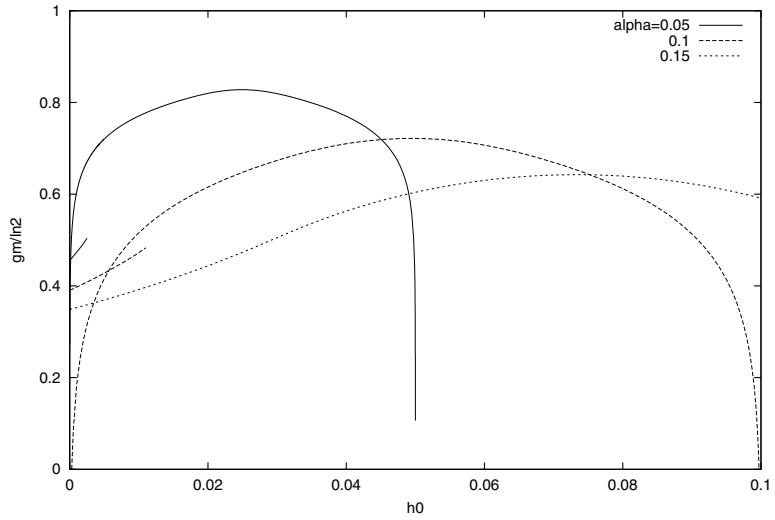

Figure 3. h_0-dependence of $g_m/\ln 2$ with $r = 0$ and $\alpha = 0.05$, 0.1, 0.15. For $\alpha = 0.05$ and 0.1, $g_m/\ln 2$ shows maxima approximately at $\alpha/2$ and they vanish beyond $h_0 \sim \alpha$. Note that for $\alpha = 0.05$ and 0.1 $g_m/\ln 2$ has other branches which continuously become the values in figure 1 for $h_0 \to 0$.

for $\alpha = 0.2$, g_m with $r = 0$ increases to a maximum as h_0 increases to h_m and then decreases as $h_0 \to h_b$. In this subsection, we are interested in the h_0-dependence of g_m for smaller α.

For $r = 0$, figure 3 shows the h_0-dependence of g_m for $\alpha = 0.05$, 0.1, 0.15. The values at $h_0 = 0$ were already shown in figure 1. g_m with $\alpha = 0.15$ is similar to that with $\alpha = 0.2$ in figure 2, while g_m for $\alpha = 0.05$, 0.1 has two branches; one tends to the values in figure 1 for $h_0 \to 0$ and other decreases drastically for $h_0 \sim 0$ and $\sim h_b$ with the maxima at $h_m \sim \alpha/2$. The crossovers between the two branches tend to zero faster than α as α decreases. This is because h_0 should be much smaller than $\sqrt{2\bar{A}} \sim \alpha/\sqrt{\ln(1/\alpha)}$ to give $w(0) \sim 0$. We can say that finite h_0 changes g_m drastically for small α even after eliminating a scale factor α.

Figures 1 and 3 imply that, for $r = 0$, these crossover points are always located in the region $h_0 > 0$, whereas for small positive r, they appear on the line $h_0 = 0$ as shown in figure 1. Thus the crossover points for $r = 0$ should form a line which ends at the point $h_0 = 0$ and $\alpha = 0$, while this line crosses the line $h_0 = 0$ at some positive α for small but positive r.

The existence of the crossover points implies that the contributing regions in the configuration space change discontinuously around these points since the saddle point with larger g_m mainly contributes to $\langle G \rangle$. According to the definition in appendix, C is the average of η_i^2 over local minimum states, which equals the average of ρ_i for $r = 0$. For example, we found numerically that the crossover point for $\alpha = 0.1$ is located at $h_0 \sim 0.0055$ with $g_m/\ln 2 \sim 0.43$, and C is 0.623 for the branch connected to $h_0 = 0$ and 0.490 for the branch connected to the maximum. Note that these values should change continuously to $2/3$ and $1/2$, respectively, for $\alpha \to 0$ according to the results in subsection 4.1. Note $C = 1/2$ implies that $\rho_i = 0$, 1 appear with the same probability, while $C = 2/3$ implies $\rho_i = 0$, 1 appear with probability $1/3$ and $2/3$, respectively. On the other hand, these values change continuously with each other in the large α region. This property is similar to the gas–liquid phase transition.

We have done some preliminary studies of local minimum states for $r = 0$ by simulations. Taking $N \sim 50$ and $P \sim 5$, local minimum states were generated randomly by gradient dynamics for various h_0. The regions which give the $C \neq 0$ local minimum states are found

to be limited by $h_0 \leqslant \alpha$ in agreement with the mean-field results. However, the averages of C are not in good agreement except for $h_0 \sim \alpha/2$. We suspect that a random search is too naive to study biased local minimum states and an exhaustive search will be needed to study the averages of C especially for $h_0 \sim 0$ and h_b.

4.4. g_m on the maximum points $h_0 = h_m$

To conclude this section, we study the behaviour of g_m along the maximum points $h_0 = h_m$ as $\alpha \to 0$. The relation (17) simplifies the saddle point equations greatly, giving

$$A = -\frac{1}{z}\frac{\bar{B}}{2\bar{A}\sqrt{2\bar{A}}}\Phi'(w(-r))\ \exp(r^2\bar{C}) \tag{19}$$

$$B = \frac{1}{z}\frac{1}{\sqrt{2\bar{A}}}\Phi'(w(-r))\ \exp(r^2\bar{C}) \tag{20}$$

which yields

$$A = -\frac{\bar{B}}{2\bar{A}}B \tag{21}$$

as in the case $h_0 = 0$. With the asymptotic relations (10), we have $B \sim 1$ for $\alpha \to 0$, which gives

$$2\bar{A} \sim \frac{1}{8\pi}\exp\left\{-\frac{(r\bar{B}+h_0)^2}{2\bar{A}}\right\}. \tag{22}$$

Using again $\bar{A} \sim -\alpha/2A$, $\bar{B} \sim \alpha$, and setting $h_0 = h_m \sim (1-2r)\alpha/2$, we have

$$|A| \sim 8\pi\alpha\exp\left\{\frac{\alpha|A|}{4}\right\}. \tag{23}$$

By this equation, we obtain $|A| \sim (8/\alpha)\ln(1/\alpha)$ asymptotically. Putting these results in g, we obtain

$$g_m \sim \ln 2 - \frac{1}{2}\alpha\ln\left\{4 + C\frac{8}{\alpha}\ln\frac{1}{\alpha}\right\} + \frac{1}{2}\alpha. \tag{24}$$

This result is similar to the Ising spin case without external fields, which is recovered by setting $C = 1/4$. Note that for $\alpha = 0$, C only varies from $1/4$ to $1/2$. Thus, for binary spin variables, $g_m \to \ln 2$ in a similar manner as in the Ising case if h_0 is set to $h_m \sim (1-2r)\,\alpha/2$.

Let us briefly comment on the situation $h_0 \neq h_m$ but close to h_m. Figure 3 implies that g_m also tends to ln 2 even in such a situation. With $h_0 \neq h_m$, one of the two terms in A and B mainly contributes for $\alpha \to 0$. Since the contributing term in A and B is determined by the common factor $\Phi'(w(\eta))$, we can write an equation similar to (21) if h_0 is scaled as α. This will allow us to make calculations similar to the ones presented above. In particular, both the terms in z will tend to 1, giving $g_m \sim \ln 2$ for $\alpha \to 0$.

5. Discussion

As discussed in sections 1 and 2, the Hopfield model with opposite interactional sign can be viewed as a network made of many different kinds of molecules, which are characterized by P kinds of charges. We call this model the multi-charge network model.

In this paper, to evaluate the memory effect of the MCN model, we have studied the number of local minimum states in the form of $\exp(Ng_m)$ with binary spin variables $\rho_i = 0, 1$ and external fields h_0. We have introduced spin variables $\eta_i = \rho_i - r$, which correspond to the Ising spins for $r = 1/2$ and pure binary variables for $r = 0$. When $h_0 = 0$ and $0 < r \leqslant 1/2$, $g_m \to \ln 2$ for $\alpha \to 0$, while when $h_0 = 0$ and $r = 0$, $g_m \to \ln(3/2)$ for $\alpha \to 0$ as is

shown in figure 1. However, finite h_0 changes the situation drastically. In particular, on the line $h_0 = h_m \sim (1-2r)\,\alpha/2$, g_m behaves similarly to the Ising spin case. This property will hold for $0 < h_0 < h_b \sim (1-r)\,\alpha$ and h_0 not close to both ends. We should note that $g_m = \ln 2$ does not mean that all configurations become local minimum states. There should be some corrections for g_m which disappear in the thermodynamic limit.

Let us give some remarks on the local minimum energies. Although the following argument will be possible for $r > 0$, we concentrate on the case $r = 0$ for the sake of simplicity. For $r = 0$, we first notice that the energy function (3) implies $H = 0$ for $\eta_i = 0$. For the non-trivial configuration, which is $\sum_i \eta_i > 0$, we expect that $a_\mu = \sum_i \xi_i^\mu \eta_i/\sqrt{N}$ can be very small for small α. This implies that the absolute minimum energy is approximately given by the linear terms of η_i. Thus, the energies of the non-trivial local minimum states are higher than zero for $h_0 > \alpha/2$, while some of them will become negative for $h_0 < \alpha/2$. Accordingly, the absolute minimum state is given by $\eta_i = 0$ from large h_0 down to $h_0 \sim \alpha/2$, below which it is given by non-trivial η_i. These observations suggest that the properties in finite temperature can be quite different between the two sides of $h_0 = \alpha/2$, which is not seen in the behaviour of g_m.

The neutralization effect will be closely related to the ability of recognition. This point will be clearly seen by studying the responses to external systems. Let us imagine that the system consists of two parts; a system A made of N spins and an external system B made of M spins, where spins in system B are assumed to be fixed and external fields are imposed on system A. Here h_0 is set to $(1-2r)\,\alpha/2$ for simplicity. For the whole system, we have the energy function given by

$$\begin{aligned} H_X &= \frac{1}{2N}\sum_\mu \left(\sum_{i\in A}\xi_i^\mu \eta_i + X^\mu\right)^2 \\ &= H + \sum_{i\in A} h_{Bi}\eta_i + \frac{1}{2N}\sum_\mu (X^\mu)^2 \end{aligned}$$

where H is the energy function of system A and $h_{Bi} \equiv \sum_\mu X^\mu \xi_i^\mu/N$. We should note that these external fields cannot be replaced by some random external fields since h_{Bi}, h_{Bj} and J_{ij} are correlated just like the correlations among J_{ij}. The first line of H_X explicitly shows that system A will recognize charges X^μ by minimizing H_X since low-energy states will be given by $\sum_{i\in A}\xi_i^\mu\eta_i \sim -X^\mu$ for all μ, implying the neutralization of the whole system. In this way, system A recognizes system B by adjusting η_i. This set of equations is similar to the perceptron problem [15] if η_i are identified with synaptic couplings. Note that, unlike the perceptron problem, the energy function of the MCN model has a simple physical meaning of networks. This set of equations can also be regarded as a replication of charges $-X^\mu$ by the charges of system A. In short, the MCN model perceives external systems by the replication of complementary shapes. In these arguments, we have assumed that H_X/N can be very close to zero by some gradient dynamics, which remains to be studied.

In our arguments, a set of quenched variables ξ_i^μ plays a very important role. We may ask if ξ_i^μ really reflects the properties of protein molecules in reality. To answer this question, we need to start from the statistical sum of molecules which have complicated shapes. This will clarify the meaning of ξ_i^μ as well as h_0, if they really have physical origins. In this respect, we should note that the interactions among protein molecules should be characterized by 'site randomness' instead of 'bond randomness'. In other words, interactions should be expressed by quenched variables defined for each kind of molecule. The MCN model may be one of the simple realizations of this idea.

Appendix

This appendix is devoted to the description of the derivation of (7) from (6). Introducing integral representations for the delta functions, (6) is expressed as

$$G = \sum_{\{\eta\}} \prod_i \int_0^\infty \int_{-i\infty}^{i\infty} \exp \phi_i \left(\eta_i' \left(\sum_i J_{ij} \eta_j - h_0 \right) - h_i \right) \frac{\mathrm{d}\phi_i \, \mathrm{d}h_i}{2\pi \mathrm{i}}. \tag{A.1}$$

In this expression, η_i' are variables which are positive for $\rho_i = 1$ and negative for $\rho_i = 0$. We set $\eta_i' = \eta_i$ for simplicity. In the exponential, summation over i and j gives

$$\sum_i \phi_i \left(\eta_i \sum_j J_{ij} \eta_j \right) = -\sum_\mu a_\mu b_\mu + \alpha \sum_i \phi_i \eta_i^2 \tag{A.2}$$

where $a_\mu = \sum_i \xi_i^\mu \eta_i / \sqrt{N}$ and $b_\mu = \sum_i \xi_i^\mu \phi_i \eta_i / \sqrt{N}$. For each μ, we introduce Gaussian variables x_μ and y_μ and write

$$\exp(-ab) = \int\int \exp \left\{ -\frac{1}{2}(x^2 + y^2) + x \frac{a-b}{\sqrt{2}} + \mathrm{i}y \frac{a+b}{\sqrt{2}} \right\} \frac{\mathrm{d}x \, \mathrm{d}y}{2\pi}$$

where index μ is dropped for simplicity. Introducing $t = (x + \mathrm{i}y)/\sqrt{2}$ and $\bar{t} = (x - \mathrm{i}y)/\sqrt{2}$, and after ξ_i^μ averages, we obtain

$$\langle \exp(-ab) \rangle = \int \exp \left\{ -t\bar{t} + \frac{1}{2}(At^2 - 2Bt\bar{t} + C\bar{t}^2) \right\} \frac{\mathrm{d}x \, \mathrm{d}y}{2\pi}$$

where $A = \sum_i \phi_i^2 \eta_i^2 / N$, $B = \sum_i \phi_i \eta_i^2 / N$ and $C = \sum_i \eta_i^2 / N$. After integrating over x_μ and y_μ, we obtain

$$\langle G \rangle = \sum_{\{\eta\}} \int_0^\infty \int_{-i\infty}^{i\infty} \exp \left(-\frac{1}{2} P \ln\{(1+B)^2 - AC\} + PB - h_0 \sum_i \phi_i \eta_i - \sum_i \phi_i h_i \right) \times \prod_i \frac{\mathrm{d}\phi_i \mathrm{d}h_i}{2\pi \mathrm{i}}.$$

Then, by expressing 1 by the delta function

$$\int \delta \left(\sum_i \phi_i^2 \eta_i^2 - NA \right) N \, \mathrm{d}A = \int\int \exp \bar{A} \left(\sum_i \phi_i^2 \eta_i^2 - NA \right) \frac{N \, \mathrm{d}A \, \mathrm{d}\bar{A}}{2\pi \mathrm{i}}$$

and writing similar equations for B, $\bar{B}$ and C, $\bar{C}$, we obtain a one site problem for ϕ_i and η_i. Replacing ϕ_i by $\mathrm{i}\phi_i$ and summing over $\eta_i = -r, 1 - r$, we obtain

$$g = -\tfrac{1}{2}\alpha \ln\{(1+B)^2 - AC\} + \alpha B - \bar{A}A - \bar{B}B - \bar{C}C + \ln z \tag{A.3}$$

where

$$z = \sum_{\eta = -r, 1-r} \Phi(w(\eta)) \exp(\eta^2 \bar{C})$$

with

$$\Phi(x) = \int_{-\infty}^{x} \exp\left(-\tfrac{1}{2}t^2\right) \frac{\mathrm{d}t}{\sqrt{2\pi}}$$

and

$$w(\eta) = \frac{\eta^2 \bar{B} - \eta h_0}{\sqrt{2\bar{A}\eta^2}}$$

where irrelevant constants are dropped in (A.3). This completes the derivation of the mean-field approximation for $\langle G \rangle$.

References

[1] Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1792
[2] Mézard M, Parisi G and Virasioro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[3] Hopfield J J 1982 *Proc. Natl. Acad. Sci. USA* **79** 2554
[4] Nokura K 1998 *J. Phys. A: Math. Gen.* **31** 7447
[5] Nokura K 2001 *J. Phys. A: Math. Gen.* **34** 1005
[6] Crick F and Mitchison G 1983 *Nature* **304** 111
[7] Hopfield J J, Feinstein D I and Palmer R G 1983 *Nature* **304** 158
[8] Jerne N K 1974 *Ann. Immunol. (Paris)* C **125** 1127
[9] Parisi G 1990 *Proc. Natl Acad. Sci. USA* **87** 429
[10] Lefèvre O and Parisi G 1993 *Network* **4** 39
[11] Marinari M, Parisi G and Ritort F 1994 *J. Phys. A: Math. Gen.* **27** 7615
[12] Marinari M, Parisi G and Ritort F 1994 *J. Phys. A: Math. Gen.* **27** 7647
[13] Parisi G and Potters M 1995 *J. Phys. A: Math. Gen.* **28** 5267
[14] Tanaka F and Edwards S F 1980 *J. Phys. F: Met. Phys.* **10** 2769
[15] Gardner E 1988 *J. Phys. A: Math. Gen.* **21** 257