PHYSICAL REVIEW E VOLUME 51, NUMBER 4 APRIL 1995

Finite-dimensional neural networks storing structured patterns

Hidetoshi Nishimori,[1,2] W. Whyte,[2] and D. Sherrington[2]
[1] *Department of Physics, Tokyo Institute of Technology, Oh-okayama, Meguro-ku, Tokyo 152, Japan*
[2] *Department of Physics, University of Oxford, Theoretical Physics, 1 Keble Road, Oxford OX1 3NP, United Kingdom*
(Received 21 November 1994)

We investigate thermodynamic properties of neural networks defined on a finite-dimensional lattice designed to store and retrieve patterns with structure. Our aim is to draw phase diagrams with axes of temperature and a parameter controlling the structure of patterns. Gauge symmetry is used to derive various exact or rigorous results on the properties of the system. These results put strong constraints on the possible phase diagrams. We also use Peierls arguments to prove the existence of a ferromagnetic phase and of a phase with finite overlap order in certain regions of the phase diagram. Our conclusion on the phase diagram is that, first, if the number of embedded patterns is smaller than a critical value, the system has in general three phases: a paramagnetic phase, a retrieval phase, and a ferromagnetic phase accompanied by finite overlap order. For larger numbers of embedded patterns, a ferromagnetic phase without overlap order appears in addition. The retrieval phase without ferromagnetic order may be replaced by a spin glass phase for large numbers of embedded patterns.

PACS number(s): 87.10.+e, 05.50.+q

I. INTRODUCTION

Neural networks represent a rich many-body system with a number of interesting features associated with a multiplicity of nonequivalent dynamical attractors, some related fairly straightforwardly to stored global microstates, others quasirandom, and with phase transitions between attractor types. Those with detailed balance yield asymptotic Boltzmann-like microstate distributions and thus have been a target of statistical-mechanical analysis for almost a decade [1–4]. An important prototypical model is the Hopfield model [1] whose features include binary-state neurons, represented by Ising spins, interacting through infinite-range synapses, equivalent to exchange interactions in the Ising formulation. The exchange interactions code the effect of embedded random patterns via the Hebb rule $J_{ij} = N^{-1}\sum_\mu \xi_i^\mu \xi_j^\mu$, where the i label the neurons, the μ the patterns, and ξ_j^μ take the quenched values ± 1. These features enabled detailed statistical-mechanical analysis of the system [2–4].

It is sometimes useful and interesting to modify or generalize some of the special characteristics of the Hopfield model. Although there are many possibilities of generalization (see, for example [3]), let us consider in particular the effects of nonrandomness in patterns embedded in the network. Schlüter and Wagner [5] considered the problem of embedding patterns with spatial structure generated stochastically according to the Boltzmann factor of the nearest-neighbor Ising model. They used the conventional Hebb rule for the synaptic efficacies, an infinite-range interaction in the space of the network itself, and calculated the storage capacity. They also proposed a modification of the Hebb rule to increase the capacity. Monasson [6] analyzed the same problem in the limit of weak spatial correlations. The dynamics of networks storing patterns with structure has been discussed by Coolen [7] for certain special cases. The issue of learning such structured patterns in perceptron-type networks has also been discussed by Monasson [6] and by Tarkowski and Lewenstein [8].

It is also of interest to consider the finiteness of spatial dimensionality not only of the patterns but also of the network itself, because spatial correlations inherent in patterns with structure make sense mainly in finite-dimensional systems with finite-range interactions. We are thus led to consider the problem of patterns, corresponding to snapshots of Monte Carlo simulations of a finite-dimensional Ising model with short-range interactions at finite temperature, embedded in a finite-range Ising model with the same range of interactions on the same finite-dimensional lattice.

The aim of the present paper is to provide statistical-mechanical treatments of such finite-dimensional neural networks embedded with structured patterns. We classify possible thermodynamic phases and draw the phase diagram as precisely as possible using rigorous analytical methods.

This paper is organized as follows. We give the precise definition of the system in Sec. II. General expectations on its behavior are then discussed, such as the phase diagram and the possible number of patterns to be successfully stored in such a system. A few limiting cases are investigated in some detail. Section III describes exact solutions and rigorous inequalities on various thermodynamic quantities. The results include exact expressions for the internal energy and the overlap order parameter on a special line in the phase diagram. Gauge symmetry is shown to play crucial roles in the calculations. The existence of ordered phases is proved in Sec. IV. The Peierls argument, modified appropriately to take account of quenched randomness, is used to show the existence of the ferromagnetic and finite-overlap states in certain regions of the phase diagram. We analyze the limit of infinite-range interactions in Sec. V. This infinite-range model corresponds to a generalization of the conventional Hopfield model to the case with an additional parameter controlling the randomness of patterns. The final section, Sec. VI, is devoted to discussions on possible phase diagrams with the results in previous sections taken into account.

II. PRELIMINARY ANALYSIS

A. Definition of the system

The Hamiltonian to be treated in this paper is

$$H = -\frac{1}{\sqrt{p}} \sum_{\langle ij \rangle} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu S_i S_j, \quad (1)$$

where the S_i, ξ_i^μ are variables of the Ising type taking values ± 1, the S_i being the usual variables, the $\xi_i^\mu = \pm 1$ quenched randomly according to a probability distribution specified below. The summation $\langle ij \rangle$ in (1) extends over nearest neighbors on hypercubic lattices, although the results of Sec. III are valid without this restriction. The number of patterns p is taken to be finite throughout this paper.

Let us follow Schlüter and Wagner [5] in choosing the probability distribution of the quenched random variables

$$P(\{\xi_i^\mu\}) = c \prod_{\mu=1}^{p} \exp\left(\frac{J_0}{\sqrt{p}} \sum_{\langle ij \rangle} \xi_i^\mu \xi_j^\mu \right). \quad (2)$$

The parameter J_0 controls the degree of order (structure) within each pattern, but there is no explicit correlation between patterns. The normalization factor c is given by $\{Z_0(J_0/\sqrt{p})\}^{-p}$, where Z_0 is the partition function of the ferromagnetic Ising model defined on the same lattice as in (1):

$$Z_0\left(\frac{J_0}{\sqrt{p}}\right) = \sum_{s} \exp\left(\frac{J_0}{\sqrt{p}} \sum_{\langle ij \rangle} S_i S_j \right).$$

Equations (1) and (2) may be interpreted that patterns are taken from snapshots of equilibrium Monte Carlo simulations of the ferromagnetic Ising model with nearest-neighbor exchange interactions and the inverse temperature in the ratio $J_0/\sqrt{p}$, and those patterns are embedded in the network following the "short-range" Hebb rule

$$J_{ij} = \frac{1}{\sqrt{p}} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \quad (3)$$

for the nearest-neighbor pair $\langle ij \rangle$ and $J_{ij} = 0$ otherwise. The factor $1/\sqrt{p}$ in (1) and (2) is useful in considering the limit $p \to \infty$, as will be explained later.

B. General expectations on the phase diagram

Embedded patterns are completely random if $J_0 = 0$. This case can be considered a short-range version of the Hopfield model. Such a system may be able to store and retrieve embedded patterns if their number is not too large and if the level of stochastic agitation, the temperature in the usual statistical-mechanical sense, is not too high; when $J_0 = 0$, a retrieval of pattern μ is considered to occur if the overlap $m_\mu = N^{-1} \sum_i \xi_i^\mu \langle S_i \rangle$, where $\langle\,\rangle$ represents a thermodynamic average, is of $O(1)$ for only one μ [9]. For sufficiently large p, the retrieval phase will not exist and a spin glass phase may instead appear. This conjecture comes partly from experience in the infinite-range Hopfield model in which the system cannot retrieve embedded patterns if their number per spin $\alpha = p/N$ exceeds a threshold α_c [2–4]. In the present finite-dimensional model, the critical number of patterns to be successfully stored is likely to be finite, p_c, rather than being proportional to the number of neurons N. This finiteness of p_c is anticipated from a consideration of networks with randomly diluted synapses, in which p_c is proportional to the average connectivity per neuron, a finite number [10]. The following argument also predicts the finiteness of p_c.

Naively, the capacity p_c is expected to be much lower than in the infinite-range counterpart; the total number of exchange interactions, where information is stored, is proportional to N (the number of neurons or spins) as compared to N^2 in the case of the infinite-range model, while each pattern requires $O(N)$ bits of information. More specifically, suppose that the number of possible values of a given exchange interaction J_{ij} is a finite number a. Then the possible number of configurations of bond values connected to a given single site is bounded from above by a^z, where z is the coordination number and hence the maximum information content per site is $z\ln a$. This shows that the amount of information to be stored and retrieved is proportional to the coordination number, which is finite in finite-dimensional lattices. For the Hebb rule under discussion, a increases with p but is bounded by $2p$, leaving intact the conclusion that p_c is finite for finite z. Therefore the maximum number of patterns to be successfully stored p_c will be finite.

In the opposite limit of $J_0 \to \infty$, fluctuations in the pattern randomness are completely suppressed, the ξ_i^μ becoming all $+1$ or all-1, and the exchange interaction J_{ij} of (3) is simply given as $\sqrt{p}$. The network system is then a pure ferromagnetic Ising model and there exists a ferromagnetic phase at low temperatures if the spatial dimensionality is larger than unity.

The phase diagram is therefore expected to be given generally as in Fig. 1, categorizing the phases into paramagnetic, ferromagnetic, and retrieval (or spin glass, depending upon whether p is less or greater than p_c). There are several other possibilities within these simple categorizations, such as the mixed phase (a ferromagnetic phase with spin glass characteristics), absence of the retrieval or spin glass phase near $J_0 = 0$ (which may be the case in low dimensions), and a ferromagnetic phase with finite

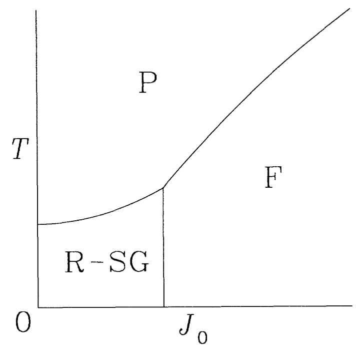

FIG. 1. Generic qualitative phase diagram of our network embedding patterns with structures. There are in general three phases: paramagnetic, ferromagnetic, and retrieval (or spin glass) phases. There are possibilities of a mixed phase state and of a finite-overlap state within the ferromagnetic phase.

values of the overlap order parameter. The last possibility will be discussed in detail in the following sections.

C. Large-p limit

In the limit of large p, the present system is likely to reduce to the Edwards-Anderson model of spin glasses [11] with a Gaussian distribution of exchange interactions. The first step of the argument is to note that the exchange interaction (3) is composed of independent terms, the number of which grows with p. The central limit theorem thus applies and the distribution of J_{ij} approaches Gaussian. The average and variance are easily calculated and the results are $[J_{ij}] = \sqrt{p}c_1(J_0/\sqrt{p})$ and $[(\Delta J_{ij})^2] = 1-c_1(J_0/\sqrt{p})^2$, where the square brackets denote the average over the distribution (2) and $c_1(J_0/\sqrt{p})$ represents the correlation function between nearest-neighbor sites of the ferromagnetic Ising model with (2) as the Boltzmann factor. In the limit $p \to \infty$, only the leading term of the high-temperature expansion is important in evaluating $c_1(J_0/\sqrt{p})$ because $J_0/\sqrt{p} \ll 1$. Thus we have $[J_{ij}] \approx \sqrt{p}\tanh(J_0/\sqrt{p}) \approx J_0$ and $[(\Delta J_{ij})^2] \approx 1$. This means that J_{ij} becomes a Gaussian variable with mean J_0 and variance unity.

It is further necessary to show the independence of J_{ij} from other exchange interactions to complete the argument for reduction to the spin glass model of Edwards and Anderson. We have obtained evidence for such a reduction by showing that the covariance of neighboring interactions vanishes in the large-p limit. For example, the covariance of exchange interactions in Fig. 2(a) is found to be

$$[\Delta J_{12}\Delta J_{23}] = c_2\left(\frac{J_0}{\sqrt{p}}\right) - c_1\left(\frac{J_0}{\sqrt{p}}\right)^2 \to 0$$

as $p \to \infty$, where c_2 is the correlation between sites 1 and 3 in Fig. 2(a). Similarly, the three-body covariance of the closed-loop configuration in Fig. 2(b) (which is present on the triangular lattice, for example) is given as

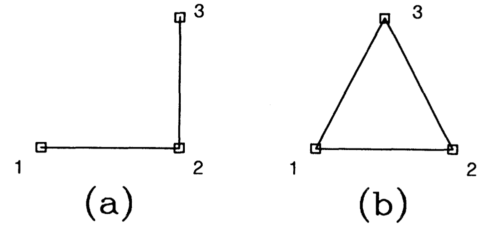

FIG. 2. Configurations of (a) two exchange interactions J_{12} and J_{23} and (b) a closed loop for which the covariances are calculated.

$$[\Delta J_{12}\Delta J_{23}\Delta J_{31}] = \frac{1}{\sqrt{p}} - \frac{3}{\sqrt{p}}c_1\left(\frac{J_0}{\sqrt{p}}\right)^2 + \frac{2}{\sqrt{p}}c_1\left(\frac{J_0}{\sqrt{p}}\right)^3 \to 0.$$

These results indicate the independence of J_{ij} in the large-p limit.

D. Two patterns embedded

The special case of $p = 2$ with $J_0 = 0$ is equivalent to the site-diluted ferromagnetic Ising model with dilution probability $\frac{1}{2}$. The proof consists of a simple gauge transformation

$$Z = \sum_s \exp\left\{\frac{\beta}{\sqrt{2}}\sum_{\langle ij\rangle}(\xi_i^1\xi_j^1 + \xi_i^2\xi_j^2)S_iS_j\right\} = \sum_s \exp\left\{\frac{\beta}{\sqrt{2}}\sum_{\langle ij\rangle}(1+\eta_i\eta_j)S_iS_j\right\}, \quad (4)$$

where β is the inverse temperature and $\eta_i = \xi_i^1\xi_i^2$. Since η_i is ± 1, some of the interactions $(1+\eta_i\eta_j)/\sqrt{2}$ are vanishing in (4) while others have the strength $\sqrt{2}$. If we classify the sites into two groups, one (group A) with $\eta_i = 1$ and the other (group B) with $\eta_i = -1$, then there are no interactions between A sites and B sites. Interactions within a group are uniform and ferromagnetic. Hence spins on B sites can be completely ignored for the spin summation over A sites and vice versa,

$$Z = Z_A Z_B.$$

The free energy averaged over the quenched distribution (2) with $J_0 = 0$ is then written as

$$-\beta F = [\ln Z] = [\ln Z_A] + [\ln Z_B] = 2[\ln Z_A].$$

The last equation stems from equivalence of A and B groups on average. It should now be clear that $[\ln Z_A]$ is the free energy of the site-diluted ferromagnetic Ising model with dilution probability $\frac{1}{2}$.

The overlap order parameter in the original representation of variables $m_1 = N^{-1}\sum_i\langle\xi_i^1 S_i\rangle$ is gauge transformed to the spontaneous magnetization of the site-diluted ferromagnet. Hence, if the dilution model has a ferromagnetic phase at low temperatures, the original system has a retrieval phase for the same parameters. Critical concentrations of the site-dilution problem are known from series expansions. In two dimensions, they are either equal to $\frac{1}{2}$ (triangular lattice) or larger than $\frac{1}{2}$ (other lattices) [12]. In three dimensions, any regular lattice has a critical concentration lower than $\frac{1}{2}$ [13]. Therefore we conclude that two-dimensional systems do not have a retrieval phase, while the three-dimensional (and probably higher dimensional) counterparts have retrieval phases at finite temperatures when $J_0 = 0$.

The existence of a retrieval phase in three (and higher) dimensions will hold also for $J_0 \approx 0$ because expansions of thermodynamic quantities around $J_0 = 0$ resemble high-temperature expansions, as seen from the form of the probability distribution (2). High-temperature expansions generally have finite ranges of convergence, implying a smooth continuation of properties of the $J_0 = 0$ case to finite (but small) J_0 situations.

III. EXACT RESULTS BY THE METHOD OF GAUGE TRANSFORMATION

We now turn to symmetry arguments leading to a number of exact or rigorous results on thermodynamic quantities of the present system. The essential idea is the same as that for the spin glass problem [14]. All the results in this section hold for any lattice with arbitrary interaction ranges.

A. Internal energy

The calculation of the internal energy provides a good example of the method. We therefore describe this in some detail. The internal energy, or the thermal average of the Hamiltonian, for a given $\{\xi_i^\mu\}$ is expressed as

$$\langle E\rangle = -\frac{\partial}{\partial\beta} Z\left(\frac{\beta}{\sqrt{p}}\right) \Big/ Z\left(\frac{\beta}{\sqrt{p}}\right), \qquad (5)$$

where Z is the partition function

$$Z\left(\frac{\beta}{\sqrt{p}}\right) = \sum_s \exp\left(\frac{\beta}{\sqrt{p}}\sum_{\langle ij\rangle}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu S_i S_j\right). \qquad (6)$$

Its configurational average is given by

$$[\langle E\rangle] = c\sum_\xi \exp\left(\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle}\sum_\mu \xi_i^\mu \xi_j^\mu\right)\langle E\rangle.$$

We consider the application to this expression of a gauge transformation

$$S_i \to S_i\sigma_i, \quad \xi_i^\mu \to \xi_i^\mu \sigma_i, \qquad (7)$$

where σ_i is arbitrarily fixed to 1 or -1 at each site. The Hamiltonian (1), the partition function (6), and the thermal average (5) are all invariant under the gauge transformation (7). The average internal energy then acquires a new expression

$$[\langle E\rangle] = c\sum_\xi \exp\left(\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle}\sum_\mu \xi_i^\mu \xi_j^\mu \sigma_i\sigma_j\right)\langle E\rangle,$$

which is independent of $\{\sigma_i\}$. Summation over $\{\sigma_i\}$ and division by 2^N does not change the value, so

$$[\langle E\rangle] = -\frac{c}{2^N}\sum_\xi Z\left(\frac{J_0}{\sqrt{p}}\right)\frac{\partial}{\partial\beta} Z\left(\frac{\beta}{\sqrt{p}}\right) \Big/ Z\left(\frac{\beta}{\sqrt{p}}\right). \qquad (8)$$

If J_0 is equal to β, this expression simplifies considerably,

$$\begin{aligned}[\langle E\rangle] &= -\frac{c}{2^N}\sum_\xi \frac{\partial}{\partial\beta} Z\left(\frac{\beta}{\sqrt{p}}\right)\\ &= -\frac{c}{2^N}\frac{\partial}{\partial\beta}\sum_s\sum_\xi \exp\left(\frac{\beta}{\sqrt{p}}\sum_{\langle ij\rangle}\sum_\mu \xi_i^\mu\xi_j^\mu S_i S_j\right)\\ &= -\frac{c}{2^N}\frac{\partial}{\partial\beta}2^N\left\{Z_0\left(\frac{\beta}{\sqrt{p}}\right)\right\}^p\\ &= pE_0\left(\frac{\beta}{\sqrt{p}}\right),\end{aligned} \qquad (9)$$

where $E_0(\beta/\sqrt{p})$ is the internal energy of the ferromagnetic Ising model corresponding to the partition function $Z_0(\beta/\sqrt{p})$. This expression is exact and applies to any lattice and any interaction range $\langle ij\rangle$, including the infinite-range model.

The internal energy of the ferromagnetic Ising model in (9) generally has a singularity at some critical point when the spatial dimensionality exceeds one. Let us suppose that $E_0(K)$ is singular at $K = K_c$. Then (9) means that the average internal energy of our neural network has the same singularity at $\beta/\sqrt{p} = K_c$ or $J_0 = \beta = \sqrt{p}K_c$ (Fig. 3). This of course implies a phase transition at this point. The line $J_0 = \beta$ should cross a phase boundary

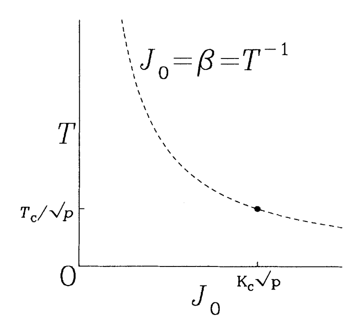

FIG. 3. The line $J_0 = \beta$ in the phase diagram (shown dashed) on which the internal energy is obtained exactly. There is a singularity in the internal energy at the point $J_0 = \beta = \sqrt{p}K_c$ indicated on the line, implying a phase transition.

there. One may often expect a multicritical point there, as discussed below.

There is a geometrical reason for the existence of this singularity. The embedded patterns $\{\xi_i^\mu\}$ undergo a sudden change of state at $J_0 = \sqrt{p}K_c$ because the patterns are generated according to the Boltzmann factor (2). More explicitly, $[\xi_i^\mu] \neq 0$ for $J_0 > \sqrt{p}K_c$ and $[\xi_i^\mu] = 0$ if $J_0 \leq \sqrt{p}K_c$ under appropriate boundary conditions (all boundary variables are $+1$, for example). The patterns develop long-range order for large J_0. Thus the singularity in the average internal energy at $J_0 = \sqrt{p}K_c$ is caused by this sudden change of embedded patterns.

The absence of other singularities in $E_0(K)$ does not mean that the point $J_0 = \sqrt{p}K_c$ is the only location where the line $J_0 = \beta$ goes across phase boundaries. It is in fact explicitly proved in Sec. IV that a boundary between ferromagnetic and nonferromagnetic phases is located somewhere in the region $J_0 < \sqrt{p}K_c$ if p is large. However, this phase transition at $J_0 < \sqrt{p}K_c$ does not show up as a singularity in the average internal energy. The same situation was encountered in the spin glass problem [14]. It may be useful to recall here that the present model reduces to the Edwards-Anderson model of spin glasses in the $p \to \infty$ limit, as discussed in Sec. II C.

B. Specific heat

It is not possible to calculate the specific heat

$$C = \beta^2[\langle E^2\rangle - \langle E\rangle^2] \tag{10}$$

explicitly like the internal energy. The reason is that the partition function $Z(\beta/\sqrt{p})$ appears squared in the denominator of the second term on the right-hand side of (10), which prevents the cancellation of the denominator and numerator as in (8) and (9) under the condition $J_0 = \beta$. We can nevertheless estimate an upper bound on the line $J_0 = \beta$ as

$$C \leq \beta^2 \left\{[\langle E^2\rangle] - [\langle E\rangle]^2\right\} = pC_0\left(\frac{\beta}{\sqrt{p}}\right), \tag{11}$$

where $C_0(\beta/\sqrt{p})$ is the specific heat corresponding to the internal energy $E_0(\beta/\sqrt{p})$. Equation (11) shows that the specific heat does not diverge on the line $J_0 = \beta$, except possibly at $J_0 = \beta = \sqrt{p}K_c$.

C. Correlation functions

The two-point correlation function representing the overlap of the network state with one of the embedded patterns, the first pattern, for example, is defined by

$$C_{R,r} \equiv [\langle \xi_0^1 S_0 \xi_r^1 S_r\rangle].$$

This quantity is gauge invariant and can be calculated explicitly by the same method as for the internal energy. The answer under the condition $J_0 = \beta$ is

$$C_{R,r} = c_{0,r}\left(\frac{\beta}{\sqrt{p}}\right),$$

where $c_{0,r}$ is the correlation function of the ferromagnetic Ising model. This indicates that there is a long-range correlation in the overlap ordering if $J_0 = \beta > \sqrt{p}K_c$, while there is no finite long-range overlap ordering for $J_0 = \beta \leq \sqrt{p}K_c$. It is clear from the discussion in Sec. III A that this long-range order is caused by the long-range ferromagnetic order in the embedded patterns.

The ferromagnetic correlation function $\langle S_0 S_r\rangle$ is not gauge invariant. However, the gauge transformation (7) gives a useful new expression for the configurational average of this quantity

$$C_{F,r} \equiv [\langle S_0 S_r\rangle] = [\langle S_0 S_r\rangle_{J_0}\langle S_0 S_r\rangle_\beta], \tag{12}$$

where $\langle S_0 S_r\rangle_{J_0}$ is thermal average for a given randomness $\{\xi_i^\mu\}$ with the effective coupling $J_0/\sqrt{p}$. The other quantity $\langle S_0 S_r\rangle_\beta$ has the same meaning, only with J_0 replaced by the usual β as in (6) and in the middle expression of (12) (where it is not explicitly written). Unlike gauge invariant quantities, it is not possible to derive an explicit closed expression for this correlation function even under the condition $J_0 = \beta$. However, we can develop useful arguments concerning the possible phase boundaries using the above relation (12).

Let us first consider the case of $J_0 = \beta$ in (12):

$$[\langle S_0 S_r\rangle_\beta] = [\langle S_0 S_r\rangle_\beta^2]. \tag{13}$$

In the long-distance limit $r \to \infty$, this relation yields $m_F = Q$, where m_F is the ferromagnetic order parameter and Q is the Edwards-Anderson spin glass order parameter. This readily leads to the conclusion that there is no retrieval or spin glass phase (with $m_F = 0, Q > 0$) on the line $J_0 = \beta$ (see Fig. 4).

Another restriction on the phase diagram comes from the following inequality derived from (12):

$$\begin{aligned} |[\langle S_0 S_r\rangle_\beta]| &\leq [|\langle S_0 S_r\rangle_{J_0}|\ |\langle S_0 S_r\rangle_\beta|] \\ &\leq [|\langle S_0 S_r\rangle_{J_0}|]\,. \end{aligned} \tag{14}$$

In the limit $r \to \infty$, the left-hand side reduces to m_F^2 at a given point (J_0, β) in the phase diagram and the last expression becomes an order parameter measured at (J_0, J_0) located on the line $J_0 = \beta$. Whatever the physical meaning of this latter order parameter is, this quantity should vanish if the point (J_0, J_0) is in the paramagnetic phase. The left-hand side then vanishes, implying the absence of a ferromagnetic phase at any inverse temperature β for such J_0. The situation is explained in Fig. 4, which shows various possible phase diagrams.

Both Figs. 4(a) and 4(b) are compatible with the above-mentioned constraints. However, experience in the spin glass case [15] and the mean-field analysis in Sec.V suggests that the multicritical point is in general located on the line $J_0 = \beta$ as in Fig. 4(a).

Another consequence of a gauge transformation applied to the ferromagnetic correlation function is a proof of absence of ferromagnetic order in a finite region near $J_0 = 0$. Following Horiguchi and Morita [16], we can

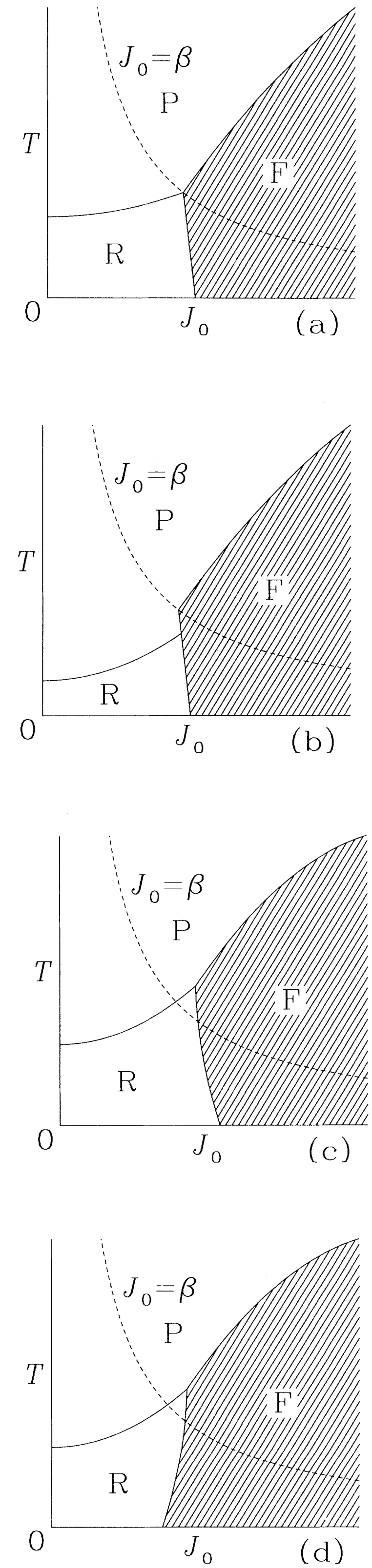

FIG. 4. (a) and (b) Phase diagrams compatible with the constraints derived from gauge transformations. The topology of (c) is forbidden by (13) and (d) is also forbidden by (14).

derive a slightly different version of (12) under boundary conditions of all spins up,

$$[\langle S_0\rangle_\beta] = [\langle S_0\rangle_{J_0}\langle S_0\rangle_\beta]\,.$$

The right-hand side is bounded from above by the spontaneous magnetization of the ferromagnetic Ising model corresponding to $Z_0(\sqrt{p}J_0)$ [16]:

$$[\langle S_0\rangle_\beta] \le m_F(\sqrt{p}J_0)m_F(\sqrt{p}\beta).$$

Therefore there is no spontaneous magnetization in the original model if $\sqrt{p}J_0 \le K_c$.

D. Dynamical correlation functions

Ozeki [17] recently applied the method of gauge transformation to dynamical correlation functions of the Ising spin glass. We show here that his idea is useful also in our neural network. Calculations will be shown in some detail for completeness.

His argument starts with the observation of gauge invariance of the transition probability appearing in the master equation for the microstate distribution at time t, $\rho_t(\mathbf{S})$,

$$\frac{d\rho_t(\mathbf{S})}{dt} = \sum_{\mathbf{S}'} W_{\mathbf{SS}'}\rho_t(\mathbf{S}'),$$

where $\mathbf{S}$ and $\mathbf{S}'$ represent sets of spin states. To show the gauge invariance of $W_{\mathbf{SS}'}$, we write this quantity as

$$W_{\mathbf{SS}'} = \frac{w_{\mathbf{SS}'}}{\rho_e(\mathbf{S}')} - \delta_{\mathbf{SS}'}\sum_{\mathbf{S}''}\frac{w_{\mathbf{S}''\mathbf{S}}}{\rho_e(\mathbf{S})}, \qquad (15)$$

where ρ_e is the normalized equilibrium distribution (the Boltzmann factor divided by the partition function) with the effective coupling $\beta/\sqrt{p}$ as employed in (6). The explicit form of $w_{\mathbf{SS}'}$ depends upon the type of dynamics. Glauber dynamics for a system controlled by the interaction (3) at temperature β^{-1} has

$$w_{\mathbf{SS}'} = w_0\delta_1(\mathbf{S},\mathbf{S}')\frac{\sqrt{\rho_e(\mathbf{S})\rho_e(\mathbf{S}')}}{\cosh\left(\beta\Delta(\mathbf{S},\mathbf{S}')/2\right)}, \qquad (16)$$

whereas Metropolis dynamics is defined by

$$w_{\mathbf{SS}'} = w_0\delta_1(\mathbf{S},\mathbf{S}')\rho_e(\mathbf{S})^{\theta(\Delta[\mathbf{S},\mathbf{S}'])}\rho_e(\mathbf{S}')^{\theta(\Delta[\mathbf{S}',\mathbf{S}])}. \quad (17)$$

Here w_0 is a constant, δ_1 is a single-spin flip operator

$$\delta_1(\mathbf{S},\mathbf{S}') = \delta\left(1, \frac{1}{2}\sum_{i=1}^{N}(1 - S_iS_i')\right),$$

$\Delta[\mathbf{S},\mathbf{S}']$ is the energy change $H(\mathbf{S}) - H(\mathbf{S}')$, and $\theta(x)$ represents the step function $\theta(x) = 1$ if $x > 0$ and 0 otherwise. It is clear from (15)–(17) that the transition probability $W_{\mathbf{SS}'}$ is invariant under

$$S_i \to S_i\sigma_i, \quad S_i' \to S_i'\sigma_i, \quad \xi_i^\mu \to \xi_i^\mu\sigma_i. \qquad (18)$$

With this gauge invariance in mind, we rewrite the autocorrelation function and uniform magnetization to prove the equivalence of these two quantities under certain initial conditions.

The autocorrelation function of the system having an initial condition of being at equilibrium with effective coupling $J_0/\sqrt{p}$ but evolving with effective coupling $\beta/\sqrt{p}$ is written as

$$\left[\langle S_i(t)S_i(0)\rangle_\beta^{J_0}\right] = \left[\sum_{S,S'} S_i \left(e^{tW}\right)_{S,S'} \rho_e(\mathbf{S}', J_0/\sqrt{p}) S_i'\right], \tag{19}$$

where the transition probability W is understood to have the effective coupling $\beta/\sqrt{p}$. The expression inside the square brackets on the right-hand side of (19) is gauge invariant. We apply the gauge transformation (18), sum over the gauge variables $\{\sigma_i\}$, and divide the result by 2^N. By using the explicit form of the normalized equilibrium distribution $\rho_e(\mathbf{S}', J_0/\sqrt{p})$, we obtain a new expression

$$\begin{aligned}\left[\langle S_i(t)S_i(0)\rangle_\beta^{J_0}\right] &= \frac{1}{2^N[Z_0(J_0/\sqrt{p})]^p} \\ &\times \sum_\xi \sum_{S,S'} S_i S_i' \left(e^{tW}\right)_{S,S'} \\ &\times \exp\left(\frac{J_0}{\sqrt{p}} \sum_{\langle ij\rangle} \sum_\mu \xi_i^\mu \xi_j^\mu S_i' S_j'\right). \end{aligned} \tag{20}$$

Next, the uniform magnetization is assumed to be time-evolved starting from the perfectly ferromagnetic state, denoted F, at $t = 0$:

$$[\langle S_i(t)\rangle_\beta^F] = \left[\sum_S \left(e^{tW}\right)_{S,F} S_i\right]. \tag{21}$$

Let us write each part of the gauge transformation as follows:

$$U: S_i \to S_i\sigma_i,$$
$$U': S_i' \to S_i'\sigma_i,$$
$$V: \xi_i^\mu \to \xi_i^\mu\sigma_i.$$

Then the quantity inside the large square brackets of (21) is transformed as

$$\begin{aligned}\sum_S \left(e^{tW}\right)_{S,F} S_i \to V \sum_S \{U\left(e^{tW}\right)_{S,F}\}\{US_i\} &= \sum_S \{VU\left(e^{tW}\right)_{S,F}\} S_i\sigma_i \\ &= \sum_S \{U'\left(e^{tW}\right)_{S,F}\} S_i\sigma_i \\ &= \sum_S \left(e^{tW}\right)_{S,\sigma} S_i\sigma_i,\end{aligned}$$

where we have used the gauge invariance of the transition probability $VUU'e^{tW} = e^{tW}$. Thus the configurational average (21) becomes

$$\begin{aligned}[\langle S_i(t)\rangle_\beta^F] &= \frac{1}{2^N Z_0(J_0/\sqrt{p})^p} \\ &\times \sum_\xi \sum_\sigma \exp\left(\frac{J_0}{\sqrt{p}} \sum_{\langle ij\rangle} \sum_\mu \xi_i^\mu \xi_j^\mu \sigma_i\sigma_j\right) \\ &\times \sum_S \left(e^{tW}\right)_{S,\sigma} S_i\sigma_i. \end{aligned} \tag{22}$$

A comparison of (20) and (22) shows their equivalence

$$\left[\langle S_i(t)S_i(0)\rangle_\beta^{J_0}\right] = [\langle S_i(t)\rangle_\beta^F]. \tag{23}$$

This result (23) is the same as in the spin glass case [17] and can thus be interpreted in the same way, as follows.

The left-hand side is the autocorrelation function for a system prepared in equilibrium at a point on the line $J_0 = \beta$ and quenched at $t = 0$ to an arbitrary point in the phase diagram (Fig. 5, from the tail to the head of arrow a). The right-hand side represents the remanent magnetization prepared as the perfect ferromagnetic state, quenched at $t = 0$ to the same point as the left-hand side (Fig. 5, from the tail to the head of arrow b). Equation (23) means that relaxations of autocorrelation and magnetization are equivalent to each other, on average, under the above initial conditions.

A similar argument proves the following relation between overlap order parameters with different initial conditions:

$$\left[\xi_i^\mu\langle S_i(t)\rangle_\beta^{J_0}\right] = [\xi_i^\mu\langle S_i(t)\rangle_\beta^F].$$

The overlap order parameter has the same value at any time, on average, for the two types of quenching shown as arrows a and b in Fig. 5.

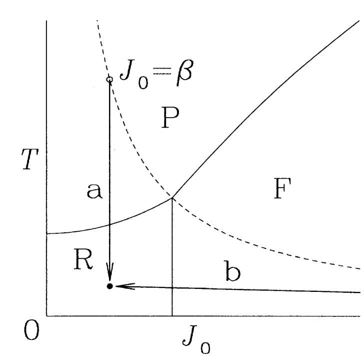

FIG. 5. The system is quenched to an arbitrary point from an equilibrium state on the line $J_0 = \beta$, as indicated by arrow a, or brought to the same point from the perfectly ferromagnetic state, as indicated by arrow b.

IV. EXISTENCE OF ORDERED PHASES

The existence of ferromagnetic states has been proved by the Peierls argument for the nonrandom Ising model in finite dimensions [18,19]. The application of this method to spin glasses has been discussed by one of the present authors [20] and by Horiguchi and Morita [16]: the existence of ferromagnetic states has been shown for spin glasses with sufficiently large mean values of the exchange interactions. In this section, we use the same methods to show the existence of ferromagnetic and finite-overlap phases for our finite-dimensional neural network with nearest-neighbor interactions under certain conditions.

A. Ferromagnetic phase

Under the boundary condition that all spins are up, the positiveness of magnetization is written as

$$[m_F] = 1 - \frac{2\,[\langle N_-\rangle]}{N} > 0, \tag{24}$$

where N_- is the number of down spins. To prove that (24) actually holds in certain regions of the phase diagram, it is necessary to demonstrate a positive lower bound of the right-hand side of the equality in (24). An upper bound on $[\langle N_-\rangle]$ is estimated as [19,16]

$$\frac{[\langle N_-\rangle]}{N} \le \sum_{l_1=1}^{\infty}\cdots\sum_{l_d=1}^{\infty}\left(\prod_{i=1}^{d} l_i^{1/(d-1)}3^{2l_i-1}\right) \times\left[\exp\left(-2\beta\sum_c J_{ij}\right)\right], \tag{25}$$

where d is the spatial dimensionality and the summation over l_i's represents a summation of contributions from various Peierls contours (polyhedrons separating down spins immediately inside and up spins immediately outside) with $2l_i$ as the number of faces orthogonal to the ith coordinate axis. The summation $\sum_c$ in the exponent runs over the bonds orthogonal to the faces of a contour with given $\{l_i\}$. From (24), a sufficient condition for $[m_F] > 0$ is that the right-hand side of (25) is less than $\frac{1}{2}$.

An upper bound on the configurational average of the exponential function in (25)

$$\left[\exp\left(-\frac{2\beta}{\sqrt{p}}\sum_c\sum_\mu \xi_i^\mu\xi_j^\mu\right)\right] = \left\{\frac{1}{Z_0(J_0/\sqrt{p})}\sum_\xi \exp\left(\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle}\xi_i^\mu\xi_j^\mu - \frac{2\beta}{\sqrt{p}}\sum_c \xi_i^\mu\xi_j^\mu\right)\right\}^p \tag{26}$$

is given as follows for $J_0 \ge 2\beta$. The numerator on the right-hand side of (26) can be regarded as a partition function if it is written as

$$Z_b \equiv \sum_s \exp\left(\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle} c_{ij}s_is_j\right), \tag{27}$$

where $c_{ij} = 1 - 2\beta/J_0$ (≥ 0) if the bond $\langle ij\rangle$ is on the Peierls contour c in (26) and $c_{ij} = 1$ otherwise. The denominator may then be reexpressed as

$$\begin{aligned} Z_0\left(\frac{J_0}{\sqrt{p}}\right) &= \sum_s \exp\left(\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle} s_is_j\right) \\ &= \sum_s \exp\left(\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle} c_{ij}s_is_j + \frac{2\beta}{\sqrt{p}}\sum_c s_is_j\right) \\ &= \left\langle \exp\left(\frac{2\beta}{\sqrt{p}}\sum_c s_is_j\right)\right\rangle_b Z_b, \end{aligned}$$

where the angular brackets with the subscript b denote the average with respect to the Boltzmann factor corresponding to the partition function Z_b. Since all couplings in (27) are non-negative, all correlation functions $\langle S_i \cdots S_k\rangle_b$ are also non-negative [18]. Therefore,

$$\begin{aligned} &\left\langle \exp\left(\frac{2\beta}{\sqrt{p}}\sum_c s_is_j\right)\right\rangle_b \\ &= \left\langle \prod_c \left(\cosh\frac{2\beta}{\sqrt{p}} + s_is_j\sinh\frac{2\beta}{\sqrt{p}}\right)\right\rangle_b \\ &= \left(\cosh\frac{2\beta}{\sqrt{p}}\right)^L \left\langle \prod_c \left(1 + s_is_j\tanh\frac{2\beta}{\sqrt{p}}\right)\right\rangle_b \\ &\ge \left(\cosh\frac{2\beta}{\sqrt{p}}\right)^L, \end{aligned} \tag{28}$$

where $L = \sum_i 2l_i$. From (26)–(28), we have

$$\left[\exp\left(-\frac{2\beta}{\sqrt{p}}\sum_c\sum_\mu \xi_i^\mu\xi_j^\mu\right)\right] \le \left(\cosh\frac{2\beta}{\sqrt{p}}\right)^{-pL}$$

under the condition $J_0 \ge 2\beta$. The sufficient condition (24) with (25) then reads

$$\left\{\sum_{l=1}^{\infty} l^{1/(d-1)}3^{2l-1}\left(\cosh\frac{2\beta}{\sqrt{p}}\right)^{-2pl}\right\}^d < \frac{1}{2}, \tag{29}$$

which is clearly satisfied by sufficiently large β if $d > 1$.

There are special simplifications in two dimensions [16], leading to the following sufficient condition yielding a larger region in which the existence of ferromagnetic order is proved:

$$\begin{aligned} &\frac{2}{4^2 3^2}\sum_{b=4,6,\ldots}^{\infty} b3^b\left[\exp\left(-\frac{2\beta}{\sqrt{p}}\sum_c\sum_\mu \xi_i^\mu\xi_j^\mu\right)\right] \\ &\le \frac{1}{72}\sum_{b=4,6,\ldots}^{\infty} b3^b\left(\cosh\frac{2\beta}{\sqrt{p}}\right)^{-pb} \\ &= \frac{u^2(2-u)}{36(1-u)^2} < \frac{1}{2}, \end{aligned} \tag{30}$$

where

$$u = 9\left(\cosh\frac{2\beta}{\sqrt{p}}\right)^{-2p}.$$

The relation (30) is satisfied if $u < 0.7944$ for positive u, thereby proving a (p-dependent) limit on β and J_0 ($\geq 2\beta$) for which a ferromagnetic state is demonstrated to exist by this analysis.

Another sufficient condition for a ferromagnetic phase plays a complementary role in determining the region with ferromagnetic order under the opposite condition on J_0 and β, $2\beta \geq J_0$. The sufficient condition (24) with (25) for ferromagnetic order can be replaced by [16]

$$\sum_{l_1=1}^{\infty}\cdots\sum_{l_d=1}^{\infty}\left(\prod_{i=1}^{d} l_i^{1/(d-1)}3^{2l_i-1}\right)$$
$$\times\left[\theta(v)\exp\left(-\frac{2\beta v}{\sqrt{p}}\right)+\theta(-v)\right] < \frac{1}{2}, \quad (31)$$

where $v = \sum_c\sum_\mu \xi_i^\mu\xi_j^\mu$. The step function $\theta(v)$ is bounded from above by $\exp(av/\sqrt{p})$ for an arbitrary non-negative constant a. The first term in the square brackets of the left-hand side of (31) is thus bounded by

$$\exp\left(-\frac{2\beta - a}{\sqrt{p}}v\right).$$

Then the same argument as before leads us to the upper bound on the first term on the left-hand side of (31)

$$\left\{\sum_{l=1}^{\infty} l^{1/(d-1)}3^{2l-1}\left(\cosh\frac{2\beta - a}{\sqrt{p}}\right)^{-2pl}\right\}^d \quad (32)$$

under the condition $J_0 \geq 2\beta - a$. The parameter a is arbitrary as long as it is non-negative. We choose $a = 2\beta - J_0$ because (32) assumes the smallest value for this a. With this a, (32) reads

$$\left\{\sum_{l=1}^{\infty} l^{1/(d-1)}3^{2l-1}\left(\cosh\frac{J_0}{\sqrt{p}}\right)^{-2pl}\right\}^d \quad (33)$$

with the condition $2\beta \geq J_0$, which comes from $a \geq 0$.

The second term in the square brackets of the left-hand side of (31) is bounded from above as

$$[\theta(-v)] \leq \left[\exp\left(-\frac{2\tilde{a}v}{\sqrt{p}}\right)\right],$$

where $\tilde{a}$ is again a non-negative constant. The same estimation as above yields a bound

$$\left[\exp\left(-\frac{2\tilde{a}v}{\sqrt{p}}\right)\right] \leq \left\{\sum_{l=1}^{\infty} l^{1/(d-1)}3^{2l-1}\left(\cosh\frac{2\tilde{a}}{\sqrt{p}}\right)^{-2pl}\right\}^d. \quad (34)$$

The parameter $\tilde{a}$ should satisfy $J_0 \geq 2\tilde{a}$ for this upper bound to hold. The best result, or the smallest value on the right-hand side of (34), is obtained when $2\tilde{a}$ takes its largest possible value J_0. Then a comparison of (33) and (34) shows that an upper bound on the second term on the left-hand side of (31) reduces to the same expression as that for the first term. Our final sufficient condition thus reads

$$\left\{\sum_{l=1}^{\infty} l^{1/(d-1)}3^{2l-1}\left(\cosh\frac{J_0}{\sqrt{p}}\right)^{-2pl}\right\}^d < \frac{1}{4}. \quad (35)$$

This result is valid for $2\beta \geq J_0$, the opposite of the inequality $J_0 \geq 2\beta$ in the first sufficient condition (29) for the ferromagnetic phase. Equation (35) is satisfied by sufficiently large J_0.

Again, in the case of the two-dimensional square lattice, a more compact formula can be derived [16]. Corresponding to (30), we have

$$\frac{u^2(2-u)}{36(1-u)^2} < \frac{1}{4} \quad (36)$$

with

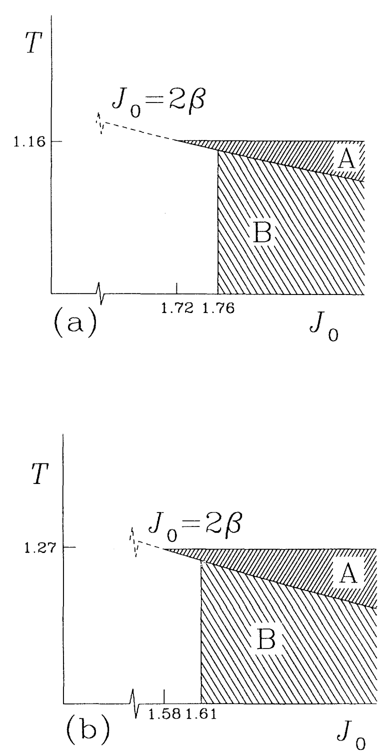

FIG. 6. Regions in the phase diagram where the existence of ferromagnetic order is proved for (a) $p = 2$ and (b) $p = 16$. The sufficient condition (30) is satisfied in A and (36) is satisfied in B.

$$u = 9\left(\cosh\frac{J_0}{\sqrt{p}}\right)^{-2p}$$

as a sufficient condition for ferromagnetic order in the region $2\beta \geq J_0$.

We show the regions in the phase diagram where either the first sufficient condition (30) (Fig. 6, region A) or the second (36) (Fig. 6, region B) is satisfied for the square lattice. A similar figure can be drawn for any hypercubic lattice.

B. Finite-overlap phase

The overlap order parameter was calculated explicitly in Sec. II C on the line $J_0 = \beta$. It was shown that this order parameter is finite for $J_0 > \sqrt{p}K_c$. We now use the Peierls argument to prove the existence of the overlap order in a finite region, not only on the line $J_0 = \beta$, with sufficiently large J_0 and β. The overlap order parameter can be expressed as a spontaneous magnetization by a simple gauge transformation:

$$[m_R] = \left[\frac{1}{N}\sum_k \frac{\sum_s S_k\xi_k^1 \exp\left(\beta/\sqrt{p}\sum_{\langle ij\rangle}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu S_i S_j\right)}{\sum_s \exp\left(\beta/\sqrt{p}\sum_{\langle ij\rangle}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu S_i S_j\right)}\right]$$

$$= \left[\frac{1}{N}\sum_k \frac{\sum_s S_k \exp\left\{\beta/\sqrt{p}\sum_{\langle ij\rangle}\left(1+\sum_{\mu=2}^{p}\xi_i^\mu\xi_j^\mu\xi_i^1\xi_j^1\right)S_i S_j\right\}}{\sum_s \exp\left\{\beta/\sqrt{p}\sum_{\langle ij\rangle}\left(1+\sum_{\mu=2}^{p}\xi_i^\mu\xi_j^\mu\xi_i^1\xi_j^1\right)S_i S_j\right\}}\right].$$

All boundary spins are supposed to be up in the last expression. A sufficient condition for $[m_R] > 0$ is, similarly to (24) and (25),

$$\sum_{l_1=1}^{\infty}\cdots\sum_{l_d=1}^{\infty}\left(\prod_{i=1}^{d} l_i^{1/(d-1)}3^{2l_i-1}\right)\left[\exp\left\{-\frac{2\beta}{\sqrt{p}}\sum_c\left(1+\sum_{\mu=2}^{p}\xi_i^\mu\xi_j^\mu\xi_i^1\xi_j^1\right)\right\}\right] < \frac{1}{2}. \tag{37}$$

The first step of our evaluation of the configurational average appearing in (37), denoted R_1 below, is to decouple the first pattern variables from the rest:

$$R_1 = \frac{1}{Z_0(J_0/\sqrt{p})^p}\sum_\xi \exp\left\{\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu - \frac{2\beta}{\sqrt{p}}\sum_c\left(1+\sum_{\mu=2}^{p}\xi_i^\mu\xi_j^\mu\xi_i^1\xi_j^1\right)\right\}$$

$$= \frac{e^{-2\beta L/\sqrt{p}}}{Z_0(J_0/\sqrt{p})}\sum_{\xi^1}\exp\left(\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle}\xi_i^1\xi_j^1\right)\left\{\frac{\sum_\xi \exp\left(J_0\sum_{\langle ij\rangle}\xi_i\xi_j/\sqrt{p} - 2\beta\sum_c \xi_i\xi_j\xi_i^1\xi_j^1/\sqrt{p}\right)}{Z_0(J_0/\sqrt{p})}\right\}^{p-1}, \tag{38}$$

where $L = \sum_i 2l_i$. To obtain an upper bound on the quantity in the curly brackets $\{\}$ of (38), to be denoted R_2, we regard the exponential function in the numerator of R_2 as a Boltzmann factor and represent the average with this factor as $\langle\rangle_{\xi^1}$ (note that ξ^1 is fixed in the calculation of R_2). Then R_2 is bounded from above, using the convexity of the exponential function,

$$R_2 = \left\langle \exp\left(\frac{2\beta}{\sqrt{p}}\sum_c \xi_i^1\xi_j^1\xi_i\xi_j\right)\right\rangle_{\xi^1}^{-1}$$

$$\leq \exp\left(-\frac{2\beta}{\sqrt{p}}\sum_c \xi_i^1\xi_j^1\langle\xi_i\xi_j\rangle_{\xi^1}\right).$$

Hence

$$R_1 \leq \frac{e^{-2\beta L/\sqrt{p}}}{Z_0(J_0/\sqrt{p})}\sum_{\xi^1}\exp\left\{\frac{J_0}{\sqrt{p}}\sum_{\langle ij\rangle}\xi_i^1\xi_j^1 - \frac{2\beta(p-1)}{\sqrt{p}}\sum_c \xi_i^1\xi_j^1\langle\xi_i\xi_j\rangle_{\xi^1}\right\}. \tag{39}$$

The final step of upper bound evaluation is to regard the exponential function in (39) as a Boltzmann factor and to denote the average with this Boltzmann factor by $\langle\rangle_f$. The above (39) is then rewritten as

$$R_1 \leq e^{-2\beta L/\sqrt{p}} \left\langle \exp\left(\frac{2\beta(p-1)}{\sqrt{p}} \sum_c \xi_i^1 \xi_j^1 \langle \xi_i \xi_j \rangle_{\xi^1}\right)\right\rangle_f^{-1} .$$

If $J_0 \geq 2\beta(p-1)$, the system corresponding to the average $\langle\rangle_f$ is a ferromagnetic Ising model. We thus find

$$\begin{aligned} R_1 &\leq e^{-2\beta L/\sqrt{p}} \left\langle \prod_c \cosh \frac{2\beta(p-1)}{\sqrt{p}} \langle \xi_i \xi_j \rangle_{\xi^1} \right\rangle_f^{-1} \\ &\leq e^{-2\beta L/\sqrt{p}}. \end{aligned}$$

A sufficient condition for the existence of a finite-overlap phase is therefore

$$\left\{\sum_{l=1}^{\infty} l^{1/(d-1)} 3^{2l-1} \exp\left(\frac{-4\beta l}{\sqrt{p}}\right)\right\}^d < \frac{1}{2}, \tag{40}$$

and J_0 must satisfy $J_0 \geq 2\beta(p-1)$. This inequality is satisfied by sufficiently large β.

The two-dimensional version of the sufficient condition turns out to be

$$\frac{u^2(2-u)}{36(1-u)^2} < \frac{1}{2}, \tag{41}$$

where $u = 9e^{-4\beta/\sqrt{p}}$ and J_0 must satisfy $J_0 \geq 2\beta(p-1)$. The inequality (41) is equivalent to $T < 1.648/\sqrt{p}$. For $p = 2$, the region where the above inequalities ($T < 1.165, J_0 \geq 2\beta$) are satisfied almost coincides with (but included in) region A in Fig. 6(a), where the ferromagnetic order has been proved to exist. If $p = 16$, the region satisfying the inequality above lies deep inside region A of Fig. 6(b).

Another sufficient condition for finite overlap is given similarly to the second condition for the ferromagnetic order given earlier. The quantity in the square brackets of (37) is replaced by

$$\left[\theta(v)\exp\left(-\frac{2\beta v}{\sqrt{p}}\right) + \theta(-v)\right]$$

as in (31). $\theta(v)$ is bounded from above by an exponential function. The problem then reduces to an estimation of the configurational average as in the previous sufficient condition. The same argument as before leads to

$$\left\{\sum_{l=1}^{\infty} l^{1/(d-1)} 3^{2l-1} \exp\left(-\frac{2lJ_0}{\sqrt{p}(p-1)}\right)\right\}^d < \frac{1}{4}. \tag{42}$$

There is a constraint $J_0 \leq 2\beta(p-1)$. There exists sufficiently large J_0 satisfying the inequality (42). The two-dimensional version of (42) is

$$\frac{u^2(2-u)}{36(1-u)^2} < \frac{1}{4},$$

where $u = 9e^{-2J_0/\sqrt{p}(p-1)}$. This inequality reads, in concrete numbers, $J_0 > 1.26\sqrt{p}(p-1)$, or $J_0 > 1.78$ and $J_0 \leq 2\beta$ for $p = 2$, and $J_0 > 75.5$ and $J_0 \leq 30\beta$ for $p = 16$. The region where these conditions are satisfied in Fig. 6(a) and 6(b) are included in the regions marked B in the figures. For $p = 2$, the present region almost coincides with B, while when $p = 16$, the present region is deep inside B. Qualitatively the same results apply to any hypercubic lattice.

In general, as p becomes larger, the region where we can prove finite overlap order shrinks while the region in which we can prove ferromagnetic order remains roughly unchanged. This is not necessarily a consequence of technical difficulties in the proof of the existence of a finite-overlap phase. Rather, the phase with finite overlap actually shrinks as p grows, as will be discussed in Sec. VI.

V. INFINITE-RANGE MODEL

The ferromagnetic Ising model with nearest-neighbor interactions on a d-dimensional hypercubic lattice reduces to the infinite-range model in the limit $d \to \infty$. Analogously, the infinite-range version of the present model (1) is expected to give useful hints on the behavior of the system with short-range interactions in high spatial dimensions. For this reason, we consider the Hamiltonian

$$H = -\frac{1}{2}\sum_{i \neq j} J_{ij} S_i S_j, \tag{43}$$

where

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

The number of patterns p is finite. The probability distribution of patterns is given as

$$P(\{\xi_i^\mu\}) = c\prod_{\mu=1}^{p} \exp\left(\frac{J_0}{2N}\sum_{i \neq j} \xi_i^\mu \xi_j^\mu\right), \tag{44}$$

where c is the normalization factor.

The free energy can be calculated in the same manner as in the case of perfectly random patterns [2] corresponding to $J_0 = 0$ and the result is

$$f = \frac{1}{2}\mathbf{m}^2 - T\left[\ln 2\cosh(\beta \mathbf{m}\cdot\boldsymbol{\xi})\right], \tag{45}$$

where the overlap order parameter $\mathbf{m} = (m_1, m_2, \ldots, m_p)$ satisfies

$$m_\mu = \left[\xi^\mu \tanh(\beta \mathbf{m}\cdot\boldsymbol{\xi})\right]. \tag{46}$$

The site index i for $\boldsymbol{\xi} = (\xi^1, \ldots, \xi^p)$ was omitted in (46) because of the self-averaging property: The configurational average of a function $[g(\boldsymbol{\xi}_i)]$, appearing in (45) and (46), is independent of i. It is proved in the Appendix that the probability distribution (44) is equivalent to the site-independent distribution

$$P(\{\xi_i^\mu\}) = \prod_\mu \frac{1}{2^p} \sum_{\eta^\mu=\pm 1} \prod_i \left(\frac{1+\eta^\mu m_{MF}}{2} \delta(\xi_i^\mu - 1) + \frac{1-\eta^\mu m_{MF}}{2} \delta(\xi_i^\mu + 1) \right) \quad (47)$$

in the calculation of the average $[g(\boldsymbol{\xi}_i)]$. The parameter $m_{MF}(\geq 0)$ is the magnetization of the infinite-range ferromagnetic Ising model satisfying the self-consistent equation

$$m_{MF} = \tanh(J_0 m_{MF}). \quad (48)$$

When $J_0 < 1$, the only solution of this equation is $m_{MF} = 0$. This means that the probability distribution (47) represents perfectly random patterns. Thus the properties of the system are independent of J_0 as long as this parameter is less than 1: The retrieval phase exists below the critical point $T_c = 1$ and various dynamically stable mixed states appear at lower temperatures [2].

Nontrivial results emerge when $J_0 > 1$. The parameter m_{MF} becomes finite, the η^μ symmetry in (47) is macroscopically broken, effectively freezing the system into one of the possible choices $\{\eta^\mu = \pm 1\}$, and the equation of state (46) has solutions with nonvanishing values of m_μ. In particular, all components of $\mathbf{m}$ are nonvanishing. To prove this statement, we first assume that $m_\mu \neq 0$ for $\mu = 1, \ldots, k$ $(< p)$, with $\mathrm{sgn} m_\mu = \eta^\mu$, where $\{\eta^\mu\}$ is the chosen symmetry-broken state for the $\{\boldsymbol{\xi}\}$ and $m_\mu = 0$ for $\mu = k+1, \ldots, p$. Then, (46) for $\mu = k+1$ reads

$$\begin{aligned} m_{k+1} &= \left[\xi^{k+1} \tanh\beta\left(m_1\xi^1 + \cdots + m_k\xi^k\right)\right]_{\{\eta^\mu\}} \\ &= \left[\xi^{k+1}\right]\left[\tanh\beta\left(m_1\xi^1 + \cdots + m_k\xi^k\right)\right]_{\{\eta^\mu\}} \neq 0, \end{aligned}$$

where $[\,]_{\{\eta^\mu\}}$ denotes that the average is taken with the particular choice of $\{\eta^\mu\}$ rather than the full sum of (47). By considering first $k = 1$ and then iteratively higher values it follows that the assumption $\mathrm{sgn} m_\mu = \eta^\mu$ is consistent but the assumption that some $m_\mu = 0$ is not.

A "uniform" solution with $m_\mu = \eta^\mu m$ satisfies

$$\begin{aligned} m &= \left[\eta^1\xi^1 \tanh\beta m\left(\eta^1\xi^1 + \cdots + \eta^p\xi^p\right)\right]_{\{\eta^\mu\}} \\ &= \left[\tanh\beta m\left(1 + \eta^1\eta^2\xi^1\xi^2 + \cdots + \eta^1\eta^p\xi^1\xi^p\right)\right]_{\{\eta^\mu\}}, \end{aligned} \quad (49)$$

which has a nonvanishing solution for large β. The critical point of this uniform-overlap state is obtained by expanding the right-hand side of (49):

$$m \approx \beta m\left\{1 + (p-1)m_{MF}^2\right\},$$

yielding $T_c = 1 + (p-1)m_{MF}^2$. Above this T_c, which was obtained from the analysis of the uniform solution, there is no solution with finite m_μ including nonuniform overlap solutions ($|m_\mu| \neq |m_\nu|$ for some μ and ν). This can be verified by expanding the free energy (45) to second order of m_μ:

$$\begin{aligned} f &\approx -T\ln 2 + \frac{1}{2}(1-\beta)\sum_{\mu=1}^p m_\mu^2 - \frac{1}{2}\beta m_{MF}^2 \sum_{\mu\neq\nu} m_\mu m_\nu \\ &\equiv -T\ln 2 + \frac{1}{2}\sum_{\mu\nu} f_{\mu\nu} m_\mu m_\nu. \end{aligned} \quad (50)$$

The eigenvalues of the matrix $f_{\mu\nu}$, $\lambda_1 = 1 - \beta\left\{1 + (p-1)m_{MF}^2\right\}$ and $\lambda_2 = 1 - \beta(1 - m_{MF}^2)$, determine the stability of the paramagnetic phase. As the temperature is lowered, λ_1 first becomes negative, signaling instability, at the critical point $T_c = 1 + (p-1)m_{MF}^2$.

The ferromagnetic order parameter is given by the formula

$$\begin{aligned} m_F &= \left[\tanh\beta\left(m_1\eta^1\xi^1 + \cdots + m_p\eta^p\xi^p\right)\right]_{\{\eta^\mu\}} \\ &= \left[\tanh\beta\left(m_1\xi^1 + \cdots + m_p\xi^p\right)\right]_{\{\eta^\mu=1\}}, \end{aligned} \quad (51)$$

which is nonvanishing for $\mathbf{m} \neq \mathbf{0}$. Thus the phase below T_c is ferromagnetic.

The final phase diagram is depicted in Fig. 7. There are three phases: paramagnetic (P), retrieval (R), and ferromagnetic (F') phases. The ferromagnetic phase is indicated with a prime (F') because the ferromagnetism in the present model is induced by the overlap ordering rather than appearing spontaneously, as implied in (51).

VI. DISCUSSIONS

Let us now summarize our results. We discussed the problem of neural networks with nearest-neighbor interactions on a finite-dimensional hypercubic lattice. Patterns to be embedded were generated according to the Boltzmann factor of the nearest-neighbor ferromagnetic Ising model on the same hypercubic lattice with a finite effective temperature.

This system has a gauge symmetry, which leads to the exact values of the internal energy and the overlap order parameter along a line defined by $J_0 = \beta$ in the phase diagram. The overlap order parameter was found to be vanishing in the range $J_0 \leq \sqrt{p}K_c$ on this line and finite for $J_0 > \sqrt{p}K_c$.

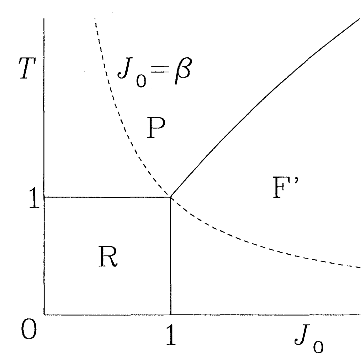

FIG. 7. The phase diagram of the infinite-range model with a finite number of patterns embedded. The line $J_0 = \beta$ goes across the multicritical point. The ferromagnetic phase F' has a finite overlap with all patterns ($m_\mu > 0$ for all μ).

We proved the existence of ferromagnetic order in the large-J_0, large-β region in the phase diagram by using the Peierls argument. The overlap order was also proved to exist in a region inside the ferromagnetic phase.

The infinite-range model was solved explicitly. The paramagnetic, retrieval, and ferromagnetic (plus finite-overlap) phases were identified in appropriate regions of the phase diagram.

These exact or rigorous results form the basis of our conjecture about the plausible structure of the phase diagram depending upon the number of embedded patterns p. We first consider the case of small p, taking the $p = 2$ model on the square lattice as an example [Fig. 8(a)]. The overlap order parameter is finite on the line $J_0 = \beta$ in the range $J_0 > \sqrt{p}K_c = 0.62$, whereas the ferromagnetic order is proved to exist for $J_0 > 1.76$ at small temperatures. We conjecture that the ferromagnetic order coexists with the overlap order for small temperatures in the range $J_0 > \sqrt{p}K_c = 0.62$ for the following reason. The critical point $J_0 = \sqrt{p}K_c$ marks the onset of ferromagnetic order in the embedded patterns $\{\xi_i^\mu\}$. This means that the exchange interaction (3) is essentially ferromagnetic for $J_0 > \sqrt{p}K_c$, leading to a ferromagnetic phase in this region.

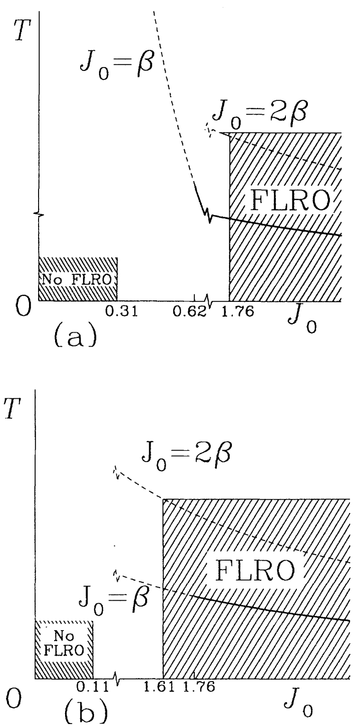

FIG. 8. Summary of constraints on the phase diagram in the case of the square lattice with (a) $p = 2$ and (b) $p = 16$. The existence of ferromagnetic long-range order (FLRO) and finite-overlap order is proved in the lower right-hand part. The absence of ferromagnetic order is proved near $J_0 = 0$. On the line $J_0 = \beta$, the overlap order parameter m_R is given explicitly in terms of the spontaneous magnetization of the ferromagnetic Ising model. m_R is finite on the lower right-hand side (solid part) and is vanishing on the other part (dashed part).

We next discuss the large-p case. Figure 8(b) shows the results of our analysis in the previous sections for $p = 16$ on the square lattice. The ferromagnetic phase is proved to exist in the low-temperature region for $J_0 > 1.61$, while the overlap order vanishes in the range $J_0 < 1.76$ on the line $J_0 = \beta$. Thus these two types of order coexist in the large-J_0 region ($J_0 > 1.76$), whereas only the ferromagnetic order exists in the intermediate region $1.61 < J_0 < \sqrt{p}K_c = 1.76$. The value of the lower limit

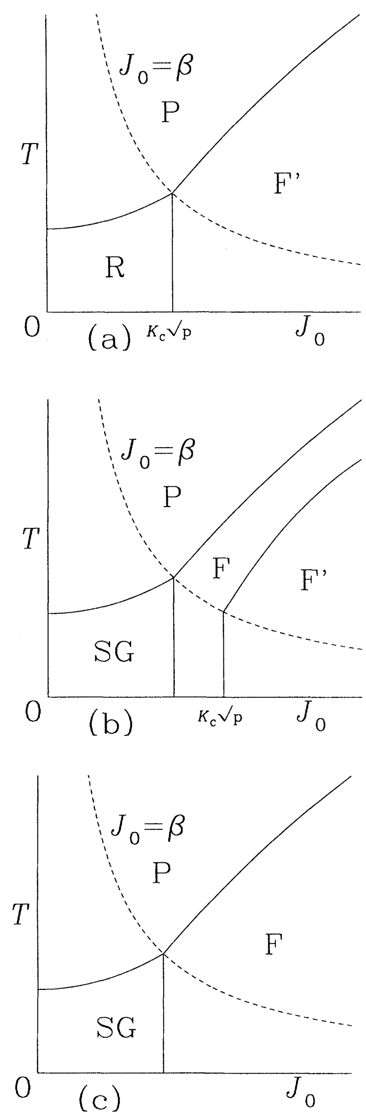

FIG. 9. Plausible phase diagrams. (a) For p small, three phases (paramagnetic P, retrieval R, and ferromagnetic with finite overlap F') meet at a multicritical point located on the line $J_0 = \beta$. (b) For large p, there appears a ferromagnetic phase without retrieval order (F) at intermediate values of J_0. This ferromagnetic phase keeps its existence in the limit $p \to \infty$ (c) where the system reduces to the spin glass model. The retrieval phase in the small-p case (a) is replaced by the spin glass phase if p is large (b) and (c). There may exist a mixed phase (a ferromagnetic phase with spin glass characteristics) and other variations.

1.61 here reflects the technical details in the Peierls argument and the ferromagnetic phase without overlap order is likely to continue to exist for smaller J_0. The upper bound $\sqrt{p}K_c = 1.76$, on the other hand, is the exact lower limit of the existence of the overlap order. Therefore, for the $p = 16$ model on the square lattice, there is a ferromagnetic phase without overlap order for intermediate values of J_0 at least on the line $J_0 = \beta$ (but probably also in a finite region for this range of J_0). This is true also for $p > 16$ on the same square lattice because the bound 1.61 remains approximately unchanged as p increases whereas $\sqrt{p}K_c$ increases indefinitely. We therefore expect that there exists a critical number of patterns p_c below which ($p < p_c$) the ferromagnetic phase coexists with the overlap order [denoted F' in Fig. 9(a)], while for $p > p_c$ the ferromagnetic phase is separated into two parts, with and without overlap order [denoted F' and F in Fig. 9(b), respectively].

In the region $J_0 \approx 0$, the retrieval phase would exist for small p as discussed in Sec. II B. This is actually the case for $p = 2$ in three dimensions (but not in two dimensions) as shown in Sec. II D. When p is small, this retrieval phase around $J_0 \approx 0$ will probably extend to $J_0 = \sqrt{p}K_c$ where the ferromagnetic phase with overlap order takes over. The retrieval phase around $J_0 \approx 0$ disappears for large p as discussed in Secs. II B and II D. The overlap order also ceases to exist in the intermediate J_0 for large p as discussed above (separation of F and F'). It is unlikely that the overlap order is finite for $J_0 \approx 0$ when this order is vanishing in the intermediate values of J_0 because smaller J_0 (≈ 0) means more random embedded patterns, harder to be stored and retrieved by the Hebb rule (3). Hence, as soon as the ferromagnetic phase splits into two regions (with and without overlap order) as p exceeds p_c, the retrieval phase near $J_0 = 0$ disappears. (Note that the overlap order can be identified with the retrieval order for $J_0 \approx 0$.) This implies that the critical number of patterns p_c for the disappearance of the retrieval phase near $J_0 = 0$ and that for the splitting of the ferromagnetic phase into two regions coincide. The ordered phase near $J_0 = 0$ for $p > p_c$, if any, would be spin-glass-like in the sense that the Edwards-Anderson order parameter $[\langle S_i \rangle^2]$ is finite with vanishing overlap and ferromagnetic order parameters. These conjectures are summarized in Figs. 9(a) and 9(b). In the limit $p \to \infty$ (after the thermodynamic limit), the model reduces to the Edwards-Anderson spin glass as discussed in Sec. II C. The ferromagnetic phase with overlap order disappears in this limit and the phase diagram looks like Fig. 9(c).

The phase diagram of the infinite-range model (Fig. 7) is consistent with these conjectures: The critical number of patterns p_c is infinite for the infinite-range model as suggested from the argument of Sec. II B about the dependence of p_c on the coordination number (which is infinite for the infinite-range model). This means that the phase diagram for any finite p should have the structure of Fig. 9(a), which is actually the case as seen in Fig. 7.

In conclusion, the system has three phases: paramagnetic, retrieval, and ferromagnetic accompanied by finite overlap phases if $p < p_c$. For p larger than p_c, the ferromagnetic phase separates into two parts, with and without overlap order for large J_0 and intermediate J_0, respectively. The ordered phase near $J_0 = 0$, if any, would be a retrieval phase ($p < p_c$) or a spin glass phase ($p > p_c$).

APPENDIX

We show that the distribution function (44) for the infinite-range model is equivalent to the single-site distribution (47) when one calculates the average of a function of single-site variables $[g(\boldsymbol{\xi}_k)]$. Ignoring the trivial normalization factor, we have

$$\begin{aligned}[g(\boldsymbol{\xi}_k)] &\propto \sum_{\{\xi_i^\mu\}} \exp\left(\frac{J_0}{2N}\sum_\mu \sum_{i\neq j} \xi_i^\mu \xi_j^\mu\right) g(\boldsymbol{\xi}_k) \\ &\propto \int \prod_\mu dx_\mu \exp\left(-\frac{J_0 N}{2}\sum_\mu x_\mu^2\right) \\ &\quad \times \sum_{\xi_i^\mu, i\neq k} \exp\left(J_0 \sum_{i\neq k}\sum_\mu x_\mu \xi_i^\mu\right) \\ &\quad \times \sum_{\xi_k^\mu=\pm 1} g(\boldsymbol{\xi}_k) \exp\left(J_0 \sum_\mu \xi_k^\mu x_\mu\right).\end{aligned}$$

The extremum of the integral over x_μ gives the self-consistent equation (48) with $x_\mu = \eta^\mu m_{MF}; \eta^\mu = \pm 1$. Spontaneous symmetry breaking leads to an arbitrary choice among the values of η^μ in the case that $m_{MF} \neq 0$. In any such symmetry broken state $\{\eta^\mu\}$

$$[g(\boldsymbol{\xi}_k)]_{\{\eta^\mu\}} \propto \sum_{\xi_k^\mu=\pm 1} g(\boldsymbol{\xi}_k) \prod_\mu \exp\left(J_0 \xi_k^\mu \eta^\mu m_{MF}\right),$$

which implies that any average is independent under the replacement of the distribution of (44) by that of (47).

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

[2] D. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985); Phys. Rev. Lett. **55**, 1530 (1985); Ann. Phys. (N.Y.) **173**, 30 (1987).

[3] J. Hertz, A. Krogh, and R.G. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, Reading, MA, 1991).

[4] D. Amit, *Modelling Brain Function* (Cambridge University Press, Cambridge, 1989).

[5] M. Schlüter and E. Wagner, Phys. Rev. E **49**, 1690 (1994).

[6] R. Monasson, J. Phys. (France) I **3**, 1141 (1993).
[7] A.C.C. Coolen, in *Statistical Mechanics of Neural Networks*, edited by L. Garrido, Lecture Notes in Physics Vol. 368 (Springer, Berlin, 1990).
[8] W. Tarkowski and M. Lewenstein, J. Phys. A **25**, 6251 (1992); **26**, 3669 (1993).
[9] For J_0 equal to or close to 0, the finiteness of the overlap $m_\mu > 0$ for only one μ can be identified with retrieval of the μth pattern. However, when J_0 is large, the patterns themselves develop long-range ferromagnetic order. In such a situation, the finiteness of m_μ should not be confused with retrieval. We thus avoid calling m_μ the retrieval order parameter.
[10] B. Derrida, E. Gardner, and A. Zippelius, Europhys. Lett. **4**, 167 (1987).
[11] S.F. Edwards and P.W. Anderson, J. Phys. F **5**, 965 (1975).
[12] M.F. Sykes, D.S. Gaunt, and M. Glen, J. Phys. A **9**, 97 (1976).
[13] D.S. Gaunt and M.F. Sykes, J. Phys. A **16**, 783 (1983).
[14] H. Nishimori, Prog. Theor. Phys. **66**, 1169 (1981).
[15] P. Le Doussal and A.B. Harris, Phys. Rev. B **40**, 9249 (1989).
[16] T. Horiguchi and T. Morita, Phys. Lett. **74**A, 340 (1979); J. Phys. A **15**, L75 (1981); **15**, 3551 (1982).
[17] Y. Ozeki (unpublished).
[18] R.B. Griffiths, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M.S. Green (Academic, London, 1972), Vol. 1.
[19] D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969), p. 113.
[20] H. Nishimori, Dr.Sc. thesis, University of Tokyo, 1981.