Self-averaging in a class of generalized Hopfield models

1994 J. Phys. A: Math. Gen. 27 7069

(http://iopscience.iop.org/0305-4470/27/21/021)

J. Phys. A. Math. Gen. 27 (1994) 7069–7077. Printed in the UK

Self-averaging in a class of generalized Hopfield models*

Anton Bovier†

Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, D-10117 Berlin, Germany

Received 20 July 1994, in final form 16 September 1994

Abstract. We prove the almost sure convergence to zero of the fluctuations of the free energy in a class of disordered mean-field spin systems that generalize the Hopfield model in two ways: (i) Multi-spin interactions are permitted and (ii) the random variables ξ_i^μ describing the 'patterns' can have arbitrary distributions with mean zero and finite $(4+\epsilon)$th moments. The number of patterns, M, is allowed to be an arbitrary multiple of the system size. This generalizes a previous result of Bovier, Gayrard and Picco for the standard Hopfield model, and improves a result of Feng and Tirozzi that required M to be a finite constant. Note that the convergence of the mean of the free energy is *not* proven.

1. Introduction

Over the past few years some interesting properties of 'self-averaging' have been observed in two classes of 'spin-glass' type models of the mean-field type, the Sherrington–Kirkpatrick model [SK] and the Hopfield model [FP, Ho]. The latter, largely used in the context of neural networks, may be of particular interest, as it contains a parameter, the number M of stored patterns as a function of the size of the system N, which can be adjusted to alter the properties of the model. In a paper by Pastur and Shcherbina [PS], it was observed that the variance of the free energy of a finite system of size N in the SK model tends to zero as $1/N$, implying the convergence to zero in probability of the difference between the free energy and its mean. This result was later generalized to the Hopfield model by Shcherbina and Tirozzi [ST] under the assumption that the ratio $\alpha = M/N$ remains bounded as $N \uparrow \infty$. Further results of this type can be found in an interesting paper by Pastur, Shcherbina and Tirozzi [PST]. Self-averaging properties of the large deviation rate function as a function of the macroscopic parameters of the model (the so-called 'overlap parameters', see below) were used crucially in two papers by Bovier, Gayrard and Picco [BGP2, BGP3]. There, sharper than variance estimates were needed, and as a consequence [BGP3] contains in particular a proof of the almost sure convergence to zero of the difference between the free energy and its mean, both in the Hopfield model under the assumption that M/N be bounded, and in the SK model. Independently, Feng and Tirozzi [FT] have recently proven such a result in a class of generalized Hopfield models, but under the very restrictive assumption that M itself be a bounded function of N. The purpose of the present paper is to show that such a condition is in fact unnecessary.

Let us describe the class of models we will consider. We denote by $\mathcal{S}_N = \{-1, 1\}^N$ the space of functions $\sigma : \Lambda \to \{-1, 1\}$. We call σ a *spin configuration* on Λ. $\mathcal{S} \equiv \{-1, 1\}^{\mathbb{N}}$

* Work partially supported by the Commission of the European Communities under contract CHRX-CT93-0411.

† E-mail: bovier@iaas-berlin.d400.de

0305-4470/94/217069+09$19.50

denotes the space of half-infinite sequences equipped with the product topology of the discrete topology on $\{-1, 1\}$. We denote by $\mathcal{B}_\Lambda$ and $\mathcal{B}$ the corresponding Borel sigma algebras. We will define a random Hamiltonian function on the spaces $\mathcal{S}_\Lambda$ as follows. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space. Let $\xi \equiv \{\xi_i^\mu\}_{i,\mu\in\mathbb{N}}$ be a two-parameter family of independent, random variables on this space. We will specify our assumptions on their distribution later. In the context of neural networks, one assumes usually that $\mathbb{P}(\xi_i^\mu = 1) = \mathbb{P}(\xi_i^\mu = -1) = \frac{1}{2}$, but here we aim for more general distributions. We consider Hamiltonians of the form

$$H_N(\sigma) \equiv -\frac{1}{N^{r-1}} \sum_{\mu=1}^{M(N)} \sum_{i_1,\ldots,i_r=1}^{N} \xi_{i_1}^\mu \cdots \xi_{i_r}^\mu \sigma_{i_1} \cdots \sigma_{i_r} . \tag{1.1}$$

Here $r \geqslant 2$ is some chosen integer. The case $r = 2$ corresponds to the usual Hopfield model, and models with general r were introduced by Lee *et al* [Lee] and Peretto and Niez [PN]. Feng and Tirozzi [FT] also studied these models, but removed the terms in the sum where two or more indices coincide, which actually amounts to adding a term of the order of a constant to H which does not alter the free energy. One may actually consider more general models in which the Hamiltonian is given as a linear combination of terms of the type (1.1) with different values of r. This only complicates, but does not really alter, the proofs, and our results can easily be extended to this situation.

Let us introduce the so-called 'overlap parameters'. This is the M-dimensional vector $m_N(\sigma)$ whose components are given by

$$m_N^\mu(\sigma) \equiv \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \sigma_i . \tag{1.2}$$

In terms of these quantities, the Hamiltonian can be written in the very convenient form

$$H_N(\sigma) = -N \, \|m_N(\sigma)\|_r^r . \tag{1.3}$$

We define the partition function

$$Z_N \equiv \frac{1}{2^N} \sum_{\sigma\in\mathcal{S}_N} \mathrm{e}^{-\beta H_N(\sigma)} \tag{1.4}$$

and the free energy

$$F_N(\beta) \equiv -\frac{1}{\beta N} \ln Z_N(\beta) . \tag{1.5}$$

It will be important to realize that

$$\|m_N(\sigma)\|_2^2 \leqslant \|A(N)\| \tag{1.6}$$

where $A(N)$ is the $N \times N$ matrix whose elements are

$$A_{ij}(N) \equiv \frac{1}{N} \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu . \tag{1.7}$$

Properties of the maximal eigenvalues of this matrix will be crucial for us. The eigenvalue distribution of this matrix was first analysed by Marchenko and Pastur [MP]. Girko [Gi] proved that, under the hypothesis of theorem 1, the maximal eigenvalue of $A(N)$ converges to $(1 + \sqrt{\alpha})^2$ in probability. Adding the ideas used by Bai and Yin [BY] one can easily show that this convergence also takes place almost surely, and even in the case where only the fourth moment of ξ_i^μ is finite. We will need additional estimates on the moments of $\|A(N)\|$ which we are only able to prove if we have a little more than four moments. The relevant estimate is formulated in the following lemma.

Lemma 1.1. Assume that $\mathbb{E}\xi_i^\mu = 0$, $\mathbb{E}\left(\xi_i^\mu\right)^2 = 1$ and $\mathbb{E}\left(\xi_i^\mu\right)^{4+\epsilon} \leqslant c < \infty$, for some $\epsilon > 0$. Then, for any $\eta \geqslant 6$ and any $\delta > 0$, if N is sufficiently large

$$\mathbb{P}\left[\|A(N)\| \geqslant (1+\sqrt{\alpha})^2(1+z)\right] \leqslant N(1+z)^{-N^{\gamma/\eta}\delta^{-1/\eta}} + \frac{c\alpha}{N^{\epsilon/2}\delta^{4+\epsilon}} \tag{1.8}$$

where $\alpha = \frac{M}{N}$, $\gamma = \frac{\epsilon}{4(4+\epsilon)}$.

Remark. The proof of lemma 1.1 is in fact an adaptation of the the truncation idea in [BY] and fairly standard estimates on the traces of powers of A, as in [BY] (but see also [BGP1]). We will therefore not give the details of the proof of lemma 1.1, but only mention that the second term is a bound on the probability that any of the ξ_i^μ exceeds the value $\sqrt{N}\delta$, while the first is a bound one would obtain *if* all ξ_i^μ satisfied this condition.

With this in mind we define

$$\tilde{f}_N(\beta) \equiv -\beta^{-1} \ln Z_N(\beta)\, \mathbb{I}_{\{\|A\| \leqslant 2(1+\alpha)^2\}} \tag{1.9}$$

We will prove:

Theorem 1. Assume that $\lim \frac{M(N)}{N} = \alpha < \infty$ and ξ satisfies the assumptions of lemma 1.1. Then

(i) If $r = 2$, for all $n < \infty$ there exists $\tau_n < \infty$, such that for all $\tau \geqslant \tau_n$, and for N sufficiently large

$$\mathbb{P}\left[\left|\tilde{f}_N(\beta) - \mathbb{E}\tilde{f}_N(\beta)\right| \geqslant \tau(\ln N)^{3/2}N^{1/2}\right] \leqslant N^{-n}\,. \tag{1.10}$$

(ii) If $r \geqslant 3$, then there exist constants $C, c, c' > 0$ such that

$$\mathbb{P}\left[\left|\tilde{f}_N(\beta) - \mathbb{E}\tilde{f}_N(\beta)\right| \geqslant zN\right] \leqslant \begin{cases} \mathrm{e}^{-cNz^2} & \text{if } 0 \leqslant z < C \\ \mathrm{e}^{-Nc'z} & \text{if } z \geqslant C\,. \end{cases} \tag{1.11}$$

We prove theorem 1 in the next section. Before doing that, we will show that it implies the following theorem.

Theorem 2. Under the assumptions of theorem 1

$$\lim_{N\uparrow\infty} |F_N(\beta) - \mathbb{E}F_N(\beta)| = 0 \qquad \text{a.s.} \tag{1.12}$$

Remark. Theorem 2 was proven under the additional assumption that $\xi_i^\mu = \pm 1$ for the case $r = 2$ in [BGP3]. In [FT] theorem 2 was proven under the hypothesis $M(N) \leqslant M_0 < \infty$ and that $\mathbb{E}\left(\xi_i^\mu\right)^4 < \infty$.

Remark. Theorem 2 may in some way be regarded as a *strong law of large numbers*. We are, however, reluctant to employ this term, because the convergence of $\mathbb{E}F_N(\beta)$ to a limit is, in general, not proven. In the standard Hopfield model this was proven under the assumption $\lim_{N\uparrow\infty} \frac{M(N)}{N} = 0$ by Koch [K] (see also [BG]).

We conclude the introduction by giving the proof of theorem 2, assuming theorem 1.

Proof of theorem 2. Set $2(1+\sqrt{\alpha})^2 = \rho$. Obviously,

$$F_N(\beta) = \frac{1}{N}\tilde{f}_n(\beta) + F_N(\beta)\mathbb{1}_{\{\|A(N)\|>\rho\}}\,. \tag{1.13}$$

By theorem 1 and the first Borel–Cantelli lemma it follows that

$$\lim_{N\uparrow\infty}\frac{1}{N}\left|\tilde{f}_N(\beta) - \mathbb{E}\tilde{f}_N(\beta)\right| = 0 \qquad \text{a.s.} \tag{1.14}$$

Thus theorem 2 will be proven if we can show that $F_N(\beta)\mathbb{1}_{\{\|A(N)\|>\rho\}} \downarrow 0$ both almost surely and in mean. The almost sure convergence follows easily, since

$$\mathbb{P}\left[F_N(\beta)\mathbb{1}_{\{\|A(N)\|>\rho\}} \neq 0 \text{ i.o.}\right] \leqslant \mathbb{P}[\|A(N)\| > \rho \text{ i.o.}] = 0 \tag{1.15}$$

where the last equality follows from applying lemma 3.1 from Bai and Yin [BY]. Finally, to prove convergence of the mean, we use first of all that

$$\begin{aligned}|H_N(\sigma)| &\leqslant N\|m_N(\sigma)\|_2^2\|m_N(\sigma)\|_\infty^{r-2}\\ &\leqslant N\|m_N(\sigma)\|_2^r\\ &\leqslant N\|A(N)\|^{r/2}\end{aligned} \tag{1.16}$$

and therefore

$$\begin{aligned}|F_N(\beta)|\,\mathbb{1}_{\{\|A(N)\|>\rho\}} &\leqslant \frac{1}{\beta N}|\ln Z_N|\mathbb{1}_{\{\|A(N)\|>\rho\}}\\ &\leqslant \|A(N)\|^{r/2}\mathbb{1}_{\{\|A(N)\|>\rho\}}\,.\end{aligned} \tag{1.17}$$

But

$$\begin{aligned}\mathbb{E}\|A(N)\|^x\mathbb{1}_{\{\|A(N)\|>\rho\}} &= \rho^x\mathbb{P}[\|A(N)\| > \rho] + \int_\rho^\infty xy^{x-1}\mathbb{P}\left[\|A(N)\| > y\right]\mathrm{d}y\\ &\leqslant 2^x(1+\sqrt{\alpha})^{2x}\left(N2^{-N^{\gamma/6}} + \frac{c\alpha}{N^{\epsilon/2}}\right) + 2^x(1+\sqrt{\alpha})^{2x}\\ &\quad\times\int_1^\infty x(1+y)^{x-1}\left(N\exp\left(-yN^{\gamma/\eta(x)}y^{-x/(4\eta(x))}\right) + \frac{c\alpha}{N^{\epsilon/2}y^{x+\epsilon/2}}\right)\mathrm{d}y\,.\end{aligned} \tag{1.18}$$

To obtain the last expression we used lemma 1.1 and made the choice $\delta = \delta(y,x) = y^{x/4}$ and $\eta(x) = \max(6, x/8)$. Obviously, the right-hand side of (1.18) tends to zero as $N\uparrow\infty$, as desired. This concludes the proof of theorem 2, assuming theorem 1. □

Remark. Note that the estimate in (1.18) implies in particular that

$$\mathbb{E}\|A(N)\| \leqslant C(1+\sqrt{\alpha})^2 \tag{1.19}$$

for some constant depending only on ϵ. This is relevant for proving theorem 1 in the case $r = 2$ (see [BGP3]).

2. Proof of theorem 1

The basic idea of the proof is the same as in [BGP3] where the case $r = 2$ has been considered, but some modifications are necessary, in particular to avoid any restrictions on the value of α.

The fact that sharper estimates can be obtained in the case $r \geqslant 3$ may justify the presentation of the details of the proof in that case.

We first introduce the decreasing sequence of sigma-algebras $\mathcal{F}_k$ that are generated by the random variables $\{\xi_i^\mu\}_{i\geqslant k}^{\mu\in\mathbb{N}}$. Since the variables $\tilde{f}_N(\beta)$ are non-zero only if $\|A\| \leqslant 2(1+\sqrt{\alpha})^2$, we may introduce the event

$$\mathcal{A} \equiv \{\|A\| \leqslant 2(1+\sqrt{\alpha})^2\} \subset \mathcal{F} \tag{2.1}$$

and the corresponding trace-sigma algebras $\tilde{\mathcal{F}} \equiv \mathcal{F} \cap \mathcal{A}$. This allows us to introduce the corresponding Martingale difference sequence [Yu]

$$\tilde{f}_N^{(k)}(\beta) \equiv \mathbb{E}\left[f_N(\beta)\middle|\tilde{\mathcal{F}}_k\right] - \mathbb{E}\left[f_N(\beta)\middle|\tilde{\mathcal{F}}_{k+1}\right]. \tag{2.2}$$

Notice that we have the identity

$$\tilde{f}_N(\beta) - \mathbb{E}\tilde{f}_N(\beta) \equiv \sum_{k=1}^{N} \tilde{f}_N^{(k)}(\beta)\mathbb{P}[\mathcal{A}] \tag{2.3}$$

by the definition of conditional expectations. The factor $\mathbb{P}[\mathcal{A}]$ tends to 1 as $\mathrm{N} \uparrow 1$, so that we just have to control the sum of the $\tilde{f}_N^{(k)}(\beta)$. To get the sharpest possible estimates, we want to use an exponential inequality. To this end we observe that [BGP3]

$$\begin{aligned}\mathbb{P}\left[\left|\sum_{k=1}^{N}\tilde{f}_N^{(k)}(\beta)\right| \geqslant Nz\right] &\leqslant 2\inf_{t\in\mathbb{R}} \mathrm{e}^{-|t|Nz}\mathbb{E}\exp\left\{t\sum_{k=1}^{N}\tilde{f}_N^{(k)}(\beta)\right\} \\ &= 2\inf_{t\in\mathbb{R}} \mathrm{e}^{-|t|Nz}\mathbb{E}\left[\mathbb{E}\left[\cdots\mathbb{E}\left[\mathrm{e}^{t\tilde{f}_N^{(1)}(\beta)}\middle|\tilde{\mathcal{F}}_2\right]\mathrm{e}^{t\tilde{f}_N^{(2)}(\beta)}\middle|\tilde{\mathcal{F}}_3\right]\cdots\mathrm{e}^{t\tilde{f}_N^{(N)}(\beta)}\middle|\tilde{\mathcal{F}}_{N+1}\right].\end{aligned} \tag{2.4}$$

To make use of this inequality, we need bounds on the conditional Laplace transforms; namely, *if* we can show that, for some function $\mathcal{L}^{(k)}(t)$, $\ln\mathbb{E}\left[\mathrm{e}^{t\tilde{f}_N^{(k)}(\beta)}\middle|\tilde{\mathcal{F}}_{k+1}\right] \leqslant \mathcal{L}^{(k)}(t)$, uniformly in $\mathcal{F}_{k+1}$, then we obtain that

$$\mathbb{P}\left[\left|\sum_{k=1}^{N}\tilde{f}_N^{(k)}(\beta)\right| \geqslant Nz\right] \leqslant 2\inf_{t\in\mathbb{R}}\exp\left(-|t|Nz + \sum_{k=1}^{N}\mathcal{L}^{(k)}(t)\right). \tag{2.5}$$

Note that this construction is so far completely model-independent. In the estimation of the conditional Laplace transforms, a conventional trick [PS] is to introduce a continuous family of Hamiltonians, $\tilde{H}_N^{(k)}(\sigma, u)$, that are equal to the original one for $u = 1$ and are independent of ξ_k for $u = 0$. We first introduce the $M(N)$-dimensional vectors

$$m_N^{(k)}(\sigma, u) \equiv \frac{1}{N}\left(\sum_{\substack{i \\ i\neq k}}\xi_i\sigma_i + u\xi_k\sigma_k\right) \tag{2.6}$$

and then define

$$\tilde{H}_N^{(k)}(\sigma, u) = -N \left\| m_N^{(k)}(\sigma, u) \right\|_r^r . \tag{2.7}$$

Note that this procedure can of course be used in all cases where the Hamiltonian is a function of the macroscopic order parameters. Naturally, we set

$$Z_N^{(k)}(\beta, u) \equiv \frac{1}{2^N} \sum_{\sigma \in \mathcal{S}_N} \mathrm{e}^{-\beta \tilde{H}_N^{(k)}(\sigma,u)} \tag{2.8}$$

and finally

$$f_N^{(k)}(\beta, u) = -\beta^{-1} \left(\ln Z_N^{(k)}(\beta, u) - \ln Z_N^{(k)}(\beta, 0) \right) . \tag{2.9}$$

Since for the remainder of the proof, β as well as N will be fixed values, to simplify our notations we will write $f_k(u) \equiv f_N^{(k)}(\beta, u)$. It relates to $\tilde{f}_N^{(k)}(\beta)$ via

$$\tilde{f}_N^{(k)}(\beta) = \mathbb{E}\left[f_k(1) | \tilde{\mathcal{F}}_k \right] - \mathbb{E}\left[f_k(1) | \tilde{\mathcal{F}}_{k+1} \right] . \tag{2.10}$$

To bound the Laplace transform, we use that, for all $x \in \mathbb{R}$,

$$\mathrm{e}^x \leqslant 1 + x + \tfrac{1}{2} x^2 \mathrm{e}^{|x|} \tag{2.11}$$

so that

$$\mathbb{E}\left[\mathrm{e}^{t \tilde{f}_N^{(k)}(\beta)} \Big| \tilde{\mathcal{F}}_{k+1} \right] \leqslant 1 + \tfrac{1}{2} t^2 \mathbb{E}\left[\left(\tilde{f}_N^{(k)}(\beta) \right)^2 \mathrm{e}^{|t \tilde{f}_N^{(k)}(\beta)|} \Big| \tilde{\mathcal{F}}_{k+1} \right] . \tag{2.12}$$

Our strategy in [BGP3] was to use a rather *poor* uniform bound on $\tilde{f}_N^{(k)}(\beta)$ in the exponent but to prove a better estimate on the remaining conditioned expectation of the square. Here it will do the same. Notice that $f_k(u)$ is convex, $f_k(0) = 0$, and therefore $|f_k(1)| \leqslant \max \left(|f_k'(0)|, |f'(1)| \right)$. But

$$f_k'(u) = \mathcal{E}_{k,u} \left(\frac{\partial}{\partial u} H_N^{(k)}(\sigma, u) \right) \tag{2.13}$$

where $\mathcal{E}_{k,u}$ denotes the expectation w.r.t. the probability measure

$$\frac{1}{Z_N^{(k)}(\beta, u)} \mathrm{e}^{-\beta \tilde{H}_N^{(k)}(\sigma,u)} d\sigma . \tag{2.14}$$

One easily verifies that $f_k'(0) = 0$, so that we can use in the following that $|f_k(1)| \leqslant |f_k'(1)|$. Obviously

$$\left| f_k'(1) \right| \leqslant \mathcal{E}_{k,u} \left| \frac{\partial}{\partial u} H_N^{(k)}(\sigma, u) \right| . \tag{2.15}$$

Computing the derivative, we obtain

$$\begin{aligned}
\left|\frac{\partial}{\partial u} H_N^{(k)}(\sigma, u=1)\right| &= \left|\sum_{\mu=1}^{M(N)} r\xi_k^\mu \sigma_k \left[m_N^\mu(\sigma)\right]^{r-1}\right| \\
&\leqslant r\left|\sum_{\mu=1}^{M(N)} \xi_k^\mu \sigma_k \left[m_N^\mu(\sigma)\right]^{r-1}\right| \\
&\leqslant r\,\|m_N(\sigma)\|_\infty^{r-3} \sum_{\mu=1}^{M(N)} \left[m_N^\mu(\sigma)\right]^2 \\
&\leqslant r\,\|m_N(\sigma)\|_\infty^{r-3}\, \|m_N^{(k)}(\sigma, u)\|_2^2 .
\end{aligned} \tag{2.16}$$

In the usual case, where $|\xi_i^\mu| \leqslant 1$, we can bound the sup-norms appearing in (2.16) by $\left\|m_N(\sigma)\right\|_\infty \leqslant 1$; in the case of unbounded ξ, we can still use that $\|m_N(\sigma)\|_\infty \leqslant \|m_N(\sigma)\|_2$. On $\mathcal{A}$, the latter is bounded by $2(1+\sqrt{\alpha})^2$. Of course here we assumed that $r \geqslant 3$. In this case, therefore, on $\mathcal{A}$,

$$\begin{aligned}
\mathcal{E}_{k,1}\left|\frac{\partial}{\partial u} H_N^{(k)}(\sigma, u=1)\right| &\leqslant r\mathcal{E}_{k,1}\|m_N(\sigma)\|_2^2 \\
&\leqslant r2\left(1+\sqrt{\alpha}\right)^2
\end{aligned} \tag{2.17}$$

if $|\xi_i^\mu| \leqslant 1$, and

$$\begin{aligned}
\mathcal{E}_{k,1}\left|\frac{\partial}{\partial u} H_N^{(k)}(\sigma, u=1)\right| &\leqslant r\mathcal{E}_{k,1}\|m_N(\sigma)\|_2^{r-1} \\
&\leqslant r2\left(\left(1+\sqrt{\alpha}\right)^2\right)^{(r-1)/2}
\end{aligned} \tag{2.18}$$

in general. Using these estimates, we see that on $\mathcal{A}$ we have

$$\left|\tilde{f}_N^{(k)}(\beta)\right| \leqslant C \tag{2.19}$$

where $C \equiv C(\alpha)$ is some finite constant depending on $\frac{M}{N}$. Using this bound in (2.12) we see that

$$\Lambda L^{(k)}(t) \leqslant \frac{C^2}{2} t^2 \mathrm{e}^{C|t|} . \tag{2.20}$$

To obtain (1.11), we insert this bound into (2.5) and bound the infimum over t by its value for $t = z/C^2$, if $z < \ln 2C$, and by its value for $t = C^{-1}$, if $z \geqslant C \ln 2$. This concludes (ii) of theorem 1. □

This leaves us with the case $r = 2$. The new difficulties have been treated in [BGP3] and here we just recall the main steps. Instead of (2.22) we have here that

$$\begin{aligned}
\left|\frac{\partial}{\partial u} H_N^{(k)}(\sigma, u)\right| &\leqslant \sum_{\mu=1}^{M} \left|m_N^{(k),\mu}(\sigma, u)\right| \\
&= \left\|m_N^{(k)}(\sigma, u)\right\|_1 \leqslant \sqrt{M}\left\|m_N^{(k)}(\sigma, u)\right\|_2 .
\end{aligned} \tag{2.21}$$

Using this, we get

$$\begin{aligned}\mathbb{E}\left[e^{t\tilde{f}_N^{(k)}(\beta)}\big|\tilde{\mathcal{F}}_{k+1}\right] &\leqslant 1+\tfrac{1}{2}t^2e^{4(1+\sqrt{\alpha})^2|t|\sqrt{M}}\mathbb{E}\left[\left(\tilde{f}_N^{(k)}(\tilde{m})\right)^2\big|\tilde{\mathcal{F}}_{k+1}\right]\\ &\leqslant \exp\left(\tfrac{1}{2}t^2e^{4(1+\sqrt{\alpha})^2|t|\sqrt{M}}\mathbb{E}\left[\left(\tilde{f}_N^{(k)}(\tilde{m})\right)^2\big|\tilde{\mathcal{F}}_{k+1}\right]\right).\end{aligned} \tag{2.22}$$

Then, just as in [BGP3] one easily verifies (using the convexity of $f_k(u)$) that

$$\mathbb{E}\left[\left(\tilde{f}_N^{(k)}(\tilde{m})\right)^2\big|\tilde{\mathcal{F}}_{k+1}\right]\leqslant \mathbb{E}\left[(f_k'(1))^2\big|\tilde{\mathcal{F}}_{k+1}\right]. \tag{2.23}$$

Symmetrizing with respect to the ξ_i^μ that are integrated over in (2.23) one obtains from here [BGP3]

$$\mathbb{E}\left[(f_k'(1))^2\big|\tilde{\mathcal{F}}_{k+1}\right]\leqslant 2(1+\sqrt{\alpha})^2\mathbb{E}\left[\|B^{(k)}\|\,|\mathcal{A}\right] \tag{2.24}$$

where $B^{(k)}$ is the random matrix with elements

$$B_{\mu\nu}^{(k)}\equiv\frac{1}{k}\sum_{j=1}^{k}\xi_j^\mu\xi_j^\nu \tag{2.25}$$

Note that the conditioning on $\mathcal{A}$ in (2.24) is essentially irrelevant here, since the probability of $\mathcal{A}$ is close to 1. Now it is easy to see that $\|B^{(k)}\| = \|A(k)\|$, and so, by the estimate (1.18), we have that

$$\mathbb{E}\left[\|B^{(k)}\|\right]\leqslant c\left(1+\sqrt{\tfrac{M}{k}}\right)^2. \tag{2.26}$$

Therefore we have

$$\mathbb{E}\exp\left\{t\sum_{k=1}^{N}\tilde{f}_N^{(k)}(\tilde{m})\right\}\leqslant\exp\left\{ct^2\epsilon^{C|t|\sqrt{M}}N(1+4\sqrt{\alpha}+\alpha\ln N)\right\}. \tag{2.27}$$

Inserting this in (2.4) and choosing $t = n\frac{\ln N}{zN}$ then gives estimate (i) of theorem 1. □

Let us conclude this paper with some final remarks. We have shown in this paper how sharp estimates on the fluctuations of the free energy (as a function of the 'overlap parameters') can be obtained in a very wide class of disordered mean-field models generalizing the Hopfield model (note that the case of the Sherrington–Kirkpatrick model can be treated in much the same way, see [BGP3]). In particular, we have shown that typical fluctuations are of order $1/\sqrt{N}$ and converge to zero almost surely. We should again stress, however, that this does not imply that the free energy itself converges almost surely to some value in the thermodynamic limit. The problem here is the average of the free energy. Note that in most of the literature on disordered systems, one tries to compute this average, tentatively *assuming* the self-averaging. But, although heuristic techniques and in particular the replica-trick allows one to do this to some extent, there is in general no *rigorous* argument that would ensure that the average of the finite-volume free energy converges. For the models considered in this paper, the average of the free energy is uniformly *bounded* between two constants; this follows from the bound (1.19). But nothing does, in principle, exclude that the free energy is a very irregular, oscillating function of the volume. Currently, only in the case $\alpha(N)\downarrow 0$, or at high temperatures, can this be rigorously excluded.

References

[BG] Bovier A and Gayrard V 1993 Rigorous results on the thermodynamics of the dilute Hopfield model *J. Stat. Phys.* **69** 597–627

[BGP1] Bovier A, Gayrard V and Picco P 1994 Gibbs states of the Hopfield model in the regime of perfect memory *Prob. Theor. Rel. Fields* **99** at press

[BGP2] Bovier A, Gayrard V and Picco P 1994 Large deviation principles for the Hopfield model and the Kac-Hopfield model *IAAS preprint* 95, *Prob. Theor. Rel. Fields* to appear

[BGP3] Bovier A, Gayrard V and Picco P 1994 Gibbs states of the Hopfield model with extensively many patterns *IAAS preprint* 97, *J. Stat. Phys.* to appear

[BY] Bai Z D and Yin Y Q 1988 Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix *Ann. Prob.* **16** 1729–41

[FP] Pastur L A and Figotin A L 1977 Exactly soluble model of a spin glass *Sov. J. Low Temp. Phys.* **3** 378–83; 1978 On the theory of disordered spin systems *Theor. Math. Phys.* **35** 403–14

[FT] Feng J and Tirozzi B 1994 The SLLN for the free energy in the Hopfield type model and Little model with finite number of patterns *Preprint*

[Gi] Girko V L 1989 Limit theorems for maximal and minimal eigenvalues of random matrices *Theor. Prob. Appl.* **35** 680–95

[Ho] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities *Proc. Natl Acad. Sci. USA* **79** 2554–8

[K] Koch H 1993 A free energy bound for the Hopfield model *J. Phys. A: Math Gen.* **26** L353–5

[Lee] Lee Y C, Doolen G, Chen H H, Sun G Z, Maxwell T, Lee H Y and Gilles C L 1986 Machine learning using higher order correlation networks *Physica* **22D** 276–306

[MP] Marchenko V A and Pastur L A 1967 Eigenvalue distribution in some ensembles of random matrices *Math. USSR Sbornik* **72** 536–67

[PN] Peretto P and Niez J J 1986 Long term memory storage capacity of multiconnected neural networks *Biol. Cybern.* **54** 53–63

[PS] Pastur L and Shcherbina M 1991 Absence of self-averaging of the order parameter in the Sherrington–Kirkpatrick model *J. Stat. Phys.* **62** 1–19

[PST] Pastur L, Shcherbina M and Tirozzi B 1994 The replica symmetric solution without the replica trick for the Hopfield model *J. Stat. Phys.* **74** 1161–83

[SK] Sherrington D and Kirkpatrick S 1972 Solvable model of a spin glass *Phys. Rev. Lett.* **35** 1792–6

[ST] Shcherbina M and Tirozzi B 1992 The free energy for a class of Hopfield models *J. Stat. Phys.* **72** 113–25

[Yu] Yurinskii V V 1976 Exponential inequalities for sums of random vectors *J. Multivariate Anal.* 6 473–99