GIBBS STATES OF THE HOPFIELD MODEL IN THE REGIME OF PERFECT MEMORY#

Anton Bovier [1 2 3], Véronique Gayrard[4], and Pierre Picco[5]

Abstract: We study the thermodynamic properties of the Hopfield model of an autoassociative memory. If N denotes the number of neurons and $M(N)$ the number of stored patterns, we prove the following results: If $\frac{M}{N} \downarrow 0$ as $N \uparrow \infty$, then there exists an infinite number of infinite volume Gibbs measures for all temperatures $T < 1$ concentrated on spin configurations that have overlap with exactly one specific pattern. Moreover, the measures induced on the overlap parameters are Dirac measures concentrated on a single point. If $\frac{M}{N} \to \alpha$, as $N \uparrow \infty$ for α small enough, we show that for temperatures T smaller than some $T(\alpha) < 1$, the induced measures can have support only on a disjoint union of balls around the previous points, but we cannot construct the infinite volume measures through convergent sequences of measures.

AMS Subject Classification Numbers: 60K35, 82B44, 82C32

April, 1993
CPT-93/P. 2894

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

[1] Institut für Angewandte Analysis und Stochastik, Hausvogteiplatz 5-7, O-1086 Berlin, Germany

[2] and PHYMAT, Département de Mathématiques, Université de Toulon et du Var, B.P. 132 - F-83957 La Garde Cedex, France

[3] e-mail: bovier@iaas-berlin.dbp.de

[4] e-mail: gayrard@marcptsu1.univ-mrs.fr

[5] e-mail: picco@cptsu2.univ-mrs.fr

I. Introduction

Disordered spin systems are one of the topics of highest current interest in mathematical statistical mechanics. Generally speaking, depending on the particular types of models, the effects of disorder may be either weak, in the sense that the model can be reasonably well approximated by a ordered one, or strong, in the sense that genuinely new phenomena appear that are pertinent to randomness. Prototypical systems of the latter kind are spin glasses, and in particular their low-temperature properties. While on the heuristic level a rather coherent theory has been developed at least for a mean field version, the Sherrington-Kirkpatrick model [SK] (for a review see [M]), from a mathematical point of view these models are extremely difficult to analyse and virtually no results are know, even on the mean field level. In order to progress in the understanding of such models, it is highly desirable to have a class of models that intermediate between simple 'almost ordered' models and spin glasses and that are amenable at least in part to a rigorous analysis. Such models are in fact provided by what is commonly known as the Hopfield model [Ho], and their analysis has attracted increasing attention of mathematical physicists over the last years.

Let us describe this model. We set $\Lambda \equiv \{1, \dots, N\}$ and $\mathcal{S}_\Lambda = \{-1,1\}^N$ the space of functions $\sigma : \Lambda \to \{-1,1\}$. We call σ a *spin configuration* on Λ. We shall write $\mathcal{S} \equiv \{-1,1\}^{\mathbb{N}}$ for the space of half infinite sequences equipped with the product topology of 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, identically distributed random variables on this space such that $\mathbb{P}(\xi_i^\mu = 1) = \mathbb{P}(\xi_i^\mu = -1) = \frac{1}{2}$. For a given non-decreasing integer valued function $M : \mathbb{N} \to \mathbb{N}$ we denote by $\mathcal{F}_N$ the sub-sigma algebra generated by the random variables $\{\xi_i^\mu\}_{1\leq i\leq N}^{1\leq\mu\leq M(N)}$. We will occasionally denote this sub-family of random variables by $\xi|_N$. The Hopfield Hamiltonian on $\mathcal{S}_\Lambda$ is then given by

$$H_N[\omega](\sigma) = -\frac{1}{2N} \sum_{(i,j)\in\Lambda\times\Lambda} \sum_{\mu=1}^{M(N)} \xi_i^\mu(\omega)\xi_j^\mu(\omega)\sigma_i\sigma_j \tag{1.1}$$

Note that of course $H_N[\omega](\sigma)$ is $\mathcal{F}_N$-measurable. In (1.1) we have made all dependences on the random parameter ω explicit. In the sequel we will drop this whenever no confusion may arise.

The Hopfield model has in fact been proposed in the context of neural networks as a model for autoassociative memory. The interpretation of the above objects in this context is the following: Λ is a (completely connected) set of N neurons, each of which can be in two states, $+1$ or -1. An element of $\mathcal{S}_\Lambda$ then describes the state of the neural network. The M families of random variables $\{\xi_i^\mu\}_{i\in\Lambda}$ are thus M randomly chosen states of the network, called *'patterns'*. Functioning of the memory is interpreted (see e.g. [A]) in that a Markovian time evolution set up in such a way that its invariant measure is a Gibbs state[1] of the Hamiltonian (1.1) should have a long time behaviour that allows to discern whether the initial condition was close to one of the patterns or not. This phenomena is clearly related to the question of breaking of ergodicity in the infinite volume limit (i.e. in the limit as $N \uparrow \infty$) of this system and thus to the existence and nature of the infinite volume Gibbs states associated to the Hopfield Hamiltonian. Thermodynamic properties of this model have thus a direct interpretation in the neural context.

The properties of this model depend crucially on the choice of the function $M(N)$. If $M(N) \equiv 1$, it is trivially equivalent to the Curie-Weiss ferromagnet by a simple change of variables. If $M(N)$ remains bounded, rather standard methods can still be applied to give a complete characterization of the Gibbs measures that in fact show the desired features of a perfect memory below the critical temperature [AGS1,H]. We would like to point out here that this case had previously been treated extensively and with mathematical rigor in papers by Pastur and Figotin in 1977 [FP1,FP2]. These remarkable articles which appear to have been fallen largely to oblivion are by the way to our knowledge the first in which the models like (1.1) have been proposed as simplified models of spin glasses.

For unbounded M the situation becomes more complicated, and the results will depend on the allow rate of growth. Koch and Piasko [KP] and Gayrard [G] have proven that essentially the same results as for bounded M can be proven if $M(N) \leq \frac{\ln N}{\ln 2}$. For faster growth rates, no rigorous results on the Gibbs states are available, but heuristic results indicate that the memory should function properly as long as $M(N) \leq \alpha N$, for small enough α. This idea is also supported by rigorous results of Newman [N] and Komlós and Paturi [KPa] on the structure of the local minima of the Hamiltonian. It is also believed that the precise structure of the Gibbs states in this regime is already fairly complicated and will depend on the precise growth properties of $M(N)$ [AGS2]. If $M(N) \geq \alpha_c N$, it is expected that the picture changes qualitatively completely. In fact, the faster $M(N)$ is allowed to grow, the more the model resembles a spin glass, since the effective couplings $J_{ij} \equiv \frac{1}{\sqrt{M}} \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu$ converge to i.i.d. Gaussian r.v.'s as $M \uparrow \infty$. It is in this sense that the Hopfield model provides a family of models intermediating between ferromagnets and spin glasses.

Recently, Shcherbina and Tirozzi [ST] have proven that if $\frac{M}{N} \downarrow 0$ as $N \uparrow \infty$, the free energy of this model converges in probability to the free energy of the Curie-Weiss model (see e.g. [E] for mathematical results concerning this model). Later, Koch [K] has obtained interesting upper and lower bounds on the expectation of the free energy that imply in particular the convergence of this quantity to the free energy of the Curie-Weiss model as $N \uparrow \infty$ under the condition that $M(N)/N \downarrow 0$ (without hypothesis on the speed). As has been noted already in [BG], his proof can easily be modified to yield the $I\!P$-almost sure convergence of the free energy under this hypothesis,

[1] We follow common practice and use the terms *Gibbs state* and *Gibbs measure* synonymously.

the proof being considerably simpler than the one in [ST].

The purpose of the present paper is to provide a complete analysis of the Gibbs states of the Hopfield model under the same hypothesis on M. We will also give a somewhat weaker results on the Gibbs measures in the regime $M(N) = \alpha N$, with $\alpha \leq \alpha_0$, for some (ridiculously) small α_0. To give a precise formulation of our results, let us fix our notations.

For $\eta \in I\!N$, we denote by $\mathcal{G}^{\eta}_{N,\beta,h}[\omega]$ the random probability measure on $(\mathcal{S}_\Lambda, \mathcal{B}(\mathcal{S}_\Lambda))$ that assigns to each $\sigma \in \mathcal{S}_\Lambda$ the mass

$$\mathcal{G}^{\eta}_{N,\beta,h}[\omega](\sigma) \equiv \frac{1}{Z^{\eta}_{N,\beta,h}[\omega]} e^{-\beta H_N[\omega](\sigma) - \beta h \sum_{i\in\Lambda} \xi^{\eta}_i(\omega)\sigma_i} \tag{1.2}$$

where $Z^{\eta}_{N,\beta,h}[\omega]$ is a normalizing factor usually called *partition function*. The quantity

$$f^{\eta}_{N,\beta,h}[\omega] \equiv -\frac{1}{\beta N} \ln Z^{\eta}_{N,\beta,h}[\omega] \tag{1.3}$$

is called the *free energy*. Note that all these quantities are $\mathcal{F}_N$-measurable. The parameter β is the inverse temperature and h is called a *magnetic field* aligned on the pattern ξ^{η}, and $\mathcal{G}^{\eta}_{N,\beta,h}[\omega]$ is called a *finite volume Gibbs state with magnetic field.* An important observation is that the value of the measure $\mathcal{G}^{\eta}_{N,\beta,h}[\omega](\sigma)$ does depend on σ only through the quantities

$$m^{\mu}_N[\omega](\sigma) \equiv \frac{1}{N}\sum_{i=1}^{N} \xi^{\mu}_i[\omega]\sigma_i, \quad \mu = 1,\dots,M \tag{1.4}$$

called *overlap parameters*, since the Hamiltonian may be written in the form

$$H_N[\omega](\sigma) = -\beta N \sum_{\mu=1}^{M} \left(m^{\mu}_N[\omega](\sigma)\right)^2 \tag{1.5}$$

This suggests to define the random map

$$\begin{aligned} \mathcal{M}_N[\omega] : \mathcal{S}_\Lambda &\to I\!R^M \\ \sigma &\to \mathcal{M}_N[\omega](\sigma) \equiv \left(m^1_N[\omega](\sigma), \dots, m^M_N[\omega](\sigma)\right) \end{aligned} \tag{1.6}$$

and the measures $\mathcal{Q}^{\eta}_{N,\beta,h}[\omega]$ on $(I\!R^M, \mathcal{B}(I\!R^M))$ that are induced by $\mathcal{G}^{\eta}_{N,\beta,h}[\omega]$ through the map $\mathcal{M}_N[\omega]$, i.e.

$$\mathcal{Q}^{\eta}_{N,\beta,h}[\omega] \equiv \mathcal{G}^{\eta}_{N,\beta,h}[\omega] \circ \mathcal{M}_N[\omega]^{-1} \tag{1.7}$$

In fact, these induced measures determine the original measures uniquely, since

$$\mathcal{G}^{\eta}_{N,\beta,h}[\omega](\sigma) = \frac{1}{|\mathcal{M}_N[\omega]^{-1}(\mathcal{M}_N[\omega](\sigma))|} \mathcal{Q}^{\eta}_{N,\beta,h}[\omega](\mathcal{M}_N[\omega](\sigma)) \tag{1.8}$$

while on the other hand, as will become evident, the induced measures are "less random" in some sense and thus easier to control. (E.g., if $M = 1$, the induced measures are entirely deterministic and the dependence on the random parameter is only through the map $\mathcal{M}$). Note that of course for fixed N, $\mathcal{M}_N[\omega]$ takes only values in the set $\{-1, -1+\frac{2}{N}, -1+\frac{4}{N}, \dots, 1-\frac{2}{N}, 1\}^M$ and $\mathcal{Q}^\eta_{N,\beta,h}[\omega]$ is an atomic measure concentrated on this set.

We want to study our model in the limit as $N \uparrow \infty$, and are primarily interested in the case where M is unbounded. We therefore want to work on the measure space $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$, with $\mathbb{R}^{\mathbb{N}}$ understood to be the space of infinite sequences equipped with the product topology of the euclidean topology of $\mathbb{R}$. For notational simplicity we will identify the measures $\mathcal{Q}^\eta_{N,\beta,h}[\omega]$ with their extensions to $\mathbb{R}^{\mathbb{N}}$ obtained by tensoring them with the Dirac-measure concentrated at 0.

Before we formulate our Theorems, some general remarks on Gibbs states in mean field models need to be made. As is well known, as opposed to 'normal' models of statistical mechanics, there is no neat characterization of infinite volume Gibbs states as solutions of the DLR-equations (see e.g. [Ge]) and therefore they can only be constructed as limit points of sequences of finite volume measures. To ensure convergence and to lift degeneracies, that is 'to pick out its *extremal* measures', it is customary to add 'magnetic fields' which are taken to zero *after* the infinite volume limit is performed. This was also the reason for defining the measures in (1.2). As we also want to know whether we have found *all* such measures, we have to give a more precise definition of what we shall mean by a *limiting Gibbs measure* for a Hamiltonian H. To do this, let $f : \mathbb{N} \to \mathbb{R}$ be a real-valued function on the integers with bounded sup-norm. For such a function we define the measures

$$\mathcal{G}^{(f)}_{N,\beta}[\omega](\sigma) \equiv \frac{1}{Z^\eta_{N,\beta,h}[\omega]} e^{-\beta H_N[\omega](\sigma) - \beta \sum_{i\in\Lambda} f_i \sigma_i} \tag{1.10}$$

where f_i denotes the value of f at site i. The measures defined in (1.2) are of course particular examples where $f_i = h\xi^\eta_i[\omega]$. We denote by $\mathcal{Q}^{(f)}_{N,\beta}[\omega]$ the corresponding induced measures.

We will say that $\mathcal{G}_\beta$ (resp. $\mathcal{Q}_\beta$) is a *limiting Gibbs measure* (resp. *limiting induced measure*) for the Hamiltonian H, if there exists a sequence of integers N_l tending to infinity as $l \uparrow \infty$ and a sequence of functions $f^{(k)}$ such that $\|f^{(k)}\|_\infty \downarrow 0$ as $k \uparrow \infty$, such that

$$\mathcal{G}_\beta[\omega] = w - \lim_{k\uparrow\infty} \lim_{l\uparrow\infty} \mathcal{G}^{(f^{(k)})}_{\beta,N_l}[\omega] \tag{1.11}$$

respectively

$$\mathcal{Q}_\beta[\omega] = w - \lim_{k\uparrow\infty} \lim_{l\uparrow\infty} \mathcal{Q}^{(f^{(k)})}_{\beta,N_l}[\omega] \tag{1.12}$$

(We use the symbol $w-\lim$ to denote the weak limits of probability measures). Note that both the sequences N_l and $f^{(k)}$ may be random variables depending strongly on ξ (and in general at least one of them will have to be random for the limits to exist).

With these notions, we are ready to announce our first Theorem:

Theorem 1: *Assume that M is non-decreasing and satisfies $\lim_{N\uparrow\infty} \frac{M(N)}{N} = 0$. Let $a^{\pm}(\beta)$ denote the largest (resp. smallest) solution of $a = \tanh(\beta a)$. Then, for all $\beta \geq 0$,*

(i)

$$w - \lim_{h\to 0\pm} \lim_{N\uparrow\infty} \mathcal{Q}^{\eta}_{N,\beta,h}[\omega] = \delta^{\infty}_{a^{\pm}(\beta)e^{\eta}}, \quad I\!P - \text{almost surely} \tag{1.13}$$

where the limits are understood in the sense of weak convergence of probability distributions; $\delta^{\infty}_{a^{\pm}(\beta)e^{\eta}}$ denotes the Dirac-measure concentrated on $a^{\pm}(\beta)e^{\eta}$ and ϵ^{η} is the η-th unit vector in $I\!R^{I\!N}$.

(ii) Moreover, any limiting induced measure for the Hamiltonian (1.1) is a convex combination of the measures in (1.13)

Remark: Note that for $\beta \leq 1$, $a^{+}(\beta) = a^{-}(\beta) = 0$ so that in this case there is a unique limiting measure. For $\beta > 1$, the measures for different η and different signs of h are all distinct and by the second statement of the theorem can reasonably be seen as the *extremal* measures. Note however that here as in general for mean field models, not all convex combinations of the extremal measures are themselves limiting measures.

We believe that our condition on M is the weakest possible under which the conclusions of Theorem 1 can hold.

Our next theorem will be concerned with the case where $\frac{M}{N} > 0$. For $\delta > 0$, we will write $a(\delta,\beta)$ for the largest solution of the equation

$$\delta a = \tanh(\beta a) \tag{1.14}$$

We denote by $\|\cdot\|$ the ℓ^2-norm on $I\!R^{I\!N}$. Given that $\lim_{N\uparrow\infty} \frac{M(N)}{N} = \alpha$, we set, for fixed β,

$$B^{(\nu,s)}_{\rho} \equiv \left\{ x \in I\!R^{I\!N} \,\middle|\, \|x - sa(1-2\sqrt{\alpha},\beta)e^{\nu}\| < \rho \right\} \tag{1.15}$$

Finally, we put

$$B_{\rho} \equiv \bigcup_{(\nu,s)\in I\!N\times\{-1,+1\}} B^{(\nu,s)}_{\rho} \tag{1.16}$$

With this notation we can announce our second theorem

Theorem 2: *There exists $\alpha_0 > 0$ such that if $\lim \frac{M(N)}{N} = \alpha$, with $\alpha \leq \alpha_0$, then, for all $\beta > 1 + 3\sqrt{\alpha}$, if $\rho^2 > C\left(a(1-2\sqrt{\alpha},\beta)\right)^{3/2}\alpha^{1/8}|\ln\alpha|^{1/4}$, for almost all ω,*

$$\lim_{N\uparrow\infty} \mathcal{Q}_{N,\beta,h=0}[\omega]\,(B_{\rho}) = 1 \tag{1.17}$$

Remark: Theorem 2 suggests of course that there should exist limiting Gibbs states with support in just one of the balls $B_\rho^{(s,\nu)}$; in the case $\alpha = 0$ we have constructed these as limits by adding a small magnetic field. Unfortunately, if $\alpha > 0$, we have not been able to do this, and the construction of the limiting measures remains an interesting open problem. However, Theorem 2 *excludes* in particular that any of the so-called *mixed states* (which have been shown to be associated to local minima of the Hamiltonian; see [AGS,N,KPa]) give rise to Gibbs states in this regime of parameters.

Remark: From the properties of the solutions of equation (1.14) (in particular $a(\delta,\beta) \sim (\beta-\delta)$ for $\beta-\delta$ small) it follows that the set B_ρ, with the minimal allowed value of ρ inserted is a union of disjoint balls as long as $\beta > \frac{1}{1-c\alpha^{1/4}}$. Note that the power 1/4 in this equation may not be optimal; indeed it is expected from the heuristic analysis of Amit et al. [AGS2] that for $\beta > \frac{1}{1-c\sqrt{\alpha}}$ the model should show 'perfect memory', i.e disjoint Gibbs states for each pattern. The region $\beta < \frac{1}{1+\sqrt{\alpha}}$ is the paramagnetic phase where uniqueness of the Gibbs state is expected to hold. For rigorous results on this domain see [ScT]. The more complicated region in-between is what Amit et al. call a spin glass phase.

Remark: In [BG] the analogue of Theorem 1 for the dilute Hopfield model has been proved under the hypothesis that $M < \frac{\ln N}{\ln 2}$ and that the dilution rate, $p(N)$, satisfies $p(N)N \uparrow \infty$. As has already been pointed out in [BG], the fact that here we have proven Theorem 1 under the weaker hypothesis $\frac{M}{N} \downarrow 0$ implies that the conclusions of Theorem 1 hold for the dilute model under the conditions $p(N)N \uparrow \infty$ and $\frac{M}{p(N)N} \downarrow 0$.

An obvious question that remains is of course that of the nature of the limiting Gibbs measures as measures on the spin-space, i.e. on $(\mathcal{S},\mathcal{B})$. In mean field models one is used to the fact that these are product measures, and under the hypothesis of Theorem 1 this is indeed the case here:

Theorem 3: *Under the assumptions and with the notation of Theorem 1,*

$$w - \lim_{h\to\pm 0} \lim_{N\uparrow\infty} \mathcal{G}^\eta_{\beta,N,h}[\omega] = B^\eta_{\pm a(\beta)}[\omega] \quad I\!P - \textit{almost surely} \tag{1.18}$$

where $B^\eta_a[\omega]$ denotes the product measure on $\{-1,1\}^{I\!N}$ with the marginal measure on σ_i given by the Bernoulli measure on $\{-1,1\}$ with mean $\xi^\eta_i a$.

Remark: Theorem 3 is all but a Corollary of Theorem 1, and follows, as we will see with very little work from the estimates we will use to proof Theorem 1. The crucial point that is needed to obtain the product structure of the limiting Gibbs measures is that the limiting induced measures are degenerate (i.e. concentrated on a single point). It is an interesting question whether this property does or does not extend to small but finite values of α. We should like to mention that factorization of the limiting measures is also tied to the so-called 'self-averaging' of the Edwards-

Anderson parameter (the spatial average of the square of the Gibbsian expectation of the σ_i) and thus, via recent work of Pastur, Shcherbina and Tirozzi [PST] to the validity of the so-called 'replica symmetric solutions' of [AGS]. We do not want to enter into any details here but refer the interested reader to [PST][2]

Under the assumption that $M(N)$ is bounded, the statement of Theorem 3 has previously been obtained by Comets [Co].

The remainder of this article is organized as follows. In the next section we will introduce a smooth version of the induced measure for finite N that will converge to the same limit as $\mathcal{Q}$ but will be easier to treat using Laplace's method. We will write an explicit expression for the density of this measure in the form $\exp\left(-\beta N\Phi(x)\right)/Z$ for an explicitly given random function Φ. In Section 3 we analyse the structure of the global minima of this function, and in Section 4 we use these results to prove the theorems. An appendix contains the proof of a bound on eigenvalues of a certain random matrix that is used frequently.

Aknowledgements: We would like to thank the Institut für Angewandte Analysis und Stochastik, Berlin, the Centre de Physique Théorique, Marseille, and the Université de Toulon et du Var for their kind hospitality that made this collaboration possible.

II. Some technical preparations

To prove the theorems announced in Section 1 we will introduce a smooth version of the induced measures by convoluting $\mathcal{Q}$ with a Gaussian measure. We denote as usual by $\mathcal{N}_M(\mu,\sigma)$ the Gaussian measure on $I\!R^M$ with mean μ and variance σ. We will also identify this measure with the measure on $I\!R^{I\!N}$ obtained by tensoring it with the Dirac-measure concentrated at zero. We define

$$\widetilde{\mathcal{Q}}^\eta_{N,\beta,h}[\omega] \equiv \mathcal{N}_{M(N)}(he^\eta, [\beta N]^{-1}1\!\!1) \star \mathcal{Q}^\eta_{N,\beta,h}[\omega] \tag{2.1}$$

(In the physics literature, this is known as the Hubbard-Stratonovich transformation [HS]). The point here is that since $\mathcal{N}_{M(N)}(he^\eta,[\beta N]^{-1}1\!\!1) \xrightarrow{w} \delta_0^\infty$ as $N\uparrow\infty$, and $h\to 0$, the convergence properties of $\mathcal{Q}^\eta_{N,\beta,h}[\omega]$ are the same as those of $\widetilde{\mathcal{Q}}^\eta_{N,\beta,h}[\omega]$, i.e. we see immediately that

Lemma 2.1: *Let $\mathcal{A}\in\mathcal{B}(I\!R^{I\!N})$ be any cylinder set. Then,*

$$\lim_{h\to 0}\lim_{N\uparrow\infty} \mathcal{Q}^\eta_{N,\beta,h}[\omega](\mathcal{A}) = \lim_{h\to 0}\lim_{N\uparrow\infty} \widetilde{\mathcal{Q}}^\eta_{N,\beta,h}[\omega](\mathcal{A}) \tag{2.2}$$

provided that one of the two limits exists.

[2] We are grateful to L. Pastur and M. Shcherbina for discussions on this issue.

A slightly sharper different result will be used to prove Theorem 2.

On the other hand, $\widetilde{\mathcal{Q}}^{\eta}_{N,\beta,h}[\omega](\mathcal{A})$ is absolutely continuous (as a measure on $I\!R^M$) and we get an explicit expression for its density.

Lemma 2.2: *Let $Q^{\eta}_{N,\beta,h}[\omega](x)$ denote the density of $\widetilde{\mathcal{Q}}^{\eta}_{N,\beta,h}[\omega]$ w.r.t. Lebesgue-measure on $I\!R^M$. Then*

$$Q^{\eta}_{N,\beta,h}[\omega](x) = \frac{\exp\left(-\beta N \Phi^{\eta}_{N,\beta,h}[\omega](x)\right)}{\int d^M x \exp\left(-\beta N \Phi^{\eta}_{N,\beta,h}[\omega](x)\right)} \tag{2.3}$$

where

$$\Phi^{\eta}_{N,\beta,h}[\omega](x) = \frac{1}{2}(x - he^{\eta}, x - he^{\eta}) - \frac{1}{\beta N}\sum_{i=1}^{N} \ln\cosh\left(\beta(\xi(\omega)x)_i\right) \tag{2.4}$$

Remark: In (2.4) and in the sequel we introduced some convenient short-hand notations: $(\cdot,\cdot)$ stands for the inner product in $I\!R^M$ and ξ is regarded as a linear map from $I\!R^M$ to $I\!R^N$ (i.e. as an M by N matrix) when acting on a vector $x \in I\!R^M$ (that is, $(\xi(\omega)x)_i \equiv \sum_\mu \xi^\mu_1 x_\mu$). We will also write ξ^t for the transpose of this matrix.

Proof: To prove the lemma, just note that for any $\mathcal{A} \in \mathcal{B}(I\!R^M)$,

$$\begin{aligned}
&\mathcal{N}_{M(N)}(he^{\eta}, [\beta N]^{-1}) \star \mathcal{Q}^{\eta}_{N,\beta,h}[\omega](\mathcal{A}) \\
&= \int d^M y \; \mathcal{N}_{M(N)}(he^{\eta}, [\beta N]^{-1})(\mathcal{A} - y)\mathcal{Q}^{\eta}_{N,\beta,h}[\omega](dy) \\
&= \frac{1}{2^N}\sum_{\sigma\in\mathcal{S}_N} \mathcal{N}_{M(N)}(he^{\eta}, [\beta N]^{-1})(\mathcal{A} - \mathcal{M}_N[\omega](\sigma))\mathcal{G}^{\eta}_{N,\beta,h}[\omega](\sigma) \\
&= \frac{1}{2^N}\sum_{\sigma\in\mathcal{S}_N} \left(\frac{\beta N}{2\pi}\right)^{M/2} \int d^M x \; e^{-\frac{\beta N}{2}(x - \mathcal{M}_N[\omega](\sigma) - he^{\eta})^2}\, 1\!\!\mathrm{I}_{\mathcal{A}}(x)\mathcal{G}^{\eta}_{N,\beta,h}[\omega](\sigma) \\
&= \frac{1}{Z^{\eta}_{N,\beta,h}[\omega]}\left(\frac{\beta N}{2\pi}\right)^{M/2} \int d^M x \; 1\!\!\mathrm{I}_{\mathcal{A}}(x)\frac{1}{2^N}\sum_{\sigma\in\mathcal{S}_N} e^{-\frac{\beta N}{2}\sum_{\mu=1}^{M}\left(x^\mu - m^\mu_N[\omega](\sigma) - h\delta_{\mu\eta}\right)^2} \\
&\qquad\qquad \times e^{\frac{\beta N}{2}\sum_{\mu=1}^{M}\left(m^\mu_N[\omega](\sigma)\right)^2 + \beta h m^{\eta}_N[\omega](\sigma)} \qquad\qquad (2.5)\\
&= \frac{1}{Z^{\eta}_{N,\beta,h}[\omega]}\left(\frac{\beta N}{2\pi}\right)^{M/2} \int d^M x \; 1\!\!\mathrm{I}_{\mathcal{A}}(x) e^{-\frac{\beta N}{2}\sum_{\mu=1}^{M}(x^\mu - h\delta_{\mu\eta})^2} \\
&\qquad\qquad \times \frac{1}{2^N}\sum_{\sigma\in\mathcal{S}_N} e^{\beta\sum_{\mu=1}^{M} x^\mu \sum_{i=1}^{N}\xi^\mu_i(\omega)\sigma_i} \\
&= \frac{1}{Z^{\eta}_{N,\beta,h}[\omega]}\left(\frac{\beta N}{2\pi}\right)^{M/2} \int d^M x \; 1\!\!\mathrm{I}_{\mathcal{A}}(x) e^{-\frac{\beta N}{2}\sum_{\mu=1}^{M}(x^\mu - h\delta_{\mu\eta})^2 + \sum_{i=1}^{N}\ln\cosh\left(\beta\sum_{\mu=1}^{M}\xi^\mu_i(\omega)x^\mu\right)}
\end{aligned}$$

Taking into account that $\tilde{\mathcal{Q}}$ is a probability measure to express Z as an integral, we arrive at the form claimed in Lemma 2.2. $\diamondsuit$

Note that the functional form (2.3) of the density Q with the explicit large parameter N in the exponent suggests naturally that the limiting measure will be concentrated at the minima of the function Φ. The main difficulty we have to deal with here is that the dimension of the domain of Φ_N tends to infinity with N, if $M(N)$ tends to infinity. Otherwise, i.e. if M is bounded, by the strong law of large numbers (SLLN) in Banach spaces ([LT], page 178), $\Phi^\eta_{N,\beta,h}[\omega]$ converges a.s. to the non-random, N-independent function

$$\frac{1}{2}(x-he^\eta,x-he^\eta)-\frac{1}{\beta}\mathbb{E}\ln\cosh\left(\beta(\xi x)_1\right)$$

and the analysis of the limiting measures is a standard exercise in the application of Laplace's method of saddle point integration (see e.g. [AGS1]). In the cases we are interested in, the SLLN cannot be applied (and actually fails), and the analysis of the structure of the minima of the random function Φ is the main problem that we will have to solve in the next section.

We remark here that the weak convergence of the infinite dimensional measures is of course equivalent to the weak convergence of the finite dimensional marginals they induce on $\mathbb{R}^k$, for all $k<\infty$ (see e.g. Billingsley [Bi], page 30). Thus we denote by π_k the natural projection $\pi_k:\mathbb{R}^{\mathbb{N}}\to\mathbb{R}^k$ and by

$$\widetilde{\mathcal{Q}}^{\eta,k}_{N,\beta,h}[\omega]\equiv\widetilde{\mathcal{Q}}^{\eta}_{N,\beta,h}[\omega]\circ\pi_k^{-1}\tag{2.6}$$

the marginals induced on $(\mathbb{R}^k,\mathcal{B}(\mathbb{R}^k))$. Note that for $M(N)\geq k$, these marginals are always absolutely continuous, and their densities are obtained from those given in Lemma 2.2 simply by integrating over the coordinates x^μ with $M\geq\mu>k$. Thus to prove Theorems 1 and 2, we have just to prove the analogous statements for the finite dimensional measures $\widetilde{\mathcal{Q}}^{\eta,k}_{N,\beta,h}[\omega]$ for all finite k.

It will turn out useful to rewrite the function $\Phi^\eta_{N,\beta,h}$ in a somewhat different form. Namely, for arbitrarily chosen δ, adding and subtracting a term $\frac{\delta}{2}(x-he^\eta,\frac{\xi^t\xi}{N}(x-he^\eta))$, (recall that ξ^t stands for the transpose of the $M\times N$-matrix ξ, that is $\xi^t\xi$ is the $M\times M$-matrix whose elements are $(\xi^t\xi)_{\mu\nu}=\sum_{i=1}^N\xi_i^\mu\xi_i^\nu$) we get that

$$\Phi^\eta_{N,\beta,h}[\omega](x)=q_{N,\delta}[\omega](x-he^\eta)+\Psi^\eta_{N,\beta,h,\delta}[\omega]\left(\xi(\omega)x\right)+\frac{\delta h^2}{2}\tag{2.7}$$

where

$$q_{N,\delta}[\omega](x)\equiv\frac{1}{2}\left(x,\left[\mathbb{1}-\delta\frac{\xi^t(\omega)\xi(\omega)}{N}\right]x\right)\tag{2.8}$$

and $\Psi^\eta_{N,\beta,h,\delta}[\omega]$ is a function from $\mathbb{R}^N$ to $\mathbb{R}$ that is simply given by

$$\Psi^\eta_{N,\beta,h,\delta}[\omega](z)\equiv\frac{1}{N}\sum_{i=1}^N\phi_{\beta,h\xi_i^\eta,\delta}(z_i)\tag{2.9}$$

with

$$\phi_{\beta,h,\delta}(y) \equiv \delta\frac{y^2}{2} - \frac{1}{\beta}\ln\cosh(\beta y) - \delta h y \tag{2.10}$$

The point here is to choose the parameter δ in such a way that the quadratic form $q_{N,\delta}[\omega]$ is positive definite with probability tending to one, and to use the fact that $\Psi^{\eta}_{N,\beta,h,\delta}$ is a very simple function whose minima are realized for just those z whose components are the minima of the functions $\phi_{\beta,h\xi_i^\eta,\delta}$. The difficult problem that remains (and that will be studied in the next section) is of course to find those x that are mapped to these z by the random mapping ξ. Before turning to this, let us note that already now we can extract an interesting result on the free energy of the Hopfield model. Namely, if we define the finite volume free energy (we set the external field to zero for convenience)

$$f_{N,\beta}[\omega] \equiv -\frac{1}{\beta N}\ln Z_{N,\beta,h=0}[\omega] \tag{2.11}$$

it is easy to prove the following

Proposition 2.3: *Set $g_\delta(\beta) \equiv \min_{y\in I\!R}\phi_{\beta,\delta}$ and $\alpha \equiv \lim_{N\uparrow\infty}\frac{M(N)}{N}$. Then for all $\beta \geq 0$ and for all $\delta < \frac{1}{1+2\sqrt{\alpha}}$ and for $I\!P$-almost all ω,*

$$g_\delta(\beta) + \frac{\alpha}{2\beta}\ln\left(\delta - 2\sqrt{\alpha}(1-\delta)\right) \leq \liminf_{N\uparrow\infty} f_{N,\beta}[\omega] \leq \limsup_{N\uparrow\infty} f_{N,\beta}[\omega] \leq g_0(\beta) \tag{2.12}$$

In particular, if $\alpha = 0$, the free energy converges almost surely and $\lim_{N\uparrow\infty} f_{N,\beta}[\omega] = g_0(\beta)$ which equals the free energy of the Curie-Weiss model.

Remark: The idea of the proof of this proposition, and in particular to write Φ in the form (2.7) is originally due to Koch [K]. He actually proved (2.12) with $f_{N,\beta}[\omega]$ replaced by $I\!E f_{N,\beta}$, but it is very easy to get rid of the expectation, as has already been shown in [BG]. In the case $\alpha = 0$, a result similar to Proposition 2.3 has also been given by Shcherbina and Tirozzi [ST]; in fact, they prove convergence of the free energy in probability, with bounds that are too weak to conclude almost sure convergence. Although their proof can certainly be improved to yield a.s. convergence, its main drawback is that it is unnecessarily complicated.

In fact, the proof of Proposition 2.3 is immediate once we know that $q_{N,\delta}$ is positive definite for $\delta < \frac{1}{1+2\sqrt{\alpha}}$, with probability sufficiently close to one. This information is contained in the following theorem on maximal eigenvalues of random matrices.

Theorem 2.4: *Assume that ξ_i^μ are i.i.d. random variables satisfying $I\!E\xi_i^\mu = 0$ and $I\!E\left(\xi_i^\mu\right)^k \leq 1$, for all $k > 1$. Let B denote the $M \times M$-matrix whose elements are*

$$B_{\mu\nu} \equiv (1-\delta_{\mu,\nu})\frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu\xi_i^\nu \tag{2.13}$$

Then, for any $c \geq 0$, $M \leq N$ and for N sufficiently large

$$\mathbb{P}\left(\|B\| > \sqrt{\alpha}(2+\sqrt{\alpha}) + cN^{-1/6}\ln N\right) \leq NN^{-c/\sqrt{\alpha}(2+\sqrt{\alpha})} \tag{2.14}$$

where $\alpha \equiv \frac{M}{N}$.

This theorem will be used again in the next section, and we will give a simple proof in an appendix. The proof of Proposition 2.3 with the help of Theorem 2.4 is left as an exercise (or see [BG]).

In the remainder of this section we state some properties of the function $\Psi^{\eta}_{N,\beta,h,\delta}[\omega]$. They follow in fact from the following lemma on the function $\phi_{\beta,h,\delta}$:

Lemma 2.5: *The function $\phi_{\beta,h,\delta} : \mathbb{R} \to \mathbb{R}$ has the following properties:*

(i) For all β, δ and h, $\phi_{\beta,-h,\delta}(y) = \phi_{\beta,h,\delta}(-y)$ for all $y \in \mathbb{R}$.

(ii) Assume that $h \geq 0$ and let $a(\delta,\beta,h)$ denote the largest solution of the equation $\delta(y-h) = \tanh(\beta y)$ (the case $h \leq 0$ follows by applying (i)). Then

(ii.1) If $0 < \beta \leq \delta$, for all $h \geq 0$, $\phi_{\beta,h,\delta}$ has a unique minimum at $y = a(\delta,\beta,h)$. $a(\delta,\beta,h)$ is a continuous function of h and $\lim_{h\to 0} a(\delta,\beta,h) = 0$.

(ii.2) If $\beta > \delta$, $h > 0$ and h sufficiently small, then $\phi_{\beta,h,\delta}$ has two local minima, but a unique global minimum which is taken on at the point $a(\delta,\beta,h) > 0$.

(ii.3) If $\delta < \beta < \infty$ and $h = 0$, then $\phi_{\beta,h=0,\delta}$ has two global minima denoted by $a^{\pm}(\delta,\beta)$, which are the strictly positive and negative solutions of $\delta y = \tanh(\beta y)$. Moreover, $a^-(\delta,\beta) = -a^+(\delta,\beta)$ and $\lim_{h\to 0+} a(\delta,\beta,h) = a^+(\delta,\beta)$. Note also that $a^+(\delta,\beta) \to +1/\delta$, as $\beta \uparrow \infty$.

(iii) For $\beta \geq \delta$ and $h \geq 0$ sufficiently small, there exist strictly positive constants, $\infty > c^+(\delta,\beta,h) \geq c^-(\delta,\beta,h) > 0$, such that for all $y \in \mathbb{R}$,

$$\phi_{\beta,h,\delta}(y) - \inf_{y\in\mathbb{R}} \phi_{\beta,h,\delta}(y) \geq \frac{c^-(\delta,\beta,h)}{2} \inf_{y_0\in\{-a(\delta,\beta,h),a(\delta,\beta,h)\}} (y-y_0)^2 \tag{2.15}$$

and

$$\phi_{\beta,h,\delta}(y) - \inf_{y\in\mathbb{R}} \phi_{\beta,h,\delta}(y) \leq \frac{c^+(\delta,\beta,h)}{2}(y - a(\delta,\beta,h))^2 \tag{2.16}$$

(iv) For all δ_+ and $\beta \geq \delta_+$, $h \geq 0$ small enough there exists a finite constant $c > 0$ such that for all $\delta_- \leq \delta_+$,

$$\left|\inf_{y\in\mathbb{R}} \phi_{\beta,h,\delta_+}(y) - \inf_{y\in\mathbb{R}} \phi_{\beta,h,\delta_-}(y)\right| \leq |\delta_+ - \delta_-| c_3 a(\delta_-,\beta,h) \tag{2.17}$$

The proof of this lemma is of course just a standard exercise in real analysis. For $\delta = 1$ it can be found e.g. in the book by Ellis [E]; the modifications necessary to accommodate $\delta \neq 1$ are trivial; note in particular that $a(\delta, \beta, h) = a(1, \beta/\delta, h\delta)/\delta$.

Let us define the sets

$$\mathcal{E}_N(a) \equiv \left\{ z \in I\!R^N \, : \forall_i z_i \in \{-a, a\} \right\} \tag{2.18}$$

Taking into account the definition of $\Psi^\eta_{N,\beta,\delta,h}$ given by (2.9) we immediately have the following

Corollary 2.6: *With the notation from Lemma 2.5 we have that*

(i) For all $1 \leq \eta \leq M(N)$ and for all $h > 0$, $\Psi^\eta_{N,\beta,\delta,h}$ has a unique global minimum which is realized for

$$z = a(\delta, \beta, h)\xi^\eta \tag{2.19}$$

(ii) If $h = 0$, then $\Psi^\eta_{N,\beta,\delta,h=0}$ takes on its absolute minimum for all points $z \in \mathcal{E}_N(a^+(\delta, \beta))$.

(iii) For all $h \geq 0$ and for $\beta > \delta$, we have that

$$\Psi^\eta_{N,\beta,\delta,h}(z) - \inf_{y \in I\!R} \phi_{\beta,h,\delta}(y) \leq \frac{c^+(\delta, \beta, h)}{2N} \|z - a(\delta, \beta, h)\xi^\eta\|^2 \tag{2.20}$$

and

$$\Psi^\eta_{N,\beta,\delta,h}(z) - \inf_{y \in I\!R} \phi_{\beta,h,\delta}(y) \geq \frac{c^-(\delta, \beta, h)}{2N} \inf_{\zeta \in \mathcal{E}_N(a(\delta,\beta,h))} \|z - \zeta\|^2 \tag{2.21}$$

The lower bound (2.21) will be used in the next section to obtain probabilistic bounds on the minima of the function Φ.

III. The global minima of the function Φ

The present section contains the central estimates of this paper. In particular we will localize the positions of the absolute minima of our function $\Phi^\eta_{N,\beta,h}[\omega]$ and control its behaviour near them. All these results will hold with probability tending to one fast enough. To simplify our notation, we will drop in this section the arguments ω in all random functions, as well as indices referring to the system size N, whenever no confusion may arise. In particular, we will write simply ξ for the matrix $\xi|_N$.

From the results of the preceeding section, in particular Corollary 2.6, it is clear that the minima of this function should be near those points $x \in I\!R^M$ for which $\xi x \in \mathcal{E}(a)$, where $a = a(\delta, \beta, h)$. Of course, we can immediately identify $2M$ such points, namely $x^{(\nu,s)} \equiv sa(\delta_{\nu 1}, \ldots, \delta_{\nu\mu}, \ldots, \delta_{\nu M}) \in$

$I\!R^M$, with $\nu \in \{1, \dots, M\}$ and $s \in \{-1, +1\}$. The first result of this section is that if M/N is sufficiently small, all possible solutions must be very close to these. The precise formulation of this statement is the following

Proposition 3.1: *For $\rho \geq 0$, define the set*

$$\mathcal{R}_\rho \equiv \left\{ x \in I\!R^M : \inf_{\nu,s} \|x - x^{(\nu,s)}\| > \rho \right\}$$

Given a, $0 < a \leq 1$, there exists strictly positive constants $\alpha_0, \epsilon_0, c_1$, and c_2 such that if $\frac{M}{N} \leq \alpha_0$, then for all $\epsilon \leq \epsilon_0$ and $\rho^2 \geq a^2 \left(c_1 \left(\frac{\epsilon}{a^2} \log \frac{128}{\sqrt{\epsilon}} \right)^{1/4} + 2\sqrt{\frac{M}{N}} \right)$

$$I\!P\left(\exists_{x \in \mathcal{R}_\rho} \inf_{z \in \mathcal{E}(a)} \frac{1}{N} \|\xi x - z\|^2 < \epsilon \right) \leq e^{-Nc_2 \left(\frac{\epsilon}{a^2} \log \frac{128}{\sqrt{\epsilon}} \right)^{1/4}} + p(N) \tag{3.1}$$

where $p(N)$ is such that $\sum_{N=1}^{\infty} p(N) < \infty$.

In particular, if $\lim_{N\uparrow\infty} \frac{M}{N} = 0$, there exists strictly positive constants ϵ_0 and c_1 such that if $0 < \epsilon \leq \epsilon_0$ and $\rho^2 > a^{3/2} c_1 (\epsilon \log \frac{128}{\sqrt{\epsilon}})^{1/4}$ then

$$I\!P\left(\exists_{x \in \mathcal{R}_\rho} \inf_{z \in \mathcal{E}(a)} \frac{1}{N} \|\xi x - z\|^2 < \epsilon \quad i.o. \right) = 0 \tag{3.2}$$

Proof: The first step in the proof of (3.1) consists in getting an a priori estimate on the modulus of those x for which the event considered may possibly occur. Here we will make use of Theorem 2.4. Note first that by the Schwarz inequality, $\|\xi x - z\|^2 < \epsilon$ implies that, since $\|z\| = a\sqrt{N}$,

$$a - \sqrt{\epsilon} \leq \frac{1}{\sqrt{N}} \|\xi x\| \leq a + \sqrt{\epsilon} \tag{3.3}$$

By Theorem 2.4, we may choose $r^*(M,N) = 2\sqrt{\frac{M}{N}} + O(N^{-1/6} \ln N)$ such that on a subset $\mathcal{A}_N \subset \Omega$, satisfying $I\!P(\mathcal{A}_N) \geq 1 - \hat{p}(N)$ with $\sum_N \hat{p}(N) < \infty$,

$$\left\| 1\!\!1 - \frac{\xi^T \xi}{N} \right\| \leq r^*(M,N) \tag{3.4}$$

Therefore, (3.3) implies that if $\omega \in \mathcal{A}_N$, then the event in (2.1) can only be realized for x satisfying

$$\frac{a - \sqrt{\epsilon}}{\sqrt{1 + r^*}} \leq \|x\| \leq \frac{a + \sqrt{\epsilon}}{\sqrt{1 - r^*}} \tag{3.5}$$

Define

$$\mathcal{R}_\epsilon \equiv \left\{ x \in I\!R^M : \frac{a - \sqrt{\epsilon}}{\sqrt{1 + r^*}} \leq \|x\| \leq \frac{a + \sqrt{\epsilon}}{\sqrt{1 - r^*}} \right\} \tag{3.6}$$

and set $\mathcal{R}_{\epsilon,\rho} \equiv \mathcal{R}_\epsilon \cap \mathcal{R}_\rho$. By the above argument, we may conclude that

$$I\!P\left(\exists_{x\in\mathcal{R}_\rho} \inf_{z\in\mathcal{E}(a)} \frac{1}{N}\|\xi x - z\|^2 < \epsilon\right) \le \hat{p}(N) + I\!P\left(\exists_{x\in\mathcal{R}_{\rho,\epsilon}} \inf_{z\in\mathcal{E}(a)} \frac{1}{N}\|\xi x - z\|^2 < \epsilon, \mathcal{A}_N\right) \tag{3.7}$$

Although in the left hand side of (3.7) we have considerably reduced the range of x-values we need to control, the fact that we need to control an event for a continuous set of points x still poses a problem. The idea to overcome this difficulty is to regard the inequality $\|\xi x - z\|^2 < \epsilon N$ as a set of N approximate equations and to use the first $[\gamma N]$ of them to fix x up to a small error. Thus, given $\gamma \in]0,1]$ we decompose the $M \times N$ matrix ξ into an $M \times [\gamma N]$ matrix $\overline{\xi}$, $(\overline{\xi}_i^\mu \equiv \xi_i^\mu; i \in \{1,\dots,[\gamma N]\}, \mu \in \{1,\dots,M\})$ and an
$M \times (N-[\gamma N])$ matrix $\underline{\xi}$, $(\underline{\xi}_i^\mu \equiv \xi_{i+[\gamma N]}^\mu; i \in \{1,\dots,N-[\gamma N]\}, \mu \in \{1,\dots,M\})$. Using Theorem 2.4 just as before, we see that we may choose $r^*(M,\gamma N) = 2\sqrt{\frac{M}{\gamma N}} + O(N^{-1/6}\ln N)$ and $r^*(M,(1-\gamma)N) = 2\sqrt{\frac{M}{(1-\gamma)N}} + O(N^{-1/6}\ln N)$, such that on a subset $\Omega \supset \tilde{\mathcal{A}}_N \supset \mathcal{A}_N$ satisfying $I\!P(\tilde{\mathcal{A}}_N) \ge 1-\tilde{p}(N)$ with $\sum_N \tilde{p}(N) < \infty$ we have

$$\begin{aligned} \left\|1\!\!1 - \frac{\overline{\xi}^T\overline{\xi}}{\gamma N}\right\| &\le r^*(M,\gamma N) \\ \left\|1\!\!1 - \frac{\underline{\xi}^T\underline{\xi}}{(1-\gamma)N}\right\| &\le r^*(M,(1-\gamma)N) \end{aligned} \tag{3.8}$$

Let us also define

$$\overline{\mathcal{E}}(a) \equiv \left\{z \in I\!R^{[\gamma N]} : \forall i \in \{1,\dots,[\gamma N]\}, z_i \in \{-a,a\}\right\},$$

$$\underline{\mathcal{E}}(a) \equiv \left\{z \in I\!R^{N-[\gamma N]} : \forall i \in \{1,\dots,N-[\gamma N]\}, z_i \in \{-a,a\}\right\}$$

With this notation we get that

$$\begin{aligned} &I\!P\left(\exists_{x\in\mathcal{R}_{\rho,\epsilon}} \inf_{z\in\mathcal{E}(a)} \frac{1}{N}\|\xi x - z\|^2 < \epsilon, \mathcal{A}_N\right) \le \\ &I\!P\left(\exists_{x\in\mathcal{R}_{\rho,\epsilon}} \left\{\inf_{\overline{z}\in\overline{\mathcal{E}}(a)} \|\overline{\xi}x - \overline{z}\|^2 < \epsilon N\right\} \cap \left\{\inf_{\underline{z}\in\underline{\mathcal{E}}(a)} \|\underline{\xi}x - \underline{z}\|^2 < \epsilon N\right\}, \tilde{\mathcal{A}}_N\right) + \tilde{p}(N) \\ &\le \sum_{\overline{z}\in\overline{\mathcal{E}}(a)} I\!P\left(\exists_{x\in\mathcal{R}_{\rho,\epsilon}}\left\{\|\overline{\xi}x - \overline{z}\|^2 < \epsilon N\right\} \cap \left\{\inf_{\underline{z}\in\underline{\mathcal{E}}(a)} \|\underline{\xi}x - \underline{z}\|^2 < \epsilon N\right\}, \tilde{\mathcal{A}}_N\right) + \tilde{p}(N) \end{aligned} \tag{3.9}$$

Given $\overline{z} \in \overline{\mathcal{E}}(a)$ let us choose $x_0 = x_0(\overline{z}) \in \mathcal{R}_{\rho,\epsilon}$ such that $\|\overline{\xi}x_0 - \overline{z}\|^2 < \epsilon N$. Notice that if such an x_0 does not exist, then the set $\{\exists_{x\in\mathcal{R}_{\rho,\epsilon}}; \|\overline{\xi}x - \overline{z}\|^2 < \epsilon N\}$ is empty. By construction, $x_0(\bar{z})$ is a random variable which is $\sigma(\overline{\xi})$-mesurable. We will now show that all points x that verify

$\|\bar{\xi}x - \bar{z}\|^2 < \epsilon N$ are close to x_0. Set $x \equiv x_0(\overline{z}) + \delta x$. using the Schwarz inequality, we see that if $\|\bar{\xi}x - \bar{z}\|^2 < \epsilon N$ and $\|\bar{\xi}x_0 - \bar{z}\|^2 < \epsilon N$, then

$$\|\overline{\xi}\delta x\| \leq \|\overline{\xi}(x_0 + \delta x) - \overline{z}\| + \|\overline{z} - \overline{\xi}x_0\| \leq 2\sqrt{\epsilon N} \tag{3.10}$$

On the other hand, on $\tilde{\mathcal{A}}_N$ we have that

$$\frac{\|\overline{\xi}\delta x\|^2}{\gamma N} = \left(\delta x, \frac{\overline{\xi}^T\overline{\xi}}{\gamma N}\delta x\right) \geq \|\delta x\|^2(1 - r^*(M, \gamma N)) \tag{3.11}$$

Therefore, on $\tilde{\mathcal{A}}_N$, (3.10) implies

$$\|\delta x\|^2 \leq \frac{4\epsilon}{\gamma(1 - r^*(M, \gamma N))} \equiv \hat{\epsilon} \tag{3.12}$$

and so

$$\begin{aligned}
&I\!P\left(\exists_{x\in\mathcal{R}_{\rho,\epsilon}}\left\{\|\overline{\xi}x - \overline{z}\|^2 < \epsilon N\right\} \cap \left\{\inf_{\underline{z}\in\underline{\mathcal{E}}(a)} \|\underline{\xi}x - \underline{z}\|^2 < \epsilon N\right\}, \tilde{\mathcal{A}}_N\right)\\
&\leq I\!P\left(\exists_{\delta x: x_0+\delta x\in\mathcal{R}_{\rho,\epsilon}}\left\{\|\overline{\xi}(x_0 + \delta x) - \overline{z}\|^2 < \epsilon N\right\} \cap \left\{\inf_{\underline{z}\in\underline{\mathcal{E}}(a)} \|\underline{\xi}(x_0 + \delta x) - \underline{z}\|^2 < \epsilon N\right\}, \tilde{\mathcal{A}}_N\right)\\
&\leq I\!P\left(\exists_{\delta x: x_0+\delta x\in\mathcal{R}_{\rho,\epsilon}\,\wedge\,\|\delta x\|^2\leq\hat{\epsilon}} \inf_{\underline{z}\in\underline{\mathcal{E}}(a)} \|\underline{\xi}(x_0 + \delta x) - \underline{z}\|^2 < \epsilon N, \tilde{\mathcal{A}}_N\right)\\
&\leq I\!P\left(\inf_{\underline{z}\in\underline{\mathcal{E}}(a)} \|\underline{\xi}x_0 - \underline{z}\|^2 < 2N\epsilon\hat{\gamma}, \tilde{\mathcal{A}}_N\right)
\end{aligned} \tag{3.13}$$

where in the last line we have used that on the set $\tilde{\mathcal{A}}_N$, if $\|\underline{\xi}(x_0 + \delta x) - \underline{z}\|^2 \leq \epsilon N$ and $\|\delta x\|^2 \leq \hat{\epsilon}$ are satisfied,

$$\begin{aligned}
\|\underline{\xi}x_0 - \underline{z}\|^2 &\leq \left(\sqrt{\epsilon N} + \sqrt{\hat{\epsilon}(1-\gamma)N(1 + r^*(M, (1-\gamma)N))}\right)^2\\
&\leq 2N\left[\epsilon + \hat{\epsilon}(1-\gamma)N(1 + r^*(M, (1-\gamma)N))\right]\\
&= 2N\epsilon\left[1 + \frac{4(1-\gamma)(1 + r^*(M, (1-\gamma)N))}{\gamma(1 - r^*(M, \gamma N))}\right]\\
&\equiv 2N\epsilon\hat{\gamma}.
\end{aligned} \tag{3.14}$$

To estimate the last line in (3.13) we will use the following elementary observation:

Lemma 3.2: *If $(Y_i)_{i=1}^K$ is a family of positive random variables and $C \geq 2$ then*

$$\left\{\frac{1}{K}\sum_{i=1}^K Y_i \leq \epsilon\right\} \subset \left\{\#\left\{i; Y_i \leq C\epsilon\right\} > (1 - \frac{1}{C})K\right\} \tag{3.15}$$

Using this Lemma for C to be chosen later, we get

$$\begin{aligned}
&I\!P\left(\inf_{\underline{z}\in\underline{\mathcal{E}}(a)} \|\underline{\xi}x_0 - \underline{z}\|^2 < 2N\epsilon\hat{\gamma}, \tilde{\mathcal{A}}_N\right)\\
&\leq I\!P\left(\exists_{I:|I|>(1-\frac{1}{C})N(1-\gamma)} \forall_{i\in I}\left[\left\{|(\underline{\xi}x_0)_i - a|^2 \leq 2C\epsilon\hat{\gamma}\right\} \cup \left\{|(\underline{\xi}x_0)_i + a|^2 \leq 2C\epsilon\hat{\gamma}\right\}\right]\right)\\
&\leq \binom{(1-\gamma)N}{(1-\gamma)\frac{N}{C}}\left[I\!P\left(\left\{|(\underline{\xi}x_0)_1 - a|^2 \leq 2C\epsilon\hat{\gamma}\right\} \cup \left\{|(\underline{\xi}x_0)_1 + a|^2 \leq 2C\epsilon\hat{\gamma}\right\}\right)\right]^{(1-\gamma)(1-\frac{1}{C})N}
\end{aligned} \tag{3.16}$$

where in the last step, we have used that, for any subset I of $\{1, \ldots, N-[\gamma N]\}$, given $\overline{\xi}$, the random variables $((\underline{\xi} x_0)_i)_{i \in I}$ form a family of independent identically distributed random variables. This is true given $\overline{\xi}$, since x_0 is a random variable which is $\sigma(\overline{\xi})$-measurable by construction. Thus we are left to bound the probability in the right hand side of (3.16). Let us state this as a Lemma:

Lemma 3.3: *There exists $v \equiv v(\epsilon, a, \gamma, C, \frac{M}{N}) > 0$ such that:*

$$\mathbb{P}\left(\left\{|(\underline{\xi} x_0)_1 - a|^2 \leq 2C\epsilon\hat{\gamma}\right\} \cup \left\{|(\underline{\xi} x_0)_1 + a|^2 \leq 2C\epsilon\hat{\gamma}\right\}\right) \leq e^{-\frac{1}{4}\left(\frac{v^2}{2} - 2\frac{\epsilon'}{a}\right)} \tag{3.17}$$

where $\epsilon' = \sqrt{2C\epsilon\hat{\gamma}}$.

Proof: Without loss of generality we can assume that $x_0 \equiv (x_0^i)_{i=1}^M$ satisfies:
$x_0^1 \geq x_0^2 \geq \ldots \geq x_0^M \geq 0$. This follows since x_0 is $\sigma(\overline{\xi})$-measurable and since for any $i \in \{1, \ldots, N-[\gamma N]\}$ the random variables $(\underline{\xi}_i^\mu)_{\mu=1}^M$ that appear in $(\underline{\xi} x_0)_i = \sum_{\mu=1}^M \underline{\xi}_i^\mu x_0^\mu$ are independent of $\overline{\xi}$. Moreover given x_0, the random variables $(\underline{\xi} x_0)_i$ for $i \in \{1, \ldots, N-[\gamma N]\}$ have the same distribution as $\sum_{\mu=1}^m \xi_i^\mu |x_0^\mu|$. Given $0 < v < 1$, we define the $\sigma(\overline{\xi})$-measurable random variable, t, by

$$t \equiv \sup\left\{s > 0 : \sum_{\mu=1}^{s} (x_0^\mu)^2 \leq a^2(1-v^2)\right\} \tag{3.18}$$

if such a t exists. To ensure the existence of t we will have to impose conditions on our parameters, which we derive now. Obviously, t exists if and only if $(x_0^1)^2 \leq a^2(1-v^2)$. Since $x_0 \in \mathcal{R}_{\rho,\epsilon}$,

$$\begin{aligned} \rho^2 \leq \|x_0 - x^{(1,1)}\|^2 &= (x_0^1 - a)^2 + \sum_{\mu \geq 2} (x_0^\mu)^2 \\ &= a^2 - 2x_0^1 a + \|x_0\|^2 \\ &\leq -2a x_0^1 + a^2 + a^2(1+\hat{\epsilon}) \end{aligned} \tag{3.19}$$

where we have set $1 + \hat{\epsilon} \equiv \frac{(1+\sqrt{\frac{\epsilon}{a^2}})^2}{1-r^*(M,N)}$. Solving this inequality for x_0^1, we see that this condition is certainly satisfied if

$$\rho^2 \geq a^2\left[1 - 2\sqrt{1-v^2} + 1 + \hat{\epsilon}\right] \tag{3.20}$$

Let us now show that if (3.20) holds, then

$$x_0^{t+1} \leq a\sqrt{\frac{1+\hat{\epsilon}}{2}} \tag{3.21}$$

Obviously, under our assumptions

$$(t+1)\left(x_0^{t+1}\right)^2 \leq \sum_{\mu=1}^{t+1} (x_0^\mu)^2 \leq \|x_0\|^2 \leq a^2(1+\hat{\epsilon}) \tag{3.22}$$

and since (3.20) implies that $t \geq 1$, (3.21) follows.

We assume from now on (3.20) and define the random variables

$$
\begin{aligned}
X_t &\equiv \sum_{\mu=1}^{t} x_0^\mu \xi_1^\mu \\
Y_t &\equiv \sum_{\mu=t+2}^{m} x_0^\mu \xi_1^\mu \\
Z_t &\equiv x_0^{t+1} \xi_1^{t+1}.
\end{aligned}
\tag{3.23}
$$

It follows from (3.23) and the definition of t that these random variables satisfy

$$
\begin{aligned}
I\!E X_t^2 &\leq a^2(1-v^2) \\
I\!E Y_t^2 &\leq a^2(\hat{\epsilon}+v^2) \\
|Z_t| &\leq a\sqrt{\frac{1+\hat{\epsilon}}{2}}
\end{aligned}
\tag{3.24}
$$

It remains to estimate

$$
I\!P\left(X_t+Y_t+Z_t \in [-a-\epsilon', -a+\epsilon'] \cup [a-\epsilon', a+\epsilon']\right)
$$

with $\epsilon' \equiv \sqrt{2C\epsilon\hat{\gamma}}$. Let us introduce the random ($\sigma(\overline{\xi})$-measurable) intervals $I^+ \equiv [a - x_0^{t+1} - \epsilon', a - x_0^{t+1} + \epsilon']$ and $I^- \equiv [-a - x_0^{t+1} - \epsilon', -a - x_0^{t+1} + \epsilon']$. notice that since ξ_1^{t+1}, X_t and Y_t are symmetric random variables we have:

$$
I\!P\left(X_t+Y_t+Z_t \in [-a-\epsilon', -a+\epsilon'] \cup [a-\epsilon', a+\epsilon']\right) = I\!P\left(X_t+Y_t \in I^- \cup I^+\right) \tag{3.25}
$$

By conditioning on the events $\{X_t \geq 0, Y_t > 0\}$, $\{X_t \geq 0, Y_t \leq 0\}$, $\{X_t < 0, Y_t \leq 0\}$ and $\{X_t < 0, Y_t > 0\}$ and using the trivial upper bound 1 for the conditional expectations in the first three cases we obtain:

$$
\begin{aligned}
&I\!P\left(X_t+Y_t \in I^- \cup I^+\right) \leq \frac{1}{4}\left\{3 + I\!P\left(X_t+Y_t \in I^- \cup I^+ | X_t < 0, Y_t > 0\right)\right\} \\
&\leq \frac{3}{4} + \frac{1}{4} I\!P\left(X_t+Y_t \in I^+ | X_t < 0, Y_t > 0\right) + \frac{1}{4} I\!P\left(X_t+Y_t \in I^- | X_t < 0, Y_t > 0\right)
\end{aligned}
\tag{3.26}
$$

where we have chosen ϵ' small enough in order that $I^+ \cap I^- = \emptyset$. The exponential Markov inequality and (3.24) yield the bound

$$
\begin{aligned}
I\!P\left(X_t+Y_t \in I^+ | X_t < 0, Y_t > 0\right) &\leq I\!P\left(Y_t > a - x_0^{t+1} - \epsilon' | Y_t > 0\right) \\
&\leq 2\exp\left(-\frac{\left(1-\sqrt{(1+\hat{\epsilon})/2} - \epsilon'/a\right)^2}{2(\hat{\epsilon}+v^2)}\right)
\end{aligned}
\tag{3.27}
$$

Because the X_t are symmetric r.v.'s, the Chebyshev inequality can be used to get the bound

$$
I\!P\left(X_t+Y_t \in I^- | X_t < 0, Y_t > 0\right) \leq I\!P\left(X_t < -a - x_o^{t+1} + \epsilon' | X_t < 0\right) \leq \frac{(1-v^2)a^2}{(a-\epsilon')^2} \tag{3.28}
$$

If finally v is chosen such that

$$2\exp\left(-\frac{\left(1-\sqrt{(1+\hat\epsilon)/2}-\epsilon'/a\right)^2}{2(\hat\epsilon+v^2)}\right)\le\frac{1}{2}\frac{v^2a^2}{(a-\epsilon')^2} \tag{3.29}$$

and if ϵ' is small enough, combining (3.27) and (3.28) we arrive at

$$I\!P\left(X_t+Y_t\in I^-\cup I^+\right)\le 1-\frac{1}{4}\left(\frac{v^2}{2}-2\frac{\epsilon'}{a}\right)\le e^{-\frac{1}{4}\left(\frac{v^2}{2}-2\frac{\epsilon'}{a}\right)} \tag{3.30}$$

which proves the Lemma. $\diamondsuit$

We now continue the proof of the proposition. Inserting the bound from Lemma 2.3 into (3.16) it follows that

$$\begin{aligned}
&I\!P\left(\exists_{x\in\mathcal{R}_{\rho,\epsilon}}\inf_{z\in\mathcal{E}(a)}\frac{1}{N}\|\xi x-z\|^2<\epsilon,\mathcal{A}_N\right)\\
&\le 2^{\gamma N}\binom{(1-\gamma)N}{(1-\gamma)\frac{N}{C}}e^{-\frac{1}{4}\left(\frac{v^2}{2}-2\frac{\epsilon'}{a}\right)(1-\gamma)(1-\frac{1}{C})N}+\hat p(N)\\
&\le e^{N\left[\gamma\ln 2+(1-\gamma)\left(\frac{1}{C}\ln C+\frac{C-1}{C}\ln\frac{C}{C-1}\right)-\frac{1}{4}\left(\frac{v^2}{2}-2\frac{\epsilon'}{a}\right)(1-\gamma)(1-\frac{1}{C})\right]}+\hat p(N)
\end{aligned} \tag{3.31}$$

where the factor $2^{\gamma N}$ takes into account the sum over $\bar z$, and in the last line we have made use of Stirling's formula [Ro] to bound the binomial coefficient. We want show that under the conditions on ϵ and ρ stated in the proposition, we can choose the parameters C,γ and v such that the coefficient of N in the exponent in the last line of (3.31) is negative. Obviously, this requires first of all that $\frac{v^2}{2}-2\frac{\epsilon'}{a}$ be positive and sufficiently large. It is thus reasonable to fix $v^2=8\epsilon'/a$. Given this choice, we then choose γ and C such that, the positive terms in the exponential are balanced by, say, one half of the negative terms. It is easy to verify that this is true with $C=\frac{256}{v^2}\log\frac{128}{v^2}$ and $\gamma=\frac{v^2}{64}$, so that with this choice we get

$$I\!P\left(\exists_{x\in\mathcal{R}_{\rho,\epsilon}}\inf_{z\in\mathcal{E}(a)}\frac{1}{N}\|\xi x-z\|^2<\epsilon,\mathcal{A}_N\right)\le e^{-Nv^2/64}+\hat p(N) \tag{3.32}$$

Recalling that ϵ' still depends on γ and C, we obtain as explicit relation between v and ϵ, for our choices

$$\epsilon=\frac{a^2v^8}{2^{22}\ln\frac{128}{v^2}}\quad\text{or}\quad v^2=\left(\frac{2^{22}\epsilon}{a^2}\ln\frac{a^2}{2^{15}\epsilon}\right)^{1/4} \tag{3.33}$$

where the constants 2^l have not to be taken too seriously. To conclude the proof, we only have to verify that under the hypothesis on ρ and ϵ stated with these choices Eqs. (3.20) and (3.29) are satisfied. To make the analysis of these conditions transparent, we will assume and use that α and ϵ/a^2 are small compared to one and keep only the leading terms in these quantities. This means that in particular $\hat\epsilon$ (defined after (3.19)) is approximately given by

$$\hat\epsilon\approx 2\sqrt{\frac{\epsilon}{a^2}}+2\sqrt{\frac{M}{N}} \tag{3.34}$$

Using this, inserting (3.34) into (3.20) we get the final bound on ρ as

$$\rho^2 \geq a^2 \left[\left(\frac{2^{22}\epsilon}{a^2} \ln \frac{a^2}{2^{15}\epsilon} \right)^{1/4} + 2\sqrt{\frac{M}{N}} \right] \tag{3.35}$$

Proceeding in the same way with (3.29), one finds that it is satisfied if

$$\frac{1}{4} \ln \frac{\epsilon}{a^2} \left(\left(\frac{2^{22}\epsilon}{a^2} \ln \frac{a^2}{2^{15}\epsilon} \right)^{1/4} + 4\sqrt{\frac{M}{N}} \right) \leq \left(1 - \frac{1}{\sqrt{2}} \right)^2 \tag{3.36}$$

This gives in general an upper bound on ϵ, and if $\alpha \neq 0$ also seems to imply a lower bound on ϵ. This is, however, only due to our choice of v^2 which was done as to allow ϵ to be as large as possible and it is easy to verify that Proposition 3.1 really holds for all $\epsilon \leq \epsilon_0$, where ϵ_0 is determined by (3.36). However, as we will see later, in our applications ϵ will have to be taken larger than $\sqrt{\alpha}$, (3.36) will in particular be the reason for an upper bound on admissible values of α. This ends the proof of Proposition 3.1. $\diamondsuit$

Proposition 3.1 will be used directly in the next section to prove our main theorems.

IV. Proof of the Theorems

In this chapter we will prove the theorems announced in Section I using the results obtained in the last two sections. Our strategy is to first prove the analogues of Theorems 1 and 2 for the measures $\tilde{\mathcal{Q}}$. Statement (i) of Theorem 1 then follows by Lemma 2.1, while Theorem 2 will be the issue of a short computation, with statement (ii) of Theorem 1 essentially a simple Corollary. Finally we will proof Theorem 3. We find it convenient to work with the Laplace transforms of the finite dimensional marginals of the measures $\tilde{\mathcal{Q}}^{\eta}_{N,\beta,h}$ which we denote by $\mathcal{L}^{\eta,k}_{N,\beta,h}[\omega](t)$, i.e.

$$\begin{aligned} \mathcal{L}^{\eta,k}_{N,\beta,h}[\omega](t) &= \int_{I\!R^k} d^k x \; e^{(t,x)} Q^{\eta,k}_{N,\beta,h}[\omega](x) \\ &= \int_{I\!R^M} d^M x \; e^{(t,\pi_k x)} Q^{\eta}_{N,\beta,h}[\omega](x) \end{aligned} \tag{4.1}$$

Recall that $a^+(\beta)$ denotes the largest solution of $a = \tanh(\beta a)$ and $\pi_k : I\!R^{I\!N} \to I\!R^k$ is the projection on the first k coordinates: $x \mapsto \pi_k(x) = (x_1, \dots, x_k)$. Throughout this section we will write $r^* \equiv r^*(M,N)$ for the upper bound on $\|\mathbb{1} - \frac{\xi^t \xi}{N}\|$ on $\mathcal{A}_N$ (see (3.4)). Also we will write $\overline{a}$ shorthand for $a(\delta, \beta, h)$, unless we want to point explicitly to the dependence of this quantity on the parameters.

Proposition 4.1: *Under the hypothesis of Theorem 1, $\forall \eta \in I\!N, \forall k \in I\!N$ and $\forall t \in I\!R^k$, for almost all ω we have*

$$\lim_{h \downarrow 0} \lim_{N \uparrow \infty} \mathcal{L}^{\eta,k}_{N,\beta,h}[\omega](t) = e^{a^+(\beta)(t,\pi_k e^{\eta})} \tag{4.2}$$

To prove the Proposition 4.1 we need the following lemma which in fact provides all the crucial estimates needed to prove our theorems.

Lemma 4.2: *For almost all ω, for all but a finite number of indices N we have*

(i) $\forall (\nu, s) \neq (\eta, 1)$

$$0 \leq \widetilde{\mathcal{Q}}^{\eta}_{N,\beta,h}[\omega]\left(B^{(\nu,s)}_{\rho}\right) \leq e^{-\bar{c}_1 \beta N} \tag{4.3}$$

where

$$\bar{c}_1 = h\bar{a} - \rho\,(2h + 4|\bar{a} - h|r^*) + \rho^2\left(2h + (1 + r^*)c^+(\delta, \beta, h)\right) \tag{4.4}$$

and $c^+ = c^+(\delta, \beta, h)$ is the same constant as in Lemma 2.5.

(ii) Let $\frac{M}{N}$, ρ and ϵ are such that the hypothesis of Theorem 1 are satisfied. Set $\delta_- = \frac{1-\frac{M}{N}}{1+r^*}$ and $\delta_+ = \frac{1}{1-r^*}$. Then, $\forall t \in I\!R^k$

$$0 \leq \int_{\mathcal{R}_\rho} d^M x\, e^{(t,\pi_k x)} Q^{\eta}_{N,\beta,h}[\omega](x) \leq e^{-\beta N \bar{c}_2} e^{h(t,\pi_k e^{\eta})} e^{+\frac{t^2}{2\beta M}} \tag{4.5}$$

where

$$\bar{c}_2 \equiv \left[\frac{1}{2}\epsilon c^-(\delta_-, \beta, h) - \bar{c}_3 a(\delta_-, \beta, h) 2r^* - \frac{M}{2\beta N}\ln\left(c^+(\delta_+, \beta, h)(1 + r^*)\right)\right] \tag{4.6}$$

with $\bar{c}_3 > 0$ a constant related to the constant c_3 in Lemma 2.5. c^- is the same as in Lemma 2.5.

An immediate consequence of the lemma is the following

Corollary 4.3: *For almost all ω, for all but a finite number of indices N,*

$$1 - (2M - 1)e^{-\bar{c}_1 \beta N} - e^{-\bar{c}_2 \beta N} \leq \widetilde{\mathcal{Q}}^{\eta}_{N,\beta,h}[\omega]\left(B^{(\eta,1)}_{\rho}\right) \leq 1 \tag{4.7}$$

We now give the proofs of Proposition 4.1 and of Theorem 2, assuming Lemma 4.2. The proof of the lemma will be given after that.

Proof of the Proposition 4.1: We consider the case $\beta > 1$; the case $\beta < 1$ is much simpler and will be left to the reader. Since we assume that $\frac{M}{N} \downarrow 0$, this implies that for all N large enough, $\beta > \delta_+$. Let us write

$$\mathcal{L}^{\eta,k}_{N,\beta,h}[\omega](t) \equiv I^c + \sum_{(\nu,s)} I^{(\nu,s)} \tag{4.8}$$

where

$$I^{(\nu,s)} \equiv \int_{B^{(\nu,s)}_{\rho}} d^M x\; e^{(t,\pi_k x)} Q^{\eta}_{N,\beta,h}[\omega](x) \tag{4.9}$$

and

$$I^c \equiv \int_{\mathcal{R}_\rho} d^M x \; e^{(t,\pi_k x)} Q^\eta_{N,\beta,h}[\omega](x). \tag{4.10}$$

Since for $x \in B_\rho^{(\nu,s)}$, $|(t,\pi_k x) - s\bar{a}(t,\pi_k e^\nu)| \leq \rho \|t\|_2$,

$$e^{-\rho\|t\|_2 + s\bar{a}(t,\pi_k e^\nu)} \widetilde{\mathcal{Q}}^\eta_{N,\beta,h}\left(B_\rho^{(\nu,s)}\right) \leq I^{(\nu,s)} \leq e^{\rho\|t\|_2 + s\bar{a}(t,\pi_k e^\nu)} \widetilde{\mathcal{Q}}^\eta_{N,\beta,h}\left(B_\rho^{(\nu,s)}\right) \tag{4.11}$$

Therefore, using (4.3) we get

$$\begin{aligned} \sum_{(\nu,s)\neq(\eta,1)} I^{(\nu,s)} &\leq e^{-\bar{c}_1\beta N} e^{\rho\|t\|_2} \sum_{(\nu,s)\neq(\eta,1)} e^{s\bar{a}(t,\pi_k e^\nu)} \\ &\leq 2M e^{-\bar{c}_1\beta N} e^{(\rho+\bar{a})\|t\|_2} \end{aligned} \tag{4.12}$$

For I^c, we have of course the bound given by (4.5).

Finally, using Corollary 4.3 and the bounds (4.11),

$$e^{-\rho\|t\|_2} e^{\bar{a}(t,\pi_k e^\eta)}(1 - 2Me^{-\bar{c}_1\beta N} - e^{-\bar{c}_1\beta N}) \leq I^{(\eta,1)} \leq e^{\rho\|t\|_2} e^{\bar{a}(t,\pi_k e^\eta)} \tag{4.13}$$

From these bounds we see that the proposition will be proven if we can choose our parameters such that $\bar{c}_1$ and $\bar{c}_2$ are sufficiently large while at the same time $\rho \equiv \rho_N$ tends to zero. Let us first consider $\bar{c}_2$. It is reasonable to fix

$$\epsilon \equiv \epsilon_N = 3\bar{c}_3 \frac{a(\delta_-,\beta,h)}{c^-(\delta_-,\beta,h)} r^* \tag{4.14}$$

in which case $\bar{c}_2 \geq a(\delta_-,\beta,h)\bar{c}_3 r^*$. At the same time, the hypothesis of Proposition 4.1 are satisfied if we choose $\rho \equiv \rho_N$ as

$$\begin{aligned} \rho_N^2 &= a^{3/2}(\delta_-,\beta,h)\left[c_1\left(3\bar{c}_3\frac{a(\delta_-,\beta,h)}{c^-(\delta_-,\beta,h)}\right)^{1/4}(r^*)^{1/4}\left(\ln\frac{c_4}{r^*}\right)^{1/4} + r^* a^{1/2}(\delta_-,\beta,h)\right] \\ &= a^{3/2} O\left((r^*)^{1/4}|\ln r^*|\right) \end{aligned} \tag{4.15}$$

which tends to zero as desired. Finally, since ρ_N tends to zero, for N large enough, $\bar{c}_1 \geq \frac{1}{2}ha(1,\beta,h) > 0$. From these observations, (4.2) now follows immediately. $\diamondsuit$

Proof of Theorem 2: We show first that the conclusion of Theorem 2 holds for the measures $\tilde{\mathcal{Q}}_{N,\beta,h=0}$. To do this, all we have to do is to show that, almost surely, $\widetilde{\mathcal{Q}}_{N,\beta,h=0}[\omega]\left(\mathcal{R}_\rho\right) \downarrow 0$, as $N \uparrow \infty$, for suitably chosen ρ. For this we can of course simply use the bound (4.5) from Lemma 4.2, with h and t set to zero. That is, we must estimate the minimal ρ for which $e^{-\beta N\bar{c}_2}$ tends to zero. Following the same arguments as in the proof of Proposition 4.1 we see that this ρ can be chosen as

$$\rho^2 = Ca^{3/2}(\delta_+,\beta)\alpha^{1/8}|\ln\alpha|^{1/4} \tag{4.16}$$

for some positive constant C, if α is sufficiently small and $\beta > \delta_+ + \alpha$ (this condition on β is needed in order that the term $\frac{M}{N}\ln(c^+(\delta_+,\beta,h)$ in (4.6) to be negligible). We recall that $\delta_+ = \frac{1}{1-2\sqrt{\alpha}} \approx 1+2\sqrt{\alpha}$.

To extend this result to the original measures $\mathcal{Q}_{N,\beta,h=0}$, we use the fact that the gaussian measures $\mathcal{N}(0,\frac{1}{\beta N})$ is in the limit $N\uparrow\infty$ concentrated on a ball of radius $O(\alpha)$, if $\limsup\frac{M}{N}=\alpha$. Namely, having established that

$$\tilde{\mathcal{Q}}_{N,\beta}(B_\rho) \geq 1-e^{-cN} \tag{4.17}$$

we define

$$\tilde{B}_\rho \equiv \left\{x\in I\!R^{I\!N}\middle|\operatorname{dist}(x,B_\rho)\leq\alpha\right\} \tag{4.18}$$

and by the definition of $\tilde{\mathcal{Q}}$, we have thus

$$\begin{aligned}
1-e^{-cN} &\leq \tilde{\mathcal{Q}}_{N,\beta}(B_\rho)= \\
&= \left(\frac{\beta N}{2\pi}\right)^{M/2}\int_{I\!R^M} d^Mx\,\mathbb{1}_{\{x\in B_\rho\}}(x)\int_{I\!R^M} e^{-\frac{\beta N}{2}\|x-m\|^2}\mathcal{Q}_{N,\beta}(dm)\\
&= \left(\frac{\beta N}{2\pi}\right)^{M/2}\int_{I\!R^M} d^Mx\,\mathbb{1}_{\{x\in B_\rho\}}(x)\int_{I\!R^M} e^{-\frac{\beta N}{2}\|x-m\|^2}\\
&\quad\times\left[\mathbb{1}_{\{\|x-m\|\leq\alpha\}}(m)+\mathbb{1}_{\{\|x-m\|>\alpha\}}(m)\right]\mathcal{Q}_{N,\beta}(dm)\\
&\equiv I+II
\end{aligned} \tag{4.19}$$

where I and II refer to the two summands with the respective characteristic functions. We will see that the first term is what we want, while the second is small; in fact

$$\begin{aligned}
II &\leq \left(\frac{\beta N}{2\pi}\right)^{M/2}\int_{I\!R^M} d^Mx\int_{I\!R^M} e^{-\frac{\beta N}{2}\gamma\|x-m\|^2}e^{-\frac{\beta N}{2}(1-\gamma)\|x-m\|^2}\mathbb{1}_{\{\|x-m\|>\alpha\}}(m)\mathcal{Q}_{N,\beta}(dm)\\
&\leq e^{-\beta N\gamma\alpha}\left(\frac{\beta N}{2\pi}\right)^{M/2}\int_{I\!R^M} d^Mx\int_{I\!R^M} e^{-\frac{\beta N}{2}(1-\gamma)\|x-m\|^2}\mathcal{Q}_{N,\beta}(dm)\\
&= e^{-\beta N\gamma\alpha}\left[\frac{1}{1-\gamma}\right]^{M/2}\\
&\leq e^{-N\gamma\alpha/2}
\end{aligned} \tag{4.20}$$

for small enough γ. Note that we assume $\beta>1$. If $\lim\frac{M}{N}=0$, this argument has to be modified slightly by replacing α in (4.18) by a suitably chosen $\epsilon(M,N)$. We leave the details to the reader.

To deal with the term I, we use that

$$\begin{aligned}
I &= \left(\frac{\beta N}{2\pi}\right)^{M/2}\int_{I\!R^M} d^Mx\int_{I\!R^M} e^{-\frac{\beta N}{2}\|x\|^2}\mathbb{1}_{\{x+m\in B_\rho\}}(x)\mathbb{1}_{\{\|x\|\leq\alpha\}}(m)\mathcal{Q}_{N,\beta}(dm)\\
&\leq \left(\frac{\beta N}{2\pi}\right)^{M/2}\int_{I\!R^M} d^Mx\int_{I\!R^M} e^{-\frac{\beta N}{2}\|x\|^2}\mathbb{1}_{\{m\in\tilde{B}_\rho\}}(m)\mathcal{Q}_{N,\beta}(dm)\\
&= \mathcal{Q}(\tilde{B}_\rho)
\end{aligned} \tag{4.21}$$

Since $\tilde{B}_\rho \subset B_{\rho+\alpha}$, this implies

$$\mathcal{Q}_{N,\beta}(B_{\rho+\alpha}) \geq 1 - e^{-\beta N \gamma \alpha/2} - e^{-cN} \tag{4.22}$$

and since α is much smaller that the ρ from (4.16), this implies immediately Theorem 2 for $\mathcal{Q}$. $\diamondsuit$

Proof of Theorem 1: Since (i) of Theorem 1 follows immediately from Proposition 4.1 and Lemma 2.1, to complete the proof of Theorem 1 we only have to verify statement (ii). But this is essentially a special case of Theorem 2. Since under its assumptions $M/N \downarrow 0$, we can choose, according to (4.16), $\rho = \rho(N)$ N-dependent and tending to zero as N goes to infinity. Thus Theorem 2 implies for such M that $\lim_{N\uparrow\infty} \mathcal{Q}_{N,\beta,h=0}(B_0) = 1$, $I\!\!P$ – almost surely. Remembering the definition of limiting induced measures in Section 1, we require the slightly stronger statement that

$$\lim_{\|f\|_\infty \downarrow 0} \lim_{N\uparrow\infty} \mathcal{Q}^{(f)}_{N,\beta}(B_0) = 1, \ I\!\!P - \text{almost surely} \tag{4.23}$$

To show this, one introduces as above for the general f-dependent measures the corresponding functions $\Phi^{(f)}_{N,\beta}(z)$. The point is that these are uniformly continuous in f. In fact a simple computation shows that they even satisfy

$$\left|\Phi^{(f)}_{N,\beta}(z) - \Phi_{N,\beta}(z)\right| \leq \|f\|_\infty \tag{4.24}$$

From this (4.23) is obtained easily by the same estimates as before. $\diamondsuit$

Theorems 1 and 2 are now proven up to the proof of Lemma 4.2 which we present now:

Proof of Lemma 4.2: Let us recall that

$$\Phi^\eta_{N,\beta,h}[\omega](x) = q_{N,\delta}[\omega](x - he^\eta) + \Psi^\eta_{N,\beta,h,\delta}[\omega](\xi x) + \frac{\delta h^2}{2}$$

We choose $\delta \equiv \delta_N = \frac{1}{1+r^*}$, with r^* as in Section 3, so that, for $\omega \in \mathcal{A}_N$, the quadratic form $q_{N,\delta}[\omega]$ is positive definite. In the rest of the proof all statements concerning random variables will be understood to hold for $\omega \in \mathcal{A}_N$. Let us first prove part *(i)*. Recalling the definition of Ψ (2.9,10), it is easy to see that

$$\Phi^\eta_{N,\beta,h}[\omega](x) = \Phi^\nu_{N,\beta,sh}[\omega](x) + h\left(x, (se^\nu - e^\eta)\right) \tag{4.25}$$

Using this observation we may write

$$\begin{aligned}\widetilde{\mathcal{Q}}^\eta_{N,\beta,h}[\omega]\left(B^{(\nu,s)}_\rho\right) &= \frac{\int_{B^{(\nu,s)}_\rho} d^M x\, e^{-\beta N\left\{\Phi^\nu_{N,\beta,sh}[\omega](x) + h(x,(se^\nu - e^\eta))\right\}}}{\int_{I\!\!R^M} d^M\, x e^{-\beta N\left\{\Phi^\eta_{N,\beta,h}[\omega](x)\right\}}} \\ &\leq \frac{\int_{B^{(\nu,s)}_\rho} d^M x\, e^{-\beta N\left\{\Phi^\nu_{N,\beta,sh}[\omega](x) + h(x,(se^\nu - e^\eta))\right\}}}{\int_{B^{(\eta,1)}} d^M x\, e^{-\beta N\left\{\Phi^\eta_{N,\beta,h}[\omega](x)\right\}}}\end{aligned} \tag{4.26}$$

To get an upper bound on this term, notice first that $\forall(\nu,s)\neq(\eta,1)$ and $\forall x\in B_\rho^{(\nu,s)}$

$$h(\overline{a}-2\rho)\leq h(x,(se^\nu-e^\eta))\leq h(\overline{a}+2\rho) \tag{4.27}$$

On the other hand, writing $x-she^\nu = x-s\overline{a}e^\nu + s(\overline{a}-h)e^\nu$, the quadratic term can be written as

$$\begin{aligned} q_{N,\delta}(x-she^\nu) &= q_{N,\delta}((\overline{a}-h)e^\nu)+q_{N,\delta}(x-s\overline{a}e^\nu)+\left(s(\overline{a}-h)e^\nu,\left[\mathbb{1}-\delta\frac{\xi^t\xi}{N}\right](x-s\overline{a}e^\nu)\right)\\ &= \frac{1}{2}(\overline{a}-h)^2(1-\delta)+R \end{aligned} \tag{4.28}$$

where

$$|R|\leq\left(\frac{\rho^2}{2}+\rho|\overline{a}-h|\right)\right)\left\|\mathbb{1}-\delta\frac{\xi^t\xi}{N}\right\|\leq\left(\frac{\rho^2}{2}+\rho|\overline{a}-h|\right)\right)(2r^*) \tag{4.29}$$

Using this bound in the numerator and denominator of (4.25), we get

$$\widetilde{\mathcal{Q}}^\eta_{N,\beta,h}[\omega]\left(B_\rho^{(\nu,s)}\right)\leq e^{-\beta N h(\overline{a}-2\rho)}e^{\beta N(\rho|\overline{a}-h|4r^*)}\mathcal{J}^{(\nu,s)} \tag{4.30}$$

where

$$\mathcal{J}^{(\nu,s)}\equiv\frac{\int_{B_\rho^{(\nu,s)}}d^Mx\,e^{-\beta N\left\{\Psi^\nu_{N,\beta,sh,\delta}(\xi x)\right\}}}{\int_{B_\rho^{(\eta,1)}}d^Mx\,e^{-\beta N\left\{\Psi^\eta_{N,\beta,h,\delta}(\xi x)\right\}}} \tag{4.31}$$

To estimate the last quantity, we change coordinates in the integrals in the numerator and denominator to $x'=x-s\overline{a}e^\nu$ and $x'=x-\overline{a}e^\eta$, respectively. Moreover, we will use that

$$\Psi^\nu_{N,\beta,sh,\delta}(\xi x)=\Psi_{N,\beta,h=0,\delta}(\xi x)+\frac{\delta h}{N}(s\xi^\nu,\xi x) \tag{4.32}$$

Let B_ρ^0 denote the ball of radius ρ centered the origin. We may then write

$$\mathcal{J}^{(\nu,s)}=\frac{\int_{B_\rho^0}d^Mx\,e^{-\beta N\left\{\Psi_{N,\beta,\delta}(\overline{a}\xi^\eta+\xi x)+[\Psi_{N,\beta,\delta}(s\overline{a}\xi^\nu+\xi x)-\Psi_{N,\beta,\delta}(\overline{a}\xi^\eta+\xi x)]-\delta h\left(e^\nu,\frac{\xi^t\xi}{N}x\right)\right\}}}{\int_{B_\rho^0}d^Mx\,e^{-\beta N\left\{\Psi_{N,\beta,\delta}(\overline{a}\xi^\eta+\xi x)-\delta h\left(e^\eta,\frac{\xi^t\xi}{N}x\right)\right\}}} \tag{4.33}$$

Now

$$\left|\delta h\left(e^\nu,\frac{\xi^t\xi}{N}x\right)\right|\leq\delta h(1+r^*)\|x\|^2 \tag{4.34}$$

while Corollary 2.6 implies that

$$|\Psi_{N,\beta,\delta}(s\overline{a}\xi^\nu+\xi x)-\Psi_{N,\beta,\delta}(\overline{a}\xi^\eta+\xi x)|\leq\frac{c^+(\delta,\beta,h)}{N}\|\xi x\|^2\leq c^+(\delta,\beta,h)(1+r^*)\|x\|^2 \tag{4.35}$$

Inserting these two bounds in (4.33) and, using that in the domain of integration $\|x\|\leq\rho$, we get

$$\mathcal{J}^{(\nu,s)}\leq e^{\beta N(1+r^*)\rho^2\left[c^+(\delta,\beta,h)+2\delta h\right]} \tag{4.36}$$

and so finally

$$\widetilde{\mathcal{Q}}^{\eta}_{N,\beta,h}[\omega]\left(B^{(\nu,s)}_{\rho}\right) \leq e^{-\beta N\left\{h(\overline{a}-2\rho)-4\rho|\overline{a}-h|r^{*}-\rho^{2}(1+\rho^{*})(c^{+}(\delta,\beta,h)+2\delta h)\right\}} \tag{4.37}$$

which proves part (i) of Lemma 4.2.

Let us now prove part (ii). We have to estimate

$$\int_{\mathcal{R}_\rho} d^M x\, e^{(t,\pi_k x)} \mathcal{Q}^{\eta}_{N,\beta,h}[\omega](x) = \frac{\int_{\mathcal{R}_\rho} d^M x\, e^{(t,\pi_k x)} e^{-\beta N \Phi^{\eta}_{N,\beta,h,\delta}[\omega](x)}}{\int_{I\!R^M} d^M x\, e^{-\beta N \Phi^{\eta}_{N,\beta,h,\delta}[\omega](x)}} \tag{4.38}$$

We treat the numerator and denominator in (4.38) separately. In particular, we will use a different choice for δ in each of them. Consider first the denominator. Here we choose $\delta = \delta_+$ such that q_{N,δ_+} is strictly negative, which will be the case if $\delta_+ = \frac{1}{1-r^*}$. This gives

$$\begin{aligned}
\int_{I\!R^M} d^M x\, e^{-\beta N \Phi^{\eta}_{N,\beta,h,\delta}[\omega](x)} &= \int_{I\!R^M} d^M x\, e^{-\beta N\left\{q_{N,\delta_+}(x-he^{\eta})+\Psi^{\eta}_{N,\beta,h,\delta_+}(\xi x)\right\}} \\
&\geq e^{-\beta N \varphi_0(\delta_+)} \int_{I\!R^M} d^M x\, e^{-\beta N\left\{\Psi^{\eta}_{N,\beta,h,\delta_+}(\xi x)-\varphi_0(\delta_+)\right\}} \\
&\geq e^{-\beta N \varphi_0(\delta_+)} \int_{I\!R^M} d^M x\, e^{-\beta N \frac{c^{+}(\delta_+,\beta,h)}{2N}\|\xi(x-\overline{a}e^{\eta})\|^2} \\
&\geq e^{-\beta N \varphi_0(\delta_+)} \int_{I\!R^M} d^M x\, e^{-\beta N \frac{c^{+}(\delta_+,\beta,h)}{2}(1+r^*)\|x\|^2}
\end{aligned} \tag{4.39}$$

where we have used (2.20) from Corollary 2.6 and where we have set

$$\varphi_0(\delta) \equiv \inf_{y \in I\!R} \phi_{\beta,h,\delta}(y)$$

We now turn to the numerator in (4.38). Here we choose $\delta = \delta_- \equiv \frac{1-\tau_N}{1+r^*}$ with $\tau_N > 0$ to be chosen later. This choice implies that $q_{N,\delta_-}(x) \geq \frac{\tau_N}{2}\|x\|^2$. Using this time the bound (2.21) from Corollary 2.6, this yields

$$\begin{aligned}
&\int_{\mathcal{R}_\rho} d^M x\, e^{(t,\pi_k x)} e^{-\beta N \Phi^{\eta}_{N,\beta,h,\delta}[\omega](x)} \\
&= e^{-\beta N \varphi_0(\delta_-)} \int_{\mathcal{R}_\rho} d^M x\, e^{(t,\pi_k x)} e^{-\beta N\left\{q_{N,\delta}(x-he^{\eta})+\Psi^{\eta}_{N,\beta,h,\delta}(\xi x)-\varphi_0(\delta_-)\right\}} \\
&\leq e^{-\beta N \varphi_0(\delta_-)} \int_{\mathcal{R}_\rho} d^M x\, e^{(t,\pi_k x)} e^{-\beta N\left\{\frac{\tau_N}{2}\|x-he^{\eta}\|^2+\frac{c^{-}(\delta_-,\beta,h)}{2N}\inf_{\zeta\in\mathcal{E}_N(\overline{a})}\|\xi x-\zeta\|^2\right\}} \\
&\leq e^{-\beta N \varphi_0(\delta_-)} e^{-\epsilon\beta N \frac{c^{-}(\delta_-,\beta,h)}{2}} \int_{I\!R^M} d^M x\, e^{(t,\pi_k x)} e^{-\beta N \frac{\tau_N}{2}\|x-he^{\eta}\|^2}
\end{aligned} \tag{4.40}$$

where in the last line we have made use of the bound (3.1) from Proposition 3.1. To do this, we assume that the hypothesis of this proposition are satisfied. In particular, ϵ depends of course on ρ.

Also, (4.40) holds on a subset $\tilde{\mathcal{A}}_N \subset \Omega$ whose complement has a probability that was bounded in (3.1). This estimate and the first Borel-Cantelli lemma will imply that the bounds we are proving are true for $I\!\!P$-almost all ω and for all but a finite number of values of N.

Combining these two bounds, we arrive at

$$
\begin{aligned}
&\int_{\mathcal{R}_\rho} d^M x\, e^{(t,\pi_k x)} Q^\eta_{N,\beta,h}[\omega](x) \\
&\leq e^{-\beta N\left(\frac{1}{2}\epsilon c^-(\delta_-,\beta,h)-(\varphi_0(\delta_+)-\varphi_0(\delta_-))\right)} \frac{\int_{I\!\!R^M} d^M x\, e^{(t,\pi_k x)} e^{-\beta N\frac{\tau_N}{2}\|x-he^\eta\|^2}}{\int_{I\!\!R^M} d^M x\, e^{-\beta N\frac{c^+(\delta_+,\beta,h)}{2}(1+r^*)\|x\|^2}} \\
&= e^{-\beta N\left(\frac{1}{2}\epsilon c^-(\delta_-,\beta,h)-(\varphi_0(\delta_+)-\varphi_0(\delta_-))\right)} \left[\frac{c^+(\delta_+,\beta,h)(1+r^*)}{\tau_N}\right]^{M/2} e^{h(t,\pi_k e^\eta)} e^{+\frac{t^2}{2\beta N\tau_N}}
\end{aligned}
\tag{4.41}
$$

Using point (iv) of Lemma 2.5, we see that

$$
\begin{aligned}
(\varphi_0(\delta_+)-\varphi_0(\delta_-)) &\leq c_3 a(\delta_-,\beta,h)(\delta_+-\delta_-) = c_3 a(\delta_-,\beta,h)\frac{2r^*+\tau_N(1-r^*)}{1-(r^*)^2} \\
&\leq \bar{c}_3 a(\delta_-,\beta,h)(2r^*+\tau_N)
\end{aligned}
\tag{4.42}
$$

Thus,

$$
\begin{aligned}
\int_{\mathcal{R}_\rho} d^M x\, e^{(t,\pi_k x)} Q^\eta_{N,\beta,h}[\omega](x) &\leq e^{h(t,\pi_k e^\eta)} e^{+\frac{t^2}{2\beta N\tau_N}} \\
&\times e^{-\beta N\left[\frac{1}{2}\epsilon c^-(\delta_-,\beta,h)-\bar{c}_3 a(\delta_-,\beta,h)(2r^*+\tau_N)-\frac{M}{2\beta N}\left\{\ln\left(c^+(\delta_+,\beta,h)(1+r^*)\right)+|\ln\tau_N|\right\}\right]}
\end{aligned}
\tag{4.43}
$$

From this expression we see that $\tau_N = \frac{M}{N}$ is a reasonable choice (we exclude the trivial case M bounded where we would of course choose a larger τ_N); it means in particular that for small $\frac{M}{N}$ it can be neglected in the first exponential in (4.43) (on the expense of slightly enlarging the constant $\bar{c}_3$). This yields (4.5) and concludes the proof of Lemma 4.2. $\diamondsuit$

Finally, we come to the

Proof of Theorem 3: What we have to do here is to show that the marginals of the Gibbs measure on any cylinder generated by a *finite* subset of the spin-variables converge to a product measure as stated. Thus we select a finite subset $V \subset I\!\!N$ and let $\mathcal{A} \subset \mathcal{B}_V$ denote a cylinder event. Of course, without restriction of generality we can assume that $\mathcal{A}$ is the event $\mathcal{A} = \{\sigma_i = s_i,\ \forall i \in V\}$. We will assume in the sequel that N is so large that $V \subset \{1,\dots,N\}$. We denote by V^c the complement of V in $\{1,\dots,N\}$. Then

$$
\begin{aligned}
\mathcal{G}^\eta_{N,\beta,h}(\mathcal{A}) =& \frac{1}{Z^\eta_{N,\beta,h}} \frac{1}{2^N} \sum_{\sigma\in\mathcal{S}_N} 1\!\!\mathrm{I}_{\mathcal{A}}(\sigma) \\
& e^{\frac{\beta}{2N}\sum_{i,j\in V}\sigma_i\sigma_j J_{ij}+\frac{\beta}{2N}\sum_{i,j\in V^c}\sigma_i\sigma_j J_{ij}+\frac{\beta}{N}\sum_{i\in V,j\in V^c}\sigma_i\sigma_j J_{ij}+h\beta\sum_{i\in V}\xi^\eta_i\sigma_i+h\beta\sum_{i\in V^c}\xi^\eta_i\sigma_i}
\end{aligned}
\tag{4.44}
$$

where we have used the abbreviation $J_{ij} \equiv \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu$. In terms of the 'local overlaps' $m_V^\alpha(\sigma) \equiv \frac{1}{|V|}\sum_{i\in V} \xi_i^\alpha \sigma_i$, this can be re-expressed in the form

$$\begin{aligned}\mathcal{G}^\eta_{N,\beta,h}(\mathcal{A}) =& \frac{1}{Z^\eta_{N,\beta,h}} \sum_{\sigma_V \subset \mathcal{S}_V} \mathbb{1}_{\mathcal{A}}(\sigma_V) e^{\beta \frac{|V|^2}{N}\frac{1}{2}\|m_V(\sigma_V)\|^2 + \beta h |V| m_V^\eta(\sigma_V)} \\ &\times \sum_{\sigma_{V^c} \subset \mathcal{S}_{V^c}} e^{\beta N' \left(\frac{1}{2}\|m_{V^c}(\sigma_{V^c})\|^2 + h' m^\eta_{V^c}(\sigma_{V^c}) + \frac{|V|}{|V^c|}(m_V(\sigma_V), m_{V^c}(\sigma_{V^c}))\right)}\end{aligned} \tag{4.45}$$

Here $N' \equiv \frac{|V^c|^2}{N}$ and $h' \equiv h\frac{N}{|V^c|}$. Of course the distinction between N, N' and $|V^c|$ is completely irrelevant in the limit $N \uparrow \infty$. Now

$$\begin{aligned}&\frac{1}{2^{|V^c|}} \sum_{\sigma_{V^c} \subset \mathcal{S}_{V^c}} e^{\beta N' \left(\frac{1}{2}\|m_{V^c}(\sigma_{V^c})\|^2 + h' m^\eta_{V^c}(\sigma_{V^c}) + \frac{|V|}{|V^c|}(m_V(\sigma_V), m_{V^c}(\sigma_{V^c}))\right)} \\ &= \left(\frac{\beta N'}{2\pi}\right)^{M/2} \int d^M x \, e^{-\frac{\beta N'}{2}(x,x)} \frac{1}{2^{|V^c|}} \sum_{\sigma_{V^c} \subset \mathcal{S}_{V^c}} e^{\beta N' \left(m_{V^c}(\sigma_{V^c}), (x + \frac{|V|}{|V^c|} m_V(\sigma_V) + h' e^\eta)\right)} \\ &= \left(\frac{\beta N'}{2\pi}\right)^{M/2} \int d^M x \, e^{-\frac{\beta N'}{2}\left(x - \frac{|V|}{|V^c|} m_V(\sigma_V) - h' e^\eta, x - \frac{|V|}{|V^c|} m_V(\sigma_V) - h' e^\eta\right)} \frac{1}{2^{|V^c|}} \sum_{\sigma_{V^c} \subset \mathcal{S}_{V^c}} e^{\beta N' (m_{V^c}(\sigma_{V^c}), x)} \\ &= e^{-\beta \frac{|V|^2}{N}\frac{1}{2}\|m_V(\sigma_V)\|^2 - \beta h |V| m_V^\eta(\sigma_V) - \beta N h^2} \left(\frac{\beta N'}{2\pi}\right)^{M/2} \\ &\times \int d^M x \, e^{\beta N' \frac{|V|}{|V^c|}(x, m_V(\sigma_V)) + \beta N' \left\{-\frac{1}{2}\|x - h' e^\eta\|^2 + \frac{1}{\beta N'}\sum_{i \in V^c} \ln\cosh\beta(\xi x)_i\right\}}\end{aligned} \tag{4.46}$$

Thus

$$\mathcal{G}^\eta_{N,\beta,h}(\mathcal{A}) = \frac{\sum_{\sigma_V \subset \mathcal{S}_V} \mathbb{1}_{\mathcal{A}}(\sigma_V) \int d^M x \, e^{\beta N' \frac{|V|}{|V^c|}(x, m_V(\sigma_V)) + \beta N' \left\{-\frac{1}{2}\|x - h' e^\eta\|^2 + \frac{1}{\beta N'}\sum_{i \in V^c} \ln\cosh\beta(\xi x)_i\right\}}}{\sum_{\sigma_V \subset \mathcal{S}_V} \int d^M x \, e^{\beta N' \frac{|V|}{|V^c|}(x, m_V(\sigma_V)) + \beta N' \left\{-\frac{1}{2}\|x - h' e^\eta\|^2 + \frac{1}{\beta N'}\sum_{i \in V^c} \ln\cosh\beta(\xi x)_i\right\}}} \tag{4.47}$$

Now the integrals in (4.47) are exactly those we had to deal with in the proof of Proposition 4.1 (excepting trivial modifications) and a re-run of that proof shows yields that, $I\!P$-almost surely,

$$\begin{aligned}\lim_{h \to 0} \lim_{N \uparrow \infty} \mathcal{G}^\eta_{N,\beta,h}(\mathcal{A}) =& \frac{\sum_{\sigma_V \subset \mathcal{S}_V} \mathbb{1}_{\mathcal{A}}(\sigma_V) \prod_{i \in V} e^{\beta a(\beta) \xi_i^\eta \sigma_i}}{\sum_{\sigma_V \subset \mathcal{S}_V} \prod_{i \in V} e^{\beta a(\beta) \xi_i^\eta \sigma_i}} \\ =& \sum_{\sigma_V \subset \mathcal{S}_V} \mathbb{1}_{\mathcal{A}}(\sigma_V) \prod_{i \in V} \frac{e^{\beta a(\beta) \xi_i^\eta \sigma_i}}{2\cosh\beta a(\beta)}\end{aligned} \tag{4.48}$$

which gives, together with the fact that $a(\beta) = \tanh\beta a(\beta)$ gives the statement of Theorem 3. $\diamond\diamond$

Remark: It is interesting to notice that in the proof of Theorem 3 the measures $\tilde{\mathcal{Q}}$ enter directly and not the actual induced measures $\mathcal{Q}$.

Appendix

In this appendix we present a proof of Theorem 2.4 on the eigenvalues of the matrix $\xi^t\xi$, namely

Theorem 2.4: *Assume that ξ_i^μ are i.i.d. random variables satisfying $I\!E\xi_i^\mu = 0$ and $I\!E\left(\xi_i^\mu\right)^k \leq 1$, for all $k > 1$. Let B denote the $M \times M$-matrix whose elements are*

$$B_{\mu\nu} \equiv (1-\delta_{\mu,\nu})\frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu\xi_i^\nu \tag{A.1}$$

Then, for any $z \geq 0$, $M \leq N$ and for N sufficiently large

$$I\!P\left(\|B\| > \sqrt{\alpha}(2+\sqrt{\alpha}) + zN^{-1/6}\ln N\right) \leq NN^{-z/\sqrt{\alpha}(2+\sqrt{\alpha})} \tag{A.2}$$

Remark: It may be noted that (A.2) implies, in particular that, for all $\epsilon \geq 0$,

$$I\!P\left(\|B\| > \sqrt{\alpha}(2+\sqrt{\alpha})(1+\epsilon)\right) \leq 4Ne^{-\epsilon N^{1/6}} \tag{A.2'}$$

Remark: Note that the conditions on ξ are satisfied for symmetric Bernoulli variables. Up to trivial rescaling, they accommodate also all bounded centered r.v.'s. In the special case of symmetric Bernoulli variables, (A.2) holds with $N^{1/6}$ replaced by $N^{1/4}$ everywhere.

In [BG] we have proven a related result for the matrix $A = B + \mathbb{1}$, namely that $\|A\| \leq e^{2\sqrt{M/N}}$ with large probability. The proof of the present, somewhat sharper result uses the same techniques and some of the combinatorial facts that are proven there. The general strategy follows that of Füredi and Komlos [FK].

The main ingredient is a bound on the expectation of the trace of high powers of B. It reads

Lemma A.1: *Let B be the matrix defined in Theorem 2.4. Then, for all $k \leq N^{1/6}$,*

$$I\!E trB^k \leq CN\left[\sqrt{\alpha}(2+\sqrt{\alpha})\right]^{k+1} \tag{A.3}$$

Remark: Similar results have been given by Koch [K] and by Tirozzi and Shcherbina [TS]. We present an independent proof here.

Proof: (of Lemma A.1) First we write, with $\mathcal{M} \equiv \{1,\dots,M\}$ and $\mathcal{N} \equiv \{1,\dots,N\}$

$$I\!E\mathrm{tr}B^k = \frac{1}{N^k}\sum_{\substack{\mu_0,\dots,\mu_{k-1}\in\mathcal{M}\\ \forall_l \mu_l\neq\mu_{l-1}}}\ \sum_{i_0,\dots,i_{k-1}\in\mathcal{N}} I\!E\left(\xi_{i_0}^{\mu_0}\xi_{i_0}^{\mu_1}\xi_{i_1}^{\mu_1}\dots\xi_{i_{k-1}}^{\mu_{k-1}}\xi_{i_{k-1}}^{\mu_0}\right) \tag{A.4}$$

We think of the two sums as sums over sequences $(\mu_0, \dots, \mu_{k-1}) \in \mathcal{M}^k$, etc. For such a sequence we will denote by

$$\{(\mu_0, \dots, \mu_{k-1})\} \equiv \left\{t \in \mathcal{M} \middle| \exists_{0 \le l \le k-1} s.t. \, \mu_l = t\right\} \tag{A.5}$$

the set of different values the sequence runs through. We may then arrange the sums in (A.4) in such a way as to first sum over all possible subsets $\Gamma_1 \subset \mathcal{M}$ and $\Gamma_2 \subset \mathcal{N}$ and then over all sequences for which the values run through exactly these subsets. Moreover, these sums will not depend on the exact subsets Γ_1, Γ_2, but only on their cardinalities. Thus

$$I\!E \mathrm{tr} B^k = \frac{1}{N^k} \sum_{r=1}^{min(k,M)} \sum_{s=1}^{min(k,N)} \binom{M}{r} \binom{N}{s} E_{k,r,s} \tag{A.6}$$

where

$$E_{k,r,s} \equiv \sum_{\substack{(\mu_0,\dots,\mu_{k-1}) \in \mathcal{M}^k \\ \{(\mu_0,\dots,\mu_{k-1})\} = \{1,\dots,r\} \\ \forall_l \, \mu_l \neq \mu_{l-1}}} \sum_{\substack{(i_0,\dots,i_{k-1}) \in \mathcal{N}^k \\ \{(i_0,\dots,i_{k-1})\} = \{1,\dots,s\}}} I\!E \left(\xi_{i_0}^{\mu_0} \xi_{i_0}^{\mu_1} \xi_{i_1}^{\mu_1} \dots \xi_{i_{k-1}}^{\mu_{k-1}} \xi_{i_{k-1}}^{\mu_0} \right) \tag{A.7}$$

where the combinatorial factors in (A.6) count the number of subsets of given cardinality. Note that $E_{k,r,s}$ does not depend on M or N.

To estimate these last quantities, we think of the sums in (A.7) in a slightly different way. Let us denote by $\mathcal{G}_{r,s}$ the complete bipartite graph with vertex sets labelled by $\mathcal{R} \equiv \{1, \dots, r\}$ and $\mathcal{S} \equiv \{1', \dots, s'\}$ (here the prime indicate that the points in the two sets are understood to be distinct), i.e. the graph with vertex set $\mathcal{R} \cup \mathcal{S}$ and edge set $\mathcal{R} \times \mathcal{S}$. Associating with each ξ_i^μ appearing in the expectation in (A.7) an edge (i, μ) of $\mathcal{G}_{r,s}$, each term in the sum (A.7) corresponds to a walk, ω, of length $2k$ on this bipartite graph (i.e. a sequence of edges linking alternately the sets $\mathcal{R}$ and $\mathcal{S}$) with the property that each vertex of $\mathcal{G}_{r,s}$ is visited at least once. Moreover, it is clear that any walk which passes over any given edge of $\mathcal{G}_{r,s}$ exactly once will give a zero contribution as the expectation of the corresponding product of ξ_i^μ's vanishes by assumption on the distribution of the ξ. Finally, the constraint $\mu_l \neq \mu_{l-1}$ in the first sum forbids that a walk after arriving at a point $i \in \mathcal{S}$ from a point $\mu \in \mathcal{R}$ returns immediately to the same point μ. We denote by $\widetilde{\mathcal{W}}_k(r, s)$ the set of walks that give a non-zero contribution. By our assumptions, we have that

$$E_{k,r,s} \le \left|\widetilde{\mathcal{W}}_k(r, s)\right| \tag{A.8}$$

The only new feature compared to the proof in [BG] is now the constraint on the walk not to return along itself after visiting a point in $\mathcal{S}$. As a main effect, this introduces a constraint on the admissible values r and s for given k as we will see now. Let ω_k be a walk in $\widetilde{\mathcal{W}}_k(r, s)$. Then the set of edges in $\mathcal{G}_{r,s}$ over which ω_k passes form a connected, bipartite graph on $\mathcal{R} \times \mathcal{S}$ which we will denote G_k. Let $(d_1, \dots, d_r)$ and $(c_1, \dots, c_s)$ denote the co-ordination numbers of the vertices of G_k in the sets $\mathcal{R}$ and $\mathcal{S}$, respectively. Due to the constraint on the walk we these numbers must satisfy

$$d_l \ge 1 \quad \text{and} \quad c_l \ge 2 \tag{A.9}$$

On the other hand, if L denotes the number of edges in G_k, then we have the following relation:

$$\sum_{l=1}^{s} c_l = \sum_{l=1}^{r} d_l = L \geq r+s-1 \tag{A.10}$$

On the other hand,

$$\sum_{l=1}^{s} c_l \geq 2s, \quad \sum_{l=1}^{r} d_l \geq r, \quad L \leq k \tag{A.11}$$

which combine to

$$2s \leq k, \quad r \leq k \quad \text{and} \quad r+s-1 \leq k \tag{A.12}$$

Let us now first consider the case where $r+s-1=k$. We will moreover assume, for notational simplicity, that k is even. Then it is clear that the graph G_k is in fact a bipartite tree. In [BG] it was proven that

Lemma A.2: *Let t_k be a bipartite tree on $\mathcal{R}\times\mathcal{S}$ with $k=r+s-1$. Then the number, $W(c_1,\dots,d_r)$, of walks ω on t_k starting in $\mathcal{R}$ and passing through each edge of t_k exactly twice times the number of such trees with given coordination numbers $c_1,\dots,c_s,d_1,\dots,d_r$ is given by*

$$W(c_1,\dots,d_r) = k(r-1)!(s-1)! \tag{A.13}$$

Using this result and noting that for any bipartite tree $\sum_{i=1}^{r} d_i = \sum_{i=1}^{s} c_i = r+s-1$, we get immediately that

$$\begin{aligned} |\widetilde{\mathcal{W}}_k(r,s)| &= \sum_{\substack{d_1,\dots,d_r\geq 1\\ \sum_{i=1}^{r} d_i=r+s-1}} \; \sum_{\substack{c_1,\dots,c_s\geq 2\\ \sum_{i=1}^{s} c_i=r+s-1}} k(r-1)!(s-1)! \\ &= k\binom{r+s-2}{r-1}\binom{r-2}{s-1}k(r-1)!(s-1)! = k!\binom{k-s-1}{s-1} \end{aligned} \tag{A.14}$$

Let us now turn to the case where $r+s-1<k$. In this case we follow Füredi and Komlos [FK] in associating to each walk ω in $\widetilde{\mathcal{G}}_k(r,s)$ a code sequence consisting of k signs $+$, k signs $-$ and $2(k-r-s+1)$ labels $(x,y)\in\mathcal{R}\times\mathcal{S}$, in the following way: Following the walk, we label an edge passed by the walk $+$ if the walk arrives at a point not previously visited. We put a label $-$ if the walk passes an edge for the second time that had previously been labelled $+$; in all other cases we call the step a 'jump' and put a label (x,y), with x the starting and y the endpoint of the jump. It is clear that the edges labelled $+$ form a bipartite tree. The important observation here is that if we are given the order in which the sites are visited by the walk and the code sequence, then it is possible to reconstruct the walk. For, at any given time step, if our label is $+$, we have to go to the next site; if the label is (x,y), we go from x to y and if it is $-$, we go in the direction of the

starting point of the next jump along the existing $+$-edges. Thus counting the number of possible code-sequences amounts to counting the number of walks. To do this, we first fix the times and labels of the jumps. Quite clearly, there are no more than

$$\binom{2k}{2(r+s-1)}(sr)^{2(k-r-s+1)} \tag{A.15}$$

possibilities (if the random variables ξ_i^μ are symmetric and not only centered, the factor $(sr)^{2(k-r-s+1)}$ may be replaced by $(2rs)^{k-r-s+1}$, since all jumps, too, must occur use a given edge an even number of times). This implies the slightly sharper mentioned in the remark following Theorem 2.4). Then, observe that the sequence of $\pm$-labels together with the order in which the walk visits the sites for the first time correspond exactly to a walk $\tilde{\omega}$ of length $2(r+s-1)$, and we can use Lemma A.1 to count these walks. The only remaining problem is to sum over the possible coordination numbers of the trees associated to the walks $\tilde{\omega}$. Now it is easy to convince oneself that any site in $\mathcal{S}$ that is not a starting point of a jump can have coordination number one. Moreover, it is clear that the number of jumps from $\mathcal{S}$ to $\mathcal{R}$ is equal to half the total number of jumps. Setting $l = k-r-s+1$, we see therefore that the sum over the coordination numbers in $\mathcal{S}$ yields a factor that is bounded by

$$\sum_{\substack{c_1,\dots,c_{s-l}\geq 2\\ c_{s-l+1},\dots,c_s\geq 1\\ \sum_{i=1}^s c_i = r+s-1}} 1 = \sum_{\substack{c_1,\dots,c_s\geq 1\\ \sum_{i=1}^s c_i = r+l-1}} 1 = \binom{r+l-2}{s-1} \tag{A.16}$$

where of course we assume $l \leq s$ (and replace l by s in (A.16) otherwise). Putting these observations together we arrive at

$$|\widetilde{\mathcal{W}}_k(r,s)| \leq (r+s-1)![rs]^{2(k-r-s+1)}\binom{k-s-1}{s-1} \tag{A.17}$$

Now put

$$S_{N,M,k,r,s} \equiv \binom{N}{s}\binom{M}{r}(r+s-1)![rs]^{2(k-r-s+1)}\binom{k-s-1}{s-1} \tag{A.18}$$

Our aim is to bound (A.18) uniformly in the allowed values of r and s for suitably chosen k. It is easy to see that if $k^6 \leq N$, the terms with $s+r-1=k$ dominate to such extend that the sum over all other choices of $r+s$ can be bounded by twice these terms, i.e.

$$\sum_{r+s\leq k+1} S_{N,M,k,r,s} \leq 2\sum_{r+s=k+1} S_{N,M,k,r,s} = 2\sum_{s=1}^{k/2}\binom{N}{k+1-r}\binom{M}{r}k!\binom{r-2}{k+1-r} \tag{A.19}$$

This last sum is finally bounded by $k/2$ times its maximal term. Using Sterling's formula one finds that this maximum is taken for $r=\gamma k$, where $\gamma = \frac{1+\sqrt{\alpha}}{2+\sqrt{\alpha}}$, for N large, and that the maximal value is given (up to irrelevant factors), by $r^{-1}N^{k+1}\left[\sqrt{\alpha}(2+\sqrt{\alpha})\right]^k$ (where $\alpha \equiv M/N$). From this Lemma A.1 follows immediately.$\diamondsuit$

Proof: (of Theorem 2.4) Theorem 2.4 follows from Lemma A.1 by a simple application of the Tchebychev inequality. Namely, since the matrix B is symmetric, we have that

$$\begin{aligned} I\!P\left(\|B\| > \sqrt{\alpha}(2+\sqrt{\alpha}) + zN^{-1/6}\ln N\right) &\leq I\!P\left(\mathrm{tr}(B^k) > \left(1\sqrt{\alpha}(2+\sqrt{\alpha}) + zN^{-1/6}\ln N\right)^k\right) \\ &\leq \frac{I\!E\mathrm{tr}(B^k)}{\left(\sqrt{\alpha}(2+\sqrt{\alpha}) + zN^{-1/6}\ln N\right)^k} \end{aligned} \tag{A.20}$$

Inserting the bound from Lemma A.1 into (A.20) yields (A.2) after some simple algebra. $\diamondsuit\diamondsuit$

References

[A] D.J. Amit, "Modelling brain: the world of attractor neural networks", Cambridge University Press, Cambridge, 1989.

[AGS1] D.J. Amit, H. Gutfreund and H. Sompolinsky, "Spin-glass models of neural networks", Phys. Rev. **A 32**: 1007-1018 (1985).

[AGS2] D.J. Amit, H. Gutfreund and H. Sompolinsky, "Storing infinite numbers of patterns in a spin glass model of neural networks", Phys. Rev. Letts. **55**: 1530-1533 (1985).

[Bi] P. Billingsley, "Convergence of Probability Measures", Wiley, New York (1968).

[BG] A. Bovier and V. Gayrard, "Rigorous results on the thermodynamics of the dilute Hopfield model", J. Stat. Phys. **72**: 79-112 (1993).

[Co] F. Comets, Large deviation estimates for a conditional probability distribution. Applications to random Gibbs measures, Prob. Theor. Rel. Fields **80**: 407-432 (1989).

[E] R.S. Ellis, Entropy, large deviations, and statistical mechanics, *Springer-Verlag*, Berlin (1985).

[FK] Z. Füredi and J. Komlós, "The eigenvalues of random symmetric matrices", Combinatorica **1**: 233-241 (1981).

[FP1] L.A. Pastur and A.L. Figotin, "Exactly soluble model of a spin glass", Sov. J. Low Temp. Phys. **3(6)**: 378-383 (1977).

[FP2] L.A. Pastur and A.L. Figotin, "On the theory of disordered spin systems", Theor. Math. Phys. **35** 403-414 (1978).

[G] V. Gayrard, The thermodynamic limit of the q-state Potts-Hopfield model with infinitely many patterns, *J. Stat. Phys.***68**: 977-1011 (1992).

[Ge] H.-O. Georgii, Gibbs measures and phase transitions, (Walter de Gruyter, Berlin, 1988).

[H] J.L. van Hemmen, "Spin glass models of a neural network", Phys. Rev. A **34**: 3435-3445 (1986).

[Ho] J.J. Hopfield,"Neural networks and physical systems with emergent collective computational abilities", Proc. Natl. Acad. Sci. USA **79**: 2554-2558 (1982).

[HS] R.L. Stratonovich, "On a method of calculating quantum distribution functions", Doklady Akad. Nauk S.S.S.R.**115**: 1097 (1957)[translation: Soviet Phys. Doklady **2**: 416-419 (1958),

J. Hubbard, "Calculation of partition functions", Phys. Rev. Lett. **3**: 77-78 (1959).

[K] H. Koch, "A free energy bound for the Hopfield model", J. Phys. A : Math Gen. **26** L353-L355 (1993).

[KPa] J. Komlos and R. Paturi, "Convergence results in a autoassociative memory model", Neural Networks **1**:239-250 (1988).

[KP] H. Koch and J. Piasko, "Some rigorous results on the Hopfield neural network model", J. Stat. Phys. **55**: 903 (1989).

[LT] Michel Ledoux and Michel Talagrand, "Probability in Banach Spaces", Springer, Berlin-Heidelberg-New York, 1991.

[MPV] M. Mézard, G. Parisi and M.A. Virasoro, Spin-glass theory and beyond, *World Scientific*, Singapore (1988).

[N] Ch.M. Newman, "Memory capacity in neural network models: Rigorous lower bounds", Neural Networks **1**: 223-238 (1988).

[PST] L. Pastur, M. Shcherbina and B. Tirozzi, "The replica symmetric solution without the replica trick for the Hopfield model", to appear in J. Stat. Phys. (1994)

[Ro] H. Robbins, "A remark on Stirling's formula", Amer. Math. Monthly **62**: 26-29 (1955).

[SK] D. Sherrington and S. Kirkpatrick, "Solvable model of a spin glass", Phys. Rev. Lett. **35**: 1792-1796 (1972).

[ST] M. Sherbina and B. Tirozzi, "The free energy for a class of Hopfield models", J. Stat. Phys. **72**, 113 (1992).

[ScT] E. Scacciatelli and B. Tirozzi, "Fluctuation of the free energy in the Hopfield model", J. Stat. Phys. **67**: 981-1008 (1992).