Probab. Theory Relat. Fields 101, 511–546 (1995)

Probability Theory and Related Fields

Large deviation principles for the Hopfield model and the Kac-Hopfield model*

Anton Bovier[1], **Véronique Gayrard**[2], **Pierre Picco**[2]

[1] Weierstraß-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, D-10117 Berlin, Germany
(e-mail: bovier@iaas-berlin.d400.de)

[2] Centre de Physique Théorique-CNRS, Luminy, Case 907, F-13288 Marseille Cedex 9, France
(e-mail: gayrard@cpt.univ-mrs.fr, picco @cpt.univ-mrs.fr)

Received: 27 April 1994/In revised form:

Summary. We study the Kac version of the Hopfield model and prove a Lebowitz-Penrose theorem for the distribution of the overlap parameters. At the same time, we prove a large deviation principle for the standard Hopfield model with infinitely many patterns.

Mathematics Subject Classification: 60K35, 82B44, 82C32

1 Introduction

In 1977 Figotin and Pastur [11, 12] introduced a class of simplified and exactly solvable models for mean-field spin-glasses, in which the random interaction J_{ij} between two spins was of the form $J_{ij} = \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu$, with ξ_i^μ, $i \in \mathbb{N}$, $\mu \in \{1, \dots, p\}$ a family of independent, identically distributed random variables, taking, in the simplest case, the values $+1$ and -1 with equal probability. While these at first did not receive much attention, the same model was reintroduced in 1983 by Hopfield [18] as a model for autoassociative memory. The interpretation of a disordered spin-system in the context of neuroscience initiated the continuing wave of interest of the physics community in the field of "neural networks". An important new ingredient in Hopfield's version of the model was, however, the interpretation of the vectors

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

ξ^μ, $\mu = 1, \ldots, p$ as a family of "patterns" to be memorized and the fact that the parameter p, the number of stored patterns, is allowed to depend on the size of the system, N. In an important paper, Amit et al. [1], using the replica method for a heuristic analysis of the thermodynamic properties of the model, discovered crucial changes in the behaviour of the model depending on the speed with which $p(N)$ grows to infinity. In particular, they pointed out that there should be a transition to a truly spin-glass like behaviour (interpreted as "failure of the memory"), if $p(N)$ grows faster than $\alpha_c N$, with $\alpha_c \approx 0.14$. Overall, it appeared that, using the speed of growth as a model-parameter, the Hopfield model yields an interesting class of models intermediating between ferromagnets and spin-glasses. Over the last few years, a fairly good mathematical understanding of the thermodynamic properties of this model has been developed, albeit under more restrictive conditions on the growth of $p(N)$ [19, 26, 2–4, 27].

From the point of view of spin systems, the Hopfield model is a mean field model and thus plagued with the typical pathologies of all such models, in particular the non-convexity of thermodynamic functions or the impossibility of implementing the DLR scheme to define the infinite volume Gibbs measures, etc. To overcome these pathologies and to give a natural interpretation of mean field models in terms of limits of "standard" models of statistical mechanics, Kac et al. [22] proposed a model with long, but finite, range interactions, known as the Kac model. Taking the infinite volume limit for such a model first, and then considering the limit as the range of interactions tends to infinity while appropriately rescaling the interaction strength, one then recovers mean field theory. The most precise and complete form of this asymptotic relation was later proven by Lebowitz and Penrose [23]. They showed that the rate function for the total mean magnetization in the Kac model converges, in the limit of infinite interaction range, to the convex hull of the corresponding rate function in the Curie-Weiss model. Such results were later recovered for more complicated mean field models, such as the Curie–Weiss–Potts model (see e.g. [21] for a recent survey).

Nothing is more natural than to consider the same question in the context of the Hopfield model; in fact, the first to introduce and study the Kac-version of the Hopfield model were Figotin and Pastur [13]. They proved, assuming that the number of stored patterns is bounded, the convergence of the free energy of the Kac-version to that of the mean-field Hopfield model (and hence to that of the Curie-Weiss model). The main purpose of the present paper is to extend this result in two ways: First, we want to allow the number of patterns to be an unbounded function of the interaction range, and second, we want to prove the convergence on the level of the rate functions. For an exposition of both the theory of large deviations and mean field models, we refer in particular to the book by Ellis [10].

To do this, we are of course confronted with the problem of proving a large deviation principle (LDP) for the Hopfield model itself. In the case where the number of patterns is bounded, this is not a problem as the existence of an LDP is essentially covered by the classical Gärtner–Ellis theorem [15, 10, 8].

In explicit form this can be found in [17] and in mathematically rigorous form in [6]. As soon as the number of pattern is an unbounded function of the system size, however, standard theorems do not apply anymore, and up to now no LDP was available in this case. An important task of the present paper is therefore to establish a large deviation principle for the Hopfield model with unbounded number of patterns. Before entering into the precise formulation of our results, let us mention one curious fact. We will actually be able to prove directly a large definition principle for the Hopfield model only under the condition that the number of patterns grows more slowly than the logarithm of the system size. By relating the Hopfield model to its Kac-version, it will, however, be possible to extend this result to much more rapidly growing number of patterns, in the sense that at least the convex hull of the rate function still exists in this case. This fact was, for us at least, quite surprising.

Beyond the large deviation results for the mean magnetization, there are a lot of interesting questions to be answered concerning in particular the Gibbs states of the Kac-version of the Hopfield model. Even in the case of the standard Kac model, there are fairly interesting problems related to this, as can be seen in the recent paper by Cassandro et al. [7]. In the case of the Hopfield model, this promises to shed light on various aspects of the properties of spin-glass type models. An investigation of these questions is under way and results will be published elsewhere [5].

We will refer to the Kac version of the Hopfield model as the FHKP-model. Let us now give a precise definition of this model. Since the results we are aiming for in the present paper will be independent of the dimension of the underlying lattice, to simplify notations we will work here in dimension one. For the same reason, we will work only with free boundary conditions.[1] We denote by Λ the set of integers $\Lambda \equiv \{-N, -N+1, \dots, N\}$ and by $\mathcal{S}_\Lambda \equiv \{-1, 1\}^\Lambda$ the space of functions $\sigma : \Lambda \to \{-1, 1\}$. We call σ a *spin configuration* on Λ. We shall write $\mathcal{S} \equiv \{-1, 1\}^{\mathbb{Z}}$ for the space of 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 $\xi = \{\xi_i^\mu\}_{i \in \mathbb{Z}, \mu \in \mathbb{N}}$ be a two-parameter family of independent, identically distributed random variables on this space such that $\mathbb{P}(\xi_i^\mu = 1) = \mathbb{P}(\xi_i^\mu = -1) = \frac{1}{2}$. The Hamiltonian with free boundary conditions on $\mathcal{S}_\Lambda$ is then given by

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

[1] Note, however, that the dimensionality will be important for the properties of the Gibbs states of the model and that more general boundary conditions will have to be considered to study them

where $J_\gamma(i-j) \equiv \frac{\gamma}{2} J(\gamma|i-j|)$, and

$$J(x) = \begin{cases} 1 & \text{if } |x| \leqq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

(Note that other choices for the function $J(x)$ are possible. They must satisfy the conditions $J(x) \geqq 0$, $\int dx\, J(x) = 2$, and must decay rapidly to zero on a scale of order unity. For example, the original choice of Kac was $J(x) = e^{-|x|}$. For us, the choice of the characteristic function is particularly convenient).

We see that the spins in this Hamiltonian interact over a distance γ^{-1}, and we will obtain results for the limit when γ tends to zero. Note that in the FKHP-model we have denoted the number of patterns by M, rather than p. We reserve the name $p \equiv p(N)$ for the number of patterns as a function of the system-size and $M \equiv M(\gamma)$ as the number of patterns as a function of the parameter γ. We are interested in the case where $M(\gamma)\uparrow \infty$, as $\gamma \downarrow 0$. We will set $\alpha(\gamma) \equiv \gamma M(\gamma)$.

The finite volume Gibbs measure for our model is defined by assigning to each $\sigma \in \mathscr{S}_\Lambda$ the mass

$$\mathscr{G}_{\Lambda,\beta,\gamma}[\omega](\sigma) \equiv \frac{1}{Z_{\Lambda,\beta,\gamma}[\omega]} e^{-\beta H_{\Lambda,\gamma}[\omega](\sigma)}, \tag{1.3}$$

where $Z_{\Lambda,\beta,\gamma}[\omega]$ is a normalizing factor usually called *partition function*. We will drop the explicit ω dependence of various random variables when no confusion may arise. For any subset $\Delta \subset \mathbb{Z}$, we define the M-dimensional vector of "overlaps" $m_\Delta[\omega](\sigma)$ whose components are

$$m^\mu_\Delta[\omega](\sigma) = \frac{1}{|\Delta|} \sum_{i\in\Delta} \xi^\mu_i[\omega]\sigma_i, \quad \mu = 1, \ldots, M. \tag{1.4}$$

The main object we will study in this paper are the distributions of $m_\Lambda(\sigma)$ under the Gibbs measure, i.e.

$$\mathscr{Q}_{\Lambda,\beta,\gamma}[\omega](m) = \mathscr{G}_{\Lambda,\beta,\gamma}(\{m_\Lambda[\omega](\sigma) = m\}). \tag{1.5}$$

$\mathscr{Q}_{\Lambda,\beta,\gamma}[\omega]$ defines a random probability measure on $(\mathbb{R}^{M(\gamma)}, \mathscr{B}(\mathbb{R}^{M(\gamma)}))$. For fixed $\gamma > 0$, this sequence of probability measures satisfies a large deviation principle in the sense that for instance the limit[2]

$$\lim_{\varepsilon\downarrow 0} \lim_{N\uparrow\infty} \frac{1}{2N+1} \ln(\mathscr{Q}_{\Lambda,\beta,\gamma}[\omega]\,[\| m_\Lambda - \tilde{m} \|_2^2 \leqq \varepsilon]) \equiv -\beta F_{\beta,\gamma}(\tilde{m}) \tag{1.6}$$

exists almost surely by the subadditive ergodic theorem and is independent of the random parameter ω. Moreover, $F_{\beta,\gamma}(\tilde{m})$ is a convex function of its argument. We will be interested in the limiting behaviour of $F_{\beta,\gamma}$ as $\gamma\downarrow 0$. Since

[2] We comment below on the equivalence of this definition with the conventional one in our case. We find this formulation particularly convenient for our purposes

the domain of this function depends on γ via $M(\gamma)$, it is natural to consider restrictions to finite dimensional cylinders. Thus, let $I \subset \mathbb{N}$ be a finite set and denote by $\Pi_I : \mathbb{R}^M \to \mathbb{R}^I$, for any M such that $I \subset \{1, \ldots, M\}$, the orthogonal projector on the components m^μ, with $\mu \in I$, of a vector $m \in \mathbb{R}^M$. We set, for $\tilde{m} \in \mathbb{R}^I$,

$$F^I_{\beta,\gamma}(\tilde{m}) \equiv -\beta^{-1} \lim_{\varepsilon \downarrow 0} \lim_{N \uparrow \infty} \frac{1}{2N+1} \ln(\mathcal{Q}_{\Lambda,\beta,\gamma}[\omega][\|\Pi_I m_\Lambda - \tilde{m}\|_2^2 \leqq \varepsilon]), \quad (1.7)$$

which enjoy the same properties as $F_{\beta,\gamma}$ itself, and which *potentially* converge to a limit $\gamma \downarrow 0$. The Lebowitz–Penrose theory relates the limit of these quantities to the analogous ones in the corresponding mean-field model, i.e. in our case the standard Hopfield model. Recall that this model is defined by the Hamiltonian in the volume $\Lambda = \{1, \ldots, N\}$,

$$H^{\text{Hopf}}_\Lambda[\omega](\sigma) = -\frac{1}{2N} \sum_{(i,j)\in\Lambda\times\Lambda} \sum_{\mu=1}^{p(N)} \xi_i^\mu[\omega]\xi_j^\mu[\omega]\sigma_i\sigma_j. \quad (1.8)$$

We denote by $\mathcal{G}^{\text{Hopf}}_{N,\beta}[\omega]$ the corresponding Gibbs measure, and by $\mathcal{Q}^{\text{Hopf}}_{N,\beta}[\omega]$ the induced distribution of the overlap parameters $m^\mu_N[\omega](\sigma) \equiv \frac{1}{N}\sum_{i=1}^N \xi_i^\mu[\omega]\sigma_i$. We also write

$$F^{\text{Hopf},I}_\beta(\tilde{m}) \equiv -\beta^{-1} \lim_{\varepsilon \downarrow 0} \lim_{N \uparrow \infty} \frac{1}{N} \ln(\mathcal{Q}^{\text{Hopf}}_{N,\beta}[\omega][\|\Pi_I m_N - \tilde{m}\|_2^2 \leqq \varepsilon]) \quad (1.9)$$

provided this limit exists. Note that, contrary to the case of the Kac-model, there are no simple arguments that prove existence and non-randomness of the limit, and even if it exists, we cannot expect it to be a convex function at low temperature, i.e. if $\beta > 1$. In fact, our results on the FHKP-model will turn out to be extremely useful in order to obtain some partial information on these questions. Let us define the convex functions $C^{\text{Hopf},I}_\beta$, which, if $F^{\text{Hopf},I}_\beta$ exists, are the convex hulls of these functions. Recall that the convex hull, Conv$f$, of a function f, is the largest convex function that is pointwise smaller than or equal to f. We set

$$C^{\text{Hopf},I}_\beta(\tilde{m}) \equiv \beta^{-1} \lim_{\varepsilon \downarrow 0} \lim_{N \uparrow \infty} \frac{1}{N} \text{Conv}(-\ln(\mathcal{Q}^{\text{Hopf}}_{N,\beta}[\omega][\|\Pi_I m_\Lambda - \tilde{m}\|_2^2 \leqq \varepsilon])). \quad (1.10)$$

Notice that the functions $C^{\text{Hopf},I}_\beta$ depend on the asymptotic behaviour of the function $p(N)$. Let $D(I) \subset \mathbb{R}^I$ denote the set of values $\tilde{m}$ for which the limsup on the right hand side of (1.9) is bounded. Our following convergence results hold for $\tilde{m} \in \text{int}\, D(I)$.

Theorem 1 *Suppose that* $p(N)$ *is such that* $\lim_{N\uparrow\infty} p(N) = +\infty$ *and* $\lim_{N\uparrow\infty} p(N)/N = 0$. *Then,*

(*i*) *For almost all* ω, $C^{\text{Hopf},I}_\beta(\tilde{m})$ *defined through* (1.10) *exists for any finite set* $I \subset \mathbb{N}$ *and is independent of* ω, ε, *and the function* $p(N)$.

(*ii*) *If, moreover,* $\lim_{N\uparrow\infty} 2^{p(N)}/N = 0$, *then, almost surely,* $F^{\text{Hopf},I}_\beta(\tilde{m})$ *defined through* (1.9) *exists and is independent of* $p(N)$.

Remark. We will give an explicit expression for the function $C_\beta^{\text{Hopf},I}(\tilde{m})$ in terms of a variational formula in Sect. 3. The independence of this function on the precise behaviour of $p(N)$ is certainly quite surprising and indeed crucial for proving the theorem under such weak conditions on $p(N)$. In fact, we will be able to give an elementary proof only of the statement (ii), while (i) will then follow by passing to the Kac-version of the model.

For the FHKP-model, we obtain the Lebowitz–Penrose theorem:

Theorem 2 *Assume that* $M(\gamma)$ *satisfies* $\lim_{\gamma\downarrow 0} M(\gamma) = +\infty$ *and* $\lim_{\gamma\downarrow 0} \gamma M(\gamma) = 0$. *Then, for any* β, *and any finite subset* I, *for almost all* ω,

$$-\beta^{-1} \lim_{\varepsilon\downarrow 0} \lim_{\gamma\downarrow 0} \lim_{N\uparrow\infty} \frac{1}{2N+1} \ln(\mathscr{Q}_{\Lambda,\beta,\gamma}[\omega][\|\Pi_I m_\Lambda - \tilde{m}\|_2^2 \leqq \varepsilon]) = C_\beta^{\text{Hopf},I}(\tilde{m}). \tag{1.11}$$

Let us make a short comment on our definitions of the rate functions through limits over balls of radius ε. Since all the measures we are considering here have actually compact support, the families of measures are in particular exponentially tight. Thus to prove a strong LDP it is enough to prove a weak one. By appropriately covering closed sets with balls of radius ε, respectively fitting such balls into open sets, one can easily proof the weak LDP with rate functions given as defined above, provided they exist. Also, of course, one can easily obtain the corresponding level-2 LDP by standard arguments. We refrain from entering into these technicalities.

As a simple special result we note that the free energy, $F_{\Lambda,\beta,\gamma}[\omega] \equiv -\frac{1}{\beta N} \ln Z_{\Lambda,\beta,\gamma}[\omega]$ satisfies

Proposition 1.1 *Assume that* $\lim_{\gamma\downarrow 0} \gamma M(\gamma) = 0$. *Then, for almost all* ω,

$$\lim_{\gamma\uparrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} F_{\Lambda,\beta,\gamma}[\omega] = \lim_{N\uparrow\infty} F_{N,\beta}^{\text{Hopf}}[\omega] = F_\beta^{\text{CW}}, \tag{1.12}$$

with $F_\beta^{\text{CW}} \equiv \inf_{x\in\mathbb{R}}(x^2/2 - \beta^{-1} \ln\cosh\beta x)$ *the free energy of the Curie–Weiss model.*

It should be noted that for $\beta > 1$ the rate function for the Hopfield model will not be convex and therefore, that of the FHKP-model will contain a 'flat' horizontal piece. In fact it is known [4] and also follows quite easily from the estimates we will give in Sect. 3 that $F_\beta^{\text{Hopf},I}(\tilde{m})$ takes on its absolute minimum for vectors $\tilde{m}$ which have only one non-zero component of values $\pm a(\beta)$, where $a(\beta)$ is the maximal solution of the equation $x = \tanh\beta x$. Obviously, these are the extremal points of the convex polytop described by the equation

$$\|\tilde{m}\|_1 \leqq a(\beta) \tag{1.13}$$

and it is therefore this polytop on which the limiting rate function of the FHKP-model takes on its minimum value. This information will not be enough to obtain the complete characterization of the Gibbs states, just as in the Hopfield model the mere knowledge of the convex hull of the rate function would not suffice. While we have not been able to prove existence of the rate function itself in the Hopfield model if $p(N) \geqq \ln N/\ln 2$, it is possible to get lower bounds that suffice to determine the Gibbs states [4]. This information will be necessary again for the analysis of the states in the FHKP-model.

The remainder of this paper is organized as follows. In Sect. 2 we construct a block-spin approximation for the Kac-Hamiltonian and give a probabilistic estimate on the error term. As an almost immediate application, we also prove Proposition 1.1. Section 3 is devoted to results on the Hopfield model itself. We prove an exponential estimate on the deviation of the rate function from its mean value and prove statement (ii) of Theorem 1. In Sect. 4 we combine all these results to prove Theorem 2 and statement (i) of Theorem 1.

2 A block-spin approximation

The main step in the analysis of Kac-type models is always to exhibit the dominant part of the Hamiltonian as an effective model on local spin averages ('block-spins') and to show that the remainder gives no contribution to the leading asymptotic behaviour of $\mathcal{Q}_{\Lambda,\beta,\gamma}$ when γ tends to zero. The purpose of the present section is to do this in the case of the FHKP-model.

We introduce a new scale $l(\gamma)$, with the property $l(\gamma)\uparrow\infty$, as $\gamma\downarrow 0$. Further conditions on $l(\gamma)$ will be imposed later. We partition the volume Λ into blocks $\Lambda(x)$, $x\in\Gamma\equiv\{-L,-L+1,\dots,L\}$, of length $l(\gamma)$: $\Lambda=\bigcup_{x=-L}^{L}\Lambda(x)$ where $(2L+1)l(\gamma)=2N+1$ (Here we assume that $(2N+1)/l(\gamma)$ is an integer; thus, in principle, we must choose $l_N(\gamma)$ depending on N in such a way that this is true while $l_N(\gamma)$ converges to $l(\gamma)$ as $N\uparrow\infty$. To simplify our notation, we shall not make this explicit).
This allows us to write

$$H_\Lambda(\sigma)=-\frac{1}{2}\sum_{(x,y)\in\Gamma\times\Gamma}\sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(y)}}J_\gamma(i-j)(\xi_i,\xi_j)\sigma_i\sigma_j, \tag{2.1}$$

where $(\xi_i,\xi_j)\equiv\sum_{\mu=1}^{M(\gamma)}\xi_i^\mu\xi_j^\mu$. Now for $i\in\Lambda(x)$ and $j\in\Lambda(y)$ we write

$$\begin{aligned}J_\gamma(i-j)&=J_\gamma(l(\gamma)(x-y))+(J_\gamma(i-j)-J_\gamma(l(\gamma)(x-y)))\\ &=\frac{1}{l(\gamma)}J_{\gamma l(\gamma)}(x-y)+\Delta J_\gamma(i,j).\end{aligned} \tag{2.2}$$

Using this, we decompose H_Λ into

$$H_\Lambda(\sigma)=H_\Lambda^0(\sigma)+\Delta H_\Lambda(\sigma), \tag{2.3}$$

where

$$H_\Lambda^0(\sigma)\equiv-\frac{1}{2}\sum_{(x,y)\in\Gamma\times\Gamma}l(\gamma)^{-1}J_{\gamma l(\gamma)}(x-y)\sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(y)}}(\xi_i,\xi_j)\sigma_i\sigma_j \tag{2.4}$$

and

$$\Delta H_\Lambda(\sigma) \equiv -\frac{1}{2} \sum_{(x,y)\in\Gamma\times\Gamma} \sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(y)}} \Delta J_\gamma(i-j)(\xi_i,\xi_j)\sigma_i\sigma_j. \tag{2.5}$$

Using that

$$\sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(y)}} (\xi_i,\xi_j)\sigma_i\sigma_j = \sum_{\mu=1}^{M(\gamma)} \left(\sum_{i\in\Lambda(x)} \xi_i^\mu\sigma_i\right)\left(\sum_{j\in\Lambda(y)} \xi_j^\mu\sigma_j\right) = l(\gamma)^2 \sum_{\mu=1}^{M(\gamma)} m^\mu_{\Lambda(x)}(\sigma) m^\mu_{\Lambda(y)}(\sigma), \tag{2.6}$$

we can write H_Λ^0 as

$$H_\Lambda^0(\sigma) \equiv -\frac{1}{2} l(\gamma) \sum_{(x,y)\in\Gamma\times\Gamma} J_{\gamma l(\gamma)}(x-y) \sum_{\mu=1}^{M(\gamma)} m^\mu_{\Lambda(x)}(\sigma) m^\mu_{\Lambda(y)}(\sigma). \tag{2.7}$$

Equation (2.7) makes explicit that H_Λ^0 depends on σ only through the block-variables $m_{\Lambda(x)}(\sigma)$ and thus has the desired form. We will now show that the remainder $\Delta H_\Lambda(\sigma)$ is asymptotically negligible.

The decomposition given here was already used by Figotin and Pastur [13]. They also showed, under the assumption that $M(\gamma) \leqq M < \infty$, that $|\Delta H_\Lambda(\sigma)\|\leqq \text{const.}\, \gamma l(\gamma) MN$, uniformly in σ and uniformly in ξ, which implies that ΔH is negligible if $l(\gamma)$ is chosen such that $\gamma l(\gamma)\downarrow 0$. In order to obtain results for $M(\gamma)$ that tend to infinity with optimal conditions on the allowed speed of growth, we will have to improve on this bound; this will require in particular to replace the *uniform* bound in ξ by an almost sure one. Precisely, we show that:

Lemma 2.1 *For all* $\varepsilon > 0$

$$\mathbb{P}\left[\sup_{\sigma\in\Sigma_\Lambda} \frac{1}{2N+1} |\Delta H_\Lambda(\sigma)| > \varepsilon\right] \leqq 16 \exp\left\{-(2N+1)\left[\frac{\varepsilon}{4\sqrt{2}\gamma l(\gamma)} - \ln 2\right]\right\} \times \exp\left\{\frac{M(\gamma)L}{2}\right\}. \tag{2.8}$$

An immediate consequence of this estimate is the

Corollary 2.2 *There exists a subset* $\Omega_\gamma \subset \Omega$ *of probability one, such that for all* $\omega \in \Omega_\gamma$,

$$\limsup_{N\uparrow\infty} \sup_{\sigma\in\mathscr{S}_\Lambda} \frac{1}{2N+1} |\Delta H_\Lambda[\omega](\sigma)| \leqq \gamma l(\gamma) 4\sqrt{2}\log 2 + \sqrt{2}\gamma M(\gamma). \tag{2.9}$$

Proof. If we choose ε in Lemma 2.1 as $\varepsilon = 4\sqrt{2}\gamma l(\gamma)(\log 2 + M(\gamma)/4l(\gamma) + \delta)$ for some $\delta > 0$, we get

$$\mathbb{P}\left[\frac{1}{2N+1} |\Delta H_\Lambda(\sigma)| \geqq \gamma l(\gamma) 4\sqrt{2}(\log 2 + \delta) + \sqrt{2}\gamma M(\gamma)\right] \leqq 16 e^{-\delta(2N+1)} \tag{2.10}$$

from which 2.9 follows by the first Borel–Cantelli lemma. □

Proof of Lemma 2.1 In order to estimate $\Delta H_\Lambda(\sigma)$, we notice first that

$$J_\gamma(i-j) - l(\gamma)^{-1} J_{\gamma l(\gamma)}(x-y) = \frac{\gamma}{2}\{\mathbb{1}_{\{|i-j|\leqq\gamma^{-1}\}}\mathbb{1}_{\{|x-y|>(\gamma l(\gamma))^{-1}\}} - \mathbb{1}_{\{|i-j|>\gamma^{-1}\}}\mathbb{1}_{\{|x-y|\leqq(\gamma l(\gamma))^{-1}\}}\}. \quad (2.11)$$

Moreover,

$$\mathbb{1}_{\{|i-j|\leqq\gamma^{-1}\}}\mathbb{1}_{\{|x-y|>(\gamma l(\gamma))^{-1}\}} = \mathbb{1}_{\{|i-j|\leqq\gamma^{-1}\}}\mathbb{1}_{\{(\gamma l(\gamma))^{-1}+1\geqq|x-y|>(\gamma l(\gamma))^{-1}\}} \quad (2.12)$$

and

$$\mathbb{1}_{\{|i-j|>\gamma^{-1}\}}\mathbb{1}_{\{|x-y|\leqq(\gamma l(\gamma))^{-1}\}} = \mathbb{1}_{\{|i-j|>\gamma^{-1}\}}\mathbb{1}_{\{(\gamma l(\gamma))^{-1}\geqq|x-y|>(\gamma l(\gamma))^{-1}-1\}}. \quad (2.13)$$

We now write $\Delta H_\Lambda(\sigma) = (\gamma/2)[\Delta^1 H_\Lambda(\sigma) - \Delta^2 H_\Lambda(\sigma)]$ with

$$\Delta^1 H_\Lambda(\sigma) \equiv -\frac{1}{2}\sum_{(x,y)\in\Gamma\times\Gamma}\sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(y)}}\mathbb{1}_{\{|i-j|\leqq\gamma^{-1}\}} \times \mathbb{1}_{\{(\gamma l(\gamma))^{-1}+1\geqq|x-y|>(\gamma l(\gamma))^{-1}\}}(\xi_i,\xi_j)\sigma_i\sigma_j \quad (2.14)$$

and

$$\Delta^2 H_\Lambda(\sigma) \equiv -\frac{1}{2}\sum_{(x,y)\in\Gamma\times\Gamma}\sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(y)}}\mathbb{1}_{\{|i-j|>\gamma^{-1}\}} \times \mathbb{1}_{\{(\gamma l(\gamma))^{-1}\geqq|x-y|>(\gamma l(\gamma))^{-1}-1\}}(\xi_i,\xi_j)\sigma_i\sigma_j. \quad (2.15)$$

We only present the estimate of $\Delta^1 H_\Lambda(\sigma)$; $\Delta^2 H_\Lambda(\sigma)$ can be treated in exactly the same way. We have

$$\Delta^1 H_\Lambda(\sigma) = -\sum_{x=-L}^{L-[(\gamma l(\gamma))^{-1}]-1}\sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(x+[(\gamma l(\gamma))^{-1}]+1)}} \times \mathbb{1}_{\{|i-j|\leqq\gamma^{-1}\}}(\xi_i,\xi_j)\sigma_i\sigma_j. \quad (2.16)$$

Let us set

$$f^\mu(x) \equiv \sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(x+[(\gamma l(\gamma))^{-1}]+1)}}\mathbb{1}_{\{|i-j|\leqq\gamma^{-1}\}}\xi_i^\mu\xi_j^\mu\sigma_i\sigma_j \quad (2.17)$$

and

$$f(x) \equiv \sum_{\mu=1}^{M(\gamma)} f^\mu(x). \quad (2.18)$$

Since $L \equiv [(\gamma l(\gamma))^{-1}]n + r$ with $0 \leqq r < [(\gamma l(\gamma))^{-1}]$ for some positive integer n, where $[x]$ denotes the integer part of x, we can rewrite (2.16) as

$$\sum_{x=-L}^{L-[(\gamma l(\gamma))^{-1}]-1} f(x) = \sum_{x=-[(\gamma l(\gamma))^{-1}]n}^{[(\gamma l(\gamma))^{-1}](n-1)} f(x) + \sum_{x=-L}^{-[(\gamma l(\gamma))^{-1}]n-1} f(x) + \sum_{x=[(\gamma l(\gamma))^{-1}](n-1)+1}^{L-[(\gamma l(\gamma))^{-1}]-1} f(x). \quad (2.19)$$

Let us consider the first sum in (2.19) first. In order to take into account the independence of (ξ_i, ξ_j) and (ξ'_i, ξ'_j) when $i \neq i'$ and $j \neq j'$ we first decompose the first sum in (2.19) in the following way:

$$\sum_{x=-[(\gamma l(\gamma))^{-1}]n}^{[(\gamma l(\gamma))^{-1}](n-1)} f(x) = \sum_{k=-[n/2]}^{[n/2]-1} \sum_{s=0}^{[1/(\gamma l(\gamma))]-1} f\left(\left[\frac{1}{\gamma l(\gamma)}\right] 2k + s\right)$$

$$+ \sum_{k=-[(n+1)/2]}^{[(n+1)/2]-2} \sum_{s=0}^{[1/(\gamma l(\gamma))]-1} f\left(\left[\frac{1}{\gamma l(\gamma)}\right](2k+1) + s\right). \tag{2.20}$$

The important point to observe here is that each of the two terms in (2.20) is now a sum of independent random variables. Let us denote these two terms by S_1 and S_2, respectively. We have

$$\mathbb{P}\left[\sup_{\sigma \in \mathcal{S}_\Lambda} \frac{1}{2N+1} \frac{\gamma}{2} |S_1| \geqq \frac{\varepsilon}{8}\right]$$

$$\leqq 2 2^{2N+1} \mathbb{P}\left[\sum_{\mu=1}^{M(\gamma)} \sum_{k=-[n/2]}^{[n/2]-1} \sum_{s=0}^{[1/(\gamma l(\gamma))]-1} f^\mu\left(\left[\frac{1}{\gamma l(\gamma)}\right] 2k + s\right)\right.$$

$$\left. \geqq \frac{\varepsilon}{4} \gamma^{-1}(2N+1)\right] \tag{2.21}$$

where the probability on the right-hand side is in fact independent of the chosen spin configuration σ. Using the exponential Markov inequality together with the independence, we get that

$$\mathbb{P}\left[\sup_{\sigma \in \mathcal{S}_\Lambda} \frac{1}{2N+1} \frac{\gamma}{2} |S_1| \geqq \frac{\varepsilon}{8}\right] \leqq 2^{2N+2} \inf_{t \geqq 0} e^{-t\varepsilon\gamma^{-1}(2N+1)/4}$$

$$\times [\mathbb{E} e^{tf^1(0)}]^{2M[n/2][1/\gamma l(\gamma)]}$$

$$\leqq 2^{2N+2} \inf_{t \geqq 0} e^{-t\varepsilon\gamma^{-1}(2N+1)/4} [\mathbb{E} e^{tf^1(0)}]^{ML}. \tag{2.22}$$

Thus we have to estimate the Laplace-transform of $f^1(0)$. We write

$$\mathbb{E} e^{tf^1(0)} = \mathbb{E} \exp\left\{ t \sum_{i \in \Lambda(0)} \xi_i^1 \sum_{\substack{j \in \Lambda([1/(\gamma l(\gamma))]+1) \\ |i-j| < \gamma^{-1}}} \xi_j^1 \right\}. \tag{2.23}$$

Notice that all the ξ_i^1 with $i \in \Lambda(0)$ are independent of the ξ_j^1 with $j \in \Lambda([1/(\gamma l(\gamma))] + 1)$, and that, conditioned on these latter variables, the variables $\xi_i^1 \sum_{j \in \Lambda([1/(\gamma l(\gamma))]+1), |i-j| < \gamma^{-1}} \xi_j^1$ are independent. If we denote by $\mathbb{E}_i$ the expectation w.r.t. ξ_i^1, this allows us to write

$$\mathbb{E}e^{tf^1(0)} = \mathbb{E}\prod_{i\in\Lambda(0)}\mathbb{E}_i \exp\left\{t\xi_i^1 \sum_{\substack{j\in\Lambda[1/(\gamma l(\gamma))]+1) \\ |i-j|<\gamma^{-1}}} \xi_j^1\right\}$$

$$\leqq \mathbb{E}\prod_{i\in\Lambda(0)}\exp\left\{\frac{1}{2}t^2\left(\sum_{\substack{j\in\Lambda([1/(\gamma l(\gamma))]+1) \\ |i-j|<\gamma^{-1}}} \xi_j^1\right)^2\right\}, \tag{2.24}$$

where we have used that $\ln\cosh x \leqq \frac{1}{2}x^2$. Using the Hölder-inequality on the last line, we arrive at

$$\mathbb{E}e^{tf^1(0)} \leqq \prod_{i\in\Lambda(0)}\left[\mathbb{E}\exp\left\{\tfrac{1}{2}l(\gamma)t^2\left(\sum_{\substack{j\in\Lambda([1/(\gamma l(\gamma))]+1) \\ |i-j|<\gamma^{-1}}} \xi_j^1\right)^2\right\}\right]^{1/l(\gamma)}. \tag{2.25}$$

Now

$$\mathbb{E}\exp\left\{\tfrac{1}{2}l(\gamma)t^2\left(\sum_{\substack{j\in\Lambda([1/(\gamma l(\gamma))]+1) \\ |i-j|<\gamma^{-1}}} \xi_j^1\right)^2\right\} \leqq \mathbb{E}\exp\left\{\tfrac{1}{2}l(\gamma)t^2\left(\sum_{j\in\Lambda([1/(\gamma l(\gamma))]+1)} \xi_j^1\right)^2\right\}$$

$$\leqq \frac{1}{\sqrt{1-t^2 l(\gamma)^2}}, \tag{2.26}$$

where we have used the Khintchine inequality and the fact that $|\Lambda([1/\gamma l(\gamma)]+1)| \leqq l(\gamma)$. Since for $0 \leqq x \leqq \frac{1}{2}$, $1/\sqrt{1-x} \leqq e^x$, for $t^2 \leqq 1/2l(\gamma)^2$, we can replace (2.26) by the more convenient bound

$$\mathbb{E}e^{tf^1(0)} \leqq e^{t^2 l(\gamma)^2}. \tag{2.27}$$

Therefore, choosing $t = 1/\sqrt{2}l(\gamma)$ in (2.22), we obtain

$$\mathbb{P}\left[\sup_{\sigma\in\mathcal{S}_\Lambda}\frac{1}{2N+1}\frac{\gamma}{2}|S_1| \geqq \frac{\varepsilon}{8}\right] \leqq 2^{2N+2}\exp\left\{-\frac{1}{4\sqrt{2}l(\gamma)\gamma}\varepsilon(2N+1)+\tfrac{1}{2}ML\right\}. \tag{2.28}$$

By the same procedure, one obtains exactly the same bound for S_2. It remains to consider the two last sums in (2.19). Obviously, they are much smaller than S_1 or S_2 and can be treated in the same manner. Finally, $\Delta^2 H_\Lambda$ is decomposed in the same way, so that we end up with eight terms all of which satisfy bounds like (2.28). Putting these together concludes the proof of Lemma 2.1. □

To understand the need for Corollary 2.2, let us anticipate that we may be able to treat H^0 further provided $M(\gamma)/l(\gamma)\downarrow 0$. Then, if only $\gamma M(\gamma)\downarrow 0$, we can choose e.g. $l(\gamma) = \sqrt{M(\gamma)/\gamma}$ to achieve that $M(\gamma)/l(\gamma)\downarrow 0$ while at the same time $|\Delta H(\sigma)|/N\downarrow 0$, a.s. by Corollary 2.2. If, on the other hand, we had only the uniform bound $\gamma l(\gamma) M(\gamma) N$ on ΔH, then we would have to demand $\gamma [M(\gamma)]^2 \downarrow 0$, which is a much stronger, and quite unnatural, restriction on the number of patterns.

Thus the Hamiltonian of the FHKP-model is asymptotically equivalent to a block-spin Hamiltonian if $\gamma M(\gamma) \downarrow 0$. But it is more or less clear that the bounds in Lemma 2.1 cannot be substantially improved, and that therefore, once this condition is no longer satisfied, such an approximation breaks down. This sheds doubt on whether in such situations (which would also include real spin-glasses), mean field models can be seen as limits of ordinary models with diverging interaction range!

As a simple first application of Corollary 2.1, let us give at this point a short proof of Proposition 1.1.

Proof of Proposition 1.1 By Corollary 2.2, it will be enough if we can compute the behaviour of

$$\tilde{Z}_{\Lambda,\beta,\gamma}[\omega] \equiv 2^{-(2N+1)} \sum_{\sigma \in \mathscr{S}_\Lambda} e^{-\beta H_\Lambda^0(\sigma)}. \tag{2.29}$$

Using that $J_{xy} = J_{\gamma l(\gamma)}(x-y)$ is a positive definite quadratic form, it follows that

$$\begin{aligned}
\tilde{Z}_{\Lambda,\beta,\gamma}[\omega] &\geqq 2^{-(2N+1)} \sum_{\sigma \in \mathscr{S}_\Lambda} \exp\left(\tfrac{1}{2} \sum_{(x,y)\in\Gamma\times\Gamma} l(\gamma)^{-1} J_{\gamma l(\gamma)}(x-y) \sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(y)}} \xi_i^1 \xi_j^1 \sigma_i \sigma_j \right) \\
&= 2^{-(2N+1)} \sum_{\sigma \in \mathscr{S}_\Lambda} \exp\left(\tfrac{1}{2} \sum_{(x,y)\in\Gamma\times\Gamma} l(\gamma)^{-1} J_{\gamma l(\gamma)}(x-y) \sum_{\substack{i\in\Lambda(x)\\ j\in\Lambda(y)}} \sigma_i \sigma_j \right) \\
&\geqq \sum_{m\in\mathscr{W}_{l(\gamma)}} \prod_{x\in\Gamma} 2^{-l(\gamma)} \sum_{\sigma\in\mathscr{S}_{\Lambda(x)}} \mathbb{1}_{\{2^{-l(\gamma)}\sum_{i\in\Lambda(x)}\sigma_i = m\}} e^{\beta l(\gamma)(1/2)m^2} \\
&\geqq \sup_{m\in\mathscr{W}_{(l\gamma)}} [Z^{\mathrm{CW}}_{l(\gamma),\beta}(m)]^{2L+1},
\end{aligned} \tag{2.30}$$

where $Z^{\mathrm{CW}}_{l,\beta}(m) \equiv \sum_{\sigma\in\mathscr{S}_l} \mathbb{1}_{\{2^{-l}\sum_{i=1}^l \sigma_i = m\}} e^{\beta l(1/2)m^2}$ is the restricted partition function of the Curie–Weiss model with volume l and $\mathscr{W}_l = \{-1, -1+2/l, \ldots, 1-2/l, 1\}$ denotes the set of possible values of $m_l(\sigma) = 1/l \sum_{i=1}^l \sigma_i$. This yields

$$\liminf_{\Lambda\uparrow\mathbb{Z}} \frac{1}{2N+1} \ln \tilde{Z}_{\Lambda,\beta,\gamma}[\omega] \geqq \operatorname*{Sup}_{m\in\mathscr{W}_{l(\gamma)}} \frac{1}{l(\gamma)} \ln Z^{\mathrm{CW}}_{l(\gamma),\beta}(m). \tag{2.31}$$

On the other hand, using the fact that

$$m^\mu_{\Lambda(x)}(\sigma) m^\mu_{\Lambda(y)}(\sigma) \leqq \tfrac{1}{2}(m^\mu_{\Lambda(x)}(\sigma))^2 + \tfrac{1}{2}(m^\mu_{\Lambda(y)}(\sigma))^2 \tag{2.32}$$

and (2.7), we see that

$$\begin{aligned}
H_\Lambda^0(\sigma) &\geqq -\frac{1}{2} l(\gamma) \sum_{(x,y)\in\Gamma\times\Gamma} J_{\gamma l(\gamma)}(x-y) \sum_{\mu=1}^{M(\gamma)} \left(\tfrac{1}{2}(m^\mu_{\Lambda(x)}(\sigma))^2 + \tfrac{1}{2}(m^\mu_{\Lambda(y)}(\sigma))^2\right) \\
&= -\frac{1}{2} l(\gamma) \sum_{x\in\Gamma} \sum_{\mu=1}^{M(\gamma)} (m^\mu_{\Lambda(x)}(\sigma))^2,
\end{aligned} \tag{2.33}$$

so that

$$\tilde{Z}_{\Lambda,\beta,\gamma}[\omega] \leqq 2^{-(2N+1)} \sum_{\sigma\in\mathcal{S}_\Lambda} \prod_{x\in\Gamma} \exp\left(\beta\tfrac{1}{2}l(\gamma)\sum_{\mu=1}^{M(\gamma)}(m^{\mu}_{\Lambda(x)}(\sigma))^2\right)$$

$$= \prod_{x\in\Gamma} 2^{-l(\gamma)} \sum_{\sigma\in\mathcal{S}_\Lambda(x)} \exp\left(\beta\tfrac{1}{2}l(\gamma)\sum_{\mu=1}^{M(\gamma)}(m^{\mu}_{\Lambda(x)}(\sigma))^2\right)$$

$$= \prod_{x\in\Gamma} Z^{\text{Hopf}}_{l(\gamma),\beta}[\omega_x]. \tag{2.34}$$

But this implies that

$$\limsup_{\Lambda\uparrow\mathbb{Z}} \frac{1}{2N+1} \ln \tilde{Z}_{\Lambda,\beta,\gamma}[\omega] \leqq \mathbb{E}\frac{1}{l(\gamma)} \ln Z^{\text{Hopf}}_{l(\gamma),\beta}[\omega], \tag{2.35}$$

where we have used the strong law of large numbers to replace the spatial average over Γ by the expectation over ξ. If now $l(\gamma)$ is chosen such that $M(\gamma)/l(\gamma)\downarrow 0$, while $l(\gamma)\uparrow\infty$, by a result of Koch [19], the right-hand side of (2.35) converges, as $\gamma\downarrow 0$, to the negative of the free energy of the Curie-Weiss model, as does, obviously, the right hand side of (2.31) (see [10]). This proves that

$$-\beta^{-1} \lim_{\gamma\downarrow\infty} \lim_{\Lambda\uparrow\mathbb{Z}} \frac{1}{2N+1} \ln \tilde{Z}_{\Lambda,\beta,\gamma}[\omega] = F^{\text{CW}}_{\beta}. \tag{2.36}$$

Corollary 2.2 on the other hand implies immediately that for any sequence γ_n tending to zero as $n\uparrow\infty$,

$$\lim_{n\uparrow\infty} |F_{\beta,\gamma_n}[\omega] - \tilde{F}_{\beta,\gamma_n}[\omega]| = 0 \tag{2.37}$$

for almost all ω (namely for those in the set $\bigcap_{n\in\mathbb{N}}\Omega_{\gamma_n}$), from which one obtains (1.12) for the subsequence γ_n. There remains in principle the possibility that (2.37) holds with probability one for any given sequence, while it *fails* with probability one for *some* (random) sequence. However, this is excluded by the following

Lemma 2.3

$$\lim_{n\uparrow\infty} \sup_{1/n\leqq\gamma\leqq 1/n+1} \sup_{\omega\in\Omega} \sup_{\Lambda} \sup_{\sigma\in\mathcal{S}_\Lambda} \frac{1}{2N+1} |H_{\Lambda,\gamma}[\omega](\sigma) - H_{\Lambda,1/n}[\omega](\sigma)| = 0. \tag{2.38}$$

Proof. To prove (2.38), notice that

$$|H_\Lambda[\omega](\sigma) - H^0_\Lambda[\omega](\sigma)|$$

$$\leqq \frac{1}{2} \sum_{(i,j)\in\Lambda\times\Lambda} \left| \sum_{\mu=1}^{M(\gamma)} \xi^{\mu}_i[\omega]\xi^{\mu}_j[\omega]J_\gamma(i-j)\sigma_i\sigma_j \right.$$

$$\left. - \sum_{\mu=1}^{M(1/n)} \xi^{\mu}_i[\omega]\xi^{\mu}_j[\omega]J_{1/n}(i-j)\sigma_i\sigma_j \right|$$

$$\leqq (2N+1)\Bigg[|M(1/(n+1) - M(1/n)| + M(1/n)n\left(\frac{1}{n} - \frac{1}{n+1}\right) + \frac{M(1/n)}{n}\Bigg]. \tag{2.39}$$

The coefficients of N in all three terms vanish, as $N\uparrow\infty$ under our assumptions on $M(\gamma)$ which proves (2.38) and the lemma. □

Combining Lemma 2.3 with (2.37) and (2.36) gives immediately Proposition 1.1. □

3 Large deviation results for the Hopfield model

In this chapter we provide the large deviation results for the standard Hopfield model that will be needed to obtain the analogous ones for the FHKP-model. At a later stage, this will in turn allow us to improve the results for the Hopfield model itself. There are two results that will be given here. The first is a result on the self-averaging properties of the large deviation rate function under the assumption $p(N)/N\downarrow 0$, as $N\uparrow\infty$. The second is a large deviation theorem for the Hopfield model under the strong assumption $p(N) < \ln N/\ln 2$. However, to do this we first have to provide some a priori large deviation estimates for the Hopfield model.

3.1 Large deviation estimates for the Hopfield model

Let us consider the quantities

$$Z^{\text{Hopf}}_{N,\beta,\rho}(m) \equiv 2^{-N} \sum_{\sigma\in\mathcal{S}_N} e^{\frac{\beta N}{2}\|m_N(\sigma)\|_2^2}\mathbb{1}_{\{\|m_N(\sigma)-m\|_2\leqq\rho\}} \tag{3.1}$$

We first proof large deviation upper and lower bounds. We define the functions

$$\Psi_{N,\beta}(m,t) \equiv (m,t) - \frac{1}{2}\|m\|_2^2 - \frac{1}{\beta N}\sum_{i=1}^{N}\ln\cosh\beta(\xi_i,t) \tag{3.2}$$

and

$$\Phi_{N,\beta}(m) \equiv \sup_{t\in\mathbb{R}^p}\Psi_{N,\beta}(m,t) \tag{3.3}$$

We also define $t^*(m)$ through $\Psi_{N,\beta}(m,t^*(m)) = \sup_{t\in\mathbb{R}^p}\Psi_{N,\beta}(m,t)$, if such a t^* exists, while otherwise $\|t^*\|\equiv\infty$. Properties of $t^*(m)$ will be discussed in Lemma 3.2. With these notations we have

Lemma 3.1 *Set* $f_{N,\beta,\rho}(m) \equiv -1/\beta N\ln Z_{N,\beta,\rho}(m)$. *Then*

$$f_{N,\beta,\rho}(m) \geqq \Phi_{N,\beta}(m) - \rho(\|t^*\|_2 + \|m\|_2 + \rho) \tag{3.4}$$

and for $\rho \geqq \sqrt{2\alpha}$,

$$f_{N,\beta,\rho}(m) \leqq \Phi_{N,\beta}(m) + \rho(\|t^*\|_2 + \|m\|_2 + \rho) + \frac{\ln 2}{\beta N} \tag{3.5}$$

Proof. We first prove the lower bound. Note that for arbitrary $t \in \mathbb{R}^p$,

$$\mathbb{1}_{\{\|m_N(\sigma) - m\|_2 \leqq \rho\}} \leqq \mathbb{1}_{\{\|m_N(\sigma) - m\|_2 \leqq \rho\}}\, e^{\beta N(t,(m_N(\sigma) - m)) + \rho\beta N \|t\|_2} \tag{3.6}$$

Thus

$$\begin{aligned} Z_{N,\beta,\rho}(m) &\leqq \inf_{t \in \mathbb{R}^p} 2^{-N} \sum_{\sigma \in \mathcal{S}_N} e^{\beta N \frac{1}{2}(\|m\|_2^2 + 2\rho\|m\|_2 + \rho^2)}\, e^{\beta N(t,(m_N(\sigma) - m)) + \beta N\rho\|t\|_2} \\ &\leqq \inf_{t \in \mathbb{R}^p} e^{\beta N[\frac{1}{2}\|m\|_2^2 - (m,t) + \frac{1}{\beta N}\sum_{i=1}^N \ln\cosh(\beta(\xi_i,t))]} e^{\beta N\rho(\|m\|_2 + \|t\|_2 + \rho)} \end{aligned} \tag{3.7}$$

This gives immediately (3.4). □

Remark. Note that to determine the supremum in (3.3) we would need to solve the system of equations

$$m^\mu = \frac{1}{N}\sum_{i=1}^N \xi_i^\mu \tanh\beta(\xi_i, t) \tag{3.8}$$

which is quite difficult. Note that at the critical points of $\Phi_{N,\beta}(m)$ we have that $t^*(m) = m$.

Next we prove the upper bound. It is of course enough to consider the case where $t^*(m)$ is finite. Thus we define, for $t^* \in \mathbb{R}^p$, the probability measures $\tilde{\mathbb{P}}$ on $\{-1,1\}^N$ through their expectation $\tilde{\mathbb{E}}_\sigma$, given by

$$\tilde{\mathbb{E}}_\sigma(\cdot) \equiv \frac{\mathbb{E}_\sigma e^{\beta N(t^*, m_N(\sigma))}(\cdot)}{\mathbb{E}_\sigma e^{\beta N(t^*, m_N(\sigma))}} \tag{3.9}$$

we have obviously that

$$\begin{aligned} Z_{N,\beta,\rho}(m) &= \tilde{\mathbb{E}}_\sigma e^{\frac{\beta N}{2}\|m_N(\sigma)\|_2^2 - \beta N(t^*, m_N(\sigma))}\mathbb{1}_{\{\|m_N(\sigma) - m\|_2 \leqq \rho\}} \mathbb{E}_\sigma e^{\beta N(t^*, m_N(\sigma))} \\ &\geqq e^{-\beta N(t^*,m) - \beta N(\|t^*\|_2\rho - \frac{1}{2}\|m\|_2^2 + \rho\|m\|_2 + \rho^2/2)}\mathbb{E}_\sigma e^{\beta N(t^*, m_N(\sigma))}\tilde{\mathbb{E}}_\sigma \mathbb{1}_{\{\|m_N(\sigma) - m\|_2 \leqq \rho\}} \\ &= e^{\beta N(\frac{1}{2}\|m\|_2^2 - (t^*,m) + \frac{1}{\beta N}\sum_{i=1}^N \ln\cosh\beta(\xi_i, t^*))} e^{-\beta N\rho(\|t^*\|_2 + \|m\|_2 + \rho/2)} \\ &\quad \times \tilde{\mathbb{P}}[\|m_N(\sigma) - m\|_2 \leqq \rho] \end{aligned} \tag{3.10}$$

But, using Chebychev's inequality, we have that

$$\tilde{\mathbb{P}}[\|m_N(\sigma) - m\|_2 \leqq \rho] \geqq 1 - \frac{1}{\rho^2}\tilde{\mathbb{E}}_\sigma\|m_N(\sigma) - m\|_2^2 \tag{3.11}$$

We choose $t^*(m)$ that satisfies Eq. (3.7): Then it is easy to compute

$$\tilde{\mathbb{E}}\|m_N(\sigma) - m\|_2^2 = \frac{p}{N}\left(1 - \frac{1}{N}\sum_{i=1}^N \tanh^2(\beta(\xi_i, t^*(m)))\right) \tag{3.12}$$

from which the lemma follows. □

In the following lemma we collect a few properties of $\Phi_{N,\beta}(m)$ that arise from convexity. We set $T \equiv \{m \in \mathbb{R}^M | t^*(m) \text{ exists}\}$, $D \equiv \{m \in \mathbb{R}^M | \Phi_{N,\beta}(m) < \infty\}$, and we denote by $\operatorname{int} D$ the interior of D. We moreover denote by $I(x) \equiv \sup_{t \in \mathbb{R}}(tx - \ln\cosh t)$ the Legendre transform of the function $\ln\cosh t$. A simple computation shows that

$$I(x) = \begin{cases} \frac{1+x}{2}\ln(1+x) + \frac{1-x}{2}\ln(1-x), & \text{if } |x| \leqq 1 \\ +\infty, & \text{otherwise} \end{cases} \tag{3.13}$$

Lemma 3.2

i)

$$\Phi_{N,\beta}(m) = -\frac{1}{2}\|m\|_2^2 + \inf_{y \in \mathbb{R}^N : m_N(y) = m} \frac{1}{\beta N}\sum_{i=1}^N I(y_i) \tag{3.14}$$

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

ii)

$$D = \{m \in \mathbb{R}^M \,|\, \exists y \in [-1,1]^N \, s.t. \, m_N(y) = m\} \tag{3.15}$$

iii) $\Phi_{N,\beta}(m)$ is continuous relative to $\operatorname{int} D$

iv) $T = \operatorname{int} D$

Note that point i) of Lemma 3.2 provides the following alternative formula for the variational formula (3.3),

$$\Phi_{N,\beta}(m) = \inf_{y \in \mathbb{R}^N : m_N(y) = m}\left(-\frac{1}{2}\|m_N(y)\|_2^2 + \frac{1}{\beta N}\sum_{i=1}^N I(y_i)\right). \tag{3.16}$$

Proof. Note that the function $g(t) \equiv \frac{1}{\beta N}\sum_{i=1}^N \ln\cosh\beta(\xi_i, t)$ is a proper convex function on $\mathbb{R}^M$. Denoting by $h(m) \equiv \sup_{t \in \mathbb{R}^M}\{(m,t) - g(t)\}$ its Legendre transform, it follows from standard results of convex analysis that $h(m)$ is a proper convex function on $\mathbb{R}^M$ and that

$$h(m) = \inf_{y \in \mathbb{R}^N : m_N(y) = m} \frac{1}{\beta N}\sum_{i=1}^N I(y_i) \tag{3.17}$$

where for each $m \in \mathbb{R}^M$ the infimum is either attained or is $+\infty$. This immediately yields i). Denoting by $\operatorname{dom} h \equiv \{x \in \mathbb{R}^M \,|\, h(m) < \infty\}$ the effective domain of h, we have, by (3.13), that $\operatorname{dom} h$ equals the right hand side of (3.15), and since $-\|m\|_2^2 \leqq 0$, ii) is proven. iii) simply follows from the fact that h being convex, it is continuous relative to the interior of $\operatorname{dom} h$. Finally, to prove iv), we will make use of the following two important results of convex analysis. First, the subgradiant of h at m, $\partial h(m)$, is a non empty set if and only if m belongs to the interior of $\operatorname{dom} h$, i.e., $m \in \operatorname{int} D$. $\partial h(m)$ is moreover a bounded convex set. Next, $(m,t) - g(t)$ achieves its supremum at $t^* \equiv t^*(m)$ if and only if $t^* \in \partial h(m)$. This concludes the proof of the lemma. □

3.2 Self averaging of the rate function

We will now derive the self-averaging properties of the large deviation rates introduced above, or rather of some slightly modified versions of them. These are the results actually needed in the proofs of our Theorems, while the estimates above are needed to establish them.

Let us make a few remarks on the questions of self-averaging in general. The central rôle played by 'self-averaging' properties in disordered systems has particularly been emphasized by Pastur and Shcherbina in their work on the Sherrington-Kirkpatrick model [25]. Shcherbina and Tirozzi [27], and Pastur et al. [26] have recently performed the same analysis for the Hopfield model. Basically, they prove two types of results:

(i) The free energy, $F^{\text{Hopf}}_{N,\beta}$ as well as its derivatives, such as the Gibbsian mean of $m^{\mu}_N(\sigma)$ are self-averaging for all choices of $p(N)$ in the sense that their variance behaves like N^{-1}.
(ii) If the so-called 'Edwards-Anderson' parameter $q_N \equiv \frac{1}{N}\sum_{i=1}^{N} \langle \sigma_i \rangle^2_{N,\beta}$, where $\langle \cdot \rangle_{N,\beta}$ denotes expectation w.r.t. the Gibbs measure of the Hopfield model, has variance of order $1/N$, then a certain set of 'mean-field equations', that can be formally derived using the so-called 'replicatrick' (see [1]), are exact.

As we will see later, here we are in need of self-averaging results on the rate functions. While it is fairly easy to prove results on level of the variances for fixed argument, along the same lines as in the above cited papers, such estimates would not be sharp enough for our purposes, since we will require results that hold uniformly in the arguments. Thus, we must extend the variance estimate to exponential estimates.
For technical reasons, we consider here somewhat softened versions of the functions introduced earlier in which the characteristic function $\mathbb{1}_{\{\|m_A - m\|_2 \leqq \rho\}}$ is replaced by a smooth version of this function. We let $\chi_\rho(x)$ be a differentiable function satisfying:

(1) $\chi_\rho(x) \geqq 0$,

(2) $|\chi'_\rho(x)| \leqq 2N\beta \mathbb{1}_{\{\rho \leqq x \leqq \rho + 1/(\beta N)\}}$,

(3) $\mathbb{1}_{\{|x| \leqq \rho\}} \leqq \chi_\rho(x) \leqq \mathbb{1}_{\{|x| \leqq \rho + 1/(\beta N)\}}$

(4) $\ln \chi(x)$ is concave.

Let us introduce, for $m \in \mathbb{R}^{p(N)}$, the functions

$$\tilde{F}^{\text{Hopf}}_{N,\beta,\rho}(m) \equiv -\frac{1}{\beta N} \ln\left(\frac{1}{2^N} \sum_{\sigma \in \mathcal{S}_N} e^{-\beta H_N(\sigma)} \chi_\rho(\|m_N(\sigma) - m\|_2)\right) \tag{3.18}$$

These are the non-normalized versions of the rate function that differ from the corresponding normalized functions $F^{\text{Hopf}}_{N,\beta,\rho}(m)$ by the free energy. We remark at this point that we will be interested in this quantity only for very (but not too) small ρ (in fact $\rho \sim p/N$). In this case, we need only to consider m with, e.g., $\|m\|_2 \leqq 2$. For, $F^{\mathrm{Hopf}}_{N,\beta,\rho}(m)$ can only be different from $+\infty$ if there exists a $\sigma \in \mathscr{S}_N$ such that $N^{-1}\sum_i \xi_i \sigma_i \approx m$. But this implies that $\|m\|_2^2 \approx \sum_\mu \sum_{i,j} \sigma_i \frac{\xi_i^\mu \xi_j^\mu}{N^2} \sigma_j \leqq \|B\|$, where B is the random matrix with elements $B_{ij} \equiv \sum_\mu \frac{\xi_i^\mu \xi_j^\mu}{N^2}$. It has been proved for instance in [27, 2]) that with probability larger than $1 - e^{-N^{1/6}r}$, $\|B\| \leqq 1 + 2r\sqrt{p/N}$.

Lemma 3.3 *For all $m \in \mathbb{R}^{p(N)}$ with $\|\tilde{m}\|_2 \leqq 2$ that satisfy $\|t^*(m)\| < \infty$, there exists ρ with $\sqrt{\frac{p}{N}} \leqq \rho \leqq 4\sqrt{\frac{p}{N}}$, (possibly depending on m), and a constant $C(\tilde{m}) < \infty$ such that and for $z > 0$,*

$$\mathbb{P}\left[|\tilde{F}^{\mathrm{Hopf}}_{N,\beta,\rho}(m) - \mathbb{E}\tilde{F}^{\mathrm{Hopf}}_{N,\beta,\rho}(m)| > z\right] \leqq \begin{cases} e^{-NC(m)z^2}, & \text{if } z \leqq 1 \\ e^{-NC(m)z}, & \text{if } z > 1. \end{cases} \tag{3.19}$$

Moreover, the constant $C(m)$ depends only on $\|t^*(\tilde{m})\|_2$ and $\|m\|_2$.

Proof. The proof of this lemma is based on the classical method of Yurinskii [28]. An exposition can be found in particular in the beautiful book by Ledoux and Talagrand [24]. Earlier applications of Yurinskii's method in the context of spin glasses and the Hopfield model can be found in [25, 26, 27]. Let us set for later use

$$Z_N(m) \equiv Z_{N,\beta,\rho,\delta}(m) \equiv \frac{1}{2^N} \sum_{\sigma \in \mathscr{S}_N} e^{-\beta H_N(\sigma)} \chi_{\rho,\delta}(\|m_N(\sigma) - m\|_2) \tag{3.20}$$

and write for simplicity

$$f_N(m) \equiv -\beta^{-1} \ln Z_N(m) \tag{3.21}$$

We now introduce the decreasing sequence of sigma-algebras $\mathscr{F}_{k,\kappa}$ that are generated by the random variables $\{\xi_i^\mu\}_{i \geqq k+1}^{1 \leqq \mu \leqq p} \cup \{\xi_k^\mu\}^{\mu \geqq \kappa}$, and the corresponding martingale difference sequence

$$\tilde{f}_N^{(k,\kappa)}(m) \equiv \mathbb{E}\left[f_N(m) | \mathscr{F}_{k,\kappa}\right] - \mathbb{E}\left[f_N(m) | \mathscr{F}^+_{k,\kappa}\right] \tag{3.22}$$

where for notational convenience we have set

$$\mathscr{F}^+_{k,\kappa} = \begin{cases} \mathscr{F}_{k,\kappa+1}, & \text{if } \kappa < p \\ \mathscr{F}_{k+1,1} & \text{if } \kappa = p. \end{cases} \tag{3.23}$$

Notice that we have the identity

$$f_N(m) - \mathbb{E}f_N(m) \equiv \sum_{k=1}^{N} \sum_{\kappa=1}^{p} \tilde{f}_N^{(k,k)}(m). \tag{3.24}$$

Note that we use a finer filtration for the construction of the martingale than [26, 27] which allows us to get much sharper estimates.
Our aim is to use an exponential Markov inequality for martingales. This requires in particular bounds on the conditional Laplace transforms of the

martingale differences. Namely, we clearly have that

$$\mathbb{P}\left[\left|\sum_{k=1}^{N}\sum_{\kappa=1}^{p}\tilde{f}_N^{(k,\kappa)}(m)\right| \geqq Nz\right] \leqq 2\inf_{t\in\mathbb{R}} e^{-|t|Nz}\mathbb{E}\exp\left\{t\sum_{k=1}^{N}\sum_{\kappa=1}^{p}\tilde{f}_N^{(k,\kappa)}(m)\right\} \tag{3.25}$$

$$= 2\inf_{t\in\mathbb{R}} e^{-|t|Nz}\mathbb{E}[\mathbb{E}[\ldots\mathbb{E}[e^{t\tilde{f}_N^{(1,1)}}(m)|\mathcal{F}_{1,1}^{+}]\, e^{t\tilde{f}_N^{(1,2)}}(m)|\mathcal{F}_{1,2}^{+}]\ldots e^{t\tilde{f}_N^{(N,p)(m)}}|\mathcal{F}_{N,p}^{+}]$$

To bound the conditional Laplace transforms, we introduce for $u\in[-1,1]$ the p-dimensional vectors $m_N^{k,\kappa}(\sigma,u)$ with components

$$m_N^{(k,\kappa),\mu}(\sigma,u) \equiv \begin{cases} \frac{1}{N}(\sum_{i\neq h}^{i}\xi_i^\kappa\sigma_i + u\xi_k^\kappa\sigma_k), & \text{if } \mu = k \\ m_N^\mu(\sigma), & \text{if } \mu \neq \kappa \end{cases} \tag{3.26}$$

and define

$$\tilde{H}_N^{(k,k)}(\sigma,u) = \frac{N}{2}\|m_N^{(k,\kappa)}(\sigma,u)\|_2^2 \tag{3.27}$$

Naturally, we set

$$Z_N^{(k,\kappa)}(m,u) \equiv \frac{1}{2^N}\sum_{\sigma\in\mathcal{S}_N} e^{-\beta\tilde{H}_N^{(k,\kappa)}(\sigma,u)}\chi_\rho(\|m_N^{(k,\kappa)}(\sigma,u)-m\|_2) \tag{3.28}$$

and finally

$$\tilde{f}_N^{(k,\kappa)}(m,u) = -\beta^{-1}\ln Z_N^{(k,\kappa)}(m,u). \tag{3.29}$$

Since for the remainder of the proof, m as well as N will be fixed values, to simplify our notations we will write simply $f_{k,\kappa}(u) \equiv f_N^{(k,k)}(m,u)$.
The point behind this definition is that

$$\begin{aligned}\tilde{f}_N^{(k,\kappa)}(m) &= \mathbb{E}[f_{k,\kappa}(1)|f_{(k,\kappa)}] - \mathbb{E}[f_{k,\kappa}(1)|f_{k,\kappa}^{+}] \\ &= \frac{1}{2}\mathbb{E}[f_{k,\kappa}(1) - f_{k,\kappa}(-1)|\mathcal{F}_{k,\kappa}].\end{aligned} \tag{3.30}$$

To obtain the last line we have used the fact that $f_{k,\kappa}(u)$ is actually a function of the product $u\xi_k^\kappa$.
On the other hand, $f_{k,\kappa}(u)$ is a concave function of u (hence the condition (4) for χ_ρ). This implies that

$$\begin{aligned}&|f_{k,\kappa}(1) - f_{(k,\kappa)}(-1)| \leqq 2\max(|f'_{k,\kappa}(1)|, |f'_{(k,\kappa)}(-1)|) \\ &\leqq 2|f'_{k,\kappa}(1)| + 2|f'_{(k,\kappa)}(-1)|.\end{aligned} \tag{3.31}$$

Now,

$$\begin{aligned}f'_{k,\kappa}(u) = \mathcal{E}_{k,\kappa,u}\Bigg(& m_N^{(k,\kappa),\kappa}(\sigma,u)\xi_k^\kappa\sigma_k + \frac{1}{\beta N}\frac{\chi'_\rho(\|m_N^{(k,\kappa)}(\sigma,u)-\tilde{m}\|_2)}{\chi_\rho(\|m_N^{(k,\kappa)}(\sigma,u)-m\|_2)} \\ & \times\frac{(m_N^{(k,\kappa),\kappa}(\sigma,u)-\tilde{m}^\kappa)\xi_k^\kappa\sigma_k}{\|m_N^{(k,\kappa)}(\sigma,u)-m\|_2}\Bigg)\end{aligned} \tag{3.32}$$

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

$$\frac{1}{Z_N^{(k,\kappa)}(m,u)}\chi_\rho(\|m_N^{(k,\kappa)}(\sigma,u)-\tilde{m}\|_2)\,e^{-\beta\tilde{H}_N^{(k,\kappa)}}(\sigma,u)\,d\sigma \tag{3.33}$$

Using the standard inequalities $e^x \leqq 1+x+\frac{x^2}{2}e^{|x|}$ and $1+y\leqq e^y$ we get

$$\mathbb{E}\,[e^{t\tilde{f}_N^{(k,\kappa)}(m)}|\mathscr{F}_{k,\kappa}^+] \leqq \exp\left(\frac{t^2}{2}\mathbb{E}\left[\left(\tilde{f}_N^{(k,\kappa)}(m)\right)^2 e^{2|t||\tilde{f}_N^{(k,\kappa)}(\tilde{m})|}\middle|\mathscr{F}_{k,\kappa}^+\right]\right) \tag{3.34}$$

To bound the $\tilde{f}_N^{(k,\kappa)}(m)$ in the exponent, we will simply use a uniform upper bound on $|f'_{k,\kappa}(\pm 1)|$. As the functions for ± 1 have the same distribution, it is enough to consider the case $+1$. Note that

$$\begin{aligned}|f'_{(k,\kappa)}(1)| &\leqq \mathscr{E}|m_M^{(k,\kappa)}(\sigma)| + \frac{1}{\beta N}\mathscr{E}\left|\frac{\chi'_\rho(\|m_N(\sigma)-\tilde{m}\|_2)}{\chi_\rho(\|m_N(\sigma)-m\|_2)}\right| \\ &\leqq 1+\frac{1}{\beta N}\mathscr{E}\frac{|\chi'_\rho(\|m_N(\sigma)-\tilde{m}\|_2)|}{\chi_\rho(\|m_N(\sigma)-m\|_2)}\end{aligned} \tag{3.35}$$

where we wrote simply $\mathscr{E}$ for $\mathscr{E}_{k,\kappa,1}$, as this measure is independent of (k,κ). To use this bound, we need the following lemma.

Lemma 3.4 *Let $c(m)\equiv 2(\|m\|_2+\|t^*(\tilde{m})\|_2)$ be finite. Then there exists ρ satisfying $\sqrt{\alpha}\leqq\rho\leqq 4\sqrt{\alpha}$ (where $\alpha\equiv\frac{p}{N}$) such that*

$$\frac{1}{\beta N}\mathscr{E}\frac{|\chi'_\rho(\|m_N(\sigma)-\tilde{m}\|_2)|}{\chi_\rho(\|m_N(\sigma)-m\|_2)}\leqq 2e^{3c(m)}-2 \tag{3.36}$$

(*Note that ρ may in principle depend on m*).

Proof. Note that the conditions on χ_ρ imply that

$$\begin{aligned}\frac{1}{\beta N}\mathscr{E}\frac{|\chi'_\rho(\|m_N(\sigma)-\tilde{m}\|_2)|}{\chi_\rho(\|m_N(\sigma)-m\|_2)} &\leqq 2\frac{Z_{N,\beta,\rho+1/(\beta N)}(m)-Z_{N,\beta,\rho}(m)}{Z_{N,\beta,\rho}(m)} \\ &= 2\left(\frac{Z_{N,\beta,\rho+1/(\beta N)}(m)}{f_{N,\beta,\rho}(m)}-1\right).\end{aligned} \tag{3.37}$$

Let us set $g(x)\equiv\frac{1}{\beta N}\ln Z_{N,\beta,x}(m)$. Obviously, $g(x)$ is a monotone increasing function of x. From the large deviation estimates in Lemma 3.1, it follows that, for $x\geqq 2\sqrt{\alpha}$,

$$g(n\sqrt{\alpha})-g(2\sqrt{\alpha})\leqq c(m)(n+2)\sqrt{\alpha}. \tag{3.38}$$

Therefore

$$c(m)(n+2)\sqrt{\alpha}\geqq\sum_{k=0}^{\sqrt{\alpha}(n-2)N\beta}\left(g(\sqrt{\alpha}+(k+1)/(\beta N))-g(\sqrt{\alpha}+k/(\beta N))\right) \tag{3.39}$$

But this implies that the number of values k in the sum for which

$$(g(\sqrt{\alpha}+(k+1)/(\beta N))-g(\sqrt{\alpha}+k/(\beta N)))>\frac{c'}{N\beta} \tag{3.40}$$

is bounded by $\frac{c(m)(n+2)\sqrt{\alpha}N\beta}{c'}$; thus if we choose $c'=\frac{n+2}{n-2}c(m)$, we can deduce that there exists a smallest value $k_0 \leqq \sqrt{\alpha}\beta N(n-2)$ for which

$$0 \leqq (g(\sqrt{\alpha}+(k_0+1)/(\beta N))-g(\sqrt{\alpha}+k_0/(\beta N))) \leqq c'/(\beta N) \tag{3.41}$$

Thus choosing $\rho = N\beta k_0 + 2\sqrt{\alpha}$, we get for this value the bound

$$\frac{1}{\beta N}\mathscr{E}\frac{|\chi_\rho'(\|m_N(\sigma)-\tilde{m}\|_2)|}{\chi_\rho(\|m_N(\sigma)-m\|_2)} \leqq 2(e^{c'}-2) \tag{3.42}$$

Choosing $n = 4$ concludes the proof of the lemma. □

Putting these observations together, we see already that $|\tilde{f}_N^{(k,\kappa)}(m)| \leqq C$ (where C depends on $\tilde{m}$, if ρ is chosen appropriately (but of order $\sqrt{\alpha}$). Hence we can write

$$\mathbb{E}\left[e^{t\tilde{f}_N^{(k,\kappa)}(m)}\middle|\mathscr{F}_{k,\kappa}^+\right] \leqq \exp\left(\frac{t^2}{2}e^{2C|t|}\mathbb{E}\left[\left(\tilde{f}_N^{(k,\kappa)}(\tilde{m})\right)^2\middle|\mathscr{F}_{k,\kappa}^+\right]\right) \tag{3.43}$$

Using the bounds (3.30) and (3.31) we get further that

$$\begin{aligned}
&\mathbb{E}\left[\left(\tilde{f}_N^{(k,\kappa)}(\tilde{m})\right)^2\middle|\mathscr{F}_{k,\kappa}^+\right] \\
&\quad \leqq 2\mathbb{E}\left[(\mathbb{E}[|f'_{k,\kappa}(1)||\mathscr{F}_{k,\kappa}])^2+(\mathbb{E}[|f'_{k,\kappa}(-1)||\mathscr{F}_{k,\kappa}])^2|\mathscr{F}_{k,\kappa}^+\right] \\
&\quad \leqq 2\mathbb{E}\left[\mathbb{E}[|f'_{k,\kappa}(1)|^2|\mathscr{F}_{k,\kappa}]+\mathbb{E}[|f'_{k,\kappa}(-1)|^2|\mathscr{F}_{k,\kappa}]|\mathscr{F}_{k,\kappa}^+\right] \\
&\quad \leqq 2\mathbb{E}\left[|f'_{k,\kappa}(1)|^2+|f'_{k,\kappa}(-1)|^2|\mathscr{F}_{k,\kappa}^+\right] \\
&\quad = 4\mathbb{E}\left[|f'_{k,\kappa}(1)|^2|\mathscr{F}_{k,\kappa}^+\right]
\end{aligned} \tag{3.44}$$

Inserting the resulting bound into (3.25), using Jensen's inequality $e^{\mathbb{E}[X|\mathscr{F}]} \leqq \mathbb{E}[e^X|\mathscr{F}]$, we get

$$\begin{aligned}
&\mathbb{P}[f_N(|m)-\mathbb{E}f_N(m)| \geqq Nx] \\
&\quad \leqq 2\inf_{t\geqq 0} e^{-tNx}\mathbb{E}\left[\exp\left(4t^2e^{2tC}\sum_{k=1}^{N}\sum_{\kappa=1}^{p}(f'_{k,\kappa}(1))^2\right)\right]
\end{aligned} \tag{3.45}$$

Now,

$$\begin{aligned}
(f'_{k,\kappa}(1))^2 &\leqq 2\left(\mathscr{E}\left[\frac{1}{N}\sum_i \xi_i^\kappa\xi_k^\kappa\sigma_i\sigma_k\right]\right)^2 \\
&\quad + 2\left(\mathscr{E}\left[\frac{(m_N^\kappa(\sigma)-\tilde{m}^\kappa)}{\|m_N(\sigma)-m\|_2}\frac{1}{\beta N}\frac{|\chi_\rho'(\|m_N(\sigma)-\tilde{m}\|_2)|}{\chi_\rho(\|m_N(\sigma)-m\|_2)}\right]\right)^2
\end{aligned}$$

$$\leqq 2\mathscr{E}\left(\frac{1}{N}\sum_i \xi_i^\kappa \sigma_i\right)^2 + 2\mathscr{E}\left[\frac{m_N^\kappa(\sigma) - \tilde{m}^\kappa)^2}{\| m_N(\sigma) - m \|_2}\right]$$

$$\times \frac{1}{(\beta N)^2}\mathscr{E}\left[\left(\frac{|\chi_\rho'(\| m_N(\sigma) - \tilde{m} \|_2)|}{\chi_\rho(\| m_N(\sigma) - m \|_2)}\right)^2\right]$$

$$\leqq 2\mathscr{E}\left(\frac{1}{N}\sum_i \xi_i^\kappa \sigma_i\right)^2 + 2(e^{3c(m)} - 1)\mathscr{E}\left[\frac{(m_N^\kappa(\sigma) - \tilde{m}^\kappa)^2}{\| m_N(\sigma) - m \|_2}\right] \tag{3.46}$$

where we have used Lemma 3.4. Thus

$$\mathscr{E}\left[\sum_{k=1}^{N}\sum_{\kappa=1}^{p}(f_{k,\kappa}'(1))^2\right] \leqq 2N\mathscr{E}\sum_{\kappa=1}^{p}\left(\frac{1}{N}\sum_i \xi_i^\kappa \sigma_i\right)^2 + 2N(e^{3c(m)} - 1)$$

$$= 4\mathscr{E}H_N(\sigma) + 2N(e^{3c(m)} - 1) \tag{3.47}$$

Note that this formula is quite remarkable, in that it relates the variance of the local free energy to the distribution of the (local) internal energy of the system. For our present purposes it will, however, suffice to bound $\mathscr{E}H_N(\sigma)$ by the supremum of the Hamiltonian over all spin configurations.

Lemma 3.5. *There exist a constant $c \geqq 1 - 2^{-1/(2\alpha)}2$ such that for all x,*

$$\mathbb{P}\left[\sup_{\sigma \in \mathscr{S}_N} H_N(\sigma) \geqq xN\right] \leqq e^{-c(x - 2\ln 2)N}. \tag{3.48}$$

Proof. Just use that

$$\mathbb{P}\left[\sup_{\sigma \in \mathscr{S}_N} 2H_N(\sigma) \geqq xN\right] \leqq 2^N \mathbb{P}\left[\frac{1}{2}\sum_{\kappa=1}^{p}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_i^\kappa\right)^2 \geqq xN\right] \tag{3.49}$$

and the fact that $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\xi_i^\kappa$ is subgaussian. □

From this estimate it follows easily that

$$\mathbb{E}e^{s\mathscr{E}[\sum_{\kappa=1}^{N}\sum_{\kappa=1}^{y}(f_{(k,\kappa)}'(1))^2]} \leqq e^{cNs} \tag{3.50}$$

for some constant $c' < \infty$ if s is smaller than some constant. Lemma 3.3 now follows by choosing t appropriately. □

Remark. One has to keep in mind that the constants in the estimate in Lemma 3.3 depend on m in a way that is difficult to control explicitly. The self-averaging result concerns thus the rate functions only at points where this constant is bounded, i.e. in the interior of the set where the rate function itself is finite. It should be noted that this includes in particular the critical points, for which $t^*(m) = m$.

Let us introduce the sets

$$\mathscr{W}_{N,p}(c) \equiv \{m \in \mathbb{R}^p | \forall_\mu m^\mu \in$$

$$\{-1, -1 + 2/\sqrt{N}, \ldots, 1 - 2/\sqrt{N}, 1\} \cap c(m) \leqq c\} \tag{3.51}$$

Proposition 3.6 *Assume that* $\lim_{N\uparrow\infty} \frac{p(N)}{N} = 0$. *Then, for almost all* ω, *there exist a sequences* $\rho = \rho_N(m)$ *satisfying* $\sqrt{\frac{p}{N}} \leqq \rho \leqq 4\sqrt{\frac{p}{N}}$ *such that*

$$\lim_{N\uparrow\infty} \sup_{m\in\mathcal{W}_{N,p(N)}(c)} \left| F^{\text{Hopf}}_{N,\beta,\rho(m)}(m) - \mathbb{E}F^{\text{Hopf}}_{N,\beta,\rho(m)}(m) \right| = 0 \tag{3.52}$$

for any $c < \infty$.

Proof. By Lemma 3.3, there exist $\rho(m)$ such that

$$\begin{aligned}
&\mathbb{P}\left[\sup_{m\in\mathcal{W}_{N,p(N)}(c)} \left| \tilde{F}^{\text{Hopf}}_{N,\beta,\rho(m)}(m) - \mathbb{E}\tilde{F}^{\text{Hopf}}_{N,\beta,\rho(m)}(m) \right| > z \right] \\
&\leqq \sum_{m\in\mathcal{W}_{N,p(N)}(c)} \mathbb{P}\left[\left| \tilde{F}^{\text{Hopf}}_{N,\beta,\rho(m)}(m) - \mathbb{E}\tilde{F}^{\text{Hopf}}_{N,\beta,\rho(m)}(m) \right| > z \right] \\
&\leqq e^{c'' N\frac{p}{N}\ln\frac{p}{N} - Nc'z^2}
\end{aligned} \tag{3.53}$$

Where we have used the well-known fact that the number of lattice with spacing $1/\sqrt{N}$ in the p-dimensional unit ball is bounded by $\exp(c'' p \ln pN)$ for some numerical constant c''. Since by our assumptions on $p(N)$ for any $z > 0$, this bound is exponentially small in N for N sufficiently large, the proposition follows from the first Borel-Cantelli Lemma. □

3.3 Two variational formulae for the rate function

We now turn to the second result on the Hopfield model, which is a large deviation principle under a very strong condition on $p(N)$, namely that $p(N) < \ln N/\ln 2$. This makes use of a very nice technique introduced first by Grensing and Kühn [16] and later used by Koch and Piasko [20] and Gayrard [14] to compute the free energy and to construct the Gibbs states of the model in this regime.

Let us denote by $I \subset \mathbb{N}$ a finite set; we will always assume that N is so large that $I \subset \{1, \dots, p(N)\}$ (we exclude the trivial case $p(N)$ bounded). We denote by Π_I the orthogonal projection from $\mathbb{R}^{p(N)}$ to $\mathbb{R}^I$. Let us introduce, for $\tilde{m} \in [-1, 1]^I$, the quantities

$$Z^I_{N,\beta,\varepsilon}[\omega](\tilde{m}) \equiv 2^{-N} \sum_{\sigma\in\mathcal{S}_N} e^{-\beta H_N(\sigma)} \mathbb{1}_{\{\|\Pi_I m_N(\sigma) - \tilde{m}\|_2^2 \leqq \varepsilon\}}. \tag{3.54}$$

We introduce the family of vectors $e_\gamma \in \{-1, 1\}^p$, for $\gamma = 1, 2, \dots, 2^p$ which represent a complete enumeration of all vectors in $\mathbb{R}^p$ whose components take only the values ± 1. We set

$$v_\gamma \equiv \{i \in \Lambda \mid \xi_i^\mu = e_\gamma^\mu, \quad \forall \mu = 1, \dots, p\}. \tag{3.55}$$

The v_γ are of course random quantities depending on the ξ_i^μ, however, their volumes $|v_\gamma|$ almost deterministic in the sense that there exist a subset $\tilde{\Omega} \subset \Omega$

of probability one, and functions δ_N tending to zero as $N \uparrow \infty$, such that for all but a finite number of indices N,

$$||v_\gamma| - 2^{-p}N| \leqq \delta_N 2^{-p(N)}N, \quad \forall \gamma \in \{1, \dots, p(N)\} \tag{3.56}$$

provided that $p(N)$ satisfies the assumption of Lemma 3.4. A proof of this fact can be found in [14] (Proposition 4.1). Let us remark further that the vectors e_γ have the property that

$$2^{-p} \sum_{\gamma=1}^{2^p} e_\gamma^\mu e_\gamma^\nu = \delta_{\mu,\nu}, \tag{3.57}$$

where δ here is the Kronecker symbol. Let us denote $\alpha_\gamma(\sigma) \equiv (1/|v_\gamma|)\sum_{i \in v_\gamma} \sigma_i$. We then have

$$m_N^\mu(\sigma) = 2^{-p} \sum_{\gamma=1}^{2^p} e_\gamma^\mu \alpha_\gamma(\sigma) \equiv m_p(\alpha(\sigma))). \tag{3.58}$$

up to a negligible error of order δ_N.

Lemma 3.7 *Assume that* $\lim_{N\uparrow\infty} 2^{p(N)} \ln N/N = 0$. *Then there exists a set* $\tilde{\Omega} \subset \Omega$ *with* $\mathbb{P}(\tilde{\Omega}) = 1$, *such that, for all* $\omega \in \tilde{\Omega}$, *for all but a finite number of indices* N,

$$\frac{1}{\beta N} \ln Z_{N,\beta,\varepsilon}^I[\omega](\tilde{m}) = \sup_{\substack{\alpha \in [-1,1]^{2^{p(N)}} \\ \|\Pi_I m_\gamma(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon}} \frac{1}{2} \left\| 2^{-p} \sum_{\gamma=1}^{2^p} e_\gamma \alpha_\gamma \right\|_2^2 - \frac{1}{\beta 2^p} \sum_{\gamma=1}^{2^p} I(\alpha_\gamma) + o(1). \tag{3.59}$$

Proof. We have

$$Z_{N,\beta,\varepsilon}^I(\tilde{m}) = 2^{-N} \sum_{\alpha_\gamma \in \mathscr{W}_{|v_\gamma|}, \gamma = 1, \dots, 2^p} \mathbb{1}_{\{\|\Pi_I m_p(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon\}} e^{\beta N (1/2) \|m_p(\alpha)\|_2^2} \sum_{\sigma \in \mathscr{S}_N} \mathbb{1}_{\{\alpha_\gamma(\sigma) = \alpha_\gamma, \gamma = 1, \dots, 2^p\}}, \tag{3.60}$$

where $\mathscr{W}_{|v_\gamma|} \equiv \{-1, -1 + 2/|v_\gamma|, \dots, 1 - 2/|v_\gamma|, 1\}$. The last sum over the σ is easy to compute. Namely

$$2^{-N} \sum_{\sigma \in \mathscr{S}_N} \mathbb{1}_{\{\alpha_\gamma(\sigma) = \alpha_\gamma, \gamma = 1, \dots, 2^p\}} = 2^{-N} \prod_{\gamma=1}^{2^p} \binom{|v_\gamma|}{|v_\gamma| \frac{1+\alpha_\gamma}{2}} = \exp\left(- \sum_{\gamma=1}^{2^p} (|v_\gamma| I(\alpha_\gamma) + O(\ln |v_\gamma|)) \right), \tag{3.61}$$

with I the well-known entropy function

$$I(x) = \begin{cases} \frac{1+x}{2} \ln(1+x) + \frac{1-x}{2} \ln(1-x) & \text{if } |x| \leqq 1, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.62}$$

The term $\sum_{\gamma=1}^{2^p} O(\ln|v_\gamma|)$ is of order $2^p \ln N$ and therefore totally negligible. Using (3.42) we obtain

$$Z_{N,\beta,\varepsilon}^I(\tilde{m}) = \sum_{\alpha_\gamma,\, \gamma=1,\,\dots,\,2^p} \mathbb{1}_{\{\|\Pi_I m_p(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon\}} \exp\Bigg(\beta N \tfrac{1}{2} \|m_p(\alpha)\|_2^2 - \sum_{\gamma=1}^{2^p} (|v_\gamma| I(\alpha_\gamma) + O(\ln|v_\gamma|)) + NO(\delta_N)\Bigg). \tag{3.63}$$

Observe that the α_γ take values in the set $\mathscr{W}_{|v_\gamma|}$ so that the total number of terms in the sum $\sum_{\alpha_\gamma,\gamma=1,\dots,2^p}$ is bounded by $\prod_{\gamma=1}^{2^p} |v_\gamma| < e^{2^p \ln N}$. Therefore, it suffices to use the upper and lower bounds

$$\sup_{\substack{\alpha_\gamma \in \mathscr{W}_{|v_\gamma|},\, \gamma=1,\,\dots,\,2^p \\ \|\Pi_I m_p(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon}} \frac{1}{2}\|m_p(\alpha)\|_2^2 - \frac{1}{\beta N}\sum_{\gamma=1}^{2^p} (|v_\gamma| I(\alpha_\gamma) - O(\ln|v_\gamma|)) - NO(\delta_N)$$

$$\leqq \frac{1}{\beta N} \ln Z_{N,\beta,\varepsilon}^I(\tilde{m}) \leqq \sup_{\substack{\alpha_\gamma \in \mathscr{W}_{|v_\gamma|},\, \gamma=1,\,\dots,\,2^p \\ \|\Pi_I m_p(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon}} \frac{1}{2}\|m_p(\alpha)\|_2^2$$

$$- \frac{1}{\beta N}\sum_{\gamma=1}^{2^p} (|v_\gamma| I(\alpha_\gamma) + O(\ln|v_\gamma|)) + NO(\delta_N) + \beta^{-1}\frac{2^p \ln N}{N}. \tag{3.64}$$

Since we have assumed that $2^p \ln N/N \downarrow 0$ as $N \uparrow \infty$, and using (3.56) we see that on $\tilde{\Omega}$ both the upper and the lower bound in (3.64) only differ by terms that converge to zero as $N \uparrow \infty$ from the quantity

$$\Phi_{N,\beta,\varepsilon}(\tilde{m}) = \sup_{\substack{\alpha_\gamma \in [-1,1],\, \gamma=1,\,\dots,\,2^p \\ \|\Pi_I m_p(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon}} \frac{1}{2}\|m_p(\alpha)\|_2^2 - \frac{1}{\beta 2^p}\sum_{\gamma=1}^{2^p} I(\alpha_\gamma). \tag{3.65}$$

But this proves the lemma. □

If $p(N)$ were bounded, we would now be done. For in such a case, the limit as N tends to infinity of $\Phi_{N,\beta,\varepsilon}(\tilde{m})$ clearly exists and yields the desired large deviation rate function in terms of a variational formula. In our situation, since the dimension of the space over which the α vary diverges, it is not *a priori* evident that the limit exists and can be expressed through a variational principle. To prove it we need some notation. Let us first observe that the vectors e^μ can be chosen in the following explicit form,

$$e_\gamma^\mu = (-1)^{[\gamma 2^{-(\mu-1)}]}, \tag{3.66}$$

where $[x]$ denotes the integer part of x. Let us define the sets

$$\mathscr{A}_d^p \equiv \{\alpha \in [-1,1]^{2^p} \mid \forall \gamma,\, \alpha_\gamma = \alpha_{\gamma+2^d}\}. \tag{3.67}$$

Obviously,

$$\mathscr{A}_0^p \subset \mathscr{A}_1^p \subset \cdots \subset \mathscr{A}_{p-1}^p \subset \mathscr{A}_p^p = [-1,1]^{2^p}. \tag{3.68}$$

The points to notice are now the following: *If* $\alpha \in \mathscr{A}_d^p$, with $d < p$, then

(i) $m_p^\nu(\alpha) = 0$, if $\nu > d$ and

(ii) $m_p^\mu(\alpha) = m_d^\mu(\alpha)$, if $\mu \leqq d$.

Let us set

$$\Theta_p(\alpha) = \frac{1}{2} \| m_p(\alpha) \|_2^2 - \frac{1}{\beta 2^p} \sum_{\gamma=1}^{2^p} I(\alpha_\gamma) \tag{3.69}$$

and

$$\Upsilon_{p,\varepsilon}(\tilde{m}) = \sup_{\substack{\alpha \in \mathscr{A}_p^p \\ \|\Pi_I m_p(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon}} \Theta_p(\alpha). \tag{3.70}$$

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

$$\Upsilon_{p,\varepsilon}(\tilde{m}) \geqq \sup_{\substack{\alpha \in \mathscr{A}_d^p \\ \|\Pi_I m_d(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon}} \Theta_p(\alpha) = \sup_{\substack{\alpha \in \mathscr{A}_d^d \\ \|\Pi_I m_d(\alpha) - \tilde{m}\|_2^2 \leqq \varepsilon}} \Theta_d(\alpha) = \Upsilon_{d,\varepsilon}(\tilde{m}). \tag{3.71}$$

Hence $\Upsilon_{p,\varepsilon}(\tilde{m})$ is an increasing sequence in p and being bounded from above, converges. With these preparations, we are ready to prove the following proposition.

Proposition 3.8 *Assume that* $\lim_{N \uparrow \infty} 2^{p(N)} \ln N/N = 0$. *Then for almost all* ω, *the induced measures* $\mathscr{Q}_{N,\beta}^{\mathrm{Hopf}}[\omega]$ *satisfy a large deviation principle with rate function* $F_\beta^{\mathrm{Hopf},I}(\tilde{m})$ *given by the following variational formula*:

$$F_\beta^{\mathrm{Hopf},I}(\tilde{m}) = -\sup_{p \in \mathbb{N}} \sup_{\substack{\alpha \in [-1,+1]^{2^p} \\ \Pi_I 2^{-p} \sum_{\gamma=1}^{2^p} e_\gamma \alpha_\gamma = \tilde{m}}} \frac{1}{2} \left\| 2^{-p} \sum_{\gamma=1}^{2^p} e_\gamma \alpha_\gamma \right\|_2^2 - \frac{1}{\beta 2^p} \sum_{\gamma=1}^{2^p} I(\alpha_\gamma)$$
$$+ \sup_{x \in [-1,1]} \left(\frac{x^2}{2} - \beta^{-1} I(x) \right). \tag{3.72}$$

If Γ_p *denotes the set*

$$\Gamma_p \equiv \{\tilde{m} \in \mathbb{R}^I \mid \exists_{\alpha \in (-1,1)^p} \colon \Pi_I m_p(\alpha) = \tilde{m}\},$$

then $F_\beta^{\mathrm{Hopf},I}(\tilde{m})$ *is uniformly bounded on* $\bar{\Gamma}_{|I|}$ *and* $F_\beta^{\mathrm{Hopf},I}(\tilde{m}) = +\infty$ *if* $m \notin \bar{\Gamma}_{|I|}$. $F_\beta^{\mathrm{Hopf},I}(\tilde{m})$ *is lower-semicontinuous, and Lipshitz continuous on* $\Gamma_{|I|}$.

Proof. We have shown above that $\Upsilon_{p,\varepsilon}(\tilde{m})$ converges for fixed ε, from which we obtain immediately that $F_{\beta,\varepsilon}^{\mathrm{Hopf}}(\tilde{m})$ exists and is given by the variational formula

$$F_{\beta,\varepsilon}^{\mathrm{Hopf},I}(\tilde{m}) = -\sup_{p \in \mathbb{N}} \sup_{\substack{\alpha \in [-1,+1]^{2^p} \\ \|\Pi_I 2^{-p} \sum_{\gamma=1}^{2^p} e_\gamma \alpha_\gamma - \tilde{m}\|_2^2 \leqq \varepsilon}} \frac{1}{2} \| m_p(\alpha) \|_2^2 - \frac{1}{\beta 2^p} \sum_{\gamma=1}^{2^p} I(\alpha_\gamma)$$
$$+ \sup_{x \in [-1,1]} \left(\frac{x^2}{2} - \beta^{-1} I(x) \right). \tag{3.73}$$

From this it is obvious that $F_{\beta,\varepsilon}^{\mathrm{Hopf},I}$ converges, as $\varepsilon\downarrow 0$ to a lower-semicontinuous function and that

$$\lim_{\varepsilon\downarrow 0} F_{\beta,\varepsilon}^{\mathrm{Hopf},I}(\tilde{m}) = \lim_{\varepsilon\downarrow 0} \inf_{m:\|m-\tilde{m}\|_2^2 \leqq \varepsilon} F_{\beta}^{\mathrm{Hopf},I}(m), \tag{3.74}$$

with $F_{\beta}^{\mathrm{Hopf},I}(m)$ understood to be *defined* by (3.53). This will imply

$$\lim_{\varepsilon\downarrow 0} F_{\beta,\varepsilon}^{\mathrm{Hopf},I}(\tilde{m}) = F_{\beta}^{\mathrm{Hopf},I}(\tilde{m}) \tag{3.75}$$

whenever $F_{\beta}^{\mathrm{Hopf},I}$ is continuous in a neighborhood of $\tilde{m}$.

Recall that $I(x)$ is uniformly bounded on $[-1,1]$, and continuous with bounded derivative on $(-1,1)$. Therefore, $\Theta_p(\alpha)$ enjoys the same properties on $[-1,1]^p$ and $(-1,1)^p$, respectively. Moreover, a straightforward computation shows that on compact subsets of its domain of continuity, Θ_p is in fact uniformly Lipshitz with constant $C2^{-p/2}$, i.e. for $\alpha,\alpha' \in (-1,1)^p$,

$$|\Theta_p(\alpha) - \Theta_p(\alpha')| \leqq C2^{-p/2}\|\alpha - \alpha'\|_2. \tag{3.76}$$

It is clear that if there exists p and $\alpha \in [-1,1]^p$ such that $\Pi_I m_p(\alpha) = \tilde{m}$, then $F_{\beta}^{\mathrm{Hopf},I}(\tilde{m}) < +\infty$. This shows that $F_{\beta}^{\mathrm{Hopf},I}$ is bounded on $\Gamma_{|I|}$. But it is not difficult to see that $\Gamma_{|I|} \supset \Gamma_p$ for all $p \geqq |I|$, so that $\Gamma_{|I|}$, so that $\Gamma_{|I|}$ is the domain of finiteness of $F_{\beta}^{\mathrm{Hopf},I}$.

Notice that if $\alpha \in [-1,1]^p$ is such that $\Pi_I m_p(\alpha) = m$ then α' defined through $\alpha'_\gamma = \alpha_\gamma + \sum_{\mu\in I} e_\gamma^\mu(\tilde{m}^\mu - m^\mu)$ satisfies $\Pi_I m_p(\alpha) = \tilde{m}$. Clearly $\|\alpha - \alpha'\|_2 = 2^{p/2}\|\tilde{m} - m\|_2$. Using this fact and (3.57), we find that $\Upsilon_{p,0}$ is actually uniformly Lipshitz continuous on compact subsets of $\Gamma_{|I|}$ with constant C, independent of p. But this implies by a simple three epsilon argument the Lipshitz continuity of $F_{\beta}^{\mathrm{Hopf},I}$ on the interior of its domain of finiteness. This concludes the proof of Proposition 3.5. □

Remark. From the rate functions for the marginal distributions one can of course, by standard arguments construct the rate functions for the infinite dimensional distributions through an inductive limit, as in the Dawson–Gärtner theorem [9] (see e.g. [8]).

It may be of interest to give an alternative expression for the variational formula (3.53) which allows to obtain some interesting bounds. To this end, we notice that the function $I(x)$ is the Legendre transform of the function $\ln\cosh(t)$, i.e. that

$$I(x) = \sup_{t\in\mathbb{R}} (tx - \ln\cosh t). \tag{3.77}$$

Let us first rewrite

$$\Phi_{N,\beta,\varepsilon}(\tilde{m}) = \sup_{\substack{m\in[-1,1]^p \\ \|\Pi_I m - \tilde{m}\|_2^2 \leqq \varepsilon}} \sup_{\substack{\alpha\in[-1,1]^{2^p} \\ 2^{-p}\sum_{\gamma=1}^{2^p} e_\gamma\alpha_\gamma = m}} \frac{1}{2}\|m\|_2^2 - \frac{1}{\beta 2^p}\sum_{\gamma=1}^{2^p} I(\alpha_\gamma). \tag{3.78}$$

To find the suprema under the constraints $2^{-p}\sum_{\gamma=1}^{2^p} e_\gamma\alpha_\gamma = m$, we introduce the corresponding Lagrange multipliers t_μ, $\mu = 1,\ldots,p$. The resulting function

$$L(m, \alpha, t) \equiv \frac{1}{2}\| m \|_2^2 - \frac{1}{\beta 2^p} \sum_{\gamma=1}^{2^p} I(\alpha_\gamma) + \sum_{\mu=1}^{p} t_\mu \left(2^{-p} \sum_{\gamma=1}^{2^p} e_\gamma^\mu \alpha_\gamma - m^\mu \right) \quad (3.79)$$

is quadratic in m^ν, and $(d/dm^\nu)L(m, \alpha, t) = 0$ if and only if $t^\nu = m^\nu$. Thus for the component ν, with $\nu \in I^c$, over which the supremum over m^ν is taken unconditioned, we must have that $t_\nu = m^\nu$. Therefore,

$$\begin{aligned}
&\Phi_{N,\beta,\varepsilon}(\tilde{m}) \\
&= \sup_{\substack{m \in [-1,1]^p \\ \|\Pi_I m - \tilde{m}\|_2^2 \leq \varepsilon}} \inf_{t \in \mathbb{R}^p} \sup_{\alpha \in [-1,1]^{2^p}} \sum_\mu \left(\frac{1}{2}(m^\mu)^2 - t_\mu m^\mu \right) \\
&\quad + \frac{1}{\beta 2^p} \sum_{\gamma=1}^{2^p} \left(\sum_{\mu=1}^{p} \alpha_\gamma e_\gamma^\mu t_\mu - \frac{1}{\beta} I(\alpha_\gamma) \right) \\
&= \sup_{\substack{w \in [-1,1]^I \\ \|w - \tilde{m}\|_2^2 \leq \varepsilon}} \sup_{\tilde{t} \in \mathbb{R}^{I^c}} \inf_{t \in \mathbb{R}^I} \sum_{\mu \in I} \frac{1}{2}(w^\mu - t_\mu)^2 \\
&\quad - \sum_{\nu=1}^{p} \frac{1}{2}\tilde{t}_\nu^2 + \frac{1}{\beta 2^p} \sum_{\gamma=1}^{2^p} \ln\cosh\left(\beta \sum_{\mu=1}^{p} e_\gamma^\mu t_\mu \right) \\
&= \sup_{\substack{w \in [-1,1]^I \\ \|w - \tilde{m}\|_2^2 \leq \varepsilon}} \sup_{\tilde{t} \in \mathbb{R}^{I^c}} \inf_{t \in \mathbb{R}^I} \sum_{\mu \in I} \frac{1}{2}(w^\mu - t_\mu)^2 \\
&\quad + \frac{1}{2^p} \sum_{\gamma=1}^{2^p} \phi_\beta\left(\sum_{\mu \in I} e_\gamma^\mu t_\mu + \sum_{\nu \in I^c} e_\gamma^\nu \tilde{t}_\nu \right)
\end{aligned} \quad (3.80)$$

where $\phi_\beta(z) \equiv -z^2/2 + (1/\beta)\ln\cosh(\beta z)$ and where to get the last line we have used the orthogonality relations (3.57).

From (3.80) we can derive the following alternative variational formula for the rate function $F_\beta^{\mathrm{Hopf},I}(\tilde{m})$:

$$\begin{aligned}
F_\beta^{\mathrm{Hopf},I}(\tilde{m}) &= -\sup_{p \in \mathbb{N}} \sup_{\tilde{t} \in \mathbb{R}^{\{1,\dots,p\}\setminus I}} \inf_{t \in \mathbb{R}^I} \sum_{\mu \in I} \frac{1}{2}(\tilde{m}^\mu - t_\mu)^2 \\
&\quad + \frac{1}{2^p} \sum_{\gamma=1}^{2^p} \phi_\beta\left(\sum_{\mu \in I} e_\gamma^\mu t_\mu + \sum_{\nu \in I^c} e_\gamma^\nu \tilde{t}_\nu \right) \\
&\quad - \inf_{x \in \mathbb{R}} \left(\frac{x^2}{2} - \beta^{-1} \ln\cosh \beta x \right).
\end{aligned} \quad (3.81)$$

In fact, to obtain (3.81), we have to show, like in the proof of Proposition 3.5, that the limit of (3.80) as N tends to infinity exists. To do this, let us define, in complete analogy to the proof of Proposition 3.5,

$$\Xi_p(\tilde{t}, w) \equiv \inf_{t \in \mathbb{R}^I} \sum_{\mu \in I} \frac{1}{2}(w^\mu - t_\mu)^2 + \frac{1}{2^p} \sum_{\gamma=1}^{2^p} \phi_\beta\left(\sum_{\mu \in I} e_\gamma^\mu t_\mu + \sum_{\nu \in I^c} e_\gamma^\nu \tilde{t}_\nu \right). \quad (3.82)$$

Note that $\Phi_{N,\beta,\varepsilon}(w)$ depends on N only through $p(N)$. This suggests to define

$$\Upsilon_p(w) \equiv \sup_{\tilde{t}\in\mathbb{R}^{\{1,\dots,p\}\setminus I}} \Xi_p(\tilde{t}, w). \tag{3.83}$$

To compute this supremum, we can first compute the suprema over subspaces in which only the first d components of $\tilde{t}$ are allowed to take non-zero values, and then take the supremum over $d \geqq p$. But notice that for such $\tilde{t}, \Xi_p(\tilde{t}, w) = \Xi_d(\tilde{t}, w)$, where obviously in the second function $\tilde{t}$ is understood to be the projection of the original $\tilde{t}$ onto the subspace $\mathbb{R}^{\{1,\dots,d\}\setminus I}$. Thus

$$\begin{aligned}\Upsilon_p(w) &= \sup_{d\leqq p} \sup_{\substack{\tilde{t}\in\mathbb{R}^{\{1,\dots,p\}\setminus I}\\ \tilde{t}_\nu=0\,\forall_\nu>d}} \Xi_p(\tilde{t}, w)\\ &= \sup_{d\leqq p} \sup_{\tilde{t}\in\mathbb{R}^{\{1,\dots,d\}\setminus I}} \Xi_d(\tilde{t}, w)\\ &= \sup_{d\leqq p} \Upsilon_d(w).\end{aligned} \tag{3.84}$$

But this implies that $\Upsilon_p(w)$ is a monotone increasing sequence in p. Since it is bounded from above (see. e.g. (3.87), it therefore converges to a finite limit. Thus

$$\lim_{N\uparrow\infty} \Phi_{N,\beta,\varepsilon}(\tilde{m}) = \sup_{\substack{w\in[-1,1]^I\\ \|w-\tilde{m}\|_2^2\leqq\varepsilon}} \lim_{p\uparrow\infty} \Upsilon_p(w). \tag{3.85}$$

From this it is obvious that the expression (3.81) represents the rate function.

From (3.80), it is possible to derive two bounds that involve suprema in a finite number of variables only. First we obtain the obvious lower bound

$$\Phi_{N,\beta,\varepsilon}(\tilde{m}) \geqq \sup_{\substack{w\in[-1,1]^I\\ \|w-\tilde{m}\|_2^2\leqq\varepsilon}} \inf_{t\in\mathbb{R}^I} \sum_{\mu\in I} \frac{1}{2}(w^\mu - t_\mu)^2 + \frac{1}{2^{|I|}} \sum_{\gamma=1}^{2^{|I|}} \phi_\beta\Big(\sum_{\mu\in I} e_\gamma^\mu t_\mu\Big) \tag{3.86}$$

by bounding the sup over $\tilde{t}$ by its value for $\tilde{t} = 0$. On the other hand, we get an upper bound

$$\begin{aligned}\Phi_{N,\beta,\varepsilon}(\tilde{m}) &\geqq \sup_{\substack{w\in[-1,1]^I\\ \|w-\tilde{m}\|_2^2\leqq\varepsilon}} \sup_{\tilde{t}\in\mathbb{R}^{I^c}} \inf_{t\in\mathbb{R}^I} \sum_{\mu\in I} \frac{1}{2}(w^\mu - t_\mu)^2\\ &\quad + \frac{1}{2^p} \sum_{\gamma=1}^{2^p} \operatorname{Conc}\phi_\beta\Big(\sum_{\mu\in I} e_\gamma^\mu t_\mu + \sum_{\nu\in I^c} e_\gamma^\nu \tilde{t}_\nu\Big)\\ &\leqq \sup_{\substack{w\in[-1,1]^I\\ \|w-\tilde{m}\|_2^2\leqq\varepsilon}} \inf_{t\in\mathbb{R}^I} \sum_{\mu\in I} \frac{1}{2}(w^\mu - t_\mu)^2\\ &\quad + \frac{1}{2^{|I|}} \sum_{\gamma=1}^{2^{|I|}} \operatorname{Conc}\phi_\beta\Big(\sum_{\mu\in I} e_\gamma^\mu t_\mu\Big).\end{aligned} \tag{3.87}$$

Here Conc f denotes the concave hull of the function f, i.e. the smallest concave function that is pointwise larger or equal to f. To obtain (3.87) we used two facts: First, the sum of the concave hulls of the functions ϕ_β is in fact the concave hull of the sum, as a function of the t and $\tilde{t}$. Moreover, the function appearing in the first line of (3.87) is symmetric in $\tilde{t}$. Being also concave in $\tilde{t}$, its supremum must be taken on at zero.

Remark. In the case $|I| = 1$ one can easily show that this upper bound coincides with the concave hull of the lower bound (3.86), and one may think that this could be true in general, but we cannot prove this.

4 A Lebowitz–Penrose theorem

Having reduced the FHKP-model effectively to an interacting local mean field model in Sect. 2 we will now use the results on the rate function for the Hopfield model obtained in Sect. 3 to show that the large deviation properties of the total overlap parameters m_Λ of the FHKP-model can be found in terms of those of the usual mean field Hopfield model. This is an analogue of the Lebowitz–Penrose theory [23] of the Kac-model.

As usual, we need to proof upper and lower bounds on the non-normalized versions of the quantities $\mathcal{Q}_{\Lambda,\beta,\gamma}[\omega](\tilde{m})$, that is we define for $\tilde{m} \in \mathbb{R}^I$

$$Z_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \equiv 2^{-(2N+1)} \sum_{\sigma \in \mathcal{S}_\Lambda} \mathbb{1}_{\{\|\Pi_I m_\Lambda(\sigma) - \tilde{m}\|_2^2 \leq \varepsilon\}} e^{-\beta H_{\Lambda,\gamma}[\omega](\sigma)}. \tag{4.1}$$

Using Corollary 2.2, we see immediately that, for almost all ω, for all but a finite number of indices N, this quantity differs from

$$\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \equiv 2^{-(2N+1)} \sum_{\sigma \in \mathcal{S}_\Lambda} \mathbb{1}_{\{\|\Pi_I m_\Lambda(\sigma) - \tilde{m}\|_2^2 \leq \varepsilon\}} e^{-\beta H^0_{\Lambda,\gamma}[\omega](\sigma)} \tag{4.2}$$

only by a factor $e^{\pm(2N+1)\gamma l(\gamma) 4\sqrt{2}\log 2 + \sqrt{2}\gamma M(\gamma)}$ which, if $\gamma l(\gamma)\downarrow 0$ and $\gamma M(\gamma)\downarrow 0$, will give a negligible contribution in the limit $\gamma\downarrow 0$. We thus have only to get bounds on $\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m})$.

In this section we will need the self averaging properties of Sect. 3.2. Due to the continuity problems on the boundary of the set on which $\Phi(m)$ is finite, we have to impose some restrictions on the set of admissible $\tilde{m}$. Given $\tilde{m}$, we defined m^* through $\Phi(m^*) = \inf_{m:\Pi_I m = \tilde{m}} \Phi(m)$. Set $T(I) \equiv \{\tilde{m} | \exists_{c<\infty} \lim\sup_{N\uparrow\infty} \|t^*(m^*)\|_2 \leqq c\}$. In the remainder of this section we assume $\tilde{m} \in T(I)$. $T(I)$ is constant on almost all of Ω and by Lemma 3.2 coincides with the interior of the set $D(I)$. For fixed $\tilde{m} \in T(I)$, we write in the sequel (see (3.51)) $\mathcal{W}_{l,M} \equiv \mathcal{W}_{l,M}(c(\tilde{m}))$.

Let us begin with the lower bound. For this, we write, using (2.7)

$$\begin{aligned}\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) = 2^{-(2N+1)} &\sum_{\sigma \in \mathcal{S}_\Lambda} \mathbb{1}_{\{\|\Pi_I m_\Lambda(\sigma) - \tilde{m}\|_2^2 \leq \varepsilon\}} \\ &\times \exp\Bigg\{\beta \frac{1}{2} l(\gamma) \sum_{(x,y)\in\Gamma\times\Gamma} J_{\gamma l(\gamma)}(x-y) \\ &\times \sum_{\mu=1}^{M(\gamma)} m^\mu_{\Lambda(x)}(\sigma) m^\mu_{\Lambda(y)}(\sigma)\Bigg\}\end{aligned}$$

$$\geqq \sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde m\|_2^2\leqq\delta(\rho)}} 2^{-(2N+1)}\sum_{\sigma\in\mathscr{S}_\Lambda}\prod_{x\in\Gamma}\mathbb{1}_{\{\|m_{A(x)}(\sigma)-m\|_2\leqq\rho\}}$$
$$\times\exp\left\{\beta\frac{1}{2}l(\gamma)\sup_{\|\Pi_I m-\tilde m\|_2^2\leqq\delta(\rho)}\sum_{(x,y)\in\Gamma\times\Gamma}J_{\gamma l(\gamma)}(x-y)\right.$$
$$\left.\times\sum_{\mu=1}^{M(\gamma)}m^\mu_{A(x)}(\sigma)m^\mu_{A(y)}(\sigma)\right\},\tag{4.3}$$

where $\delta(\rho)\equiv(\sqrt{\varepsilon}-\rho)^2$. Equation (4.3) holds for arbitrary $\rho>\sqrt{\alpha}$ but we will later choose $\rho=\rho_\gamma$ that tends to zero with γ in a suitable way. Since under the characteristic functions in the last expression

$$\sum_{(x,y)\in\Gamma\times\Gamma}J_{\gamma l(\gamma)}(x-y)\sum_{\mu=1}^{M(\gamma)}m^\mu_{A(x)}(\sigma)m^\mu_{A(y)}(\sigma)$$
$$=\sum_{(x,y)\in\Gamma\times\Gamma}J_{\gamma l(\gamma)}(x-y)\sum_{\mu=1}^{M(\gamma)}((m^\mu)^2+m^\mu(m^\mu_{A(x)}(\sigma)-m^\mu)+m^\mu(m^\mu_{A(y)}(\sigma)-m^\mu)$$
$$+(m^\mu_{A(x)}(\sigma)-m^\mu)(m^\mu_{A(y)}(\sigma)-m^\mu))$$
$$\geqq\sum_{x\in\Gamma}\|m\|_2^2-2\rho\|m\|_2-\rho^2,\tag{4.4}$$

we get from (4.3)

$$\tilde Z_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde m)\geqq\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde m\|_2^2\leqq\delta(\rho)}}\prod_{x\in\Gamma}2^{-l(\gamma)}\sum_{\sigma\in\mathscr{S}_{Ax}}\mathbb{1}_{\{\|m_{A(x)}(\sigma)-m\|_2\leqq\rho\}}$$
$$\times\exp\left\{\beta\frac{1}{2}l(\gamma)\|m\|_2^2-2\beta l(\gamma)\rho-\beta l(\gamma)\rho^2\right\}.\tag{4.5}$$

In the last line we recognize the function

$$Z^{\mathrm{Hopf}}_{l(\gamma),\beta,\rho}[\omega](m)\equiv 2^{-l(\gamma)}\sum_{\sigma\in\mathscr{S}_{l(\gamma)}}e^{-\beta H^{\mathrm{Hopf}}_{l(\gamma)}(\sigma)}\mathbb{1}_{\{\|m_l(\sigma)-m\|_2\leqq\rho\}},\tag{4.6}$$

so that our lower bound can be expressed in the form

$$Z_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde m)\geqq\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde m\|_2^2\leqq\delta(\rho)}}\prod_{x\in\Gamma}Z^{\mathrm{Hopf}}_{l(\gamma),\beta,\rho}[\omega_x](m)e^{(-2\beta l(\gamma)\rho-\beta l(\gamma)\rho^2)},\tag{4.7}$$

where ω_x is defined, in a slightly abusive way, through the relation that for $i\in\{1,\dots,l(\gamma)\}$, $\xi_i(\omega_x)\equiv\xi_{l(\gamma)x+i}(\omega)$. (Sorry!) Thus, using Kolmogorov's strong law of large numbers, we see that for fixed γ, for almost all ω,

$$\lim_{N\uparrow\infty}\frac{1}{2N+1}\ln\tilde Z_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde m)\geqq\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde m\|_2^2\leqq\delta(\rho)}}\lim_{N\uparrow\infty}\frac{1}{2N+1}$$
$$\times\sum_{x\in\Gamma}\ln Z^{\mathrm{Hopf}}_{l(\gamma),\beta,\rho}[\omega_x](m)-2\beta\rho-\beta\rho^2$$
$$=\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde m\|_2^2\leqq\delta(\rho)}}\frac{1}{l(\gamma)}\mathbb{E}[\ln Z^{\mathrm{Hopf}}_{l(\gamma),\beta,\rho}[\omega](m)]$$
$$-2\beta\rho-\beta\rho^2.\tag{4.8}$$

Furthermore, we can write

$$\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde{m}\|_2^2\leqq\delta(\rho)}} \frac{1}{l(\gamma)}\mathbb{E}[\ln Z^{\text{Hopf}}_{l(\gamma),\beta,\rho}[\omega](m)]$$
$$\geqq \mathbb{E}\left[\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde{m}\|_2^2\leqq\delta(\rho)}} \frac{1}{l(\gamma)}\ln Z^{\text{Hopf}}_{l(\gamma),\beta,\rho}[\omega](m)\right]$$
$$-\mathbb{E}\left(\sup_{m\in\mathscr{W}_{l,M}}\frac{1}{l(\gamma)}|\ln Z^{\text{Hopf}}_{l(\gamma),\beta,\rho}[\omega](m)-\mathbb{E}\ln Z^{\text{Hopf}}_{l(\gamma),\beta,\rho}[\omega](m)|\right). \tag{4.9}$$

The first term in (4.9) is what we want. To control the second, we can use Lemma 3.4. By the bound provided there, a simple calculation shows that

$$\mathbb{E}\left[\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde{m}\|_2^2\leqq\delta(\rho)}} \frac{1}{l(\gamma)}|\ln Z^{\text{Hopf}}_{l(\gamma),\beta,\rho}[\omega](m)-\mathbb{E}\ln Z^{\text{Hopf}}_{l(\gamma),\beta,\rho}[\omega](m)|\right]$$
$$\leqq c\sqrt{\frac{M(\gamma)}{l(\gamma)}} \tag{4.10}$$

for some positive constant c. Here ρ is choosen as in Proposition 3.7. To simplify the notation we will set $\delta=\delta(\rho)$. Taking furthermore advantage of the fact that $\lim_{N\uparrow\infty}(1/2N+1)\ln\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m})$ is necessarily a concave function, we arrive at the lower bound

$$\lim_{N\uparrow\infty}\frac{1}{2N+1}\ln\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m})$$
$$\geqq \operatorname{Conc}\mathbb{E}\left[\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde{m}\|_2^2\leqq\delta}} \frac{1}{l(\gamma)}\ln Z^{\text{Hopf}}_{l(\gamma),\beta,1/M(\gamma)}[\omega](m)\right]$$
$$-c\sqrt{\frac{M(\gamma)}{l(\gamma)}} \tag{4.11}$$

Finally, by using the trivial bound

$$Z^{\text{Hopf}}_{l(\gamma),\beta,\tilde{\varepsilon}}[\omega](\tilde{m})\equiv 2^{-l(\gamma)}\sum_{\sigma\in\mathscr{S}_{l(\gamma)}} e^{-\beta H^{\text{Hopf}}_{l(\gamma)}(\sigma)}\mathbb{1}_{\{\|\Pi_I m_l(\sigma)-\tilde{m}\|_2^2\leqq\tilde{\varepsilon}\}}$$
$$\leqq e^{c'M(\gamma)\ln(l(\gamma)/M(\gamma))}\sup_{\substack{m\in\mathscr{W}_{l,M}\\ \|\Pi_I m-\tilde{m}\|_2^2\leqq\delta}} Z^{\text{Hopf}}_{l(\gamma),\beta,1/M(\gamma)}[\omega](m), \tag{4.12}$$

this becomes

$$\lim_{N\uparrow\infty}\frac{1}{2N+1}\ln\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m})$$
$$\geqq \operatorname{Conc}\mathbb{E}\left[\frac{1}{l(\gamma)}\ln Z^{\text{Hopf},I}_{l(\gamma),\beta,\tilde{\varepsilon}}[\omega](\tilde{m})\right]-c\sqrt{\frac{M(\gamma)}{l(\gamma)}}. \tag{4.13}$$

This is in fact the desired form of the lower bound.

We now derive the upper bound. Here we use the simple fact that

$$m^{\mu}_{\Lambda(x)}(\sigma)m^{\mu}_{\Lambda(y)}(\sigma) \leqq \tfrac{1}{2}(m^{\mu}_{\Lambda(x)}(\sigma))^2 + \tfrac{1}{2}(m^{\mu}_{\Lambda(y)}(\sigma))^2 \tag{4.14}$$

to write

$$\begin{aligned}
&\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \\
&\leqq 2^{-(2N+1)} \sum_{\sigma\in\mathcal{S}_\Lambda} \mathbb{1}_{\{\|\Pi_I m_\Lambda(\sigma)-\tilde{m}\|_2^2 \leqq \varepsilon\}} \\
&\quad \times \exp\left\{\beta\frac{1}{2}l(\gamma)\sum_{x\in\Gamma}\|m_{\Lambda(x)}(\sigma)\|_2^2\right\} \\
&\leqq \sum_{m_x^\mu,\, x\in\Gamma,\, \mu\in I} \mathbb{1}_{\{\|\Pi_I \frac{1}{2L+1}\sum_{x\in\Gamma} m_x - m_x\|_2^2 \leqq \varepsilon\}} \\
&\quad \times \prod_{x\in\Gamma} 2^{-l(\gamma)} \sum_{\sigma\in\mathcal{S}_{l(\gamma)}} \mathbb{1}_{\{\|\Pi_I m_{l(\gamma)}(\sigma)-\tilde{m}\|_2 \leqq \rho\}} \exp\left\{\beta\frac{1}{2}l(\gamma)\sum_{x\in\Gamma}\|m_x\|_2^2\right\} \\
&\leqq \sum_{m_x^\mu,\, x\in\Gamma,\, \mu\in I} \mathbb{1}_{\{\|\Pi_I \frac{1}{2L+1}\sum_{x\in\Gamma} m_x - \tilde{m}\|_2^2 \leqq \varepsilon\}} \\
&\quad \times \exp\left\{\sum_{x\in\Gamma} \ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\rho}[\omega_x](m_x)\right\}.
\end{aligned} \tag{4.15}$$

Since the number of terms in the sum over the m_x^μ is bounded by $[l(\gamma)]^{(2L+1)|I|} = e^{(2N+1)(\ln l(\gamma)/l(\gamma))|I|}$ we can bound (4.15) by the number of terms in the sum times the maximal term. This gives

$$\begin{aligned}
\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) &\leqq e^{(2N+1)\frac{\ln l(\gamma)}{l(\gamma)}|I|} \\
&\quad \times \sup_{\substack{m_x^\mu,\, x\in\Gamma,\, \mu\in I \\ \|\Pi_I \frac{1}{2L+1}\sum_{x\in\Gamma} m_x - \tilde{m}\|_2^2 \leqq \varepsilon}} \exp\left\{\sum_{x\in\Gamma} \ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\rho}[\omega_x](m_x)\right\}.
\end{aligned} \tag{4.16}$$

Therefore, with $\varepsilon' = \sqrt{\varepsilon^2 + \rho}$

$$\begin{aligned}
&\lim_{N\uparrow\infty} \frac{1}{2N+1} \ln \tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) - \frac{\ln l(\gamma)}{l(\gamma)}|I| \\
&\leqq \lim_{L\uparrow\infty} \frac{1}{2L+1} \sup_{\substack{m_x^\mu,\, x\in\Gamma,\, \mu\in I \\ \|\Pi_I \frac{1}{2L+1}\sum_{x\in\Gamma} m_x - \tilde{m}\|_2^2 \leqq \varepsilon}} \sum_{x\in\Gamma} \frac{1}{l(\gamma)} \ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\rho}[\omega_x](m_x) \\
&\leqq \mathrm{Conc}\, \frac{1}{l(\gamma)} \mathbb{E} \ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\varepsilon'}[\omega](\tilde{m}) \\
&\quad + \lim_{L\uparrow\infty} \frac{1}{2L+1} \sum_{x\in\Gamma} \frac{1}{l(\gamma)} \sup_{m_x^\mu,\, \mu\in I} |\ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\rho}[\omega_x](m_x) \\
&\quad - \mathbb{E} \ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\rho}[\omega_x](m_x)|
\end{aligned}$$

$$\leqq \operatorname{Conc} \frac{1}{l(\gamma)} \mathbb{E} \ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\varepsilon'}[\omega](\tilde{m})$$

$$+ \lim_{L\uparrow\infty} \frac{1}{l(\gamma)} \mathbb{E}\left[\sup_{m^{\mu}_{x},\, \mu\in I} \left|\ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\rho}[\omega_x](m_x) - \mathbb{E} \ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\rho}[\omega_x](m_x)\right| \right] \tag{4.17}$$

almost surely, by the Kolmogorov law of large numbers. Moreover, using Lemma 3.1 and the inequality (4.12), we can bound the expectation of the supremum in the last line just as in the case of the lower bound by $c\sqrt{M(\gamma)/l(\gamma)}$. Thus, we have that almost surely,

$$\lim_{N\uparrow\infty} \frac{1}{2N+1} \ln \tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \leqq \operatorname{Conc} \frac{1}{l(\gamma)} \mathbb{E} \ln Z^{\mathrm{Hopf},I}_{l(\gamma),\beta,\varepsilon}[\omega](\tilde{m})$$

$$+ c\sqrt{\frac{M(\gamma)}{l(\gamma)}} + \frac{\ln l(\gamma)}{l(\gamma)}|I|. \tag{4.18}$$

A consequence of these bounds will be the following:

Lemma 4.1. *Assume that* $M(\gamma), l(\gamma)$ *are such that* $\lim_{\gamma\downarrow 0} M(\gamma)\ln l(\gamma)/l(\gamma) = 0$. *Then, for all* $\varepsilon' < \varepsilon''$ *and for almost all* ω,

$$\limsup_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} \tilde{F}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \leqq \limsup_{l\uparrow\infty} \operatorname{Conv} \mathbb{E} F^{\mathrm{Hopf},I}_{l,\beta,\varepsilon'}(\tilde{m}), \tag{4.19}$$

and

$$\liminf_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} \tilde{F}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \geqq \liminf_{l\uparrow\infty} \operatorname{Conv} \mathbb{E} F^{\mathrm{Hopf},I}_{l,\beta,\varepsilon''}(\tilde{m}), \tag{4.20}$$

where

$$\tilde{F}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \equiv -\beta^{-1} \frac{1}{2N+1} \ln \frac{\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m})}{\tilde{Z}_{\Lambda,\beta,\gamma,\varepsilon}[\omega]}.$$

In particular, if the limit $\lim_{l\uparrow\infty} \operatorname{Conv} \mathbb{E} F^{\mathrm{Hopf},I}_{l,\beta,\varepsilon}(\tilde{m}) \equiv C^{\mathrm{Hopf},I}_{\beta,\varepsilon}(\tilde{m})$ *exists and is continuous in* ε, *then, for almost all* ω,

$$\lim_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} \tilde{F}_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) = C^{\mathrm{Hopf},I}_{\beta,\varepsilon}(\tilde{m}). \tag{4.21}$$

If in addition $\lim_{\gamma\downarrow 0} \gamma l(\gamma) = 0$ *then for almost all* ω,

$$\lim_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} F_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) = C^{\mathrm{Hopf},I}_{\beta,\varepsilon}(\tilde{m}). \tag{4.22}$$

Proof. Let us first remark that due to Lemma 2.3, if the statements concerning the limits $\gamma\downarrow 0$ hold for subsequences $\gamma_n = 1/n$ with probability one, then they hold for with probability one for *all* subsequences. By this remark, Eqs. (4.19) and (4.20) follow directly from the bounds (4.18) and (4.13). (4.21) is a direct consequence of (4.19) and (4.20) under the additional assumptions on the existence and continuity of $C^{\mathrm{Hopf},I}_{\beta,\varepsilon}(\tilde{m})$. Finally, the use of the estimates of Sect. 2, in particular Corollary 2.2 allows to replace $\tilde{F}$ by F and thus to obtain (4.22) under the additional assumption on $l(\gamma)$. □

We will now use this lemma to prove Theorems 1 and 2.

Proof of Theorem 1. We consider the situation where $M(\gamma) = |\ln \gamma|/\ln 3$. If we choose $l(\gamma)$ such that $l(\gamma)|\ln l(\gamma)| = \gamma$, then (4.22) relates this FHKP-model to the Hopfield model with $p(l) = \ln(l \ln l)/\ln 3 = \ln l + \ln \ln l/\ln 3$. But this function satisfies the assumption of Proposition 3.5, so that

$$\lim_{l\uparrow\infty} F_{l,\beta,\varepsilon}^{\mathrm{Hopf},I}(\tilde{m}) = \inf_{m:\, \| m - \tilde{m} \|_2^2 \leqq \varepsilon} F_{\beta}^{\mathrm{Hopf},I}(m), \tag{4.23}$$

where $F_{\beta}^{\mathrm{Hopf},I}$ is given in Proposition 3.5. In particular, the continuity of this function on $\Gamma_{|I|}$ implies immediately the continuity of the left-hand side of (4.23) in ε for all $\tilde{m} \in \Gamma_{|I|}$. Thus, under these assumption, (4.22) holds and, moreover,

$$F_{\beta}(\tilde{m}) \equiv \lim_{\varepsilon\downarrow 0} \lim_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} F_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) = \mathrm{Conv}\, F_{\beta}^{\mathrm{Hopf},I}(\tilde{m}) \tag{4.24}$$

exists and is given by the convex hull of the function (3.53).

Now the left hand sides of (4.22) and (4.24) do not depend on the choice of $l(\gamma)$. Therefore, we can make a different choice of $l(\gamma)$, to relate the same FHKP-model to a Hopfield model with different $p(l)$. For any function $p(l)$, such that $p(l) = l\, q(l)$, where $q(l)\downarrow 0$, we just have to choose $l(\gamma)$ in such a way that

$$l(\gamma)\, q(l(\gamma)) = \frac{|\ln \gamma|}{\ln 3} = M(\gamma). \tag{4.25}$$

Then the assumption of Lemma 4.1 on $M(\gamma)$ and $l(\gamma)$ are still satisfied, but the rate functions of the FHKP-model with this $M(\gamma)$ will be related to those of the Hopfield model with the chosen $p(l)$. Now, instead of (4.19) and (4.20), we can derive from (4.13) and (4.18) that, for all $\varepsilon'' > \varepsilon$, for almost all ω,

$$\limsup_{l\uparrow\infty} \mathrm{Conv}\, \mathbb{E}\, F_{l,\beta,\varepsilon}^{\mathrm{Hopf},I}(\tilde{m}) \leqq \lim_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} F_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \tag{4.26}$$

and

$$\liminf_{l\uparrow\infty} \mathrm{Conv}\, \mathbb{E}\, F_{l,\beta,\varepsilon}^{\mathrm{Hopf},I}(\tilde{m}) \geqq \lim_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} F_{\Lambda,\beta,\gamma,\varepsilon''}[\omega](\tilde{m}) \tag{4.27}$$

But the right-hand sides are continuous in ε, and so the limit

$$\lim_{l\uparrow\infty} \mathrm{Conv}\, \mathbb{E}\, F_{l,\beta,\varepsilon}^{\mathrm{Hopf},I}(\tilde{m}) = \lim_{\gamma\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}} F_{\Lambda,\beta,\gamma,\varepsilon}[\omega](\tilde{m}) \tag{4.28}$$

actually exists, almost surely, and is a continuous function of ε. In fact,

$$\lim_{\varepsilon\downarrow 0} \lim_{l\uparrow\infty} \mathrm{Conv}\, \mathbb{E}\, F_{l,\beta,\varepsilon}^{\mathrm{Hopf},I}(\tilde{m}) = \mathrm{Conv}\, F_{\beta}^{\mathrm{Hopf},I}(\tilde{m}), \tag{4.29}$$

with the left-hand side independent of the function $p(l)$. This concludes the proof of Theorem 1. □

Proof of Theorem 2. We have actually just established that the requirements for (4.20) are in fact satisfied as long as $p(l) = l q(l)$, with $q(l)$ tending to zero arbitrarily slowly. Making the choice $l(\gamma) = \sqrt{M(\gamma)\gamma}$, we see that (4.21) and thus (4.25) hold as long as $M(\gamma)$ satisfies $\gamma M(\gamma)\downarrow 0$. But this proves Theorem 2. □

References

1. Amit, D.J., Gutfreund, H., Sompolinsky, H.: Statistical mechanics of neural networks near saturation. Ann. Phys. **173** 30–67 (1987)
2. Bovier, A., Gayrard, V.: Rigorous results on the thermodynamics of the dilute Hopfield model. J. Stat. Phys. **69** 597–627 (1993)
3. Bovier, A., Gayrard, V.: Rigorous results on the Hopfield model of neural networks. Resenhas do IME-USP **1** 161–172 (1994)
4. Bovier, A., Gayrard, V., Picco, P.: Gibbs states of the Hopfield model in the regime of perfect memory. Probab. Theory Relat. Fields **100** 329–363 (1994)
5. Bovier, A., Gayrard, V., Picco, P.: (in preparation)
6. Comets, F.: Large deviation estimates for a conditional probability distribution. Applications to random Gibbs measures. Probab. Theory Relat. Fields **80** 407–432 (1989)
7. Cassandro, M., Orlandi, E., Presutti, E.: Interfaces and typical Gibbs configurations for one-dimensional Kac potentials. Probab. Theory Relat. Fields **96** 57–96 (1993)
8. Deuschel, J.-D., Stroock, D.: Large deviations. Boston: Academic Press 1989
9. Dawson, D.A., Gärtner, J.: Large deviations from the McKeane-Vlasov limit for weakly interacting diffusions. Stochastics **20** 247–308 (1987)
10. Ellis, R.S.: Entropy, large deviations, and statistical mechanics. Berlin Heidelberg New York: 1985
11. Pastur, L.A., Figotin, A.L.: Exactly soluble model of a spin glass. Sov. J. Low Temp. Phys **3**(6) 378–383 (1977)
12. Pastur, L.A., Figotin, A.L.: On the theory of disordered spin systems. Theor. Math. Phys. **35** 403–414 (1978)
13. Pastur, L.A., Figotin, A.L.: Infinite range limit for a class of disordered spin systems. Theor. Math. Phys. **51** 564–569 (1982)
14. Gayrard, V.: The thermodynamic limit of the q-state Potts-Hopfield model with infinitely many patterns. J. Stat. Phys. **68**: 977–1011 (1992)
15. Gärtner, J.: Large deviations from the invariant measure. Theory Probab. Appl. **22** 24–39 (1977)
16. Grensing, D., Kühn, K.: On classical spin-glass models. J. Phys. **48** 713–721 (1987)
17. van Hemmen, J.L.: Spin glass model of a neural network. Phys. Rev. A **34**, 3435–3445 (1986)
18. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA **79**, 2554–2558 (1982)
19. Koch, H.: A free energy bound for the Hopfield model. J. Phys. A: Math Gen. **26**, L353–L355 (1993)
20. Koch, H., Piasko, J.: Some rigorous results on the Hopfield neural network model. J. Stat. Phys. **55**, 903–928 (1989)
21. Kesten, H., Schonmann, R.: Behaviour in large dimensions of the Potts and Heisenberg model, Rev. Math. Phys. **1**, 147–182 (1990)
22. Kac, M., Uhlenbeck, G. and Hemmer, P.C.: On the van der Waals theory of vapour-liquid equilibrium. I. Discussion of a one-dimensional model, J. Math. Phys. **4**, 216–228 (1963); II. Discussion of the distribution functions, J. Math. Phys. **4**, 229–247 (1963); III. Discussion of the critical region, J. Math. Phys. **5**, 60–74 (1964)
23. Lebowitz, J., Penrose, O.: Rigorous treatment of the Van der Waals Maxwell theory of the liquid-vapour transition, J. Math. Phys. **7**, 98–113 (1966)
24. Ledoux, M., Talagrand, M.: Probability in Banach spaces, Springer, Berlin-Heidelberg-New York, 1991
25. Pastur, L., Shcherbina, M.: Absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model, J. Stat. Phys. **62**, 1–19 (1991)
26. Pastur, L., Shcherbina, M. and Tirozzi, B.: The replica symmetric solution without the replica trick for the Hopfield model, J. Stat. Phys. **74**, 1161–1183 (1994)
27. Shcherbina, M., Tirozzi, B.: The free energy for a class of Hopfield models, J. Stat. Phys. **72**, 113–125 (1992)
28. Yurinskii, V.V.: Exponential inequalities for sums of random vectors, J. Multivariate. Anal. **6**, 473–499 (1976)