The Annals of Probability
1996, Vol. 24, No. 3, 1444–1475

AN ALMOST SURE LARGE DEVIATION PRINCIPLE FOR THE HOPFIELD MODEL[1]

BY ANTON BOVIER AND VÉRONIQUE GAYRARD

Weierstrass-Institute and
Centre de Physique Théorique

We prove a large deviation principle for the finite-dimensional marginals of the Gibbs distribution of the macroscopic "overlap" parameters in the Hopfield model in the case where the number of random "patterns" M, as a function of the system size N, satisfies $\limsup M(N)/N = 0$. In this case, the rate function is independent of the disorder for almost all realizations of the patterns.

1. Introduction. Mean field models in statistical mechanics furnish nice examples for the interpretation of thermodynamics as the theory of large deviation for Gibbs measures of microscopically defined statistical mechanics systems [9]. Roughly speaking, in such models the Hamiltonian is only a function of (extensive) "macroscopic" quantities (density, magnetization, etc.) of the system. In the thermodynamic limit, the distribution of these quantities is expected to be concentrated on a sharp value and to satisfy a large deviation principle. The corresponding rate functions are then the thermodynamic potentials (free energy, pressure) that govern the macroscopic response of the system to external (intensive) conditions. The classical paradigm of the theory is that the number of relevant macroscopic variables is excessively small (order of 10) compared to the number of microscopic variables (order of 10^{23}).

Over the last decade, the formalism of statistical mechanics and thermodynamics has found increasing applications in systems in which the macroscopic behaviour is far more complex and described by a "large' number of variables. Such systems can be found in biology (heteropolymers, neural networks), but also in the domain of disordered solids and, in particular, spin glasses. Some fundamental aspects of these ideas are discussed in an interesting recent paper by Parisi [17]. For such systems, many basic problems are not very well understood, and many technical aspects defy a mathematical investigation at the present time. An interesting toy model (that nonetheless also has practical relevance) where this situation can be studied and for which mathematical results are available is the Hopfield model [11, 18, 19]. This model is a mean field spin system in the sense explained above. However, the Hamiltonian, instead of being a function of a few macroscopic variables is a function of

Received May 1995; revised November 1995.
[1]Work partially supported by the Commission of the European Union under contract CHRX-CT93-0411.
AMS 1991 *subject classifications*. 60F10, 82B44, 82C32.
Key words and phrases. Hopfield model, neural networks, self-averaging, large deviations.

macroscopic variables that are random functions of the microscopic ones, and those number tends to infinity with the size of the system in a controllable way. More specifically, the model is defined as follows.

Let $\mathcal{S}_N \equiv \{-1, 1\}^N$ denote the set of functions σ: $\{1, \dots, N\} \to \{-1, 1\}$ and set $\mathcal{S} \equiv \{-1, 1\}^{\mathbb{N}}$. We call σ a spin configuration and denote by σ_i the value of σ at i. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space and let ξ_i^μ, $i, \mu \in \mathbb{N}$, denote a family of independent identically distributed random variables on this space. For the purposes of this paper we will assume that $\mathbb{P}[\xi_i^\mu = \pm 1] = \frac{1}{2}$, but more general distributions can be considered. We will write $\xi^\mu[\omega]$ for the N-dimensional random vector whose ith component is given by $\xi_i^\mu[\omega]$ and call such a vector a pattern. On the other hand, we use the notation $\xi_i[\omega]$ for the M-dimensional vector with the same components. M will be chosen as a function of N and the function $M(N)$ is an important parameter of the model. We will generally set $\alpha \equiv \alpha(N) \equiv (M(N))/N$. When we write $\xi[\omega]$ without indices, we frequently will consider it as an $N \times M$ matrix and we write $\xi^t[\omega]$ for the transpose of this matrix. Thus, $\xi^t[\omega]\xi[\omega]$ is the $M \times M$ matrix whose elements are $\sum_{i=1}^N \xi_i^\mu[\omega]\xi_i^\nu[\omega]$. With this in mind, we will use throughout the paper a vector notation with $(\cdot, \cdot)$ standing for the scalar product in whatever space the argument may lie. For example, the expression (y, ξ_i) stands for $\sum_{\mu=1}^M \xi_i^\mu y_\mu$ and so forth.

We define random maps $m_N^\mu[\omega]$: $\mathcal{S}_N \to [-1, 1]$ through

$$m_N^\mu[\omega](\sigma) \equiv \frac{1}{N} \sum_{i=1}^N \xi_i^\mu[\omega]\sigma_i. \tag{1.1}$$

(We will make the dependence of random quantities on the random parameter ω explicit by an added $[\omega]$ whenever we want to stress it; otherwise, we will frequently drop the reference to ω to simplify the notation.) Naturally, these maps "compare" the configuration σ globally to the random configuration $\xi^\mu[\omega]$. A Hamiltonian is now defined as the simplest negative function of these variables, namely,

$$H_N[\omega](\sigma) \equiv -\frac{N}{2} \sum_{\mu=1}^{M(N)} \left(m_N^\mu[\omega](\sigma)\right)^2, \tag{1.2}$$

where $M(N)$ is some, generally increasing, function that crucially influences the properties of the model. With $\|\cdot\|_2$ denoting the l_2-norm in $\mathbb{R}^M$, (1.2) can be written in the compact form

$$H_N[\omega](\sigma) = -\frac{N}{2} \|m_N[\omega](\sigma)\|_2^2. \tag{1.3}$$

Through this Hamiltonian we define, in a natural way, finite volume Gibbs measures on $\mathcal{S}_N$ via

$$\mu_{N,\beta}[\omega](\sigma) \equiv \frac{1}{Z_{N,\beta}[\omega]} \exp(-\beta H_N[\omega](\sigma)) \tag{1.4}$$

and the induced distribution of the overlap parameters via

$$\mathcal{Q}_{N,\beta}[\omega] \equiv \mu_{N,\beta}[\omega] \circ m_N[\omega]^{-1}. \tag{1.5}$$

The normalizing factor $Z_{N,\beta}[\omega]$, given by

$$Z_{N,\beta}[\omega] \equiv 2^{-N} \sum_{\sigma \in \mathcal{S}_N} \exp(-\beta H_N[\omega](\sigma)) \equiv \mathbb{E}_\sigma \exp(-\beta H_N[\omega](\sigma)), \tag{1.6}$$

is called the partition function.

This model has been studied very heavily in the physics literature. As a basic introduction to what is commonly believed about its properties, we refer to the seminal paper by Amit, Gutfreund and Sompolinsky [1]. Over the last few years, a considerable amount of mathematically rigorous results on these measures has been established [3–6, 10, 12–14, 16, 20, 22]. It is known that under the hypothesis that $\limsup_{N\uparrow\infty} M(N)/N = 0$, weak limits can be constructed for which the $\mathcal{Q}_N$ converge to Dirac measures in $\mathbb{R}^\infty$ [4]. Disjoint weak limits have also been constructed in the case where $\limsup_{N\uparrow\infty} M(N)/N = \alpha > 0$, for small α in [6]. In this note we restrict our attention to the case $\alpha = 0$ and the question to what extent a large deviation principle (LDP) for the distribution of the macroscopic overlaps can be proven.

A first step in this direction had been taken already in [5]. There, a LDP was proven, but only under the restrictive assumption $M(N) < \ln N/\ln 2$, while only a weaker result concerning the existence of the convex hull of the rate function was proven in the general case $\alpha = 0$ in a rather indirect way. The first LDP in the Hopfield model was proven earlier by Comets [7] for the case of a finite number of patterns. Here we prove a LDP under more natural, and probably optimal, assumptions.

Since the overlap parameters form a vector in a space of unbounded dimension, the most natural setting for a LDP is to consider the finite-dimensional marginals. Let $I \subset \mathbb{N}$ be a finite set of integers, let $\mathbb{R}^I \subset \mathbb{R}^{\mathbb{N}}$ denote the corresponding subspace and, finally, let Π_I denote the canonical projection from $\mathbb{R}^J$ onto $\mathbb{R}^I$ for all $J \subset \mathbb{N}$ such that $I \subset J$. Without loss of generality, we can and will assume in the sequel that $I = \{1, \dots, |I|\}$. Let us introduce the maps $n_p\colon [-1,1]^{2^p} \to [-1,1]^p$ through

$$n_p(y) \equiv 2^{-p} \sum_{\gamma=1}^{2^p} e_\gamma y_\gamma, \tag{1.7}$$

where e_γ, $\gamma = 1, \dots, 2^p$, is some enumeration of all 2^p vectors in $\mathbb{R}^p$ whose components take values ± 1 only. Given $I \subset \mathbb{N}$, we define the set $D_{|I|}$ as the set

$$D_{|I|} \equiv \{m \in \mathbb{R}^{|I|} \mid \exists y \in [-1,+1]^{2^{|I|}}\colon\ n_{|I|}(y) = m\}. \tag{1.8}$$

THEOREM 1. *Assume that* $\limsup_{N\uparrow\infty}(M/N) = 0$. *Then for any finite* $I \subset \mathbb{N}$ *and for all* $0 < \beta < \infty$, *the family of distributions* $\mathcal{Q}_{N,\beta}[\omega] \circ \Pi_I^{-1}$ *satisfies a*

LDP for almost all $\omega \in \Omega$ with rate function F_β^I given by

$$F_\beta^I(\tilde{m}) = -\sup_{p\in\mathbb{N}} \sup_{\substack{y\in[-1,1]^{2^p}\\ \Pi_I n_p(y)=\tilde{m}}} \left[\tfrac{1}{2}\|n_p(y)\|_2^2 - \beta^{-1}2^{-p}\sum_{\gamma=1}^{2^p} I(y_\gamma)\right] + \sup_{y\in\mathbb{R}}\left(\tfrac{1}{2}y^2 - \beta^{-1}I(y)\right), \tag{1.9}$$

where

$$I(y) \equiv \begin{cases} \dfrac{1+y}{2}\ln(1+y) + \dfrac{1-y}{2}\ln(1-y), & \text{if } |y| \le 1, \\ +\infty, & \text{otherwise}, \end{cases} \tag{1.10}$$

F_β^I is lower semicontinuous, Lipshitz continuous on the interior of $D_{|I|}$, bounded on $D_{|I|}$ and equal to $+\infty$ on $D_{|I|}^c$.

REMARK. Note that F_β^I is *not* convex in general.

To prove Theorem 1 we will define, for $\tilde{m} \in \mathbb{R}^I$,

$$F_{N,\beta,\varepsilon}^I(\tilde{m}) \equiv -\frac{1}{\beta N}\ln \mu_{N,\beta}[\omega](\|\Pi_I m_N(\sigma) - \tilde{m}\|_2 \le \varepsilon) \tag{1.11}$$

and show that (i) if $\tilde{m} \in D_{|I|}$, then

$$\lim_{\varepsilon\downarrow 0}\lim_{N\uparrow\infty} F_{N,\beta,\varepsilon}^I(\tilde{m}) = F_\beta^I(\tilde{m}) \tag{1.12}$$

almost surely and (ii) if $\tilde{m} \in D_{|I|}^c$, then

$$\lim_{\varepsilon\downarrow 0}\lim_{N\uparrow\infty} F_{N,\beta,\varepsilon}^I(\tilde{m}) = +\infty \tag{1.13}$$

almost surely.

From these two equations it follows from standard arguments (see, e.g., [8]) that for almost all ω, for all Borel sets $\mathscr{A} \subset \mathscr{B}(\mathbb{R}^I)$,

$$\begin{aligned} -\inf_{\tilde{m}\in \operatorname{int}\mathscr{A}} F_\beta^I(\tilde{m}) &\le \liminf_{N\uparrow\infty} \frac{1}{\beta N}\ln \mathscr{Q}_{N,\beta}[\omega]\circ\Pi_I^{-1}(\mathscr{A}) \\ &\le \limsup_{N\uparrow\infty}\frac{1}{\beta N}\ln \mathscr{Q}_{N,\beta}[\omega]\circ\Pi_I^{-1}(\mathscr{A}) \le -\inf_{\tilde{m}\in \operatorname{cl}\mathscr{A}} F_\beta^I(\tilde{m}), \end{aligned} \tag{1.14}$$

where $\operatorname{int}\mathscr{A}$ and $\operatorname{cl}\mathscr{A}$ denote the interior and the closure of the set $\mathscr{A}$, respectively. The properties of the rate function will be established directly from its explicit form (1.9).

An important feature is that the rate function is nonrandom. This means that under the conditions of the theorem, the thermodynamics of this disordered system is described in terms of completely deterministic potentials. From the thermodynamic point of view discussed above, this is an extremely satisfactory result. Namely, in these terms it means that although the Hamiltonian of our model is a function of an unbounded number of random macroscopic quantities, we may select any finite subset of these in which we may be interested and can be assured that there will exist, with probability 1, in the infinite volume limit, thermodynamic potentials that are functions of these variables only and which are, moreover, completely deterministic. The sole condition for this to hold is that the number of macroscopic variables goes to infinity with a sublinear rate.

In the remainder of this article we will present the proof of Theorem 1. There will be three important steps. First, we prove large deviation estimates for the mass of small balls in $\mathbb{R}^M$, using fairly standard techniques. The resulting bounds are expressed in terms of a certain random function. The crucial step is to show that in a strong sense this function is "self-averaging." The proof of this fact uses the Yurinskii [24] martingale representation and exponential estimates. These are finally combined to obtain deterministic estimates on cylinder events from which the convergence result (1.12) then follows rather easily.

2. The basic large deviation estimates. In this section we recall exponential upper and lower bounds that have already been derived in [5]. They provide the starting point of our analysis. Let us consider the quantities

$$Z_{N,\beta,\rho}[\omega](m) \equiv \mu_{N,\beta}[\omega](\|m_N(\sigma) - m\|_2 \le \rho) Z_{N,\beta}[\omega]. \tag{2.1}$$

We first prove a large deviation *upper* bound.

LEMMA 2.1.

$$\frac{1}{\beta N} \ln Z_{N,\beta,\rho}(m) \le \Phi_{N,\beta}(m) + \rho\left(\|t^*\|_2 + \|m\|_2 + \frac{\rho}{2}\right), \tag{2.2}$$

where

$$\Phi_{N,\beta}(m) \equiv \inf_{t\in\mathbb{R}^M} \Psi_{N,\beta}(m,t) \tag{2.3}$$

with

$$\Psi_{N,\beta}(m,t) \equiv -(m,t) + \frac{1}{2}\|m\|_2^2 + \frac{1}{\beta N}\sum_{i=1}^{N} \ln\cosh\beta(\xi_i, t) \tag{2.4}$$

and $t^* \equiv t^*(m)$ *is defined through* $\Psi_{N,\beta}(m, t^*(m)) = \inf_{t\in\mathbb{R}^M} \Psi_{N,\beta}(m,t)$, *if such a* t^* *exists, while otherwise* $\|t^*\| \equiv \infty$.

PROOF. Note that for arbitrary $t \in \mathbb{R}^M$,

$$\mathbb{1}_{\{\|m_N(\sigma)-m\|_2\le\rho\}} \le \mathbb{1}_{\{\|m_N(\sigma)-m\|_2\le\rho\}} \exp(\beta N(t, (m_N(\sigma) - m)) + \rho\beta N\|t\|_2). \tag{2.5}$$

Thus

$$
\begin{aligned}
Z_{N,\beta,\rho}(m) &= \mathbb{E}_\sigma \exp\left(\frac{\beta N}{2}\|m_N(\sigma)\|_2^2\right)\mathbb{1}_{\{\|m_N(\sigma)-m\|_2\le\rho\}}\\
&\le \inf_{t\in\mathbb{R}^M}\mathbb{E}_\sigma\exp\left(\beta N\frac{1}{2}(\|m\|_2^2+2\rho\|m\|_2+\rho^2)\right)\\
&\qquad\times\exp(\beta N(t,(m_N(\sigma)-m))+\beta N\rho\|t\|_2)\\
&\le \inf_{t\in\mathbb{R}^M}\exp\left(\beta N\left[\frac{1}{2}\|m\|_2^2-(m,t)+\frac{1}{\beta N}\sum_{i=1}^N\ln\cosh(\beta(\xi_i,t))\right]\right)\\
&\qquad\times\exp\left(\beta N\rho\left(\|m\|_2+\|t\|_2+\frac{\rho}{2}\right)\right).
\end{aligned}
\tag{2.6}
$$

This gives immediately the bound of Lemma 2.1. □

REMARK. Note that if a finite $t^*(m)$ exists, then it is the solution of the system of equations

$$
m^\mu = \frac{1}{N}\sum_{i=1}^N \xi_i^\mu \tanh\beta(\xi_i,t). \tag{2.7}
$$

Next we prove a corresponding *lower* bound.

LEMMA 2.2. *For $\rho\ge\sqrt{2(M/N)}$, we have that*

$$
\frac{1}{\beta N}\ln Z_{N,\beta,\rho}(m)\ge\Phi_{N,\beta}(m)-\rho\left(\|m\|_2+\|t^*(m)\|_2-\frac{\rho}{2}\right)-\frac{\ln 2}{\beta N}, \tag{2.8}
$$

where the notations are the same as in Lemma 2.1.

PROOF. The technique to prove this bound is the standard one to prove a Cramér-type lower bound (see, e.g., [23]). It is of course enough to consider the case where $\|t^*\|_2<\infty$. We define, for $t^*\in\mathbb{R}^M$, the probability measures $\tilde{\mathbb{P}}$ on $\{-1,1\}^N$ through their expectation $\tilde{\mathbb{E}}_\sigma$, given by

$$
\tilde{\mathbb{E}}_\sigma(\cdot)\equiv\frac{\mathbb{E}_\sigma\exp(\beta N(t^*,m_N(\sigma)))(\cdot)}{\mathbb{E}_\sigma\exp(\beta N(t^*,m_N(\sigma)))}. \tag{2.9}
$$

We have obviously that

$$
\begin{aligned}
Z_{N,\beta,\rho}(m) &= \tilde{\mathbb{E}}_\sigma \exp\left(\frac{\beta N}{2}\|m_N(\sigma)\|_2^2 - \beta N(t^*, m_N(\sigma))\right)\mathbb{1}_{\{\|m_N(\sigma)-m\|_2\le\rho\}} \\
&\quad \times \mathbb{E}_\sigma \exp(\beta N(t^*, m_N(\sigma))) \\
&\ge \exp\left(-\beta N(t^*, m) - \beta N\left(\rho\|t^*\|_2 - \frac{1}{2}\|m\|_2^2 + \rho\|m\|_2 - \frac{\rho^2}{2}\right)\right) \\
&\quad \times \mathbb{E}_\sigma \exp(\beta N(t^*, m_N(\sigma)))\tilde{\mathbb{E}}_\sigma \mathbb{1}_{\{\|m_N(\sigma)-m\|_2\le\rho\}} \\
&= \exp\left(\beta N\left(\frac{1}{2}\|m\|_2^2 - (t^*, m) + \frac{1}{\beta N}\sum_{i=1}^N \ln\cosh\beta(\xi_i, t^*)\right)\right) \\
&\quad \times \exp\left(-\beta N\rho\left(\|t^*\|_2 + \|m\|_2 - \frac{\rho}{2}\right)\right)\tilde{\mathbb{P}}_\sigma[\|m_N(\sigma) - m\|_2 \le \rho],
\end{aligned}
\tag{2.10}
$$

but, using Chebychev's inequality, we have that

$$
\begin{aligned}
\tilde{\mathbb{P}}_\sigma[\|m_N(\sigma) - m\|_2 \le \rho] &= 1 - \tilde{\mathbb{P}}_\sigma[\|m_N(\sigma) - m\|_2 \ge \rho] \\
&\ge 1 - \frac{1}{\rho^2}\tilde{\mathbb{E}}_\sigma\|m_N(\sigma) - m\|_2^2.
\end{aligned}
\tag{2.11}
$$

We choose $t^*(m)$ that satisfies (2.7). Then it is easy to compute

$$
\tilde{\mathbb{E}}\|m_N(\sigma) - m\|_2^2 = \frac{M}{N}\left(1 - \frac{1}{N}\sum_{i=1}^N \tanh^2(\beta(\xi_i, t^*(m)))\right),
\tag{2.12}
$$

from which the lemma follows. □

In the following lemma we collect a few properties of $\Phi_{N,\beta}(m)$ that arise from convexity. We set $\Gamma \equiv \{m \in \mathbb{R}^M \mid \|t^*(m)\|_2 < \infty\}$, where $t^*(m)$ is defined in Lemma 2.1, $D \equiv \{m \in \mathbb{R}^M \mid \Phi_{N,\beta}(m) > -\infty\}$, and we denote by $\mathrm{ri}\, D$ the relative interior of D (see, e.g., [21], page 44). We moreover denote by $I(x) \equiv \sup_{t\in\mathbb{R}}(tx - \ln\cosh t)$ the Legendre transform of the function $\ln\cosh t$. A simple computation shows that $I(x)$ coincides with the function defined in (1.10).

LEMMA 2.3. (i)

$$
\Phi_{N,\beta}(m) = \frac{1}{2}\|m\|_2^2 - \inf_{y\in\mathbb{R}^N:\, m_N(y)=m} \frac{1}{\beta N}\sum_{i=1}^N I(y_i),
\tag{2.13}
$$

where for each $m \in \mathbb{R}^M$ *the infimum is attained or is* $+\infty$ *vacuously.*

(ii)

$$
D = \{m \in \mathbb{R}^M \mid \exists y \in [-1, 1]^N \text{ s.t. } m_N(y) = m\}.
\tag{2.14}
$$

(iii) $\Phi_{N,\beta}(m)$ *is continuous relative to* $\mathrm{ri}\, D$.

(iv) $\Gamma = \mathrm{int}\, D$ *if* $\det(\xi^t\xi/N) \neq 0$.

(v) *If* t^* *is defined as in Lemma* 2.1 *and* y^* *realizes the infimum in* (2.13), *then*

$$\beta^2\left(t^*, \frac{\xi^t\xi}{N}t^*\right) = \frac{1}{N}\sum_{i=1}^{N}[I'(y_i^*)]^2. \tag{2.15}$$

REMARK. Note that Lemma 2.3(i) provides an alternative formula for the variational formula (2.3).

REMARK. Under the condition $\det(\xi^t\xi/N) \neq 0$, the relative interior in (iii) can be replaced by the interior. In the situation where we want to apply the lemma, this condition is satisfied with probability greater than $1 - \exp(-cN^{1/6})$.

PROOF OF LEMMA 2.3. Note that the function

$$g(t) \equiv 1/(\beta N)) \sum_{i=1}^{N} \ln\cosh\beta(\xi_i, t)$$

is a proper convex function on $\mathbb{R}^M$. Denoting by $h(m) \equiv \sup_{t\in\mathbb{R}^M}\{(m,t) - g(t)\}$ its Legendre transform, it follows from standard results of convex analysis (cf. [21], page 142, Theorem 16.3 and, in particular, the second illustration of that theorem on page 144) that $h(m)$ is a proper convex function on $\mathbb{R}^M$ and that

$$h(m) = \inf_{y\in\mathbb{R}^N:\, m_N(y)=m} \frac{1}{\beta N}\sum_{i=1}^{N} I(y_i), \tag{2.16}$$

where for each $m \in \mathbb{R}^M$ the infimum is either attained or is $+\infty$. This immediately yields (i). Denoting by $\operatorname{dom} h \equiv \{x \in \mathbb{R}^M \mid h(m) < +\infty\}$ the effective domain of h, we have, by (1.10), that $\operatorname{dom} h$ equals the right-hand side of (2.14), and since $\|m\|_2^2 \geq 0$, (ii) is proven. Part (iii) simply follows from the fact that h being convex, it is continuous relative to $\operatorname{ri}(\operatorname{dom} h)$ ([21], page 82, Theorem 10.1). Finally, to prove (iv), note first that the condition $\det(\xi^t\xi/N) \neq 0$ implies that $\operatorname{int} D \neq \varnothing$. Thus we can make use of the following two results of convex analysis ([21], page 218, Theorem 23.5): First, the subgradient of h at m, $\partial h(m)$, is a nonempty set if and only if m belongs to the interior of $\operatorname{dom} h$, that is, $m \in \operatorname{int} D$. Moreover, $\partial h(m)$ is a bounded convex set. Next, $(m,t) - g(t)$ achieves its supremum at $t^* \equiv t^*(m)$ if and only if $t^* \in \partial h(m)$. To prove (v) we only have to consider the case where t^* exists and, consequently, $|y_i^*| < 1$ for all i. To prove (2.15), introduce Lagrange multipliers $t \in \mathbb{R}^M$ for the constraint variational problem in (2.13). The corresponding Euler equations are then

$$\begin{aligned} &\frac{1}{\beta}I'(y_i) = (\xi_i, t), \qquad i = 1, \dots, N, \\ &m_N^\mu(y) = m^\mu, \qquad \mu = 1, \dots, M. \end{aligned} \tag{2.17}$$

Using the fact that $I'(x) = \tanh^{-1}(x)$, one sees that the t^* that solves these equations is identical to the solution of (2.7); from this, formula (2.15) follows immediately. This concludes the proof of the lemma. □

We see that as long as ρ can be chosen as a function of N that tends to zero as N goes to infinity, Lemmas 2.1 and 2.2 seem to provide asymptotically coinciding upper and lower bounds, at least for such m for which $t^*(m)$ is bounded. The unpleasant feature in these bounds is that $\Psi_{N,\beta}$ is a rather complicated random function and that the $\Phi_{N,\beta}$ is defined through an infimum of such a function. In the next section we analyse this problem and show that this function is essentially nonrandom.

3. Self-averaging. We show now that the random upper and lower bounds derived in the last section are actually, with large probability, independent of the realization of the randomness. In fact we will prove this under the restriction that m should be such that, at least on a subspace of full measure, $t^*(m)$ has a uniformly bounded l_2-norm. With this in mind, the result will follow from the next proposition. In the sequel, let $\Omega_1 \subset \Omega$ denote the subspace for which $\|\xi^t[\omega]\xi[\omega]/N\| = \|\xi[\omega]\xi^t[\omega]/N\| \le (1+\sqrt{\alpha})^2(1+\varepsilon)$ holds for some fixed small ε ($\varepsilon = 1$ will be a suitable choice). Recall that $\alpha \equiv M/N$. By Theorem 2.4 of [4] (see also [3] and [22]), $\mathbb{P}[\Omega_1] \ge 1 - 4N\exp(-\varepsilon N^{1/6})$.

PROPOSITION 3.1. *For any $R < \infty$ there exists $0 < \delta < 1/2$ and a set $\Omega_2 \subset \Omega$ with $\mathbb{P}[\Omega_2] \ge 1 - \exp(-N\alpha^{1-2\delta}/R)$, such that for all $\omega \in \Omega_1 \cap \Omega_2$,*

$$\sup_{t:\,\|t\|_2 \le R} \left|\Psi[\omega](m,t) - \mathbb{E}\Psi(m,t)\right| \le \alpha^{1/2-\delta}(6 + 2\|m\|_2). \tag{3.1}$$

REMARK. The subspace Ω_2 does *not* depend on m.

Note that an immediate corollary to Proposition 3.1 is that, under its assumptions,

$$\left|\inf_{t:\,\|t\|_2 \le R} \Psi[\omega](m,t) - \inf_{t:\,\|t\|_2 \le R} \mathbb{E}\Psi(m,t)\right| \le \alpha^{1/2-\delta}(6 + 2\|m\|_2). \tag{3.2}$$

REMARK. An obvious consequence of (3.2) is the observation that if $m \in \mathbb{R}^M$ and $\omega \in \Omega_1 \cap \Omega_2$ are such that

$$\inf_{t\in\mathbb{R}^M} \Psi[\omega](m,t) = \inf_{t:\,\|t\|_2 \le R} \Psi[\omega](m,t) \tag{3.3}$$

and

$$\inf_{t\in\mathbb{R}^M} \mathbb{E}\Psi(m,t) = \inf_{t:\,\|t\|_2 \le R} \mathbb{E}\Psi[\omega](m,t), \tag{3.4}$$

then

$$\left|\Phi[\omega](m) - \inf_t \mathbb{E}\Psi(m,t)\right| \le c\alpha^{1/2-\delta}. \tag{3.5}$$

PROOF OF PROPOSITION 3.1. The proof of the proposition follows from the fact that for bounded values of t, $\Psi(m,t)$ differs uniformly only little from its expectation. This will be proven by first controlling a lattice supremum and then using some a priori Lipshitz bound on $\Psi(m,t)$. We prove the Lipshitz bound first.

LEMMA 3.2. *Assume that* $\omega \in \Omega_1$. *Then*

$$|\Psi[\omega](m,t) - \Psi[\omega](m,s)| \le ((1+\sqrt{\alpha})(1+\varepsilon) + \|m\|_2)\|t-s\|_2. \tag{3.6}$$

PROOF. Note that

$$\begin{aligned}|\Psi(m,t) - \Psi(m,s)| & \\ &\le \left| -(m, t-s) + \frac{1}{\beta N}\sum_i [\ln\cosh(\beta(\xi_i,t)) - \ln\cosh(\beta(\xi_i,s))]\right| \\ &\le \|m\|_2\|t-s\|_2 + \left|\frac{1}{\beta N}\sum_i [\ln\cosh(\beta(\xi_i,t)) - \ln\cosh(\beta(\xi_i,s))]\right|.\end{aligned} \tag{3.7}$$

On the other hand, by the mean-value theorem, there exists $\tilde{t}$ such that

$$\begin{aligned}&\left|\frac{1}{\beta N}\sum_i [\ln\cosh(\beta(\xi_i,t)) - \ln\cosh(\beta(\xi_i,s))]\right| \\ &\qquad = \left|\left(t-s, \frac{1}{N}\sum_i \xi_i \tanh(\beta(\xi_i,\tilde{t}))\right)\right| \\ &\qquad = \left|\frac{1}{N}\sum_i (t-s,\xi_i)\tanh(\beta(\xi_i,\tilde{t}))\right|.\end{aligned} \tag{3.8}$$

Using the Schwarz inequality, we have that

$$\begin{aligned}\left|\frac{1}{N}\sum_i (t-s,\xi_i)\tanh(\beta(\xi_i,\tilde{t}))\right| &\le \frac{1}{N}\sqrt{\sum_i (t-s,\xi_i)^2}\sqrt{\sum_i \tanh^2(\beta(\xi_i,\tilde{t}))} \\ &\le \sqrt{\left(\left(s-t, \sum_i \frac{\xi_i^t\xi_i}{N}(s-t)\right)\right)} \\ &\le \sqrt{\left\|\frac{\xi^t\xi}{N}\right\|}\|t-s\|_2,\end{aligned} \tag{3.9}$$

But this implies the lemma. □

Let us now introduce a lattice $\mathcal{W}_{N,M}$ with spacing $1/\sqrt{N}$ in $\mathbb{R}^M$. We also denote by $\mathcal{W}_{N,M}(R)$ the intersection of this lattice with the ball of radius R. The point is that first, for any $t \in \mathbb{R}^M$, there exists a lattice point $s \in \mathcal{W}_{N,M}$ such that $\|s-t\|_2 \le \sqrt{\alpha}$, while on the other hand,

$$|\mathcal{W}_{N,M}(R)| \le \exp(\alpha N(\ln(R/\alpha))). \tag{3.10}$$

LEMMA 3.3.

$$\mathbb{P}\Big[\sup_{t\in\mathscr{W}_{N,M}(R)} |\Psi(m,t)-\mathbb{E}\Psi(m,t)| > x\Big] \le \exp\Big(-N\Big(\frac{x^2}{R}\Big(1-\frac{1}{2}e^{x/R}\Big)-\alpha\ln\Big(\frac{R}{\alpha}\Big)\Big)\Big). \tag{3.11}$$

PROOF. Clearly we only have to prove that for all $t \in \mathscr{W}_{N,M}(R)$,

$$\mathbb{P}[|\Psi(m,t)-\mathbb{E}\Psi(m,t)| > x] \le \exp\Big(-N\frac{x^2}{R}\Big(1-\frac{1}{2}e^{x/R}\Big)\Big). \tag{3.12}$$

To do this we write $\Psi(m,t)-\mathbb{E}\Psi(m,t)$ as a sum of martingale differences and use an exponential Markov inequality for martingales. Note first that

$$\Psi(m,t)-\mathbb{E}\Psi(m,t) = \frac{1}{\beta N}\sum_{i=1}^{N}(\ln\cosh(\beta(\xi_i,t)) - \mathbb{E}\ln\cosh(\beta(\xi_i,t))). \tag{3.13}$$

We introduce the decreasing sequence of sigma algebras $\mathscr{F}_{k,\kappa}$ that are generated by the random variables $\{\xi_i^\mu\}_{i\ge k+1}^{1\le\mu\le M} \cup \{\xi_k^\mu\}^{\mu\ge\kappa}$. We set

$$\tilde f_N^{(k,\kappa)} \equiv \mathbb{E}\Big[\beta^{-1}\sum_i \ln\cosh(\beta(\xi_i,t))\big|\mathscr{F}_{k,\kappa}\Big] - \mathbb{E}\Big[\beta^{-1}\sum_i \ln\cosh(\beta(\xi_i,t))\big|\mathscr{F}_{k,\kappa}^{+}\Big], \tag{3.14}$$

where for notational convenience we have set

$$\mathscr{F}_{k,\kappa}^{+} = \begin{cases} \mathscr{F}_{k,\kappa+1}, & \text{if } \kappa < M,\\ \mathscr{F}_{k+1,1}, & \text{if } \kappa = M.\end{cases} \tag{3.15}$$

Notice that we have the identity

$$\Psi(m,t)-\mathbb{E}\Psi(m,t) \equiv \frac{1}{N}\sum_{k=1}^{N}\sum_{\kappa=1}^{M}\tilde f_N^{(k,\kappa)}. \tag{3.16}$$

Our aim is to use an exponential Markov inequality for martingales. This requires, in particular, bounds on the conditional Laplace transforms of the martingale differences (see, e.g., [15], Chapter 1.3, Lemma 1.5). Namely,

$$\begin{aligned}\mathbb{P}\Big[\Big|\sum_{k=1}^{N}\sum_{\kappa=1}^{M}\tilde f_N^{(k,\kappa)}\Big| \ge Nx\Big] &\le 2\inf_{u\in\mathbb{R}}\exp(-|u|Nx)\mathbb{E}\exp\Big\{u\sum_{k=1}^{N}\sum_{\kappa=1}^{M}\tilde f_N^{(k,\kappa)}\Big\}\\ &= 2\inf_{u\in\mathbb{R}}\exp(-|u|Nx)\mathbb{E}\Big[\mathbb{E}\Big[\dots\mathbb{E}\Big[\exp\Big(u\tilde f_N^{(1,1)}\Big)\big|\mathscr{F}_{1,1}^{+}\Big]\\ &\quad\times\exp\Big(u\tilde f_N^{(1,2)}\Big)\big|\mathscr{F}_{1,2}^{+}\Big]\dots\exp\Big(u\tilde f_N^{(N,M)}\Big)\big|\mathscr{F}_{N,M}^{+}\Big],\end{aligned} \tag{3.17}$$

where the first inequality is nothing but the exponential Markov inequality. Now notice that

$$
\begin{aligned}
\tilde{f}_N^{(k,\kappa)} &= \mathbb{E}\Big[\beta^{-1}\sum_i \ln\cosh(\beta(\xi_i, t))|\mathcal{F}_{k,\kappa}\Big] - \mathbb{E}\Big[\beta^{-1}\sum_i \ln\cosh(\beta(\xi_i, t))|\mathcal{F}^+_{k,\kappa}\Big] \\
&= \mathbb{E}\big[\beta^{-1}\ln\cosh(\beta(\xi_k, t))|\mathcal{F}_{k,\kappa}\big] - \mathbb{E}\big[\beta^{-1}\ln\cosh(\beta(\xi_k, t))|\mathcal{F}^+_{k,\kappa}m\big] \\
&= \mathbb{E}\Big[\beta^{-1}\ln\cosh\Big(\beta\Big(\sum_{\mu\neq\kappa}\xi_k^\mu t_\mu + \xi_k^\kappa t_\kappa\Big)\Big)\Big|\mathcal{F}_{k,\kappa}\Big] \\
&\qquad - \mathbb{E}\Big[\beta^{-1}\ln\cosh\Big(\beta\Big(\sum_{\mu\neq\kappa}\xi_k^\mu t_\mu + \xi_k^\kappa t_\kappa\Big)\Big)\Big|\mathcal{F}^+_{k,\kappa}\Big] \qquad (3.18)\\
&= \tfrac{1}{2}\beta^{-1}\mathbb{E}\Big[\ln\cosh\Big(\beta\Big(\sum_{\mu\neq\kappa}\xi_k^\mu t_\mu + \xi_k^\kappa t_\kappa\Big)\Big) \\
&\qquad\qquad - \ln\cosh\Big(\beta\Big(\sum_{\mu\neq\kappa}\xi_k^\mu t_\mu - \xi_k^\kappa t_\kappa\Big)\Big)\Big|\mathcal{F}_{k,\kappa}\Big].
\end{aligned}
$$

Now we use the fact that

$$
\frac{\cosh(a+b)}{\cosh(a-b)} = \frac{1+\tanh a\tanh b}{1-\tanh a\tanh b} \leq \frac{1+\tanh|b|}{1-\tanh|b|} \leq e^{2|b|} \qquad (3.19)
$$

to deduce from (3.18) that

$$
|\tilde{f}_N^{(k,\kappa)}| \leq |t_\kappa|. \qquad (3.20)
$$

Using the standard inequalities $e^x \leq 1 + x + (x^2/2)e^{|x|}$ and $1 + y \leq e^y$, we get, therefore,

$$
\mathbb{E}\Big[\exp\Big(u\tilde{f}_N^{(k,\kappa)}(\tilde{m})\Big)\Big|\mathcal{F}^+_{k,\kappa}\Big] \leq \exp\Big(\frac{u^2}{2}t_\kappa^2\exp(|u||t_\kappa|)\Big). \qquad (3.21)
$$

From this and (3.17), we get now

$$
\begin{aligned}
&\mathbb{P}[|\Psi(m,t) - \mathbb{E}\Psi(m,t)| > x] \\
&\quad \leq 2\inf_u \exp\Big(-uNx + \frac{u^2}{2}N\|t\|_2^2\exp(|u|\|t\|_\infty)\Big) \\
&\quad \leq \begin{cases} 2\exp\Big(-N\dfrac{x^2}{\|t\|_2^2}\Big(1-\dfrac{1}{2}\exp(x/\|t\|_2)\Big)\Big), & \text{if } \|t\|_2 \geq 1, \\ 2\exp\Big(-Nx^2\Big(1-\dfrac{1}{2}e^x\Big)\Big), & \text{if } \|t\|_2 < 1, \end{cases}
\end{aligned} \qquad (3.22)
$$

where the last inequality is obtained by choosing $u = x/\|t\|_2^2$ in the first and $u = x/\|t\|_2$ in the second case. This gives the lemma. □

We can now continue the proof of Proposition 3.1. Choose $0 < \delta < 1/2$ and define Ω_2 to be the set of $\omega \in \Omega$ for which

$$\sup_{t \in \mathscr{W}_{N,M}(R)} |\Psi(m,t) - \mathbb{E}\Psi(m,t)| \le \alpha^{1/2-\delta}. \tag{3.23}$$

By Lemma 3.3,

$$\begin{aligned} \mathbb{P}[\Omega_2] &\ge 1 - \exp\biggl(-N\frac{\alpha^{1-2\delta}}{R}\biggl(1 - \frac{1}{2}\exp\biggl(\frac{\alpha^{1/2-\delta}}{R}\biggr)\biggr) + N\alpha\ln\biggl(\frac{R}{\alpha}\biggr)\biggr) \\ &= 1 - \exp\biggl(-NO\biggl(\frac{\alpha^{1-2\delta}}{R}\biggr)\biggr). \end{aligned} \tag{3.24}$$

Combining Lemma 3.2 with (3.23) and taking into account the remark preceding Lemma 3.3, we see that on $\Omega_1 \cap \Omega_2$,

$$\begin{aligned} \sup_{t: \|t\|_2 \le R} |\Psi(m,t) - \mathbb{E}\Psi(m,t)| &\le \alpha^{1/2-\delta} + 2\sqrt{\alpha}(\|m\|_2 + (1+\sqrt{\alpha})(1+\varepsilon)) \\ &\le \alpha^{1/2-\delta}(6 + \|m\|_2) \end{aligned} \tag{3.25}$$

for α small, which proves Proposition 3.1. □

4. Proof of Theorem 1. The results of Sections 2.1 and 3.1 can now be combined to get a large deviation principle in the product topology. The point here is that, apart from the possibility that $t^*(m)$ may become unbounded, the estimates in Lemmas 2.1 and 2.2 together with Proposition 3.1 tell us that, up to corrections that tend to zero with N, the quantity $(\beta N)^{-1} \ln Z_{N,\beta,\rho}(m)$ is given by the infimum over t of the completely nonrandom function $\mathbb{E}\Psi_{N,\beta}(m,t)$. We will first prove that, for all $\tilde{m} \in D_{|I|}$, (1.12) holds. The main step in the proof of this fact is the following theorem.

THEOREM 4.1. *Assume that* $\limsup_{N\uparrow\infty}(M(N)/N) = 0$ *and that* $0 < \beta < \infty$. *Then there exists a set* $\tilde{\Omega} \subset \Omega$ *with* $\mathbb{P}[\tilde{\Omega}] = 1$ *such that, for all finite subsets* $I \subset \mathbb{N}$ *and for all* $\tilde{m} \in [-1,1]^I$ *such that, for all* $\varepsilon > 0$, *there exists* $c = c(\tilde{m}, \varepsilon) < \infty$, $\exists N_0 \le \infty$, $\forall N \ge N_0$,

$$\sup_{m: \|\Pi_I m - \tilde{m}\|_2 \le \varepsilon} \inf_{t \in \mathbb{R}^M} \mathbb{E}\Psi_{N,\beta}(m,t) = \sup_{m: \|\Pi_I m - \tilde{m}\|_2 \le \varepsilon} \inf_{t \in \mathbb{R}^M: \|t\|_2 \le c} \mathbb{E}\Psi_{N,\beta}(m,t), \tag{4.1}$$

it holds that, for all $\omega \in \tilde{\Omega}$,

$$\begin{aligned} \lim_{\varepsilon\downarrow 0} \lim_{N\uparrow\infty} F^I_{N,\beta,\varepsilon}[\omega](\tilde{m}) = &-\sup_{p\in\mathbb{N}} \sup_{\substack{y \in [-1,1]^{2^p} \\ \Pi_I n_p(y) = \tilde{m}}} \biggl[\tfrac{1}{2}\|n_p(y)\|_2^2 - \beta^{-1} 2^{-p} \sum_{\gamma=1}^{2^p} I(y_\gamma)\biggr] \\ &+ \sup_{y\in\mathbb{R}} \bigl(\tfrac{1}{2}y^2 - \beta^{-1} I(y)\bigr). \end{aligned} \tag{4.2}$$

REMARK. The assumption in Theorem 4.1 looks horrible at first glance. The reader will observe that it is made in order to allow us to apply the self-averaging results from the last section. We will show later, however, that the set of values $\tilde{m}$ for which it is satisfied can be constructed explicitly and is nothing other than $D_{|I|}$.

PROOF OF THEOREM 4.1. We will first establish an upper bound for the quantity

$$Z^I_{N,\beta,\varepsilon}[\omega](\tilde{m}) \equiv \mu_{N,\beta}[\omega](\|\Pi_I m_N(\sigma) - \tilde{m}\|_2 \le \varepsilon) Z_{N,\beta}[\omega]. \tag{4.3}$$

To do so, notice that on Ω_1, $\|m_N(\sigma)\|_2 \le (1+\sqrt{\alpha})\sqrt{(1+\varepsilon)} < 2$ for all σ. We may cover the ball of radius 2 with balls of radius $\rho > \sqrt{\alpha}$, centered at the lattice points in $\mathscr{W}_{N,M}(2)$. We then have that on Ω_1,

$$\begin{aligned} Z^I_{N,\beta,\varepsilon}[\omega](\tilde{m}) &\le \sum_{\substack{m\in\mathscr{W}_{N,M}(2)\\ \|\Pi_I m-\tilde{m}\|_2\le\varepsilon}} Z_{N,\beta,\rho}[\omega](m) \\ &\le \sup_{\substack{m\in\mathscr{W}_{N,M}(2)\\ \|\Pi_I m-\tilde{m}\|_2\le\varepsilon}} Z_{N,\beta,\rho}[\omega](m) \sum_{\substack{m\in\mathscr{W}_{N,M}(2)\\ \|\Pi_I m-\tilde{m}\|_2\le\varepsilon}} 1 \\ &\le \sup_{\substack{m:\|m\|_2<2\\ \|\Pi_I m-\tilde{m}\|_2\le\varepsilon}} Z_{N,\beta,\rho}[\omega](m)\exp(\alpha N(\ln 2/\alpha)). \end{aligned} \tag{4.4}$$

As long as $\alpha \downarrow 0$, the factor $\exp(\alpha N(\ln 2/\alpha))$ in the upper bound is irrelevant for the exponential asymptotic, as is the difference between ε and $\varepsilon-\rho$. Using the estimates used in the proof of Lemma 2.1, we can replace $Z_{N,\beta,\rho}[\omega](m)$ in (4.4) by its upper bound in terms of the function Ψ. Namely,

$$\begin{aligned} &\frac{1}{\beta N}\ln Z^I_{N,\beta,\varepsilon}[\omega](\tilde{m}) \\ &\quad \le \sup_{\substack{m:\|m\|_2<2\\ \|\Pi_I m-\tilde{m}\|_2\le\varepsilon}} \inf_{\substack{t\in\mathbb{R}^M\\ \|t\|_2\le c}} \Psi_N[\omega](m,t) + \rho\left(c+2+\frac{\rho}{2}\right) + \beta^{-1}\alpha\ln\frac{2}{\alpha}. \end{aligned} \tag{4.5}$$

Finally, combining (4.5) with (3.2), we get that, for $\omega \in \Omega_1 \cap \Omega_2$ and for any c,

$$\begin{aligned} \frac{1}{\beta N}\ln Z^I_{N,\beta,\varepsilon}[\omega](\tilde{m}) \le \sup_{m:\|\Pi_I m-\tilde{m}\|_2\le\varepsilon} \inf_{\substack{t\in\mathbb{R}^M\\ \|t\|_2\le c}} \mathbb{E}\Psi_N(m,t) + 10\alpha^{1/2-\delta} \\ + \rho\left(\frac{c+2+\rho}{2}\right) + \beta^{-1}\alpha\ln\frac{2}{\alpha}. \end{aligned} \tag{4.6}$$

By assumption, there exists a value $c < \infty$, such that the true minimax over $\mathbb{E}\Psi_N(m,t)$ is taken for a value of t with norm bounded uniformly in N by some constant c. The constant c in (4.6) is then chosen as this same constant, and then the restriction $\|t\|_2 \le c$ is actually void, and the minimax is taken for some values (m^*, t^*) which depend only on $\tilde{m}$ and ε. This is already essentially the desired form of the upper bound.

We now turn to the more subtle problem of obtaining the corresponding form of the lower bound. Trivially,

$$Z^{I}_{N,\beta,\varepsilon+\rho}[\omega](\tilde{m}) \geq Z_{N,\beta,\rho}[\omega](m^*). \tag{4.7}$$

We will modify slightly the derivation of the lower bound for $Z_{N,\beta,\rho}[\omega](m^*)$. Namely, instead of defining the shifted measure $\tilde{\mathbb{P}}$ with respect to the random value of t that realizes the infimum of $\Psi_N[\omega](m^*, t)$, we do this with the deterministic value t^* that realizes the infimum of $\mathbb{E}\Psi_N(m^*, t)$. This changes nothing in the computations in (2.10) and (2.11). What changes, however, is the estimate on $\tilde{\mathbb{E}}_\sigma\|m_N(\sigma)-m^*\|_2^2$, since t^* does not satisfy (2.7), but is instead a solution of the equations

$$m^*_\mu = \mathbb{E}\xi_1^\mu \tanh(\beta(\xi_1, t^*)). \tag{4.8}$$

Thus, in place of (2.12) we get

$$\begin{aligned}
&\tilde{\mathbb{E}}_\sigma\|m_N(\sigma) - m^*\|_2^2 \\
&\quad = \mathbb{E}_\sigma \prod_{i=1}^N \exp(\beta(t^*, \xi_i\sigma_i)) \\
&\qquad \times \sum_\nu \biggl(N^{-2}\sum_{j,k}\xi_j^\nu\xi_k^\nu\sigma_j\sigma_k - 2m^*_\nu N^{-1}\sum_j \xi_j^\nu\sigma_j + (m^*_\nu)^2\biggr) \\
&\qquad \times \biggl[\prod_{i=1}^N \cosh\beta(\xi_i, t^*)\biggr]^{-1} \\
&\quad = \frac{1}{N^2}\sum_\nu\sum_j 1 + \frac{1}{N^2}\sum_\nu\sum_{j\neq k}\tanh(\beta(t^*, \xi_j))\tanh(\beta(t^*, \xi_k))\xi_j^\nu\xi_k^\nu \\
&\qquad - 2\frac{1}{N}\sum_j\sum_\nu m^*_\nu \tanh(\beta(t^*, \xi_j))\xi_j^\nu + \sum_\nu (m^*_\nu)^2 \\
&\quad = \frac{M}{N}\biggl(1 - \frac{1}{N}\sum_i \tanh^2(\beta(t^*, \xi_i))\biggr) \\
&\qquad + \sum_\nu\biggl(\frac{1}{N}\sum_i \xi_i^\nu \tanh(\beta(t^*, \xi_i)) - m^*_\nu\biggr)^2.
\end{aligned} \tag{4.9}$$

The first summand in (4.9) is bounded by α and we have to control the second. To do so we use (4.8) to write

$$\begin{aligned}
&\sum_\nu\biggl(\frac{1}{N}\sum_i \xi_i^\nu \tanh(\beta(t^*, \xi_i)) - m^*_\nu\biggr)^2 \\
&\quad = \sum_\nu\biggl(\frac{1}{N}\sum_i \xi_i^\nu \tanh(\beta(t^*, \xi_i)) - \mathbb{E}\xi_1^\nu\tanh(\beta(\xi_1, t^*))\biggr)^2 \\
&\quad = \sum_\nu\biggl(\frac{1}{N}\sum_i \xi_i^\nu \tanh(\beta(t^*, \xi_i)) - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\nu\tanh(\beta(t^*, \xi_i))\biggr)^2 \\
&\quad \equiv G_N(t^*).
\end{aligned} \tag{4.10}$$

We will now prove, in analogy to Proposition 3.1, that $G_N(t)$ is actually small with large probability. This will be slightly more complicated than in Proposition 3.1 and will, in fact, consist of two steps. The first step is a fairly crude bound on $G_N(t)$ that will be used in a second step to obtain a refined bound.

LEMMA 4.2. *For all* $\omega \in \Omega_1$,

$$G_N[\omega](t) \leq 6. \tag{4.11}$$

PROOF. For notational simplicity, let us set $T_i \equiv \tanh(\beta(\xi_i, t))$. We have that

$$\begin{aligned} G_N(t) &\leq 2\sum_{\mu=1}^{M}\left(\left[\frac{1}{N}\sum_i \xi_i^\mu T_i\right]^2 + \left[\frac{1}{N}\sum_i \mathbb{E}\xi_i^\mu T_i\right]^2\right) \\ &= \frac{2}{N^2}\sum_{\mu=1}^{M}\sum_{i,j}(\xi_i^\mu \xi_j^\mu T_i T_j + \mathbb{E}(\xi_i^\mu T_i)\mathbb{E}(\xi_j^\mu T_j)). \end{aligned} \tag{4.12}$$

For the first term, we can use simply that

$$\frac{2}{N^2}\sum_{\mu=1}^{M}\sum_{i,j}\xi_i^\mu \xi_j^\mu T_i T_j \leq 2\left\|\frac{\xi\xi^t}{N}\right\|\left(\frac{1}{N}\sum_i T_i^2\right) \leq 2\left\|\frac{\xi\xi^t}{N}\right\|, \tag{4.13}$$

but on Ω_1, the norm in the last term is bounded by $(1+\sqrt{\alpha})^2(1+\varepsilon)$. To bound the second term in (4.12), we use the independence of both ξ_i^μ and T_i for different indices i to write

$$\begin{aligned} &\frac{2}{N^2}\sum_{\mu=1}^{M}\sum_{i,j}\mathbb{E}(\xi_i^\mu T_i)\mathbb{E}(\xi_j^\mu T_j) \\ &\quad= \frac{2}{N^2}\sum_{\mu=1}^{M}\sum_{i,j}\mathbb{E}(\xi_i^\mu T_i \xi_j^\mu T_j) + \frac{2}{N^2}\sum_{\mu=1}^{M}\sum_{i}((\mathbb{E}\xi_i^\mu T_i)^2 - \mathbb{E}(T_i)^2) \\ &\quad\leq 2\mathbb{E}\left\|\frac{\xi\xi^t}{N}\right\| + \frac{2M}{N} \\ &\quad\leq 2\alpha + 2(1+\sqrt{\alpha})^2(1+\varepsilon). \end{aligned} \tag{4.14}$$

Combining these two bounds we get (4.11). □

Lemma 4.2 tells us that $G_N(t)$ is bounded, but not yet that it is small. To do this, we observe first that its mean value is small.

LEMMA 4.3.

$$0 \leq \mathbb{E}G_N(t) \leq \alpha. \tag{4.15}$$

PROOF.

$$\begin{aligned}\mathbb{E}G_N(t) &= \sum_{\nu=1}^{M}\mathbb{E}\bigg[\frac{1}{N}\sum_i \xi_i^\nu \tanh(\beta(t,\xi_i)) - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\nu \tanh(\beta(t,\xi_i))\bigg]^2\\ &= \sum_{\nu=1}^{M}\bigg(\mathbb{E}\bigg[\frac{1}{N}\sum_i \xi_i^\nu \tanh(\beta(t,\xi_i))\bigg]^2 - \bigg[\frac{1}{N}\sum_i \mathbb{E}\xi_i^\nu \tanh(\beta(t,\xi_i))\bigg]^2\bigg)\\ &= \sum_{\nu=1}^{M}\bigg(\frac{1}{N^2}\sum_i \mathbb{E}\tanh^2(\beta(t,\xi_i)) - \frac{1}{N^2}\sum_i [\mathbb{E}\xi_i^\nu \tanh(\beta(t,\xi_i))]^2\bigg),\\ &\le \frac{M}{N},\end{aligned} \tag{4.16}$$

where we have used the independence of the summands for different indices i. □

In the sequel we will need that the mean value of $G_N(t)$ does not differ much from its conditional expectation relative to Ω_1. Namely,

$$|\mathbb{E}G_N(t) - \mathbb{E}[G_N(t)|\Omega_1]| \le 2M\exp(-\varepsilon N^{1/6}) \tag{4.17}$$

is arbitrarily small.

Finally, we will show that on Ω_1, with large probability, $G_N(t)$ differs only little from its conditional expectation relative to Ω_1.

LEMMA 4.4. *Assume that* $x \gg (\ln N)/\sqrt{N}$. *Then*

$$\mathbb{P}[|G_N(t) - \mathbb{E}[G_N(t)|\Omega_1]| \ge x\,|\Omega_1] \le \exp(-b\sqrt{N}x) \tag{4.18}$$

for some positive constant b.

PROOF. Basically the proof of this lemma relies on the same technique as that of Proposition 3.1. However, a number of details are modified. In particular, we use a coarser filtration of $\mathscr{F}$ to define our martingale differences. Namely, we denote by $\mathscr{F}_k$ the sigma algebra generated by the random variables $\{\xi_i^\mu\}_{i\ge k}^{\mu\in\mathbb{N}}$. We also introduce the trace sigma algebra $\tilde{\mathscr{F}} \equiv \mathscr{F}\cap\Omega_1$ and by $\tilde{\mathscr{F}}_k \equiv \mathscr{F}_k \cap \Omega_1$ the corresponding filtration of the trace sigma algebra. We set

$$f_N^{(k)} \equiv \mathbb{E}[G_N(t)|\tilde{\mathscr{F}}_k] - \mathbb{E}[G_N(t)|\tilde{\mathscr{F}}_{k+1}]. \tag{4.19}$$

Obviously, we have, for $\omega \in \Omega_1$,

$$G_N[\omega](t) - \mathbb{E}[G_N(t)|\Omega_1] = \sum_{k=1}^{N} f_N^{(k)}. \tag{4.20}$$

Thus the lemma will be proven if we can prove an estimate of the form (4.18) for the sum of the $f_N^{(k)}$. This goes just as in the proof of Proposition 3.1, that is, it relies on uniform bounds on the conditional Laplace transforms

$$\mathbb{E}\Big[\exp\Big(uf_N^{(k)}\Big)|\tilde{\mathscr{F}}_{k+1}\Big]. \tag{4.21}$$

The strategy to get the uniform bounds is very similar to the strategy used in [6] and [2]. We introduce

$$G_N^{(k)}(t, z) \equiv \sum_{\mu}\left(\frac{1}{N}\sum_{i\neq k}\xi_i^{\mu}T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^{\mu}T_i + \frac{z}{N}\xi_k^{\mu}T_k\right)^2 \tag{4.22}$$

and set

$$g_k(z) \equiv G_N^{(k)}(t, z) - G_N^{(k)}(t, 0). \tag{4.23}$$

We then have that

$$f_N^{(k)} = \mathbb{E}[g_k(1)|\tilde{\mathscr{F}}_k] - \mathbb{E}[g_k(1)|\tilde{\mathscr{F}}_{k+1}] \tag{4.24}$$

since $G_N^{(k)}(t, 0)$ is independent of the random variables ξ_k. On the other hand,

$$g_k(1) = \int_0^1 dz\, g_k'(z) \tag{4.25}$$

and

$$g_k'(z) = 2\sum_{\nu=1}^{M}\left[\frac{1}{N}\sum_{i\neq k}\xi_i^{\nu}T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^{\nu}T_i + \frac{z}{N}\xi_k^{\nu}T_k\right]\frac{1}{N}\xi_k^{\nu}T_k. \tag{4.26}$$

Let us first get a uniform bound on $|f_N^{(k)}|$ on Ω_1. From the formulas above it follows clearly that

$$|f_N^{(k)}| \le 2\sup_z |g_k'(z)|, \tag{4.27}$$

but, using the Schwarz inequality,

$$\begin{aligned} |g_k'(z)| &\le \frac{2}{N}\sum_{\mu}\left|\frac{1}{N}\sum_{i\neq k}\xi_i^{\nu}T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^{\nu}T_i + \frac{z}{N}\xi_k^{\mu}T_k\right| \\ &\le \frac{2}{N}\sqrt{M}\sqrt{\sum_{\mu}\left[\frac{1}{N}\sum_{i\neq k}\xi_i^{\nu}T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^{\nu}T_i + \frac{z}{N}\xi_k^{\mu}T_k\right]^2} \\ &= \frac{2\sqrt{M}}{N}\sqrt{G_N^{(k)}(t, z)}. \end{aligned} \tag{4.28}$$

However, on Ω_1 it is trivial to check that $G_N^{(k)}(t, z)$ satisfies, for $z \in [0, 1]$, the same bound as $G_N(t)$, so that on Ω_1,

$$|g_k'(z)| \le \frac{12\sqrt{M}}{N}. \tag{4.29}$$

Now we turn to the estimation of the conditional Laplace transform. Using the standard inequality

$$e^x \le 1 + x + \tfrac{1}{2}x^2 e^{|x|}, \tag{4.30}$$

we get

$$\begin{aligned}\mathbb{E}\Big[\exp\Big(uf_N^{(k)}\Big)|\tilde{\mathscr{F}}_{k+1}\Big] &\leq 1+\frac{1}{2}u^2\mathbb{E}\Big[\Big(f_N^{(k)}\Big)^2\exp\Big(|u||f_N^{(k)}|\Big)|\tilde{\mathscr{F}}_{k+1}\Big] \\ &\leq 1+\frac{1}{2}u^2\exp\Big(|u|\frac{12\sqrt{M}}{N}\Big)\mathbb{E}\Big[\Big(f_N^{(k)}\Big)^2|\tilde{\mathscr{F}}_{k+1}\Big].\end{aligned} \tag{4.31}$$

A simple computation (see [6]) shows that

$$\begin{aligned}\mathbb{E}\Big[\Big(f_N^{(k)}\Big)^2|\tilde{\mathscr{F}}_{k+1}\Big] &\leq \mathbb{E}[(g_k(1))^2|\tilde{\mathscr{F}}_{k+1}] \\ &= \mathbb{E}\bigg[\bigg(\int_0^1 dz\, g_k'(z)\bigg)^2\bigg|\tilde{\mathscr{F}}_{k+1}\bigg] \\ &\leq \mathbb{E}\bigg[\int_0^1 dz\,(g_k'(z))^2\bigg|\tilde{\mathscr{F}}_{k+1}\bigg] \\ &\leq \sup_{0\leq z\leq 1}\mathbb{E}\big[(g_k'(z))^2|\tilde{\mathscr{F}}_{k+1}\big].\end{aligned} \tag{4.32}$$

Let us write

$$\begin{aligned}g_k'(z) &= 2\sum_{\nu=1}^{M}\bigg[\frac{1}{N}\sum_i \xi_i^\nu T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\nu T_i\bigg]\frac{1}{N}\xi_k^\nu T_k \\ &\quad + 2\sum_{\nu=1}^{M}\frac{z-1}{N^2}T_k^2.\end{aligned} \tag{4.33}$$

Thus

$$\begin{aligned}(g_k'(z))^2 &\leq 8\bigg(\sum_{\nu=1}^{M}\bigg[\frac{1}{N}\sum_i \xi_i^\nu T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\nu T_i\bigg]\frac{1}{N}\xi_k^\nu T_k\bigg)^2 \\ &\quad + 8T_k^4(z-1)^2\frac{M^2}{N^4}.\end{aligned} \tag{4.34}$$

Let us abbreviate the two summands in (4.34) by (I) and (II). The term (II) is of order $\alpha^2 N^{-2}$ and thus can simply be bounded uniformly. We have to work a little more to control the conditional expectation of the first summand. We write

$$\begin{aligned}\mathbb{E}[(I)|\tilde{\mathscr{F}}_{k+1}] &= \frac{8}{N^2}\mathbb{E}\bigg[\sum_{\mu,\nu}\xi_k^\mu\xi_k^\nu T_k^2\bigg[\frac{1}{N}\sum_i \xi_i^\mu T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\mu T_i\bigg] \\ &\quad \times\bigg[\frac{1}{N}\sum_i \xi_i^\nu T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\nu T_i\bigg]\bigg|\tilde{\mathscr{F}}_{k+1}\bigg].\end{aligned} \tag{4.35}$$

We observe that under the expectation conditioned on $\tilde{\mathscr{F}}_{k+1}$ we may interchange the indices of $1\leq j\leq k$ and use this to symmetrize the expression

(4.35). [Notice that this is the reason why we separated the z-dependent contribution in (4.34).] This gives

$$\begin{aligned}\mathbb{E}[(I)|\tilde{\mathscr{F}}_{k+1}] &= \frac{8}{N^2}\mathbb{E}\bigg[\sum_{\mu,\nu}\sum_{j=1}^{k}\frac{\xi_j^\mu\xi_j^\nu}{k}T_j^2\bigg[\frac{1}{N}\sum_i \xi_i^\mu T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\mu T_i\bigg] \\ &\qquad \times \bigg[\frac{1}{N}\sum_i \xi_i^\nu T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\nu T_i\bigg]\bigg|\tilde{\mathscr{F}}_{k+1}\bigg] \\ &\le \frac{8}{N^2}\mathbb{E}\bigg[\bigg\|\sum_{j=1}^{k}\frac{\xi_j\xi_j^t}{k}T_j^2\bigg\|\sum_{\mu=1}^{M}\bigg[\frac{1}{N}\sum_i \xi_i^\mu T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\mu T_i\bigg]^2\bigg|\tilde{\mathscr{F}}_{k+1}\bigg].\end{aligned} \tag{4.36}$$

However, by Lemma 4.2, on Ω_1,

$$\sum_{\mu=1}^{M}\bigg[\frac{1}{N}\sum_i \xi_i^\mu T_i - \mathbb{E}\frac{1}{N}\sum_i \xi_i^\mu T_i\bigg]^2 = G_N(t) \le 6 \tag{4.37}$$

and since

$$\bigg\|\sum_{j=1}^{k}\frac{\xi_j\xi_j^t}{k}T_j^2\bigg\| \le \bigg\|\sum_{j=1}^{k}\frac{\xi_j\xi_j^t}{k}\bigg\| \equiv \|B(k)\|, \tag{4.38}$$

we get that

$$\mathbb{E}[(\mathrm{I})|\tilde{\mathscr{F}}_{k+1}] \le \frac{48}{N^2}\mathbb{E}[\|B(k)\|\,|\Omega_1] \le \frac{48}{N^2}\frac{\mathbb{E}\|B(k)\|}{\mathbb{P}[\Omega_1]}. \tag{4.39}$$

It is easy to show (see [2]) that

$$\mathbb{E}\|B(k)\| \le c\Big(1+\sqrt{M/k}\Big)^2 \tag{4.40}$$

for some constant $2 > c > 1$. Collecting our estimates and using that $1+x \le e^x$, we arrive at

$$\begin{aligned}&\mathbb{E}\Big[\exp\Big(uf_N^{(k)}\Big)\big|\tilde{\mathscr{F}}_{k+1}\Big] \\ &\qquad \le \exp\big(\tfrac{1}{2}u^2\exp\big(|u|12\sqrt{M}/N\big)N^{-2}\big[8\alpha^2 + 76(1+\sqrt{M/k})^2\big]\big).\end{aligned} \tag{4.41}$$

Since

$$\sum_{k=1}^{N}\big(1+\sqrt{M/k}\big)^2 = N + \sqrt{MN} + M\ln N = N\big(1+4\sqrt{\alpha}+\alpha\ln N\big), \tag{4.42}$$

this yields that

$$\begin{aligned}&\mathbb{P}\bigg[\sum_{k=1}^{N} f_N^{(k)} \ge x\big|\Omega_1\bigg] \\ &\qquad \le \inf_u \exp\bigg(-ux + \frac{u^2}{2N}\exp\bigg(\frac{|u|12\sqrt{M}}{N}\bigg) \\ &\qquad\qquad \times \big[8\alpha^2 + 76 + 304\sqrt{\alpha} + 76\alpha\ln N\big]\bigg).\end{aligned} \tag{4.43}$$

In order to perform the infimum over u in (4.43), we must distinguish two cases. First, if $\alpha \le 1/\ln N$, we may chose $u = \sqrt{N}$, which yields

$$\mathbb{P}\left[\sum_{k=1}^{N} f_N^{(k)} \ge x\right] \le \exp(-\sqrt{N}x + c_1) \tag{4.44}$$

for some positive constant c_1. If now α goes to zero with N more slowly than $1/\ln N$, a good estimate of the infimum is obtained by choosing $u = N/12\sqrt{M}$. This gives

$$\begin{aligned}\mathbb{P}\left[\sum_{k=1}^{N} f_N^{(k)} \ge x\right] &\le \exp\left(-\sqrt{N}\frac{x}{12\sqrt{\alpha}}\right)\exp\left\{\frac{e}{36}\left[\alpha + \frac{12}{\alpha} + \frac{48}{\sqrt{\alpha}} + 2\ln N\right]\right\}\\ &\le \exp\left(\frac{-\sqrt{N}x}{12 + c_2 \ln N}\right)\end{aligned} \tag{4.45}$$

for some positive constant c_2. From here the lemma follows immediately. □

COROLLARY 4.5. *There exists a set* $\Omega_3(t^*) \subset \Omega_1$ *with* $\mathbb{P}[\Omega_1 \backslash \Omega_3] \le \exp(-bN^{1/4})$ *such that, for all* $\omega \in \Omega_3(t^*)$,

$$\tilde{\mathbb{P}}_\sigma[\|m_N(\sigma) - m^*\|_2 \le [2(2\alpha + N^{-1/4})]^{1/2}] \ge \tfrac{1}{2}. \tag{4.46}$$

PROOF. This follows from combining (4.9) and (4.10) with Lemmas 4.2, 4.3 and 4.4 and choosing $x = N^{-1/4}$ in the latter. □

To be able to use Corollary 4.5, we will choose from now on $\rho > [2(2\alpha + N^{-1/4})]^{1/2}$.

Now except on a subspace Ω_4^c of probability smaller than $4Ne^{-cN^{1/6}}$ $\|\xi^t\xi/N - \mathbb{1}\| \le \sqrt{\alpha}(2+\sqrt{\alpha})(1+c)$ (see the appendix of [4]) which implies in particular that on O_4, $\det \xi^t\xi/N \ne 0$. Thus on Ω_4, Lemma 2.3(iv) implies that if $\|t^*\|_2$ is bounded, then $m^* \in \operatorname{int} D$, that is, there exists $y^* \in [-1,1]^N$ such that $m^* = m_N(y^*)$. However, $\|m_N(y^*)\|_2^2 \le \|\xi^t\xi/N\|\|y^*\|_2^2/N \le \|\xi^t\xi/N\|$. Since by assumption $\|t^*\|_2 < c$, we see that on $\Omega_1 \cap \Omega_4$, $\|m^*\|_2 \le 2$. As a consequence, putting together Proposition 3.1, Corollary 4.5 and (2.10), we find that on $\Omega_3(t^*)$,

$$\begin{aligned}&\frac{1}{\beta N}\ln Z^I_{N,\beta,\varepsilon+\rho}[\omega](\tilde{m})\\ &\qquad \ge \mathbb{E}\Psi_N(m^*, t^*) - 10\alpha^{1/2-\delta} - \rho\left(\frac{c+2-\rho}{2}\right) - \frac{\ln 2}{\beta N},\end{aligned} \tag{4.47}$$

which is the desired form of the lower bound.

Finally, by a simple Borel–Cantelli argument, it follows from the estimates on the probabilities of the sets Ω_1, Ω_2 and $\Omega_3(t^*)$ that there exits a set $\tilde{\Omega}$ of measure 1 on which

$$\limsup_{N\uparrow\infty} \frac{1}{\beta N}\ln Z^I_{N,\beta,\varepsilon}[\omega](\tilde{m}) \le \limsup_{N\uparrow\infty} \sup_{m:\|\Pi_I m - \tilde{m}\|_2 \le \varepsilon} \inf_{t\in\mathbb{R}^M} \mathbb{E}\Psi_N(m,t) \tag{4.48}$$

and

$$\liminf_{N\uparrow\infty} \frac{1}{\beta N} \ln Z^I_{N,\beta,\varepsilon}[\omega](\tilde m) \geq \liminf_{N\uparrow\infty} \sup_{m:\, \|\Pi_I m - \tilde m\|_2 \leq \varepsilon-\rho} \inf_{t\in\mathbb{R}^M} \mathbb{E}\Psi_N(m,t). \tag{4.49}$$

It remains to show that the lim sups and the lim infs on the right-hand sides of (4.48) and (4.49) coincide. From here on there is no difference to the procedure in the case $M < \ln N/\ln 2$ that was treated in [5]. We repeat the outline for the convenience of the reader. We write $\mathbb{E}\Psi_N(m,t)$ in its explicit form as

$$\mathbb{E}\Psi_N(m,t) = \tfrac{1}{2}\|m\|_2^2 - (m,t) + \beta^{-1}2^{-M}\sum_{\gamma=1}^{2^M} \ln\cosh(\beta(e_\gamma,t)), \tag{4.50}$$

where the vectors e_γ, $\gamma = 1,\dots,2^M$, form a complete enumeration of all vectors with components ± 1 in $\mathbb{R}^M$. They can be conveniently chosen as

$$e^\mu_\gamma = (-1)^{[\gamma 2^{1-\mu}]}, \tag{4.51}$$

where $[x]$ denotes the smaller integer greater than or equal to x. Note that $\mathbb{E}\Psi_N(m,t)$ depends on N only through $M(N)$. We may use Lemma 2.3 to show that

$$\inf_{t\in\mathbb{R}^M} \mathbb{E}\Psi_N(m,t) = \tfrac{1}{2}\|m\|_2^2 - \inf_{y\in\mathbb{R}^{2^M}:\, n_M(y)=m} \beta^{-1}2^{-M}\sum_{\gamma=1}^{2^M} I(y_\gamma) \tag{4.52}$$

and hence

$$\begin{aligned} &\sup_{m:\, \|\Pi_I m-\tilde m\|_2\leq\varepsilon} \inf_{t\in\mathbb{R}^M} \mathbb{E}\Psi_N(m,t) \\ &\quad = \sup_{y\in\mathbb{R}^{2^M}:\, \|\Pi_I n_M(y)-\tilde m\|_2\leq\varepsilon} \tfrac{1}{2}\|n_M(y)\|_2^2 - \beta^{-1}2^{-M}\sum_{\gamma=1}^{2^M} I(y_\gamma). \end{aligned} \tag{4.53}$$

To prove that this expression converges as N (or rather M) tends to infinity, we define, for any integers d, p with $d \leq p$, the sets

$$\mathscr{A}^p_d \equiv \left\{ y \in [-1,1]^{2^p} \mid y_\gamma = y_{\gamma+2^d} \right\}. \tag{4.54}$$

Obviously,

$$\mathscr{A}^p_0 \subset \mathscr{A}^p_1 \subset \cdots \subset \mathscr{A}^p_{p-1} \subset \mathscr{A}^p_p = [-1,1]^{2^p}. \tag{4.55}$$

The definition of these sets implies the following fact: If $y \in \mathscr{A}^p_d$ with $d < p$, then (i) $n^\nu_p(y) = 0$ if $\nu > d$ and (ii) $n^\mu_p(y) = n^\mu_d(y)$ if $\mu \leq d$.

Let us set

$$\Theta_p(y) = \tfrac{1}{2}\|n_p(y)\|_2^2 - \beta^{-1}2^{-p}\sum_{\gamma=1}^{2^p} I(y_\gamma) \tag{4.56}$$

and

$$\Upsilon_{p,\varepsilon}(\tilde m) = \sup_{\substack{y\in\mathscr{A}^p_p \\ \|\Pi_I n_p(y)-\tilde m\|_2\leq\varepsilon}} \Theta_p(y). \tag{4.57}$$

Therefore, for $y \in \mathscr{A}_d^p$, $\Theta_p(y) = \Theta_d(y)$, while at the same time the constraint in the sup is satisfied simultaneously w.r.t. n_p or n_d, as soon as d is large enough such that $I \subset \{1, \ldots, d\}$. Therefore,

$$\Upsilon_{p,\varepsilon}(\tilde{m}) \geq \sup_{\substack{y \in \mathscr{A}_d^p \\ \|\Pi_I n_p(y) - \tilde{m}\|_2 \leq \varepsilon}} \Theta_p(y) = \sup_{\substack{y \in \mathscr{A}_d^d \\ \|\Pi_I n_d(y) - \tilde{m}\|_2 \leq \varepsilon}} \Theta_d(y) = \Upsilon_{d,\varepsilon}(\tilde{m}). \tag{4.58}$$

Hence $\Upsilon_{p,\varepsilon}(\tilde{m})$ is an increasing sequence in p. Moreover, $\Upsilon_{p,\varepsilon}(\tilde{m}) \leq \sup_{y \in \mathscr{A}_p^p, \|\Pi_I n_p(y) - \tilde{m}\|_2 \leq \varepsilon} \frac{1}{2}\|n_p(y)\|_2^2 \leq 1$ and so because it is bounded from above, it converges. Thus

$$\begin{aligned} \lim_{N\uparrow\infty} \sup_{m:\|\Pi_I m - \tilde{m}\|_2 \leq \varepsilon} \inf_{t \in \mathbb{R}^M} \mathbb{E}\Psi_N(m,t) &= \lim_{N\uparrow\infty} \Upsilon_{M,\varepsilon}(\tilde{m}) \\ &= \sup_p \Upsilon_{p,\varepsilon}(\tilde{m}). \end{aligned} \tag{4.59}$$

It remains to consider the limit $\varepsilon \downarrow 0$. It is clear that $\sup_p \Upsilon_{p,\varepsilon}(\tilde{m})$ converges to a lower-semicontinuous function and that

$$\lim_{\varepsilon\downarrow 0} \sup_p \Upsilon_{p,\varepsilon}(\tilde{m}) = \lim_{\varepsilon\downarrow 0} \sup_{m:\|\Pi_I m - \tilde{m}\|_2 \leq \varepsilon} \sup_p \Upsilon_{p,0}(m). \tag{4.60}$$

Thus if $\sup_p \Upsilon_{p,0}(\tilde{m})$ is continuous in a neighborhood of $\tilde{m}$, we get

$$\lim_{\varepsilon\downarrow 0} \sup_p \Upsilon_{p,\varepsilon}(\tilde{m}) = \sup_p \Upsilon_{p,0}(\tilde{m}) \tag{4.61}$$

as desired. However, as was shown in [5], from the explicit form of Υ, one shows easily that $\sup_p \Upsilon_{p,0}(\tilde{m})$ is Lipshitz continuous in the interior of the set on which it is bounded. This proves Theorem 4.1. □

We will show next that a sufficient condition for condition (4.1) to hold is that $\tilde{m}$ belongs to $D_{|I|}$. While this appears intuitively "clear," the rigorous proof is surprisingly tedious. Let us first introduce some notation and results.

Let E_p be the $2^p \times p$ matrix whose rows are given by the vectors e_γ, $\gamma = 1, \ldots, 2^p$, which, for convenience, are ordered according to (4.51). We will denote by e^μ, $\mu = 1, \ldots, p$, the column vectors of E_p and by E_p^t its transpose. It can easily be verified that

$$2^{-p}(e^\mu, e^\nu) = \begin{cases} 1, & \text{if } \mu = \nu, \\ 0, & \text{otherwise.} \end{cases} \tag{4.62}$$

Thus, the $2^p \times 2^p$ matrix $2^{-p} E_p E_p^t$ is a projector that projects on the subspace spanned by the orthogonal vectors $\{e^\mu\}_{\mu=1}^p$, and $2^{-p} E_p^t E_p$ is the identity in $\mathbb{R}^p$. Given a linear transformation A from R^p to R^q, we define

$$AC = \{Ax \mid x \in C\} \quad \text{for } C \subset \mathbb{R}^p. \tag{4.63}$$

With this notations the vector $n_p(y)$ and the set D_p, defined in (1.7) and (1.8), can be rewritten as

$$n_p(y) = 2^{-p} E_p^t y, \qquad D_p = 2^{-p} E_p^t [-1, 1]^{2^p}. \tag{4.64}$$

Moreover, for any set $I \subset \{1, \dots, p\}$, we have the property

$$\Pi_I D_p = D_{|I|}. \tag{4.65}$$

Finally, let us remark that, of course, the statements of Lemma 2.3 apply also to the deterministic function $\inf_{t\in\mathbb{R}^M} \mathbb{E}\Psi_{N,\beta}(m, t)$. All references to Lemma 2.3 in the sequel are to be understood as referring to properties of this latter function, that is given explicitly in (4.52).

By Lemma 2.3, the condition (4.1) of Theorem 4.1 is satisfied if and only if the supremum in the l.h.s of (4.1) is taken on at a point m in $\operatorname{int} D_M$. More precisely, by (2.15), this condition is equivalent to demanding that for all $\varepsilon > 0$ and all p, the supremum over y s.t. $\|\Pi_I n_p(y) - \tilde{m}\|_2 \le \varepsilon$ of $\Theta_p(y)$ is taken on at a point y^* such that

$$2^{-p} \sum_{\gamma=1}^{2^p} [I'(y^*_\gamma)]^2 \le c. \tag{4.66}$$

We set

$$\mathscr{A}_\varepsilon(\tilde{m}) \equiv \left\{ y \in [-1, 1]^{2^M} \colon \|\Pi_I n_M(y) - \tilde{m}\|_2 \le \varepsilon \right\}. \tag{4.67}$$

LEMMA 4.6. *Assume that* $0 < \beta < \infty$. *Then, for all* $\tilde{m} \in D_{|I|}$ *and* $\varepsilon > 0$, *there exists* $c(\tilde{m}, \varepsilon) < \infty$ *such that, for all* $p \ge |I|$,

$$\sup_{\substack{y\in[-1,1]^{2^p}\\ \Pi_I n_p(y) \in B_\varepsilon(\tilde{m})}} \Theta_p(y) = \Theta_p(y^*), \tag{4.68}$$

where

$$T_p(y^*) \equiv 2^{-p} \sum_{\gamma=1}^{2^p} \left[I'(y^*_\gamma) \right]^2 \le c(\tilde{m}, \varepsilon). \tag{4.69}$$

PROOF. The proof proceeds by showing that if y does not satisfy condition (4.69), then we can find a δy such that $y + \delta y \in \mathscr{A}_\varepsilon(\tilde{m})$ and $\Theta_p(y+\delta y) > \Theta_p(y)$, so that y cannot be the desired y^*. Let us first note that

$$\begin{aligned} \Theta_p(y + \delta y) - \Theta_p(y) &= \tfrac{1}{2}[\|n_p(y + \delta y)\|_2^2 - \|n_p(y)\|_2^2] \\ &\quad + 2^{-p}\beta^{-1} \sum_{\gamma=1}^{2^p} [I(y_\gamma) - I(y_\gamma + \delta y_\gamma)]. \end{aligned} \tag{4.70}$$

Using the properties of the matrix E_p and the fact that $y \in [-1,1]^{2^p}$ we can bound the difference of the quadratic terms as

$$\begin{aligned}\|n_p(y+\delta y)\|_2^2 - \|n_p(y)\|_2^2 &= \|n_p(\delta y)\|_2^2 + 2^{-p+1}(\delta y, 2^{-p}E_pE_p^T y)\\ &\geq -2^{-p/2+1}\|\delta y\|_2.\end{aligned} \tag{4.71}$$

Thus we can show that $\Theta_p(y+\delta y) > \Theta_p(y)$ holds by showing that

$$2^{-p}\beta^{-1}\sum_{\gamma=1}^{2^p}[I(y_\gamma) - I(y_\gamma + \delta y_\gamma)] > 2^{-p/2}\|\delta y\|_2. \tag{4.72}$$

Define the map Y from $[-1,1]^{2^p}$ to $[-1,1]^{2^{|I|}}$ by

$$Y_\gamma(y) \equiv 2^{-p+|I|}\sum_{\tilde\gamma=0}^{2^{p-|I|}-1} y_{\gamma+\tilde\gamma 2^{|I|}} \qquad \gamma = 1,\dots,2^{|I|}. \tag{4.73}$$

Using (4.64) we get that

$$\begin{aligned}\Pi_{|I|}n_p(y) &= 2^{-|I|}E_{|I|}^t\left(2^{-p+|I|}\sum_{\gamma=1}^{2^{p-|I|}}\Pi_{\{(\gamma-1)2^{|I|}+1,\dots,\gamma 2^{|I|}\}}y\right)\\ &= 2^{-|I|}E_{|I|}^t Y(y).\end{aligned} \tag{4.74}$$

Therefore, the property that $y \in \mathscr{A}_\varepsilon(\tilde m)$ depends only on the quantity $Y(y)$.

Notice that if $\tilde m \in D_{|I|}$ and $\varepsilon > 0$, then there exists $X \in (-1,1)^{2^{|I|}}$ such that $\|n_I(X) - \tilde m\|_2 \leq \varepsilon$. This implies that for any p, the vector $x \in \mathbb{R}^{2^p}$ with components $x_\gamma \equiv X_{\gamma \bmod 2^{|I|}}$ lies also in $\mathscr{A}_\varepsilon(\tilde m)$. Moreover,

$$\max_\gamma |x_\gamma| = \max_\gamma |X_\gamma| \equiv 1 - d < 1 \tag{4.75}$$

and, therefore,

$$T_p(x) \leq [I'(1-d)]^2 \tag{4.76}$$

is some finite p-independent constant. We will use this fact to construct our δy. We may of course choose an optimal X, that is, one for which d is maximal. In the sequel, X and x will refer to this vector. Let now y be a vector in $\mathscr{A}_\varepsilon(\tilde m)$ for which $T_p(y) > c$ for some large constant c. We will show that this cannot minimize Θ_p. We will distinguish two cases:

Case 1. Let us introduce two parameters, $0 < \eta \ll d$ and $0 < \lambda < 1$, that will be appropriately chosen later. In this first case we assume that y is such that

$$\sum_{\gamma=1}^{2^p}\mathbb{1}_{\{|y_\gamma|\geq 1-\eta\}} \geq (1-\lambda)2^{p-|I|} \tag{4.77}$$

and we choose

$$\delta y \equiv \rho(x - y), \tag{4.78}$$

where $0 < \rho < 1$ will be determined later. It then trivially follows from the definition of x and the convexity of the set $\mathscr{A}_\varepsilon$ that $y + \rho(x-y) \in \mathscr{A}_\varepsilon$ and that $y + \rho(x-y) \in [-1+\rho d, 1-\rho d]^{2^p}$. Thus, if we can show that with this choice, and with a ρ such that $\rho d > \eta$, (4.72) holds, we can exclude that the infimum is taken on for such a y.

Let us first consider components y_γ such that $|y_\gamma| > 1-d$. Since $|x_\gamma| \leq 1-d$, we have, for those components, $\operatorname{sgn} \delta y_\gamma = -\operatorname{sgn} y_\gamma$ and thus $I(y_\gamma) - I((y+\delta y)_\gamma) > 0$. This fact together with (4.77) entails

$$\begin{aligned} &2^{-p} \sum_{\gamma=1}^{2^p} [I(y_\gamma) - I((y+\delta y)_\gamma)] \mathbb{1}_{\{|y_\gamma| \geq 1-d\}} \\ &\quad \geq 2^{-p} \sum_{\gamma=1}^{2^p} [I(y_\gamma) - I((y+\delta y)_\gamma)] \mathbb{1}_{\{|y_\gamma| \geq 1-\eta\}} \\ &\quad \geq \inf_{\substack{|y_\gamma| \geq 1-\eta \\ |x_\gamma| \leq 1-d}} (1-\lambda) 2^{-|I|} [I(y_\gamma) - I((y+\delta y)_\gamma)]. \end{aligned} \tag{4.79}$$

Note that $I(z)$ is symmetric with respect to zero and is a strictly increasing function of z for $z > 0$. Thus $I((y+\delta y)_\gamma)$ is maximized over $|x_\gamma| \leq 1-d$ for $x_\gamma = (1-d) \operatorname{sgn} y_\gamma$. From this we get

$$\begin{aligned} &\inf_{|x_\gamma| \leq 1-d} [I(y_\gamma) - I((y+\delta y)_\gamma)] \\ &\qquad \geq [I(y_\gamma) - I(|y_\gamma| + \rho((1-d) - |y_\gamma|))] \end{aligned} \tag{4.80}$$

and the infimum over $|y_\gamma| \geq 1-\eta$ in the r.h.s. of (4.80) is easily seen to be taken on for $|y_\gamma| = 1-\eta$. Thus

$$\begin{aligned} &\inf_{\substack{|y_\gamma| \geq 1-\eta \\ |x_\gamma| \leq 1-d}} (1-\lambda) 2^{-|I|} [I(y_\gamma) - I((y+\delta y)_\gamma)] \\ &\quad \geq (1-\lambda) 2^{-|I|} [I(1-\eta) - I(1-\eta-\rho(d-\eta))] \\ &\quad \geq (1-\lambda) 2^{-|I|} \rho(d-\eta) I'(1-\eta-\rho(d-\eta)) \\ &\quad \geq (1-\lambda) 2^{-|I|} \rho(d-\eta) \tfrac{1}{2} \ln(\eta + \rho(d-\eta))|, \end{aligned} \tag{4.81}$$

where we have used the convexity of I and the bound $I'(1-x) \geq \frac{1}{2} |\ln x|$ for $0 < x < 1$.

We now have to consider the components y_γ with $|y_\gamma| \leq 1-d$. Here the entropy difference $I(y_\gamma) - I((y+\delta y)_\gamma)$ can of course be negative. To get a lower bound on this difference, we use (4.80) and perform the change of variable

$|y_\gamma| = (1-d) - z_\gamma$ to write

$$
\begin{aligned}
&\inf_{\substack{|y_\gamma|\leq 1-d\\ |x_\gamma|\leq 1-d}}[I(y_\gamma) - I((y+\delta y)_\gamma)]\\
&\qquad = \inf_{0\leq z_\gamma\leq 1-d} I((1-d) - z_\gamma + \rho z_\gamma) - I((1-d) - z_\gamma)\\
&\qquad \geq \inf_{0\leq z_\gamma\leq 1-d} -\rho z_\gamma I'((1-d) - z_\gamma + \rho z_\gamma)\\
&\qquad = \inf_{0\leq z_\gamma\leq 1-d} -\rho z_\gamma \frac{1}{2}\ln\left(\frac{2-d-z_\gamma+\rho z_\gamma}{d+z_\gamma-\rho z_\gamma}\right)\\
&\qquad \geq -\rho(1-d)\frac{1}{2}\ln\left(\frac{2-d}{d}\right)\\
&\qquad \geq -\frac{\rho}{2}\ln\frac{2}{d},
\end{aligned}
\tag{4.82}
$$

and putting together (4.82) and (4.77) yields

$$
\begin{aligned}
&2^{-p}\sum_{\gamma=1}^{2^p}[I(y_\gamma) - I((y+\delta y)_\gamma)]\mathbb{1}_{\{|y_\gamma|<1-d\}}\\
&\qquad \geq -(1-(1-\lambda)2^{-|I|})\frac{\rho}{2}\ln\left(\frac{2}{d}\right).
\end{aligned}
\tag{4.83}
$$

Therefore, (4.83) together with (4.79) and (4.81) gives

$$
\begin{aligned}
&\beta^{-1}2^{-p}\sum_{\gamma=1}^{2^p}[I(y_\gamma) - I((y+\delta y)_\gamma)]\\
&\qquad \geq \beta^{-1}\rho\biggl\{(1-\lambda)2^{-|I|}(d-\eta)\frac{1}{2}|\ln(\eta+\rho(d-\eta))|\\
&\qquad\qquad - (1-(1-\lambda)2^{-|I|})\frac{1}{2}\ln\left(\frac{2}{d}\right)\biggr\}.
\end{aligned}
\tag{4.84}
$$

On the other hand, we have

$$
2^{-p/2}\|\delta y\|_2 \leq 2\rho. \tag{4.85}
$$

Consequently, (4.72) holds if we can choose λ, η, and ρ so that the following inequality holds,

$$
\begin{aligned}
&\beta^{-1}\biggl\{(1-\lambda)2^{-|I|}(d-\eta)\frac{1}{2}|\ln(\eta+\rho(d-\eta))|\\
&\qquad - (1-(1-\lambda)2^{-|I|})\frac{1}{2}\ln\left(\frac{2}{d}\right)\biggr\} > 2.
\end{aligned}
\tag{4.86}
$$

However, this is always possible by taking, for example, $\lambda < 1$, $\eta \equiv \rho d/2$ and $\rho \equiv d^K$, where $K \equiv K(d, |I|, \lambda) > 1$ is chosen sufficiently large as to satisfy

$$(1-\lambda)2^{-|I|}d\left(\frac{1-d^K}{2}\right)\frac{K+1}{2}|\ln d| > 4 + \frac{1}{2}|\ln d|. \tag{4.87}$$

Case 2. We will assume that $\lambda < 1$ and that η and ρ are chosen as in Case 1. We can then assume that

$$\sum_{\gamma=1}^{2^p} \mathbb{1}_{\{|y_\gamma|\geq 1-\eta\}} < (1-\lambda)2^{p-|I|}. \tag{4.88}$$

We assume further that

$$T_p(y) > c \tag{4.89}$$

for c sufficiently large to be chosen later. Here we will choose δy such that

$$Y(\delta y) \equiv 0, \tag{4.90}$$

so that trivially $y + \delta y \in \mathscr{A}_\varepsilon(\tilde{m})$. Let us introduce a parameter $0 < \zeta < \eta$, that we will choose appropriately later, and let us set, for $\gamma \in \{1, \dots, 2^{|I|}\}$,

$$\mathscr{K}_\gamma^+ \equiv \{\tilde{\gamma} \in \{1, \dots, 2^{p-|I|}\} \mid |y_{\gamma+(\tilde{\gamma}-1)2^{|I|}}| \geq 1-\zeta\} \tag{4.91}$$

and

$$\mathscr{K}_\gamma^- \equiv \{\tilde{\gamma} \in \{1, \dots, 2^{p-|I|}\} \mid |y_{\gamma+(\tilde{\gamma}-1)2^{|I|}}| \leq 1-\eta\}.$$

For all indices γ such that $\mathscr{K}_\gamma^+ = \varnothing$, we simply set $\delta y_{\gamma+(\tilde{\gamma}-1)2^{|I|}} \equiv 0$ for all $\tilde{\gamma} \in \{1, \dots, 2^{p-|I|}\}$. If $\mathscr{K}_\gamma^+$ were empty for *all* γ, then $T_p(y) \leq [I'(1-\zeta)]^2$, which contradicts our assumption (4.89), for suitably large c (depending only on ζ). Thus we consider now the remaining indices γ for which $\mathscr{K}_\gamma^+ \neq \varnothing$.

First note that (4.88) implies that $|\mathscr{K}_\gamma^+| < (1-\lambda)2^{p-|I|}$ and that $\mathscr{K}_\gamma^- > \lambda 2^{p-|I|}$, so that choosing $1 > \lambda > \frac{1}{2}$, we have $|\mathscr{K}_\gamma^+| < |\mathscr{K}_\gamma^-|$. Our strategy will be to find δy in such a way as to decrease the moduli of the components in $\mathscr{K}_\gamma^+$ at the expense of possibly increasing them on $\mathscr{K}_\gamma^-$ in such a way as to leave $Y(y + \delta y) = Y(y)$.

In the sequel, we will consider the case where there is only one index γ, for example, $\gamma = 1$, for which $\mathscr{K}_\gamma^+$ is nonempty. The general case is treated essentially by iterating the same procedure. We will use the simplified notation $y_{1+2^{|I|}\tilde{\gamma}} \equiv y_{\tilde{\gamma}}$ and $\delta y_{1+2^{|I|}\tilde{\gamma}} \equiv \delta y_{\tilde{\gamma}}$ and also set $\mathscr{K}_1^\pm \equiv \mathscr{K}^\pm$. We will assume, moreover, that all components $y_{\tilde{\gamma}}$ are positive, as this is the worst situation. We will chose δy such that $\delta y_{\tilde{\gamma}} = 0$ if $\tilde{\gamma} \in \{\mathscr{K}^+ \cup \mathscr{K}^-\}^c$ and $\delta y_{\tilde{\gamma}} < 0$ if $\tilde{\gamma} \in \mathscr{K}^+$. For each $\tilde{\gamma} \in \mathscr{K}^+$ we will choose a unique and distinct $\tilde{\gamma}' \in \mathscr{K}^-$ and set $\delta y_{\tilde{\gamma}'} = -\delta y_{\tilde{\gamma}}$. This ensures that $Y(\delta y) = 0$. We will also make sure that for all $\tilde{\gamma}$, $|\delta y_{\tilde{\gamma}}| \leq \eta/2 - \zeta$.

We have to construct $\delta y_{\tilde\gamma}$ for $\tilde\gamma \in \mathscr{K}^+$. In this process we have to consider the following three functionals:

1. The change in the quadratic term of Θ_p. This is bounded by

$$\delta E(\delta y) \equiv 2^{-p/2+1}\sqrt{2\sum_{\tilde\gamma\in\mathscr{K}^+}\delta y_{\tilde\gamma}^2}\,. \tag{4.92}$$

2. The change in the entropy term,

$$\begin{aligned}\delta I(\delta y) &\equiv 2^{-p}\sum_{\tilde\gamma\in\mathscr{K}^+}\big(I(y_{\tilde\gamma}+\delta y_{\tilde\gamma})-I(y_{\tilde\gamma})\big)\\ &\qquad+2^{-p}\sum_{\tilde\gamma\in\mathscr{K}^-}\big(I(y_{\tilde\gamma}+\delta y_{\tilde\gamma})-I(y_{\tilde\gamma})\big)\\ &\geq 2^{-p}\sum_{\tilde\gamma\in\mathscr{K}^+}|\delta y_{\tilde\gamma}|I'(y_{\tilde\gamma}+\delta y_{\tilde\gamma})-2^{-p}\sum_{\tilde\gamma\in\mathscr{K}^-}\delta y_{\tilde\gamma}|\ln\eta|/2\\ &=2^{-p}\sum_{\tilde\gamma\in\mathscr{K}^+}|\delta y_{\tilde\gamma}|\big(I'(y_{\tilde\gamma}+\delta y_{\tilde\gamma})-|\ln\eta|/2\big)\\ &\geq 2^{-p-1}\sum_{\tilde\gamma\in\mathscr{K}^+}|\delta y_{\tilde\gamma}|I'(y_{\tilde\gamma}+\delta y_{\tilde\gamma}),\end{aligned} \tag{4.93}$$

   where we have used that for $1\geq|x|\gg|y|\gg 0.9$, $I(x)-I(y)\approx|x-y||\ln(1-y)|$ and that, under our assumption, for $\tilde\gamma\in\mathscr{K}^+$, $y_{\tilde\gamma}+\delta y_{\tilde\gamma}\geq 1-\eta/2$.
3. Finally, we have that

$$\begin{aligned}T_p(y+\delta y) &\leq 2^{-p}\sum_{\gamma\notin\mathscr{K}^+}[I'(y_{\tilde\gamma}+\delta y_{\tilde\gamma})]^2+2^{-p}\sum_{\tilde\gamma\in\mathscr{K}^+}[I'(y_{\tilde\gamma}+\delta y_{\tilde\gamma})]^2\\ &\leq [I'(1-\eta/2)]^2+2^{-p}\sum_{\tilde\gamma\in\mathscr{K}^+}[I'(y_{\tilde\gamma}+\delta y_{\tilde\gamma})]^2.\end{aligned} \tag{4.94}$$

Looking at these three functionals suggests choosing $\delta y_{\tilde\gamma}$ for $\tilde\gamma\in\mathscr{K}^+$ as the solution of the equation

$$-\delta y_{\tilde\gamma}=\tau I'(y_{\tilde\gamma}+\delta y_{\tilde\gamma}). \tag{4.95}$$

The point is that with this choice, (4.94) yields [we set for simplicity $\delta E(\delta y(\tau))\equiv\delta E(\tau)$, etc.]

$$\delta I(\tau)\geq\frac{1}{8\tau}(\delta E(\tau))^2 \tag{4.96}$$

while

$$T_p(\tau)\leq[I'(1-\zeta)]^2+\tau^{-2}(\delta E(\tau))^2. \tag{4.97}$$

Thus we can ensure that the entropy gain dominates the potential loss in the quadratic term provided we can choose $\tau<\delta E(\tau)/8$. However, we know that

$T_p(\tau)$ is a continuous function of t and $T_p(0) \geq c$. Thus there exists $\tau_0 > 0$ such that for all $\tau \leq \tau_0$, $T_p(\tau) \geq c/2$, and so by (4.97),

$$\tau^{-1}\delta E(\tau) \geq \sqrt{c/2 - [I'(1-\zeta)]^2}, \tag{4.98}$$

which inserted into (4.97) yields that

$$\delta I(\tau) \geq (\ln 2/4)\sqrt{c/2 - [I'(1-\zeta)]^2}\,\delta E(\tau). \tag{4.99}$$

It is clear that if c is chosen large enough ("large" depending only on ζ), this gives $\delta I(t) > \delta E(t)$, as desired. Finally, it is easy to see that $|\delta y_{\tilde{\gamma}}|$ is bounded from above by the solution of the equation

$$x = \tau I'(1-x), \tag{4.100}$$

which is of the order of $x \approx \tau|\ln\tau|$. If ζ is chosen, for example, $\zeta = \eta/4$, we see from this that for small enough τ, $|\delta y_{\tilde{\gamma}}| \leq \eta/2 - \zeta$, so that all our conditions can be satisfied. Thus, there exist $c < \infty$ depending only on η (which in turn depends only on $\tilde{m}$ and ε) such that any y that satisfies the assumptions of Case 2 with this choice of c in (4.89) cannot realize the infimum of Θ_p. The two cases combined prove the lemma. □

To conclude the proof of Theorem 1 we show that, for $\tilde{m} \in D_I^c$, (1.13) holds. This turns out to be rather simple. The main idea is that if $\tilde{m} \in D_{|I|}^c$, then on a subset of Ω of probability 1, for N large enough and ε small enough, the set $\{\sigma \in \mathcal{S}_N \mid \|\Pi_I m_N(\sigma) - \tilde{m}\|_2 \leq \varepsilon\}$ is empty.

To do so we will first show that uniformly in the configurations σ, the vector $\Pi_I m_N(\sigma)$ can be rewritten as the sum of a vector in $D_{|I|}$ and a vector whose norm goes to zero as N goes to infinity. Let e_γ, $\gamma = 1, \dots, 2^{|I|}$, be the column vectors of the matrix $E_{|I|}^t$. We set

$$v_\gamma \equiv \{i \in \{1, \dots, N\} \mid \xi_i^\mu = e_\gamma^\mu,\ \forall \mu \in I\}. \tag{4.101}$$

These sets are random sets, depending on the realization of the random variables ξ_i^μ. Their cardinality, however, remains very close to their mean value. More precisely, let λ_γ denote the fluctuation of $|v_\gamma|$ about its mean:

$$\lambda_\gamma \equiv 2^{|I|}N^{-1}\big||v_\gamma| - 2^{-|I|}N\big|. \tag{4.102}$$

There exists a subset $\Omega_4 \in \Omega$ of probability 1 and a function δ_N, tending to zero as N tends to infinity, such that, for all but a finite number of indices,

$$|\lambda_\gamma| < \delta_N, \qquad \gamma = 1, \dots, 2^{|I|}. \tag{4.103}$$

This fact has been proven in [10]. Using (4.101), $\Pi_I m_N(\sigma)$ can be rewritten as

$$\Pi_I m_N(\sigma) = 2^{-|I|}E_{|I|}^t(X(\sigma) + \delta X(\sigma)), \tag{4.104}$$

where $X(\sigma)$ and $(\delta X)(\sigma)$ are, respectively, the vectors with components $X_\gamma(\sigma) \equiv |v_\gamma|^{-1} \sum_{i \in v_\gamma} \sigma_i \in [-1, 1]$, $(\delta X)_\gamma(\sigma) \equiv \lambda_\gamma X_\gamma(\sigma)$, $\gamma = 1, \dots, 2^{|I|}$. It then follows from the properties of the matrix $E^t_{|I|}$ and (4.103) that, on Ω_4,

$$\|\Pi_I m_N(\sigma) - n_{|I|}(X(\sigma))\|_2 < \delta_N. \tag{4.105}$$

Now, by assumption, $\tilde{m} \in D^c_{|I|}$, that is, there exists $\tilde{\varepsilon} > 0$ such that $\{x \in \mathbb{R}^{|I|} \mid \|x - \tilde{m}\|_2 \leq \tilde{\varepsilon}\} \subset (D_{|I|})^c$. Therefore, since $n_{|I|}(X(\sigma)) \in D_{|I|}$, we have $\|n_{|I|}(X(\sigma)) - \tilde{m}\|_2 > \tilde{\varepsilon}$. From this and (4.105) it follows that on Ω_4, $\|\Pi_I m_N(\sigma) - \tilde{m}\|_2 > \tilde{\varepsilon} - \delta_N$. Finally, for N large enough and ε small enough, we get

$$\{\sigma \in \mathscr{S}_N \mid \|\Pi_I m_N(\sigma) - \tilde{m}\|_2 \leq \varepsilon\} = \varnothing. \tag{4.106}$$

From this, (1.13) easily follows. This concludes the proof of Theorem 1. □

Acknowledgments. We would like to thank an anonymous referee for a careful reading of the manuscript, for pointing out a number of errors and for helpful suggestions concerning the presentation of the paper.

REFERENCES

[1] AMIT, D. J., GUTFREUND, H. and SOMPOLINSKY, H. (1987). Statistical mechanics of neural networks near saturation. *Ann. Phys.* **173** 30–67.

[2] BOVIER, A. (1994). Self-averaging in a class of generalized Hopfield models. *J. Phys. A* **27** 7069–7077.

[3] BOVIER, A. and GAYRARD, V. (1993). Rigorous results on the thermodynamics of the dilute Hopfield model. *J. Statist. Phys.* **69** 597–627.

[4] BOVIER, A., GAYRARD, V. and PICCO, P. (1994). Gibbs states of the Hopfield model in the regime of perfect memory. *Probab. Theory Related Fields* **100** 329–363.

[5] BOVIER, A., GAYRARD, V. and PICCO, P. (1995). Large deviation principles for the Hopfield model and the Kac–Hopfield model. *Probab. Theory Related Fields* **101** 511–546.

[6] BOVIER, A., GAYRARD, V. and PICCO, P. (1995). Gibbs states of the Hopfield model with extensively many patterns. *J. Statist. Phys.* **79** 395–414.

[7] COMETS, F. (1989). Large deviation estimates for a conditional probability distribution. Applications to random Gibbs measures. *Probab. Theory Related Fields* **80** 407–432.

[8] DEMBO, A. and ZEITOUNI, O. (1992). *Large Deviation Techniques and Applications*. Jones and Bartlett, Boston.

[9] ELLIS, R. S. (1985). *Entropy, Large Deviations, and Statistical Mechanics*. Springer, Berlin.

[10] GAYRARD, V. (1992). Thermodynamic limit of the q-state Potts–Hopfield model with infinitely many patterns. *J. Statist. Phys.* **68** 977–1011.

[11] HOPFIELD, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci. U.S.A.* **79** 2554–2558.

[12] KOCH, H. (1993). A free energy bound for the Hopfield model. *J. Phys. A* **26** L353–L355.

[13] KOCH, H. and PIASKO, J. (1989). Some rigorous results on the Hopfield neural network model. *J. Statist. Phys.* **55** 903–928.

[14] KOMLÓS, J. and PATURI, R. (1988). Convergence results in a autoassociative memory model. *Neural Networks* **1** 239–250.

[15] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces*. Springer, Berlin.

[16] NEWMAN, CH. M. (1988). Memory capacity in neural network models: rigorous lower bounds. *Neural Networks* **1** 223–238.

[17] PARISI, G. (1995). Complexity in biology: the point of view of a physicist. Preprint, Univ. Rome.

[18] PASTUR, L. A. and FIGOTIN, A. L. (1977). Exactly soluble model of a spin glass. *Soviet Journal of Low Temperature Physics* **3** 378–383.
[19] PASTUR, L. A. and FIGOTIN, A. L. (1978). On the theory of disordered spin systems. *Theoret. and Math. Phys.* **35** 403–414.
[20] PASTUR, L., SHCHERBINA, M. and TIROZZI, B. (1994). The replica symmetric solution without the replica trick for the Hopfield model. *J. Statist. Phys.* **74** 1161–1183.
[21] ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton Univ. Press.
[22] SHCHERBINA, M. and TIROZZI, B. (1992). The free energy for a class of Hopfield models. *J. Statist. Phys.* **72** 113–125.
[23] VARADHAN, S. R. S. (1984). *Large deviations and applications*. SIAM, Philadelphia.
[24] YURINSKII, V. V. (1976). Exponential inequalities for sums of random vectors. *J. Multivariate Anal.* **6** 473–499.

WEIERSTRAß-INSTITUT
FÜR ANGEWANDTE ANALYSIS UND STOCHASTIK
MOHRENSTRASSE 39
D-10117 BERLIN
GERMANY

CENTRE DE PHYSIQUE THÉORIQUE-CNRS
LUNINY, CASE 907
F-13288 MARSEILLE CEDEX 9
FRANCE