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

The Free Energy of a Class of Hopfield Models

M. Shcherbina[1] **and B. Tirozzi**[2]

Received December 9, 1991; final January 29, 1993

We show that in the limit $p \to \infty$, $N \to \infty$, $\alpha = p/N \to 0$ the limit free energy of the Hopfield model equals in probability the Curie–Weiss free energy. We prove also that the free energy of the Hopfield model is self-averaging for any finite α.

KEY WORDS: Free energy; overlaps; self-averaging.

1. INTRODUCTION

Consider the model of the form (the Hopfield model[1-4])

$$H_N = -\frac{1}{2} \sum_{i \neq j}^{N} J_{ij} S_i S_j \tag{1.1}$$

where

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

and $S_1, \ldots, S_N$ are Ising spins $S_i = \pm 1$ and ξ_i^μ are independent random variables. This model, which was proposed as a model of associative memory in the neural network theory,[7,12] has many features of the spin-glass model with very interesting properties (see ref. 5 and references therein). The most interesting case is when the number of random vectors is proportional to N:

$$\frac{p}{N} \to \alpha > 0 \qquad (N \to \infty) \tag{1.2}$$

[1] Institute of Low Temperature Physics and Engineering, Kharkov, 310164, Ukraine.
[2] Department of Mathematics, University of Rome "La Sapienza," 00185 Rome, Italy.

0022-4715/93/0700-0113$07.00/0

in this case the model has a spin-glass phase which disturbs the retrieval property.[12] While many physics papers are devoted to the analysis of the Hopfield model in the case (1.2) (see, e.g., ref. 5), unfortunately, few of them are rigorous mathematically. The main part of the rigorous results is concerned with the Hopfield model in the case

$$\frac{p}{N} \to 0 \qquad (N \to \infty) \tag{1.3}$$

The case of p independent of N was studied in refs. 6–9. The case $p \sim \ln N$ was considered in ref. 10.

All these versions of the Hopfield model have, in the thermodynamic limit, a free energy which coincides in probability with that of the Curie–Weiss model [i.e., the model (1.1) with $p=1$].

In the present paper we consider the free energy of the model (1.1) in the limit with the constraint (1.3). We hope that this will be the first step to obtaining an expansion in the parameter α for the free energy of the model (1.1). In Section 2 we give the main definitions and results, which are proven in Section 3; in Section 4 we give some auxiliary lemmas.

2. DEFINITIONS AND RESULTS

We will use other Hamiltonians besides the one defined in (1.1):

$$H(\gamma) = H_N - N^{-1/2}\varepsilon \sum_{\mu=1}^{p} \gamma^\mu \sum_{i=1}^{N} \xi_i^\mu S_i \tag{2.1}$$

$$\begin{aligned} H^a(\gamma, \mathbf{c}) &= H(\gamma) + \frac{N}{2} \sum_{\mu=1}^{p} \left(\frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu S_i - c^\mu \right)^2 \\ &= -\sum_{\mu=1}^{p} \left(c^\mu + \frac{\varepsilon \gamma^\mu}{\sqrt{N}} \right) \sum_{i=1}^{N} \xi_i^\mu S_i + \frac{N}{2} \sum_{\mu=1}^{p} (c^\mu)^2 + \frac{p}{2} \end{aligned} \tag{2.2}$$

where $\gamma = (\gamma^1, ..., \gamma^p)$, $\mathbf{c} = (c^1, ..., c^p)$, and γ^μ are independent Gaussian random variables with zero mean and variance 1. The c^μ are some real numbers and ε is a given parameter. E will denote the expectation with respect to the random variables (γ, ξ_i^μ, $i=1,..., N$, $\mu=1,..., p$) and we will use the symbols $\langle \cdot \rangle$ and $\langle \cdot \rangle_{H(\gamma)}$ for the expectations with respect to the Gibbs distribution of the N-spin system generated by the Hamiltonians H_N and $H(\gamma)$, respectively.

Let us introduce also the free energy of the N-spin system:

$$f_N(H_N) = -\frac{1}{\beta N} \ln \sum_{S_1, ..., S_N} e^{-\beta H_N} \tag{2.3}$$

and the real function

$$f^* = -\frac{1}{\beta} \ln 2 \cosh \beta c + \frac{c^2}{2} \tag{2.4}$$

We will consider also the overlap parameters, which are the typical order parameters of the neural networks:

$$m^\mu = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \langle S_i \rangle, \qquad \mu = 1,..., p \tag{2.5}$$

We will say that the function $\phi_N(\xi)$, $\xi = \xi_i^\mu$, $i = 1,..., N$, $\mu = 1,..., p$, has the self-averaging property (s.a.) if

$$\lim_{N \to \infty} E(\phi_N(\xi) - E(\phi_N(\xi))^2 = 0 \tag{2.6}$$

The first result is as follows.

Theorem 1. $f_N(H_N)$ is s.a.

The s.a. property of the free energy of the spin glass has been shown in ref. 11 and we will prove it in the case of the Hopfield model using the ideas of refs. 11 and 13. The main result is the following theorem.

Theorem 2. The free energy of the model (1.1) converges in probability to the free energy of the Curie–Weiss theory:

$$\lim_{N \to \infty} f_N(H) = \min_c f^*(c) \tag{2.7}$$

under the condition (1.3).

3. PROOFS

Proof of Theorem 1. Let $\mathscr{F}_k$ be the Σ-algebras generated by the sets $(\xi \mid \xi_i^\mu)_{i \geqslant k}^{\mu=1,..., p}$ fixed),[11,16,17] and let F_N^k be the conditional expectation of F_N with respect to $\mathscr{F}_k$:

$$F_N^k = E\left(-\frac{1}{\beta} \ln Z_N \,\middle|\, \mathscr{F}_k \right) \equiv E_{<k}\left(-\frac{1}{\beta} \ln Z_N \right) \tag{3.1}$$

where $E_{<k}$ is defined by

$$E_{<k} = \sum_{(\xi_i^\mu)_{i \leqslant k}^{\mu=1,..., p}} \prod_{\mu, i} P(\xi_i^\mu) \tag{3.2}$$

From the definition we have

$$F_N^1 = -\frac{1}{\beta} \ln Z_N, \qquad F_N^{N+1} = E\left(-\frac{1}{\beta} \ln Z_N\right) \tag{3.3}$$

It is also evident that

$$E(F_N^k \mid \mathscr{F}_l) = \begin{cases} F_N^k & \text{if } k > l \\ F_N^l & \text{if } k \leqslant l \end{cases} \tag{3.4}$$

We define also

$$\Psi^k = F_N^k - F_N^{k+1}$$

Then

$$f_N - E(f_N) = \frac{1}{N} \sum_{k=1}^{N} \Psi_k$$

and

$$E(f_N - E(f_N))^2 = \frac{1}{N^2} \sum_{k=1}^{N} E\Psi_k^2 + \frac{2}{N} \sum_{k<l}^{N} E\Psi_k \Psi_l \tag{3.5}$$

Using the properties of the conditional expectation, we have ($k < l$)

$$E\Psi_k \Psi_l = E(E(\Psi_k \Psi_l \mid \mathscr{F}_l)) = E(\Psi_l E(\Psi_k \mid \mathscr{F}_l))$$

But

$$E(\Psi_k \mid \mathscr{F}_l) = E(F_N^k - F_N^{k+1} \mid \mathscr{F}_l) = F_N^l - F_N^l = 0$$

Thus, in order to obtain the result, it is enough to show that

$$E\Psi_k^2 < C$$

where C is a constant not depending on k.

Let us define the following useful Hamiltonians:

$$H_k = -\frac{1}{2N} \sum_{\mu} \sum_{i,j,i \neq j; i,j \neq k} \xi_i^\mu \xi_j^\mu S_i S_j \tag{3.6}$$

$$R_k = -\frac{1}{N} \sum_{\mu} \sum_{l \neq k} \xi_k^\mu \xi_l^\mu S_k S_l \tag{3.7}$$

$$\tilde{H}_k(t) = H_k + t R_k$$

$$\tilde{f}_k(t) = -\frac{1}{\beta} \{\ln Z_N(\tilde{H}_k(t)) - \ln Z_N(\tilde{H}_k(0))\}$$

$$\tilde{H}_k(1) = H$$

We have

$$\Psi_k = E_{<k}\tilde{f}_k(1) - E_{<k+1}\tilde{f}_k(1)$$
$$E\{\Psi_k^2\} \leqslant 2E\{(\tilde{f}_k(1))^2\}$$

But since $\tilde{f}_k(0)=0$ and

$$\frac{d^2}{dt^2}\tilde{f}_k(t) \leqslant 0$$

then

$$-\left\langle \frac{1}{N}\sum_{l\neq k}\sum_{\mu} \xi_k^\mu \xi_l^\mu S_k S_l \right\rangle_H = \tilde{f}_k(1)' \leqslant \tilde{f}_k(1) \leqslant \tilde{f}_k(1)'$$
$$= -\left\langle \frac{1}{N}\sum_{l\neq k}\sum_{\mu} \xi_k^\mu \xi_l^\mu S_k S_l \right\rangle_{\tilde{H}(0)}$$

Therefore

$$E\{(\tilde{f}_k(1))^2\} \leqslant E\{(\tilde{f}_k(1)')^2\} + E\{(\tilde{f}_k(0)')^2\}$$
$$\leqslant E\left\{\left\langle \sum_{l_1,l_2\neq k} J_{kl_1}J_{kl_2}S_{l_1}S_{l_2}\right\rangle_H\right\}$$
$$+\frac{1}{N^2}E\left\{\sum_{l_1,l_2\neq k}\sum_{\mu_1,\mu_2} \xi_k^{\mu_1}\xi_{l_1}^{\mu_1}\xi_k^{\mu_2}\xi_{l_2}^{\mu_2}\langle S_{l_1}S_{l_2}\rangle_{\tilde{H}(0)}\right\} \tag{3.8}$$

In order to estimate the first term on the r.h.s. of this formula let us use the fact that H is symmetrical with respect to all indexes and therefore

$$E\left\{\left\langle \sum_{l_1,l_2\neq k} J_{kl_1}J_{kl_2}S_{l_1}S_{l_2}\right\rangle_H\right\} = E\left\{\sum_{l_1,l_2,k}\frac{1}{N}J_{kl_1}J_{kl_2}\langle S_{l_1}S_{l_2}\rangle_H\right\}$$
$$\leqslant E\{\|J\|^2\} \leqslant 4 + O\left(\frac{1}{N}\right)$$

Here we use also Lemma 4.2. Since $\tilde{H}(0)$ is independent on ξ_k^μ the second term in the r.h.s. of (3.8) can be estimated as

$$\alpha^2 + \frac{1}{N^2}E\left\{\sum_{l_1,l_2}\sum_{\mu}\xi_{l_1}^\mu\xi_{l_2}^\mu\langle S_{l_1}S_{l_2}\rangle_{\tilde{H}(0)}\right\} \leqslant \alpha^2 + E\{\|J\|\}$$
$$\leqslant \alpha^2 + 2 + O\left(\frac{1}{N}\right)$$

Theorem 1 is proven.

Proof of Theorem 2. From the result of Theorem 1 it is sufficient to find the limit of $E(f_N)$. By using the Bogolubov inequality, we obtain, calling $f(H^a(\gamma, \mathbf{c}))$ and $f(H(\gamma))$ the free energy of the Hamiltonians (2.1) and (2.2), respectively,

$$0 \leqslant E(f(H^a(\gamma, \mathbf{c}))) - E(f(H(\gamma)))$$
$$\leqslant \frac{1}{2} E\left(\sum_{\mu=1}^{p} \left\langle \left(\frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu S_i - c^\mu \right)^2 \right\rangle_{H(\gamma)} \right)$$

and thus, choosing the c^μ equal to the overlaps

$$m^\mu = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \langle S_i \rangle_{H(\gamma)}$$

one gets

$$0 \leqslant E(\min_{\mathbf{c} \in R^p} f(H^a(\gamma, \mathbf{c}))) - E(f(H(\gamma)))$$
$$\leqslant E\left(\frac{1}{2N^2} \left\langle \sum_{\mu=1}^{p} \sum_{i,j=1}^{N} \xi_i^\mu \xi_j^\mu (S_i - \langle S_i \rangle_{H(\gamma)})(S_j - \langle S_j \rangle_{H(\gamma)}) \right\rangle_{H(\gamma)} \right) \tag{3.9}$$

On the other hand, from the formula of integration by parts it follows that

$$\frac{1}{N^{3/2}} \sum_{\mu=1}^{p} E\left(\gamma^\mu \sum_{i=1}^{N} \xi_i^\mu \langle S_i \rangle_{H(\gamma)} \right)$$
$$= \frac{\beta}{N^2} \sum_{\mu=1}^{p} E\left(\sum_{i,j=1}^{N} \xi_i^\mu \xi_j^\mu \langle (S_i - \langle S_i \rangle_{H(\gamma)})(S_j - \langle S_j \rangle_{H(\gamma)}) \rangle_{H(\gamma)} \right)$$

Inserting in (3.9) and using Schwartz's inequality, we have

$$0 \leqslant E(\min_{\mathbf{c} \in R^p} f(H^a(\gamma, \mathbf{c}))) - E(f(H(\gamma)))$$
$$\leqslant \frac{1}{2\beta} E\left(\frac{1}{N^{3/2}} \sum_{\mu=1}^{p} \gamma^\mu \sum_{i=1}^{N} \xi_i^\mu \langle S_i \rangle_{H(\gamma)} \right)$$
$$\leqslant \frac{1}{2\beta} \left[E \frac{1}{N} \sum_{\mu=1}^{p} (\gamma^\mu)^2 \right]^{1/2}$$
$$\times \left[E \frac{1}{2N^2} \sum_{\mu=1}^{p} \sum_{i,j=1}^{N} \xi_i^\mu \xi_j^\mu \langle S_i \rangle_{H(\gamma)} \langle S_j \rangle_{H(\gamma)} \right]^{1/2} \tag{3.10}$$

The last factor in (3.10) can be written as

$$\frac{1}{N}\sum_{i,j=1}^{N} J_{ij}\langle S_i\rangle_{H(\gamma)}\langle S_j\rangle_{H(\gamma)} \leqslant \|J\|$$

and can be bounded on the basis of Lemma 4.2.

Therefore on the basis of (3.10) we get

$$0 \leqslant E(\min_{\mathbf{c}\in R^p} f(H^a(\gamma, \mathbf{c}))) - E(f(H(\gamma))) \leqslant \text{const}(p/N)^{1/2}$$

Let us denote by $H^a(0, \mathbf{c})$ the Hamiltonian $H^a(\gamma, \mathbf{c})$ with $\varepsilon = 0$; then we have

$$\begin{aligned}
0 &\leqslant E(\min_{\mathbf{c}\in R^p} f(H^a(0, \mathbf{c}))) - E(f(H(\gamma))) \\
&\leqslant E(\min_{\mathbf{c}\in R^p} f(H^a(\gamma, \mathbf{c}))) - E(f(H(\gamma))) \\
&\quad + \left| E\left(\frac{1}{N^{3/2}}\sum_{\mu=1}^{p} \gamma^\mu \sum_{i=1}^{N} \xi_i^\mu \langle S_i\rangle_{H^a(0,\mathbf{c})}\right)\right| \\
&\quad + \left| E\left(\frac{1}{N^{3/2}}\sum_{\mu=1}^{p} \gamma^\mu \sum_{i=1}^{N} \xi_i^\mu \langle S_i\rangle_{H}\right)\right| \qquad (3.11)
\end{aligned}$$

Now it is easy to estimate the expressions

$$\left| E\left(\frac{1}{N^{3/2}}\sum_{\mu=1}^{p} \gamma^\mu \sum_{i=1}^{N} \xi_i^\mu \langle S_i\rangle_{H^a(0,\mathbf{c}),H}\right)\right|$$

in the right-hand side of (3.11) in the same way as in (3.10). Then we have proved the following fact:

$$\lim_{N\to\infty} E(f(H)) = \lim_{N\to\infty} E(\min_{\mathbf{c}\in R^p} f(H^a(0, \mathbf{c})))$$

It is easy to see that

$$\begin{aligned}
&\lim_{N\to\infty} E(\min_{\mathbf{c}\in R^p} f(H^a(0, \mathbf{c}))) \\
&\qquad \leqslant \lim_{N\to\infty} E(\min_{\mathbf{c}=(c^1,0,\ldots,0)} f(H^a(0, \mathbf{c}))) = \min_{c\in R} f^*(c) = f^*(c^*) \qquad (3.12)
\end{aligned}$$

where c^* is the point at which f^* gets its global minimum. Now we will show that

$$\lim_{N\to\infty} E(\min_{\mathbf{c}\in R^p} f(H^a(0, \mathbf{c}))) \geqslant \min_{c\in R} f^*(c) \qquad (3.13)$$

To this aim we use the method developed in ref. 8. Let us note first that, for finite N, $f(H^a(0, \mathbf{c})) \to \infty$ if $\sum_{\mu=1}^{p} (c^\mu)^2 \to \infty$ and that therefore the function $f(H^a(0, \mathbf{c}))$ takes its minimum value at some critical point which satisfies the equations

$$c^\mu = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \langle S_i \rangle_{H^a(0,\mathbf{c})} \tag{3.14}$$

Let C be the set of all such points. Then

$$\begin{aligned} &E(\min_{\mathbf{c} \in R^p} f(H^a(0, \mathbf{c}))) \\ &\quad = E\left(\min_{\mathbf{c} \in C} \left[-\frac{1}{\beta N} \sum_{i=1}^{N} \ln 2 \cosh \beta \left(\sum_{\mu=1}^{p} \xi_i^\mu c^\mu \right) + \frac{1}{2} \sum_{\mu=1}^{p} (c^\mu)^2 \right]\right) \\ &\quad = E\left(\min_{\mathbf{c} \in C} \left[\frac{1}{N} \sum_{i=1}^{N} f^*(h_i) - \frac{1}{2} \sum_{\mu \neq \nu}^{p} c^\mu c^\nu \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \xi_i^\nu \right]\right) \\ &\quad = E\left(\min_{\mathbf{c} \in C} \left[\frac{1}{N} \sum_{i=1}^{N} f^*(h_i) - \frac{1}{2} \sum_{\mu,\nu}^{p} \mathscr{A}_{\mu\nu} c^\mu c^\nu \right]\right) \\ &\quad \geqslant E\left(\min_{\mathbf{c} \in C} \left[\frac{1}{N} \sum_{i=1}^{N} f^*(h_i) - \|\mathscr{A}\| \sum_{\mu=1}^{p} (c^\mu)^2 \right]\right) \end{aligned} \tag{3.15}$$

where $h_i = \sum_{\mu=1}^{p} \xi_i^\mu c^\mu$ and the matrix $\mathscr{A}$ is defined in Lemma 4.1. But, for $\mathbf{c} \in C$,

$$\sum_{\mu=1}^{p} (c^\mu)^2 = \frac{1}{N} \sum_{i,j=1}^{N} J_{ij} \langle S_i \rangle_{H^a(0,\mathbf{c})} \langle S_j \rangle_{H^a(0,\mathbf{c})} \leqslant \|J\|$$

Therefore (3.15) gives

$$E(\min_{\mathbf{c} \in R^p} f(H^a(0, \mathbf{c}))) \geqslant \min_{h \in R} f^*(h) - E^{1/2}(\|J\|^2)\, E^{1/2}(\|\mathscr{A}\|^2)$$

and by the use of Lemmas 4.1 and 4.2 we obtain (3.13).

4. AUXILIARY RESULTS

Lemma 4.1. Let $\mathscr{A}_{\mu\nu}$ be defined by the formula

$$\mathscr{A}_{\mu\nu} = (1 - \delta_{\mu\nu}) \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \xi_i^\nu \qquad (\mu, \nu = 1,..., p)$$

and let $p/N \to \alpha \geqslant 0$ $(p, N \to \infty)$, then, for any $\varepsilon > 0$,

$$\mathrm{Prob}(\|\mathscr{A}\| > (\alpha + 2\sqrt{\alpha})(1+\varepsilon)) \leqslant \exp\left[-\frac{\sqrt{\alpha}}{4(\alpha+2\sqrt{\alpha})}\varepsilon^{4/3}(2p)^{2/3}\right]$$

$$E(\|\mathscr{A}\|^2) \leqslant 4\alpha + o(\alpha) \qquad (\alpha \to 0)$$

Proof. Let us define the sequence of numbers a_n by the relations

$$a_0 = 1$$

$$a_n = E\left(\frac{1}{p}\operatorname{Tr}\mathscr{A}^n\right) \tag{4.1}$$

$$= \frac{1}{pN^n} \sum_{\mu_1,\dots,\mu_n,i_1,\dots,i_n} E(\xi_{i_1}^{\mu_1}\xi_{i_1}^{\mu_2}\xi_{i_2}^{\mu_2}\xi_{i_2}^{\mu_3}\cdots\xi_{i_n}^{\mu_n}\xi_{i_n}^{\mu_1})$$

Since

$$E\xi_{i_1}^{\mu_1} = 0$$

we have nonzero terms in the last sum only if for some $k \geqslant 2$, $i_k = i_1$ and $\mu_k = \mu_1$ or $i_k = i_1$ and $\mu_{k+1} = \mu_1$. In addition, all the terms in this sums are positive and therefore if we take into account some of them more than once the sum will increase. So for $n \geqslant 3$

$$a_n \leqslant \sum_{k=0}^{n-2} \frac{1}{pN^2} \sum_{\mu_1 \neq \mu_2, \mu_1 \neq \mu_3} E(\xi_{i_1}^{\mu_1}\xi_{i_1}^{\mu_2}(\mathscr{A}^k)_{\mu_2\mu_1}\xi_{i_1}^{\mu_1}\xi_{i_1}^{\mu_3}(\mathscr{A}^{n-k-2})_{\mu_3\mu_1}$$

$$+\ \xi_{i_1}^{\mu_1}\xi_{i_1}^{\mu_2}(\mathscr{A}^k)_{\mu_2\mu_3}\xi_{i_1}^{\mu_3}\xi_{i_1}^{\mu_1}(\mathscr{A}^{n-k-2})_{\mu_1\mu_1})$$

$$\leqslant \sum_{k=0}^{n-2}\left[\frac{1}{N}(a_{n-2} + a_{n-1}) + E\left(\frac{1}{Np}\operatorname{Tr}(\mathscr{A}^k + \mathscr{A}^{k+1})\operatorname{Tr}\mathscr{A}^{n-k-2}\right)\right] \tag{4.2}$$

But since for l, $k \geqslant 1$,

$$E\left(\frac{1}{p^2}\operatorname{Tr}\mathscr{A}^k\operatorname{Tr}\mathscr{A}^l\right) - a_k a_l$$

$$\leqslant \frac{lk}{p^2N^2} \sum_{\mu_1 \neq \nu_2, \mu_2 \neq \mu_1, j_1} E(\xi_{j_1}^{\mu_1}\xi_{j_1}^{\mu_2}(\mathscr{A}^{k-1})_{\mu_2\mu_1}\xi_{j_1}^{\mu_1}\xi_{j_1}^{\nu_2}(\mathscr{A}^{l-1})_{\nu_2\mu_2})$$

$$\leqslant \frac{lk}{pN} a_{l+k-1} \tag{4.3}$$

on the basis of (4.2) we have

$$a_n \leqslant \alpha \sum_{k=0}^{n-2} a_k a_{n-k-2} + \alpha \sum_{k=1}^{n-1} a_k a_{n-k-1} + \frac{n}{N}(a_{n-2} + a_{n-1}) + \frac{n^3}{N^2}(a_{n-2} + a_{n-3})$$
$$= \alpha \sum_{k=1}^{n-2} a_k a_{n-k-2} + \alpha \sum_{k=3}^{n-2} a_k a_{n-k-1} + a_{n-1}\left(\alpha a_0 + \frac{n}{N}\right) + a_{n-2}\left(\alpha a_0 + \frac{n}{N} + \frac{n^3}{N^2}\right) + a_{n-3}\left(\alpha a_2 + \frac{n^3}{N^2}\right) \tag{4.4}$$

and if $n^3/N^2 < \alpha^2 \varepsilon/2$ [that is, $n < (\alpha^2 N^2 \varepsilon/2)^{1/3}$], then taking into account that $a_0 = 1$, $a_1 = 0$, $a_2 = \alpha$, we get from (4.4)

$$a_n \leqslant \alpha \sum_{k=1}^{n-2} a_k a_{n-k-2} + \alpha \sum_{k=3}^{n-2} a_k a_{n-k-1} + (a_{n-1} a_0 + a_{n-2} a_0 + a_{n-3} a_2)\, \alpha(1+\varepsilon) \tag{4.5}$$

Now let us define the sequence a_n^* by the initial conditions

$$a_0^* = 1, \qquad a_1^* = 0, \qquad a_2^* = \alpha(1+\varepsilon) \tag{4.6}$$

and the recurrence formula

$$a_n^* = \alpha(1+\varepsilon)\left(\sum_{k=0}^{n-2} a_k^* a_{n-k-2}^* + \sum_{k=1}^{n-1} a_k^* a_{n-k-1}^*\right) \tag{4.7}$$

Since from the initial conditions (4.6) it follows that

$$a_{0,1,2} \leqslant a_{0,1,2}^*$$

then on the basis of (4.5) and (4.7) by induction it is easy to obtain for $n < (\alpha^2 N^2 \varepsilon/2)^{1/3}$ that

$$a_n \leqslant a_n^* \tag{4.8}$$

On the other hand, if we define $\varphi(x)$ as

$$\varphi(x) = \sum_{n=0}^{\infty} a_n^* x^n \tag{4.9}$$

then from formulas (4.6) and (4.7) it follows that $\varphi(x)$ satisfies the equation

$$\varphi(x)=\alpha(1+\varepsilon)(x+x^2)\,\varphi(x)^2-x\alpha(1+\varepsilon)\,\varphi(x)+1$$

by resolving the quadratic equation for $\varphi(x)$ and choosing the minus sign in front of the square root in order to satisfy the condition $\varphi(0)=a_0^*=1$, we get

$$\varphi(x)=\frac{\alpha(1+\varepsilon)x+1-\{[\alpha(1+\varepsilon)x+1]^2-4\alpha(1+\varepsilon)(x+x^2)\}^{1/2}}{2\alpha(1+\varepsilon)(x+x^2)}$$

since the singularity in $x=0$ can be removed, this function is analytic for $0\leqslant x<\{\alpha(1+\varepsilon)+2[\alpha(1+\varepsilon)]^{1/2}\}^{-1}$, then it is uniformly bounded in this region and we get

$$\begin{aligned}\varphi(x)&=\sum_{n=0}^{\infty}a_n^*x^n\\&\leqslant\varphi\left(\frac{1}{\alpha(1+\varepsilon)+2[\alpha(1+\varepsilon)]^{1/2}}\right)\\&=\frac{2+[\alpha(1+\varepsilon)]^{1/2}}{1+[\alpha(1+\varepsilon)]^{1/2}}<2\end{aligned}$$

and since $a_n^*\geqslant 0$, then we get, for $n<(\alpha N^2\varepsilon/2)^{1/3}$,

$$a_n^*\leqslant 2\{\alpha(1+\varepsilon)+2[\alpha(1+\varepsilon)]^{1/2}\}^n \tag{4.10}$$

But on the other hand

$$\begin{aligned}pa_n&\geqslant E(\|\mathscr{A}\|^n)=\int\lambda^n\,dP(\lambda)\\&\geqslant(\alpha+2\sqrt{\alpha})^n\,(1+\varepsilon)^n\,P((\alpha+2\sqrt{\alpha})(1+\varepsilon))\end{aligned} \tag{4.11}$$

where $P(\lambda)=\mathrm{Prob}(\|\mathscr{A}\|>\lambda)$. Therefore (4.10) and (4.11) give

$$\begin{aligned}&\mathrm{Prob}(\|\mathscr{A}\|>(\alpha+2\sqrt{\alpha})(1+\varepsilon))\\&\quad\leqslant 2p\left(\frac{\alpha(1+\varepsilon)+2[\alpha(1+\varepsilon)]^{1/2}}{(\alpha+2\sqrt{\alpha})(1+\varepsilon)}\right)^{(\alpha^2N^2\varepsilon/2)^{1/3}}\\&\quad\leqslant\exp\left(-\frac{\sqrt{\alpha}\,\varepsilon^{4/3}(2p)^{2/3}}{4(\alpha+2\sqrt{\alpha})}\right)\\&\quad E(\|\mathscr{A}\|^2)\leqslant 4\alpha+o(\alpha)\qquad(\alpha\to 0)\end{aligned}$$

Lemma 4.2. If J_{ij} is defined by the formula (1.1) and $p/N \to \alpha > 0$ $(N, p \to \infty)$, then

$$\operatorname{Prob}(\|J\| > (1+\sqrt{\alpha})^2 + \eta) \leqslant \exp[-\mathscr{M}\eta^{4/3}(2N)^{2/3}]$$

where $\mathscr{M}$ and η depend only on α.

Proof. Lemma 4.2 follows from Lemma 4.1 if we observe that

$$J = \alpha I + \alpha \tilde{\mathscr{A}}$$

where $\tilde{\mathscr{A}}$ is of the same form as $\mathscr{A}$ with $\tilde{p} = N$ and $\tilde{N} = p$ and

$$\tilde{\alpha} = \tilde{p}/\tilde{N} = 1/\alpha$$

Since $\|\alpha I + \alpha \tilde{\mathscr{A}}\| \leqslant \alpha + \alpha\, \|\tilde{\mathscr{A}}\|$, we have

$$\begin{aligned}
&\operatorname{Prob}(\|\alpha I + \alpha \tilde{\mathscr{A}}\| \leqslant \alpha + \alpha(\tilde{\alpha} + 2\tilde{\alpha}^{1/2})(1+\varepsilon)) \\
&\quad \geqslant \operatorname{Prob}(\|\tilde{\mathscr{A}}\| \leqslant (\tilde{\alpha} + 2\tilde{\alpha}^{1/2})(1+\varepsilon)) \\
&\quad = 1 - \operatorname{Prob}(\|\tilde{\mathscr{A}}\| \geqslant (\tilde{\alpha} + 2\tilde{\alpha}^{1/2})(1+\varepsilon))
\end{aligned} \tag{4.12}$$

On the other hand

$$\begin{aligned}
&\operatorname{Prob}(\|\alpha I + \alpha \tilde{\mathscr{A}}\| \leqslant \alpha + \alpha(\tilde{\alpha} + 2\tilde{\alpha}^{1/2})(1+\varepsilon)) \\
&\quad = 1 - \operatorname{Prob}(\|\alpha I + \alpha \tilde{\mathscr{A}}\| \geqslant \alpha + \alpha(\tilde{\alpha} + 2\tilde{\alpha}^{1/2})(1+\varepsilon))
\end{aligned} \tag{4.13}$$

Therefore from (4.12) and (4.13) we obtain

$$\begin{aligned}
&\operatorname{Prob}(\|\alpha I + \alpha \tilde{\mathscr{A}}\| \geqslant \alpha + \alpha(\tilde{\alpha} + 2\tilde{\alpha}^{1/2})(1+\varepsilon)) \\
&\quad \leqslant \operatorname{Prob}(\|\tilde{\mathscr{A}}\| \geqslant (\tilde{\alpha} + 2\tilde{\alpha}^{1/2})(1+\varepsilon)) \\
&\quad \leqslant \exp\left[-\frac{\tilde{\alpha}^{1/2}\varepsilon^{4/3}(2\tilde{p})^{2/3}}{4(\tilde{\alpha} + 2\tilde{\alpha}^{1/2})}\right] \\
&\quad = \exp\left[-\frac{\sqrt{\alpha}\,\varepsilon^{4/3}(2N)^{2/3}}{4(1 + 2\sqrt{\alpha})}\right]
\end{aligned}$$

Finally, if $\eta = \varepsilon(1 + 2\sqrt{\alpha})$, then

$$\operatorname{Prob}(\|J\| \geqslant (1+\sqrt{\alpha})^2 + \eta) \leqslant \exp\left[-\frac{\sqrt{\alpha}\,\eta^{4/3}(2N)^{2/3}}{4(1+2\sqrt{\alpha})^{7/3}}\right]$$

Remark. We have formulated Lemma 4.2 for $\alpha > 0$, but since J is the sum of p projectors, then $\|J\|$ is an increasing function of α and in the case $p/N \to 0$ we can use, e.g., the inequality

$$\mathrm{Prob}(\|J\| > 2) \leqslant \exp(-\mathscr{M} N^{2/3})$$

where $\mathscr{M}$ does not depend on N, p and since $\|J\| \leqslant p$, we have

$$E(\|J\|^2) = \int_0^p \lambda^2 \, dP(\lambda) \leqslant 4 + \int_2^p \lambda^2 \, dP(\lambda)$$
$$\leqslant 4 + p^2 \exp(-\mathscr{M} N^{2/3})$$

Note Added. When this paper was finished we received a preprint by H. Koch[13] where our result is obtained in a different manner.

ACKNOWLEDGMENTS

M.S. thanks the Department of Mathematics of Rome University for kind hospitality and B.T. thanks the Institute of Low Temperature Physics and Engineering of Kharkov for kind hospitality.

REFERENCES

1. W. A. Little, *Math. Biosci.* **13**:13101 (1974).
2. W. A. Little, *Math. Biosci.* **39**:13281 (1974).
3. J. J. Hopfield, *Proc. Natl. Acad. Sci. USA* **79**:2554 (1982).
4. J. J. Hopfield, *Proc. Natl. Acad. Sci. USA* **81**:3088 (1982).
5. M. Mezard, G. Parisi, and M. Virasoro, *Spin-Glass Theory and Beyond* (World Scientific Singapore, 1987).
6. L. A. Pastur and A. L. Figotin, *Teor. Mat. Fis.* **35**:193 (1978).
7. D. J. Amit, H. Gutfreund, and W. Sompolinsky, *Phys. Rev. A* **32**:1007 (1985).
8. C. Procesi and B. Tirozzi, *Int. J. Mod. Phys.* **4**:143 (1989).
9. S. Albeverio, B. Tirozzi, and B. Zegarlinksy, Preprint.
10. H. Koch and J. Piasko, *J. Stat. Phys.* **55**:903 (1989).
11. L. A. Pastur and M. Shcherbina, *J. Stat. Phys.* **62**:1 (1991).
12. D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Ann. Phys.* **173**:30 (1987).
13. M. Shcherbina, Preprint.
14. R. Der, V. Dotsenko and B. Tirozzi, *J. Phys. A: Math. Gen.* **25**:2843 (1992).
15. H. Koch, Preprint.
16. V. V. Yurinskii, Exponential inequalities for sums of random vectors, *J. Multivariate Anal.* **6**:473 (1976).
17. M. Loeve, *Probability Theory* (Princeton University Press, Princeton, New York, 1955).