Journal of Statistical Physics, Vol. 72, Nos. 1/2, 1993

Rigorous Results on the Thermodynamics of the Dilute Hopfield Model

Anton Bovier[1] **and Véronique Gayrard**[2]

Received November 10, 1992; final February 2, 1993

We study the Hopfield model of an autoassociative memory on a random graph on N vertices where the probability of two vertices being joined by a link is $p(N)$. Assuming that $p(N)$ goes to zero more slowly than $O(1/N)$, we prove the following results: (1) If the number of stored patterns $m(N)$ is small enough such that $m(N)/Np(N)\downarrow 0$, as $N\uparrow\infty$, then the free energy of this model converges, upon proper rescaling, to that of the standard Curie–Weiss model, for almost all choices of the random graph and the random patterns. (2) If in addition $m(N) < \ln N/\ln 2$, we prove that there exists, for $T<1$, a Gibbs measure associated to each original pattern, whereas for higher temperatures the Gibbs measure is unique. The basic technical result in the proofs is a uniform bound on the difference between the Hamiltonian on a random graph and its mean value.

KEY WORDS: Neural networks; Hopfield model; random graphs; mean-field theory.

1. INTRODUCTION

The Hopfield model of an autoassociative memory[12] is described by a Hamiltonian function

$$H_N(\xi;\sigma) = -\frac{1}{N}\sum_{\substack{(i,j)\in\Lambda\times\Lambda\\ i\neq j}}\sum_{\mu=1}^{m}\xi_i^\mu\xi_j^\mu\sigma_i\sigma_j \tag{1.1}$$

on the space $\mathcal{S}^N$ of spin configurations $\sigma\equiv\{\sigma_i\}_{i\in\Lambda}$, where, for a given positive integer N, $\Lambda\equiv\{1,\ldots,N\}$ and the spins variables $\sigma_i\in\mathcal{S}\equiv\{-1,+1\}$ indicate

[1] Institut für Angewandte Analysis und Stochastik, Hausvogteiplatz 5-7, O-1086 Berlin, Germany.
[2] Centre de Physique Théorique-CNRS, Luminy, Case 907, F-13288 Marseille Cedex, France.

822/72/1-2-6

the excitation state of the neuron i. The variables $\{\xi_i^\mu\}_{i\in\Lambda}^{\mu=1,\dots,m} \in \mathcal{S}^{m\times N}$ describe the m patterns the system is supposed to memorize. It is generally assumed that these patterns are "random," i.e., the components ξ_i^μ form a family of mN independent, identically distributed random variables. Typically, one is interested in choosing m as a function of N as large as possible under the condition that certain crucial properties of the system are retained ("memory capacity").

It was noticed very early (see, e.g., ref. 1) that this model formally resembles closely a mean-field model of a spin glass, the Sherrington–Kirkpatrick model,[18] which has been heavily investigated by physicists (see for a review ref. 16). Therefore, tools from spin-glass theory, such as the replica method, have been employed to study this model. More recently, it has been realized that the Hopfield model is in fact much easier to handle in a mathematically rigorous way than real spin-glass models, at least if the number of stored patterns m is subject to certain restrictions. Moreover, m may serve as a parameter that allows one to continuously drive the system from an essentially trivial regime ($m=1$) to a highly complex and unpredictable "spin-glass" regime ($m>N$). From this point of view, the Hopfield model does represent in fact an extremely interesting disordered mean-field model.

Let us describe some of the main results so far obtained (for a more extensive account, see the articles in ref. 8): In 1989, Koch and Piasko gave in a remarkable article[14] a complete analysis of the thermodynamic limit of this model under the constraint that m is allowed to grow with the system size N no faster than $(\ln N)/\ln 2$, using a method originally introduced by Grensing and Kühn.[11] Their construction implied the almost sure convergence of the free energy to a calculable limit (which is simply the free energy of the standard Curie–Weiss model) as well as that of the distribution with respect to the Gibbs measures of the so-called overlap parameters

$$m_N^\mu(\xi;\sigma) = \frac{1}{N}\sum_i \xi_i^\mu \sigma_i \tag{1.2}$$

A detailed description of these results will be given later. These results have been sharpened and generalized to the q-state Potts version of this same model by Gayrard.[10] More recently, Koch[15] obtained a further very interesting result. He proved bounds on the free energy for all finite N that in particular imply that if m is chosen such that $\lim_{N\uparrow\infty}(m/N)=0$, then the expectation of the free energy with respect to the distribution of the patterns converges to the free energy of the Curie–Weiss model. As a matter of fact, it is very easy to extend his results to obtain the almost sure convergence of the free energy (see Section 3 of this article). It should be noted that this result holds for all temperatures.

The Hopfield model as given by (1.1) can be seen as a spin system on the complete graph on $\{1,..., N\}$. Both from the point of view of applications in the context of neural networks and from that of the theory of disordered systems, it is desirable to study generalizations of the Hopfield model on more general graphs, and in particular on random graphs; still more generally, one may even wish to study this model when the interaction between sites i and j is not only governed by the matrix $\sum_\mu \xi_i^\mu \xi_j^\mu$ but is modulated by a random variable ε_{ij}. This model is then called the "dilute Hopfield model" and is given by the Hamiltonian

$$H_N(\xi; \varepsilon; \sigma) = -\frac{1}{Np} \sum_{\substack{(i,j) \in \Lambda \times \Lambda \\ i \neq j}} \varepsilon_{ij} \sum_{\mu=1}^{m} \xi_i^\mu \xi_j^\mu \sigma_i \sigma_j \tag{1.3}$$

where $p = \mathbb{E}(\varepsilon_{ij}) > 0$. Dilute neural network models are frequently studied in the regime where $p = c/N$ with $c \ll \ln N$ (so-called highly dilute model). There it has been noted by Derrida *et al.*[7] that the dynamics of this model with asymmetric ε_{ij} (i.e., ε_{ij} independent of ε_{ji}) can be solved exactly in the limit $N \uparrow \infty$ if the number of patterns is kept proportional to c. The reason for this is that in this limit the underlying graph has essentially the structure of a (disconnected) tree (see, e.g., ref. 5). An undesirable feature in this situation is the instability of this model against mixing of patterns and thus noisy dynamics. This last point is very easily understood in terms of the Hopfield Hamiltonian (1.3). Namely, if the underlying graph has the structure of a tree, then by cutting any edge it becomes disconnected, and choosing σ to equal one pattern on one of the components and another on the second, one finds that this configuration differs in energy only by a finite amount from the original patterns. Moreover, one may construct an infinite number of such mixtures.

Diluted networks are of interest not only if they are easier to analyze, but also for pragmatic reasons of network architecture. In very large networks, maintaining full connectivity is clearly undesirable if not impractical for technical reasons. It is thus natural to ask how such a model behaves if it is less highly diluted, and in particular one may ask how much the network may be diluted if the properties of the fully connected network are to be retained. This has been done recently[3] in the regime where $m = \alpha Np$, where it has been shown that rigorous lower bounds on the storage capacity as proven first by Newman[17] for the model (1.1) can be recovered in this situation, provided that $p \geq [(\ln N)/N]^{1/2}$.

In the present paper we study this model from the point of view of mean-field theory in the regime where $m \ll N$. As we will see, the mean-field results prove fairly robust against the effect of dilution and can be reproven under fairly weak assumption on the ε_{ij}, although we must always require p to be much larger than in Derrida's model.

To be able to make precise statements, we need to introduce some notation. First, let $\Omega_\xi \equiv \{-1, +1\}^{\mathbb{N}\times\mathbb{N}}$, $\mathscr{F}_\xi$ the corresponding Borel σ-algebra, and let $\mathbb{P}_\xi$ denote the product measure on Ω_ξ such that $\xi \equiv \{\xi_i^\mu\}_{i\in\mathbb{N}}^{\mu\in\mathbb{N}}$ is a family of independent, identically distributed random variables with $\mathbb{P}_\xi(\xi_i^\mu = \pm 1) = 1/2$. Note that for a given, nondecreasing function m: $\mathbb{N} \to \mathbb{N}$ we will denote by $\xi(N)$ the family $\{\xi_i^\mu\}_{i=1,\dots,N}^{\mu=1,\dots,m(N)}$.

To define the probability space for the dilution variables ε, we need to be slightly more sophisticated due to the fact that we want the marginal distributions of the ε_{ij} to depend on the size of the network while at the same time define all Hamiltonians for different N on the same probability space. A way of doing this is the following.[3] Let $(\Omega_\varepsilon, \mathscr{F}_\varepsilon, \mathbb{P}_\varepsilon)$ be some abstract probability space and let $\{W_{ij}\}_{i,j\in\mathbb{N}\times\mathbb{N}}$ be a family of i.i.d. real-valued random variables on this space whose common distribution is the uniform distribution on the interval $[0, 1]$. Then for a given function p: $\mathbb{N} \to [0, 1]$ define

$$\varepsilon_{ij}(N) \equiv \mathbb{1}_{\{W_{ij} \leqslant p(N)\}} \tag{1.4}$$

where $\mathbb{1}_A$ is the characteristic function of the event A, i.e., is equal to one if A holds and equal to zero otherwise. Obviously, then,

$$\mathbb{P}_\varepsilon(\varepsilon_{ij}(N) = 1) = 1 - \mathbb{P}_\varepsilon(\varepsilon_{ij}(N) = 0) = p(N) \tag{1.5}$$

as desired, and for fixed N, $\varepsilon_{ij}(N)$ are i.i.d. random variables. Moreover, one may check that for fixed i, j, the family $\{\varepsilon_{ij}(N)\}_{N\in\mathbb{N}}$ forms a Markov chain and for given marginals (1.5) it has the property that $\mathbb{P}_\varepsilon(\varepsilon_{ij}(N) = \varepsilon_{ij}(N-1))$ is maximized.

Remark. Obviously, such a construction can easily be generalized for $\varepsilon_{ij}(N)$ taking values in a more general space than $\{0, 1\}$.

Let us now define the finite-volume partition functions and free energy of our model through

$$Z_{N,\beta}(\xi; \varepsilon) \equiv \sum_{\sigma \in \mathscr{S}^N} \frac{1}{2^N} e^{-\beta H_N(\xi;\varepsilon;\sigma)} \tag{1.6}$$

and

$$f_{N,\beta}(\xi; \varepsilon) \equiv -\frac{1}{\beta N} \ln Z_{N,\beta}(\xi; \varepsilon) \tag{1.7}$$

[3] We are indebted to Chuck Newman for suggesting this form of presentation, which is, although essentially equivalent, far more elegant than our original version.

Let us further denote by $f_{\mathrm{CW}}(\beta)$ the free energy of the Curie–Weiss model,[9] i.e.,

$$f_{\mathrm{CW}}(\beta) = \inf_{x \in \mathbb{R}} \left(-\frac{1}{\beta} \ln \cosh(\beta x) + \frac{x^2}{2} \right) \tag{1.8}$$

Then, we have the following result.

Theorem 1. Let p: $\mathbb{N} \to (0, 1]$ be a decreasing function such that $p(N)N \uparrow \infty$, as $N \uparrow \infty$, and let m: $\mathbb{N} \to \mathbb{N}$ be an increasing function such that $m(N)/p(N)N \downarrow 0$, as $N \uparrow \infty$. Then, for all $0 \leqslant \beta < \infty$,

$$\lim_{N \uparrow \infty} f_{N,\beta}(\xi; \varepsilon) = f_{\mathrm{CW}}(\beta), \qquad \mathbb{P}_\varepsilon \times \mathbb{P}_\xi\text{-a.s.} \tag{1.9}$$

As in the standard Hopfield model, when the number of patterns m is small enough, the extremal infinite-volume Gibbs states of the dilute Hopfield model are expected to be measures $\mathcal{G}^\alpha$ concentrated near the original patterns ξ^α. Here what we will in fact be interested in is the limiting distributions of the overlap parameters (1.2) with respect to the measure $\mathcal{G}^\alpha$. More precisely, in order to construct the measure $\mathcal{G}^\alpha$, we add to the Hamiltonian $H_{N,\varepsilon,\xi}(\sigma)$ a "magnetic field" h coupling to the pattern ξ^α, that is, we write

$$H^\alpha_{N,h}(\varepsilon; \xi; \sigma) = H_N(\varepsilon; \xi; \sigma) - h \sum_{i=1}^{N} \sigma_i \xi^\alpha_i \tag{1.10}$$

We denote by $\mathcal{G}^\alpha_{N,\beta,h}(\varepsilon; \xi)$ the finite-volume Gibbs measure which assigns to the configuration $\sigma \in \mathcal{S}^N$ the probability

$$\mathcal{G}^\alpha_{N,\beta,h}(\varepsilon; \xi; \sigma) = \frac{\exp\{-\beta H^\alpha_{N,h}(\varepsilon; \xi; \sigma)\}}{\sum_{\sigma \in \mathcal{S}^N} \exp\{-\beta H^\alpha_{N,h}(\varepsilon; \xi; \sigma)\}} \tag{1.11}$$

We denote by $m^\mu_N(\xi)$ the map

$$\begin{aligned} m^\mu_m(\xi)\colon \quad & \mathcal{S}^N \to [-1, 1] \\ & \sigma \mapsto m^\mu_N(\xi; \sigma) \end{aligned} \tag{1.12}$$

and by $\mathcal{L}^\alpha[m^\mu_N(\xi)]$ the law of $m^\mu_N(\xi)$ under $\mathcal{G}^\alpha_{N,\beta,h}(\varepsilon; \xi)$. Let $a^+(\beta)$, respectively $a^-(\beta)$, be the largest and smallest solutions of the equation $a = \tanh(\beta a)$ and define $\tilde{m}^\pm_{\alpha,\mu}(\beta) \equiv a^\pm(\beta)\delta_{\alpha,\mu}$, where $\delta_{\alpha,\mu}$ is the Kronecker symbol. Then, denoting by $\delta_{\{x\}}$ the Dirac measure on $\mathbb{R}$ concentrated at the point x, we have the following result.

Theorem 2. Suppose that all the assumptions of Theorem 1 are satisfied and that in addition $m < \ln N/\ln 2$. Let α, μ be fixed. Then, for all $0 \leqslant \beta < \infty$,

$$\lim_{h \downarrow 0} \lim_{N \uparrow \infty} \mathscr{L}^{\alpha}[m_N^{\mu}(\xi)] = \delta_{\{\tilde{m}^{+}_{\alpha,\mu}(\beta)\}} \qquad \mathbb{P}_{\xi} \times \mathbb{P}_{\varepsilon}\text{-a.s.} \tag{1.13}$$

The same result holds for $h \uparrow 0$ with $\tilde{m}^{+}_{\alpha,\mu}(\beta)$ replaced by $\tilde{m}^{-}_{\alpha,\mu}(\beta)$.

Remark. The restriction on the number of patterns in Theorem 2 is due to the fact that even in the standard Hopfield model, the analog of Theorem 2 has only been proven under this hypothesis. If (1.13) holds in the standard Hopfield model under weaker restrictions on m, we expect to be able to prove it also for the dilute Hopfield model under the same conditions plus those of Theorem 1.

Our proofs of Theorems 1 and 2 actually follow from the analogous results in the standard Hopfield model together with the following theorem, which really constitutes our main technical result. Let $\mathbb{P}$ denote the product measure $\mathbb{P}_{\xi} \times \mathbb{P}_{\varepsilon}$ on $(\Omega_{\xi} \times \Omega_{\varepsilon}, \mathscr{F}_{\xi} \times \mathscr{F}_{\varepsilon})$.

Theorem 3. Let m be an increasing function such that $m(N)/N \downarrow 0$ as $N \uparrow \infty$. Then there exists an event $\mathscr{A}_N \in \mathscr{F}_{\xi}$ and a constant $0 < K < \infty$ such that

$$\mathbb{P}_{\xi}(\mathscr{A}_N) \geqslant 1 - \frac{K}{N^2} \tag{1.14}$$

and such that if p satisfies $p(N)N \uparrow \infty$ as $N \uparrow \infty$, then, for any strictly decreasing function γ: $\mathbb{N} \to \mathbb{R}$ satisfying $\gamma(N) \downarrow 0$ as $N \uparrow \infty$ and $p(N)N\gamma^2(N) > c$ for some constant $0 < c < \infty$, there exists a constant $\rho > 0$ such that

$$\mathbb{P}(\forall \sigma \in \mathscr{S}^N \,|H_N(\varepsilon; \xi; \sigma) - \mathbb{E}_{\varepsilon} H_N(\varepsilon; \xi; \sigma)| < \gamma \sqrt{m} N \,|\, \mathscr{A}_N \times \Omega_{\varepsilon}) \geqslant 1 - e^{-\rho N} \tag{1.15}$$

Remark. It should be noted that our results require only the weakest plausible conditions on the dilution rate $p(N)$. In fact, in terms of the underlying random graph, this condition assures that the "giant component" of the graph is so big that the number of vertices in its complement is $o(N)$.[5] If $p(N)$ were smaller, e.g., $\overline{\lim}\, Np(N) < \infty$, then an extensive fraction of the graph would consist of finite connected components and a result like Theorem 3 could not be expected. It is also very likely that the condition in Theorem 1 on the number of stored paterns is optimal, although as yet we cannot prove this. The situation in Theorem 2 is less clear, the reason being the lack of knowledge on the structure of Gibbs states in the Hopfield model if the number of patterns exceeds ln N.

The proof of this theorem will be given in the next section. In Section 3 and 4 we will use this to prove Theorems 1 and 2, respectively.

2. A UNIFORM BOUND ON THE HAMILTONIAN

In this section we prove Theorem 3, which provides a uniform bound on the difference between the Hamiltonian of the dilute Hopfield model and its average with respect to the dilution variables ε. We have recently proven such a result for the dilute Curie–Weiss model[4] (which corresponds to the case $m \equiv 1$) and our basic strategy will be the same; however, this time the presence of the random couplings $J_{ij} \equiv \sum_\mu \xi_i^\mu \xi_j^\mu$ adds considerable complications.

Let us set $(\xi_i, \xi_j) \equiv \sum_\mu \xi_i^\mu \xi_j^\mu$. We may write the Hamiltonian as (we suppress the dependence of ε and ξ on N for simplicity of notation)

$$H_N(\varepsilon; \xi; \sigma) = -\frac{1}{pN} \sum_{i \neq j} |(\xi_i, \xi_j)| \operatorname{sign}(\xi_i, \xi_j) \varepsilon_{ij} \sigma_i \sigma_j \tag{2.1}$$

Here we choose to define

$$\operatorname{sign}(x) = \begin{cases} 1 & \text{if } x \geqslant 0 \\ -1 & \text{if } x < 0 \end{cases} \tag{2.2}$$

Now define the set Λ^+ as the set of all pairs of sites where the spins are aligned with the couplings, i.e.,

$$\Lambda^+(\sigma; \xi) \equiv \{(i, j) \in \Lambda \times \Lambda, i \neq j \mid \sigma_i \sigma_j = \operatorname{sign}(\xi_i, \xi_j)\} \tag{2.3}$$

Defining furthermore Ξ_{ij} as the indicator function of this set, i.e.,

$$\Xi_{ij} = \begin{cases} 1 & \text{if } (i, j) \in \Lambda^+(\sigma; \xi) \\ 0 & \text{otherwise} \end{cases} \tag{2.4}$$

and noticing that

$$\sigma_i \sigma_j \operatorname{sign}(\xi_i, \xi_j) = 2\Xi_{ij} - 1 \tag{2.5}$$

we may rewrite our Hamiltonian as

$$H_N(\varepsilon; \xi; \sigma) = \frac{1}{pN} \sum_{i \neq j} |(\xi_i, \xi_j)| \varepsilon_{ij} - \frac{2}{pN} \sum_{i \neq j} |(\xi_i, \xi_j)| \Xi_{ij} \varepsilon_{ij} \tag{2.6}$$

We want to prove that uniformly in σ, the Hamiltonian $H_N(\varepsilon; \xi; \sigma)$ is close to its expectation with respect to the distribution $\mathbb{P}_\varepsilon$. Since the first term in (2.6) is independent of σ, the real task is to show this property for the second term in (2.6). More precisely, let us consider the probability

$$\mathbb{P}_\varepsilon\left(\exists\sigma\in\mathcal{S}^N: \left|\sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\varepsilon_{ij} - p\sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\right| > p\gamma \sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\right) \tag{2.7}$$

where $\gamma \equiv \gamma(N)$ is some positive decreasing function tending to zero with N that will be chosen appropriately later. We will show that with $\mathbb{P}_\xi$-probability that tends to one as $N\uparrow\infty$, the probability (2.7) is exponentially small. Note that (2.7) is bounded from above by

$$\begin{aligned}&\mathbb{P}_\varepsilon\left(\exists\sigma\in\mathcal{S}^N: \sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\varepsilon_{ij} > p(1+\gamma)\sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\right)\\ &\quad + \mathbb{P}_\varepsilon\left(\exists\sigma\in\mathcal{S}^N: \sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\varepsilon_{ij} < p(1-\gamma)\sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\right)\end{aligned} \tag{2.8}$$

Our estimates will be the same for both terms in (2.8), so that we only concentrate on

$$Q_N(\xi) \equiv \mathbb{P}_\varepsilon\left(\exists\sigma\partial\mathcal{S}^N: \sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\varepsilon_{ij} > p(1+\gamma)\sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\right) \tag{2.9}$$

We have, bounding the probability of the union by the sum of the probabilities and then using the exponential Markov inequality,[6]

$$\begin{aligned}Q_N(\xi) &\leqslant \sum_{\sigma\in\mathcal{S}^N} \inf_{t\geqslant 0} \exp\left\{-p(1+\gamma)t\sum_{i\neq j} |(\xi_i, \xi_j)|\, \Xi_{ij}\right.\\ &\qquad \left. + \sum_{i\neq j} \ln[p(e^{|(\xi_i,\xi_j)|\, t\Xi_{ij}} - 1) + 1]\right\}\\ &\equiv \sum_{\sigma\in\mathcal{S}^N} \inf_{t\geqslant 0} \exp\{A_N(\xi, \sigma)\}\end{aligned} \tag{2.10}$$

Now

$$\ln[p(e^{|(\xi_i,\xi_j)|\, t\Xi_{ij}} - 1) + 1] = \Xi_{ij} \ln[p^{|(\xi_i,\xi_j)|\, t} - 1) + 1] \tag{2.11}$$

so that the exponent in (2.10) can be written as

$$A_N(\xi, \sigma) = \sum_{i\neq j} \Xi_{ij}\{-p(1+\gamma)t\, |(\xi_i, \xi_j)| + \ln[p(e^{|(\xi_i,\xi_j)|\, t} - 1) + 1]\} \tag{2.12}$$

Now for $t \geqslant 0$ we have the following bound:

$$\ln[p(e^{|(\xi_i,\xi_j)|t}-1)+1] \leqslant p(e^{|(\xi_i,\xi_j)|t}-1)$$
$$= p\left\{|(\xi_i,\xi_j)|\,t + \frac{|(\xi_i,\xi_j)|^2 t^2}{2} + R_3(|(\xi_i,\xi_j)|\,t)\right\} \tag{2.13}$$

where

$$R_3(x) \equiv \sum_{n=3}^{\infty} \frac{x^n}{n!} \tag{2.14}$$

Our strategy will now be the following: Anticipating that R_3 will be small, we choose t^* such that

$$-p(1+\gamma)\sum_{i\neq j}|(\xi_i,\xi_j)|\,\Xi_{ij}t + p\sum_{i\neq j}|(\xi_i,\xi_j)|\,\Xi_{ij}t + p\sum_{i\neq j}\frac{|(\xi_i,\xi_j)|^2 t^2}{2}\Xi_{ij} \tag{2.15}$$

is minimized, i.e.,

$$t^* = \gamma\frac{\sum_{i\neq j}|(\xi_i,\xi_j)|\,\Xi_{ij}}{\sum_{i\neq j}|(\xi_i,\xi_j)|^2\,\Xi_{ij}} \tag{2.16}$$

This gives the bound

$$Q_N(\xi) \leqslant \sum_{\sigma\in\mathscr{S}^N}\exp\left\{-\frac{1}{2}\gamma^2 p\frac{[\sum_{i\neq j}|(\xi_i,\xi_j)|\,\Xi_{ij}]^2}{\sum_{i\neq j}|(\xi_i,\xi_j)|^2\,\Xi_{ij}} + p\sum_{i\neq j}\Xi_{ij}R_3(t^*|(\xi_i,\xi_j)|)\right\} \tag{2.17}$$

Our aim is to get a σ-independent bound on the exponential in (2.17). To this end notice that first of all we have the trivial upper bounds

$$\sum_{i\neq j}|(\xi_i,\xi_j)|\,\Xi_{ij} \leqslant \sum_{i\neq j}|(\xi_i,\xi_j)| \tag{2.18}$$

and

$$\sum_{i\neq j}|(\xi_i,\xi_j)|^2\,\Xi_{ij} \leqslant \sum_{i\neq j}|(\xi_i,\xi_j)|^2 \tag{2.19}$$

More interestingly, we may also get corresponding lower bounds. Namely, since

$$\Xi_{ij} = \frac{\sigma_i\sigma_j\,\mathrm{sign}(\xi_i,\xi_j)+1}{2} \tag{2.20}$$

we get that

$$\sum_{i\neq j} |(\xi_i, \xi_j)| \Xi_{ij} = \tfrac{1}{2} \sum_{i\neq j} [\sigma_i\sigma_j(\xi_i, \xi_j) + |(\xi_i, \xi_j)|]$$

$$= \tfrac{1}{2}\left[\sum_{\mu=1}^{m} (\xi^\mu, \sigma)^2 - mN + \sum_{i\neq j} |(\xi_i, \xi_j)|\right] \tag{2.21}$$

and hence

$$\sum_{i\neq j} |(\xi_i, \xi_j)| \, \Xi_{ij} \geqslant \tfrac{1}{2}\left[-mN + \sum_{i\neq j} |(\xi_i, \xi_j)|\right] \tag{2.22}$$

Finally,

$$\sum_{i\neq j} |(\xi_i, \xi_j)|^2 \Xi_{ij} = \tfrac{1}{2} \sum_{i\neq j} \left[\sum_{\mu,\nu} \xi_i^\mu \xi_i^\nu \xi_j^\mu \xi_j^\nu \sigma_i \sigma_j + |(\xi_i, \xi_j)|^2\right]$$

$$= \tfrac{1}{2} \sum_{\mu,\nu} \left(\sum_i \xi_i^\mu \xi_i^\nu \sigma_i\right)^2 - \tfrac{1}{2} m^2 N + \tfrac{1}{2} \sum_{i\neq j} |(\xi_i, \xi_j)|^2 \tag{2.23}$$

and hence

$$\sum_{i\neq j} |(\xi_i, \xi_j)|^2 \Xi_{ij} \geqslant \tfrac{1}{2}\left[\sum_{i\neq j} |(\xi_i, \xi_j)|^2 - m^2 N\right] \tag{2.24}$$

Combining these four bounds, we get the following result.

Lemma 2.1. t^* as defined in (2.16) satisfies the following upper and lower bounds:

$$t^* \leqslant t_u^* \equiv 2\gamma \frac{\sum_{i\neq j} |(\xi_i, \xi_j)|}{\sum_{i\neq j} |(\xi_i, \xi_j)|^2 - m^2 N} \tag{2.25}$$

if $t_u^* > 0$, and

$$t^* \geqslant t_l^* \equiv \frac{\gamma}{2} \frac{\sum_{i\neq j} |(\xi_i, \xi_j)| - mN}{\sum_{i\neq j} |(\xi_i, \xi_j)|^2} \tag{2.26}$$

[Note that the condition $t_u^* > 0$ will be satisfied on the subspace of Ω_ξ where we will want to use (2.25).]

An immediate consequence of this lemma together with (2.17) is the following proposition, which yields the desired uniform, but still ξ-dependent bound on Q_N.

Proposition 2.2. The probability $Q_N(\xi)$ satisfies the bound

$$Q_N(\xi) \leqslant 2^N \exp\left\{-\frac{1}{8}\gamma^2 p \frac{[\sum_{i\neq j} |(\xi_i, \xi_j)| - mN]^2}{\sum_{i\neq j} |(\xi_i, \xi_j)|^2} + p \sum_{i\neq j} R_3(t_u^* |(\xi_i, \xi_j)|)\right\} \tag{2.27}$$

for all ξ such that $t_u^* > 0$.

What we now need to prove is that with very large probability, the exponential in (2.27) is sufficiently small to offset the 2^N prefactor. Note that it depends only on the quantities $\sum_{i\neq j} |(\xi_i, \xi_j)|^n$ and it is those we need to control. To see how this can be done, it is reasonable to think of the variables (ξ_i, ξ_j) as being essentially Gaussians with variance $m^{1/2}$. In fact, we have the following bounds.

Lemma 2.3. The moments of the variables $|(\xi_i, \xi_j)|$ (for $i \neq j$) satisfy the following upper and lower bounds:

(i) If $l \leqslant m/2$,

$$\frac{m!\, l!}{(m-l)!} \leqslant \mathbb{E}(\xi_i, \xi_j)^{2l} \leqslant 2^{2l} \frac{m!\, l!}{(m-l)!} \tag{2.28}$$

(ii) If $l > m/2$,

$$\begin{aligned} m!\,(k(k+1))^{l-k} &\leqslant \mathbb{E}|(\xi_i, \xi_j)|^{2l} \leqslant 2^{2l} m!\,(k(k+1))^{l-k} \\ &\qquad \text{if} \quad m = 2k \\ m!\,k(k(k+1))^{l-k-1} &\leqslant \mathbb{E}|(\xi_i, \xi_j)|^{2l} \leqslant 2^{2l} m!\,k(k+1))^{l-k-1} \\ &\qquad \text{if} \quad m = 2k-1 \end{aligned} \tag{2.29}$$

(iii) The odd moments are bounded in terms of the even ones through

$$(\mathbb{E}|(\xi_i, \xi_j)|^{2l-2})^{1+1/(2l-2)} \leqslant \mathbb{E}|(\xi_i, \xi_j)|^{2l-1} \leqslant (\mathbb{E}|(\xi_i, \xi_j)|^{2l})^{1-1/(2l)} \tag{2.30}$$

(iv) Moreover, for the first two moments we have the exact formulas

$$\mathbb{E}|(\xi_i, \xi_j|^2 = m, \qquad \mathbb{E}|(\xi_i, \xi_j)| = \frac{2m}{2^m}\binom{m-1}{[m/2]} \sim \left(\frac{2}{\pi}\right)^{1/2} m^{1/2} \tag{2.31}$$

Proof. Notice first that point (iii) simply follows from Jensen's inequality.[6] The even moments are easier to compute since in this case the absolute value may be dropped, and since for $i \neq j$, $\xi_i^\mu \xi_j^\mu$ has the same distribution as ξ_1^μ, so that

$$\mathbb{E}|(\xi_i, \xi_j)|^{2l} = \mathbb{E}\left(\sum_{\mu=1}^{m} \xi_1^\mu\right)^{2l} \tag{2.32}$$

But since ξ_1^μ are i.i.d. symmetric Bernoulli, the moment generating function for the r.v. $\sum_{\mu=1}^{m} \xi_1^\mu$ is $(\cosh x)^m$, and thus

$$\mathbb{E}\left(\sum_{\mu=1}^{m} \xi_1^\mu\right)^{2l} = \frac{d^{2l}}{dx^{2l}} \cosh^m x \bigg|_{x=0} \tag{2.33}$$

Thus, we just need to estimate the $2l$th derivative of $\cosh^m x$. Let us put $C_s(x) \equiv \cosh^{m-s} x \sinh^s x$. Since

$$\frac{d}{dx} C_s = (m-s) C_{s+1} + s C_{s-1} \tag{2.34}$$

it is natural to label each term appearing in the $2l$th derivative by a random walk ω of length $2l$ on $\{0, 1, 2,..., m\}$. Moreover, since at the end we must set $x = 0$, only such walks will give a nonzero contribution which finally produce a C_0, i.e., we count only walks starting at zero and ending at zero. Finally, we define the weight of each step of the walk by

$$w(\omega_{t+1}, \omega_t) = \begin{cases} m - \omega_t & \text{if } \omega_{t+1} - \omega_t = 1 \\ \omega_t & \text{if } \omega_{t+1} - \omega_t = -1 \end{cases} \tag{2.35}$$

Then we have that

$$\frac{d^{2l}}{dx^{2l}} \cosh^m x \bigg|_{x=0} = \sum_{\omega: 0 \to 0} \prod_{t=1}^{2l} w(\omega_t, \omega_{t-1}) \tag{2.36}$$

Now, since ω must contain the same number of steps going up as going down, we may pair them in such a way that to each step going up at, say, time t (and starting at ω_t), we associate the next step down starting at the position $\omega_{t'} = \omega_t + 1$. Notice that such a time t' will necessarily exist. Now the weight of each such pair is $(m - \omega_t)\omega_{t'} = (m - \omega_t)(\omega_t + 1)$, and the weight of the walk is the product over all pairs of these quantities. The important observation is now that the function $(m - x)(x + 1)$ attains its maximum at $x=(m-1)/2$, and therefore the walk with highest weight is simply the one for which ω_t is as close to this value as possible under the constraints that $\omega_0=\omega_{2l}=0$. It is trivial ro see that such a walk will have the weights corresponding to the lower bounds in (2.28), (2.29).

The upper bounds are simply obtained by multiplying the highest weights by the trivial upper bound 2^{2l} for the number of contributing walks [this could be slightly improved to $\binom{2l}{l}$].

Finally, the exact formulas (2.31) for the first two moments are obtained by standard computations. This concludes the proof of Lemma 2.3. ∎

The reader will now verify that if we were to replace all powers of $|(\xi_i, \xi_j)|$ in (2.27) by the respective moments, this would indeed yield an exponentially small value for Q_N. Our next step will therefore consist in proving that the fluctuations of the powers of $|(\xi_i, \xi_j)|$ about their expectations are sufficiently small. More precisely, we want to control the probabilities

$$P_n(\delta) \equiv \mathbb{P}_\xi\left(\sum_{i \neq j} [|(\xi_i, \xi_j)|^n - \mathbb{E}|(\xi_i, \xi_j)|^n] \geqslant \delta^n N^2 \mathbb{E}|(\xi_1, \xi_2)|^n\right) \quad (2.37)$$

Note that the obvious bound

$$P_n(\delta) \leqslant N(N-1)\, \mathbb{P}_\xi(|(\xi_1, \xi_2)|^n - \mathbb{E}|(\xi_1, \xi_2)|^n \geqslant \delta^n \mathbb{E}((\xi_1, \xi_2)|^n) \quad (2.38)$$

would be a disaster, as the last probabilities in (2.38) do not depend on N and thus will never offset the N^2 prefactor. To improve it, we must exhibit some independence of the terms appearing in the sum over i and j. To do so, we go only halfway toward (2.38), i.e., we notice that

$$P_n(\delta) \leqslant N \mathbb{P}_\xi\left(\sum_{i \neq 1} [|(\xi_i, \xi_1)|^n - \mathbb{E}|(\xi_i, \xi_1)|^n] \geqslant \delta^n N \mathbb{E}|(\xi_1, \xi_2)|^n\right) \quad (2.39)$$

The terms in the remaining sum are now independent. To obtain a bound on P_n that behaves like $1/N^2$, we now use the sixth-order Chebychev inequality to bound the probabilities in (2.39). This gives

$$\begin{aligned} &\mathbb{P}_\xi\left(\sum_{i \neq 1} [|(\xi_i, \xi_1)|^n - \mathbb{E}|(\xi_i, \xi_1)|^n] \geqslant \delta^n N \mathbb{E}|(\xi_1, \xi_2)|^n\right) \\ &\quad \leqslant \frac{\mathbb{E}(\sum_{i \neq 1} [|(\xi_i, \xi_1)|^n - \mathbb{E}|(\xi_i, \xi_1)|^n])^6}{N^6 \delta^{6n} (\mathbb{E}|(\xi_1, \xi_2)|^n)^6} \end{aligned} \quad (2.40)$$

Let us put $a_i \equiv |(\xi_i, \xi_1)|^n - \mathbb{E}|(\xi_i, \xi_1)|^n$. Then, since $\mathbb{E}a_i = 0$,

$$\mathbb{E}\left(\sum_i a_i\right)^6 = \sum_i \mathbb{E}a_i^6 + \binom{6}{2}\sum_{i\neq j}\mathbb{E}a_i^2\mathbb{E}a_j^4$$
$$+\binom{6}{3}\sum_{i\neq j}\mathbb{E}a_i^3\mathbb{E}a_j^3$$
$$+\binom{6}{2}\binom{4}{2}\sum_{i\neq j\neq k}\mathbb{E}a_i^2\mathbb{E}a_j^2\mathbb{E}_k^2$$
$$= N\mathbb{E}a_2^6 + 15N(N-1)\,\mathbb{E}a_2^2\mathbb{E}a_2^4 + 20N(N-1)(\mathbb{E}a_2^3)^2$$
$$+90N(N-1)(N-2)(\mathbb{E}a_2^2)^3 \tag{2.41}$$

With these preliminaries it is now an easy matter to prove the following.

Lemma 2.4. There exists a finite positive constant C such that, uniformly in m, n, and N,

$$\frac{\mathbb{E}(\sum_{i\neq 1}[|(\xi_i,\xi_1)|^n - \mathbb{E}|(\xi_i,\xi_1)|^n])^6}{(\mathbb{E}|(\xi_1,\xi_2)|^n)^6} \leqslant N^3C^n \tag{2.42}$$

Proof. Note that from (2.41) it follows that the quantity on the left of (2.42) can be expressed as a finite sum of terms of the form

$$c_{a,i}N^i\frac{\mathbb{E}|(\xi_2,\xi_1)|^{na}}{[\mathbb{E}|(\xi_2,\xi_1)|^n]^a} \tag{2.43}$$

with $a=2, 3, 4, 6$; $i=1, 2, 3$; and $c_{a,i}$ finite numerical constants. Using the upper and lower bounds from Lemma 2.3, one easily checks that the ratios of expectations in (2.43) are all bounded by const^n, uniformly in m. But this yields the claim of Lemma 2.4. ▮

From Lemma 2.4 we can now deduce the following result.

Lemma 2.5. There exist finite positive constants δ and K such that

$$\mathbb{P}_\xi\left(\exists_{n\geqslant 3}\colon \sum_{i\neq j}[|(\xi_i,\xi_j)|^n - \mathbb{E}|(\xi_i,\xi_j)|^n] \geqslant \delta^n N^2\mathbb{E}|(\xi_1,\xi_2)|^n\right) \leqslant \frac{K}{N^2} \tag{2.44}$$

Proof. Just notice that by (2.39) and Lema 2.4

$$\mathbb{P}_\xi\left(\exists_{n\geqslant 3}\colon \sum_{i\neq j}[|(\xi_i,\xi_j)|^n - \mathbb{E}|(\xi_i,\xi_j)|^n] \geqslant \delta^n N^2\mathbb{E}|(\xi_1,\xi_2)|^n\right)$$
$$\leqslant \sum_{n\geqslant 3}\left(\sum_{i\neq j}[|(\xi_i,\xi_j)|^n - \mathbb{E}|(\xi_i,\xi_j)|^n] \geqslant \delta^n N^2\mathbb{E}|(\xi_1,\xi_2)|^n\right)$$
$$\leqslant \sum_{n\geqslant 3}\frac{1}{N^2}\left(\frac{C}{\delta}\right)^n \tag{2.45}$$

from which (2.44) follows if δ is chosen such that $C/\delta < 1$. ▮

Let us now define the event $\mathscr{A}_N \in \mathscr{F}_\xi$ as follows:

Definition 2.1. $\omega \in \mathscr{A}_N$, iff $\xi \equiv \xi(\omega)$ satisfies:

(i) $\forall_{n\geqslant 3} \quad \sum_{i\neq j} |(\xi_i, \xi_j)|^n \leqslant (1+\delta^n)\, N^2 \mathbb{E}|(\xi_1, \xi_2)|^n$

with δ chosen as in Lemma 2.5.

(ii) $(1-\varepsilon)\, N^2 m \leqslant \sum_{i\neq j} |(\xi_i, \xi_j)|^2 \leqslant (1+\varepsilon)\, N^2 m$

(iii) $$(1-\varepsilon) N^2 \frac{2m}{2^m} \binom{m-1}{[m/2]} \leqslant \sum_{i\neq j} |(\xi_i, \xi_j)| \leqslant (1+\varepsilon) N^2 \frac{2m}{2^m} \binom{m-1}{[m/2]} \quad \text{for some } \varepsilon < \tfrac{1}{2}$$

We define further the set

$$\mathscr{A} \equiv \bigcup_{N_0=1}^{\infty} \bigcap_{N\geqslant N_0} \mathscr{A}_N \tag{2.46}$$

From the previous results it follows already that $\mathscr{A}_N$ has large probability:

Lemma 2.6. The exists a constant $0 < K < \infty$ such that

$$\mathbb{P}_\xi(\mathscr{A}_N) \geqslant 1 - \frac{K}{N^2} \tag{2.47}$$

Moreover,

$$\mathbb{P}_\xi(A) = 1 \tag{2.48}$$

Proof. The bound (2.47) is easily pieced together from the previous lemmata. (2.48) follows from (2.47) and the Borel–Cantelli Lemma.[6] ▌

Now the event $\mathscr{A}_N$ was constructed in such a way as to ensure that (2.27) is small. More precisely, we have the following result.

Lemma 2.7. On the set $\mathscr{A}_N$, we have, for any function γ such that $\gamma(N) \downarrow 0$, as $N \uparrow \infty$, the following:

(i) $$t_u^* \leqslant 3 \frac{2\gamma}{\sqrt{m}} \left(1 - \frac{m}{N}\right)^{-1} \tag{2.49}$$

$$\text{(ii)}\quad \frac{[\sum_{i\neq j}|(\xi_i,\xi_j)|-mN]^2}{\sum_{i\neq j}|(\xi_i,\xi_j)|^2}\geqslant\frac{1}{2}\frac{[-mN+(1-\varepsilon)(2m/\pi)^{1/2}N^2]^2}{(1+\varepsilon)\,mN^2}$$

$$\geqslant\frac{1}{6\pi}N^2+\frac{1}{3}m-\frac{2^{1/2}}{3\pi^{1/2}}Nm^{1/2} \tag{2.50}$$

$$\text{(iii)}\quad \sum_{i\neq j}R_3(t^*|(\xi_i,\xi_j)|)\leqslant\gamma^3KN^2 \tag{2.51}$$

Remark. It should be noted here that no assumptions are made in this lemma on the speed with which $\gamma(N)$ tends to zero. This is, as we will see shortly, essential in order to get the weakest possible assumptions on $p(N)$. This renders our proof somewhat more complicated. A simpler proof can be found under the assumption that $\gamma(N)(\ln N)^{1/2}\downarrow 0$.

Proof. The proofs of (i) and (ii) are fairly immediate, using point (iv) of Lemma 2.3 (and assuming m large, for simplicity). To prove (iii), just notice that on $\mathscr{A}_N$,

$$\begin{aligned}
&\sum_{i\neq j}R_3(t_u^*|(\xi_i,\xi_j)|)\\
&\quad=\sum_{n\geqslant 3}\frac{(t_u^*)^n}{n!}\sum_{i\neq j}|(\xi_i,\xi_j)|^n\\
&\quad\leqslant\sum_{n\geqslant 3}\frac{(t_u^*)^n}{n!}(1+\delta^n)\,N^2\mathbb{E}|(\xi_1,\xi_2)|^n\\
&\quad\leqslant\sum_{l=2}^{m/2}(1+\delta^{2l-1})N^2\frac{(t_u^*)^{2l-1}}{(2l-1)!}\left(2^{2l}\frac{m!\,l!}{(m-l)!}\right)^{(2l-1)/2l}\\
&\qquad+\sum_{l=2}^{m/2}(1+\delta^{2l})N^2\frac{(t_u^*)^{2l}}{(2l)!}2^{2l}\frac{m!\,l!}{(m-l)!}\\
&\qquad+\sum_{l=m/2+1}^{\infty}(1+\delta^{2l-1})N^2\frac{(t_u^*)^{2l-1}}{(2l-1)!}\left(2^{2l}m!\left(\frac{m+1}{2}\right)^{2l-m}\right)^{(2l-1)/2l}\\
&\qquad+\sum_{l=m/2+1}^{\infty}(1+\delta^{2l})N^2\frac{(t_u^*)^{2l}}{(2l)!}2^{2l}m!\left(\frac{m}{2}\right)^{2l-m}
\end{aligned} \tag{2.52}$$

Now we will always assume that m/N goes to zero as N goes to infinity. Therefore, t_u^* is effectively bounded by, say, $7\gamma/\sqrt{m}$, for N large enough. Moreover, γ will be taken to zero with N, so that we may assume it to be as small as desired. It is then a trivial matter to realize that all four sums in (2.52) converge and that moreover

$$\sum_{l=2}^{m/2} (1+\delta^{2l-1}) N^2 \frac{(t_u^*)^{2l-1}}{(2l-1)!} \left(2^{2l} \frac{m!\, l!}{(m-l)!}\right)^{(2l-1)/2l}$$
$$\leqslant N^2 \sum_{l=2}^{m/2} (1+\delta^{2l-1})\, 6^{2l-1} \gamma^{2l-1} \frac{l!^{(2l-1)/2l}}{(2l-1)!}$$
$$\leqslant C_1 N^2 \gamma^3 \tag{2.53}$$

and similarly

$$\sum_{l=2}^{m/2} (1+\delta^{2l}) N^2 \frac{(t_u^*)^{2l}}{(2l)!} 2^{2l} \frac{m!\, l!}{(m-l)!} \leqslant C_2 N^2 \gamma^4 \tag{2.54}$$

while the last two sums are bounded by

$$C_3(1+\delta^m)\, 7^m e^m \gamma^m \tag{2.55}$$

and are thus completely negligible. Combining these bounds yields (iii). ▮

We are finally ready to merge these results into a bound for Q_N:

Proposition 2.8. Assume that $m/N \downarrow 0$ and $pN \uparrow \infty$ as $N \uparrow \infty$ and choose γ such that $pN\gamma^2 > c$ for some contant $0 < c < \infty$. Then, for $\omega \in \mathscr{A}_N$ and for N sufficiently large, there exists $\rho > 0$ such that

$$Q_N(\xi) \leqslant e^{-\rho N} \tag{2.56}$$

Proof. Inserting the bounds from the previous lemma into (2.27), we get that

$$Q_N(\xi) \leqslant 2^N \exp\left(-\frac{1}{32} pN^2\gamma^2 + pN^2 K\gamma^3\right)$$
$$= \exp\left(-\frac{pN^2\gamma^2}{32}(1-32K\gamma) + N \ln 2\right) \tag{2.57}$$

Choosing now N large enough and $\gamma(N)$ such that

$$pN\gamma^2 > \frac{32}{1-32K\gamma} \ln 2 \tag{2.58}$$

we get the bound (2.56) ▮

From this proposition and Lemma 2.6 we now get immediately Theorem 3. ▮▮

3. CONVERGENCE OF THE FREE ENERGY

In this section we discuss the consequences of the uniform bounds obtained in the previous section for the convergence of the free energy of the dilute Hopfield model. Let us denote by $f_{N,\beta}(\xi)$ the free energy of the standard Hopfield model, and let us introduce

$$\Delta f_{N,\beta}(\xi;\varepsilon) \equiv |f_{N,\beta}(\xi;\varepsilon) - f_{N,\beta}(\xi)| \tag{3.1}$$

We have the following result.

Proposition 3.1. Assume that $p(N)N \uparrow \infty$ and $m(N)/p(N)N \downarrow 0$, as $N \uparrow \infty$. Then, for all β ,

$$\lim_{N \uparrow \infty} \Delta f_{N,\beta}(\xi;\varepsilon) = 0, \qquad \mathbb{P}_\xi \times \mathbb{P}_\varepsilon\text{-a.s.} \tag{3.2}$$

Proof. By Theorem 3 there exists an event $\mathscr{C}_N \in \mathscr{F}_\xi \times \mathscr{F}_\varepsilon$ such that

$$\mathbb{P}(\mathscr{C}_N) \geqslant \left(1 - \frac{K}{N^2}\right)(1 - e^{-\rho N}) \tag{3.3}$$

such that on $\mathscr{C}_N$, for all $\sigma \in \mathscr{S}^N$,

$$|H_N(\xi;\varepsilon,\sigma) - \mathbb{E}_\varepsilon H_N(\xi;\varepsilon;\sigma)| < \gamma(N)[m(N)]^{1/2} N \tag{3.4}$$

for any decreasing function γ satisfying $\gamma(N) \downarrow 0$ as $N \uparrow \infty$ and $p(N)\, N\gamma(N)^2 > c$ for some constant $0 < c < \infty$. But $\mathbb{E}_\varepsilon H_N(\xi;\varepsilon;\sigma)$ is nothing but $H_N(\xi;\sigma)$, and hence a trivial calculation shows that (3.4) implies that

$$\Delta f_{N,\beta}(\xi;\varepsilon) \leqslant \gamma(N)[m(N)]^{1/2} \tag{3.5}$$

If we moreover choose γ such that $\gamma(N)\sqrt{m} \downarrow 0$ as $N \uparrow \infty$, setting $\mathscr{C} \equiv \bigcup_{N_0 \geqslant 0} \bigcap_{N \geqslant N_0} \mathscr{C}_N$, we see immediately that on $\mathscr{C}$,

$$\lim_{N \uparrow \infty} \Delta f_{N,\beta}(\xi;\varepsilon) = \lim_{N \uparrow \infty} \gamma(N)[m(N)]^{1/2} = 0 \tag{3.6}$$

Now combining the constraints on γ gives the condition $m/pN \downarrow 0$ as $N \uparrow \infty$, while (3.3) and the Borel–Cantelli Lemma imply that $\mathbb{P}(\mathscr{C}) = 1$, which proves the proposition. ▌

Therefore, to prove Theorem 1, we just need to prove the almost sure convergence of the free energy of the standard Hopfield model. Now in a recent paper, Koch[15] has shown that under the assumption that $m/N \downarrow 0$,

$$\mathbb{E}_\xi f_{N,\beta}(\xi) \to f_{\mathrm{CW}}(\beta) \tag{3.7}$$

Tirozzi and Shcherbina[19] also very recently proved this convergence in probability (with a bound on the probabilities that cannot yield almost sure convergence). As a matter of fact, it is very easy to modify the approach of Koch to prove the almost sure convergence (this would even seem a more natural consequence of his computations). Let us state this result and give the proof for completeness.

Theorem 3.2. Assume that $m(N)/N \downarrow 0$ as $N \uparrow \infty$. Then

$$\lim_{N \uparrow \infty} f_{N,\beta}(\xi) = f_{\mathrm{CW}}(\beta), \qquad \mathbb{P}_{\xi}\text{-a.s.} \tag{3.8}$$

Proof. We follow essentially the analysis of ref. 15. The first step consists in rewriting the partition function in terms of Gaussian integrals in a standard way:

$$\begin{aligned} &2^{-N} \sum_{\sigma \in \mathcal{S}^N} \exp[-\beta H_{N,\xi}(\sigma)] \\ &\quad = \left(\frac{N\beta}{2\pi}\right)^{m/2} \int_{\mathbb{R}^m} d^m z \exp\left[-\tfrac{1}{2} N\beta(z, z) + \sum_{i=1}^{N} \ln\cosh(\beta(z, \xi_i))\right] \end{aligned} \tag{3.9}$$

where z is an m-componenty vector and $(z, \xi_i) \equiv \sum_{\mu=1}^{m} z_\mu \xi_i^\mu$, etc. Now notice that

$$\begin{aligned} -\frac{1}{2} & N\beta(z, z) + \sum_{i=1}^{N} \ln\cosh(\beta(z, \xi_i)) \\ &= -\frac{1}{2} N\beta(z, z) + \sum_{i=1}^{N} \frac{1-\delta}{2} \beta(z, \xi_i)^2 \\ &\quad + \sum_{i=1}^{N} \left(\ln\cosh(\beta(z, \xi_i)) - \frac{1-\delta}{2} \beta(z, \xi_i)^2\right) \\ &\leqslant -\frac{1}{2} N\beta(z, z) + \sum_{i=1}^{N} \frac{1-\delta}{2} \beta(z, \xi_i)^2 \\ &\quad + N \max_{x \in \mathbb{R}} \left(\ln\cosh(\beta x) - \frac{1-\delta}{2} \beta x^2\right) \\ &= -\frac{1}{2} N\beta(z, \mathbb{1} - (1-\delta) A) z) \\ &\quad + N \max_{x \in \mathbb{R}} \left(\ln\cosh(\beta x) - \frac{1-\delta}{2} \beta x^2\right) \end{aligned} \tag{3.10}$$

where the $m \times m$ matrix A has components $A_{\mu\nu} = (1/N) \sum_{i=1}^{N} \xi_i^\mu \xi_i^\nu$. Of course, this decomposition is only useful for a choice of δ such that the matrix $(\mathbb{1} - (1-\delta)A)$ is strictly positive. *If* this is the case, then inserting this inequality into (3.9) gives the following upper bound on the partition function:

$$Z_{N,\xi} \leqslant \left\{ \exp\left[N \max_x \left(\ln \cosh(\beta x) - \frac{1-\delta}{2} \beta x^2 \right) \right] \right\} \times \det(\mathbb{1} - (1-\delta)A)^{-1/2}$$
$$\leqslant \left\{ \exp\left[N \max_x \left(\ln \cosh(\beta x) - \frac{1-\delta}{2} \beta x^2 \right) \right] \right\} \times (\lambda_{\min}(\mathbb{1} - (1-\delta)A))^{-m/2} \qquad (3.11)$$

where $\lambda_{\min}(M)$ denotes the smallest eigenvalue of the matrix M. (3.11) yields immediately the lower bound

$$f_{N,\beta}(\xi) \geqslant -\beta^{-1} \max_x \left(\ln \cosh(\beta x) - \frac{1-\delta}{2} \beta x^2 \right) + \frac{m}{2\beta N} \ln(\lambda_{\min}(\mathbb{1} - (1-\delta)A)) \qquad (3.12)$$

If we could choose $\delta = \delta(N)$ in such a way that $\delta(N) \downarrow 0$ as $N \uparrow \infty$, this lower bound would converge to the Curie–Weiss free energy, and since the Curie–Weiss free energy is trivially an upper bound for the Hopfield free energy, this would give the desired convergence. The following proposition tells us that with probability one this is indeed the case.

Proposition 3.3. Let $\lambda_{\max}(A)$ denote the largest positive eigenvalue of A. Then, for any constant c and for N large enough, we have that

$$\mathbb{P}(\lambda_{\max}(A) > e^{2(m/N)^{1/2}}(1 + (1+c) N^{-1/6} \ln N)) \leqslant 2N^{-c} \qquad (3.13)$$

Bounds on largest eigenvalues of random matrices can be found, for instance, in ref. 13. They prove results like (3.13) for symmetric matrices with i.i.d. entries. Their method is in fact well suited to be adapted to the present situation. The basic input into the proof of (3.13) is the following bound on the trace of the powers of the matrix A:

Lemma 3.4. *Let* $k \leqslant N^{1/6}$. Then

$$\mathbb{E}_\xi \operatorname{tr} A^k \leqslant 2N e^{2k\rho/(1+\rho)} \qquad (3.14)$$

where $\rho = (m/N)^{1/2}$.

Remark. Koch[15] and Tirozzi and Shcherbina[19] announced analogous bounds on the traces of $(A-\mathbb{1})^k$. We present a proof the lines of ref. 13 in an appendix. Let us also note that our proof has the additional virtue that it also holds when the ξ_i^μ are centered but not necessarily symmetric random variables.

Let us now show how Lemma 3.4 implies Proposition 3.3.

Proof of Proposition 3.3 (using Lemma 3.4). Notice first that $\lambda_{\max}(A)^k \leqslant \operatorname{tr} A^k$ for all k (note that A is a positive matrix). Thus, using the Chebychev inequality and Lemma 3.4, for all $k \leqslant N^{1/6}$

$$\begin{aligned}
\mathbb{P}(\lambda_{\max}(A) > e^{2\rho/(1+\rho)}(1+x)) &\leqslant \mathbb{P}(\operatorname{tr}(A^k) > (e^{2\rho/(1+\rho)}(1+x))^k) \\
&\leqslant \frac{\mathbb{E}\operatorname{tr}(A^k)}{(e^{2\rho/(1+\rho)}(1+x))^k} \\
&\leqslant N\frac{2e^{2k\rho/(1+\rho)}}{(e^{2\rho/(1+\rho)}(1+x))^k} \\
&= 2N\left(1-\frac{x}{1+x}\right)^k \\
&\leqslant 2N\exp\left\{-k\frac{x}{1+x}\right\}
\end{aligned} \tag{3.15}$$

Now we choose k as large as possible, i.e., $k=N^{1/6}$ and $x=(1+c)N^{-1/6}\ln N$. Then (3.15) yields

$$\begin{aligned}
&\mathbb{P}(\lambda_{\max}(A) > e^{2\rho/(1+\rho)}(1+x)) \\
&\quad\leqslant 2N\exp\left\{-N^{1/6}\frac{(1+c)N^{-1/6}\ln N}{1+(1+c)N^{-1/6}\ln N}\right\} \\
&\quad\sim 2NN^{-(1+c)}
\end{aligned} \tag{3.16}$$

which proves the proposition. ∎

We are now ready to prove Theorem 3.2.

Proof of Theorem 3.2. By Proposition 3.3, and using the triangle inequality, we see that with probability greater than $1-2N^{-2}$,

$$\begin{aligned}
\lambda_{\min}(\mathbb{1}-(1-\delta)A) &\geqslant 1-(1-\delta)(e^{2(m/N)^{1/2}}(1+3N^{-1/6}\ln N)) \\
&\sim \delta-(1-\delta)(2(m/N)^{1/2}+3N^{-1/6}\ln N)
\end{aligned} \tag{3.17}$$

so that we may choose $\delta(N) = 2(m/N)^{1/2} + 3N^{-1/6} \ln N$, which tends to zero with N, and get that

$$f_{N,\xi}(\beta) \geqslant -\frac{1}{\beta} \max_{\rho} \left(\ln \cosh(\beta\rho) - \frac{1-\delta(N)}{2} \beta\rho^2 \right) + \frac{m}{N} \ln \left[2\left(\frac{m}{N}\right)^{1/2} + 3N^{-1/6} \ln N \right]^2 \tag{3.18}$$

Now the last term in (3.18) goes to zero with N, while the first converges (by continuity) to

$$\max_{\rho}(\ln \cosh(\beta\rho) - \tfrac{1}{2}\beta\rho^2) = f_{\mathrm{CW}}(\beta)$$

Since (3.18) holds on an event whose complement has summable probability, a standard Borel–Cantelli argument as in previous instances yields convergence on a set of full measure. ▮

Theorem 1 is now a direct corollary of Proposition 3.1 and Theorem 3.2. ▮▮

4. LIMIT DISTRIBUTIONS OF THE OVERLAP PARAMETERS

In this section we will assume that $m(N) < \ln N/\ln 2$, which is the restriction under which the analog of Theorem 2 could be proved in the standard Hopfield model.[14,10]

We will now give the proof of Theorem 2. Since it will closely follow that given in ref. 10 and makes direct use of some results established therein, in order to be concise, we will only stress the aspects due to the dilution and refer for details to ref. 10.

One main ingredient of the proof consists in a random partition of the set $\Lambda = \{1,..., N\}$ which can be briefly described as follows. Let us fix an arbitrary enumeration of the $d = |\mathcal{S}^m| = 2^m$ elements of the set $\mathcal{S}^m$ and write

$$\mathcal{S}^m \equiv \{e_1,..., e_k,..., e_d\} \tag{4.1}$$

with $e_k = (e_k^1,..., e_k^\mu,..., e_k^m)$. For all $\mu = 1,..., m$, we denote by e^μ the d-component vector $e^\mu = (e_1^\mu,..., e_k^\mu,..., e_d^\mu)$. Note that the vectors e^μ are orthogonal to each other, i.e.,

$$\frac{1}{d}(e^\mu, e^\nu) = \delta_{\mu,\nu} \tag{4.2}$$

Now any given realization of the r.v. ξ together with the enumeration (4.1) induces a random partition of the set Λ into d disjoint (possibly empty) subsets $\Lambda_k(\xi)$, defined as

$$\Lambda_k(\xi)=\{i\in\Lambda\colon \xi_i=e_k\}, \qquad k=1,..., d \tag{4.3}$$

The random partition (4.3) has the property that, for $m<\ln N/\ln 2$, the cardinality of each subset $\Lambda_k(\xi)$ remains close to its mean value N/d. More precisely, remembering that $\xi\equiv\xi(\omega)$ is a r.v. on the probability space $(\Omega_\xi, \mathscr{F}_\xi, \mathbb{P}_\xi)$, we recall from ref. 10 the following result.

Lemma 4.1.[10,14] Define the event $\mathscr{D}_N\in\mathscr{F}_\xi$ as

$$\mathscr{D}_N=\left\{\omega\in\Omega_\xi \,\middle|\, |\Lambda_k(\xi)|\equiv\frac{N}{d}(1+\lambda_k),\ |\lambda_k|<\delta(N),\ 1\leqslant k\leqslant d\right\} \tag{4.4}$$

where $\delta(N)=(d/N)^{1/2}\ln N$ tends to zero as N tends to infinity. Then $\exists N_0$ such that $\forall N>N_0$

$$\mathbb{P}_\xi(\mathscr{D}_N)\geqslant 1-2d\exp\left\{-\frac{N}{2d}\left(1-\frac{1}{d}\right)^{-1}\delta^2(N)\right\} \tag{4.5}$$

Now let us define the map X_ξ:

$$X_\xi\colon\ \mathscr{S}^N\mapsto\Xi_\xi=\prod_{k=1}^{d}\left\{-1,\ -1+\frac{2}{|\Lambda_k(\xi)|},...,\ 1-\frac{2}{|\Lambda_k(\xi)|},\ 1\right\}$$

$$\sigma\to\begin{cases}X_\xi(\sigma)=(X_{\xi,1}(\sigma),..., X_{\xi,k}(\sigma),..., X_{\xi,d}(\sigma))\\ X_{\xi,k}(\sigma)=\dfrac{1}{|\Lambda_k(\xi)|}\displaystyle\sum_{i\in\Lambda_k(\xi)}\sigma_i\end{cases} \tag{4.6}$$

Note that $X_\xi(\sigma)$ is a d-component vector. Using the random partition of Λ, we can rewrite the overlap parameters in terms of $X_\xi(\sigma)$ as

$$\begin{aligned}m_N^\mu(\sigma,\xi)&=\frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu\sigma_i\\&=\frac{1}{N}\sum_{k=1}^{d}e_k^\mu\sum_{i\in\Lambda_k(\xi)}\sigma_i\\&=\frac{1}{d}\sum_{k=1}^{d}e_k^\mu(1+\lambda_k)\,X_{\xi,k}(\sigma)\\&=\frac{1}{d}(e^\mu,\,[M+\mathbb{1}]\,X_\xi(\sigma))\end{aligned} \tag{4.7}$$

where M is the $d\times d$ diagonal matrix with entries $M_{kk}=\lambda_k$. Thus, given any $x\in\Xi_\xi$ the overlap $m_N^\mu(\sigma,\xi)$, regarded as a function of the configurations σ, takes the value

$$m_N^\mu(\sigma;\xi)=\frac{1}{d}(e^\mu,[M+\mathbb{1}]x) \tag{4.8}$$

for all σ in the subset $\{\sigma\in\mathcal{S}^N \mid X_\xi(\sigma)=x\}$.

From now on we assume that $h\geqslant 0$. To prove Theorem 2, it is enough to show that under its assumptions, for any continuous function g: $[-1,1]\to\mathbb{R}$,

$$\lim_{h\downarrow 0}\lim_{N\uparrow\infty}\sum_{\sigma\in\mathcal{S}^N} g(m_N^\mu(\sigma;\xi))\,\mathcal{G}_{N,\beta,h}^\alpha(\varepsilon;\xi;\sigma)$$
$$=\begin{cases} g(0) & \text{if } 0\leqslant\beta\leqslant 1\\ g(a(\beta)\delta_{\mu,\nu}) & \text{if } \beta\geqslant 1\end{cases}\quad \mathbb{P}_\xi\times\mathbb{P}_\varepsilon\text{-a.s.} \tag{4.9}$$

By (4.9) we have

$$\sum_{\sigma\in\mathcal{S}^N} g(m_N^\mu(\sigma;\xi))\,\mathcal{G}_{N,\beta,h}^\alpha(\varepsilon;\xi;\sigma)=\sum_{x\in\Xi_\xi} g\left(\frac{1}{d}(e^\mu,[M+\mathbb{1}]\,x)\right)\nu_{N,\beta,h}^\alpha(\varepsilon;\xi;x) \tag{4.10}$$

where $\nu_{N,\beta,h}^\alpha(\varepsilon;\xi)$ is the probability measure on Ξ_ξ induced by $\mathcal{G}_{N,\beta,h}^\alpha(\varepsilon;\xi)$ through the map X_ξ which to each $x\in\Xi_\xi$ assigns the probability

$$\nu_{N,\beta,h}^\alpha(\varepsilon;\xi;x)=\sum_{\substack{\sigma\in\mathcal{S}^N\\ X_\xi(\sigma)=x}}\mathcal{G}_{N,\beta,h}^\alpha(\varepsilon;\xi;\sigma) \tag{4.11}$$

Thus, to compute the expectation (4.11), we are left to study the measure $\nu_{N,\beta,h}^\alpha(\varepsilon;\xi)$.

Let us denote by $\mathcal{G}_{N,\beta,h}^\alpha(\xi;\sigma)$ the finite-volume Gibbs measure associated to the mean Hamiltonian

$$\mathbb{E}_\varepsilon H_{N,h}^\alpha(\varepsilon;\xi;\sigma)=H_N(\xi;\sigma)-h(\sigma,\xi^\alpha) \tag{4.12}$$

that is to say, the Hamiltonian of the standard Hopfield model with a magnetic field coupling to the pattern ξ^μ. Let $\tilde{\nu}_{N,\beta,h}^\alpha(\varepsilon;\xi)$ be the measure induced by $\tilde{\mathcal{G}}_{N,\beta,h}^\alpha(\xi;\sigma)$ under the map X_ξ. The following lemma presents a bond on the density $\nu_{N,\beta,h}^\alpha(\varepsilon;\xi;x)$ in terms of the density $\tilde{\nu}_{N,\beta,h}^\alpha(\xi;x)$.

Lemma 4.2. There exists an event $\mathcal{C}_N\in\mathcal{F}_\xi\times\mathcal{F}_\varepsilon$ such that on $\mathcal{C}_N$, for all $x\in\Xi_\xi$,

$$e^{-2\beta N\sqrt{m\gamma(N)}}\,\tilde{\nu}_{N,\beta,h}^\alpha(\xi;x)\leqslant\nu_{N,\beta,h}^\alpha(\varepsilon;\xi;x)\leqslant e^{2\beta N\sqrt{m\gamma(N)}}\,\tilde{\nu}_{N,\beta,h}^\alpha(\xi;x) \tag{4.13}$$

where $\gamma(N)$ is chosen as in Theorem 3. Moreover, for N large enough

$$\mathbb{P}(\mathscr{C}_N) \geqslant \left(1 - \frac{K}{N^2}\right)(1 - e^{-\rho N}) \tag{4.14}$$

where ρ and K are positive constants.

Proof. By Theorem 3 there exists an event $\mathscr{C}_N \in \mathscr{F}_\xi \times \mathscr{F}_\varepsilon$ whose probability satisfies the bound (4.14) such that on $\mathscr{C}_N$, for all $\sigma \in \mathscr{S}^N$ and for any function γ satisfying $\gamma(N) \downarrow 0$ as $N \uparrow \infty$ and $pN\gamma^2 > c$ for some constant $0 < c < \infty$,

$$|H_N(\varepsilon; \xi; \sigma) - \mathbb{E}_\varepsilon H_N(\varepsilon; \xi; \sigma)| \leqslant \gamma(N)\sqrt{m}N \tag{4.15}$$

Now note that

$$|H^\alpha_{N,h}(\varepsilon; \xi\sigma) - \mathbb{E}_\varepsilon H^\alpha_{N,h}(\varepsilon; \xi; \sigma)| = |H_N(\varepsilon; \xi; \sigma) - \mathbb{E}_\varepsilon H_N(\varepsilon; \xi; \sigma)| \tag{4.16}$$

and (4.15) and (4.16) together with the definitions of $\tilde{\nu}^\alpha_{N,\beta,h}(\xi, x)$ and $\nu^\alpha_{N,\beta,h}(\xi, x)$ easily yields (4.13), which proves the lemma. ▮

Let us now consider the measure $\tilde{\nu}^\alpha_{N,\beta,h}(\xi)$. Since the mean Hamiltonian (4.12) can be expressed in terms of the overlap parameters as

$$\mathbb{E}_\varepsilon H^\alpha_{N,h}(\varepsilon; \xi; \sigma) = -N\left\{\sum_{\mu=1}^{m} [m^\mu_N(\sigma; \xi)]^2 - hm^\alpha_N(\sigma; \xi)\right\} + mN \tag{4.17}$$

we have by (4.8) that, for any given $x \in \Xi_\xi$, the right-hand side of (4.15) takes the same value for all the configurations σ such that $X_\xi(\sigma) = x$. Therefore the density $\tilde{\nu}^\alpha_{N,\beta,h}(\xi; x)$ can be written as

$$\tilde{\nu}^\alpha_{N,\beta,h}(\xi, x) = \frac{\exp\{-NF^\alpha_{N,\beta,h,M}(x)\}}{\sum_{x \in \Xi_\xi} \exp\{-NF^\alpha_{N,\beta,h,M}(x)\}} \tag{4.18}$$

for all $x \in \Xi_\xi$, where

$$F^\alpha_{N,\beta,h,M}(x) = -\beta\left\{\sum_{\mu=1}^{m}\left[\frac{1}{d}(e^\mu, [M + \mathbb{1}]x)\right]^2 + h\frac{1}{d}(e^\alpha, [M + \mathbb{1}]x)\right\}$$
$$+ \ln|\{\sigma \in \mathscr{S}^N: X_\xi(\sigma) = x\}| \tag{4.19}$$

(4.18) has now a convenient form in that, roughly speaking, the point at which (4.19) takes its minimum value can be computed exactly and $\tilde{\nu}^\alpha_{N,\beta,h}(\xi; x)$ can be shown to be concentrated at that point. Wer collect the result we will need in the following lemma.

Lemma 4.3. [10] Let $a(\beta, h)$ be the largest solution of the equation $x = \tanh(\beta[x+h])$ and denote by $\tilde{x}^{\alpha}(h, \beta) \in [-1, 1]^d$ the vector

$$\tilde{x}^{\alpha}(h, \beta) = a(\beta, h) e^{\alpha} \tag{4.20}$$

Let $\varrho(h, N)$ and $\tilde{\varrho}(h, N)$ be two arbitrarily chosen functions that tend to zero as N tends to infinity. Define the subset $A \in \Xi_{\xi}$ as

$$A = \{x \in \Xi_{\xi} \mid \|x - \tilde{x}^{\alpha}(h, \beta)\| \leqslant \sqrt{d}\varrho(h, N)\} \tag{4.21}$$

Then, for all $\omega \in \mathscr{D}_N$ and for N large enough

$$\begin{aligned} \sum_{x \in A^c} \tilde{v}^{\alpha}_{N,\beta,h}(\xi, x) & \\ \leqslant \exp[N\{3\delta(N) - a\tilde{\varrho}^2(h, N)\}] & \\ + \exp\left[N\left\{3\delta(N) + b\beta h\tilde{\varrho}(h, N) - \frac{b\beta h}{2}(\varrho(h, N) - \tilde{\varrho}(h, N))^2\right\}\right] & \end{aligned} \tag{4.22}$$

where a and b are positive constants.

We are now ready to complete the proof of Theorem 2.

Proof of Theorem 2. Subtracting from both sides of (4.10) the term $g((1/d)(e^{\mu}, \tilde{x}^{\alpha}(h, \beta)))$, we get

$$\begin{aligned} &\left| \sum_{\sigma \in \mathscr{S}^N} \left[g(m_N^{\mu}(\sigma; \xi)) - g\left(\frac{1}{d}(e^{\mu}, \tilde{x}^{\alpha}(h, \beta))\right)\right] \mathscr{G}^{\alpha}_{N,\beta,h}(\varepsilon; \xi; \sigma)\right| \\ &\quad = \left| \sum_{x \in \Xi_{\xi}} \left[g\left(\frac{1}{d}(e^{\mu}, [M + \mathbb{1}]x)\right) - g\left(\frac{1}{d}(e^{\mu}, \tilde{x}^{\alpha}(h, \beta))\right)\right] v^{\alpha}_{N,\beta,h}(\varepsilon; \xi; x)\right| \end{aligned} \tag{4.23}$$

and decomposing the sum over $x \in \Xi_{\xi}$ as the sums over $x \in A$ and $x \in A^c$,

$$\begin{aligned} &\left| \sum_{x \in \Xi_t} \left[g\left(\frac{1}{d}(e^{\mu}, [M + \mathbb{1}]x\right) - g\left(\frac{1}{d}(e^{\mu}, \tilde{x}^{\alpha}(h, \beta))\right)\right] v^{\alpha}_{N,\beta,h}(\varepsilon; \xi; x)\right| \\ &\quad \leqslant \sum_{x \in A} \left| g\left(\frac{1}{d}(e^{\mu}, [M + \mathbb{1}]x)\right) - g\left(\frac{1}{d}(e^{\mu}, \tilde{x}^{\alpha}(h, \beta))\right)\right| v^{\alpha}_{N,\beta,h}(\varepsilon; \xi; x) \\ &\qquad + \sum_{x \in A^c} \left| g\left(\frac{1}{d}(e^{\mu}, [M + \mathbb{1}]x)\right) - g\left(\frac{1}{d}(e^{\mu}, \tilde{x}^{\alpha}(h, \beta))\right)\right| v^{\alpha}_{N,\beta,h}(\varepsilon; \xi; x) \end{aligned} \tag{4.24}$$

so that we are left to show that each of the two terms in the right-hand side of (4.24) goes to zero as N tends to infinity. To bound the former, note that for all $x \in A$ and $\omega \in \mathscr{D}_N$

$$\left|\frac{1}{d}(e^\mu, [M+\mathbb{1}]x) + \frac{1}{d}(e^\mu, \tilde{x}^\alpha(h, \beta))\right| \leqslant \delta(N) + \varrho(h, N) \tag{4.25}$$

since on $\mathscr{D}_N$, $|(e^\mu, Mx)| \leqslant d\delta(N)$, and by definition of A, $|(e^\mu, [x - \tilde{x}^\alpha(h, \beta)])| \leqslant d\varrho(h, N)$. Therefore, by continuity of g,

$$\left| g\left(\frac{1}{d}(e^\mu, [M+\mathbb{1}]x)\right) - g\left(\frac{1}{d}(e^\mu, \tilde{x}^\alpha(h, \beta))\right)7 \leqslant \zeta \tag{4.26}$$

for any arbitrarily small ζ, provided that N is sufficiently large and finally, for all $\omega \in \mathscr{D}_N$,

$$\sum_{x \in A} \left| g\left(\frac{1}{d}(e^\mu, [M+Iq]x)\right) - g\left(\frac{1}{d}(e^\mu, \tilde{x}^\alpha(h, \beta))\right)\right| v^\alpha_{N,\beta,h}(\varepsilon; \xi; x) \leqslant \zeta \tag{4.27}$$

To treat the second term in the right-hand side of (4.24) we use that since g is bounded,

$$\left| g\left(\frac{1}{d}(e^\mu, [M+Id]x)\right) - g\left(\frac{1}{d}(e^\mu, \tilde{x}_\alpha(h, \beta))\right)\right| \leqslant 2\,\|g\|_\infty$$

where $\|\cdot\|_\infty$ denotes the norm of the supremum. Thus

$$\sum_{x \in A^c} \left| g\left(\frac{1}{d}(e^\mu, [M+Id]x)\right) - g\left(\frac{1}{d}(e^\mu, \tilde{x}^\alpha(h, \beta))\right)\right| v^\alpha_{N,\beta,h}(\varepsilon; \xi; x)$$
$$\leqslant 2\,\|g\|_\infty \sum_{x \in A^c} v^\alpha_{N,\beta,h}(\varepsilon; \xi; x) \tag{4.28}$$

Inserting successively the bounds (4.13) and (4.20) of Lemmas 4.2 and 4.3, we have that, on $\mathscr{C}_N \cap \mathscr{D}_N$,

$$\sum_{x \in A^c} \tilde{v}^\alpha_{N,\beta,h}(\varepsilon; \xi; x)$$
$$\leqslant \exp[2\beta N\sqrt{m}\gamma(N)] \Big(\exp[N\{3\delta(N) - a\tilde{\varrho}^2(h, N)\}]$$
$$+ \exp\left[N\left\{3\delta(N) + b\beta h\tilde{\varrho}(h, N) - \frac{b\beta h}{2}(\varrho(h, N) - \tilde{\varrho}(h, N))^2\right\}\right]\Big) \tag{4.29}$$

and this last bound converges to zero as N tends to infinity provided that γ is chosen such that $\gamma(N)[m(N)]^{1/2} \downarrow 0$ as $N \uparrow \infty$ and that $\varrho(h, N)$ and $\tilde{\varrho}(h, N)$ are chosen such that

$$2\beta\sqrt{m}\gamma(N) - 3\delta(N) - a\tilde{\varrho}^2(h, N) < 0 \tag{4.30}$$

and

$$2\beta\sqrt{m}\gamma(N) + 3\delta(N) + b\beta h\tilde{\varrho}(h, N) - \frac{b\beta h}{2}[\rho(h, N) - \tilde{\varrho}(h, N)]^2 < 0 \tag{4.31}$$

which is possible for any a and b. Note that putting together the above constraint on γ and those of Theorem 3 imposes the condition $m(N)/p(N)N \downarrow 0$ as $N \uparrow \infty$. Now setting $\mathscr{E} \equiv \bigcup_{N_0 > 0} \bigcap_{N > N_0} \{\mathscr{C}_N \cap \mathscr{D}_N\}$, (4.5), (4.14), and the Borel–Cantelli Lemma imply $\mathbb{P}(\mathscr{E}) = 1$. Thus (4.23) and (4.24) together with the previous bounds give

$$\lim_{N \uparrow \infty} \left| \sum_{\sigma \in \mathscr{S}^N} \left[g(m_N^\mu(\sigma; \xi)) - g\left(\frac{1}{d}(e^\mu, \tilde{x}^\alpha(h, \beta))\right)\right] \mathscr{G}^\alpha_{N,\beta,h}(\varepsilon; \xi; \sigma)\right| = 0,$$
$$\mathbb{P}_\xi \times \mathbb{P}_\varepsilon\text{-a.s.} \tag{4.32}$$

Finally, using (4.2)

$$\frac{1}{d}(e^\mu, \tilde{x}^\alpha(h, \beta)) = \frac{a^+(\beta)}{d}(e^\mu, e^\alpha) = a^+(\beta)\delta_{\mu,\nu} \tag{4.33}$$

uniformly in N, and since

$$\lim_{h \downarrow 0} a^+(\beta)\delta_{\mu,\nu} = \begin{cases} 0 \\ \tilde{m}^+_{\mu,\nu}(\beta) \end{cases} \tag{4.34}$$

the case where $h \geqslant 0$ of Theorem 2 is proven. ∎

APPENDIX

In this appendix we give a proof of Lemma 3.4. Since this result may have some more general interest, we present it under more general conditions on the ξ_i^μ. Our proof is largely inspired by an estimate on largest eigenvalues of random matrices given by Komlos and Füredi.[13]

We will assume the following about the ξ_i^μ here:

(i) $\{\xi_i^\mu\}_{i=1,\dots,N}^{\mu=1,\dots,m}$ is a family of i.i.d. r.v.'s.

(ii) $\mathbb{E}\xi_i^\mu = 0$.

(iii) $\mathbb{E}(\xi_i^\mu)^l \leqslant \sigma^l$ for all $l \geqslant 2$.

We shall also, without loss of generality, assume that $m \leqslant N$. Let us define the $(N+m)\times(N+m)$ matrix B with elements

$$B_{\alpha\beta}=\begin{cases}\xi^{\beta}_{\alpha-m} & \text{if } \alpha>m \text{ and } \beta\geqslant m\\ \xi^{\beta-m}_{\alpha} & \text{if } \beta>m \text{ and } \alpha\leqslant m\\ 0 & \text{else}\end{cases} \tag{A.1}$$

Notice that then

$$(B^2)_{\alpha\beta}=\begin{cases}\sum_{i=1}^{N}\xi^{\alpha}_{i}\xi^{\beta}_{i} & \text{if } \alpha\leqslant m \text{ and } \beta\leqslant m\\ \sum_{\mu=1}^{m}\xi^{\mu}_{\alpha}\xi^{\mu}_{\beta} & \text{if } \alpha>m \text{ and } \beta>m\\ 0 & \text{else}\end{cases} \tag{A.2}$$

Clearly B^2 is the direct sum of two matrices B_1 and B_2, and the matrix A we are interested in is just $A=(1/N)B_1$. Let us introduce the two index sets $I_1=\{1,..., m\}$ and $I_2=\{m+1,..., m+N\}$. Clearly then we may write

$$\mathbb{E}\operatorname{tr} B_1^k=\sum_{\alpha_0,...,\alpha_{k-1}\in I_1}\;\sum_{\beta_0,...,\beta_{k-1}\in I_2}\mathbb{E}(B_{\alpha_0\beta_0}B_{\beta_0\alpha_1}\cdots B_{\alpha_{k-1}\beta_{k-1}}B_{\beta_{k-1}\alpha_0}) \tag{A.3}$$

We may think of the two sums as sums over sequences $(\alpha_0,..., \alpha_{k-1})\in I_1^k$, etc. For such a sequence we will denote by

$$\{(\alpha_0,...,\alpha_{k-1})\}\equiv\{t\in I_1 \mid \exists_{0\leqslant l<k} \text{ s.t. } \alpha_l=t\} \tag{A.4}$$

the set of different values the sequence runs through. We may then arrange the sums in (A.3) in such a way as to first sum over all possible subsets $\Gamma_1\subset I_1$ and $\Gamma_2\subset I_2$ and then over all sequences for which the values run through exactly these subsets. Thus

$$\mathbb{E}\operatorname{tr} B_1^k=\sum_{\Gamma_1\subset I_1}\sum_{\Gamma_2\subset I_2}\sum_{\substack{(\alpha_0,...,\alpha_{k-1})\in I_1^k\\ \{(\alpha_0,...,a_{k-1})\}=\Gamma_1}}\;\sum_{\substack{(\beta_0,...,\beta_{k-1})\in I_2^k\\ \{(\beta_0,...,\beta_{k-1})\}=\Gamma_2}}$$

$$\times\,\mathbb{E}(B_{\alpha_0\beta_0}B_{\beta_0\alpha_1}\cdots B_{\alpha_{k-1}\beta_{k-1}}B_{\beta_{k-1}\alpha_0}) \tag{A.5}$$

Now it is obvious that the sums over the sequences in (A.5) do not depend on the exact sets Γ_1 and Γ_2, but only on the cardinalities of these two sets. Thus we may write

$$\mathbb{E}\operatorname{tr} B_1^k=\sum_{r=1}^{\min(k,m)}\;\sum_{s=1}^{\min(k,N)}\binom{m}{r}\binom{N}{s}E_{k,r,s} \tag{A.6}$$

where

$$E_{k,r,s} \equiv \sum_{\substack{(\alpha_0,\ldots,\alpha_{k-1}) \in I_1^k \\ \{(\alpha_0,\ldots,a_{k-1})\} = \{1,\ldots,r\}}} \sum_{\substack{(\beta_0,\ldots,\beta_{k-1}) \in I_2^k \\ \{(\beta_0,\ldots,\beta_{k-1})\} = \{m+1,\ldots,m+s\}}} \times \mathbb{E}(B_{\alpha_0\beta_0} B_{\beta_0\alpha_1} \cdots B_{\alpha_{k-1}\beta_{k-1}} B_{\beta_{k-1}\alpha_0}) \tag{A.7}$$

and where we have used that the combinatorial factors in (A.5) count the number of subsets of given cardinality. Note that $E_{k,r,s}$ no longer depends on m or N [the appearance of m in (A.7) being completely spurious].

To estimate these last quantities, we would like to think of the sums in (A.7) in a slightly different way. Let us denote by $\mathscr{G}_{r,s}$ the complete bipartite graph with vertex sets labeled by $\mathscr{R} \equiv \{1,\ldots, r\}$ and $\mathscr{S} \equiv \{m+1,\ldots, m+s\}$, i.e., the graph with vertex set $\mathscr{R} \cup \mathscr{S}$ and edge set $\mathscr{R} \times \mathscr{S}$. Each term in the sum (A.7) corresponds to a walk of length $2k$, ω, on this bipartite graph (i.e., a sequence of edges linking alternately the sets $\mathscr{R}$ and $\mathscr{S}$) with the property that each vertex of $\mathscr{G}_{r,s}$ is visited at least once. Moreover, it is clear that any walk which passes over any given edge of $\mathscr{G}_{r,s}$ exactly once will give a zero contribution, as the expectation of the corresponding product of the $B_{\alpha\beta}$ vanishes by assumption (ii) on the distribution of the ξ. We will denote by $\mathscr{W}_k(r, s)$ the set of walks that give a nonzero contribution. By our assumptions, we then have the following result.

Lemma A.1.

$$E_{k,r,s} \leqslant \sigma^{2|r+s-1|} |\mathscr{W}_k(r, s)| \tag{A.8}$$

We are thus left to estimate the number of walks in $\mathscr{W}_k(r, s)$.

Notice first that for fixed r and s, the shortest possible walk contributing must have length $2k = 2(r+s-1)$. Let us thus first consider the case $k = r+s-1$. In this case, the walk must visit each edge either zero or two times. Moreover, the edges it does visit form a bipartite tree on $(\mathscr{R}, \mathscr{S})$. We classify all such walks according to the different trees they generate, count the number of walks for a given tree, and then enumerate all bipartite trees. We get the following result.

Lemma A.2. Let t be a bipartite tree on $(\mathscr{R}, \mathscr{S})$ with coordination numbers $d_1,\ldots, d_r$, $c_{m+1},\ldots,c_{m+s}$. Let $\Omega(t)$ denote the set of all walks in $\mathscr{W}_{r+s-1}(r, s)$ that generate t. Then

$$|\Omega(t)| = (r+s-1) \prod_{i=1}^{r} (d_i - 1)! \prod_{j=m+1}^{m+s} (c_j - 1)! \tag{A.9}$$

Proof. Let us pick a particular vertex i, say in $\mathscr{R}$. Suppose ω arrives at i at time n_0 for the first time. There are $d_i - 1$ branches emerging from i (other than the one the walk just comes from), and the walk must pass completely over them before it is allowed to return. So at the next step, there are $(d_i - 1)$ choices for the walk to continue. Given that choice, the walk will return to i at some later time n_1 after having passed exactly over the entire chosen branch. Now there remain $(d_i - 2)$ choices to continue and so on, until after the $(d_i - 1)$th visit of the vertex i it leaves it in the direction it first came from, never to return to it. Clearly, the total number of choices arising from the visits at this vertex amounted to $(d_i - 1)!$, and obviously each vertex contributes such a factor, whence the two products of factorials in (A.9). Finally, it remains to decide on the starting edge for the walk, of which there are $(r + s - 1)$, which accounts for the first factor in (A.9). ∎

By this result, we only have to know the number of bipartite trees with given coordination numbers. However, this is a standard problem of graph theory and one has the following generalization of Cayley's formula:

Lemma A.3. Let $T(r, s; d_1, \dots, d_r; c_{m+1}, \dots, c_{m+s})$ denote the number of bipartite trees with given coordination numbers d_i, c_j. Then, if $\sum_i d_i = \sum_j c_j = r + s - 1$,

$$T(r, s; d_1, \dots, d_r; c_{m+1}, \dots, c_{m+s}) = \frac{(r-1)!\,(s-1)!}{(d_1 - 1)! \cdots (d_r - 1)!\,(c_{m+1} - 1)! \cdots (c_{m+s} - 1)!} \tag{A.10}$$

and zero otherwise.

(The proof of this formula is by induction as in the standard version of Cayley's formula. See, e.g., ref. 2.)

Combining these results, we get the following.

Lemma A.4. Let $k = r + s - 1$. Then

$$|\mathscr{W}_k(r, s)| = (r + s - 1)\binom{r + s - 2}{r - 1}^2 (r-1)!\,(s-1)! \tag{A.11}$$

The proof of this formula is straightforward.

Let us now return to the general case, $k \geqslant r + s - 1$. Using the previous results, it is fairly easy to prove the following rather crude bound:

Lemma A.5.

$$|\mathscr{W}_k(r,s)| \leqslant \binom{2k}{2(r+s-1)} (sr)^{2(k-r-s+1)}(r+s-1) \times \binom{r+s-2}{r-1}^2 (r-1)!\,(s-1)! \tag{A.12}$$

Proof. To get (A.12) just note the following: First, for an arbitrary walk ω, it is still possible to construct in a unique way a bipartite tree $t(\omega)$ on $(\mathscr{R}, \mathscr{S})$. To do this, just follow the walk and include into t successively all edges that lead to a vertex not previously visited by the walk. Then we may construct a new walk $\tilde{\omega}(\omega)$ of length $2(r+s-1)$ whose associated tree is also $t(\omega)$ by again following ω and including an edge into $\tilde{\omega}$ if it is an edge from t and is visited the first or the second time. Moreover, we give it the orientation $+$ if it is visited the first time and $-$ if it is visited the second time. It is easy to verify that this creates the desired walk. Now we know how many walks $\tilde{\omega}$ exist; thus we need only to estimate the number of walks ω giving rise to the same $\tilde{\omega}$. To do this, just squeeze $2(k-r-s+1)$ edges between those of $\tilde{\omega}$. There are $\binom{2k}{2(r+s-1)}$ ways of distributing them, and there are no more then sr ways of placing each edge (in fact there are much fewer). But this gives the estimate in Lemmma A.5. ∎

Let us define the quantities

$$S_{N,m,k,r,s} \equiv \sigma^{2(r+s-1)} \binom{N}{s}\binom{m}{r}\binom{2k}{2(r+s-1)} (sr)^{2(k-r-s+1)}(r+s-1) \times \binom{r+s-2}{r-1}^2 (r-1)!\,(s-1)! \tag{A.13}$$

We clearly have that

$$E \operatorname{tr} B_1^k \leqslant \sum_r \sum_s S_{N,m,k,r,s} = \sum_{q=1}^{k+1} \sum_{r=1}^{q-1} S_{N,m,k,r,q-1} \tag{A.14}$$

Now a simple calculation shows that

$$S_{N,m,k,r,q-1-r} \leqslant \frac{5k^6}{12\sigma^2 N(1-k/N)} S_{N,m,k,r,q-r} \tag{A.15}$$

and therefore, if $k \leqslant N^{1/6}\sigma^{1/3}$,

$$S_{N,m,k,r,q-1-r} \leqslant \tfrac{1}{2} S_{N,m,k,r,q-r} \tag{A.16}$$

Thus

$$\sum_{q=1}^{k+1} \sum_{r=1}^{q-1} S_{N,m,k,r,q-r} \leqslant 2 \sum_{r=1}^{k} S_{N,m,k,r,k-r} \tag{A.17}$$

and finally we arrive at the following results.

Lemma A.6. For $k \leqslant N^{1/6}$,

$$E \operatorname{tr} B_1^k \leqslant k \max_r S_{N,m,k,r,k-r} \tag{A.18}$$

What we are left with finally is to determine the maximum in (A.18). For N large, and using that $k \ll N$, we find that the minimum is realized for $r \approx m\gamma$, where $\gamma \equiv \rho/(1+\rho)$ [remember that $\rho = (m/N)^{1/2}$]. Inserting this value, a simple calculation show that the right-hand side of (A.18) is equal to $N^{k+1}e^{2k\gamma}$, up to an irrelevant correction factor that goes to one as $N \uparrow \infty$. But from this, Lemma 3.4 is obvious. ▮

ACKNOWLEDGMENTS

V.G. would like to thank Joel Lebowitz and the Mathematics Department of Rutgers University, where part of this work was done, for their kind hospitality. We are grateful to Hans Koch for sending a copy of his work prior to publication. Finally, we would like to thank Pierre Picco for a critical reading of the manuscript and valuable comments.

Work partially supported by the Commission of the European Communities under contract No. SC1-CT91-0695.

NOTE ADDED IN PROOF

The results allowing to remove the condition $m < (\ln N)/(\ln 2)$ in Theorem 2 have meanwhile been obtained by the authors and P. Picco.[20]

REFERENCES

1. D. J. Amit, H. Gutfreund, and H. Sompolinsky, Spin-glass models of neural networks, *Phys. Rev. A* **32**:1007–1018 (1985); Storing infinite numbers of patterns in a spin glass model of neural networks, *Phys. Rev. Lett.* **55**:1530–1533 (1985).
2. C. Bergé, *Graphs and Hypergraphs* (North-Holland, Amsterdam, 1973).

3. A. Bovier and V. Gayrard, Rigorous bounds on the storage capacity of the dilute Hopfield model, *J. Stat. Phys.* **69**:597–627 (1992).
4. A. Bovier and V. Gayrard, The thermodynamics of the Curie–Weiss model with random couplings, preprint IAAS-Berlin (1992), *J. Stat. Phys.* **72** (1993), to appear.
5. B. Bollobás, *Random Graphs* (Academic Press, New York, 1985).
6. Y. S. Chow and H. Teicher, *Probability Theory* (Springer, New York, 1978).
7. B. Derrida, E. Gardner, and A. Zippelius, *Europhys. Lett.* **4**:167 (1987).
8. E. Domany, J. L. van Hemmen, and K. Schulte, *Models of Neural Networks* (Springer, 1990).
9. R. S. Ellis, *Entropy, Large, Deviations, and Statistical Mechanics* (Springer-Verlag, Berlin, 1985).
10. V. Gayrard, The thermodynamic limit of the q-state Potts–Hopfield model with infinitely many patterns, *J. Stat. Phys.* **68**:977–1011 (1992).
11. D. Grensing and R. Kühn, Random site spin glass models, *J. Phys. A* **19**:L1153 (1986).
12. J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. USA* **79**:2554–2558 (1982).
13. Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices, *Combinatorica* **1**:233–241 (1981).
14. H. Koch and J. Piasko, Some rigorous results on the Hopfield neural network model, *J. Stat. Phys.* **55**:903 (1989).
15. H. Koch, A free energy bound for the Hopfield model, *J. Phys. A* **26**:L353 (1993).
16. M. Mézard, G. Parisi, and M. A. Virasoro, *Spin-Glass Theory and Beyond* (World Scientific, Singapore, 1988).
17. Ch. M. Newman, Memory capacity in neural network models: Rigorous lower bounds, *Neural Networks* **1**:233–238 (1988).
18. D. Sherrington and S. Kirkpatrick, *Phys. Rev. Lett.* **35**:1792 (1972).
19. M. Shcherbina and B. Tirozzi, Exact mean field theory for some Hopfield model, preprint (1992).
20. A. Bovier, V. Gayrard, and P. Picco, Gibbs States of the Hopfield model in the regime of perfect memory, preprint CPT-93/P.2894 (Marseille), submitted to *Prob. Theor. Rel. Fields* (1993).