Journal of Statistical Physics, Vol. 87, Nos. 1/2, 1997

On the Free Energy of the Hopfield Model

G. R. Guerberoff[1,2] and G. A. Raggio[1,3]

Received June 12, 1996; final September 18, 1996

The general theory of inhomogeneous mean-field systems of Raggio and Werner provides a variational expression for the (almost sure) limiting free energy density of the Hopfield model

$$H^{\{\xi\}}_{N,p}(S) = -\frac{1}{2N} \sum_{i,j=1}^{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu S_i S_j$$

for Ising spins S_i and p random patterns $\xi^\mu = (\xi_1^\mu, \xi_2^\mu, ..., \xi_N^\mu)$ under the assumption that

$$\lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \delta_{\xi_i} = \lambda, \qquad \xi_i = (\xi_i^1, \xi_i^2, ..., \xi_i^p)$$

exists (almost surely) in the space of probability measures over p copies of $\{-1, 1\}$. Including an "external field" term $-\sum_{\mu=1}^{p} h^\mu \sum_{i=1}^{N} \xi_i^\mu S_i$, we give a number of general properties of the free-energy density and compute it for (a) $p=2$ in general and (b) p arbitrary when λ is uniform and at most the two components h^{μ_1} and h^{μ_2} are nonzero, obtaining the (almost sure) formula

$$f(\beta, \mathbf{h}) = \tfrac{1}{2} f^{\rm cw}(\beta, h^{\mu_1} + h^{\mu_2}) + \tfrac{1}{2} f^{\rm cw}(\beta, h^{\mu_1} - h^{\mu_2})$$

for the free energy, where $f^{\rm cw}$ denotes the limiting free energy density of the Curie–Weiss model with unit interaction constant. In both cases, we obtain explicit formulas for the limiting (almost sure) values of the so-called overlap parameters

$$m_N^\mu(\beta, \mathbf{h}) = N^{-1} \sum_{i=1}^{N} \xi_i^\mu \langle S_i \rangle$$

in terms of the Curie–Weiss magnetizations. For the general i.i.d. case with $\text{Prob}\{\xi_i^\mu = \pm 1\} = (1/2) \pm \varepsilon$, we obtain the lower bound $1 + 4\varepsilon^2(p-1)$ for the

[1] Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Ciudad Universitaria, 5000 Córdoba, Argentina.

[2] Becario CONICOR; e-mail: guerber@fis.uncor.edu.

[3] Investigador CONICET; e-mail: raggio@fis.uncor.edu.

0022-4715/97/0400-0333$12.50/0

temperature T_c separating the trivial free regime where the overlap vector is zero from the nontrivial regime where it is nonzero. This lower bound is exact for $p=2$, or $\varepsilon=0$, or $\varepsilon=\pm 1/2$. For $p=2$ we identify an intermediate temperature region between $T_*=1-4\varepsilon^2$ and $T_c=1+4\varepsilon^2$ where the overlap vector is homogeneous (i.e., all its components are equal) and nonzero. T_* marks the transition to the nonhomogeneous regime where the components of the overlap vector are distinct. We conjecture that the homogeneous nonzero regime exists for $p \geqslant 3$ and that $T_*=\max\{1-4\varepsilon^2(p-1), 0\}$.

KEY WORDS: Hopfield model; equilibrium statistical mechanics; random mean-field models.

1. INTRODUCTION

The Hopfield model for N classical Ising spins $\{S_i : i=1, 2,..., N\}$, is specified by the choice of p configurations $\{\xi^\mu = (\xi_1^\mu, \xi_2^\mu,..., \xi_N^\mu): \mu=1, 2,..., p\}$, which determine the interaction constants

$$J_{i,j}^{\{\xi\}} = -\frac{1}{2N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \qquad i, j=1, 2,..., N$$

entering the Hamiltonian

$$H_{N,p}^{\{\xi\}}(S) = \sum_{i,j=1}^{N} J_{i,j}^{\{\xi\}} S_i S_j \tag{1.1}$$

The statistical mechanics of this random mean-field model was studied in the late seventies by Pastur and Figotin,[1] although it gained notoriety and its name after Hopfield[2] proposed it as a model for an associative memory.[3] In this context the p configurations $S=\xi^\mu$ are called patterns and they become fixed points of the "linear threshold dynamics"

$$S_i(t+1) = \text{sign}\left(-2 \sum_{j=1}^{N} J_{i,j}^{\{\xi\}} S_j(t)\right)$$

as $N \to \infty$.

The equilibrium thermodynamics of the model assuming the $\{\xi_i^\mu\}$ are independent identically distributed random variables with $\text{Prob}\{\xi_i^\mu = \pm 1\} = 1/2$, and p independent of N, was analyzed by many authors after the work of Pastur and Figotin (e.g., refs. 4–7).

The analysis of the case of p growing with N but $p/N \to \alpha$ as $N \to \infty$ followed. One distinguishes the subextensive regime $\alpha=0$ (or regime of perfect memory), where satisfactory and complete results are available,[8–13] from the extensive regime $\alpha > 0$, where only partial results are known.[14–16] The present state of affairs, excluding the general "self-averaging" property of the model obtained in ref. 16, is clearly exposed in ref. 17.

Let

$$f_{N,p}^{\{\xi\}}(\beta) = (\beta N)^{-1} \ln Z_{N,p}^{\{\xi\}}(\beta)$$

be minus the free-energy density associated with the partition function

$$Z_{N,p}^{\{\xi\}}(\beta) = \sum_{S \in \{\pm 1\}^N} \exp\{-\beta H_{N,p}^{\{\xi\}}(S)\}$$

The strongest result on the thermodynamic limit under the i.i.d. assumption with symmetric distribution in the subextensive case (which includes the case of finite p) is in ref. 12: (1) $f_{N,p}^{\{\xi\}}(\beta)$ converges almost surely (i.e., with probability one) as $N \to \infty$ to the limiting negative free-energy density of the Curie–Weiss model with interaction constant 1 at the same reciprocal temperature. (2) Let $a^{\pm}(\beta)$ denote the largest $(+)$, resp. smallest $(-)$, solution of the equation $a = \tanh(\beta a)$. Upon adding $-h \sum_{i=1}^{N} \xi_i^{\nu} S_i$ to the Hamiltonian, the measures induced on the so-called overlap parameters

$$m_N^{\mu}(\beta; \{\xi\}) = N^{-1} \sum_{i=1}^{N} \xi_i^{\mu} \langle S_i \rangle$$

(here $\langle \cdot \rangle$ denotes the thermal expectation) converge as $h \to \pm 0$ weakly and almost surely to the point measure sitting at $a^{\pm}(\beta)$ times the νth unit vector in $\mathbb{R}^p$.

The purpose of this paper is, first, to point out an alternative treatment of the case of finitely many patterns using the methods of refs. 18 and 19. These papers provide a general and efficient treatment of the equilibrium thermodynamics of arbitrary mean-field systems. The last section of the second paper[19] gives a general result which can be applied directly to a general class of random mean-field models which includes the finite-p Hopfield model. Moreover, these results are obtained under an ergodic hypothesis on the random process $\{\xi_i^{\mu}\}$ which is weaker than the i.i.d. assumption: ergodicity for $N \to \infty$ of the empirical distribution of the vectors $\xi_i = (\xi_i^1, \xi_i^2, ..., \xi_i^p)$ $(i = 1, 2, ..., N)$. Second, we establish here a number of properties of the limiting free-energy density, compute it in some particular cases, and give a lower bound for the critical temperature marking the transition from zero to nonzero overlap in the general i.i.d. case (Section 5).

We recall the following familiar facts about the equilibrium thermodynamics of the Curie–Weiss model, which will often be used without mentioning them. We write $f^{\text{cw}}(\beta, h; J)$ for the limiting negative free-energy density of the Curie–Weiss model with interaction constant $J \geqslant 0$ at the reciprocal temperature β and external field h:

$$-\frac{J}{2N} \sum_{i,j=1}^{N} S_i S_j - h \sum_{i=1}^{N} S_i \tag{1.2}$$

and $m^{\text{cw}}(\beta, h; J)$ for the corresponding limiting mean magnetization. We have

$$f^{\text{cw}}(\beta, h; J) = \max_{t \in [-1,1]} g(\beta, h, t; J)$$

where

$$g(\beta, h, t; J) = \frac{J}{2} t^2 + ht + \beta^{-1} \eta \left(\frac{1+t}{2} \right) \tag{1.3}$$

and $\eta(t) = -t \ln(t) - (1-t) \ln(1-t)$ for $t \in [0, 1]$ with the usual understanding that $0 \ln(0) = 0$.

For $\beta J \leqslant 1$ or $h \neq 0$, there is a unique maximizer $t = m^{\text{cw}}(\beta, h; J)$ for g. For $\beta J > 1$ and $h = 0$ there are two maximizers $m^{\text{cw}}(\beta, 0^{\pm}; J)$ with $-m^{\text{cw}}(\beta, 0^-; J) = m^{\text{cw}}(\beta, 0^+; J) > 0$. For all $\beta J \leqslant 1$, one has $m^{\text{cw}}(\beta, 0; J) = 0$. One has $\lim_{\beta \downarrow 1/J} m^{\text{cw}}(\beta, 0^{\pm}; J) = m^{\text{cw}}(1/J, 0; J) = 0$. Now, $\beta \mapsto f^{\text{cw}}(\beta, h; J)$ is a strictly convex, strictly decreasing, and differentiable function with

$$-\beta^2 \frac{\partial f^{\text{cw}}}{\partial \beta} (\beta, h; J) = \eta \left(\frac{1 + m^{\text{cw}}(\beta, h; J)}{2} \right)$$

$h \mapsto f^{\text{cw}}(\beta, h; J)$ is strictly convex, even, and differentiable except for $h = 0$ when $\beta J > 1$. For $\beta J > 1$, the left and right derivatives $(\partial f^{\text{cw}}/\partial h)_{\pm} (\beta, 0; J)$ are given by $m^{\text{cw}}(\beta, 0^{\pm}; J)$.

2. BASIC RESULTS

We consider more general models where the configuration space $\mathcal{S}$ of the spin is a finite discrete set of reals. The positive integer p is fixed and supressed in the notation. Let σ denote the "spin function" (identity function) on $\mathcal{S}$, i.e., $\sigma(s) = s$ $(s \in \mathcal{S})$. We let $\mathcal{X}$ be the p-fold direct product of $\mathcal{S}$ provided with the product topology. By collecting the ξ_i^μ with fixed i to $\boldsymbol{\xi}_i = (\xi_i^1, \xi_i^2, ..., \xi_i^p) \in \mathcal{X}$, for $i = 1, 2, ..., N$, we obtain a random process on the spin sites taking values in $\mathcal{X}$. It is crucial for the application of the theory of refs. 18 and 19 that the randomness can be localized in the spin sites, in the present case

$$J_{i,j}^{\{\xi\}} = -\frac{1}{2N}\langle \xi_i, \xi_j \rangle, \qquad i, j = 1, 2, ..., N$$

Here and in the following $\langle \cdot, \cdot \rangle$ denotes the usual scalar product in $\mathbb{R}^p$.

Given a vector $\mathbf{h} = (h^1, h^2, ..., h^p)$ with real components h^μ, we add the term

$$-\sum_{\mu=1}^{p} h^\mu \sum_{i=1}^{N} \xi_i^\mu S_i = -\sum_{i=1}^{N} \langle \mathbf{h}, \xi_i \rangle S_i$$

to the Hamiltonian (1.1) and denote by $f_N^{\{\xi\}}(\beta, \mathbf{h})$ the corresponding (negative) free-energy density. This incorporates p "external magnetic fields" of strength h^μ in the "direction of the pattern μ," which permit the control of the overlap parameters:

$$\frac{\partial f_N^{\{\xi\}}}{\partial h^\mu}(\beta, \mathbf{h}) = m_N^\mu(\beta, \mathbf{h}; \{\xi\}) \tag{2.1}$$

We write $\delta_{\mathbf{x}}$ for the point measure sitting at $\mathbf{x} \in \mathscr{X}$. The only hypothesis on a realization of the random process $\{\xi_i^\mu\}$ used in ref. 19 is

$$\text{weak-}\lim_{N \to \infty} N^{-1} \sum_{i=1}^{N} \delta_{\xi_i} = \lambda \tag{2.2}$$

in the space of probability measures on $\mathscr{X}$. This ergodicity condition is satisfied almost surely if $\{\xi_i^\mu\}$ are i.i.d., or, more generally, if $\{\xi_i\}$ is a stationary $\mathscr{X}$-valued Markov process.

Any probability measure ϕ on $\mathscr{S}$ is specified by the numbers $\{\phi(\{s\}): s \in \mathscr{S}\}$; we denote by

$$S(\phi) = -\sum_{s \in \mathscr{S}} \phi(\{s\}) \ln(\phi(\{s\}))$$

its entropy relative to the uniform probability measure (i.e., normalized counting measure) on $\mathscr{S}$, denoted by ω.

We write Λ for the support in $\mathscr{X}$ of the probability measure λ. Consider the set Γ_λ of functions $\mathbf{x} \mapsto \phi_{\mathbf{x}}$ on Λ taking values in $\mathscr{M}(\mathscr{S})$, the probability measures on $\mathscr{S}$. If $\Phi = \{\mathbf{x} \mapsto \phi_{\mathbf{x}}\}$ is such a family, then we let $\mathbf{m}_\lambda(\Phi)$ be the vector in $\mathbb{R}^p$ with components

$$m_\lambda^\mu(\Phi) \stackrel{\text{def}}{=} \sum_{\mathbf{x} \in \Lambda} \lambda(\{\mathbf{x}\})\, x^\mu \langle \sigma \rangle_{\phi_{\mathbf{x}}}, \qquad \mu = 1, 2,..., p \tag{2.3}$$

The following result, which is a special case of a general result obtained in ref. 19, is the basis of all our work.

Theorem.[19] If (2.2) holds (resp. holds almost surely, i.e., with probability 1), then

$$\lim_{N \to \infty} f_N^{\{\xi\}}(\beta, \mathbf{h}) = \max_{\Phi \in \Gamma_\lambda} \left\{ \tfrac{1}{2} \|\mathbf{m}_\lambda(\Phi)\|^2 + \langle \mathbf{h}, \mathbf{m}_\lambda(\Phi) \rangle + \beta^{-1} \sum_{\mathbf{x} \in \Lambda} \lambda(\{\mathbf{x}\})\, S(\phi_{\mathbf{x}}) \right\} \tag{2.4}$$

(resp. almost surely).

For a family $\Phi = \{\mathbf{x} \mapsto \phi_{\mathbf{x}}\} \in \Gamma_\lambda$ we will write

$$S_\lambda(\Phi) = \sum_{\mathbf{x} \in \Lambda} \lambda(\{\mathbf{x}\})\, S(\phi_{\mathbf{x}})$$

This is a positive, strictly concave function on Γ_λ, the convex sum of two families being defined pointwise: $(\alpha\Phi + (1-\alpha)\Psi)_{\mathbf{x}} = \alpha\phi_{\mathbf{x}} + (1-\alpha)\psi_{\mathbf{x}}$. Now, S_λ assumes its maximal value $\ln(|\mathscr{S}|)$ only for the family having as constant value on Λ the uniform probability measure ω on $\mathscr{S}$. The minimal value 0 of S_λ is assumed exactly on those families Φ for which for every $\mathbf{x} \in \Lambda$ one has $\phi_{\mathbf{x}} = \delta_s$ for some $s \in \mathscr{S}$ which may depend on $\mathbf{x}$.

We write $f(\beta, \mathbf{h})$ for the nonrandom limiting (negative) free-energy density. The solution of the variational problem of (2.4) poses a rather unmanageable problem in the general case. However, in the classical Ising spin case $\mathscr{S} = \{1, -1\}$ a probability measure ϕ on $\mathscr{S}$ is uniquely determined by the expectation value $\langle \sigma \rangle_\phi$, since $\phi(\{1\}) = 1 - \phi(\{-1\}) = (1 + \langle \sigma \rangle_\phi)/2$. Thus, the entropy of ϕ is a function of this expectation

$$S(\phi) = -\phi(\{1\}) \ln[\phi(\{1\})] - \phi(\{-1\}) \ln[\phi(\{-1\})] = \eta\left(\frac{1 \pm \langle \sigma \rangle_\phi}{2}\right)$$

It follows that the functional to be maximized in (2.4) depends only on the set $\{\langle \sigma \rangle_{\phi_{\mathbf{x}}} : \mathbf{x} \in \Lambda\}$ of expectations, and one obtains a "gap" equation in terms of these expectations which is the multidimensional analogue of the equation $a = \tanh(\beta(Ja + h))$ for the Curie–Weiss model. For $\phi \in \mathscr{M}(\mathscr{S})$, we write $\hat{\phi}$ for the flipped measure determined by $\langle \sigma \rangle_{\hat{\phi}} = -\langle \sigma \rangle_\phi$; alternatively, $\hat{\phi}(\{\pm 1\}) = \phi(\{\mp 1\})$. Notice that $S(\hat{\phi}) = S(\phi)$.

Given a family $\Phi = \{\phi_{\mathbf{x}}\}$ of probability measures on $\mathscr{S} = \{1, -1\}$, we denote by $F(\Phi)$ (the λ dependence is supressed; the β and $\mathbf{h}$ dependence will be indicated by an index when needed) the functional (negative "free energy") to be maximized in (2.4):

$$F(\Phi) = 2^{-1} \|\mathbf{m}_\lambda(\Phi)\|^2 + \langle \mathbf{h}, \mathbf{m}_\lambda(\Phi)\rangle + \beta^{-1} S_\lambda(\Phi)$$

Due to (2.3), F depends only on the values of Φ on the support of λ. We observe that Γ_λ equipped with the metric $d(\Phi, \Psi) = \max_{\mathbf{x} \in \Lambda} |\langle\sigma\rangle_{\phi_{\mathbf{x}}} - \langle\sigma\rangle_{\psi_{\mathbf{x}}}|$ is a compact metric space (isomorphic to $[-1, 1]^{|\Lambda|}$ via the identification of $\phi \in \mathscr{M}(\mathscr{S})$ with $\langle\sigma\rangle_\phi \in [-1, 1]$). Moreover, F is continuous.

Proposition 1. Let $\mathscr{S} = \{1, -1\}$.

1. If Φ is a maximizing family for the variational problem (2.4), then $|\langle\sigma\rangle_{\phi_{\mathbf{x}}}| < 1$ for all $\mathbf{x} \in \Lambda$; moreover,

$$\langle\sigma\rangle_{\phi_{\mathbf{x}}} = \tanh(\beta\langle \mathbf{m}_\lambda(\Phi) + \mathbf{h}, \mathbf{x}\rangle) \tag{2.5}$$

for every $\mathbf{x} \in \Lambda$.

2. If Φ maximizes (2.4), then $\langle\sigma\rangle_{\phi_{-\mathbf{x}}} = -\langle\sigma\rangle_{\phi_{\mathbf{x}}}$ (i.e., $\phi_{-\mathbf{x}} = \widehat{\phi_{\mathbf{x}}}$), for all $\mathbf{x}$ for which $\mathbf{x} \in \Lambda$, and $-\mathbf{x} \in \Lambda$. A family with this property will be called reflexive.

3. If $\mathbf{h} = \mathbf{0}$, then Φ maximizes (2.4) iff $\hat{\Phi} = \{\mathbf{x} \to \widehat{\phi_{\mathbf{x}}}\}$ maximizes (2.4).

Proof. 1. Take such a "boundary" family Φ with $\langle\sigma\rangle_{\phi_{\mathbf{x}_o}} = \pm 1$ for some $\mathbf{x}_o \in \Lambda$, and define $\tilde{\Phi}$ by changing only the value at $\mathbf{x}_o$; i.e., $\langle\sigma\rangle_{\tilde{\phi}_{\mathbf{x}_o}} = \pm a$ for an $a \in (-1, 1)$ to be specified. We can compute the change in F and get

$$\begin{aligned} l(a) \overset{\text{def}}{=} F(\tilde{\Phi}) - F(\Phi) &= \lambda(\{\mathbf{x}_o\})\, g(\beta, W_\pm, a; p\lambda(\{\mathbf{x}_o\})) \\ &\quad - \frac{p}{2} \lambda(\{\mathbf{x}_o\})^2 - \lambda(\{\mathbf{x}_o\})\, W_\pm \end{aligned}$$

where $W_\pm = \pm\langle \mathbf{x}_o, \mathbf{h} + \mathbf{m}_\lambda(\Phi)\rangle - \lambda(\{\mathbf{x}_o\})\, p$, and g is given by (1.3). Now the maximal value of $g(\beta, \pm W, \cdot; p\lambda(\{\mathbf{x}_o\}))$ is certainly larger than its value at $a = 1$ for any $\beta > 0$, so that there is an a with $l(a) > l(1) = 0$.

Viewing F as a function of $\{\langle\sigma\rangle_{\phi_{\mathbf{x}}} : \mathbf{x} \in \Lambda\}$ and setting all first partial derivatives equal to 0 when $\lambda(\{\mathbf{x}\}) \neq 0$, we obtain the gap equations (2.5) as the condition for critical point of F. By the previous result, we have that the maximal value of F is not assumed in a family taking the boundary values where the derivatives do not exist.

The symmetry property stated in part 2 is an immediate consequence of the gap equations, but we give a direct proof.

2. Define a new family $\tilde{\Phi}$ by $\tilde{\phi}_{\mathbf{x}} = \alpha_1(\mathbf{x})\, \phi_{\mathbf{x}} + \alpha_2(\mathbf{x})\, \widehat{\phi_{-\mathbf{x}}}$, where $\alpha_1(\mathbf{x}) = \lambda(\{\mathbf{x}\})(\lambda(\{\mathbf{x}\}) + \lambda(\{-\mathbf{x}\}))^{-1}$ if $\mathbf{x} \in \Lambda \cap (-\Lambda) = \Lambda_o$; $\alpha_1(\mathbf{x}) = 1$ otherwise; and $\alpha_2(\mathbf{x}) = 1 - \alpha_1(\mathbf{x})$. Using $\alpha_1(-\mathbf{x}) = \alpha_2(\mathbf{x})$, for $\mathbf{x} \in \Lambda_o$, and the definition of $\mathbf{m}_\lambda(\cdot)$, the change of variables $\mathbf{x} \to -\mathbf{x}$ gives $\mathbf{m}_\lambda(\tilde{\Phi}) = \mathbf{m}_\lambda(\Phi)$. Strict concavity of the entropy and the same direct calculation gives $S_\lambda(\tilde{\Phi}) \geqslant S_\lambda(\Phi)$ with strict inequality iff for some $\mathbf{x}_o \in \Lambda_o$ one has $\langle \sigma \rangle_{\phi_{-\mathbf{x}_o}} \neq -\langle \sigma \rangle_{\phi_{\mathbf{x}_o}}$. Thus, $F(\tilde{\Phi}) \geqslant F(\Phi)$, with strict inequality if Φ is not reflexive.

3. $m_\lambda^\mu(\hat{\Phi}) = -m_\lambda^\mu(\Phi)$ and $S_\lambda(\hat{\Phi}) = S_\lambda(\Phi)$, so that $F(\hat{\Phi}) = F(\Phi)$ if $\mathbf{h} = \mathbf{0}$. ∎

In their seminal paper, "On the theory of disordered spin systems,"[1] Pastur and Figotin obtain

$$f(\beta, \mathbf{h}) = \max_{\mathbf{m}} F^{\mathrm{PF}}(\mathbf{m}) \tag{2.6}$$

where

$$F^{\mathrm{PF}}(\mathbf{m}) := (-1/2)\, \|\mathbf{m}\|^2 + \beta^{-1} \sum_{\mathbf{x} \in \Lambda} \lambda(\mathbf{x}) \ln[2 \cosh(\beta \langle \mathbf{m} + \mathbf{h}, \mathbf{x} \rangle)], \qquad \mathbf{m} \in [-1, 1]^p$$

This is obtained under a spatial homogeneity condition and a strong mixing condition under translations for the random patterns which imply ergodicity. Their proof appeals to Bogoliubov's approximating Hamiltonian method and it is claimed that the ergodicity properties are sufficient to extend this method to the random case. The same functional is obtained in ref. 6 using large-deviation techniques, assuming the ξ_i^μ are i.i.d. variables with symmetric distribution. The same assumption is made in refs. 5 and 7; presumably the large-deviation techniques of these papers extend to cover the case where only ergodicity is assumed (as claimed by a referee in the case of ref. 7). We obtain a proof of (2.6) using only the ergodicity condition (2.2) by studying the relationship between F^{PF} and F.

For $\mathbf{m} \in [-1, 1]^p$, we define a family $\Phi[\mathbf{m}] \in \Gamma_\lambda$ by

$$\langle \sigma \rangle_{\Phi[\mathbf{m}]_{\mathbf{x}}} = \tanh(\beta \langle \mathbf{m} + \mathbf{h}, \mathbf{x} \rangle), \qquad \mathbf{x} \in \Lambda$$

A straightforward computation using the formula

$$\eta\left(\frac{1 + \tanh(u)}{2}\right) = \ln(2 \cosh(u)) - u \tanh(u)$$

gives

$$F(\Phi[\mathbf{m}]) = F^{\mathrm{PF}}(\mathbf{m}) + (1/2)\,\|\mathbf{m} - \mathbf{m}_\lambda(\Phi[\mathbf{m}])\|^2 \tag{2.7}$$

The first derivative of F^{PF} is

$$(\nabla F^{\mathrm{PF}})(\mathbf{m}) = -\mathbf{m} + \mathbf{m}_\lambda(\Phi[\mathbf{m}])$$

Thus, $\mathbf{m}_o$ is a critical point of F^{PF} iff

$$\mathbf{m}_o = \mathbf{m}_\lambda(\Phi[\mathbf{m}_o]) \tag{2.8}$$

Notice that if $\mathbf{h} = \mathbf{0}$, then $\mathbf{m} = \mathbf{0}$ is a critical point.

Computing the derivative of F with respect to $\langle \sigma \rangle_{\Phi_{\mathbf{x}}}$, one finds that Φ^o is a critical point of F iff [cf. (2.5)]

$$\langle \sigma \rangle_{\Phi[\mathbf{m}_\lambda(\Phi^o)]_{\mathbf{x}}} = \langle \sigma \rangle_{\Phi^o_{\mathbf{x}}} \qquad \text{for every} \quad \mathbf{x} \in \Lambda \Leftrightarrow \Phi^o = \Phi[\mathbf{m}_\lambda(\Phi^o)] \tag{2.9}$$

The following result is immediate.

Proposition 2. The maps

$$\mathbf{m} \mapsto \Phi[\mathbf{m}], \qquad \Phi \mapsto \mathbf{m}_\lambda(\Phi)$$

restricted to the critical points of F^{PF} and of F, respectively, are each other's inverses, and define a bijective correspondence between critical points of F^{PF} and critical points of F. One has

$$F^{\mathrm{PF}}(\mathbf{m}) = F(\Phi[\mathbf{m}]), \qquad F^{\mathrm{PF}}(\mathbf{m}_\lambda(\Phi)) = F(\Phi)$$

if $\mathbf{m}$ (resp. Φ) is a critical point of F^{PF} (resp. F). In particular, the maximizers of F^{PF} and F are in bijective correspondence and the maximal value of the functionals is equal.

The following result collects a number of general properties of the limiting free energy density.

Proposition 3. Let $\mathscr{S} = \{1, -1\}$.

1. $0 < \beta \mapsto f(\beta, \mathbf{h})$ is strictly convex and strictly decreasing. $0 < \beta \mapsto \beta f(\beta, \mathbf{0})$ is nondecreasing. Moreover, $f(\beta, \mathbf{0}) \geqslant f^{\mathrm{cw}}(\beta, 0; 1)$.

2. $\mathbf{h} \mapsto f(\beta, \mathbf{h})$ is convex and invariant under inversion. If π is a permutation together with arbitrary changes of sign of the components of vectors in $\mathbb{R}^p$, and $\lambda(\{\pi(\mathbf{x})\}) = \lambda(\{\mathbf{x}\})$ holds for every $\mathbf{x} \in \Lambda$, then $f(\beta, \pi(\mathbf{h})) = f(\beta, \mathbf{h})$.

3. For all $\beta > 0$ and $\mathbf{h} \in \mathbb{R}^p$ one has

$$\beta^{-1} \ln(2) \leqslant \max_{\mu \in \{1,2,\dots,p\}} f^{\mathrm{cw}}(\beta, \langle \mathbf{J}(\mu), \mathbf{h} \rangle; \|\mathbf{J}(\mu)\|^2)$$
$$\leqslant f(\beta, \mathbf{h})$$
$$\leqslant \sum_{\mathbf{x} \in \Lambda} \lambda(\{\mathbf{x}\}) f^{\mathrm{cw}}(\beta, \langle \mathbf{x}, \mathbf{h} \rangle; 2^p \lambda(\{\mathbf{x}\})) \tag{2.10}$$

where

$$J^\nu(\mu) \stackrel{\mathrm{def}}{=} \sum_{\{\mathbf{x} \in \mathcal{X}: x^\mu = 1\}} x^\nu(\lambda(\{\mathbf{x}\}) + \lambda(\{-\mathbf{x}\})), \qquad \nu = 1, 2, \dots, p$$

4. Suppose $\beta \leqslant p^{-1}$. There exists a unique family $\Phi^o \in \Gamma_\lambda$ maximizing (2.4). When $\mathbf{h} = \mathbf{0}$, Φ^o is the constant family taking the value ω on Λ; and one has $\mathbf{m}_\lambda(\Phi^o) = \mathbf{0}$ and $f(\beta, \mathbf{0}) = \beta^{-1} \ln(2)$. Moreover, if $\beta < p^{-1}$, then:

(a) $\mathbf{h} \mapsto f(\beta, \mathbf{h})$ is differentiable with

$$\mathbf{m}(\beta, \mathbf{h}) \stackrel{\mathrm{def}}{=} \lim_{N \to \infty} \mathbf{m}_N(\beta, \mathbf{h}, \{\xi\}) = (\nabla_h f)(\beta, \mathbf{h}) = \mathbf{m}_\lambda(\Phi^o)$$

the limit existing (almost surely) under condition (2.2) (almost surely).

(b) $\beta \mapsto f(\beta, \mathbf{h})$ is differentiable, and if (2.2) holds (almost surely) the limiting entropy density exists (almost surely) and equals $S_\lambda(\Phi^o) = -\beta^2(\partial f/\partial \beta)(\beta, \mathbf{h})$.

5. If $\beta_1 < \beta_2$ and $f(\beta_2, \mathbf{0}) = \beta_2^{-1} \ln(2)$, then $f(\beta_1, \mathbf{0}) = \beta_1^{-1} \ln(2)$. The set $\{\beta > 0$: $f(\beta, \mathbf{0}) = \beta^{-1} \ln(2)\}$ is not empty and equal to the interval $(0, \beta_c]$, where $p^{-1} \leqslant \beta_c \leqslant 1$.

Remark. The bound p^{-1} on the critical value of β is attained in special cases: for $p = 2$ when λ is supported by two points (see Section 4), and for general p when $\mathrm{Prob}\{\xi_i^\mu = 1\}$ is either 0 or 1 (see Section 5).

The bounds of (2.10) suffice to compute f when $\mathbf{h} = \mathbf{0}$ and $\beta 2^p \lambda(\{\mathbf{x}\}) \leqslant 1$, for all $\mathbf{x} \in \Lambda$. It is known that $f^{\mathrm{cw}}(\beta, 0; J) = \beta^{-1} \ln(2)$ for $\beta J \leqslant 1$; thus under the condition on β and λ, the upper bound of (2.10) gives $f(\beta, \mathbf{0}) = \beta^{-1} \ln(2)$. This happens, for example, when λ is equidistributed on Λ and $\beta \leqslant 2^{-p} |\Lambda|$ ($\leqslant 1$).

The lower bound

$$\max_{\mu \in \{1,2,\dots,p\}} f^{\mathrm{cw}}(\beta, \langle \mathbf{J}(\mu), \mathbf{h} \rangle; \|\mathbf{J}(\mu)\|^2)$$

is exact if $\mathbf{h}$ has only one nonzero component and λ is uniform, leading to the value $f^{\mathrm{cw}}(\beta, \|\mathbf{h}\|; 1)$ for f.

The following notation will be used throughout. For $\mu = 1, 2, \ldots, p$ we define the subsets $\mathscr{A}_\mu \subset \mathscr{X}$ by

$$\mathscr{A}_\mu = \{\mathbf{x} \in \mathscr{X}: x^\mu = 1\}$$

For a subset $\mathscr{A}$ of $\mathscr{X}$ we write $-\mathscr{A}$ for the subset obtained by changing the sign of all the members of $\mathscr{A}$. We have $\mathscr{A}_\mu \cap (-\mathscr{A}_\mu) = \varnothing$ and $\mathscr{A}_\mu \cup (-\mathscr{A}_\mu) = \mathscr{X}$.

Proof. 1. To prove the strict convexity and the strict decrease of $f(\cdot, \mathbf{h})$, we use the fact that for any family Φ maximizing (2.4) one has $S_\lambda(\Phi) > 0$. This follows [see the statements about $S_\lambda(\Phi)$ made after the Theorem] from Proposition 1.

Now suppose $0 < \beta_1 < \beta_2$ and Φ_{β_2} is a maximizing family for F_{β_2}. Then,

$$\begin{aligned} f(\beta_2, \mathbf{h}) &= F_{\beta_2}(\Phi_{\beta_2}) = F_{\beta_1}(\Phi_{\beta_2}) - \beta_1^{-1}\beta_2^{-1}(\beta_2 - \beta_1)\, S_\lambda(\Phi_{\beta_2}) \\ &< F_{\beta_1}(\Phi_{\beta_2}) \leqslant f(\beta_1, \mathbf{h}) \end{aligned}$$

If $\beta_0 \neq \beta_1$ and $0 < \alpha < 1$, let $\beta_\alpha = \alpha\beta_1 + (1-\alpha)\,\beta_0$, and consider a maximizing family Φ_{β_α} for F_{β_α}. Then we have

$$\begin{aligned} f(\beta_\alpha, \mathbf{h}) &\leqslant \alpha f(\beta_1, \mathbf{h}) + (1-\alpha)\, f(\beta_0, \mathbf{h}) \\ &\quad - \alpha(1-\alpha)(\beta_1 - \beta_0)^2\, (\beta_1\beta_0\beta_\alpha)^{-1}\, S_\lambda(\Phi_{\beta_\alpha}) \end{aligned}$$

Suppose $\beta_1 < \beta_2$, and let Φ be a maximizing family for F_{β_1}. Then

$$\beta_2 f(\beta_2, \mathbf{0}) \geqslant \beta_2 F_{\beta_2}(\Phi) = \tfrac{1}{2}(\beta_2 - \beta_1)\, \|\mathbf{m}_\lambda(\Phi)\|^2 + \beta_1 F_{\beta_1}(\Phi) \geqslant \beta_1 f(\beta_1, \mathbf{0})$$

Finally, notice that for $\mathbf{h} = \mathbf{0}$ the Hamiltonian (1.1) is bounded above by the Hamiltonian $H^{\{\xi\}}_{N, p=1}$, which is equivalent to the Hamiltonian (1.2) of the Curie–Weiss model at zero field and unit interaction by the change of variables $S_i \mapsto \xi_i^1 S_i$. Thus, $f(\beta, \mathbf{0}) \geqslant f^{\mathrm{cw}}(\beta, 0; 1)$ and, since 1 is the critical temperature of this Curie–Weiss model, we deduce $\beta_c \leqslant 1$.

2. The convexity of $f(\beta, \cdot)$ is obvious from (2.4). Suppose π is a permutation together with arbitrary changes of sign of the components of vectors in $\mathbb{R}^p$. For a family Φ let $\tilde{\Phi}$ be defined by $\tilde{\phi}_\mathbf{x} = \phi_{\pi(\mathbf{x})}$. Notice that π maps $\mathscr{X}$ bijectively onto itself. Then if $\mathbf{x}$ and $\pi(\mathbf{x})$ have the same (nonzero) measure for all $x \in \Lambda$, we get $\mathbf{m}_\lambda(\tilde{\Phi}) = \pi^{-1}(\mathbf{m}_\lambda(\Phi))$. Since $S(\hat{\phi}) = S(\phi)$, the same assumption on the measure gives $S_\lambda(\tilde{\Phi}) = S_\lambda(\Phi)$; thus $F(\tilde{\Phi}, \mathbf{h}) = F(\Phi, \pi(\mathbf{h}))$.

3. The upper bound follows from (2.4) by using the Lemma of the Appendix in the case $\mathscr{A} = \mathscr{X}$ to estimate $\|\mathbf{m}_\lambda(\Phi)\|^2$ for any family Φ.

Our best lower bound on f is obtained by computing F for any family Φ which is reflexive and constant on $\mathscr{A}_\mu$, where it takes the value ϕ. One obtains $m_\lambda^\nu(\Phi)=\langle\sigma\rangle_\phi J^\nu(\mu)$ and $F(\Phi)=g(\beta, \langle \mathbf{J}(\mu), \mathbf{h}\rangle, \langle\sigma\rangle_\phi; \|\mathbf{J}(\mu)\|^2)$.

4. This proceeds by applying the contraction principle. Define a map T on Γ_λ by $\langle\sigma\rangle_{T(\Phi)_\mathbf{x}}=\tanh(\beta\langle \mathbf{m}_\lambda(\Phi)+\mathbf{h}, \mathbf{x}\rangle)$, recalling that a probability measure on $\mathscr{S}$ is uniquely determined by the expectation of σ. Now, T maps Γ_λ into itself because the range of tanh is $(-1, 1)$. Writing the tanh as the integral of $\cosh^{-2}$ and using $\cosh(\cdot)\geqslant 1$, we obtain

$$|\langle\sigma\rangle_{T(\Phi)_\mathbf{x}}-\langle\sigma\rangle_{T(\Psi)_\mathbf{x}}|\leqslant \beta p\, d(\Phi, \Psi) \qquad \text{for all} \quad \mathbf{x}\in\Lambda$$

Thus

$$d(T(\Phi), T(\Psi))\leqslant \beta p\, d(\Phi, \Psi) \tag{2.11}$$

with equality iff $\Phi=\Psi$. Thus (ref. 20, Problem 1, p. 267) the gap equations have a unique solution Φ^o when $\beta p\leqslant 1$. One also concludes that

$$\Phi^o=\lim_{n\to\infty} T^n(\Phi) \tag{2.12}$$

for every $\Phi\in\Gamma_\lambda$.

By Proposition 1, the unique solution of (2.5) existing for $\beta p\leqslant 1$ gives a maximum of F. Thus, (2.4) has a unique maximizing family Φ^o for $\beta p\leqslant 1$. An alternative direct proof of this statement proceeds by showing that $\Gamma_\lambda\ni\Phi\mapsto F(\Phi)$ is strictly concave. The Hessian of F is $B-\beta^{-1}D$ with $B_{\mathbf{x},\mathbf{y}}=\lambda(\mathbf{x})\,\lambda(\mathbf{y})\langle\mathbf{x}, \mathbf{y}\rangle$, and

$$D_{\mathbf{x},\mathbf{y}}=\delta_{\mathbf{x},\mathbf{y}}\lambda(\mathbf{x})(1-\langle\sigma\rangle_{\phi_\mathbf{x}}^2)^{-1}\geqslant\delta_{\mathbf{x},\mathbf{y}}\lambda(\mathbf{x})\overset{\text{def}}{=}D^o_{\mathbf{x},\mathbf{y}}$$

It follows that the quadratic form associated with the Hessian is bounded above by the quadratic form associated to $B-\beta^{-1}D^o$, which is independent of Φ. Moreover, for β sufficiently small, the matrix $D^o-\beta B$ is positive definite.

If $\Omega\in\Gamma_\lambda$ denotes the constant family on Λ with value ω, then $\mathbf{m}_\lambda(\Omega)=\mathbf{0}$, so that Ω is indeed the only maximizer for $\mathbf{h}=\mathbf{0}$. The formula for $f(\beta, \mathbf{0})$ follows by evaluating $F(\Omega)$ using $S(\omega)=\ln(2)$.

To prove differentiability of $f(\beta, \cdot)$, we will have to work with different $\mathbf{h}$'s and thus use an index $\mathbf{h}$ to distinguish the various quantities. The same estimate as before leads to

$$d(T_\mathbf{h}(\Phi), T_\mathbf{k}(\Phi))\leqslant\beta p\, \|\mathbf{h}-\mathbf{k}\|_\infty \tag{2.13}$$

where $\|\mathbf{h}\|_\infty=\max_\mu |h^\mu|$, and the inequality is strict as soon as $\mathbf{h}\neq\mathbf{k}$.

By the triangle inequality, (2.11), and (2.13), we have

$$d(T_{\mathbf{h}}(\Phi), T_{\mathbf{k}}(\Psi)) \leqslant \beta p \, \|\mathbf{h}-\mathbf{k}\|_{\infty} + \beta p \, d(\Phi, \Psi) \tag{2.14}$$

We write $\Phi^o(\mathbf{h})$ for the unique maximizer corresponding to the field vector $\mathbf{h}$. For any integer n and family Φ we have by the triangle inequality and (2.14)

$$d(\Phi^o(\mathbf{h}), \Phi^o(\mathbf{k})) \leqslant d(\Phi^o(\mathbf{h}), T^n_{\mathbf{h}}(\Phi)) + d(T^n_{\mathbf{k}}(\Phi), \Phi^o(\mathbf{k})) + \frac{\beta p(1-\beta^n p^n)}{1-\beta p} \|\mathbf{h}-\mathbf{k}\|_{\infty}$$

Upon taking the limit $n \to \infty$ using (2.12), we conclude that

$$d(\Phi^o(\mathbf{h}), \Phi^o(\mathbf{k})) \leqslant \frac{\beta p}{1-\beta p} \|\mathbf{h}-\mathbf{k}\|_{\infty} \tag{2.15}$$

A direct estimate on (2.3) using (2.15) now produces

$$\|\mathbf{m}_{\lambda}(\Phi^o(\mathbf{h})) - \mathbf{m}_{\lambda}(\Phi^o(\mathbf{k}))\| \leqslant \frac{\beta p^{3/2}}{1-\beta p} \|\mathbf{h}-\mathbf{k}\|_{\infty} \tag{2.16}$$

We can now prove differentiability of $f(\beta, \cdot)$ as follows. Let $\mathbf{e}$ be an arbitrary unit vector in $\mathbb{R}^p$. By virtue of convexity of $f(\beta, \cdot)$, the left and right derivatives $(f_{\mathbf{e}})_{\pm}$ in direction $\mathbf{e}$ exist and satisfy

$$\begin{aligned} &\varepsilon^{-1}[f(\beta, \mathbf{h}) - f(\beta, \mathbf{h}-\varepsilon\mathbf{e})] \\ &\qquad \leqslant (f_{\mathbf{e}})_- \leqslant (f_{\mathbf{e}})_+ \leqslant \varepsilon^{-1}[f(\beta, \mathbf{h}+\varepsilon\mathbf{e}) - f(\beta, \mathbf{h})] \end{aligned} \tag{2.17}$$

for all $\varepsilon > 0$. Since $\Phi^o(\mathbf{k})$ is the (unique) maximizer of $F_{\mathbf{k}}$, we can estimate the r.h.s. of this inequality

$$\begin{aligned} &\varepsilon^{-1}[f(\beta, \mathbf{h}+\varepsilon\mathbf{e}) - f(\beta, \mathbf{h})] \\ &\quad = \varepsilon^{-1}[F_{\mathbf{h}+\varepsilon\mathbf{e}}(\Phi^o(\mathbf{h}+\varepsilon\mathbf{e})) - F_{\mathbf{h}}(\Phi^o(\mathbf{h}))] \\ &\quad \leqslant \varepsilon^{-1}[F_{\mathbf{h}+\varepsilon\mathbf{e}}(\Phi^o(\mathbf{h}+\varepsilon\mathbf{e})) - F_{\mathbf{h}}(\Phi^o(\mathbf{h}+\varepsilon\mathbf{e}))] \\ &\quad = \varepsilon^{-1}[\langle \mathbf{h}+\varepsilon\mathbf{e}, \mathbf{m}_{\lambda}(\Phi^o(\mathbf{h}+\varepsilon\mathbf{e}))\rangle - \langle \mathbf{h}, \mathbf{m}_{\lambda}(\Phi^o(\mathbf{h}+\varepsilon\mathbf{e}))\rangle] \\ &\quad = \langle \mathbf{e}, \mathbf{m}_{\lambda}(\Phi^o(\mathbf{h}+\varepsilon\mathbf{e}))\rangle \end{aligned}$$

Proceeding analogously with the l.h.s. of (2.17), we get

$$\varepsilon^{-1}[f(\beta, \mathbf{h}) - f(\beta, \mathbf{h}-\varepsilon\mathbf{e})] \geqslant \langle \mathbf{e}, \mathbf{m}_{\lambda}(\Phi^o(\mathbf{h}-\varepsilon\mathbf{e}))\rangle$$

Thus,

$$\langle \mathbf{e}, \mathbf{m}_\lambda(\Phi''(\mathbf{h}-\varepsilon\mathbf{e}))\rangle \leqslant (f_{\mathbf{e}})_- \leqslant (f_{\mathbf{e}})_+ \leqslant \langle \mathbf{e}, \mathbf{m}_\lambda(\Phi''(\mathbf{h}+\varepsilon\mathbf{e}))\rangle$$

and $(f_{\mathbf{e}})_- = (f_{\mathbf{e}})_+ = \langle \mathbf{e}, \mathbf{m}_\lambda(\Phi''(\mathbf{h}))\rangle$ follows upon taking the limit $\varepsilon \downarrow 0$ using (2.16).

The claim about the limit of the overlap parameter vector follows by the fact that the limit of derivatives of a sequence of differentiable convex functions exists and is equal to the derivative of the limiting function when this function exists and is differentiable.

The proof of differentiability of $f(\cdot, \mathbf{h})$ is analogous.

5. Since $f(\beta, \mathbf{h}) \geqslant \beta^{-1}\ln(2)$ by part 3 of the proposition, the first claim follows from the nondecreasing property of $\beta \mapsto \beta f(\beta, \mathbf{0})$ (part 1 of the proposition). The rest of the claims follow from parts 1 and 4. ∎

3. THE CASE OF UNIFORM λ

The limiting measure λ is uniform, i.e., $\lambda(\{\mathbf{x}\}) = 2^{-p}$, when, for example, the ξ_i^μ are i.i.d. with symmetric distribution. The following result shows that the upper bound of (2.10) is attained when λ is uniform, and at most two components of $\mathbf{h}$ are nonzero.

Let $\phi^{\text{cw}}(\beta, h)$ be the unique probability measure on $\mathscr{S}$ such that $\langle\sigma\rangle_{\phi^{\text{cw}}(\beta,h)} = m^{\text{cw}}(\beta, h; 1)$.

Proposition 4. Let $\mathscr{S} = \{1, -1\}$ and suppose λ is uniform and $\mathbf{h}$ has at most two nonzero components, say h^{μ_1} and h^{μ_2}; then

$$f(\beta, \mathbf{h}) = \tfrac{1}{2} f^{\text{cw}}(\beta, h^{\mu_1} + h^{\mu_2}; 1) + \tfrac{1}{2} f^{\text{cw}}(\beta, h^{\mu_1} - h^{\mu_2}; 1) \tag{3.1}$$

In particular, $f(\beta, \mathbf{0}) = f^{\text{cw}}(\beta, 0; 1)$. Now, $0 < \beta \mapsto f(\beta, \mathbf{h})$ is differentiable. If Φ and Ψ are maximizers of (2.4), then

$$S_\lambda(\Phi) = S_\lambda(\Psi) = -\beta^2 \frac{\partial f}{\partial \beta}(\beta, \mathbf{h})$$

If (2.2) holds true (almost surely) the limiting entropy density exists (almost surely) and is equal to $-\beta^2(\partial f/\partial\beta)(\beta, \mathbf{h})$.

Moreover:

1. If $\beta > 1$ (resp. $\beta \leqslant 1$), then for $|h^{\mu_1}| \neq |h^{\mu_2}|$ (resp. arbitrary h^{μ_1} and h^{μ_2}), one has:

(a) There is a unique family $\Phi^o \in \Gamma_\lambda$ maximizing (2.4) given by $\phi^o_{\mathbf{x}} = \phi^{\mathrm{cw}}(\beta, \varepsilon_1 h^{\mu_1} + \varepsilon_2 h^{\mu_2})$ for $\mathbf{x} \in (\varepsilon_1 \mathscr{A}_{\mu_1}) \cap (\varepsilon_2 \mathscr{A}_{\mu_2})$, $\varepsilon_j \in \{\pm\}$. One has $m^\kappa_\lambda(\Phi^o) = 0$ for κ distinct from μ_1 and μ_2, and

$$m^{\mu_1}_\lambda(\Phi^o) = (1/2)(m^{\mathrm{cw}}(\beta, h^{\mu_1} + h^{\mu_2}; 1) + m^{\mathrm{cw}}(\beta, h^{\mu_1} - h^{\mu_2}; 1))$$

$$m^{\mu_2}_\lambda(\Phi^o) = (1/2)(m^{\mathrm{cw}}(\beta, h^{\mu_1} + h^{\mu_2}; 1) + m^{\mathrm{cw}}(\beta, h^{\mu_2} - h^{\mu_1}; 1))$$

(b) $(h^{\mu_1}, h^{\mu_2}) \mapsto f(\beta, \mathbf{h})$ is differentiable and

$$m^{\mu_j}(\beta, \mathbf{h}) \stackrel{\text{def}}{=} \lim_{N \to \infty} m^{\mu_j}_N(\beta, \mathbf{h}; \{\xi\}) = \left(\frac{\partial f}{\partial h^{\mu_j}}\right)(\beta, \mathbf{h}) = m^{\mu_j}_\lambda(\Phi^o)$$

the limits existing (almost surely) under condition (2.2) (almost surely) for $j = 1, 2$.

2. If $\beta > 1$, then:

(a) $(h^{\mu_1}, h^{\mu_2}) \mapsto f(\beta, \mathbf{h})$ is not differentiable on the two lines $|h^{\mu_1}| = |h^{\mu_2}|$.

(b) For $|h^{\mu_1}| = |h^{\mu_2}| \neq 0$ (resp. for $h^{\mu_1} = h^{\mu_2} = 0$) there are $2^{2^{(p-2)}}$ (resp. $2^{2^{(p-1)}}$) families Φ maximizing (2.4). For each of these one has $m^\kappa_\lambda(\Phi) = 0$ for $\mu_1 \neq \kappa \neq \mu_2$ and the vector $(m^{\mu_1}_\lambda(\Phi), m^{\mu_2}_\lambda(\Phi))$ takes two (resp. four) values.

The families of part 2(b) and the values of $\mathbf{m}_\lambda$ can be given explicitly in terms of ϕ^{cw} and m^{cw}. For uniform λ and at most two nonzero field components Proposition 4 provides a rather complete picture. For subcritical temperatures (i.e., $\beta > 1$) the subgradients of f on the two lines $|h^{\mu_1}| = |h^{\mu_2}|$ can be described explicitly to obtain the phase diagram of Fig. 1. There is a jump discontinuity in the μ_1th and μ_2th limiting overlap parameters on each line as indicated in the figure. The (h^{μ_1}, h^{μ_2}) plane splits into the four open cones delimited by the two lines; in each cone the limiting behavior of the overlap parameters as $\mathbf{h} = (h^{\mu_1}, h^{\mu_2}) \to (0, 0)$ for $\beta > 1$ is different:

$$\lim_{\mathbf{h} \to 0;\, 0 < |h^{\mu_2}| < h^{\mu_1}} (m^{\mu_1}(\beta, \mathbf{h}), m^{\mu_2}(\beta, \mathbf{h})) = (m^{\mathrm{cw}}(\beta, 0^+; 1), 0)$$

$$\lim_{\mathbf{h} \to 0;\, 0 < |h^{\mu_1}| < h^{\mu_2}} (m^{\mu_1}(\beta, \mathbf{h}), m^{\mu_2}(\beta, \mathbf{h})) = (0, m^{\mathrm{cw}}(\beta, 0^+; 1))$$

$$\lim_{\mathbf{h} \to 0;\, 0 < |h^{\mu_2}| < -h^{\mu_1}} (m^{\mu_1}(\beta, \mathbf{h}), m^{\mu_2}(\beta, \mathbf{h})) = (m^{\mathrm{cw}}(\beta, 0^-; 1), 0)$$

$$\lim_{\mathbf{h} \to 0;\, 0 < |h^{\mu_1}| < -h^{\mu_2}} (m^{\mu_1}(\beta, \mathbf{h}), m^{\mu_2}(\beta, \mathbf{h})) = (0, m^{\mathrm{cw}}(\beta, 0^-; 1))$$

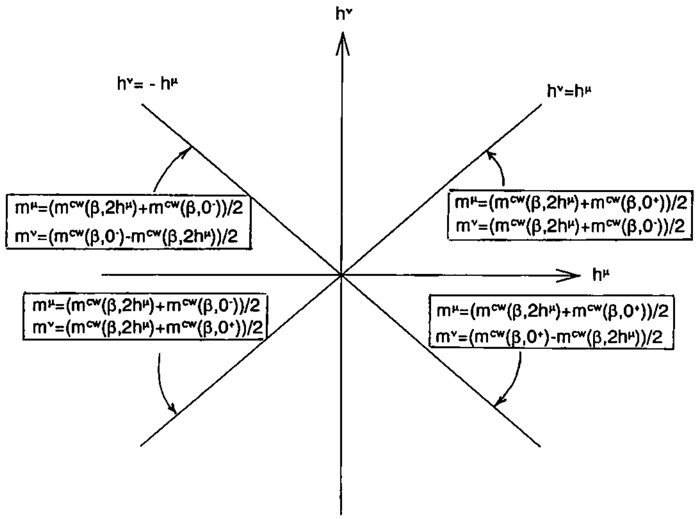

Fig. 1. Phase diagram for $\beta > 1$ in the i.i.d. case with symmetric distribution and two nonzero components h^{μ}, h^{ν} of $\mathbf{h}$.

In simple words, as $\mathbf{h} \to \mathbf{0}$, $\mathbf{m}$ follows the behavior dictated by the field component with maximal modulus.

The simple picture that emerges for uniform λ and two nonzero components is due to the fact that f is given by the upper bound of (2.10) where all $\langle \sigma \rangle_{\phi_{\mathbf{x}}}$ are decoupled. This is quantitatively wrong for three or more nonzero field components. If λ is uniform, $p = 3$, and all three components of $\mathbf{h}$ are nonzero; then for all $\beta > 0$

$$f(\beta, \mathbf{h}) < 2^{-p} \sum_{\mathbf{x} \in \mathscr{X}} f^{\mathrm{cw}}(\beta, \langle \mathbf{x}, \mathbf{h} \rangle; 1) \tag{3.2}$$

except for some particular nonzero values of the components.

Proof. Suppose $p \geqslant 2$, that λ is uniform, and that $\mathbf{h}$ has at most two nonzero components. Due to the permutational symmetry, we may assume that h^1 and h^2 are the only components which can be nonzero. Since $\Lambda = -\Lambda = \mathscr{X}$, we may (due to Proposition 1) obtain the maximal value of F by varying over reflexive families. Such a family Φ is specified uniquely by its values on $\mathscr{A}_1$, since $\phi_{\mathbf{x}} = \widehat{\phi_{-\mathbf{x}}}$ for $\mathbf{x} \in -\mathscr{A}_1$. We partition $\mathscr{A}_1$ into two sets $\mathscr{B}_1 = \mathscr{A}_1 \cap \mathscr{A}_2$ and $\mathscr{B}_2 = \mathscr{A}_1 \cap (-\mathscr{A}_2)$ of equal measure $2^{-(p-2)}$. The proof proceeds in four steps. We drop the index λ in $\mathbf{m}_{\lambda}$.

First step. If the family Φ is reflexive and has the additional property of being constant on each of the two sets $\mathscr{B}_1$ and $\mathscr{B}_2$, then, letting $\bar{\phi}_j$ be the constant value of $\phi_{\mathbf{x}}$ on $\mathscr{B}_j$ $(j=1,2)$, we have

$$m^1(\Phi)=\tfrac{1}{2}(\langle\sigma\rangle_{\bar{\phi}_1}+\langle\sigma\rangle_{\bar{\phi}_2}) \tag{3.3}$$

$$m^2(\Phi)=\tfrac{1}{2}(\langle\sigma\rangle_{\bar{\phi}_1}-\langle\sigma\rangle_{\bar{\phi}_2}) \tag{3.4}$$

$$m^\mu(\Phi)=0 \qquad \text{for all} \quad \mu\neq 1,2 \tag{3.5}$$

To prove this, partition $\mathscr{A}_1$ into four disjoint subsets $\mathscr{B}_1\cap\mathscr{A}_\mu$, $\mathscr{B}_1\cap(-\mathscr{A}_\mu)$, $\mathscr{B}_2\cap\mathscr{A}_\mu$, and $\mathscr{B}_2\cap(-\mathscr{A}_\mu)$.

Second step. For a given reflexive family Φ we construct a new family $\tilde{\Phi}$ which is reflexive and constant on $\mathscr{B}_j$ $(j=1,2)$. This involves two probability measures ρ_j on $\mathscr{B}_j$ $(j=1,2)$. We obtain a lower bound on $F(\tilde{\Phi})-F(\Phi)$ and then choose these measures in order to maximize this lower bound. Define $\tilde{\Phi}$ by putting

$$\tilde{\phi}_{\mathbf{x}}=\sum_{\mathbf{y}\in\mathscr{B}_j}\rho_j(\{\mathbf{y}\})\,\phi_{\mathbf{y}}=:\psi_j \qquad \text{if} \quad \mathbf{x}\in\mathscr{B}_j$$

for $\mathbf{x}\in\mathscr{A}_1$ and extending reflexively, i.e., $\tilde{\phi}_{\mathbf{x}}=\widehat{\tilde{\phi}_{-\mathbf{x}}}$ for $\mathbf{x}\in(-\mathscr{A}_1)$, that is, $\tilde{\phi}_{\mathbf{x}}=\widehat{\psi_j}$ for $\mathbf{x}\in-\mathscr{B}_j$. The family $\tilde{\Phi}$ is reflexive and constant on each of the subsets $\mathscr{B}_1$, $\mathscr{B}_2$, $-\mathscr{B}_1$, and $-\mathscr{B}_2$. By the first step, $\mathbf{m}(\tilde{\Phi})$ is given by Eqs. (3.3)–(3.5). The entropy part of F for $\tilde{\Phi}$ is

$$\tfrac{1}{2}S\Big(\sum_{\mathbf{y}\in\mathscr{B}_1}\rho_1(\{\mathbf{y}\})\,\phi_{\mathbf{y}}\Big)+\tfrac{1}{2}S\Big(\sum_{\mathbf{y}\in\mathscr{B}_2}\rho_2(\{\mathbf{y}\})\,\phi_{\mathbf{y}}\Big)$$

It follows that

$$F(\tilde{\Phi})=\tfrac{1}{2}g(\beta,h^1+h^2,\langle\sigma\rangle_{\psi_1})+\tfrac{1}{2}g(\beta,h^1-h^2,\langle\sigma\rangle_{\psi_2}) \tag{3.6}$$

where for real h and $\beta>0$, $g(\beta,h,\cdot)=g(\beta,h,\cdot;1)$ is the function on $[-1,1]$ defined by (1.3) for $J=1$.

We now compute $F(\Phi)$. Using reflexivity and $S(\hat{\phi})=S(\phi)$, we find for the entropy part $2^{-(p-1)}(\sum_{\mathbf{x}\in\mathscr{B}_1}S(\phi_{\mathbf{x}})+\sum_{\mathbf{x}\in\mathscr{B}_2}S(\phi_{\mathbf{x}}))$.

Now, by reflexivity, $m^\mu(\Phi)=2^{-(p-1)}\sum_{\mathbf{x}\in\mathscr{A}_1}x^\mu\langle\sigma\rangle_{\phi_{\mathbf{x}}}$, so, using the partition $\mathscr{A}_1=\mathscr{B}_1\cup\mathscr{B}_2$, we get

$$m^{1(2)}(\Phi)=2^{-(p-1)}\Big(\sum_{\mathbf{x}\in\mathscr{B}_1}\langle\sigma\rangle_{\phi_{\mathbf{x}}}+(-)\sum_{\mathbf{x}\in\mathscr{B}_2}\langle\sigma\rangle_{\phi_{\mathbf{x}}}\Big)$$

Thus, using the Lemma of the Appendix (with $\mathscr{A} = \mathscr{A}_1$) to estimate the sum of quadratic terms in F, we get

$$F(\Phi) \leqslant \tfrac{1}{2} \sum_{\mathbf{x} \in \mathscr{B}_1} 2^{-(p-2)} g(\beta, h^1 + h^2, \langle \sigma \rangle_{\phi_\mathbf{x}})$$
$$+ \tfrac{1}{2} \sum_{\mathbf{x} \in \mathscr{B}_2} 2^{-(p-2)} g(\beta, h^1 - h^2, \langle \sigma \rangle_{\phi_\mathbf{x}})$$

Our bound is then

$$F(\tilde{\Phi}) - F(\Phi) \geqslant \tfrac{1}{2} \left[g(\beta, h^1 + h^2, \langle \sigma \rangle_{\psi_1}) - 2^{-(p-2)} \sum_{\mathbf{x} \in \mathscr{B}_1} g(\beta, h^1 + h^2, \langle \sigma \rangle_{\phi_\mathbf{x}}) \right]$$
$$+ \tfrac{1}{2} \left[g(\beta, h^1 - h^2, \langle \sigma \rangle_{\psi_2}) \right.$$
$$\left. - 2^{-(p-2)} \sum_{\mathbf{x} \in \mathscr{B}_2} g(\beta, h^1 - h^2, \langle \sigma \rangle_{\phi_\mathbf{x}}) \right] \tag{3.7}$$

We now choose the measures ρ_j in order to make the r.h.s. of this inequality as large as possible. By definition, the value $\langle \sigma \rangle_{\psi_j}$ lies in the interval

$$[\min_{\mathbf{x} \in \mathscr{B}_j} \{\langle \sigma \rangle_{\phi_\mathbf{x}}\}, \max_{\mathbf{x} \in \mathscr{B}_j} \{\langle \sigma \rangle_{\phi_\mathbf{x}}\}]$$

Let t_j be a number in this interval that maximizes $g(\beta, h^1 + h^2, \cdot)$ for $j = 1$ and $g(\beta, h^1 - h^2, \cdot)$ for $j = 2$. Choose ρ_j such that $\langle \sigma \rangle_{\psi_j} = t_j$. Then the averages over $\mathscr{B}_j$ in the r.h.s. of (3.7) are not greater than $g(\beta, h^1 + h^2, t_1) = g(\beta, h^1 + h^2, \langle \sigma \rangle_{\psi_1})$ and $g(\beta, h^1 - h^2, t_2) = g(\beta, h^1 - h^2, \langle \sigma \rangle_{\psi_2})$, respectively. For this choice of measures, the r.h.s. of (3.7) is nonnegative and as large as possible.

Third step. By the first and second steps, we conclude that if λ is uniform and only the components h^{μ_1} and h^{μ_2} of $\mathbf{h}$ can be nonzero, then the maximum of F over Γ_λ is equal to the maximum of F over the reflexive families which are constant on $\mathscr{B}_1 = \mathscr{A}_{\mu_1} \cap \mathscr{A}_{\mu_2}$ and on $\mathscr{B}_2 = \mathscr{A}_{\mu_1} \cap (-\mathscr{A}_{\mu_2})$. Thus, denoting by ψ_j the values on $\mathscr{B}_j$, using (3.6) and $\mathscr{B}_1 \cap \mathscr{B}_2 = \varnothing$, we get

$$f(\beta, \mathbf{h}) = \max_{\psi_1, \psi_2 \in \mathscr{M}(\mathscr{S})} \{ \tfrac{1}{2} g(\beta, h^{\mu_1} + h^{\mu_2}, \langle \sigma \rangle_{\psi_1}) + \tfrac{1}{2} g(\beta, h^{\mu_1} - h^{\mu_2}, \langle \sigma \rangle_{\psi_2}) \}$$
$$= \tfrac{1}{2} \max_{\psi \in \mathscr{M}(\mathscr{S})} g(\beta, h^{\mu_1} + h^{\mu_2}, \langle \sigma \rangle_\psi) + \tfrac{1}{2} \max_{\psi \in \mathscr{M}(\mathscr{S})} g(\beta, h^{\mu_1} - h^{\mu_2}, \langle \sigma \rangle_\psi)$$

but the respective maxima are equal to $f^{\mathrm{cw}}(\beta, h^{\mu_1} \pm h^{\mu_2}; 1)$.

It is known that the value of t in $[-1, 1]$ maximizing $g(\beta, h, \cdot)$ is unique if $h \neq 0$ or if $\beta \leqslant 1$; we denote it by $t(\beta, h)$. It thus follows that a maximizing reflexive family which is constant on $\mathscr{B}_1$ and on $\mathscr{B}_2$ is unique if $|h^{\mu_1}| \neq |h^{\mu_2}|$ or if $\beta \leqslant 1$; denote this family by Φ^o. Now suppose that Φ is a maximizing family; by Proposition 1, Φ must be reflexive. Consider the uniformization $\tilde{\Phi}$ on $\mathscr{B}_j$ of Φ constructed in the second step. We have

$$\begin{aligned} F(\Phi) \geqslant F(\tilde{\Phi}) \geqslant F(\Phi) & \\ & + \tfrac{1}{2} \Big[g(\beta, h^{\mu_1} + h^{\mu_2}, \langle \sigma \rangle_{\psi_1}) \\ & - 2^{-(p-2)} \sum_{\mathbf{x} \in \mathscr{B}_1} g(\beta, h^{\mu_1} + h^{\mu_2}, \langle \sigma \rangle_{\phi_{\mathbf{x}}}) \Big] \\ & + \tfrac{1}{2} \Big[g(\beta, h^{\mu_1} - h^{\mu_2}, \langle \sigma \rangle_{\psi_2}) \\ & - 2^{-(p-2)} \sum_{\mathbf{x} \in \mathscr{B}_2} g(\beta, h^{\mu_1} - h^{\mu_2}, \langle \sigma \rangle_{\phi_{\mathbf{x}}}) \Big] \end{aligned}$$

where both corrections in $[\cdot]$ on the r.h.s. are nonnegative. It thus follows that $F(\Phi) = F(\tilde{\Phi})$, that $\tilde{\Phi} = \Phi^o$, and that the corrections are zero. From the latter fact, we conclude (the argument is made explicit below) that Φ is constant on $\mathscr{B}_j$ $(j = 1, 2)$, and thus $\Phi = \Phi^o$.

It is known that there are two values of t in $[-1, 1]$ maximizing $g(\beta, h, \cdot)$ when $\beta > 1$ and $h = 0$; we denote them by $t(\beta, 0^{\pm})$. Suppose that $h^{\mu_1} = h^{\mu_2} \neq 0$, and let Φ be a maximizing family which must be reflexive. The above inequality now implies that the correction term is zero, i.e.,

$$g(\beta, 2h^{\mu_1}, \langle \sigma \rangle_{\psi_1}) - 2^{-(p-2)} \sum_{\mathbf{x} \in \mathscr{B}_1} g(\beta, 2h^{\mu_1}, \langle \sigma \rangle_{\phi_{\mathbf{x}}}) = 0$$

$$g(\beta, 0, \langle \sigma \rangle_{\psi_2}) - 2^{-(p-2)} \sum_{\mathbf{x} \in \mathscr{B}_2} g(\beta, 0, \langle \sigma \rangle_{\phi_{\mathbf{x}}}) = 0$$

and that $\langle \sigma \rangle_{\psi_1} = t(\beta, 2h^{\mu_1})$ and $\langle \sigma \rangle_{\psi_2} = t(\beta, 0^{\pm})$. From the first identity we conclude that $\langle \sigma \rangle_{\phi_{\mathbf{x}}}$ is constant on $\mathscr{B}_1$ and equal to $t(\beta, 2h^{\mu_1})$; this then determines uniquely the value of Φ (and $\tilde{\Phi}$) on $\mathscr{B}_1$. From the second identity we conclude that on $\mathscr{B}_2$, $\langle \sigma \rangle_{\phi_{\mathbf{x}}}$ takes any of the two values $t(\beta, 0^{\pm})$. Since there are 2^{p-2} elements in $\mathscr{B}_2$, there are $2^{2^{(p-2)}}$ possibilities. Going through the other cases $h^{\mu_1} = -h^{\mu_2} \neq 0$ and $h^{\mu_1} = h^{\mu_2} = 0$, one obtains the claims of part 2 of the proposition.

Fourth step. The rest of the claims of Proposition 4 follow from: (1) the (familiar) solution of the variational problem $\max\{g(\beta, h, t): t \in [-1, 1]\}$; (2) the known differentiability properties of $h \mapsto f^{\mathrm{cw}}(\beta, h; J)$ and of $\beta \mapsto f^{\mathrm{cw}}(\beta, h; J)$; (3) Eq. (2.1) and the fact that the limit of derivatives of a sequence of differentiable convex functions exists and is equal to the derivative of the limiting function when this function exists and is differentiable; and (4) Eqs. (3.3) and (3.4). ∎

We now give a proof of (3.2). We assume λ is uniform and $p = 3$. By Proposition 1, it suffices to consider reflexive families. A reflexive family Φ is specified by the values it takes on the four one-point sets $\mathscr{B}_1 = \mathscr{A}_1 \cap \mathscr{A}_2 \cap \mathscr{A}_3$, $\mathscr{B}_2 = \mathscr{A}_1 \cap \mathscr{A}_2 \cap (-\mathscr{A}_3)$, $\mathscr{B}_3 = \mathscr{A}_1 \cap (-\mathscr{A}_2) \cap \mathscr{A}_3$, and $\mathscr{B}_4 = \mathscr{A}_1 \cap (-\mathscr{A}_2) \cap (-\mathscr{A}_3)$. To shorten the notation, write $\langle j \rangle = \langle \sigma \rangle_{\phi_j}$, where ϕ_j is the value of Φ on $\mathscr{B}_j$. We compute

$$m^1(\Phi) = \tfrac{1}{4}(\langle 1 \rangle + \langle 2 \rangle + \langle 3 \rangle + \langle 4 \rangle)$$
$$m^2(\Phi) = \tfrac{1}{4}(\langle 1 \rangle + \langle 2 \rangle - \langle 3 \rangle - \langle 4 \rangle)$$
$$m^3(\Phi) = \tfrac{1}{4}(\langle 1 \rangle - \langle 2 \rangle + \langle 3 \rangle - \langle 4 \rangle)$$

to obtain

$$\begin{aligned} F(\Phi) = {} & \tfrac{1}{4} g(\beta, h^1 + h^2 + h^3, \langle 1 \rangle) + \tfrac{1}{4} g(\beta, h^1 + h^2 - h^3, \langle 2 \rangle) \\ & + \tfrac{1}{4} g(\beta, h^1 - h^2 + h^3, \langle 3 \rangle) + \tfrac{1}{4} g(\beta, h^1 - h^2 - h^3, \langle 4 \rangle) \\ & - \tfrac{1}{32}(\langle 1 \rangle - \langle 2 \rangle - \langle 3 \rangle + \langle 4 \rangle)^2 \end{aligned}$$

Writing $K(\Phi) = \frac{1}{32}(\langle 1 \rangle - \langle 2 \rangle - \langle 3 \rangle + \langle 4 \rangle)^2$, we then have

$$\begin{aligned} f(\beta, \mathbf{h}) & = \max_{\Phi \text{ reflexive}} F(\Phi) \\ & \leqslant \tfrac{1}{4} f^{\mathrm{cw}}(\beta, h^1 + h^2 + h^3; 1) + \tfrac{1}{4} f^{\mathrm{cw}}(\beta, h^1 + h^2 - h^3; 1) \\ & \quad + \tfrac{1}{4} f^{\mathrm{cw}}(\beta, h^1 - h^2 + h^3; 1) + \tfrac{1}{4} f^{\mathrm{cw}}(\beta, h^1 - h^2 - h^3; 1) \\ & = 2^{-3} \sum_{\mathbf{x} \in \mathscr{X}} f^{\mathrm{cw}}(\beta, \langle \mathbf{h}, \mathbf{x} \rangle; 1) \end{aligned}$$

due to $K(\Phi) \geqslant 0$ and the fact that $f^{\mathrm{cw}}(\beta, \cdot; 1)$ is even. There is equality iff for a maximizing family Φ^o one has $K(\Phi^o) = 0$ and $\langle i \rangle$ is equal to the corresponding Curie–Weiss magnetization. In terms of the Curie–Weiss magnetizations, $K(\Phi^o) = 0$ iff

$$\begin{aligned} & m^{\mathrm{cw}}(\beta, h^1 + h^2 + h^3) - m^{\mathrm{cw}}(\beta, h^1 + h^2 - h^3) \\ & \quad - m^{\mathrm{cw}}(\beta, h^1 - h^2 + h^3) + m^{\mathrm{cw}}(\beta, h^1 - h^2 - h^3) = 0 \end{aligned}$$

This is certainly satisfied if one or more of the components of $\mathbf{h}$ is zero; it is also satisfied for certain particular nonzero values of the three components. But in general the expression is not zero.

4. THE CASE OF TWO PATTERNS

If $p=2$, then $\mathscr{X}=\{(+,+),(+,-),(-,+),(-,-)\}$ in obvious notation. A complete detailed description can be given because the $2^p=4$ equations (2.5) decouple. The limiting negative free-energy density can be described in terms of the limiting negative free-energy density at reciprocal temperature β of a Curie–Weiss model (1.2). We use the notation introduced in the Introduction, and we write accordingly $\phi^{\mathrm{cw}}(\beta, h; J)$ for the unique probability measure on $\mathscr{S}$ whose spin expectation is $m^{\mathrm{cw}}(\beta, h; J)$.

The dependence on the measure λ enters only through the parameters

$$a=\lambda(\{(+,+)\})+\lambda(\{(-,-)\}); \qquad b=\lambda(\{(+,-)\})+\lambda(\{(-,+)\})$$

After a lengthy but straightforward discussion of the solutions of (2.5) using the knowledge about the Curie–Weiss model, we obtain the following results:

$$f(\beta, \mathbf{h})=af^{\mathrm{cw}}(\beta, h^1+h^2; 2a)+bf^{\mathrm{cw}}(\beta, h^1-h^2; 2b)$$

This is differentiable in the whole (h^1, h^2) plane except for the two lines $|h^1|=|h^2|$ when $2\beta a>1$ (for the line $h^2=-h^1$) or $2\beta b>1$ (for the line $h^2=h^1$).[4] One has $(\nabla_{\mathbf{h}} f)(\beta, \mathbf{h})=\mathbf{m}(\beta, \mathbf{h})$ with

$$m^1(\beta, \mathbf{h})=am^{\mathrm{cw}}(\beta, h^1+h^2; 2a)+bm^{\mathrm{cw}}(\beta, h^1-h^2; 2b)$$

$$m^2(\beta, \mathbf{h})=am^{\mathrm{cw}}(\beta, h^1+h^2; 2a)-bm^{\mathrm{cw}}(\beta, h^1-h^2; 2b)$$

The families maximizing (2.4) can be described completely in all cases, but we give only a rough description.

When $\mathbf{h}=\mathbf{0}$ and $\beta\leqslant\min\{(2a)^{-1}, (2b)^{-1}\}$ the maximizing family is unique and $\phi^o_{\mathbf{x}}$ is the uniform probability measure.

When $\mathbf{h}=\mathbf{0}$ and $\beta>\max\{(2a)^{-1}, (2b)^{-1}\}$ there are four maximizing families corresponding to alternatives $\phi_{(++)}=\phi^{\mathrm{cw}}(\beta, 0^{\pm}; 2a)$, $\phi_{(--)}=\phi^{\mathrm{cw}}(\beta, 0^{\mp}; 2a)$, $\phi_{(+-)}=\phi^{\mathrm{cw}}(\beta, 0^{\pm}; 2b)$, and $\phi_{(-+)}=\phi^{\mathrm{cw}}(\beta, 0^{\mp}; 2b)$.

[4] Notice that the critical values $\beta=(2a)^{-1}$ and $\beta=(2b)^{-1}$ are in the interval $(1/2, \infty)$. The bound $\beta=p^{-1}$ of Proposition 3 is attained for the degenerate cases $a=1$ or $b=1$.

When $|h^1| = |h^2|$ and

$$\min\{(2a)^{-1}, (2b)^{-1}\} < \beta \leqslant \max\{(2a)^{-1}, (2b)^{-1}\}$$

there are just two maximizing families.

It is remarkable that for $\mathbf{h} = \mathbf{0}$ the two components of $\mathbf{m}_\lambda(\Phi)$ for a maximizing family Φ are *equal for* $a > b$ and *the negative of each other* for $a < b$.

When $|h^1| \neq |h^2|$ or if $\beta \leqslant \min\{(2a)^{-1}, (2b)^{-1}\}$ the maximizing family is unique.

We now consider the particular case where the ξ_i^μ are i.i.d. One then has an asymmetry parameter $\varepsilon \in [-1/2, 1/2]$, such that

$$\text{Prob}\{\xi_i^\mu = \pm 1\} = \tfrac{1}{2} \pm \varepsilon \tag{4.1}$$

In this case the characteristic parameters a and b determining the measure λ are given by $a = \frac{1}{2} + 2\varepsilon^2$ and $b = \frac{1}{2} - 2\varepsilon^2$. At zero field $\mathbf{h} = \mathbf{0}$ and for $\varepsilon \neq 0$ one has two distinct critical temperatures

$$T_c = (\beta_c)^{-1} = 1 + 4\varepsilon^2, \qquad T_* = (\beta_*)^{-1} = 1 - 4\varepsilon^2$$

and three distinct regimes:

1. *The homogeneous zero-overlap regime*: For $T \geqslant T_c$ one has $f(\beta, \mathbf{0}) = \beta^{-1} \ln(2)$ and there is a unique maximizing family. The associated overlap parameter is $\mathbf{m} = \mathbf{0}$.
2. *The homogeneous nonzero-overlap regime*: For $T_* \leqslant T < T_c$ one has $f(\beta, \mathbf{0}) = af^{\text{cw}}(\beta, 0; 2a) + b\beta^{-1} \ln(2)$; there are two maximizing families Φ^1 and Φ^2. The two associated overlap vectors $\mathbf{m}^{(1)}$ and $\mathbf{m}^{(2)}$ are the negative of each other ($-\mathbf{m}^{(1)} = \mathbf{m}^{(2)}$) and have equal nonzero components $[(m^{(j)})^1 = (m^{(j)})^2]$.
3. *The inhomogeneous overlap regime*: For $T < T_*$ there are four maximizing families and associated to them four distinct overlap vectors. Each of these vectors has the property that its two components are unequal.

Thus, T_c is the transition temperature for the regime of zero overlap to nonzero overlap, and $T_*(< T_c)$ is the transition temperature from the regime of homogeneous overlap (i.e., the two components of the overlap vectors are equal) to the nonhomogeneous overlap (i.e., the two components of the overlap vectors are distinct). The "phase diagram" is given in Fig. 2.

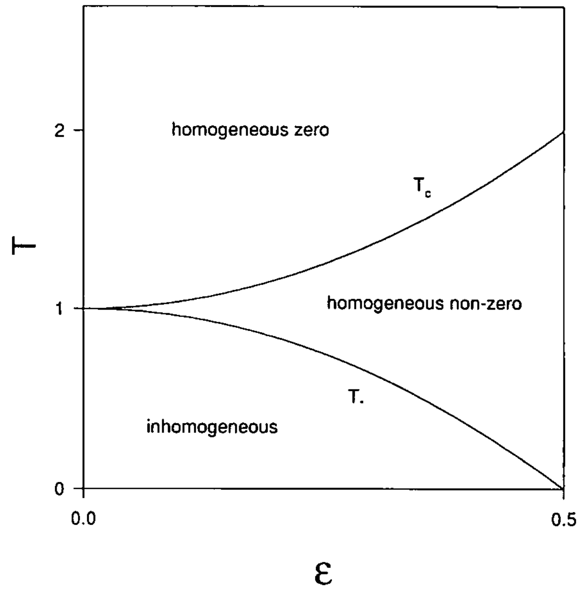

Fig. 2. Phase diagram for $p = 2$ in the general i.i.d. case: $\text{Prob}\{\xi_i^\mu = \pm 1\} = \frac{1}{2} \pm \varepsilon$.

5. THE GENERAL I.I.D. CASE

In the previous section dealing with $p = 2$ patterns and the general i.i.d. case we identified two critical temperatures $T_c > T_*$ which separate three different regimes. In this last section we collect the scanty information that we presently have for arbitrary finite p. Tamarit and Curado[21] describe results obtained (via the methods of Amit *et al.*) assuming pair-correlated patterns; here also three distinct regimes appear. A numerical study of the critical points of the functional F^{PF} has been performed in ref. 22; they distinguish a number of metastable "phases" below T_c arising from local maxima of F^{PF}.

We suppose that the ξ_i^μ are i.i.d. with distribution (4.1) and we consider the case $\mathbf{h} = \mathbf{0}$ omitting it from the notation. Since the Hamiltonian (1.1) is invariant under a change of sign of each ξ_i^μ, we conclude that $f(\beta)$ is an even function of ε. In view of part 5 of Proposition 3, we have

$$\beta_c = \sup\{\beta > 0: f(\beta) = \beta^{-1} \ln(2)\}$$

Alternatively, β_c is the supremum over the set of positive β such that F_β admits a unique maximizer Φ^o (which is then the uniform family assigning

$\langle\sigma\rangle = 0$ at each $\mathbf{x} \in \mathscr{X}$). The following lower bound is exact for every ε when $p = 2$ and for every p when $\varepsilon = 0, \pm\frac{1}{2}$.

Proposition 5. $\beta_c^{-1} = T_c \geqslant 1 + 4\varepsilon^2(p-1)$.

Proof. Notice that for $\varepsilon = 0$ we already have proved $T_c = 1$. So we assume that $\varepsilon \neq 0$.

For $\mathbf{x} \in \mathscr{X}$, let $r(\mathbf{x}) = \frac{1}{2}(\sum_{\mu=1}^{p} x^\mu + p)$, which is just the number of $+1$ in $\mathbf{x}$, so that $r(\mathbf{x}) \in \{0, 1,..., p\}$. For $r = 0, 1,..., p$ consider the subsets $\mathscr{X}_r$ of $\mathscr{X}$ of those $\mathbf{x}$ with $r(\mathbf{x}) = r$:

$$\mathscr{X}_r = \left\{\mathbf{x} \in \mathscr{X}\colon \sum_{\mu=1}^{p} x^\mu = 2r - p\right\}$$

Then $|\mathscr{X}_r| = \binom{p}{r}$ and $\lambda(\mathbf{x}) = (\frac{1}{2} + \varepsilon)^{r(\mathbf{x})}(\frac{1}{2} - \varepsilon)^{p - r(\mathbf{x})}$. Let

$$\gamma(r) = (\tfrac{1}{2} + \varepsilon)^r (\tfrac{1}{2} - \varepsilon)^{p-r} \binom{p}{r}$$

denote the binomial distribution associated to the distribution of the ξ_i^μ.

Suppose that the family $\Phi \in \Gamma_\lambda$ is uniform on each $\mathscr{X}_r$, i.e., $\langle\sigma\rangle_{\Phi_\mathbf{x}}$ is constant on $\mathscr{X}_r$ for every r; then letting s be the vector in $\mathbb{R}^{p+1}$ with components $s_r = \langle\sigma\rangle_{\Phi_\mathbf{x}}$, $\mathbf{x} \in \mathscr{X}_r$, we obtain

$$m_\lambda^\mu(\Phi) = \sum_{r=0}^{p} (\tfrac{1}{2} + \varepsilon)^r (\tfrac{1}{2} - \varepsilon)^{p-r} s_r \sum_{\mathbf{x} \in \mathscr{X}_r} x^\mu$$

Since all $\mathbf{x} \in \mathscr{X}_r$ arise from any one vector in $\mathscr{X}_r$ by permuting the components, the sum $\sum_{\mathbf{x} \in \mathscr{X}_r} x^\mu$ is independent of μ. The evaluation gives

$$\sum_{\mathbf{x} \in \mathscr{X}_r} x^\mu = \frac{2r - p}{p} \binom{p}{r}$$

The computation of $S_\lambda(\Phi)$ is direct and we obtain

$$F(\Phi) = \frac{1}{2p}\left[\sum_{r=0}^{p} \gamma(r)(2r - p)\, s_r\right]^2 + \beta^{-1} \sum_{r=0}^{p} \gamma(r)\, \eta\left(\frac{1 + s_r}{2}\right) =: \tilde{F}(\mathbf{s})$$

It follows that

$$f(\beta) \geqslant \max_{\Phi} F(\Phi) = \max_{\mathbf{s}} \tilde{F}(\mathbf{s}) = \tilde{f}(\beta)$$

where the first maximum is over the families which are uniform on each $\mathscr{X}_r$ and the second maximum is over $(p+1)$-dimensional vectors with components in the interval $[-1, 1]$.

We will solve the variational problem for $\tilde{F}$ and show that

$$\{\beta > 0: \tilde{f}(\beta) = \beta^{-1} \ln(2)\} = (0, \tilde{\beta}_c]$$

with $\tilde{\beta}_c = (1 + 4\varepsilon^2(p-1))^{-1}$. The inequality $f(\beta) \geqslant \tilde{f}(\beta) \geqslant \tilde{F}(\mathbf{0}) = \beta^{-1} \ln(2)$ then implies that $\beta_c \leqslant \tilde{\beta}_c$, which is the claim.

The critical points of $\tilde{F}$ are the solutions of

$$s_r = \tanh\left\{\beta \frac{2r-p}{\sqrt{p}} \sum_{t=0}^{p} \gamma(t) \frac{2t-p}{\sqrt{p}} s_t\right\}, \qquad r = 0, 1,..., p \tag{5.1}$$

Introduce the function $G: \mathbb{R} \to \mathbb{R}$:

$$G(m) = -\frac{1}{2p} m^2 + \frac{1}{\beta} \sum_{r=0}^{p} \gamma(r) \ln\left[2 \cosh\left(\beta m \frac{2r-p}{p}\right)\right]$$

The condition $dG(m)/dm = 0$ is

$$m = \sum_{r=0}^{p} \gamma(r)(2r-p) \tanh\left(\beta m \frac{2r-p}{p}\right) \tag{5.2}$$

For $m \in \mathbb{R}$, let $\mathbf{s}(m)$ be the vector in $\mathbb{R}^{p+1}$ with components $s(m)_r = \tanh(\beta m(2r-p)/p)$. For $\mathbf{s} \in \mathbb{R}^{p+1}$, let $\hat{m}(\mathbf{s}) = \sum_{r=0}^{p} \gamma(r)(2r-p)\, s_r$. The analogue of (2.7) is

$$\tilde{F}(\mathbf{s}(m)) = G(m) + \left(\frac{1}{2p}\right)[m - \hat{m}(\mathbf{s}(m))]^2$$

Moreover, (5.1) and (5.2) are, respectively, equivalent to $\mathbf{s} = \mathbf{s}(\hat{m}(\mathbf{s}))$ and $m = \hat{m}(\mathbf{s}(m))$. Thus we have the analogue of Proposition 2: the maps $m \mapsto \mathbf{s}(m)$ and $\mathbf{s} \mapsto \hat{m}(\mathbf{s})$ are each other's inverses when restricted to critical points of G, respectively, $\tilde{F}$; and $\tilde{F}(\mathbf{s}) = G(\hat{m}(\mathbf{s}))$ at critical points. In particular, the maximizers of $\tilde{F}$ and of G are in bijective correspondence and

$$\max_{\mathbf{s}} \tilde{F}(\mathbf{s}) = \max_{m} G(m)$$

Let us now look at solutions of (5.2); to this end we analyze the function

$$h(m) := \sum_{r=0}^{p} \gamma(r)(2r-p) \tanh\left(\beta m \frac{2r-p}{p}\right)$$

given by the r.h.s. of (5.2). The result of this analysis is the following lemma, the proof of which is elementary.

Lemma 1. h is strictly increasing with $0 < h'(m) \leqslant h'(0)$, $\forall m \in \mathbb{R}$. Moreover, h is strictly concave (resp. strictly convex) for $m > 0$ (resp. $m < 0$).

This result establishes that there are at most three solutions of the gap equation $m = h(m)$ (say $m_o = 0$, $\pm a_o$), since a strictly convex or concave function can intersect a straight line at most in two points and in the present case $m = 0$ is a solution.

Finally, let us compute the value of $\tilde{\beta}_c$, for which $m_o = 0$ ceases to be a maximizer. The second derivative of G at $m = 0$ is

$$G''(0) = -p^{-1} + \beta p^{-2} \sum_{r=0}^{p} \gamma(r)(2r - p)^2$$

so that $\tilde{\beta}_c$ is determined by the condition $\tilde{\beta}_c \sum_{r=0}^{p} \gamma(r)(2r - p)^2 = p$. With the well-known first and second moments of the binomial distribution we compute that $\tilde{\beta}_c$ has the claimed value. ▮

To attempt to define the second critical reciprocal temperature, consider the set

$$\mathscr{H} = \{\beta > 0 \colon m_\lambda^\mu(\Phi) \text{ is independent of } \mu \in \{1, 2, ..., p\}$$
$$\text{for each maximizer } \Phi \text{ of } F_\beta\}$$

of reciprocal temperatures, where every maximizing family Φ has homogeneous $\mathbf{m}_\lambda(\Phi)$. The first problem is to show that this set is an interval. Our conjecture is that this is the case and that $\beta_* = \sup \mathscr{H}$ is given by

$$(\beta_*)^{-1} = T_* = \max\{1 - 4\varepsilon^2(p-1), 0\}$$

This would imply that for given $\varepsilon \neq 0$, there is a critical value of p given by $p_c = [1 + (2\varepsilon)^{-2}]^{1/2}$ (which is greater than 2); or alternatively, for a given value of p there is a critical value of ε given by $\varepsilon_c = [4(p-1)]^{-1/2}$ (which is smaller than 1/2 for $p \geqslant 3$) such that for $p \geqslant p_c$, or alternatively $|\varepsilon| \geqslant \varepsilon_c$ the homogeneous regime persists and the nonhomogeneous regime does not appear (i.e., $T_* = 0$). The "phase diagram" for $p \geqslant 3$ would then be that of Fig. 3.

The conjectures are amply verified by numerical solution of the variational problem. The subject is presently under investigation.

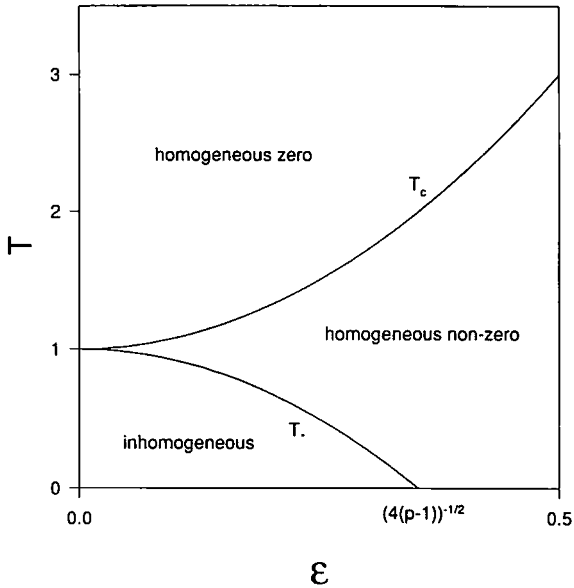

Fig. 3. Conjectured phase diagram for $p \geqslant 3$ in the general i.i.d. case.

APPENDIX

The following result is used in the proof of Proposition 4 in the particular case $\mathscr{A} = \mathscr{A}_1$ and in the proof of part 3 of Proposition 3 with $\mathscr{A} = \mathscr{X}$. In both cases the hypothesis is easily verified.

Lemma. If $\mathscr{A} \subset \mathscr{X}$ satisfies

$$\begin{aligned}&|\mathscr{A} \cap \mathscr{A}_\mu \cap \mathscr{A}_\nu| + |\mathscr{A} \cap (-\mathscr{A}_\mu) \cap (-\mathscr{A}_\nu)| \\ &\quad = |\mathscr{A} \cap \mathscr{A}_\mu \cap (-\mathscr{A}_\nu)| + |\mathscr{A} \cap (-\mathscr{A}_\mu) \cap \mathscr{A}_\nu|\end{aligned}$$

for all $\mu \neq \nu \in \{1, 2,..., p\}$; then

$$\sum_{\mu=1}^{p} \left[\sum_{\mathbf{x} \in \mathscr{A}} \lambda(\{\mathbf{x}\})\, x^\mu \langle \sigma \rangle_{\phi_\mathbf{x}} \right]^2 \leqslant |\mathscr{A}| \sum_{\mathbf{x} \in \mathscr{A}} [\lambda(\{\mathbf{x}\}) \langle \sigma \rangle_{\phi_\mathbf{x}}]^2$$

Proof. This is Bessel's inequality in $\mathbb{R}^{|\mathscr{A}|}$ applied to the vector with components $\lambda(\{\mathbf{x}\}) \langle \sigma \rangle_{\phi_\mathbf{x}}$ $(\mathbf{x} \in \mathscr{A})$, and the p vectors $\{e^\mu\}$ with components $(e^\mu)_\mathbf{x} = |\mathscr{A}|^{-1/2} x^\mu$ $(\mathbf{x} \in \mathscr{A})$, whose orthonormality is a consequence of the hypothesis on $\mathscr{A}$. ∎

ACKNOWLEDGMENTS

We thank Dr. F. A. Tamarit for informative discussions. Support from the Consejo de Investigaciones Científicas y Tecnológicas de la Provincia de Córdoba (CONICOR) and the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) is gratefully acknowledged.

REFERENCES

1. L. A. Pastur and A. L. Figotin, Exactly soluble model of a spin glass, *Sov. J. Low Temp. Phys.* **3**:378–383 (1977); On the theory of disordered spin systems, *Theoret. Math. Phys.* **35**:403–414 (1978); Infinite range limit for a class of disordered systems, *Theoret. Math. Phys.* **51**:564–569 (1982).
2. J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. USA* **79**:2554–2558 (1982).
3. D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, Cambridge, 1989).
4. D. J. Amit, H. Gutfreund, and H. Sompolinsky, Spin glass models of neural networks, *Phys. Rev. A* **32**:1007–1018 (1985).
5. J. L. van Hemmen, Spin-glass methods of a neural network, *Phys. Rev. A* **34**:3435–3445 (1986).
6. S. Albeverio, B. Tirozzi, and B. Zegarlinski, Rigorous results for the free energy in the Hopfield model, *Commun. Math. Phys.* **150**:337–373 (1992).
7. F. Comets, Large deviation estimates for a conditional probability distribution. Applications to random interaction Gibbs measures, *Prob. Theory Related Fields* **80**:407–432 (1989).
8. H. Koch and J. Piasko, Some rigorous results on the Hopfield neural network model, *J. Stat. Phys.* **55**:903–928 (1989).
9. H. Koch, A free energy bound for the Hopfield model, *J. Phys. A* **26**:353–355 (1993).
10. M. V. Scherbina and B. Tirozzi, The free energy for a class of Hopfield models, *J. Stat. Phys.* **72**:113–125 (1992).
11. A. Bovier and V. Gayrard, Rigorous results on the thermodynamics of the dilute Hopfield model, *J. Stat. Phys.* **69**:597–627 (1993).
12. A. Bovier, V. Gayrard, and P. Picco, Gibbs states of the Hopfield model in the regime of perfect memory, *Prob. Theory Related Fields* **100**:329–363 (1994).
13. A. Bovier, V. Gayrard, and P. Picco, Large deviation principles for the Hopfield model and the Kac–Hopfield model, *Prob. Theory Related Fields* (1995), to appear.
14. D. J. Amit, H. Gutfreund, and H. Sompolinsky, Storing infinite number of patterns in a spin glass model of neural networks, *Phys. Rev. Lett.* **55**:1530–1533 (1985).
15. A. Bovier, V. Gayrard, and P. Picco, Gibbs states of the Hopfield model with extensively many patterns, *J. Stat. Phys.* **79**:395–414 (1995).
16. A. Bovier, Self-averaging in a class of generalized Hopfield models, *J. Phys. A* **27**:7069–7077 (1994).
17. A. Bovier and V. Gayrard, Rigorous results on the Hopfield model of neural networks. Invited lecture at the V CLAPEM, São Paulo (Brasil), *Resenhas Inst. Mat. Estat. Univ. São Paulo* **1**:161–172 (1994).
18. G. A. Raggio and R. F. Werner, Quantum statistical mechanics of general mean field systems, *Helv. Phys. Acta* **62**:980–1003 (1989).

19. G. A. Raggio and R. F. Werner, The Gibbs variational principle for inhomogeneous mean-field systems, *Helv. Phys. Acta* **64**:633–667 (1991).
20. J. Dieudonné, *Foundations of Modern Analysis*, Vol. I (Academic Press, New York, 1969).
21. F. A. Tamarit and E. M. F. Curado, Pair-correlated patterns in the Hopfield model of neural networks, *J. Stat. Phys.* **62**:473–480 (1991).
22. J. F. Fontanari and W. K. Theumann, On the storage of correlated patterns in Hopfield's model, *J. Phys. France* **51**:375–386 (1990).