Mathematical Aspects of Hopfield Models

vorgelegt von
Diplom-Physiker
Beat Niederhauser

Vom Fachbereich 3 – Mathematik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr. rer. nat.
genehmigte Dissertation

Berichter:
PD Dr. A. Bovier
Prof. Dr. J. Gärtner
Prof. Dr. C. Newman

Tag der wissenschaftlichen Aussprache: 29. September 2000

Berlin 2000
D 83

Summary

This thesis is concerned with two models from equilibrium statistical mechanics of disordered systems. Both of them are variants of the Hopfield model, and belong to the class of mean-field models.

In the first part, we treat the case of p-spin interactions ($p \geq 4$ and even) and super-extensively many patterns (their number M scaling as αN^{p-1}). We consider two choices of the Hamiltonians. We find that there exists a critical temperature, at which the replica overlap changes from 0 to a strictly positive value. We give upper and lower bounds for its value, and show that for one choice of the Hamiltonian, both of them converge as $p \to \infty$ to the critical temperature (up to a constant factor) of the random energy model. This critical temperature coincides with the minimum temperature for which annealed free energy and mean of the quenched free energy are equal. The relation between the two results is furnished by an integration by parts formula that is proved by perturbative expansion of the Boltzmann factors. We also calculate the fluctuations of the free energy which are shown to be of the order of $N^{-1/2}$. Furthermore, we find that there exists a critical α above which with large probability the minimum of the Hamiltonian is not realized in the vicinity of any of the patterns. This means that while there is a condensation for low temperatures, the Gibbs measure does not concentrate around the patterns.

In a second part of the thesis, we prove upper bounds on the norm of certain random matrices with dependent entries. These estimates are used in Part I to prove the fluctuations of the free energy. They are proved by the Chebyshev-Markov inequality, applied to the trace of large powers of the matrices. The key ingredient is a representation of the trace of these large powers in terms of walks on an appropriate bipartite graph. This reduces the calculation of the expectation of the trace to the combinatorial problem of counting the maximum number of sub-circuits of a given circuit. The results show that the dependence between the entries cannot be neglected.

Finally, in the last part, we consider a two pattern Hopfield model with Gaussian patterns. We show that there are uncountably many pure states indexed by the circle $\mathcal{S}^1$. This symmetry is randomly broken in the sense that the metastate is supported on a continuum of pairs of pure states that are related by a global (spin-flip) symmetry. We prove these results by studying the random rate function of the induced measure on the overlap parameters. In particular, the breaking of the symmetry is shown to be due to the fluctuations of this rate function at the (degenerate) minima of its expectation. These fluctuations are described by a random process on $\mathcal{S}^1$ whose global minima determine the chosen set (eventually a pair) of states.

Zusammenfassung

Diese Dissertation behandelt zwei Modelle aus der statistischen Mechanik ungeordneter Systeme. Beide sind Varianten des Hopfield-Modells und gehören zur Klasse der Molekularfeldmodelle.

Im ersten Teil behandeln wir den Fall mit p-Spin-Wechselwirkungen ($p \geq 4$ und gerade) und superextensiv vielen Mustern (deren Anzahl M wie αN^{p-1} wächst), wobei wir zwei verschiedene Energiefunktionen betrachten. Wir beweisen die Existenz einer kritischen Temperatur, bei welcher der sogenannte Replikaüberlapp von Null auf einen strikt positiven Wert springt. Wir geben obere und untere Schranken für ihren Wert an und zeigen, daß für die eine Wahl der Hamiltonfunktion beide mit $p \to \infty$ gegen die kritische Temperatur (bis auf einen konstanten Faktor) des Random Energy Model konvergieren. Diese kritische Temperatur fällt mit der kleinsten Temperatur zusammen, für welche die ausgeglühte freie Energie und der Erwartungswert der abgeschreckten freien Energie identisch sind. Der Zusammenhang zwischen diesen beiden Resultaten wird durch eine partielle Integrationsformel geliefert, welche mit Hilfe einer Störungsentwicklung der Boltzmannfaktoren bewiesen wird. Außerdem berechnen wir die Fluktuationen der freien Energie und erhalten, daß sie von der Ordnung $N^{-1/2}$ sind. Weiterhin beweisen wir die Existenz eines kritischen α, oberhalb dessen das Minimum der Hamiltonfunktion mit großer Wahrscheinlichkeit nicht in der Nähe eines der Muster angenommen wird. Dies bedeutet, daß, obwohl sich das Gibbsmaß bei kleinen Temperaturen auf einer kleinen Teilmenge des Zustandsraumes konzentriert, dies nicht in der Nähe der Muster geschieht.

In einem zweiten Teil beweisen wir obere Schranken für die Norm von gewissen zufälligen Matrizen mit abhängigen Einträgen. Diese Abschätzungen werden im ersten Teil zur Berechnung der Fluktuationen der freien Energie benutzt. Sie werden mit der Chebyshev-Markov-Ungleichung, angewandt auf die Spur von hohen Potenzen der Matrizen, bewiesen. Das zentrale Resultat dazu ist eine Darstellung der Spur von diesen hohen Potenzen als Wege auf gewissen bipartiten Graphen. Dies transformiert das Berechnen des Erwartungswertes der Spur auf das kombinatorische Problem, die maximale Anzahl kreisförmiger Teilgraphen eines gegebenen Eulergraphen zu bestimmen. Die Resultate zeigen, dass die Abhängigkeit zwischen den Einträgen eine wichtige Rolle spielt und nicht vernachlässigt werden kann.

Im letzten Teil schließlich betrachten wir ein Hopfield-Modell mit zwei Gauß'schen Mustern. Wir zeigen, daß überabzählbar viele extremale Gibbszustände existieren, welche durch den Einheitskreis $\mathcal{S}^1$ indiziert werden. Diese Symmetrie wird zufällig gebrochen im Sinne, daß der Metazustand von einem Kontinuum von Paaren von extremalen Gibbsmaßen getragen wird, welche durch eine globale Spinsymmetrie verknüpft sind. Wir beweisen diese Resultate mit Hilfe der zufälligen Ratenfunktion des induzierten Maßes auf den Überlapparametern. Insbesondere zeigen wir, daß die Symmetriebrechung durch die Fluktuationen der Ratenfunktion auf den (entarteten) Minima ihrer Erwartung erzwungen wird. Diese Fluktuationen werden durch einen zufälligen Prozeß auf $\mathcal{S}^1$ beschrieben, dessen globale Minima die Menge (schlussendlich ein Paar) der extremalen Zustände auswählen.

To Antoinette, Erik, and Simon

CONTENTS

Part iii: Gaussian Hopfield Model

1 Introduction

1.1 Disordered Systems: Spin Glasses, Biopolymers, and Memory

Disordered systems are modeled in statistical physics by random interactions. The underlying assumption is that the disorder comes about by a process (preparation) one cannot precisely control, but that its effect is such that a typical system (in the frequentist sense of "most of the realizations") behaves as if it were sampled from an appropriate probability distribution. The precise distribution chosen should reflect the knowledge about this process that realizes the disorder. Most often one makes the quite general ('universality') assumption that the results depend only on a few parameters (such as mean and variance), and not on the finer properties of this process. One then argues that any distribution having the appropriate values will give the correct answers.

There are two qualitatively different classes of disordered systems, whose distinction is not sharp, though. The first one could be characterized vaguely by saying that its elements are in some way small perturbations of a standard, non-disordered model. For example, in a model for ferromagnetism on a lattice (Ising, for example), impurities, dislocations, insertions, and other lattice defects may be viewed as small perturbations (provided their density is not too high). Since their precise positions are unknown, one models them in the above spirit by some sort of random variables. In fact, if the results of the standard models are to be taken seriously, they should show some robustness against such small changes, since it is clear that no macroscopic lattice is completely perfect.

On the other hand, there are physical systems that show features that cannot be considered as small perturbations of homogeneous systems. Before providing a motivation for the *Hopfield model* that will be studied in this thesis, we would like to present some of these realistic examples from physics and biology, where a truly random interaction is the appropriate way to model natural phenomena.

Among the most prominent examples are the so called *spin glasses*. Typically, these are substitutional alloys of two or more metals. Examples are binary alloys of the type *noble metal – transition metal* such as AuFe, CuMn, and alloys of two *transition metals* such as FeNi (for more examples see [Cho], Appendix A). Experiments revealed that at low temperatures, the spins are frozen in a seemingly random way.[1] The existence of a phase transition is indicated by the behavior of the susceptibility as a function of temperature. Moreover, their dynamics show very peculiar features. In fact, spin glasses show the phenomenon of *aging*, which means that the dynamical properties depend strongly on the time elapsed since preparation.[2] Recent reviews of theoretical results can be found in [S] (equilibrium thermodynamics), and [BCKM] (dynamical aspects). For a broader exposition, as well as experimental results and techniques, see [Cho].

In this case, the preparation process distributes the moment carrying atoms as substitutions on the lattice. Believing that the above assumption is verified, one replaces this deterministic process by a stochastic one. This process, indexed by the sites of the lattice, indicates for each

[1] More precisely, one observes a local magnetic field characteristic of frozen spins, and an *absence* of corresponding Bragg peaks in neutron diffraction experiments, ruling out a periodic pattern.

[2] This is a consequence of the fact that the system does not truly attain an equilibrium state even on macroscopic time scales.

position the presence or absence of an atom of the particular element. One then introduces a deterministic Hamiltonian, which is supposed to model the interaction between the spins in a more or less realistic way. The disorder enters thus through site variables, and the corresponding model is said to have *site disorder*.

Conversely, one can consider the spin variables to be the same at all sites, and introduce the disorder through a random interaction between pairs of spins. The models are then said to have *bond disorder*.

A sequence of models of decreasing complexity has been introduced over the years, simplifying to the extreme the interactions, but still showing very peculiar features unknown in more classical models, and supposed to grasp some of the aspects of real spin glasses.

On one end of this sequence is the so called *RKKY-model* (after Rudermann-Kittel-Kasuya-Yoshida, see [Cho]). In this site-disorder model, the system is described by variables σ_i (taking value in some compact space) for each site i of the lattice, which interact pairwise via the coupling $J_{i,j} = G_{i,j} n_i n_j$, where n_i are the i.i.d. occupation number random variables (describing absences or presence of magnetic atoms), and $G_{i,j}$ describes the effective coupling between spins by (λ and q_F are two positive parameters)

$$G_{i,j} = \frac{1}{\lambda + |i-j|} \left(\frac{-q_F|i-j| \cos(q_F|i-j|) + \sin(q_F|i-j|)}{q_F|i-j|^3} \right).$$

This model is extremely difficult to analyze and essentially nothing is known on a rigorous mathematical level (see however [Z]).

On the other end, one considers the *Sherrington-Kirkpatrick (SK) model* [SK], which is of the *mean-field* type,[3] and has bond-disorder. In this model, the system is described by variables σ_i taking values ± 1 at each site $i \in \{1, \ldots, N\}$. Their interaction is given by the couplings $N^{-1/2} J_{i,j}$, where the $J_{i,j}$ are i.i.d. standard normal random variables (that is, each pair of spins interacts at the same scale, irrespective of their positions). Physicists predict by non-rigorous methods that this model shows a very peculiar behavior at low temperatures. However, not only the methods, but even the results are difficult to cast into a mathematical form. A nice presentation from a rigorous viewpoint can be found in [NS2].

Let us now turn to an example from biochemistry which touches upon one of the most prominent unsolved problems in this field, namely the folding of biopolymers such as RNA and proteins. These biopolymers can be thought of as a strand of basic monomers[4] whose interactions give the whole polymer its bioactive three dimensional shape (see the relevant chapters of [Sty] for thorough explanations). While it is hopeless to determine analytically the exact structure from the sequence, it is nevertheless interesting to analyze the general aspects of this folding mechanism. In particular, one tries to model the fact that for real biopolymers, there is a critical temperature, above which the polymer *denaturizes*, that is, it unfolds into a random coil. At low temperatures, it assumes its functional form.

[3] Here *mean-field* means that there is no notion of distance between sites. Models of this form are also called *neighbor models*. There exists also a more precise definition of the term mean-field in the setting of disordered systems [BG1], which however does not include the SK-model.

[4] *Amino acids* in the case of proteins, respectively *nucleotides*, each consisting of a *base*, a *sugar*, and a *phosphate group*, in the case of RNA. There are also proteins that consist of several polymers.

If we assume that the polymer building process is not biased towards certain arrangements, then one would sensibly model it by an i.i.d. sequence of letters from a finite alphabet. Obviously, this is an unrealistic assumption of the situation in nature since such a bias is present (this is the whole point of evolution). However, it can be thought of as the state in which nature was in a prebiotic age, that is, before evolution set in. One can then ask whether the above mentioned phase transition occurs also in sequences that did not undergo evolutionary selection (see for example [PPRT]).

Other aspects involve the role of evolution more explicitly. For instance, it is very interesting to note that protein folding takes places on a time scale of milliseconds, which is far less than physicists predict for a stochastic dynamics. This means in fact that the proteins that appear in nature are not only optimized for functionality, but also for folding in that they (almost) never get stuck in a local minimum which does not correspond to the functional shape. Stated differently, the bioactive form is a minima with an extremely large domain of attraction. Certainly, this has to be an effect of evolution (see e.g. [GG] and references therein). Of course, this is just a narrow aspect of protein evolution, since they also have to be optimized for other criteria (functionality, stability in the presence of other bioactive substances).

We now turn to the model which will occupy us for the rest of this work, the *Hopfield model*. It was introduced by Figotin and Pastur [FP1, FP2] as a model for a spin glass. However, it was also introduced independently by Hopfield [Ho] in the context of neural networks, and it is in this spirit that we would like to present it.

This model is not derived directly from a physical or biological system. Rather, it was introduced as simple model for a *content-addressable* (also termed *auto-associative*) memory. This means the following: one wants to store a certain amount of information, and retrieve and/or recognize it on the basis of partial or corrupted data. This is an extremely difficult task for a usual search algorithm. However, it is a task that even very simple living beings like insects are capable of. Hopfield introduced a model based on earlier work by McCullough and Pitts [MCP], and Hebb [H], who respectively proposed a model for the transmission of information by neurons through their synapses, and a rule on how these connections should be altered during the learning process. In the following, we will assume that the system has already learned the information, and we will concentrate on the retrieval mechanism.

To be precise, suppose that $\sigma_i \in \{-1,+1\} = \Sigma$ describes the state of neuron i: firing, or not firing.[5] Suppose furthermore that the system has learned M different binary *patterns* of size N, each of them described by a sequence $(\xi_i^\mu)_{i=1,\dots,N}$, where $\xi_i^\mu \in \{-1,+1\}$.

Based on Hebb's rule, he proposed to associate to the possible states $(\sigma_i)_{i=1,\dots,N} \in \Sigma^N$ of the system an energy functional H, given by

$$H(\sigma) = -\frac{1}{N} \sum_{i,j} \sigma_i \sigma_j \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu. \tag{1.1}$$

Suppose now that the system is fed a corrupted sequence $(\xi_i')_{i=1,\dots,N}$, which does not differ too

[5] These terms refer to whether an electric potential is transmitted across the synaptic interface to the connected neurons. Obviously, a two element state space is a drastic simplification of the biological reality, and is due to [MCP].

much from one of the memorized patterns $(\xi_i^\nu)_{i=1,\dots,N}$. The system should then evolve in such a way as to end up (ideally) in the non-corrupted state $(\xi_i^\nu)_{i=1,\dots,N}$.

Obviously, a gradient dynamics derived from H will typically fail to reach the desired state, since the system gets trapped in any local minima that it encounters. To escape them, one would therefore choose a stochastic dynamics, which will eventually find the global minimum. Natural candidates are dynamics of the Glauber type. These are stochastic processes which have the equilibrium Gibbs measures corresponding to the Hamiltonian (1.1) as their invariant distributions. For such an evolution to end up where it should, there should be an equilibrium Gibbs measure corresponding to the pattern $(\xi_i^\nu)_{i=1,\dots,N}$, meaning that it gives large weight only to a few (compared to the 2^N possible) configurations, which are close to this pattern.

Since we are using the notion of closeness, a word has to be said about distance in the space of configuration. One usually chooses the following function. For two configurations σ, σ', their *overlap* is given by

$$R(\sigma, \sigma') = \frac{1}{N} \sum_i \sigma_i \sigma'_i. \tag{1.2}$$

This parameter is obviously not a distance (since for identical configurations its value is 1). However, it is straightforward to check that it relates to the Hamming distance d_H by

$$R(\sigma, \sigma') = 1 - \frac{2}{N} d_H(\sigma, \sigma').$$

In the special case where σ' is the memorized pattern $(\xi_i^\mu)_i$, one denotes the corresponding overlap by $m^\mu(\sigma)$, that is,

$$m^\mu(\sigma) = \frac{1}{N} \sum_i \sigma_i \xi_i^\mu. \tag{1.3}$$

These latter parameters turn out to be quite important. In fact, the Hamiltonian can be written entirely as a function of them,

$$H(\sigma) = -N \sum_{\mu=1}^{M} \Big(m^\mu(\sigma) \Big)^2, \tag{1.4}$$

or, if one considers m^μ as the μ^{th} component of an M dimensional vector, $H(\sigma) = -N \|m(\sigma)\|_2^2$. The last expression shows that in the case of only one pattern, the model is equivalent to the Curie-Weiss model of ferromagnetism.[6]

So, where does randomness come into play? This is incorporated in the model by the following reasoning. Suppose the model should be capable of storing arbitrary patterns, with no inherent structure (neither in the patterns, nor between them), and one is interested in the behavior of the system for "typical choices" of these patterns. Then it is reasonable to choose the ξ_i^μ as i.i.d. random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\{-1, +1\}$ (obviously, if the stored information takes values in a larger space, then one should also choose the spin space accordingly[7]). If one is to model patterns that have no bias towards one of the two spin values,

[6] To see this equivalence, consider the configuration σ', obtained by the local gauge transformation $\sigma'_i = \sigma_i \xi_i^1$.

[7] However, in Part III we will treat a situation where this is not the case. The motivation for that model lies not really in the context of neural networks, though.

the appropriate choice for $\mathbb{P}$ is clearly the measure which makes the ξ's symmetric Bernoulli variables. Until further notice (that is, until the introduction to Part III), we will adhere to this choice. The introduction of this random variables turns the Hamiltonian into a random variable too. Moreover, the (finite volume) Gibbs measure, defined through

$$\mathcal{G}_N(\sigma) = \frac{1}{Z_N} e^{-\beta H_N(\sigma)}, \tag{1.5}$$

is now a random measure on the space of configurations σ. Since the normalization constant Z_N is random, so is its normalized logarithm, the *free energy*, $F_N = \frac{1}{N} \ln Z_N$.

One is generally interested in the system for large sizes N, and thus also in the thermodynamic limit $N \to \infty$. It turns out that a crucial parameter of the system is the ratio α of number of patterns M to system size N. In fact, if α tends to zero as N grows to infinity, the analysis is much simpler than in the case when α stays strictly positive.

It is beyond the scope of this introduction to overview even the rigorous mathematical results on this model. Let us just briefly indicate one of the most successful strategies to deal with it. Since the Hamiltonian H depends on σ only through the (random) parameters $m^\mu(\sigma)$, the induced distribution $\mathcal{Q}$ of these quantities contains essentially all information about the Gibbs states themselves. The study of $\mathcal{Q}$, which is a measure on the space $\mathbb{R}^M$, turns out to be simpler since it is "less random" than the distribution of the spins. This approach is in the spirit of large deviation theory, that is, one studies the random rate function for $\mathcal{Q}$ which has nice self-averaging properties. For a more detailed discussion, see Section 2.3 in [BG1].

We remark that from the point of view of statistical mechanics, this approach is rather natural if $\alpha \ll 1$. Indeed, if this is the case, the system is controlled by a few (αN) parameters, as opposed to the N spin degrees of freedom of the system.

While much is known about this model already (the verification of the replica symmetric solution [T3], concentration of $\mathcal{Q}$ on the union of small balls [BG1], the weights that are given to the different balls [BM], precise statements about the Gibbs measures [BG3], central limit theorems for the overlap parameters [BG4, GL]), we will not go into any details. The above short explanation serves just as an indication on how one can treat this case. It will turn out that the variant to be introduced shortly is not amenable to this techniques, and this is one motivation to study it.

1.2 The p-Spin Hopfield Model

Having introduced the standard Hopfield model, we now motivate the variants which will be studied in the first part of this thesis. We will then state the main results and indicate some of the auxiliary results used in their proof. This exposition is informal in style. For precise statements and more ideas and remarks, we refer to Chapter 2.

Suppose that we want to incorporate higher order synaptic connections into our Hamiltonian (1.1). A straightforward way to this is to define (compare (1.4))

$$\bar{H}_N(\sigma) = -N \sum_{\mu=1}^{M} \left| m^\mu(\sigma) \right|^p, \tag{1.6}$$

for an even number p, larger than four.[8] This Hamiltonian appeared for the first time in [Lee], respectively [PN]. A crucial role is again played by the number of patterns M. If one chooses it proportional to N, that is $M \sim \alpha N$, then one is in a situation that can be handled by the standard tool of the induced measure $\mathcal{Q}$ of the overlap parameters m^μ (see [BG1] for details). The main point here is that one expects (from numerical simulations) a good retrieval capability even for M as large as αN^{p-1}! The only rigorous results in this situation are to our knowledge due to Newman [N1], treating the question of storage capacity of such a network. More precisely, he gives bounds on the probability that the patterns are surrounded by macroscopic energy barriers at a certain (Hamming) distance. This distance measures the maximal error rate which is allowed in the retrieval process. Furthermore, he finds a relation between p and the maximal α for which the result holds. This confirmed earlier non-rigorous and numerical work.

For normalization reasons, becoming more transparent in Chapter 2, one subtracts a constant from the above Hamiltonian (its expectation) and multiplies the result by some constant s_p, so that our final choice is

$$\bar{H}_N(\sigma) = -\frac{N}{s_p} \sum_{\mu=1}^{\alpha N^{p-1}} \left(\left| m^\mu(\sigma) \right|^p - \mathbb{E} \left| m^\mu(\sigma) \right|^p \right). \tag{1.7}$$

The normalization s_p is in fact chosen in such a way that $\bar{H}$, considered as a random process indexed by the configurations σ has mean zero and covariance function

$$\mathbb{E} \bar{H}_N(\sigma) \bar{H}_N(\sigma') = \alpha N f_p(R(\sigma, \sigma')), \tag{1.8}$$

where f_p is (in leading order in N) a weighted sum of all even powers less than p of its argument, and R is the overlap parameter (distance function) defined in (1.2).

While the interaction (1.7) is the most direct generalization of the usual model, there is a second, in some sense better choice. Observe that in (1.7), the interaction not only contains couplings between groups of p spins, but in fact all multi spin interactions coupling even numbers (less than p) of spins at the same scale. These additional interactions are reflected in the function f_p appearing in the covariance (1.8) of the Hamiltonian.

Let us see what a true p-spin interaction might look like. There is already a disordered model of which has such an interaction, which in addition has Gaussian form, the *p-spin SK-model.* It has been considered recently by M. Talagrand, who made considerable progress in its analysis. In this model, the state space is the same that we consider, but its Hamiltonian is given by

$$H_N^{\mathrm{SK}}(\sigma) = -\left(\frac{p!}{N^{p-1}} \right)^{\frac{1}{2}} \sum_{1 \le i_1 < \ldots < i_p \le N} J_{i_1, \ldots, i_p} \sigma_{i_1} \cdot \ldots \cdot \sigma_{i_p}, \tag{1.9}$$

where the $J_{i_1,\ldots,i_p}$ are i.i.d. standard normal random variables. Its mean is zero and its covariance is simply

$$\mathbb{E} H_N^{\mathrm{SK}}(\sigma) H_N^{\mathrm{SK}}(\sigma') = N R^p(\sigma, \sigma'), \tag{1.10}$$

[8] We will not consider odd p, due to technical difficulties. One expects a similar behavior for this case.

where R is again the overlap defined in (1.2). We now observe that each of the quantities

$$\tilde{J}_{i_1,\dots,i_p} = \left(\frac{1}{\alpha N^{p-1}}\right)^{\frac{1}{2}} \sum_{\mu=1}^{M(N)} \xi_{i_1}^{\mu} \cdot \ldots \cdot \xi_{i_p}^{\mu}$$

converges in distribution to a standard normal random variable. However, although they are pairwise uncorrelated, they are not independent variables. Nevertheless, in analogy with the Hamiltonian (1.9), we define a new Hopfield interaction by

$$\begin{aligned} H_N(\sigma) &= -\left(\frac{p!}{N^{p-1}}\right)^{\frac{1}{2}} \sum_{1\leq i_1<\ldots<i_p\leq N} \tilde{J}_{i_1,\dots,i_p}\sigma_{i_1}\cdot\ldots\cdot\sigma_{i_p} \\ &= -\left(\frac{p!}{N^{p-1}}\right)^{\frac{1}{2}} \sum_{\mu=1}^{M(N)} \sum_{1\leq i_1<i_2<\ldots<i_p\leq N} \prod_{l=1}^{p} \xi_{i_l}^{\mu}\sigma_{i_l}. \end{aligned} \tag{1.11}$$

This function contains only those parts of the Hamiltonian (1.7) that couple exactly p spins, being therefore a pure p-spin interaction. This new Hamiltonian has mean and the covariance (compare (1.10))

$$\mathbb{E} H_N(\sigma)H_N(\sigma') = \alpha N R(\sigma,\sigma')^p, \tag{1.12}$$

in leading order in N. For the rest of this introduction, we will restrict our attention to the interaction H.

As mentioned above, we are interested in the case where M grows super-extensively. Obviously, the induced measure $\mathcal{Q}$ does not help much in this setting. Indeed, this measure now lives in a space of dimension αN^{p-1}, which is "infinitely" much larger than the number N of spin degrees of freedom. Its behavior is therefore at least as difficult to describe as the Gibbs state itself. In particular, there is no hope for a large deviation principle in this case. New tools have therefore to be found. Fortunately, and this provides another motivation to study this model, the progress made by M. Talagrand in the p-spin SK-model (the Hamiltonian (1.9), see [T4]) relies on different methods. This is necessary since in this model, no prototypical spin configurations as the patterns in the case of the Hopfield model are present. Hence, there are no induced measures to be studied either. He was therefore forced to use different methods, which we now applied to our model. However, the SK-Hamiltonian is Gaussian, and this being a very special type of process, one could at first suspect that the approach taken depended strongly on its rather particular properties and would fail to be useful in other settings.

The study of our variant of the Hopfield model thus yields the opportunity to see whether this is true, and this might be the point to announce that these methods, which essentially rely on calculations of second moments of suitably truncated partition functions seem indeed to be rather general and do not depend too strongly on the Gaussian nature of the Hamiltonian. However, the ensuing calculations are much longer than in the non-Gaussian setting of the Hopfield model (as one would expect of course).

We will try to explain the main points of these calculations in Section 2.2 in the case of the most simple disordered mean-field model, the *Random Energy Model* (introduced by Derrida [D1]). In this model, the Hamiltonian is simply i.i.d. random field. For each configuration, the energy is a Gaussian random variable with mean zero and variance N (that is, it does not depend on the precise configuration at all, and they are just used to index the process).

Let us now turn to our results. The principal object of interest is of course the sequence of (random) Gibbs measure (we will for the rest of the introduction consider mainly the interaction H)

$$\mathcal{G}_N(\sigma) = \frac{1}{Z_N} e^{-\beta H_N(\sigma)}, \tag{1.13}$$

where the random quantity Z is called the *partition function.* However, this measure is quite difficult to study as a whole, and we will thus look at it from a particular angle. Observe that for $\beta = 0$ (corresponding to infinite temperature), the measure $\mathcal{G}$ does not depend at all on the interaction, and is thus just the product measure on the on the spins. One can then pose the following, vaguely stated

Question 1: *For which values of the parameters α and β can $\mathcal{G}$ be considered a small perturbation of the product measure on the spins?*

Of course, one has to make precise the notion of closeness. A usual approach is the following: Take two copies of the system with the same realization of the disorder variables ξ, and consider the *order parameter*

$$\mathbb{E}\mathcal{G} \otimes \mathcal{G}\big[|R(\sigma, \sigma')|\big],$$

where, as usual, $\mathbb{E}$ denotes integration with respect to the disorder, and for any function f, $\mathcal{G}[f]$ is its expectation with respect to the Gibbs measure.[9] The above order parameter is conventionally called *replica overlap.*

For $\beta = 0$, one has by the weak law of large numbers,

$$\lim_{N\uparrow\infty} \mathbb{E}\mathcal{G} \otimes \mathcal{G}\big[|R(\sigma, \sigma')|\big] = 0. \tag{1.14}$$

We say therefore that a couple (α, β) lies in the *high-temperature region*, if the associated (random) Gibbs measure satisfies (1.14). The main result can then be stated by the following two partial answers.

Result 1.1: *For each $\alpha > 0$, there exists a critical β_p such that for all $\beta < \beta_p$, the couple (α, β) lies in the high-temperature regime.*

One would like to have a complementary statement, expressing the fact that for all values of β above β_p,

$$\liminf_{N\uparrow\infty} \mathbb{E}\mathcal{G} \otimes \mathcal{G}\big[|R(\sigma, \sigma')|\big] > 0. \tag{1.15}$$

Unfortunately, the result we are able to prove is slightly weaker. Namely, we have

Result 1.2: *For each α and each $\beta > \beta_p$, there exists a set $\mathcal{I} \subset (\beta, \beta_p)$ of strictly positive Lebesgue measure, on which inequality (1.15) holds.*

[9]That is, $\mathcal{G}[f] = \int_{\Sigma^N} f \, d\mathcal{G}$.

Obviously, one expects (or rather hopes) that $\beta_p + \varepsilon > \hat{\beta}_p$. But until now, I have not been able to find a (monotonicity) argument ruling out the contrary, which could be called a *reentrant phase transition.*

In the course of the proof, we obtain upper and lower bounds bounds on the critical β_p, which are both proportional to $\alpha^{-1/2}$ for large values of α, and constant for small values. An analysis of these bounds and some other straightforward calculations show moreover

Result 1.3: *The critical β_p and the mean free energy converge as $p \uparrow \infty$ to the corresponding values of the Random Energy Model at rescaled temperature, that is*

$$\lim_{p\uparrow\infty} \lim_{N\uparrow\infty} \mathbb{E} \frac{1}{N} \ln Z_{N,\beta} = \lim_{N\uparrow\infty} \mathbb{E} \frac{1}{N} \ln Z^{\mathrm{REM}}_{\alpha^{-1/2}\beta},$$

and

$$\lim_{p\uparrow\infty} \beta_p = \alpha^{-1/2} \beta_{\mathrm{REM}}.$$

The Results 1.1 and 1.2 can be expressed as follows. For small values of β, the entropy of the configurations wins against the minima of the Hamiltonian. That is, the measure $\mathcal{G}$ is "spread out" over the configurations. For large β, the measure $\mathcal{G}$ gives a high weight to a comparatively small subset of the configuration space.

It is natural to ask where this concentration[10] takes place, and in particular, whether the configurations close to one of the patterns get this extraordinary weight. Since the configuration where the global minimum of H is attained is a candidate to lie in this subset, we can ask

Question 2: *Does the extremum of the Hamiltonian lie close to one of the patterns?*

A partial answer is given by

Result 2: *For large enough α, the probability that the extremum of H lies in the vicinity of any pattern tends to zero.*

Vicinity means a ball in the Hamming distance centered at the patterns. Their diameter is increasing in α. In fact, we show slightly more: The minimum value of the Hamiltonian on the union of these balls is separated by a macroscopic difference from the absolute minimum. This implies that while we cannot be sure that the absolute minimum is assumed in the subset of large $\mathcal{G}$ measure, the single configuration $\arg\sup H$ has more weight than the union of the balls around the patterns. However, it could still be that there are secondary minima which are much flatter than the absolute one, which would imply that the measure concentrates around these subminimum configurations.

A word or two about the proofs seem to be appropriate. Result 2 follows essentially from the calculations of the fluctuations of H in the balls around the patterns in the spirit of [N1], and from estimates on the extremum of H. Result 1.1 is a consequence of the following result.

[10] *Condensation* might be a physically more appropriate term.

Result: *In leading order in N,*

$$\mathbb{E}\,\frac{\partial F}{\partial \beta} = \frac{\partial F^{\mathrm{an}}}{\partial \beta}\Big(1 - \mathbb{E}\mathcal{G}\otimes\mathcal{G}[R(\sigma,\sigma')^p]\Big), \tag{1.16}$$

where $F = \frac{1}{N}\ln Z$ *is the* free energy *and* $F^{\mathrm{an}} = \frac{1}{N}\ln \mathbb{E}Z$ *is the* annealed free energy*.*

The equivalent of relation (1.16) in the Gaussian SK-models is an exact identity and is just an integration by parts formula [ALR, T4]. Here, we will need an expansion of the Boltzmann factors to prove it. Given this result, one then compares the functions F_N and F^{an}. By Jensen's inequality, it is always true that $\mathbb{E}F_N \le F^{\mathrm{an}}$. One then defines

$$\beta_p = \sup\{\beta : \limsup_N E\,F_N = \lim_N F^{\mathrm{an}}\},$$

from which Result 1.1 follows.

The problem with the low temperature phase is the fact that (1.16) relates derivatives of functions, while one only has knowledge about the function themselves. In the regime $(\beta_p, \beta_p+\varepsilon)$, we simply use a continuity argument (which does not give any bound on ε). For $\beta > \beta_p$, we are in a better situation as we have an estimate on the derivative of $\mathbb{E}F_N$ obtained from a bound on the extremum of H.

Finally, we would like to state some open problems which seem to be worth studying. The ultimate goal is obviously to describe the Gibbs measures completely. While this is for the moment a hopeless task, one expects to gain some insight into the structure of the condensed phase. The following are some steps in this direction, motivated by the successful answers in the case of the p-spin SK-model.

Open problem 1: *Determine the fluctuations of the free energy precisely.*

We are aiming at a result of the following type:

$$\mathbb{P}[|F_N - \mathbb{E}F_N| > cN^{-1/2}] \le e^{-CN}, \tag{1.17}$$

or some other, summable (in N) function on the right, and valid for all β. The reason why one expects this, is the fact that the above result holds for the interaction $\bar{H}$. Moreover, for high temperatures, we will show that for $\bar{H}$, the fluctuations are only of the order N^{-1}. Also, recent results [BKL] show that in the Gaussian models, the order of the fluctuations of the free energy in the hight temperature regime decreases in p (for the SK-model), and is on an exponential small scale in case of the REM (see [BKL]).

If a bound of the form (1.17) is true, then it follows by Borel-Cantelli, that the free energy is self-averaging, that is, $\lim_N |F_N - \mathbb{E}F_N| = 0$, $\mathbb{P}$-almost surely (observe that in general, and in particular for the low temperature regime, it is not expected that $\mathbb{E}F_N$ itself converges).

Also, if such a bound holds, Result 2 can be sharpened to: *For large enough α, with probability one, for all but finitely many N, the minimum of H does not lie in the vicinity of any pattern.*

The main objective is obviously a precise description of the Gibbs measures themselves. The following is reasonable conjecture.

Open problem 2: *Show that in the low temperature phase, the set of configurations which essentially carries the mass of the Gibbs measure is further decomposed into disjoint subsets, termed* lumps*, and show that different lumps are orthogonal. In particular, show that there exist at least two lumps that are not related by a global spin flip.*

In fact, the existence of one lump follows from the fact that the replica overlap is strictly positive (see the construction in [T4]). Once the decomposition of the state space into these lumps is proved, the next step towards the description of the Gibbs measures is

Open problem 3: *Determine the relative weights given to the different lumps, that is, find their order statistics.*

We know that the lumps are not close to any of the patterns for large α. However, if α is small, one expects the contrary:

Open problem 4: *Show that for small α, and β larger than the critical value, the Gibbs measures give large weight to configurations that are close to one of the of the patterns ξ^μ, that is, each lump contains at least one pattern.*

1.3 Norms of Random Matrices

A second part of this thesis is devoted to the study of the norms of certain random matrices. This topic lies somewhat off the main line of this work. However, not only are these results crucial to the proofs of the fluctuations of the free energy in Part I, but the matrices appearing are rather natural and the results in our view of general interest in the context of the spectral theory of random matrices.

Random matrices were introduced by Wigner and Dyson in an attempt to describe resonances of slow neutrons and very heavy nuclei. Since it is a hopeless task to find exactly the highly excited energy levels, it was proposed to study instead an ensemble of Schrödinger operators, satisfying the symmetries prescribed by the physical system. Of primary interest was the distribution of the spectrum of these operators. In his seminal work [Wi1,Wi2], Wigner proved the famous *semi-circle law*. We refer to [Wi3] for a nice overview.

Another important question concerns the behavior of the large eigenvalues.[11] One type of result is a refined analysis of the limiting behavior of the spectral distribution at the edge of the spectrum [SnSo]. Of special is interest is also the operator norm of the matrix, that is, the largest eigenvalue. This point has been studied by Geman [Ge], Füredi and Komlos [FK], and recently by Soshnikov [So]. The types of matrices considered until now encompass principally symmetric N by N matrices with independent entries (*Wigner ensemble*) and sample covariance matrices (*Marchenko-Pastur ensemble*) [Si, BaY, YBaK, BdMS, Ba].

Estimates on the norms of sample covariance matrices have played a crucial role in the investigation of the (standard) Hopfield model [ST, Ko, BGP2, BG1,BG2]. Not surprisingly, estimates on the norms of a different type of random matrices do play a crucial role in the study of our variant of the Hopfield model. The matrices we will consider have the following form. Let $\{\xi_i^\mu\}_{\mu,i\in\mathbb{N}}$

[11] Here and in the following, *large eigenvalue* means large in absolute value.

be an array of i.i.d. Bernoulli random variables, taking values $+1$ and -1 with equal probability. Construct the $M \times M$ matrix A with entries

$$A^{\mu\nu} \equiv M^{-1} \left(\sum_{i=1}^{N} \xi_i^\mu \xi_i^\nu \right)^q = M^{-1} \sum_{(i_l)_{l=1,\dots,q}} \prod_{l=1}^{q} \xi_{i_l}^\mu \xi_{i_l}^\nu. \tag{1.18}$$

We are interested in the behavior of the norm of A when $N \to \infty$ and M scales as $N^{q'}$, that is, $MN^{-q'} \to \alpha$, for some positive constant α.

Before presenting our results, we like to give a (wrong) heuristic argument, which shows that the dependence between the off-diagonal entries of the matrix cannot be neglected. Let us for the moment look at the case $q' = q$, that is, $M = \alpha N^q$. Then the matrix elements of A can be written as

$$A^{\mu\nu} = \alpha N^{-\frac{q}{2}} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \xi_i^\mu \xi_i^\nu \right)^q = \alpha N^{-\frac{q}{2}} (w_{\mu,\nu})^q.$$

Each of the random variables $w_{\mu,\nu}$ converges in law to a standard normal random variable. Moreover, they are pairwise uncorrelated. Suppose now (it is here that we go wrong) that they are all independent. Then we are in the setting of Bai and Yin [BaY] (in particular, their moment condition is satisfied), and from their result, we get that

$$\|A\| \sim \alpha C_q N^{\frac{q}{2}}.$$

It turns out that while this heuristics gives the correct answer if q is even, it is by a factor $\sqrt{N}$ too large in the case of odd q. More precisely, we have

Result 1: *The result is that whenever $q' \geq q \geq 2$, the norm $\|A\|$ satisfies*

$$\|A\| \leq C \begin{cases} N^{\frac{q-1}{2}}, & q \text{ odd} \\ N^{\frac{q}{2}}, & q \text{ even} \end{cases} \tag{1.19}$$

on a set of probability larger than $1 - e^{-N^l}$, for some positive l.

The difference in the result for odd and even q is indeed due to the higher order correlations of the elements, as will be become apparent in the proof. We also remark that the estimates do not depend on q', as long as it is larger than q. This is due to the (deterministic) diagonal terms. Subtracting them would give a new estimate, which involves both q and q'.

A second matrix B we consider is a variant of the above. Namely, in the sum on the right of (1.18), we only retain the "completely off-diagonal" terms. That is,

$$B^{\mu\nu} = \sum_{\substack{(i_l)_{l=1,\dots,q} \\ \text{different}}} \prod_{l=1}^{q} \xi_{i_l}^\mu \xi_{i_l}^\nu, \tag{1.20}$$

where *different* indicates that no two indices have the same value. This restriction may seem harmless, since after all, most choices of values of the indices satisfy it. However, it turns out that the scale of the norm changes drastically. Indeed, we obtain

Result 2: *If $q' > q \geq 2$,*

$$\|B\| \leq C \tag{1.21}$$

on a set of probability at least $1 - e^{-C' N^{\frac{1}{2}-\varepsilon}}$ for all $\varepsilon > 0$ and N large enough.

To understand the above results, it is worthwhile to look at the idea of the proofs. The general strategy to get upper bounds on the norm of a symmetric random matrix M of dimension d is the following. The matrix being symmetric, its trace is equal to the sum of the eigenvalues. Suppose we knew that all of them are positive, then certainly the trace would be an upper bound on the largest of them, and $\frac{1}{d}\mathrm{tr}M$ a lower bound.

Now, look at a very high, even power M^k of M. Then the eigenvalues are indeed all positive. Moreover, the k^{th} power of the largest eigenvalue tends to dominate all others, and for increasing k, the trace of M^k becomes a better and better bound on it.

To get the estimate of the excess probability, one uses this observation together with the Chebyshev-Markov inequality. That is,

$$\begin{aligned} \mathbb{P}\Big[\|M\| > c\Big] &= \mathbb{P}\Big[\|M^k\| > c^k\Big] \\ &\leq \mathbb{P}\Big[\mathrm{tr}M^k > c^k\Big] \\ &\leq c^{-k}\mathbb{E}\,\mathrm{tr}\,M^k. \end{aligned} \tag{1.22}$$

The key to the proof is therefore an accurate upper bound on the expectation of the trace of M^k. In the setting where the matrix M is built up from i.i.d. random variables, one generally tries to represent the trace of M^k as a sum of walks on a graph whose edges correspond to the underlying i.i.d. variables. Taking the expectation then means counting the number of possible walks, that satisfy certain restrictions that are due to the particular distribution of the random variables.

In our case, it will be shown in Chapter 9 that $\mathbb{E}\,\mathrm{tr}\,A^k$ can be calculated by the following procedure. Let the graph $\mathcal{G}$ be a circuit[12] with k edges and r vertices. Let $\mathcal{G}^q$ be the graph obtained from $\mathcal{G}$ by replacing each edge by q edges. The main step then consists in solving

Problem 1: *Determine the maximum number of subgraphs any partition of the edge set of $\mathcal{G}^q$ into circuits can contain.*

It will turn out that one can get a sufficiently good bound $s(k, r)$ of the above quantity in terms of r and k only. Moreover, the partitions with maximum number of elements maximize the number of small subgraphs (with one or two edges). It will be shown that the expectation of the trace is then the sum over all possible graphs $\mathcal{G}$ with $r \leq k$ of the quantity $M^{r-k}N^{s(k,r)}$.

Under the condition $q' > q$, the dominant contribution will come from the term for which $r = k$. This means that by reducing r, the loss in powers of M is much larger than the possible gain due to the larger number of different graphs $\mathcal{G}$. Looking at this maximum term now allows to understand the different behavior for even and odd q. Indeed, if $r = k$, then the graph $\mathcal{G}$ is just a cycle (meaning that no vertex is visited twice). Suppose that q is even. Then we obviously

[12] A *circuit* is a graph such that each vertex has an even number of incident edges.

can decompose the edges between two adjacent vertices into $\frac{q}{2}$ circuits of length 2. As mentioned above, the maximizing partitions are just of this form. Thus $s(r=k,k)=\frac{kq}{2}$.

On the other hand, if q is odd, then only $\frac{q-1}{2}$ circuits of length two can be built between two adjacent vertices of $\mathcal{G}^q$, leaving one edge between them. These remaining edges form a graph isomorphic to $\mathcal{G}$, and so they form one single big circuit which cannot be decomposed further.[13] The total number of circuits is thus $s(r=k,k)=\frac{k(q-1)}{2}+1$. The resulting extra factor N will not play any role, since one chooses at the end k growing with N.

In the case of the matrix B, one proceeds as in (1.22). However, the condition on the values of the indices implies that one has to solve (with $\mathcal{G}$ being the same graph)

Problem 2:*Determine the maximum number of subgraphs of any partition of the edge set of $\mathcal{G}$ into circuits.*

Obviously, this problem is easy to solve once the answer to Problem 1 is known (as the graph under consideration is much simpler). In fact, the maximum number can be bounded again in terms of k and r, namely by $s'(k,r)=k-r+1$. Again, the expectation of the trace is then given by the sum over all possible graphs $\mathcal{G}$ for $r\leq k$ of the quantity $M^{r-k}N^{qs'(k,r)}$. The dominant contribution comes also from the term with $r=k$.

Finally, to actually get exponential estimates of the excess probability, one has to choose k as a power of N. Analysis of the combinatorial terms which appear in the lower order contributions shows that k has to be less than $N^{\frac{1}{2}-\varepsilon}$ for some positive ε.

Before turning to the last part of the introduction, we state again some open problems. As remarked before, the diagonal terms in the matrices prevent us from getting more accurate bounds for q' strictly larger than q (in fact, our bounds do not really involve q'). Thus, we state

Open problem 1: *Find bounds on the matrices A' and B' that are obtained by setting the diagonal entries of A, respectively B to zero.*

To get these bounds, one has to calculate finer estimates on the combinatorics in the analogues of Problem 1, respectively Problem 2.

A second natural problem which seems tractable concerns the distribution of the eigenvalues near the spectral edge. In fact, the recent work of Soshnikov and Sinai [SiSo, So] on Wigner matrices relies essentially on the calculation of very high moments of the trace (up to moments of order $\sqrt{N}$). Since in the course of our proofs we do also calculate these moments (up to almost the same order), it seems reasonable that one could get results in this direction in our case as well. We therefore state vaguely

Open problem 2: *Determine the distribution of the eigenvalues of the matrices A and B.*

1.4 Thermodynamic Limit: Metastates and Chaotic Size Dependence

In the last part of the thesis, we study a simple model of the Hopfield type to illustrate certain notions in the description of large disordered systems and their thermodynamic limit. To put this

[13] Of course, one has to prove that the maximizing partitions are indeed of this form.

into a larger context, we briefly look in this section at some fundamental aspects of equilibrium statistical mechanics of disordered systems. Our model will be introduced and discussed in Section 1.5.

Recall that one of the main goals of statistical mechanics is to describe the phenomenon of *phase transitions*. That is, one tries to solve the apparent paradox that smooth interactions give rise to discontinuous behavior of large systems (such as discontinuity of the density, magnetization etc.). It was realized that no finite system can exhibit this feature, and that the appropriate description is furnished by infinite systems. In doing so, the basic underlying assumption is the following postulate:

A system with a large number of degrees of freedom is close to an infinite system.

Of course, the above has to be given a precise meaning. This means that one has to solve the following two problems:

(a) Define a consistent notion of an infinite system.
(b) In what sense are finite systems close to an infinite system? In particular, if there are more than one infinite volume states (corresponding to a phase transition), which of them describe(s) the finite volume state most accurately?

In the case of lattice spin systems, these points have been answered in a satisfactory way. The theory, which goes back to the seminal work of Dobrushin [Do], and Lanford and Ruelle [LR], is now well developed and understood (see [G, vEFS]). Let us very briefly sketch the set-up for this theory (we follow [B3]). For the sake of an example and to keep difficulties to a minimum, we restrict our attention to models on the lattices $\mathbb{Z}^d$ with finite spin space Σ and finite range interaction $\Phi = \{\Phi_A\}_{A\subset\mathbb{Z},\text{finite}}$.[14] The configuration space $\Sigma_\infty = \Sigma^{\mathbb{Z}^d}$ is equipped with the σ-algebra $\mathcal{F}$ generated by the finite dimensional cylinder sets. We also define for any $\Delta \subset \mathbb{Z}^d$ the σ-algebra $\mathcal{F}_\Delta$, which is generated by the cylinder sets with finite basis in Δ. The measurable space $(\Sigma_\infty, \mathcal{F})$ is then given an *a priori measure* λ, which in the case of finite Σ is usually taken to be the counting measure. For a given interaction Φ, the finite volume Hamiltonians are defined by

$$H_\Lambda(\sigma) = -\sum_{A\cap\Lambda\neq\emptyset} \Phi_A(\sigma).$$

A *local specification for* Φ is then a family of probability kernels $\left\{\pi^{(\cdot)}_{\Lambda,\beta}\right\}_{\Lambda\subset\mathbb{Z}^d}$ from $(\Sigma_\infty, \mathcal{F})$ to itself such that

(i) for all Λ and all $\mathcal{A} \subset \mathcal{F}$, the function $\pi^{(\cdot)}_{\Lambda,\beta}(\mathcal{A})$ is $\mathcal{F}_{\Lambda^c}$-measurable;

[14] In general, compactness of the spin space is quite essential to existence proofs. However, the finite range condition can often be relaxed, replacing it by the notion of a *regular interaction* which means that a condition on the decay at infinity is satisfied. See [G, vEFS] for more details.

(ii) For any $\eta \in \Sigma_\infty$, $\pi^\eta_{\Lambda,\beta}$ is a probability measure on $(\Sigma_\infty, \mathcal{F})$ satisfying

$$\pi^\eta_{\Lambda,\beta}(\sigma) = \frac{e^{-\beta H_\Lambda((\sigma_\Lambda, \eta_{\Lambda^c}))}}{Z^\eta_{\Lambda,\beta}} \lambda_\Lambda(\sigma_\Lambda)\delta_{\eta_{\Lambda^c}},$$

where $(\sigma_\Lambda, \eta_{\Lambda^c})$ is the configuration that agrees with σ on Λ and with η on Λ^c, λ_Λ and σ_Λ are the restrictions to Λ of the respective objects, $Z^\eta_{\Lambda,\beta}$ is the normalization constant, and β the inverse temperature.

Local specifications satisfy compatibility relations analogous to conditional expectations. Namely, for any $\Lambda, \Lambda' \subset \mathbb{Z}^d$, with $\Lambda \subset \Lambda'$,

$$\pi^\eta_{\Lambda',\beta}(\cdot) = \int_{\Sigma_\infty} \pi^\eta_{\Lambda',\beta}(d\sigma)\pi^\sigma_{\Lambda,\beta}(\cdot) = \int_{\Sigma_\infty} \pi^\eta_{\Lambda',\beta}(d\sigma)\pi^{(\eta_{\Lambda'^c}, \sigma_{\Lambda'})}_{\Lambda,\beta}(\cdot),$$

where the second equality follows by from the definitions. This equality is abbreviated by $\pi^{(\cdot)}_{\Lambda',\beta} = \pi^{(\cdot)}_{\Lambda',\beta}\pi^{(\cdot)}_{\Lambda,\beta}$.[15] The specifications can thus be viewed as "conditional expectations waiting for a measure" (quote from [B3]). One thus defines:

A measure μ_β on $(\Sigma_\infty, \mathcal{F})$ is called *compatible* with the local specification $\left\{\pi^{(\cdot)}_{\Lambda,\beta}\right\}_{\Lambda\subset\mathbb{Z}^d}$ if for all $\Lambda \subset \mathbb{Z}^d$ and all $\mathcal{A} \in \mathcal{F}$

$$\mu_\beta(\mathcal{A}|\mathcal{F}_{\Lambda^c}) = \pi^{(\cdot)}_{\Lambda,\beta}(\mathcal{A}), \quad \mu_\beta - a.s.$$

A measure which is compatible with a local specification is called a *Gibbs measure*.

In our setting, the existence of such a measure is guaranteed by compactness. Moreover, all possible infinite volume measures appear as weak limit points in the space $\mathcal{M}_1(\Sigma_\infty)$ of probability measures[16] of the set of finite volume measures (the specifications). This means that by choosing appropriate boundary conditions η, and an increasing and absorbing sequence of finite volumes,[17] the corresponding measures converge weakly to the infinite volume limit. In this sense, both problem (a) and (b) above are solved.

Let us now see what happens in the disordered case. We still assume that the (now random) interaction $\Phi[\omega]$ is finite range, and the spin space is compact. Moreover, we suppose that the underlying probability space $(\Omega, \mathcal{B}, \mathbb{P})$ has a product structure, that is, $\Omega = \Omega_0^{\mathbb{Z}^d}$, where Ω_0 is a topological space, and $\mathcal{B}$ is the Borel σ-algebra generated by the product topology. This set up is valid for most cases of interest. A reasonable definition of a Gibbs measures is then the following:

A measurable map $\mu : \Omega \to \mathcal{M}_1(\Sigma_\infty, \mathcal{F})$ is a *random Gibbs measure for the random interaction* Φ if for almost all ω, $\mu[\omega]$ is compatible with the local specification $\left\{\pi^{(\cdot)}_\Lambda[\omega]\right\}$ for this interaction.

[15] The product of two probability kernels is a probability kernel: $(\pi_1\pi_2)(\omega, \mathcal{A}) = \int \pi_1(\omega, d\omega')\pi_2(\omega', \mathcal{A})$.

[16] The test functions for this topology are the *local functions*, that is, functions that depend only on the value of a finite number of spins.

[17] These volumes should be *smooth*, in the sense that their surface to volume ratio tends to zero along the increasing sequence. For more details on this point, see [vEFS], Section 2.4.1.

Now, in this case, the question of existence of such a measure is more subtle. Of course, by compactness, for almost all ω the finite volume measures $\mu_\Lambda^\eta[\omega]$ taken along an increasing and absorbing subsequence Λ_n has limit points. We can therefore extract subsequences of finite volume measures converging to a Gibbs measure for the interaction Φ. The delicate point here is that these subsequences will in general depend on the realization ω of the disorder, and this in turn questions the measurability of the map $\mu[\omega]$. A way out of this difficulty is to extend the local specifications, which are measures on $(\Sigma_\infty, \mathcal{F})$, to measures $K_{\Lambda,\beta}^\eta$ on the space $(\Omega \times \Sigma_\infty, \mathcal{B} \otimes \mathcal{F})$ such that

(i) the marginal distribution of $K_{\Lambda,\beta}^\eta$ on Ω is $\mathbb{P}$, and
(ii) the conditional distribution, given $\Sigma_\infty \otimes \mathcal{B}$, is the local specification $\mu_{\Lambda,\beta}^\eta[\omega]$.

This in fact suffices to show the existence of a random Gibbs measure if Σ is compact. Indeed, one can show [B3]

Theorem: *If Σ is compact, then there exists an increasing and absorbing sequence Λ_N such that the weak limit*

$$\lim_{N\uparrow\infty} K_{\Lambda_N,\beta}^\eta = K_\beta^\eta,$$

exists, and the conditional distribution

$$K_\beta^\eta(\cdot|\Sigma_\infty \otimes \mathcal{B})$$

is a random Gibbs measure.

It turns out, however, that the resulting Gibbs measure is in some sense a mixed state of systems with disorder that agrees on finite domains. This is due to the fact that the proof involves taking averages over the disorder at infinity (this means averaging over the tail σ-algebra $\mathcal{B}_\infty = \cap_{\Lambda\subset\mathbb{Z}^d}\mathcal{B}_{\Lambda^c}$). In light of question (b) above, this is certainly not an appropriate way of describing the system. A second extension, first proposed by Aizenman and Wehr [AW], and subsequently promoted by Newman and Stein [NS4], should capture in more detail the asymptotic dependence on the disorder.

The setting is the following. Let $\mathcal{M}_1(\Sigma_\infty)$ be the space of probability measures on $(\Sigma_\infty, \mathcal{F})$, equipped with the weak topology and the induced Borel σ-algebra $\mathcal{W}$. Consider the space $\Omega \times \mathcal{M}_1(\Sigma_\infty)$, equipped with the product σ-algebra of $\mathcal{B}$ and $\mathcal{W}$. For any $\Lambda \subset \mathbb{Z}^d$, let $\mathcal{K}_{\Lambda,\beta}^\eta$ be a measure on $\Omega \times \mathcal{M}_1(\Sigma_\infty)$ such that

(i) the marginal distribution on Ω is $\mathbb{P}$, that is

$$\int_{\mathcal{M}_1(\Sigma_\infty)} \mathcal{K}_{\Lambda,\beta}^\eta(d\omega, d\nu) = \mathbb{P}(d\omega);$$

(ii) the conditional measure $\kappa_{\Lambda,\beta}[\omega](\cdot)$ on $\mathcal{M}_1(\Sigma_\infty)$ given $\mathcal{F}$ is the Dirac measure on $\mu_{\Lambda,\beta}^\eta[\omega]$, that is,

$$\kappa_{\Lambda,\beta}[\omega](\cdot) \equiv \mathcal{K}_{\Lambda,\beta}^\eta(\cdot|\mathcal{M}_1(\Sigma_\infty) \otimes \mathcal{F})[\omega] = \delta_{\mu_{\Lambda,\beta}^\eta[\omega]}.$$

Again by compactness, one has existence of limit points of the above objects. More precisely, one proves [B3]

Theorem: *If Σ is compact, then there exist increasing and absorbing sequences of volumes Λ_N such that the limit*

$$\lim_{N\uparrow\infty} \mathcal{K}^{\eta}_{\Lambda_N,\beta} \equiv \mathcal{K}^{\eta}_{\beta}$$

exists. Moreover, the conditional distribution $\kappa^{\eta}_{\beta} \equiv \mathcal{K}^{\eta}_{\beta}(\cdot|\mathcal{B}\otimes\mathcal{M}_1(\Sigma_\infty))$ given $\mathcal{B}$ is a probability distribution on $\mathcal{M}_1(\Sigma_\infty)$ that for almost all ω gives full measure to the set of Gibbs measures corresponding to the underlying interaction. Furthermore,

$$K^{\eta}_{\beta}(\cdot|\mathcal{B}) = \mathcal{K}^{\eta}_{\beta}(\mu|\mathcal{B}).$$

The measure κ^{η}_{β} is called the Aizenman-Wehr (conditioned) metastate*.*

Let us look at two examples.

(i) Suppose that we have almost sure convergence of the local specifications, that is

$$\mu^{\eta}_{\Lambda_N,\beta}[\omega] \to \mu_\infty[\omega], \quad \mathbb{P}-a.s. \tag{1.23}$$

In general, almost sure convergence cannot be expected, and should be considered as exceptional. The corresponding metastate is given by

$$\kappa(\cdot)[\omega] = \delta_{\mu_\infty[\omega]}, \quad \mathbb{P}-a.s. \tag{1.24}$$

That is, if $\mu_\infty[\omega]$ does depend on the realization of the disorder (this should be the generic case), then the metastate shows a non-trivial structure even in the case of almost sure convergence. We will in fact find such a behavior in our model, where we enforce almost sure convergence by an external field.

Now, suppose that there exists an exact symmetry in the system. To be concrete, consider the standard Ising model (non-random) with free boundary conditions. There is no disorder in this model, but we can artificially introduce a degenerate measure $\mathbb{P}$ on the space of interactions. It is well known that below the critical temperature

$$\mu^{\text{free}}_{\Lambda_N,\beta} \to \frac{1}{2}\mu^{+}_{\beta} + \frac{1}{2}\mu^{-}_{\beta},$$

where μ^{+}_{β} and μ^{-}_{β} denote the extremal Gibbs measures with positive, respectively negative mean magnetization. Convergence is obviously almost sure with respect to $\mathbb{P}$ (and the limit does not depend on ω). The metastate is thus simply

$$\kappa(\cdot)[\omega] = \delta_{\frac{1}{2}\mu^{+}_{\beta}+\frac{1}{2}\mu^{-}_{\beta}}.$$

(ii) The metastate gives the most useful information, when the finite volume measures converge in law to some limiting measure, that is, if we have

$$\mu^{\eta}_{\Lambda_N,\beta} \xrightarrow{\mathcal{D}} \mu^{\eta}_{\infty,\beta}.$$

In this case, the δ distribution appearing in (1.23) is replaced by some more general distribution. Our model, which is however of the mean-field type, shows in fact this behavior. We will see that an exact symmetry (global spin flip) is present too, which implies that the corresponding metastate is a distribution on the measures

$$\frac{1}{2}\mu_\beta^+[\omega] + \frac{1}{2}\mu_\beta^-[\omega], \tag{1.25}$$

where the two measures are related by global spin flip and do depend on ω. Our results also exhibit clearly the supplementary information provided by conditioning on $\mathcal{F}$ (compare Theorem 10.3 with Corollary 10.4).

Unfortunately, more interesting, concrete examples are hard to find, and until now, they are mostly restricted to mean-field type models (random field Curie-Weiss model [Ku1, Ku2], Hopfield model [BG3]). Therefore, any new tractable model is welcome, and should be studied to increase our understanding of the mechanisms.

There is also the notion of an *empirical metastate*, introduced by Newman and Stein [NS2, NS3]: Let $\{\Lambda_N\}_N$ be an increasing and absorbing sequence of finite volumes. Define a random empirical measure on $\mathcal{M}_1(\Sigma_\infty)$ by

$$\kappa_N^{\mathrm{em}}(\cdot)[\omega] \equiv \frac{1}{N}\sum_{n=1}^{N} \delta_{\mu^\eta_{\Lambda_N,\beta}[\omega]}.$$

Convergence of this object has been studied for some models by Külske [Ku1]. He found that extremely sparse subsequences are necessary to achieve almost sure convergence, whereas for subsequences that grow more slowly, convergence in law can be shown. In our model as well, we find that for sufficiently sparse sequences, convergence in law holds.

Finally, to capture even more precisely the behavior of the measures along the sequence of increasing volumes, Bovier and Gayrard [BG3] proposed, in analogy with the invariance principle, a *superstate*: For a fixed sequence of volumes Λ_N, let

$$\mu^\eta_{\Lambda_N}(t)[\omega] \equiv (t - \frac{\lfloor tn \rfloor}{n})\mu^\eta_{\Lambda_{\lfloor tn \rfloor +1}}[\omega] + (1 - t + \frac{\lfloor tn \rfloor}{N})\mu^\eta_{\Lambda_{\lfloor tn \rfloor}}[\omega],$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x (this is just the usual linear interpolation scheme, as in the invariance principle). $\mu^\eta_{\Lambda_N}(t)$ is a stochastic process with state space $\mathcal{M}_1(\Sigma_\infty)$. Convergence of this object to some random process $\mu^\eta(t)[\omega]$ can reasonably only be expected in distribution. Thus, we are in the same situation as with the Gibbs measures themselves. One might therefore construct a Aizenman-Wehr metastate on the level of Gibbs measure valued random processes.[18] Again, there are at present only a few examples where detailed information about this object has been obtained, and it is interesting to note that Brownian motion appears in all of them. We refer to [BG3, Ku3] for details. In our case, we are stuck with a S^1 valued random process with quite peculiar features, see Chapter 13.

[18]We remark that for this construction, there is no canonical choice for the appropriate topology, see also [BG3].

This quick overview concentrated on the lattice spin setting. The constructions of AW-metastates, empirical metastates, and superstates can however also been done in the case of mean-field (neighbor) models. The only difference lies in the construction of the Gibbs measures. In particular, since there is no notion of a boundary in this case, limit points of an increasing sequence of volumes are in general mixtures of pure states (these are the extreme elements of the set of Gibbs measures). To construct these pure states, one can either apply an external field (which is taken to zero after the thermodynamic limit), or condition on certain tail events (this means that one works in the canonical ensemble). For a general discussion on the issue of limiting Gibbs measures in mean-field models, we refer to [BG1], Section 2.4, and [BG3], Section 2. With this, we finish our quick tour of general aspects of the thermodynamic limit and turn to some precise, physical questions and conjectures.

The unfortunate point about disordered lattice spin systems is the fact that concrete, mathematically worked out models are scarce (there is essentially one example, the *random field Ising model*). In particular, spin glass type models (that is, models with random multi spin interactions) have turned out to be extremely hard to analyze.

Physicists, however, have proposed a number of different scenarios for the behavior such systems. As is often the case when few rigorous results are present, there is a vigorous debate about the issue. Let us briefly present the different proposals. The main point of the discussion is the question about the number of pure states in lattice spin glasses. On one hand there are the proposals of Fisher and Huse [FH1–4], predicting the existence of only two pure states in any dimension higher than 3. Their conjecture is based on a scaling argument.

At the other extreme, Parisi and collaborators [MPV, MPR] predict an infinity of pure states in the thermodynamic limit. Their proposal is inspired by the (non-rigorous) picture of the SK-model. Although this model is of the mean-field type, it is nevertheless claimed that the situation is also correct for finite dimensional models (down to $d = 3$). In particular, their analysis concentrates on the so called overlap distribution $P(q)$.[19] The use of this order parameter (better: function) in analytical and numerical studies has been questioned in [FH4, NS5].

Intermediate scenarios have been discussed as well [BF, NS1–6, N, vE]. The main idea in the approach of Newman and Stein is to classify the possible scenarios on the basis of first principles, using only general ergodic properties using the concept of metastates described above.

In this context, in one of their most recent papers [NS6], they also conjectured that in a disordered lattice system, in any approximate decomposition of a finite volume Gibbs states into "pure states", the weights in this decomposition should be mostly concentrated on a single subset of states that are related by an exact symmetry of the system, while other states would appear with a weight that tends to zero as the volume tends to infinity. The particular subset chosen could of course be random and could depend strongly on the volume. This behavior is called *chaotic size dependence*.

The model we shall introduce shortly, illustrates these concepts in the case where the number of pure states is uncountable. While models with a finite number of pure states are common, and also a case with countably many states has been treated (the standard Hopfield model with αN patterns [BG3]), the appearance of a continuum of limiting states in a model with discrete spins

[19] The overlap distribution is the distribution on $[0, 1]$ of the replica overlap under the product of two Gibbs measures with the same realization of the disorder.

is rather rare.

1.5 Gaussian Hopfield Model: Random Symmetry Breaking

Let us state the definitions of our variant of the Hopfield model and the main quantities of interest. This is again informal in style; we refer to Chapter 10, for more precise definitions and exact results. The general set up is as in Section 1.1. The (finite) configuration space is $\Sigma_N = \{-1,+1\}^N$. The disorder is modeled by random variables $\xi_i^\mu[\omega]$, $i \in \mathbb{N}$, $\mu = 1,2$. However, in this case, we not only take only two patterns,[20] but they are also standard Gaussian variables instead of Bernoulli.

The overlap parameters $m_N^\mu[\omega](\sigma)$ are defined as in (1.3), that is

$$m_N^\mu[\omega](\sigma) = \frac{1}{N}\sum_{i=1}^{N} \xi_i^\mu[\omega]\sigma_i. \tag{1.26}$$

The Hamiltonian is

$$H_N[\omega](\sigma) = -\frac{N}{2}\sum_{\mu=1,2}\Big(m_N^\mu[\omega](\sigma)\Big)^2 = -\frac{N}{2}\|m_N[\omega](\sigma)\|_2^2. \tag{1.27}$$

This system has a peculiar feature. If we rewrite ${\xi'}_i^1 = {\xi^\alpha}_i^1 = \xi_i^1\cos(\alpha) + \xi_i^2\sin(\alpha)$ and ${\xi'}_i^2 = {\xi'}_i^\alpha = \xi_i^1\sin(\alpha) - \xi_i^2\cos(\alpha)$ the Hamiltonian has the same form in the primed variables. However, this transformation is a *statistical symmetry*, mapping one disorder realization of the model to another one, drawn from the same distribution, as opposed to for example the spin-flip symmetry which is an exact symmetry for any given realization of the disorder.

Through this Hamiltonian, finite volume Gibbs measures on Σ_N are defined by

$$\mathcal{G}_{N,\beta}[\omega](\sigma) \equiv 2^{-N}\frac{e^{-\beta H_N[\omega](\sigma)}}{Z_{N,\beta}[\omega]}. \tag{1.28}$$

We will be concerned exclusively with the low temperature region, that is $\beta > 1$. Since the number of pattern is very small compared to the system size, we base our analysis on the induced distribution of the overlap parameters (compare the remarks in Section 1.1, page 5)

$$\mathcal{Q}_{N,\beta}[\omega] \equiv \mu_{N,\beta}[\omega] \circ m_N[\omega]^{-1}. \tag{1.29}$$

The extremal Gibbs measures are constructed by *tilting* the Hamiltonian (1.27) with an *external magnetic field*, that is,

$$H_N^h[\omega](\sigma) \equiv -\frac{N}{2}\|m_N[\omega](\sigma)\|_2^2 - N(h, m_N[\omega](\sigma)), \tag{1.30}$$

where $h = (b\cos(\vartheta), b\sin(\vartheta)) \in \mathbb{R}^2$. The corresponding measures on the spins and on $\mathbb{R}^2$ are denoted by $\mathcal{G}_{N,\beta}^h[\omega]$ and $\mathcal{Q}_{N,\beta}^h[\omega]$, respectively. So, the first problem to be solved is

[20] Any finite number would do; two is the least non-trivial case and is chosen to keep technicalities minimal.

Question 1: *What is the set of extremal measures?*

The answered is found by taking first the thermodynamic limit and then relaxing the magnetic field to 0, that is, the iterated limits $\lim_{b\downarrow 0}\lim_{N\uparrow\infty}$.

Result 1.1: *For each direction of the external field, the measures $\mathcal{Q}_N^h[\omega]$ and $\mathcal{G}_N^h[\omega$ converge almost surely. The limit*

$$\lim_{\beta\downarrow 0}\lim_{N\uparrow\infty} \mathcal{Q}_N^h[\omega] = \delta_{(r^*\cos(\vartheta), r^*\sin(\vartheta))},$$

where r^* is a positive number depending on β, is independent of ω, whereas the limit of $\mathcal{G}_N^h$ does depend on ω.

This means that the AW-metastate on the level of the induced measures is just a Dirac mass on a deterministic point mass in $\mathbb{R}^2$. On the other hand, the metastate on the level of the Gibbs measures is a Dirac mass on a random measure depending on the realization of the disorder. We have here the situation (1.23), respectively (1.24). Since there is one degree of freedom in the magnetic field (its direction), one readily gets

Result 1.2: *The set of limiting induced measures is indexed by the circle Σ^1. Moreover, for each $\omega\in\Omega$, the set of limiting Gibbs measures is indexed by Σ^1.*

The more interesting problem is the case without a tilting field.

Question 2: *What are the limiting states when no external field is applied?*

It turns out that in this case we are in the situation described under point (ii) (page 18), namely that one has convergence in distribution of the measures $\mathcal{Q}_N$ and $\mathcal{G}_N$, and the corresponding metatstate is of the form (1.25).

Result 2: *Both $\mathcal{Q}_N[\omega]$ and $\mathcal{G}_N[\omega]$ converge in distribution. The AW-metastate on the level of the induced measures is the image of the uniform distribution on $[0,\pi)$ under the map*

$$[0,\pi)\ni\vartheta\mapsto \frac{1}{2}\delta_{m(\vartheta)} + \frac{1}{2}\delta_{-m(\vartheta)},$$

where $m(\vartheta) = (r^*\cos(\vartheta), \rho^*\sin(\vartheta)) \in \mathbb{R}^2$ and r^* is as in Result 1.1. The AW-metastate on the level of the Gibbs measures is the image of the uniform distribution on $[0,\pi)$ under the map

$$[0,\pi)\ni\vartheta\mapsto \mathcal{G}_{\infty,\beta,m}[\omega](\{\sigma_I = s_I\}) = \frac{1}{2}\prod_{i\in I}\frac{e^{\beta s_i(\xi_i[\omega],m)}}{2\cosh\beta(\xi_i[\omega],m)} + \frac{1}{2}\prod_{i\in I}\frac{e^{-\beta s_i(\xi_i[\omega],m)}}{2\cosh\beta(\xi_i[\omega],m)}.$$

The fact that the metastates are images of the uniform distribution on an interval is a consequence of the stochastic symmetry which was mentioned before. We also mention that the breaking of the stochastic symmetry is not universal. In particular, the standard Hopfield model with two patterns, that is, if the ξ_i^μ are i.i.d. ± 1 Bernoulli random variables, the fluctuations are too small too provoke it. The metastate on the level of the induced measures is in this case supported on point masses at all four points in $\mathbb{R}^2$ that are related by a stochastic symmetry (this points lie on the coordinate axes at a certain distance m^* from the origin, where m^* satisfies the Curie-Weiss consistency relation $m^* = \tanh \beta m^*$).

The proofs of the above results are close in spirit to large deviation theory. More precisely, one deals with a random rate function for the induced measures $\mathcal{Q}$. In particular, its fluctuations have to be controlled. The proof of Result 2 is the more delicate one. In fact, one has to consider the number of global extrema of a sequence of random processes on Σ^1 that converges in distribution to a Gaussian process. We thus prove first a result on the limiting process, and use then a strong approximation result. This means that there exists a sequence of processes that converge almost surely to the limiting process and is equal in distribution to the original sequence. This allows to get the results for the original sequence. We observe that it is absolutely necessary that we have a finite number of patterns, which implies that all the processes take value in the same space, namely $\mathbb{R}^2$ (or some other finite dimension).

So, Result 2 shows that finite systems are eventually well approximated by a pair of states. The next question is then obviously, which of the uncountably many achieve this:

Question 3: *Which limiting measures describe a finite system most accurately?*

We will show that the answer to the above question depends sensitively on the system size. In particular, there is a subsequence of volumes such that along this sequence, the best infinite pairs are distributed uniformly on the circle. On the level of the induced measures, this is the following statement.

Result 3: *There exist deterministic sequences of volumes N_k such that the empirical metastate taken along N_k converges almost surely to the law of the limiting induced measure.*

This means that the sequence of measures comes close to any of the infinite states. The system thus shows chaotic size dependence. The result relies on the fact that finite size measures are approximately independent if the sequence is sparse enough ($N_k = k!$ will do).

Finally, we want to conclude with some open problems inspired by this model. As mentioned above, it is crucial that we have only a finite number of patterns. A canonical question is thus

Open problem 1: *Consider the case of a growing number of patterns.*

The methods used in the results above depend critically on the fact that the number of patterns is finite. New ideas have therefore to be found to treat this problem.

All models treated in this thesis are of mean-field type. A generalization in a different direction is to look at a corresponding model on the lattice with long-range interactions.

Open problem 2: *Consider a Kac variant of this model. That is, take as phase space $\{-1,+1\}^{\mathbb{Z}}$ and a formal interaction,*

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

where $J_\gamma(i-j) = \gamma J(\gamma|i-j|)$, and $J(x) = \mathbb{1}_{[-1/2,1/2]}$.

Kac models have recently gained a renewed interest. In particular, the magnetization profile was considered (for example in the random field Ising model [COP], or in the standard Hopfield model [BGP3]). In our case, the relevant feature is again that the mesoscopic magnetization has uncountably many equilibrium values (indexed by S^1), being thus closer to a rotor model (in the sense that there may be arbitrarily small excitations of the ground states). With this, we conclude the introduction to this thesis.

1.6 Acknowledgments

It is a great pleasure to thank the persons who made this work possible. First and foremost, I want to express my gratitude to my advisor Anton Bovier for the collaboration over the past years. Through countless discussions, he not only taught me the basics of the Hopfield model, but transmitted intuition and enthusiasm for mathematical physics.

I am very grateful to Aernout van Enter who proposed the model studied in Part III, for the joint paper that resulted from it [BvEN], and valuable discussions. I am indebted to Jürgen Gärtner and Charles Newman for readily accepting the task of being co-examiners. Furthermore, I thank the organizers of the "Prague Phase Transition" schools, Roman Kotecký and Marek Biskup, for six interesting weeks, and the staff at Eurandom, in particular its director Willem van Zwet and the coordinator of the "Interacting Stochastic Systems" project, Frank den Hollander, for making my research stay in Eindhoven possible.

Special thanks go to Charles Pfister who introduced me to mathematical physics and without whom I would not have had the opportunity to come to Berlin. I profited enormously from discussions with Ostap Hryniv, Christof Külske, and Yvan Velenik. I thank the Graduiertenkolleg "Stochastische Prozesse und Probabilistische Analysis", and in particular its speaker Michael Scheutzow for providing an excellent working environment. Furthermore, I am grateful to my friends and colleagues Dirk Becherer, Pietro Caputo, Frank Döberlein, Daniel Egloff, Felix Esche, Urs Gruber, Michael Kirch, Thorsten Rheinländer, and Rosa Sahin. They made these three years an interesting and unforgettable period. Finally, I thank my parents and my brother for permanent moral support. Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

Part I

The Multi-Spin Interaction Model

2 Results and Relation to the REM

2.1 Definitions and Results

In this first part, we deal with the Hopfield model with p-spin interactions. We first recall the definitions and then present in detail our results. There will be a certain overlap with Section 1.2, but statements are cast into a more precise form.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space and $\{\xi_i^\mu\}_{i,\mu\in\mathbb{N}}$ a family of i.i.d. Bernoulli variables, taking values 1 and -1 with equal probability. Let $\Sigma = \{+1, -1\}$, and Σ^N the Cartesian product. We equip Σ^N with the product topology and the uniform distribution as a priori measure $\mathbb{P}_\sigma$.

Define for each $N \in \mathbb{N}$ a (finite) random *Hamiltonian*, that is, a function $H_N : \Omega \times \Sigma^N \to \mathbb{R}$ by

$$H_N[\omega](\sigma) \equiv -\left(\frac{p!}{N^{2p-2}}\right)^{\frac{1}{2}} \sum_{\mu=1}^{M(N)} \sum_{i_1<\ldots<i_p} \prod_{l=1}^{p} \xi_{i_l}^\mu \sigma_{i_l}. \tag{2.1}$$

The value of p is considered a fixed parameter of the model, and will in the following be even and at least be 4. The upper limit $M(N)$ in the first sum of (2.1) should scale as N^{p-1}, i.e.

$$\lim_{N\uparrow\infty} \frac{M(N)}{N^{p-1}} = \alpha < \infty. \tag{2.2}$$

The limit α will also turn out to be a crucial parameter for the behavior of the system. In the sequel, we will write with slight abuse of notation $M(N) = \alpha N^{p-1}$ even for finite N.

To simplify notation, we will use the following multiindex notation. For finite subsets $\mathcal{I}$ of the natural numbers, and real numbers $(x_n)_{n\in\mathbb{N}}$, let by $x_\mathcal{I} = \prod_{l\in\mathcal{I}} x_l$. The Hamiltonian (2.1) can then be written as

$$H_N[\omega](\sigma) = -\left(\frac{p!}{N^{2p-2}}\right)^{\frac{1}{2}} \sum_{\mu=1}^{M(N)} \sum_{\substack{\mathcal{I}\subset\mathcal{N} \\ |\mathcal{I}|=p}} \xi_\mathcal{I}^\mu \sigma_\mathcal{I}, \tag{2.3}$$

where $\mathcal{N} = \{1, \ldots, N\}$.

These Hamiltonians define random, finite volume Gibbs measures $\mathcal{G}_{N,\beta}[\omega]$ by assigning each configuration $\sigma \in \Sigma^N$ a weight proportional to its Boltzmann factor, that is

$$\mathcal{G}_{N,\beta}[\omega](\sigma) = 2^{-N} \frac{e^{-\beta H_N[\omega](\sigma)}}{Z_{N,\beta}[\omega]}. \tag{2.4}$$

Consider now the Hamiltonian as a random process indexed by $\sigma \in \Sigma^N$. We state the following identities, which are verified by direct calculations. The mean of H_N with respect to $\mathbb{P}$ vanishes for all σ, that is

$$\mathbb{E} H_N(\sigma) = 0, \quad \forall \sigma \in \Sigma^N, \tag{2.5}$$

whereas the variance satisfies (for some number C depending on p only)

$$\alpha N(1 - CN^{-1}) \leq \mathbb{E} H_N(\sigma)^2 = \frac{p!}{N^{2p-2}} \sum_{\mu=1}^{M(N)} \sum_{\substack{\mathcal{I}\subset\mathcal{N} \\ |\mathcal{I}|=p}} 1 \leq \alpha N, \tag{2.6}$$

which motivates our choice of normalization in the definition of H_N. The covariance is given as

$$\mathbb{E} H_N(\sigma) H_N(\sigma') = \frac{p!}{N^{2p-2}} \sum_{\mu=1}^{M(N)} \sum_{\mathcal{I}\subset\mathcal{N}} \sigma_{\mathcal{I}} \sigma'_{\mathcal{I}} = \alpha N R^p(\sigma,\sigma')(1+\mathcal{O}(N^{-1})), \tag{2.7}$$

where $R(\sigma,\sigma') \equiv \frac{1}{N}\sum_{i=1}^N \sigma_i \sigma'_i$ is the (normalized) *replica overlap*. Note that this covariance is in leading order and up to the factor α the same as the covariance for the p-spin SK-model ([T4]).

The normalizing factor Z_N in (2.4) is called *partition function* and it is given by

$$Z_{N,\beta}[\omega] = \mathbb{E}_\sigma\, e^{-\beta H_N[\omega](s)}, \tag{2.8}$$

where $\mathbb{E}_\sigma$ is the expectation with respect to the a priori distribution on Σ^N. The *annealed partition function* is the mean of Z_N under $\mathbb{P}$. Observe that for any $\sigma \in \Sigma^N$, the Hamiltonian $H_N(\sigma)$ has the same distribution, and thus $\mathbb{E}\mathbb{E}_\sigma Z_N = \mathbb{E}_\sigma \mathbb{E}\, e^{-\beta H_N} = \mathbb{E}\, e^{\beta H_N(\sigma)}$.

The *quenched free energy* $F_{N,\beta}[\omega]$ is defined as the normalized logarithm of the partition function, that is, $F_{N,\beta}[\omega] \equiv \frac{1}{N}\ln Z_{N,\beta}[\omega]$.[21] The *annealed free energy* $F^{\text{an}}_{N,\beta}$ is the normalized logarithm of the expectation of the partition function, i.e. $F^{\text{an}}_{N,\beta} = \frac{1}{N}\ln \mathbb{E} Z_{N,\beta}$. Observe that by Hölders's inequality, the quenched free energy and the annealed free energy are convex functions of β.

We also recall the second model that will be considered. On the same configuration space and with the same random variables ξ, we define macroscopic random order parameters

$$m^\mu[\omega](\sigma) \equiv \frac{1}{N}\sum_{i=1}^N \xi_i^\mu \sigma_i. \tag{2.9}$$

These parameters are considered as components of a vector in $\mathbb{R}^{M(N)}$ with $M(N)$ as in (2.2) (in fact, the resulting vector is confined to $[-1,1]^{M(N)}$). A new sequence of random Hamiltonians in now defined through

$$\bar{H}_N[\omega](\sigma) - \frac{N}{s_p}\left(\|m[\omega](\sigma)\|_p^p - \mathbb{E}\,\|m[\omega](\sigma)\|_p^p\right), \tag{2.10}$$

where $s = s_p > 0$ satisfies

$$s^2 = s_p^2 = (2p-1)!! - ((p-1)!!)^2, \tag{2.11}$$

and $n!!$ is defined as[22]

$$n!! = \begin{cases} \displaystyle\prod_{i=0}^{\frac{n-2}{2}} (n-2i), & \text{if } n \text{ is even;} \\ \displaystyle\prod_{i=0}^{\frac{n-1}{2}} (n-2i), & \text{if } n \text{ is odd.} \end{cases} \tag{2.12}$$

[21] Note that physicists often use a different normalization, $F_N = -\frac{1}{\beta N}\ln Z_N$.

[22] This is the number of possible arrangements of $n+1$ elements in pairs.

From the above definitions, one easily deduces that

$$\mathbb{E}\,\bar{H}_N(\sigma) = 0, \quad \forall\sigma \in \Sigma^N, \tag{2.13}$$

and

$$\mathbb{E}\,\bar{H}_N(\sigma)\bar{H}_N(\sigma) = \alpha N(1 + \mathcal{O}(N^{-1})). \tag{2.14}$$

The fundamental difference to the Hamiltonian H appears, when we look at the covariance of $\bar{H}$. It is given by

$$\begin{aligned}\mathbb{E}\,\bar{H}_N(\sigma)\bar{H}_N(\sigma') &= \frac{\alpha N}{s_p^2} \sum_{\substack{q=2\\ \text{even}}}^{p} ((p-q-1)!!)^2 \binom{p}{q}^2 q!\, R(\sigma,\sigma')^q (1+\mathcal{O}(N^{-1})) \\ &= \alpha N g_p(R(\sigma,\sigma'))(1+\mathcal{O}(N^{-1})),\end{aligned} \tag{2.15}$$

where the double factorial is set to one for non-positive values of the argument. From the above expression, we clearly see that the Hamiltonian $\bar{H}$ always contains a contribution corresponding to the $p = 2$ Hamiltonian. It is therefore not surprising, that the limit $p \to \infty$ behaves quite differently than in the case of the interaction H. We define the Gibbs measures and the free energies in exactly the same way as for H. The resulting quantities will be denoted by an additional bar.

The first result we prove for both choices of the Hamiltonian is that for high enough temperatures (that is, low values of β), the limit of the annealed free energy exists.

Theorem 2.1: *If $\beta < e^{-2}(p!)^{\frac{1}{2}} \equiv \beta'_p$, then the annealed free energy corresponding to H satisfies*

$$F_{N,\beta}^{\text{an}} = \frac{\alpha\beta^2}{2}(1 + \mathcal{O}(N^{-1})). \tag{2.16}$$

The corresponding result for the interaction $\bar{H}$ reads

Theorem 2.1': *Assume that $\beta < \frac{s_p}{2} = \bar{\beta}'_p$. Then*

$$\bar{F}_{N,\beta}^{\text{an}} = \frac{\alpha\beta^2}{2}(1 + \mathcal{O}(N^{-1})). \tag{2.17}$$

Note that for larger values of β, the annealed free energy diverges, that is, $\lim_N F_{N,\beta}^{\text{an}} = +\infty$. Jensen's inequality implies that the expectation of the free energy is less than the annealed free energy,

$$\mathbb{E}\,F_{N,\beta} = \frac{1}{N}\mathbb{E}\,\ln Z_{N,\beta} \le \frac{1}{N}\ln \mathbb{E}\,Z_{N,\beta} = F_{N,\beta}^{\text{an}}. \tag{2.18}$$

We define the critical temperature to be the infimum of values for which the annealed free energy exists and equality in (2.18) holds, i.e. in terms of β,

$$\beta_p \equiv \sup\left\{\beta \ge 0 : \limsup_{N\uparrow\infty} \mathbb{E}\,F_{N,\beta} = \limsup_{N\uparrow\infty} F_{N,\beta}^{\text{an}}\right\}, \tag{2.19}$$

and the analogue value β'_p for the Hamiltonian $\bar{H}$. Observe that in general $\lim_N \mathbb{E} F_N$ need not exist.

Relation (2.7) shows that up to a very small error, the covariance of the interaction H is identical to the covariance of p-spin SK Hamiltonian. This motivates a comparison with the *random energy model* (REM), introduced by Derrida [D1,D2] as a caricature of a spin-glass. Let us recall its definition. In this model, the energy associated to the spin configurations $\sigma \in \Sigma^N$ are i.i.d. Gaussian variables $\sqrt{N} X_\sigma$ with mean zero and variance N (i.e. the $\{X_\sigma\}_{\sigma \in \Sigma^N}$ are i.i.d. standard normal random variables). The partition function is then

$$Z_{N,\beta}^{\text{REM}} = \mathbb{E}_\sigma e^{\beta \sqrt{N} X_\sigma}. \tag{2.20}$$

It follows from standard results on extremes of independent random variables (for this and sharp results on the fluctuations see [BKL]) that[23]

$$f_\beta^{REM} = \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \ln Z_{N,\beta}^{\text{REM}} = \begin{cases} \beta^2/2, & \text{if } \beta \leq \sqrt{2\ln 2} \\ \beta\sqrt{2\ln 2} - \ln 2, & \text{if } \beta \geq \sqrt{2\ln 2} \end{cases} \tag{2.21}$$

whereas the annealed free energy satisfies

$$\frac{1}{N} \ln \mathbb{E}\, Z_{N,\beta}^{\text{REM}} = \frac{\beta^2}{2}, \tag{2.22}$$

for all values of β.

It will turn out that as p tends to infinity, $\sqrt{\alpha}\beta_p$ tends to the critical value $\sqrt{2\ln 2}$ of the REM. Moreover, pointwise in α, β,

$$\frac{1}{\beta} \lim_{p\to\infty} \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \ln Z_{N,\beta} = \frac{1}{\beta} f_\beta^{REM}. \tag{2.23}$$

It is a priori not clear, why this behavior should be expected. First of all, our couplings between spins are not Gaussian variables. Also, although the covariance process of our Hamiltonian converges pointwise to the covariance process of the REM Hamiltonian, it is obviously a "quite different matter to make the iterated limits $\lim_{p\uparrow\infty} \lim_{N\uparrow\infty}$, or the iterated limits $\lim_{N\uparrow\infty} \lim_{p\uparrow\infty}$" (quote from [T4], page 3).

The reason is that the great number of μ's makes the "spin couplings" behave in some sort like Gaussians. This fact is already reflected in the calculation of the annealed free energy. In fact, the bound β'_p is exactly the value for which the Gaussian approximation fails. Below this value of β, one shows that the free energy, as well as the critical β are essentially equal to those in the p-spin SK-model, for which the convergence of these quantities to the REM have been shown (compare [T4]). The precise statement is given by the next theorem. First, recall the definition of the *Cramér entropy* $I(t)$:

$$I(t) = \begin{cases} \frac{1}{2}(1-t)\ln(1-t) + \frac{1}{2}(1+t)\ln(1+t), & \text{if } |t| \leq 1, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.24}$$

[23] This is already contained in [D2]

We then have the following bounds on the critical temperature.

Theorem 2.2: *The critical value $\beta_p = \beta_p(\alpha)$ satisfies*

$$\beta_p(\alpha)^2 \geq \min\left(\frac{\beta_p'^{\,2}}{4}, \inf_{t\in[0,1]} I(t)\frac{1+t^p}{\alpha t^p}\right) \equiv \check{\beta}_p(\alpha)^2. \tag{2.25}$$

Furthermore, if $\alpha \geq \frac{8\ln 2}{p!} = \alpha_p$ then

$$\beta_p(\alpha)^2 \leq \frac{2\ln 2}{\alpha} \equiv \hat{\beta}_p(\alpha)^2. \tag{2.26}$$

If $\alpha \leq \alpha_p$, then $\beta_p^2 \leq \frac{p!}{4}$.

Remarks: (i) It is reasonable to suspect that the inequality (2.26) is strict. Indeed, in the case of the p-spin SK-model, this can be shown by a judicious bound on the supremum of the Hamiltonian which supposedly could be used in our case as well [B2].

(ii) The bounds on the critical temperature are essentially (up to a factor $\sqrt{\alpha}$) the same as for the p-spin SK-model ([T4], Theorem 1.1).[24]

It is elementary that as p tends to infinity,

$$\inf_{0\leq t\leq 1}(2(1+t^p)I(t))^{1/2} = 2\sqrt{\ln 2}\left(1-\frac{2^{-p-1}}{\ln 2}\right) + \mathcal{O}(p^3 2^{-2p}). \tag{2.27}$$

This, together with the convexity of the free energy in β, will allow us to prove the following statement.

Theorem 2.3: *As $p\to\infty$, the lower bound $\check{\beta}_p \uparrow \hat{\beta}$. Moreover, pointwise in α, β,*

$$\lim_{p\uparrow\infty}\lim_{N\uparrow\infty}\frac{1}{N}\mathbb{E}F_{N,\beta} = f^{REM}_{\beta\alpha^{-1/2}}. \tag{2.28}$$

The basic strategy used to prove these results are rather general. In Section 2.2, we will explain them by means of the analogous calculations in the REM. For now, we just mention that the hard part is to prove the lower bound (2.25), whereas the upper bound (2.26) is comparatively easy and will follow from an estimate on the ground state energy.

In the case of the Hamiltonian $\bar{H}$, we get the following bounds.

Theorem 2.2': *The critical value $\bar{\beta}_p = \bar{\beta}_p(\alpha)$ satisfies*

$$\bar{\beta}_p(\alpha)^2 \geq \min\left(\frac{\bar{\beta}_p'^{\,2}}{4}, \inf_{t\in[0,1]} I(t)\frac{1+g_p(t)}{\alpha g_p(t)}\right) \equiv \check{\beta}_p(\alpha)^2. \tag{2.29}$$

Furthermore, if $\alpha \geq \frac{8\ln 2}{p!} = \alpha_p$ then

$$\beta_p(\alpha)^2 \leq \frac{2\ln 2}{\alpha} \equiv \hat{\beta}_p(\alpha)^2. \tag{2.30}$$

[24] Observe that in [T4], the normalization of the Hamiltonian contains an extra factor $2^{-1/2}$.

If $\alpha \le \alpha_p$, *then* $\beta_p^2 \le \frac{p!}{4}$.

An important point in the study of disordered models is the question of self-averaging of the free energy. The following result settles this matter in the case of the interaction H.

Theorem 2.4: *There exists a constant* C *such that the variance of the free energy satisfies*

$$\mathbb{E}\left[(F_N - \mathbb{E}F_N)^2\right] \le CN^{-1}. \tag{2.31}$$

Furthermore, for $\varepsilon \in (0, \frac{1}{2})$, *and* $\tau > 0$,

$$\mathbb{P}\left[|F_N - \mathbb{E}F_N| \ge \tau N^{-\varepsilon}\right] \le C\tau^{-2}N^{2\varepsilon-1}. \tag{2.32}$$

This shows that, as one would expect, the fluctuations of the free energy are of the order of $N^{-1/2}$. However, the above result is very weak. At least one hopes for a result of the form (1.17). Moreover, from the results in the p-spin SK-model, the REM (for these cases, see [BKL]), and the result below on the second Hamiltonian, one suspects that the fluctuations are of much lower order, at least for small β.

Theorem 2.4': *If* $\beta < \check{\beta}_p(\alpha)$ *(as in (2.29)), then there exist* $C_1, C_2 > 0$ *such that*

$$\mathbb{P}\left[\bar{F}_N \le \frac{1}{N}\ln \mathbb{E}\bar{Z}_N - \frac{u}{N}\right] \le C_1 e^{f(u)}, \tag{2.33}$$

where

$$f(u) = \max\left\{\left(\frac{u - 2\ln 2}{\beta C_2}\right), \frac{u - 2\ln 2}{\beta C_2}\right\}. \tag{2.34}$$

Moreover, for all values of β,

$$\mathbb{P}\left[\bar{F}_N \ge \frac{1}{N}\ln \mathbb{E}\bar{Z}_N + \frac{u}{N}\right] \le e^{-u}. \tag{2.35}$$

The estimate on the fluctuations above the annealed free energy follows immediately from Chebyshev's inequality with first mean, applied to the function $\bar{Z}_N$. This argument obviously holds also for the Hamiltonian H_N.

The following results connects the critical temperature to the behavior of the order parameter $R(\sigma, \sigma')$. The first one might be called an *integration by parts formula* since in the case of Gaussian interactions (the p-spin SK-model, see [T4]), it is an easy consequence of the relation

$$\mathbb{E}[gf(g)] = \mathbb{E}[g^2]\mathbb{E}[f'(g)], \tag{2.36}$$

which holds for any centered Gaussian random variable g and any function f not growing faster than some polynomial at infinity.

Theorem 2.5: *The replica overlap* $R(\sigma, \sigma')$ *satisfies*

$$\beta\mathbb{E}\,\frac{\partial F_N}{\partial \beta} = \alpha\beta^2(1 - \mathbb{E}\mathcal{G}_{N,\beta} \otimes \mathcal{G}_{N,\beta}[R(\sigma, \sigma')^p](1 + \mathcal{O}(N^{-1})), \tag{2.37}$$

Observe that the first term on the right is in fact the derivative of the annealed free energy.[25] We then have the following consequence to Theorem 2.2 and Theorem 2.5.

Theorem 2.6: *If $\beta < \beta_p$, then*

$$\limsup_{N\uparrow\infty} \mathbb{E}\mathcal{G}_{N,\beta} \otimes \mathcal{G}_{N,\beta}[R(\sigma,\sigma')^p] = 0. \tag{2.38}$$

If $\limsup_N \mathbb{E}\,\frac{\partial F_N}{\partial \beta} < \alpha\beta$, *then*

$$\liminf_{N\uparrow\infty} \mathbb{E}\mathcal{G}_{N,\beta} \otimes \mathcal{G}_{N,\beta}[R(\sigma,\sigma')^p] > 0. \tag{2.39}$$

In particular, (2.39) holds for all $\beta \geq \hat{\beta}_p$. Moreover, for every $\varepsilon > 0$ there exists a set $I \subset (\beta_p, \beta_p + \varepsilon)$ of strictly positive Lebesgue measure on which the condition (2.39) is also satisfied.

It seems reasonable that (2.39) holds for all β above the critical β_p, but there is no a priori reason which prohibits a *reentrant phase transition.*[26]

Inequality (2.39) expresses in a strong way that below the critical temperature, the Gibbs measure condensates on a small subset of the configuration space. From the results in the $p = 2$ model, one expects this subset to be close to some of the patterns ξ^μ for small values of α. On the other hand, for large values of α, one can show that this almost never happens:

Theorem 2.7: *Suppose that α satisfies $\alpha\beta_p(\alpha) > (p!)^{-1/2}$. Then there exists a $\delta \in (0, \frac{1}{p})$ and $C > 0$ such that for all N large enough*

$$\mathbb{P}[\arg\sup |H_N(\sigma)| \in \bigcup_{\mu=1}^{M(N)} B_\delta(\xi^\mu)] \leq N^{-1}, \tag{2.40}$$

where $B_\delta(\xi^\mu)$ is the $N\delta$-ball around ξ^μ in the space $\mathbb{R}^N$ with respect to the Hamming metric. In particular, there exists an $\alpha_{sp} = \alpha_{sp}(p)$ such that (2.40) holds for all $\alpha > \alpha_{sp}$.

The proof of this result is based on the comparison between the ground state energy of the system and an estimate on the values of the Hamiltonian in the balls around the patterns. While the former increases as $\sqrt{\alpha}$, the latter is almost constant and with high probability close to $N(p!)^{-1/2}$.

While one cannot show at this point that the system condensates on a set that is close to the extremal value of the Hamiltonian, the above theorem already tells us that this single configuration has exponentially more weight than the balls around the patterns. More precisely, one expects to have different disjoint regions in configuration space (termed *lumps*) which carry the mass of the measure. And these lumps will not be close to the patterns. This means that we are in a region of disordered condensation, or a *spin glass phase.*

To conclude, we state a result for the interaction $\bar{H}_N$ which goes in the same direction as 2.6. However, the proof does not rely on an integration by parts formula similar to (2.37), but is

[25] Our order parameter (the second term on the right) is thus the difference between the derivatives of the two free energies, which from a physical point of view is a quite nice result.

[26] In fact, the set of points where this happens could be countably infinite and even have an accumulation point at β_p. Of course such a pathological behavior is not expected.

based on the extremely small fluctuations in the low temperature region (Theorem 2.4', inequality (2.34)).

Theorem 2.6': *If $\beta < \check{\beta}_p(\alpha)$ (as in (2.29), then there exist constants $\gamma, K > 0$ such that*

$$\mathbb{E}[\mu_{N,\beta} \otimes \mu_{N,\beta}[e^{\gamma N R(\sigma,\sigma')^2}]] \leq K, \tag{2.41}$$

for all $\gamma < \gamma^*$. In particular, this implies that

$$\mathbb{E}[\exp(\mu_{N,\beta} \otimes \mu_{N,\beta}[\gamma N R(\sigma,\sigma')^2])] \leq K. \tag{2.42}$$

The rest of Part I is organized as follows. In Section 2.2, we explain the ideas behind the proof of the bounds on the critical temperature by calculating the corresponding quantities in the REM. In Chapter 3, Theorem 2.1 is proved. Chapter 4 and 5 are devoted to the lower, respectively the upper bound on the critical β (as well as the proof of Corollary 2.3). In Chapter 6 the results on the fluctuations are proved. Chapter 7 deals with the result on the replica overlap (Theorem 2.5 and Corollary 2.6). In the last chapter of this part, we collect the proofs of all results on the second Hamiltonian $\bar{H}_N$. Finally, the Appendix A contains a concentration of measure result which is a slight improvement of a theorem by Ledoux [Le] and is used in the course of the proof of Theorem 2.2.

2.2 Second Moment Method: The REM

In this section, we like to comment upon and explain the methods used to prove the bounds on the critical temperature. Since the calculations are somewhat technical in our case, we illustrate the general idea in the case of the REM. The ideas are the same, but the simple structure of this models allows to understand better the main argument.

The upper bound (2.26) is in fact a rather trivial corollary of the extreme value behavior of the Hamiltonian, that is, the ground state energy, in physicists' terms. The key idea is to bound the supremum of $H_N(\sigma)$ by the supremum of independent random variables. One also knows that the derivative of the free energy with respect to β is equal to the expectation of the Hamiltonian with respect to the associated Gibbs measure. This gives an upper bound on the former quantity, from which one can directly deduce an upper bound on the critical β.

Let us explain this in the case of the REM. Recall that in this model, the Hamiltonian is a i.i.d. random process indexed by the configurations σ, distributed as $\mathcal{N}(0, N)$. By straightforward calculation, one verifies the following identity,[27]

$$\frac{\partial F_N^{\mathrm{REM}}}{\partial \beta} = -\frac{1}{N} \mathcal{G}_{N,\beta}[H_N], \tag{2.43}$$

and thus

$$\mathbb{E}\frac{\partial F_N}{\partial \beta} \leq \frac{1}{N} \mathbb{E}[\sup_\sigma |H_N(\sigma)|]. \tag{2.44}$$

[27]This relation obviously holds for any Hamiltonian. In fact, it is a paradigm of statistical physics that mean values of extensive quantities are obtained by taking the derivative of the free energy with respect to their conjugated intensive quantities.

A well known inequality (verified by integration by parts) then states that

$$\frac{1}{N}\mathbb{E}\,[\sup_\sigma H_N(\sigma)] = \int_0^\infty \mathbb{P}\,[\sup_\sigma |H_N(\sigma)| > Nu]du. \tag{2.45}$$

Estimating the excess probability of the supremum of random variables by the sum of the excess probabilities for each of the variables[28] yields immediately

$$\mathbb{P}\,[\sup_\sigma |H_N(\sigma)| > tN] \le 2^N \mathbb{P}\,[|H_N(\sigma)| > tN] \le 2^{N+1} e^{-\frac{t^2 N}{2}}. \tag{2.46}$$

We use (2.46) in (2.45) for $t \ge \sqrt{2\ln 2}$. For values less than this cutoff, we use the a priori bound $\mathbb{P}\,[\sup_\sigma |H_N(\sigma)| > tN] \le 1$. This implies that

$$\begin{aligned}\frac{1}{N}\mathbb{E}\,[\sup_\sigma H_N(\sigma)] &\le \sqrt{2\ln 2} + 2\int_{\sqrt{2\ln 2}}^{\infty} e^{-N(\frac{t^2}{2}-\ln 2)}dt \\ &\le \sqrt{2\ln 2} + CN^{-1} \equiv \beta' + CN^{-1}.\end{aligned} \tag{2.47}$$

This is the upper bound on the derivative of the expectation of the free energy. Suppose now that $\beta > \sqrt{2\ln 2} = \beta'$. The bounds (2.44) and (2.47) then imply that

$$\mathbb{E}\,F_N(\beta) \le \mathbb{E}\,F_N(\beta') + (\beta - \beta')\beta' + CN^{-1}, \tag{2.48}$$

and in the limit

$$\limsup_{N\uparrow\infty} \mathbb{E}\,F_N(\beta) \le -\frac{\beta'^2}{2} + \beta\beta' = \frac{\beta^2}{2} - \frac{1}{2}(\beta-\beta')^2 < \frac{\beta^2}{2}, \tag{2.49}$$

which by definition means that $\beta' \le \beta_{\mathrm{REM}}$. In the case of the p-spin Hopfield model, the above calculations are identical except for the bounds on the extrema of the Hamiltonian, where the non-Gaussian character induces somewhat more involved calculations.

The basic idea behind Talagrand's approach to prove the lower bound (which he did for the p-spin SK-model in [T4]), is to obtain a variance estimate on the partition function. This will imply that the expectation of the logarithm behaves like the logarithm of the expectation of this quantity. In the REM, one would naively compute

$$\begin{aligned}\mathbb{E}\,[Z_{N,\beta}^{\mathrm{REM}\,2}] &= \mathbb{E}_{\sigma,\sigma'}\,\mathbb{E}\,e^{\beta\sqrt{N}(X_\sigma + X_{\sigma'})} \\ &= 2^{-2N}\left(\sum_{\sigma\neq\sigma'} e^{N\beta^2} + \sum_\sigma e^{2N\beta^2}\right) \\ &= e^{N\beta^2}\left[(1-2^{-N}) + 2^{-N}e^{N\beta^2}\right].\end{aligned} \tag{2.50}$$

[28] Note that in the case of independent Gaussian random variables, even the distribution of the supremum is well known. See [LLR] for an introduction to extreme value theory.

The second term in the brackets is exponentially small if and only if $\beta^2 < \ln 2$, and this cannot be the critical value since it violates the upper bound β' above.[29] The point is that while in the computation of $\mathbb{E} e^{2\beta\sqrt{N}X_\sigma}$, the dominant contribution comes from the part of the distribution of X_σ around $X_\sigma = 2\beta\sqrt{N}$, whereas in $\mathbb{E} Z^{\text{REM}}_{N,\beta}$ the main part is contributed by X_σ around $\beta\sqrt{N}$. One is thus led to consider the second moment of a suitably truncated version of $Z^{\text{REM}}_{N,\beta}$. Namely, for $c > 0$,

$$\tilde{Z}^{\text{REM}}_{N,\beta}(c) = \mathbb{E}_\sigma e^{\beta\sqrt{N}X_\sigma} 1\!\!1_{\{X_\sigma < c\sqrt{N}\}}. \tag{2.51}$$

One then finds that (modulo irrelevant prefactors)

$$\mathbb{E}\,\tilde{Z}^{\text{REM}}_{N,\beta}(c) = \begin{cases} e^{\frac{\beta^2 N}{2}}, & \text{if } \beta < c, \\ \frac{1}{\sqrt{N}(\beta - c)} e^{N\beta c - \frac{Nc^2}{2}}, & \text{if } \beta > c. \end{cases} \tag{2.52}$$

Trivially, for $\beta \leq c$,

$$\mathbb{E}\,\tilde{Z}_{N,\beta}(c) \geq \mathbb{E}\,Z_{N,\beta}\left(1 - e^{-\frac{1}{2}(c-\beta)^2 N}\right) \tag{2.53}$$

On the other hand,

$$\mathbb{E}\,\tilde{Z}_{N,\beta}(c)^2 = (1 - 2^{-N})\left(\mathbb{E}\,\tilde{Z}_{N,\beta}(c)\right)^2 + 2^{-N}\mathbb{E} e^{2\beta\sqrt{N}X_\sigma} 1\!\!1_{\{X_\sigma < c\sqrt{N}\}}, \tag{2.54}$$

where the second term satisfies

$$2^{-N}\mathbb{E} e^{2\beta\sqrt{N}X_\sigma} \leq \begin{cases} 2^{-N} e^{2\beta^2 N}, & \text{if } 2\beta < c \\ 2^{-N} e^{2c\beta N - \frac{c^2 N}{2}}, & \text{otherwise,} \end{cases} \tag{2.55}$$

or

$$2^{-N}\mathbb{E} e^{2\beta\sqrt{N}X_\sigma} 1\!\!1_{\{X_\sigma < (1+\varepsilon)\beta\sqrt{N}\}} \leq (\mathbb{E}\,\tilde{Z}_{N,\beta})^2 \times \begin{cases} e^{-N(\ln 2 - \beta^2)}, & \beta < \frac{c}{2}, \\ e^{-N(c-\beta)^2 - N(\ln 2 - \frac{c^2}{2})}, & \frac{c}{2} < \beta < c. \end{cases} \tag{2.56}$$

Hence, for all $c < \sqrt{2\ln 2}$, and all $\beta \neq c$

$$\mathbb{E}\,\frac{(\tilde{Z}_{N,\beta}(c) - \mathbb{E}\,\tilde{Z}_{N,\beta}(c))^2}{\mathbb{E}[\tilde{Z}_{N,\beta}(c)^2]} \leq e^{-Ng(c,\beta)}, \tag{2.57}$$

where $g(c,\beta) > 0$. Thus, by Chebyshev's inequality, it is immediate that

$$\lim_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\ln\tilde{Z}_{N,\beta}(c) = \lim_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{E}\,\tilde{Z}_{N,\beta}(c), \quad \forall c < \sqrt{2\ln 2}. \tag{2.58}$$

Since this gives a lower bound of the free energy that is as close to the upper bound as desired, we see that the upper bound gives in fact the true value.

This is a remarkable feature of the REM: the expectation of the logarithm of the partition function coincides with the log of the expectation of a suitably truncated partition function. This is clearly rather special to the REM. However, the above method is general enough to provide lower bounds in the far more complicated situations of the p-spin SK-model (see [T4]) and, as we will show, in the p-spin Hopfield model.

[29] This is already contained in [D2]

3 Annealed Free Energy

In this chapter, the existence of the free energy for small enough β is proved (Theorem 2.1). As remarked after (2.8) above, $\mathbb{E} Z_{N,\beta} = \mathbb{E} e^{-\beta H_N[\omega](\sigma)}$ and is independent of σ. Hence,

$$\begin{aligned}
\ln \mathbb{E} Z_{N,\beta} = \ln \mathbb{E} e^{-\beta H_N[\omega](\sigma)} &= \ln \mathbb{E} \exp\left(\beta \left(\frac{p!}{N^{2p-2}} \right)^{\frac{1}{2}} \sum_{\mu=1}^{M(N)} \sum_{\substack{\mathcal{I}\subset\mathcal{N} \\ |\mathcal{I}|=p}} \xi_{\mathcal{I}}^{\mu} \right) \\
&= \sum_{\mu=1}^{M(N)} \ln \mathbb{E} \exp\left(\beta \left(\frac{p!}{N^{2p-2}} \right)^{\frac{1}{2}} \sum_{\substack{\mathcal{I}\subset\mathcal{N} \\ |\mathcal{I}|=p}} \xi_{\mathcal{I}}^{\mu} \right) \\
&= M(N) \ln \mathbb{E} \exp\left(\beta \left(\frac{p!}{N^{p-2}} \right)^{\frac{1}{2}} Y \right),
\end{aligned} \tag{3.1}$$

where we introduced the abbreviation $Y \equiv N^{-\frac{p}{2}} \sum_{\mathcal{I}\subset\mathcal{N}} \xi_{\mathcal{I}}^{1}$. We now expand the exponential function according to the bound $\left| e^{x} - 1 - x - \frac{x^2}{2} \right| < |x|^3 e^{|x|}$. Thus,

$$\begin{aligned}
&\left| \mathbb{E} \left[\exp\left(\beta \left(\frac{p!}{N^{p-2}} \right)^{\frac{1}{2}} Y \right) \right] - 1 - \frac{\beta^2 N^{2-p}}{2} \right| \\
&\qquad \leq \mathbb{E} \left[\beta^3 \left(\frac{p!}{N^{p-2}} \right)^{\frac{3}{2}} |Y|^3 \exp\left(\beta \left(\frac{p!}{N^{p-2}} \right)^{\frac{1}{2}} |Y| \right) \right] + \mathcal{O}(N^{1-p}).
\end{aligned} \tag{3.2}$$

Observe that the quadratic term is in fact just the variance of H_N. We will show in a moment that the expectation on the right-hand side of (3.2) is bounded by a constant times $N^{3-\frac{3p}{2}}$. Assuming this and recalling that $p \geq 4$, we get

$$\begin{aligned}
\ln \mathbb{E} Z_N &\leq M(N) \ln(1 + \frac{\beta^2 N^{2-p}}{2} + C_1 N^{1-p} + C_2 N^{3-\frac{3p}{2}}) \\
&\leq M(N) \left(\frac{\beta^2 N^{2-p}}{2} + C_3 N^{1-p} \right) \\
&\leq \frac{\alpha \beta^2 N}{2} (1 + C_4 N^{-1}).
\end{aligned} \tag{3.3}$$

On the other hand, for N large enough,

$$\begin{aligned}
\ln \mathbb{E} Z_N &\geq M(N) \ln(1 + \frac{\beta^2 N^{2-p}}{2} - C_1 N^{1-p} - C_5 N^{3-\frac{3p}{2}}) \\
&\geq M(N) \left(\frac{\beta^2 N^{2-p}}{2} - C_6 N^{1-p} - C_7 N^{4-2p} \right) \\
&\geq \frac{\alpha \beta^2 N}{2} (1 - C_8 N^{-1}).
\end{aligned} \tag{3.4}$$

The bounds (3.3) and (3.4) are the statement of the theorem.

We still have to show that the remainder on the right-hand side of (3.2) is indeed bounded by the claimed value. To to this, we decompose the exponent into two factors, and use on one the obvious bound $|Y| \leq (p!)^{-1} N^{p/2}$. This yields

$$\begin{aligned} \mathbb{E}\left[|Y|^3 \exp\left(\beta (p!)^{\frac{1}{2}} N^{\frac{2-p}{2}} |Y|\right)\right] &= \mathbb{E}\left[|Y|^3 \exp\left(\beta (p!)^{\frac{1}{2}} N^{\frac{2-p}{2}} |Y|^{\frac{2}{p}} |Y|^{\frac{p-2}{p}}\right)\right] \\ &\leq \mathbb{E}\left[|Y|^3 \exp\left(\beta (p!)^{\frac{2}{p}-\frac{1}{2}} |Y|^{\frac{2}{p}}\right)\right]. \end{aligned} \tag{3.5}$$

The term $|Y|^{2/p}$ behaves like the square of a Gaussian. More precisely, we have the following bound.

Lemma 3.1: *Let $\{X_i\}_{i=1,\dots,N}$ be a sequence of i.i.d. Bernoulli variables, taking values $+1$, -1 with equal probability. Then $\forall C \in (0, e^{-p})$ there exists an ε'_C (depending also on p) and an $\bar{N} \in \mathbb{N}$ such that for all $\varepsilon > \varepsilon'_C$*

$$\mathbb{P}\left[\left| N^{-p/2} \sum_{\substack{\mathcal{I} \subset \mathcal{N} \\ |\mathcal{I}|=p}} \prod_{l \in \mathcal{I}} X_l \right| > \varepsilon\right] \leq 2 \exp\left(-C^{2/p} \frac{(p!)^{\frac{2}{p}} \varepsilon^{\frac{2}{p}}}{2}\right). \tag{3.6}$$

Proof: We shall show that $\sum_{\mathcal{I} \subset \mathcal{N}} X_{\mathcal{I}}$ is a function of $\sum_{i \in \mathcal{N}} X_i$ only. Since the distribution of this latter random variable is well known, all we have to do is to find an accurate upper bound for the function relating the two quantities. And since we are only interested in the tail behavior, we can restrict our attention to large values of the sum (large meaning at least of the order of $\sqrt{N}$).

Suppose that $\sum_{i \in \mathcal{N}} X_i = N - 2l$. Then the quantity $\sum_{\mathcal{I}} X_{\mathcal{I}}$ is given by

$$\sum_{\substack{\mathcal{I} \subset \mathcal{N} \\ |\mathcal{I}|=p}} X_{\mathcal{I}} = \sum_{k=0}^{p} (-1)^k \binom{l}{k} \binom{N-l}{p-k} = \lfloor z^p \rfloor \left[(1+z)^{N-l} (1-z)^l\right], \tag{3.7}$$

where $\lfloor z^p \rfloor (\cdot) \equiv \frac{1}{p!} \frac{\partial^p}{\partial z^p} \cdot \Big|_{z=0}$ is the operator which extracts the coefficient of the term z^p from a formal power series. To evaluate this coefficient, we consider the polynomial on the right-hand side of (3.7) as a function from $\mathbb{C} \to \mathbb{C}$ (by definition, it is analytic). Then, by Cauchy's integral formula

$$\lfloor z^p \rfloor \left[(1+z)^{N-l} (1-z)^l\right] = \frac{1}{2\pi i} \oint_{\mathcal{C}} z^{-p-1} (1+z)^{N-l} (1-z)^l \, dz, \tag{3.8}$$

for any closed path $\mathcal{C}$ surrounding the origin counterclockwise. To evaluate this integral, we apply the well known saddle point method (see for instance [CH]). We choose $\mathcal{C}$ to be a circle around the origin with radius

$$r = \frac{N-2l}{2(N-p)} \left(1 - \sqrt{1 - \frac{4p(N-p)}{(N-2l)^2}}\right). \tag{3.9}$$

Suppose that $\frac{4p(N-p)}{(N-2l)^2} < \kappa < 1$. Then the argument of the square root is positive. Moreover, the following bounds for r hold,

$$\frac{p}{N-2l} \leq r \leq \frac{p}{N-2l} (1 + C_1(\kappa)), \tag{3.10}$$

where C_1 increases from zero to some finite constant as κ varies from zero to 1.

Indeed, $\sqrt{1-x}$ is C^∞ for all $|x|<1$. Therefore, for all $\kappa<1$, we can find a $C>0$ such that $\sqrt{1-x}\geq 1-\frac{x}{2}-Cx^2$, for all $|x|<\kappa$. Obviously, C tends to $\frac{1}{8}$ as κ tends to zero. This implies the upper bound. On the other hand, $\sqrt{1-x}\leq 1-\frac{x}{2}$, for all $x\geq -1$, which yields the lower bound.

The contour integral in (3.8) then becomes

$$\begin{aligned} I \equiv & \frac{1}{2\pi i}\oint_C z^{-p-1}(1+z)^{N-l}(1-z)^l\,dz \\ = & \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(-ip\vartheta\ln r+(N-l)\ln(1+re^{i\vartheta})+l\ln(1-re^{i\vartheta})\right)\,d\vartheta \\ = & \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{g(\vartheta)}\,d\vartheta. \end{aligned} \tag{3.11}$$

As usual, we expand the function g around its maximum (which happens to lie at $\vartheta=0$) and try to control the error. This yields

$$\begin{aligned} I &= \exp\left(g(0)+\frac{(2\pi)^3}{3!}\sup_{\zeta\in[-\pi,\pi)} g^{(3)}(\zeta)\right)\int_{-\pi}^{\pi} e^{\frac{\vartheta^2}{2}g^{(2)}(0)}\,d\vartheta \\ &= r^{-p}(1+r)^{N-l}(1-r)^l\exp\left(\frac{(2\pi)^3}{3!}\sup_{\zeta\in[-\pi,\pi)} g^{(3)}(\zeta)\right)\int_{-\pi}^{\pi} e^{\frac{\vartheta^2}{2}g^{(2)}(0)}\,d\vartheta \end{aligned} \tag{3.12}$$

The main contribution comes from the term $r^{-p}(1+r)^{N-l}(1-r)^l$. Using (3.10), this is bounded by

$$\begin{aligned} r^{-p}(1+r)^{N-l}(1-r)^l &= \exp\left(-p\ln r+(N-l)\ln(1+r)+l\ln(1-r)\right) \\ &\leq \exp\left(-p\ln p+p\ln(N-2l)+(N-l)r-lr\right) \\ &\leq \exp\left(-p\ln p+p\ln(N-2l)+(N-2l)r\right) \\ &\leq \frac{(N-2l)^p}{p!}\sqrt{p}e^{C_1(\kappa)p}. \end{aligned} \tag{3.13}$$

The integral in (3.12) is explicitly

$$\begin{aligned} \int_{-\pi}^{\pi} e^{\frac{\vartheta^2}{2}g^{(2)}(0)}\,d\vartheta &\leq \int_{\mathbb{R}} \exp\left(\frac{\vartheta^2}{2}\left(\frac{lr}{(1-r)^2}-\frac{(N-l)r}{(1+r)^2}\right)\right)\,d\vartheta \\ &= \left(\frac{\pi}{\frac{(N-l)r}{(1+r)^2}-\frac{lr}{(1-r)^2}}\right)^{1/2}, \end{aligned} \tag{3.14}$$

and can be bounded by (for all N large enough)

$$\left((N-l)\frac{r}{(1+r)^2}-l\frac{r}{(1-r)^2}\right)^{-\frac{1}{2}} \leq p^{-\frac{1}{2}}\left(1-\frac{p^2}{(N-2l)^2}\right)\left(1-\frac{2\kappa}{3})\right). \tag{3.15}$$

Finally, we estimate the error due to the remainder in the Taylor expansion in (3.12). One shows by a straightforward computation that for all $\kappa, \delta > 0$ there exists an $\bar{N}_{\kappa,\delta} \in \mathbb{N}$ such that

$$|g^{(3)}(\vartheta)| \leq p(1 + C_1(\kappa))\,(1 + \kappa(1 + C_1(\kappa)) + \delta) = pC_3(\kappa, \delta), \tag{3.16}$$

where $C_3 = 1$ for $\kappa = \delta = 0$. Hence, the error committed can be bounded as (if $N > \bar{N}_{\kappa,\delta}$)

$$\begin{aligned} \exp\left(\frac{(2\pi)^3}{3!} \sup_{\zeta \in [-\pi,\pi)} g^{(3)}(\zeta)\right) \\ \leq \exp\left(\frac{2\pi}{3!} p(1 + C_1(\kappa))\left(1 + \kappa(1 + C_1(\kappa)) + \frac{C_2}{N - 2l}\right)\right). \end{aligned} \tag{3.17}$$

This follows from the exact expression for $g^{(3)}$,

$$g^{(3)}(\vartheta) = ire^{i\vartheta}\left((N - l)\frac{re^{i\vartheta} - 1}{(1 + re^{i\vartheta})^3} - l\frac{1 + re^{i\vartheta}}{(re^{i\vartheta} - 1)^3}\right), \tag{3.18}$$

which one gets through straightforward derivation.[30]

Inserting the bounds (3.13), (3.15), and (3.16) into the estimate (3.12) then gives

$$I \leq \frac{(N - 2l)^p}{p!} e^{(C_1(\kappa) + C_3(\kappa,\delta))p}, \tag{3.19}$$

and thus

$$f(\sum_{i \in \mathcal{N}} X_i) \leq \frac{1}{p!} e^{(C_1(\kappa) + C_3(\kappa,\delta))p} (\sum_{i \in \mathcal{I}} X_i)^p, \quad N \geq \bar{N}_{\kappa,\delta} \tag{3.20}$$

Let $\rho(\kappa, \delta) = e^{(C_1(\kappa) + C_3(\kappa,\delta))p}$, for $\kappa \in (0, 1)$ and $\delta > 0$. Then ρ is increasing in κ and bounded below by e^p. Thus, for all $C \in (0, e^{-p})$, we can find $\tilde{\kappa} \in (0, 1)$ and $\tilde{\delta} > 0$ such that $C \leq \rho(\tilde{\kappa}, \tilde{\delta})^{-1}$. Let now

$$\varepsilon_{\kappa,\delta} \equiv \left(\frac{4p}{\kappa}\right)^{p/2} \frac{\rho(\kappa, \delta)}{p!}. \tag{3.21}$$

Suppose that $\varepsilon > \varepsilon_{\tilde{\kappa},\tilde{\delta}}$ and $N \geq \bar{N}_{\tilde{\kappa},\tilde{\delta}}$. Then, we have that

$$\mathbb{P}\left[N^{-1/2} \sum_{i \in \mathcal{N}} X_i > \left(\varepsilon p!\, \rho(\tilde{\kappa}, \tilde{\delta})^{-1}\right)^{1/p}\right] \leq \exp\left(-\frac{1}{2}\left(\varepsilon p!\, \rho(\tilde{\kappa}, \tilde{\delta})^{-1}\right)^{2/p}\right), \tag{3.22}$$

by the standard bound on sums of Bernoulli variables. On the other hand, since

$$N^{-1/2} \sum_{i \in \mathcal{N}} X_i > \left(\varepsilon p!\, \rho(\tilde{\kappa}, \tilde{\delta})^{-1}\right)^{1/p} > \left(\varepsilon_{\tilde{\kappa},\tilde{\delta}} p!\, \rho(\tilde{\kappa}, \tilde{\delta})^{-1}\right)^{1/p} = \left(\frac{4p}{\tilde{\kappa}}\right)^{1/2} \tag{3.23}$$

implies that

$$\frac{4pN}{(N - 2l)^2} < \tilde{\kappa} < 1, \tag{3.24}$$

[30] Or alternatively with Mathematica.

the condition following (3.9) is satisfied and hence the above bound on $f(\sum_{i\in\mathcal{N}} X_i)$ is valid. Thus

$$\begin{aligned}\mathbb{P}\left[N^{-1/2}\sum_{i\in\mathcal{N}} X_i > \left(\varepsilon p!\,\rho(\tilde\kappa,\tilde\delta)^{-1}\right)^{1/p}\right] &= \mathbb{P}\left[N^{-p/2}\frac{\rho(\tilde\kappa,\tilde\delta)}{p!}\left(\sum_{i\in\mathcal{N}} X_i\right)^p > \varepsilon\right] \\ &\geq \mathbb{P}\left[N^{-p/2} f(\sum_{i\in\mathcal{N}} X_i) > \varepsilon\right].\end{aligned} \tag{3.25}$$

Hence, by (3.22) and (3.25),

$$\mathbb{P}\left[N^{-p/2} f(\sum_{i\in\mathcal{I}} X_i) > \varepsilon\right] \leq \exp\left(-\frac{1}{2}\left(\varepsilon p!\,\rho(\tilde\kappa,\tilde\delta)^{-1}\right)^{2/p}\right) \leq \exp\left(-\frac{C^{2/p}}{2}(\varepsilon p!\,)^{2/p}\right). \tag{3.26}$$

Thus, we have shown that for all $C\in(0,e^{-p})$, there exists $\tilde\varepsilon_C = \varepsilon_{\tilde\kappa,\tilde\delta}$ such that (3.26) holds for all $\varepsilon > \tilde\varepsilon_C$ and all N large enough. Together with the analogue bound for the negative tails, this proves the lemma. □

To finish the proof of the theorem, let us go back to (3.5). To get the claimed bound, it is enough to show that the integral on the right-hand side is bounded uniformly in N. Indeed, since the variable Y satisfies the bound (3.6) of the lemma, we get for any $C' < e^{-p}$

$$\begin{aligned}\mathbb{E}\left[|Y|^3\exp\left(\beta(p!)^{\frac{2}{p}-\frac{1}{2}}|Y|^{\frac{2}{p}}\right)\right] &\leq \sum_{l\geq 1}\mathbb{E}\left[|Y|^3 \mathbb{1}_{\{|Y|\in[l,l+1)\}}\exp\left(\beta(p!)^{\frac{2}{p}-\frac{1}{2}}|Y|^{\frac{2}{p}}\right)\right] \\ &\leq (l+1)^3\mathbb{P}[|Y|\geq l]\exp\left(\beta(p!)^{\frac{2}{p}-\frac{1}{2}}(l+1)^{\frac{2}{p}}\right) \\ &\leq \int_0^\infty (x+1)^3\exp\left(\beta(p!)^{\frac{2}{p}-\frac{1}{2}}(x+1)^{\frac{2}{p}} - C'^{2/p}(p!)^{\frac{2}{p}}x^{\frac{2}{p}}\right)dx \\ &\quad + (\tilde\varepsilon_{p,C'}+1)^3.\end{aligned} \tag{3.27}$$

By the preceding lemma, for any $\beta \leq e^{-2}(p!)^{\frac{1}{2}}$, we can find $C' < e^{-p}$ and a corresponding $\varepsilon'_{C'}$ such that the above integral is finite. This proves the theorem. □

We observe that we could have equally well replaced H_N by in $-H_N$ in the proof of Theorem 2.1, without changing the result (since only the square of the Hamiltonian does enter). We therefore have readily the following result, which we state for further use.

Corollary 3.2: *If $\beta < \beta'_p$, then*

$$\mathbb{E}\mathbb{E}_\sigma e^{\beta H_N} = e^{\frac{\alpha\beta^2 N}{2}(1+\mathcal{O}(N^{-1}))}. \tag{3.28}$$

Proof: Completely analogous to the proof of Theorem 2.1. □

4 Critical β and Convergence to the REM

4.1 Estimates on the Truncated Partition Function

We start with a lemma which will be applied frequently.

Lemma 4.1: *If $\beta < \frac{1}{2}\beta'_p$, then there exists a constant $C > 0$ such that*

$$\mathbb{E}\left[e^{-\beta H_N(\sigma)-\beta H_N(\sigma')}\right] \leq e^{\alpha N \beta^2 (1+R(\sigma,\sigma')^p+C)}, \tag{4.1}$$

for all N large enough.

Proof: The proof is actually almost identical to the proof of Theorem 2.1. We start by expanding the exponential up to order two, with the same error as in the proof of Theorem 2.1 (inequality (3.2)). This error term is then treated similarly, by first decoupling the terms in σ and σ' with Cauchy-Schwarz. This already shows why β has to be less than half the bound of Theorem 2.1. The linear term in the expansion vanishes, whereas the quadratic term gives us the covariance term $R(\sigma,\sigma')^p$. Indeed, if we set $Y^\mu(\sigma) = N^{-p/2}\sum_{\mathcal{I}\subset\mathcal{N}} \xi^\mu_{\mathcal{I}}\sigma_{\mathcal{I}}$, we get

$$\begin{aligned}
\ln \mathbb{E}\Big[&\exp(-\beta H_N(\sigma) - \beta H_N(\sigma'))\Big] \\
&\leq \sum_{\mu=1}^{M(N)} \ln\Bigg(1 + \frac{\beta^2 p!}{2} N^{2-p}\mathbb{E}\left[\left(Y^\mu(\sigma)+Y^\mu(\sigma')\right)^2\right] \\
&\qquad + \frac{\beta^3 (p!)^{\frac{3}{2}}}{3} N^{3-\frac{3p}{2}}\mathbb{E}\Big[|Y^\mu(\sigma)+Y^\mu(\sigma')|^3 \\
&\qquad\qquad \times \exp(\beta (p!)^{\frac{1}{2}} N^{1-\frac{p}{2}}|Y^\mu(\sigma)+Y^\mu(\sigma')|)\Big]\Bigg).
\end{aligned} \tag{4.2}$$

We now apply the triangle inequality and Cauchy-Schwarz to the error term, which yields

$$\begin{aligned}
N^{3-\frac{3p}{2}}&\mathbb{E}\left[|Y^\mu(\sigma)+Y^\mu(\sigma')|^3 e^{\beta (p!)^{1/2}N^{1-p/2}|Y^\mu(\sigma)+Y^\mu(\sigma')|}\right] \\
&\leq N^{3-\frac{3p}{2}} \sum_{i=1}^{3} \left(\mathbb{E}\left[|Y^\mu(\sigma)|^{2j}\exp(2\beta(p!)^{\frac{1}{2}}N^{1-\frac{p}{2}}|Y^\mu(\sigma)|)\right]\right)^{\frac{1}{2}} \\
&\qquad \times \left(\mathbb{E}\left[|Y^\mu(\sigma')|^{6-2j}\exp(2\beta(p!)^{\frac{1}{2}}N^{1-\frac{p}{2}}|Y^\mu(\sigma')|)\right]\right)^{\frac{1}{2}} \\
&\leq C_1 N^{3-\frac{3p}{2}},
\end{aligned} \tag{4.3}$$

if $\beta < \frac{1}{2}\beta'_p$ and N large enough, by the result in the proof of Theorem 2.1 (cf. the remark after (3.2)).

The quadratic term in (4.2) is evaluated easily. One obtains (observing that the covariance of H_N appears)

$$\begin{aligned}
\mathbb{E}\left[(-Y^\mu(\sigma) - Y^\mu(\sigma'))^2\right] &= 2\mathbb{E}[Y^\mu(\sigma)^2] + 2\mathbb{E}[Y^\mu(\sigma)Y^\mu(\sigma')] \\
&= 2N^{-p}\binom{N}{p} + 2N^{-p}\sum_{\mathcal{I}\subset\mathcal{N}}\sigma_{\mathcal{I}}\sigma'_{\mathcal{I}} \\
&= \frac{2}{p!}(1+R(\sigma,\sigma')^p) + \mathcal{O}(N^{-1}).
\end{aligned} \tag{4.4}$$

Hence,

$$\begin{aligned} \ln \mathbb{E}\, e^{-\beta(H_N(\sigma)+H_N(\sigma'))} &\leq \sum_{\mu=1}^{M(N)} \ln\left(1+\frac{\beta^2}{N^{p-2}}(1+R(\sigma,\sigma')^p)+\frac{C_2}{N^{p-1}}+\frac{C_1}{N^{\frac{3p}{2}-3}}\right) \\ &\leq M(N)(\beta^2 N^{2-p}(1+R(\sigma,\sigma')^p)+C_3 N^{1-p}), \end{aligned} \tag{4.5}$$

that is,

$$\mathbb{E}\, e^{-\beta H_N(\sigma)-\beta H_N(\sigma')} \leq e^{\alpha\beta^2 N(1+R(\sigma,\sigma')^p)+C_4}. \tag{4.6}$$

This proves the lemma. □

To get the lower bound for the critical temperature, we would like to compare $\mathbb{E}\, Z_{N,\beta}^2$ and $(\mathbb{E}\, Z_{N,\beta})^2$. However, as mentioned in the introduction and explained in Section 2.2, it is essential to do this comparison for a truncated partition function. Define therefore

$$\widetilde{Z}_{N,\beta}(c) \equiv \mathbb{E}_\sigma\left[e^{-\beta H_N[\omega](\sigma)} \mathbb{1}_{\{-H_N \leq c\alpha\beta^2 N\}}\right], \tag{4.7}$$

for $c > 1$. The key observation is that the truncation has no influence on the expectation of the partition function if c is chosen appropriately. This is the content of the following lemma.

Lemma 4.2: *For all $\beta > 0$, $c > 1$ such that $\beta c < \beta_p'$ there exist $K, K' > 0$ such that*

$$\begin{aligned} \mathbb{E}\, \widetilde{Z}_{N,\beta}(c) &\geq \left(1 - K e^{K'(c-1)^2 N}\right) \mathbb{E}\, \mathbb{E}_\sigma \left[e^{-\beta H_N(\sigma)}\right] \\ &= \left(1 - K e^{-K'(c-1)^2 N}\right) \mathbb{E}\, Z_{N,\beta}. \end{aligned} \tag{4.8}$$

Proof: Let us estimate (here and in the rest of the proof, $q = q(N) \equiv \alpha\beta^2 N$),

$$\mathbb{E}\, Z_{N,\beta} - \mathbb{E}\, \widetilde{Z}_{N,\beta} = \mathbb{E}\, \mathbb{E}_\sigma \left[e^{-\beta H_N(\sigma)} \mathbb{1}_{\{-\beta H_N > cq\}}\right]. \tag{4.9}$$

Chebyshev's exponential inequality (applied to the expectation with respect to the disorder) yields

$$\mathbb{E}\, Z_{N,\beta} - \mathbb{E}\, \widetilde{Z}_{N,\beta} \leq \mathbb{E}_\sigma \inf_{t>0} e^{-tcq} \mathbb{E}\left[e^{-\beta(1+t)H_N(\sigma)}\right]. \tag{4.10}$$

We now use Theorem 2.1 with β replaced by $(1+t)\beta$. If $(1+t)\beta < \beta_p'$, we obtain

$$\inf_{t>0} e^{-tcq} \mathbb{E}\left[e^{-\beta(1+t)H_N(\sigma)}\right] \leq \inf_{t>0} e^{-tcq + \frac{(1+t)^2 q}{2} + qCN^{-1}}. \tag{4.11}$$

The exponent is minimized for $t = c-1$. For this value, the above condition is satisfied since we assumed that $\beta c < \beta_p'$, and we get

$$\begin{aligned} \inf_{t>0} e^{-tcq} \mathbb{E}\left[e^{-\beta(1+t)H_N(\sigma)}\right] &\leq e^{-\frac{q}{2}(c-1)^2 + CqN^{-1}} e^{\frac{q}{2}} \\ &\leq e^{-\frac{q}{2}(c-1)^2 + CqN^{-1}} \mathbb{E}\left[e^{-\beta H_N(\sigma)}\right], \end{aligned} \tag{4.12}$$

where the second inequality follows from (2.16) in Theorem 2.1. Plugging (4.12) into (4.9) implies the statement of the lemma. □

We now turn to the square of the truncated partition function. We bound the quantity

$$\mathbb{E} e^{-\beta H(\sigma)-\beta H(\sigma')} \mathbb{1}_{\{-H_N \leq c\alpha\beta N\}} \mathbb{1}_{\{-H_N \leq c\alpha\beta N\}} \tag{4.13}$$

by two different functions. When calculating the expectation with respect to σ and σ', we use one bound for small values of the replica overlap $R(\sigma, \sigma')$, and the other for the rest. Define therefore

$$S(b) \equiv \mathbb{E}_{\sigma,\sigma'} \left[e^{-\beta(H_N(\sigma)+H_N(\sigma'))} \mathbb{1}_{\{|R(\sigma,\sigma')|<b\}} \right] \tag{4.14}$$

and

$$T(c,b,b') \equiv \mathbb{E}_{\sigma,\sigma'} \left[e^{-\beta(H_N(\sigma)+H_N(\sigma'))} \mathbb{1}_{\{|R(\sigma,\sigma')|\in[b,b']\}} \mathbb{1}_{\{-\beta(H_N(\sigma)+H_N(\sigma'))<2c\alpha\beta^2 N\}} \right]. \tag{4.15}$$

Then

$$\widetilde{Z}_{N,\beta}(c)^2 \leq S(b) + T(c,b,1), \tag{4.16}$$

for all $b \in (0,1)$. We now control each of the terms separately. We start with $S(b)$.

Lemma 4.3: *Suppose $\beta < \frac{\beta_p'}{2}$, and b is such that*

$$\gamma \equiv \alpha\beta^2 b^{p-2} < \frac{1}{2}. \tag{4.17))$$

Then for all $\varepsilon \in (0, \frac{1}{2} - \gamma)$ there exists $N_\varepsilon \in \mathbb{N}$ such that for all $N > N_\varepsilon$,

$$\mathbb{E} S(b) \leq \frac{1}{\sqrt{1-2(\gamma+\varepsilon)}} e^{\alpha\beta^2 N}. \tag{4.18}$$

Proof: If β satisfies the above condition, we can apply Lemma 4.1 to the integrand of the right-hand side of (4.14). One obtains

$$\mathbb{E}\left[e^{-\beta(H_N(\sigma)+H_N(\sigma'))} \mathbb{1}_{\{|R(\sigma,\sigma')|<b\}} \right] \leq \mathbb{1}_{\{|R(\sigma,\sigma')|<b\}} e^{\alpha\beta^2 N(1+(R(\sigma,\sigma')^p)+CN^{-1})}. \tag{4.19}$$

Thus,

$$\begin{aligned} \mathbb{E} S(b) &\leq \mathbb{E}_{\sigma,\sigma'} \left[e^{\alpha\beta^2 N(1+R(\sigma,\sigma')^p+CN^{-1})} \mathbb{1}_{\{|R(\sigma,\sigma')|<b\}} \right] \\ &\leq \mathbb{E}_{\sigma,\sigma'} \left[e^{\alpha\beta^2 N(1+R(\sigma,\sigma')^2 b^{p-2}+CN^{-1})} \mathbb{1}_{\{|R(\sigma,\sigma')|<b\}} \right] \\ &= e^{\alpha\beta^2 N} \mathbb{E}_{\sigma,\sigma'} \left[e^{\alpha\beta^2 N(R(\sigma,\sigma')^2 b^{p-2}+CN^{-1})} \right]. \end{aligned} \tag{4.20}$$

By condition (4.17), for all $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that

$$\alpha\beta^2 N\big(R(\sigma,\sigma')^2 b^{p-2} + CN^{-1}\big) < (\gamma+\varepsilon) N R(\sigma,\sigma')^2, \tag{4.21}$$

for all $N \geq N_\varepsilon$.

We now use a convenient identity for Gaussian variables, which goes under the name of Hubbard-Stratonovich transform. In fact, for a standard normal random variable g (i.e. g is distributed with density $(2\pi)^{-1/2}\exp(-x^2/2)$ with respect to Lebesgue measure on $\mathbb{R}$), one has $e^{\frac{a^2}{2}} = \mathbb{E}\, e^{ga}$. This identity, together with inequality (4.21), allows us to bound the second factor on the right-hand side of (4.20) by

$$\begin{aligned}\mathbb{E}_{\sigma,\sigma'}\big[\exp(\alpha\beta^2 N R^2(\sigma,\sigma') b^{p-2}(1+CN^{-1}))\big] &\leq \mathbb{E}_{\sigma,\sigma'}\big[\exp((\gamma+\varepsilon)N^{-1}(\sum_{i=1}^N \sigma_i\sigma_i')^2)\big] \\ &= \mathbb{E}_{\sigma,\sigma'}\mathbb{E}_g\big[\exp(g\sqrt{\frac{2(\gamma+\varepsilon)}{N}}(\sum_{i=1}^N \sigma_i\sigma_i'))\big]\end{aligned} \tag{4.22}$$

We now use the fact that $\sum_i \sigma_i\sigma_i'$ and $\sum_i \sigma_i$ have the same distribution, which leads to

$$\begin{aligned}\mathbb{E}_g\mathbb{E}_{\sigma,\sigma'}\big[\exp(g\sqrt{\frac{2(\gamma+\varepsilon)}{N}}(\sum_{i=1}^N \sigma_i\sigma_i'))\big] &= \mathbb{E}_g\mathbb{E}_{\sigma}\big[\exp(g\sqrt{\frac{2(\gamma+\varepsilon)}{N}}(\sum_{i=1}^N \sigma_i))\big] \\ &= \mathbb{E}_g\Big[\big(\cosh(g\sqrt{\frac{2(\gamma+\varepsilon)}{N}})\big)^N\Big].\end{aligned} \tag{4.23}$$

As is easily checked by series expansion, $\cosh x \leq \exp\frac{x^2}{2}$, for all real x. Thus, for all $\varepsilon < \frac{1}{2} - \gamma$,

$$\begin{aligned}\mathbb{E}_g\Big[\big(\cosh(g\sqrt{\frac{2(\gamma+\varepsilon)}{N}})\big)^N\Big] &\leq \mathbb{E}_g\big[\exp(g^2(\gamma+\varepsilon))\big] \\ &= \sqrt{\frac{1}{1-2(\gamma+\varepsilon)}}.\end{aligned} \tag{4.24}$$

This proves the lemma. □

The next result concerns the term $T(c,b,1)$ in (4.16).

Lemma 4.4: *Let $I(t)$ be the Cramèr Entropy as defined in (2.24). Suppose that there exist $c > 1$, $d > 0$, such that*

$$\forall t \in [b,b'], \quad 2\alpha\beta^2 c\Big(1 - \frac{c}{2(1+t^p)}\Big) \leq \alpha\beta^2 + I(t) - d. \tag{4.25}$$

Then, if

$$c < \min\big(\frac{1}{2\beta}\beta_p', 1+b^p\big), \tag{4.26}$$

there exists $\bar{N} \in \mathbb{N}$ such that for all $N \geq \bar{N}$,

$$\mathbb{E}T(c,b,b') \leq e^{\alpha\beta^2 N} e^{-\frac{Nd}{2}}. \tag{4.27}$$

Proof: By definition,

$$\mathbb{E}T(c,b,b') = \mathbb{E}_{\sigma,\sigma'}\mathbb{E}\left[e^{-\beta(H(\sigma)+H(\sigma'))}\mathbb{1}_{\{|R(\sigma,\sigma')|\in[b,b']\}}\mathbb{1}_{\{-\beta(H_N(\sigma)+H_N(\sigma'))\le 2c\alpha\beta^2N\}}\right]. \tag{4.28}$$

In a first step, we bound the expectation over the disorder. Chebyshev's inequality, applied to the cut-off of the Hamiltonian (let again $q\equiv\alpha\beta^2N$), yields

$$\begin{aligned}\mathbb{E}&\left[\mathbb{1}_{\{-\beta(H_N(\sigma)+H_N(\sigma'))\le 2cq\}}e^{-\beta(H_N(\sigma)+H_N(\sigma'))}\right]\\ &\qquad\le\inf_{t>0}e^{2tcq}\mathbb{E}\left[e^{-\beta(1-t)(H_N(\sigma)+H_N(\sigma'))}\right]\end{aligned} \tag{4.29}$$

To evaluate the expectation, we now use the the bound (4.1) from Lemma 4.1, with β replaced by $\beta(1-t)$. This gives, if $\beta(1-t)<\frac{1}{2}\beta'_p$,

$$\begin{aligned}\inf_{t>0}e^{2tcq}\mathbb{E}\left[e^{-\beta(1-t)(H_N(\sigma)+H_N(\sigma'))}\right]&\le\inf_{t>0}e^{2tcq}e^{(1-t)^2q(1+R(\sigma,\sigma')^p+C_1N^{-1})}\\ &\le\inf_{t>0}e^{2tcq}e^{(1-t)^2q(1+R(\sigma,\sigma')^p)}e^{C_2N^{-1}q}.\end{aligned} \tag{4.30}$$

The infimum is attained for $(1-t)(1+R^p)=c$. The condition preceding (4.30) is then equivalent to $\beta c(1+R^p)^{-1}<\frac{1}{2}\beta'_p$, which is always satisfied by the hypothesis (this is the first term on the right-hand side in (4.26)). Moreover, the minimizing t has to be positive, which is assured by the second term in (4.26). Inserting this value into (4.30), leads to

$$\begin{aligned}\mathbb{E}&\left[e^{-\beta H_N(\sigma)-\beta H_N(\sigma')}\mathbb{1}_{\{-\beta H_N(\sigma)-\beta H_N(\sigma')\le 2c\alpha\beta N\}}\right]\\ &\qquad\le C_3\exp\left(2c\alpha\beta^2N\Big(1-\frac{c}{2(1+R(\sigma,\sigma')^p)}\Big)\right).\end{aligned} \tag{4.31}$$

Finally, we integrate over all configurations σ, σ' satisfying $|R(\sigma,\sigma')|\in[b,b']$. We observe that $R(\sigma,\sigma')$ has the same distribution as $S(\sigma)=N^{-1}\sum_{i=1}^N\sigma_i$. Hence,

$$\begin{aligned}\mathbb{E}\Big[T(c,b,b')\Big]&\le C_3\mathbb{E}_{\sigma,\sigma'}\left[\exp\left(2c\alpha\beta^2\Big(1-\frac{c}{2(1+R(\sigma,\sigma')^p)}\Big)\right)\mathbb{1}_{\{|R(\sigma,\sigma')|\in[b,b']\}}\right]\\ &=C_3\mathbb{E}_\sigma\left[\exp\left(2c\alpha\beta^2N\Big(1-\frac{c}{2(1+S(\sigma)^p)}\Big)\mathbb{1}_{\{|S(\sigma)|\in[b,b']\}}\right]\right.\\ &\le 2C_3N\exp\left(N\sup_{t\in[b,b']}\left[2\alpha\beta^2c\Big(1-\frac{c}{2(1+S(\sigma)^p)}\Big)-I(t)\right]\right)\\ &\le C_4e^{N(\alpha\beta^2(1+\frac{\ln N}{N})-d)}\le e^{N(\alpha\beta^2-\frac{d}{2})}.\end{aligned} \tag{4.32}$$

The second to last inequality follows from the hypothesis of the lemma, and the observation that we sum over at most $2N$ values of $S(\sigma)$. The last one is valid for all N larger than a certain $\bar{N}\in\mathbb{N}$. $\square$

From the preceding results, we now get a variance estimate for the truncated partition function.

Proposition 4.5: *Suppose that $\beta < \check{\beta}_p$. Then there exist constants $C > 0$ and $c > 1$ such that*

$$\mathbb{E}[\tilde{Z}_{N,\beta}(c)^2] \leq C(\mathbb{E}\tilde{Z}_{N,\beta}(c))^2. \tag{4.33}$$

Furthermore,

$$\mathbb{P}[\tilde{Z}_{N,\beta}(c) > \frac{1}{2}\mathbb{E}\tilde{Z}_{N,\beta}(c)] \geq \frac{3}{4C}. \tag{4.34}$$

Proof: We first prove that the hypothesis implies that the assumptions of Lemmas 4.2–4.4 are satisfied. Consider therefore $\beta < \frac{1}{2}\beta'_p$ such that

$$\beta^2 < \inf_{0 \leq t \leq 1} I(t)\frac{1+t^p}{\alpha t^p}. \tag{4.35}$$

Then it is immediate that

$$2\alpha\beta^2\left(1 - \frac{1}{2(1+t^p)}\right) < \alpha\beta^2 + I(t), \tag{4.36}$$

for all $t \in [0,1]$. By continuity, there exist $c^* > 1$ and $d^* > 0$ such that $\forall c \in (1, c^*)$ and $d \in (0, d^*)$

$$2c\alpha\beta^2\left(1 - \frac{c}{2(1+t^p)}\right) < \alpha\beta^2 + I(t) - d, \quad \forall t \in [0,1]. \tag{4.37}$$

This implies the hypothesis of Lemma 4.4.

We now show that $(\mathbb{E}[\tilde{Z}_N])^2$ is of the order of $\mathbb{E}[\tilde{Z}_N{}^2]$. We start by fixing the free parameters b, b', and c. Choose first b such that $\gamma(b) = \frac{1}{4}$ (or any other constant less than one half). Then choose c such that

$$c < \min\left(c^*, \frac{\beta'_p}{2\beta}, 1 + b^p\right). \tag{4.38}$$

Then the hypotheses of all preceding lemmas are fulfilled. Finally, choose $b' = 1$. By Lemmas 4.3 and 4.4, we then have

$$\mathbb{E}\left[\tilde{Z}_N{}^2\right] \leq \mathbb{E}[S(b) + T(c,b,1)] \leq (C_1 + e^{-Nd/2})e^{\alpha\beta^2 N}. \tag{4.39}$$

The right-hand side is by Theorem 2.1 bounded by

$$(C_1 + e^{-Nd/2})e^{\alpha\beta^2 N} \leq 2C_2\Big(\mathbb{E}[Z_N]\Big)^2, \tag{4.40}$$

which in turn is of the order of $(\mathbb{E}[\tilde{Z}_N])^2$ by Lemma 4.2, so that

$$(C_1 + e^{-Nd/2})e^{\alpha\beta^2 N} \leq C_3\Big(\mathbb{E}[\tilde{Z}_N]\Big)^2. \tag{4.41}$$

Therefore,

$$\mathbb{E}\left[\tilde{Z}_N{}^2\right] \leq C_3\Big(\mathbb{E}[\tilde{Z}_N]\Big)^2. \tag{4.42}$$

The second assertion of the proposition follows from the Paley-Zygmund inequality, which states that for a positive random variable Y and any positive constant g,

$$\mathbb{P}\Big[Y \geq g\mathbb{E}Y\Big] \geq (1-g^2)\frac{(\mathbb{E}Y)^2}{\mathbb{E}[Y^2]}. \tag{4.43}$$

This relation gives us a lower bound on the probability that $\tilde{Z}_N \geq g\mathbb{E}[\tilde{Z}_N]$, which is strictly greater than zero and uniform in N. Indeed, if we set $g = \frac{1}{2}$ in (4.43), then, by (4.42), we get

$$\mathbb{P}\big[\tilde{Z}_N \geq \frac{1}{2}\mathbb{E}\tilde{Z}_N\big] \geq \frac{3}{4C_3}. \tag{4.44}$$

This concludes the proof of the proposition. □

4.2 Proof of the Lower Bound

We use Proposition 4.5 together with a concentration of measure result to show that the mean of F_N cannot deviate too much from the logarithm of the mean of the partition function.

It follows from the definition, that $\tilde{Z}_N \leq Z_N$. Furthermore, Lemma 4.2 implies that for any $C_1 \in (0,1)$ there exists an N' such that for all $N \geq N'$, $\mathbb{E}\tilde{Z}_N \geq C_1\mathbb{E}Z_N$. Therefore, for any $C_2 > 0$, there exists an $\bar{N}$ such that for all $N \geq \bar{N}$,

$$\mathbb{P}\big[\tilde{Z}_N \geq \frac{1}{2}\mathbb{E}\tilde{Z}_N\big] \leq \mathbb{P}\big[Z_N \geq \frac{1}{2}C_2\mathbb{E}Z_N\big]. \tag{4.45}$$

With (2.16) from Theorem 2.1 and the definition of the free energy, we get

$$\mathbb{P}\big[Z_N \geq \frac{1}{2}C_2\mathbb{E}Z_N\big] \leq \mathbb{P}\big[F_N \geq \frac{\alpha\beta^2}{2} - N^{-1}C_3\big]. \tag{4.46}$$

Fix $k > \frac{1}{2}$, and suppose that $\mathbb{E}F_N < \alpha\beta^2/2 - N^{k-1}$, infinitely often in N. Then the right-hand side of the last inequality is bounded by

$$\begin{aligned}\mathbb{P}\big[\tilde{Z}_N \geq \frac{1}{2}\mathbb{E}\tilde{Z}_N\big] &\leq \mathbb{P}\big[F_N \geq \frac{\alpha\beta^2}{2} - C_2N^{-1}\big] \\ &= \mathbb{P}\big[F_N - \frac{\alpha\beta^2}{2} + N^{k-1} \geq N^{k-1} - C_2N^{-1}\big] \\ &\leq \mathbb{P}\big[F_N - \mathbb{E}F_N \geq N^{k-1} - C_2N^{-1}\big].\end{aligned} \tag{4.47}$$

We will deduce a contradiction to Proposition 4.5 by showing that this latter probability tends to zero with N growing. We will use a deviation inequality, which is proved in the appendix as Corollary A.4.

Let us generalize the space of the disorder. Namely, we consider $\{\xi_i^\mu\}_{i=1,\dots,N}^{\mu=1,\dots M(N)}$ as points in the space $\mathbb{R}^{M(N)\times N}$, equipped with the degenerate measure $(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes M\times N}$. The finite Hamiltonian and the free energy are then also functions on $[-1,1]^{M\times N}$. With abuse of notation, we denote by ω a point in this space, i.e. $\omega = \{\xi_i^\mu\}_{i\in\mathcal{N}}^{\mu\in\mathcal{M}}$.

Theorem 4.6: *Suppose that $G_N : [-1,1]^{L(N)} \to \mathbb{R}$ are smooth positive functions, separately convex, and satisfy the following conditions: there exist constants $c^*, t^*, \kappa, \alpha > 0$ and a $\bar{N} \in \mathbb{N}$ such that for all $N \geq \bar{N}$,*

(i) for all $c > c^*$: $\mathbb{P}[\|G_N\|_{\mathrm{Lip}} > cN^{-1/2}] \leq e^{-\delta(c)N}$, for some increasing function δ;
(ii) for all $t > t^*$: $\mathbb{P}[|G_N| > tN] \leq e^{-\kappa N(t-\alpha)}$;
(iii) the Lipschitz constant as a function of N is uniformly bounded by some polynomial function p of N.

Then for all $t > 0$ and $k \in [\frac{1}{2}, 1)$ there exists constants $C > 0$ and $\bar{N} \in \mathbb{N}$ such that

$$\mathbb{P}[NG_N \geq N \int G_N dP + tN^k] \leq e^{-\frac{N^{2k-1}t^2}{C}}, \tag{4.48}$$

for all N larger than $\bar{N}$.

Proof: The theorem is proved as Corollary A.4 in the appendix. □

All we have to do is to check that the hypotheses of Theorem 4.6 are satisfied for $G_N = F_N$. It is obvious, that the conclusion (4.48) is in contradiction with (4.44), and thus proves the lower bound on the critical temperature.

Condition (iii) is easy to check (using simply that the supremum of H does not exceed a certain power of N), as well as separate convexity in each of the variables ξ_i^μ. Positivity of the free energy is assured by the following lemma.

Lemma 4.7: *Consider $F_{N,\beta}$ as a function on $[-1,1]^{M(N)\times N}$. Then $F_{N,\beta}$ is pair and convex along each straight line passing through the origin. Moreover, $F_{N,\beta}[\omega = 0] = 0$, as well as $\frac{d}{d\beta}F_{N,\beta}[\omega]|_{\beta=0} = 0$, and hence $F_{N,\beta}$ is a non-negative function.*

Proof: Let $\omega \in [-1,1]^{M(N)\times N}$ be fixed. Parametrize the line g through ω and the origin by $\lambda \in \mathbb{R}$. Obviously, $F_{N,\beta}|_g$ is symmetric with respect to the origin and continuous in ω. Thus, without restricting the generality, we may assume $\lambda > 0$. It is easy to check that $F_{N,\beta}[\lambda\omega] = F_{N,\lambda^p\beta}[\omega]$. Moreover $F_{N,\beta}$ is convex in β, since its second derivative with respect to β is the variance of H_N with respect to the Gibbs measure $\mathcal{G}_{N,\beta}$, and thus always non-negative. Finally, λ^p is a monotone, non-negative function of λ and hence $F_{N,\lambda^p\beta}[\omega]$ is convex in λ. The same is true for all $\lambda < 0$, whence the first part of the lemma follows.

The proof that $F_{N,\beta}[\omega = 0] = 0$ and $\frac{d}{d\beta}F_{N,\beta}[\omega]|_{\beta=0} = 0$ is obvious. From this and the fact that $F_{N,\beta}$ is continuous in ω, it follows that $F_{N,\beta}[\omega] \geq 0$, for all $\omega \in [-1,1]^{M(N)\times N}$. □

We now turn to the estimate on the Lipschitz norm of F_N. Condition (iii) being checked, we now show that (i) holds as well.

Lemma 4.8: *There exist constants $c^*, C > 0$, and $\bar{N} \in \mathbb{N}$ such that for all $c > c^*$ and all $N \geq \bar{N}$ there exist sets $\widetilde{\Omega}_N \subset \Omega$ such that $F_N|_{\widetilde{\Omega}_N}$ has Lipschitz constant less than $cN^{-\frac{1}{2}}$ and $\mathbb{P}[\widetilde{\Omega}^c] \leq Ce^{-N(c-c^*)}$.*

Proof: Suppose for the moment that there exists a set $\widetilde{\Omega}_N$ such that the *Hamiltonian* restricted to this set has Lipschitz constant less than $c_1\sqrt{N}$. Then it is straight- forward that the assertion of the Lemma is true. Indeed, by definition of the free energy, it follows that (denote by $\xi' = \xi[\omega']$)

$$\begin{aligned} \frac{1}{N}\left|F_N[\omega] - F_N[\omega']\right| &= \frac{1}{N}\ln\frac{\mathbb{E}_\sigma\, e^{-\beta H_N[\omega](\sigma)}}{\mathbb{E}_\sigma\, e^{-\beta H_N[\omega'](\sigma)}} \\ &\leq \frac{1}{N}\ln\frac{\mathbb{E}_\sigma\, e^{-\beta H_N[\omega'](\sigma)}e^{\beta(H_N[\omega](\sigma)-H_N[\omega'](\sigma))}}{\mathbb{E}_\sigma e^{-\beta H_N[\omega'](\sigma)}} \\ &\leq \frac{1}{N}\ln\frac{\mathbb{E}_\sigma\, e^{-\beta H_N[\omega'](\sigma)}e^{\beta c_1\sqrt{N}\|\xi-\xi'\|_2}}{\mathbb{E}_\sigma e^{-\beta H_N[\omega'](\sigma)}} = \beta c_1 N^{-\frac{1}{2}}\|\xi-\xi'\|_2, \end{aligned} \tag{4.49}$$

uniformly in ω on $\widetilde{\Omega}_N$, which is the statement of the Lemma.

To prove that the Hamiltonian is indeed Lipschitz on a large set, we proceed as follows. Define the *overlap parameters* by $m_N^\mu[o](\sigma) \equiv \frac{1}{N}\sum_{i=1}^N \xi_i^\mu\sigma_i$. Then one shows by induction over p that the Hamiltonian can be expressed as

$$\begin{aligned} H_N[\omega](\sigma) &= \frac{(p!)^{1/2}}{N^{p-1}}\sum_{\mu=1}^{M(N)}\sum_{\substack{\mathcal{I}\subset\mathcal{N}\\|\mathcal{I}|=p}}\xi_{\mathcal{I}}^\mu\sigma_{\mathcal{I}} \\ &= N\sum_{k=1}^{p/2}C'_{p,k,N}N^{(2k-p)/2}\sum_{\mu=1}^{M(N)}|m_N^\mu[\omega](\sigma)|^{2k} + C'_{p,N}, \end{aligned} \tag{4.50}$$

where the numbers $C'_{p,k,N}$ are almost constant in the sense that $C'_{p,k,N} = C'_{p,k}(1+\mathcal{O}(N^{-1}))$ (and similarly for $C'_{p,N}$). Then it is immediate that

$$\begin{aligned} &|H_N[\omega](\sigma) - H_N[\omega'](\sigma)| \\ &\quad\leq N\sum_{k=1}^{p/2}|C_{p,k}|N^{(2k-p)/2}\left|\sum_{\mu=1}^{M(N)}\left(m_N^\mu[\omega](\sigma)^{2k} - m_N^\mu[\omega'](\sigma)^{2k}\right)\right| + C_p, \end{aligned} \tag{4.51}$$

for some constants $C_{p,k}$. It is not too difficult to bound the right-hand side. For every k,

$$\begin{aligned} &N\left|N^{(2k-p)/2}\sum_{\mu=1}^{M(N)}\left(m_N^\mu[\omega](\sigma)^{2k} - m_N^\mu[\omega'](\sigma)^{2k}\right)\right| \\ &\quad= N\left|N^{(2k-p)/2}\sum_{\mu=1}^{M(N)}\left(m_N^\mu[\omega](\sigma) - m_N^\mu[\omega'](\sigma)\right)\right. \\ &\qquad\qquad\left.\times\left(\sum_{j=0}^{2k-1}m_N^\mu[\omega](\sigma)^j m_N^\mu[\omega'](\sigma)^{2k-1-j}\right)\right| \\ &\quad\leq N\sum_{j=0}^{2k-1}C_k\left|\sum_{\mu=1}^{M(N)}\left(m_N^\mu[\omega](\sigma) - m_N^\mu[\omega'](\sigma)\right)^2\right|^{\frac{1}{2}} \\ &\qquad\qquad\times\left|\sum_{\mu=1}^{M(N)}N^{2k-p}m_N^\mu[\omega](\sigma)^{2j}m_N^\mu[\omega'](\sigma)^{4k-2j-2}\right|^{\frac{1}{2}}. \end{aligned} \tag{4.52}$$

The first factor is easy to treat. Indeed, Cauchy-Schwarz yields

$$N\left|\sum_{\mu=1}^{M(N)}\left(m_N^\mu[\omega](\sigma)-m_N^\mu[\omega'](\sigma)\right)^2\right|^{\frac{1}{2}} \leq N^{1/2}\left(\sum_{\mu=1}^{M(N)}\sum_{i=1}^{N}(\xi_i^\mu-\xi_i'^\mu)^2\right)^{\frac{1}{2}} = N^{1/2}\|\xi-\xi'\|_2. \tag{4.53}$$

It is now sufficient to show that on a set of measure exponentially close to one, the second factor in (4.52) (which is an upper bound for the Lipschitz constant of H) is bounded by some constant. In order to symmetrize the terms, we apply the Hölder inequality to each summand in this factor, that is

$$\begin{aligned}&\left|\sum_{\mu=1}^{M(N)} N^{2k-p} m_N^\mu[\omega](\sigma)^{2j} m_N^\mu[\omega'](\sigma)^{4k-2j-2}\right|^{\frac{1}{2}} \\ &\leq \left(N^{2k-p}\sum_{\mu=1}^{M(N)} m_N^\mu[\omega](\sigma)^{4k-2}\right)^{\frac{j}{4k-2}}\left(N^{2k-p}\sum_{\mu=1}^{M(N)} m_N^\mu[\omega'](\sigma)^{4k-2)}\right)^{\frac{2k-1-j}{4k-2}}.\end{aligned} \tag{4.54}$$

Define for each k the set $\Omega_{k,c}$ by

$$\Omega_{k,c} \equiv \left\{\omega\in\Omega : \sup_\sigma N^{2k-p}\sum_{\mu=1}^{M(N)} m_N^\mu[\omega](\sigma)^{4k-2} > c\right\}. \tag{4.55}$$

On the complement of the union of these sets, i.e. on $(\bigcup_k \Omega_{k,c})^c$, the Lipschitz constant of the Hamiltonian is bounded by a constant times $\sqrt{N}$. We now calculate the probability of the sets $\Omega_{k,c}$. By elementary arguments,

$$\mathbb{P}\left[\sup_\sigma N^{2k-p}\sum_{\mu=1}^{M(N)} m_N^\mu[\omega](\sigma)^{4k-2} > c\right] \leq \sum_\sigma \mathbb{P}\left[N^{2k-p}\sum_{\mu=1}^{M(N)} m_N^\mu[\omega](\sigma)^{4k-2} > c\right]. \tag{4.56}$$

To evaluate the latter probability, we use Chebyshev's exponential inequality and then expand again the exponential function. We get

$$\mathbb{P}\left[N^{2k-p}\sum_{\mu=1}^{M(N)} m_N^\mu[\omega](\sigma)^{4k-2} > c\right] \leq \inf_{q>0} e^{-qc}\prod_{\mu=1}^{M(N)}\mathbb{E}\, e^{qN^{2k-p}m_N^\mu(\sigma)^{4k-2}}. \tag{4.57}$$

Now,

$$\begin{aligned}&\left|\mathbb{E}\, e^{qN^{2k-p}m_N^\mu(\sigma)^{4k-2}} - 1 - \mathbb{E}\left[qN^{2k-p}m_N^\mu(\sigma)^{4k-2}\right]\right| \\ &\qquad \leq \frac{q^2N^{4k-2p}}{2}\mathbb{E}\left|m_N^\mu(\sigma)^{8k-4}e^{|qN^{2k-p}m_N^\mu(\sigma)^{4k-2}|}\right|.\end{aligned} \tag{4.58}$$

The term on the left-hand side is easy to treat. With the obvious definition of $s^2_{k,N}$ (which is bounded, and converges to some constant s_k as $N \uparrow \infty$), we get

$$N^{2k-p}\mathbb{E}\left[q m_N^\mu(\sigma)^{4k-2}\right] = qN^{2k-p}N^{1-2k}\mathbb{E}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \xi_i^\mu \sigma_i\right)^{4k-2} \equiv qN^{1-p}s^2_{k,N} \tag{4.59}$$

The remainder (the term on the right-hand side) in (4.58) is also easy to bound. Since we anticipate that q has to be proportional to N to counter the term 2^N coming from the sum over all configurations σ in (4.56), we let $q = \gamma N$. Then, since $|m_N^\mu| < 1$, and for all $\gamma < \frac{N^{p-2k}}{2}$,

$$\begin{aligned} q^2 N^{4k-2p}\mathbb{E}&\left[m_N^\mu(\sigma)^{4(2k-1)} e^{|qN^{2k-p}m_N^\mu(\sigma)^{2(2k-1)}|}\right] \\ &= \gamma^2 N^{4-2p}\mathbb{E}\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \xi_i^\mu\sigma_i\right)^{4(2k-1)}\right. \\ &\qquad\qquad \left. \times \exp\left(\gamma N^{2k-p}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \xi_i^\mu\sigma_i\right)^2 m_N^\mu(\sigma)^{4(k-1)}\right)\right] \\ &\le \gamma^2 N^{4-2p}\mathbb{E}\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \xi_i^\mu\sigma_i\right)^{4(2k-1)}\right. \\ &\qquad\qquad \left. \times \exp\left(\gamma N^{2k-p}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N \xi_i^\mu\sigma_i\right)^2\right)\right] \\ &\le \gamma^2 N^{4-2p}K_{p,k}, \end{aligned} \tag{4.60}$$

where the last line follows from the Gaussian tail behavior of the variable $\frac{1}{\sqrt{N}}\sum_{i=1}^N \xi_i^\mu\sigma_i$. Inserting the previous two bounds in (4.58) gives

$$\ln \mathbb{E}\, e^{qN^{2k-p}m_N^\mu(\sigma)^{2(2k-1)}} \le \ln(1 + \gamma N^{2-p}s^2_{k,N} + K_{p,k}\gamma^2 N^{4-2p}). \tag{4.61}$$

As before, we use the bound $\ln(1+x) \le x$, for all $x \ge 0$. This gives

$$\begin{aligned} \ln \mathbb{E}\, e^{qN^{2k-p}\sum_{\mu=1}^{M(N)} m_N^\mu(\sigma)} &= \sum_{\mu=1}^{M(N)} \ln \mathbb{E}\, e^{qN^{2k-p}m_N^\mu(\sigma)} \\ &\le \alpha\gamma N s^2_{k,N} + \alpha\gamma^2 K_{p,k}N^{3-p}. \end{aligned} \tag{4.62}$$

Using this result in (4.56), we get

$$\begin{aligned} \mathbb{P}&\left[\sup_\sigma N^{2k-p}\sum_{\mu=1}^{M(N)} m_N^\mu[\omega](\sigma)^{4k-2} > c\right] \\ &\qquad\qquad \le 2^N \inf_{\gamma\in(0,\frac{1}{2}N^{p-2k})} e^{-\gamma N c + \alpha\gamma s^2_{k,N}N + \alpha\gamma^2 K_{p,k}N^{3-p}}. \end{aligned} \tag{4.63}$$

Since we want the right-hand side to vanish exponentially in N, we should choose c such that

$$c > \gamma^{-1} \ln 2 + \alpha s_k^2 \geq 2 \ln 2 + \alpha s_k^2 \equiv c^* \tag{4.64}$$

If this condition is satisfied, then for any $\delta > 0$ there exists an $\bar{N} \in \mathbb{N}$ such that

$$\mathbb{P}\left[\sup_\sigma N^{2k-p} \sum_{\mu=1}^{M(N)} m_N^\mu[\omega](\sigma)^{4k-2} > c\right] \leq e^{-N(c-c^*)+\delta} \tag{4.65}$$

for all N larger than $\bar{N}$.

Define therefore

$$\tilde{\Omega}_N \equiv \left(\bigcup_k \Omega_{k,c}\right)^c \tag{4.66}$$

for all $c > c^*$. Then there exist constants $C > 0$, and $\bar{N} \in \mathbb{N}$ such that

$$\mathbb{P}\left[\tilde{\Omega}_N^c\right] \leq C e^{-N(c-c^*)}, \tag{4.67}$$

for all $N \geq \bar{N}$.

On the set $\tilde{\Omega}_N$, the Lipschitz constant of the Hamiltonian is bounded by a constant times $\sqrt{N}$, and therefore, by the derivation at the beginning of this proof, the same bound holds for the free energy. This concludes the proof of Lemma 4.8. □

We now check condition (i) of Theorem 4.6. This is a much simpler task, since we have already calculated almost everything.

Lemma 4.9: *The Hamiltonian satisfies*

$$\mathbb{P}\left[\sup_\sigma |H_N(\sigma)| > tN\right] \leq C \begin{cases} \exp\left(-N(\frac{t^2}{2\alpha} - \ln 2)\right), & \text{if } t \leq \frac{\alpha(p!)^{\frac{1}{2}}}{2}, \\ \exp\left(-N(\frac{(p!)^{\frac{1}{2}}}{2} t - \frac{\alpha p!}{8} - \ln 2)\right), & \text{otherwise.} \end{cases} \tag{4.68}$$

Proof: We start with a crude bound to extract the supremum. Standard arguments and Chebyshev's inequality in its exponential form yield

$$\mathbb{P}[\sup_\sigma |H_N(\sigma)| > tN] \leq 2^N \inf_{q>0} e^{-qtN} \mathbb{E} e^{qH_N(\sigma)} + 2^N \inf_{q>0} e^{-qtN} \mathbb{E} e^{-qH_N(\sigma)}. \tag{4.69}$$

We now use Theorem 2.1, respectively Corollary 3.2 to bound the two integrals and obtain

$$\begin{aligned} \mathbb{P}\left[\sup_\sigma |H_N(\sigma)| > tN\right] &\leq C_1 2^{N+1} \inf_{q \in (0,\beta_p')} e^{-qtN} e^{\frac{\alpha q^2 N}{2}} \\ &= C_2 \begin{cases} \exp\left(-N(\frac{t^2}{2\alpha} - \ln 2)\right), & \text{if } t \leq \frac{\alpha(p!)^{\frac{1}{2}}}{2}, \\ \exp\left(-N(\frac{(p!)^{\frac{1}{2}}}{2} t - \frac{\alpha p!}{8} - \ln 2)\right), & \text{otherwise.} \end{cases} \end{aligned} \tag{4.70}$$

This proves the lemma. □

We are finally able to prove the lower bound on the critical β.

Proof of Theorem 2.2, lower bound: By Proposition 4.5, if $\beta < \check{\beta}_p$,

$$\mathbb{P}\left[\tilde{Z}_N \geq \frac{1}{2}\mathbb{E}\,\tilde{Z}_N\right] \geq C_1. \tag{4.71}$$

On the other hand, by (4.47), if for some $k > \frac{1}{2}$

$$\mathbb{E}\,F_N \leq \frac{\alpha\beta^2}{2} - N^{k-1} \tag{4.72}$$

infinitely often in N, then

$$\mathbb{P}\left[\tilde{Z}_N \geq \frac{1}{2}\mathbb{E}\,\tilde{Z}_N]\right] \leq \mathbb{P}\left[F_N - \mathbb{E}\,F_N \geq N^{k-1} - C_2 N^{-1}\right]. \tag{4.73}$$

By Theorem 4.6, whose hypotheses are assured by Lemmata 4.7–4.9, this latter probability tends to zero exponentially fast in N^{2k-1}. This contradicts (4.71) and we therefore reject the assumption (4.72). This proves (2.25). □

4.3 Upper Bound on the Critical β

The proof of the upper bound in Theorem 2.2 is considerably simpler than the lower bound. It follows essentially from the following observation. The definitions imply that

$$N\frac{\partial F_N[\omega]}{\partial\beta} = \mathcal{G}_{N,\beta}[\omega]\left[-H_N[\omega](\sigma)\right], \tag{4.74}$$

from which we get immediately

$$N\mathbb{E}\,\frac{\partial F_N}{\partial\beta} \leq \mathbb{E}\,\sup_\sigma |H_N(\sigma)|. \tag{4.75}$$

Suppose we knew a uniform upper bound for the last expression, and thus on $\limsup_N \mathbb{E}\,\frac{\partial F_N}{\partial\beta}$. Then we only had to find the greatest value of β, for which the derivative of the annealed free energy is still lower than this bound. Let us therefore estimate $\mathbb{E}\,\sup_\sigma |H_N(\sigma)|$.

Lemma 4.10: *If $\alpha \geq \frac{8\ln 2}{p!}$, the right-hand side of (4.75) satisfies*

$$\mathbb{E}\,\sup_\sigma |H_N(\sigma)| \leq N\sqrt{2\alpha\ln 2} + C_1\sqrt{N} + C_2, \tag{4.76}$$

for some positive constant C. If $\alpha \leq \frac{8\ln 2}{p!}$, then

$$\mathbb{E}\,\sup_\sigma |H_N(\sigma)| \leq N\frac{\alpha(p!)^{1/2}}{2} + C. \tag{4.77}$$

Proof: Since $|H_N(\sigma)|$ is a positive random variable, we have

$$\begin{aligned}\mathbb{E}\sup_\sigma |H_N(\sigma)| &= \int_0^\infty \mathbb{P}[\sup_\sigma |H_N(\sigma)| > u]\, du \\ &= N\int_0^\infty \mathbb{P}[\sup_\sigma |H_N(\sigma)| > tN]\, dt\end{aligned} \tag{4.78}$$

The integrand was already estimated in Lemma 4.9. We distinguish three different intervals of t, where we bound this probability by 1, by the Gaussian, and the exponential bound, respectively. Thus,

$$\begin{aligned}\mathbb{E}\sup_\sigma |H_N(\sigma)| \le N\int_0^{t_0} dt + C_1 N\int_{t_0}^{t_1} \exp(-N(\frac{t^2}{2\alpha} - \ln 2))\, dt \\ + C_1 N\int_{t_1}^\infty \exp(-N(\frac{(p!)^{1/2}}{2}t - \frac{\alpha p!}{8} - \ln 2))\, dt,\end{aligned} \tag{4.79}$$

where $t_0 = \sqrt{2\alpha \ln 2}$ and $t_1 = \frac{\alpha(p!)^{1/2}}{2}$. If $\alpha \ge \frac{8\ln 2}{p!}$, then $t_0 \le t_1$, and hence, by standard arguments,

$$\mathbb{E}\sup_\sigma |H_N(\sigma)| \le N\sqrt{2\alpha\ln 2} + C_2\sqrt{N} + C_3. \tag{4.80}$$

If $\alpha \le \frac{8\ln 2}{p!}$, then $t_1 \le t_0$. In this case, the sum on the right-hand side of (4.79) consists of only two terms, and hence

$$\mathbb{E}\sup_\sigma |H_N(\sigma)| \le N\frac{\alpha(p!)^{1/2}}{2} + C_3. \tag{4.81}$$

This proves the lemma. □

Proof of Theorem 2.2, upper bound: The proof of the upper bound for the critical β is now straightforward. Suppose that $\alpha \ge \frac{8\ln 2}{p!}$. Let

$$\beta_\infty = \sqrt{\frac{2\ln 2}{\alpha}}. \tag{4.82}$$

Then for all $\beta \ge 0$, we have by the mean value theorem that

$$\mathbb{E}F_N(\beta) \le \mathbb{E}F_N(\beta_\infty) + (\beta - \beta_\infty)\frac{1}{N}\mathbb{E}\sup_\sigma H_N(\sigma). \tag{4.83}$$

Suppose that $\beta > \beta_\infty$. Then

$$\begin{aligned}
\limsup_{N\uparrow\infty} \mathbb{E} F_N(\beta) &\leq \limsup_{N\uparrow\infty} \mathbb{E} F_N(\beta_\infty) + (\beta-\beta_\infty)\sqrt{2\alpha\ln 2} \\
&\leq \frac{\alpha\beta_\infty^2}{2} + (\beta-\beta_\infty)\alpha\beta_\infty \\
&= \alpha\beta\beta_\infty - \frac{\alpha\beta_\infty^2}{2} \\
&= \frac{\alpha\beta^2}{2} - \frac{1}{2}(\beta-\beta_\infty)^2 \\
&< \frac{\alpha\beta^2}{2},
\end{aligned} \tag{4.84}$$

which by definition means that $\beta_p \leq \beta_\infty$.

If $\alpha \leq \frac{8\ln 2}{p!}$, we proceed as above, but we define

$$\beta_\infty = \sqrt{\frac{p!}{4}}, \tag{4.85}$$

and use the bound (4.77) for the supremum of the Hamiltonian. This concludes the proof of the upper bound. □

4.4 Convergence to the REM: Proof of Theorem 2.3

Suppose that $\beta < \check{\beta}_p$. In this case, $\lim_N F_N^{an} = f_{\beta\alpha^{-1/2}}^{REM}$. Thus, if $\mathbb{E} F_N$ did not converge to $f_{\beta\alpha^{-1/2}}^{REM}$, then in particular it would satisfy the condition (4.72) infinitely often in N. But this has been shown to be false for all $\beta < \check{\beta}_p$. Thus the claim is valid for these values of β. Furthermore, since

$$\lim_{p\uparrow\infty} \check{\beta}_p = \hat{\beta} = \sqrt{\frac{2\ln 2}{\alpha}} \tag{4.86}$$

the assertion is true for all $\beta < \hat{\beta}$.

Assume that $\beta \geq \hat{\beta}$. By convexity,

$$\mathbb{E} F_{N,\beta} \geq \mathbb{E} F_{N,\beta_p} + \alpha\beta_p(\beta-\beta_p), \tag{4.87}$$

and thus

$$\begin{aligned}
\lim_{p\uparrow\infty} \liminf_{N\uparrow\infty} \mathbb{E} F_{N,\beta} &\geq \lim_{p\uparrow\infty} \liminf_{N\uparrow\infty} \mathbb{E} F_{N,\beta} + \alpha \lim_{p\uparrow\infty} \beta_p(\beta-\beta_p) \\
&= \frac{\alpha\hat{\beta}^2}{2} + \alpha\hat{\beta}(\beta-\hat{\beta}) \\
&= \sqrt{2\alpha\ln 2}\beta - \ln 2 = \mathbb{E} f_{\beta\alpha^{-1/2}}^{REM}.
\end{aligned} \tag{4.88}$$

On the other hand, the REM free energy is always an upper bound. Indeed, the third line of (4.84) already gives the desired bound (note that the third line in (4.84) is also valid for $\beta = \hat{\beta}$),

$$\limsup_{N\uparrow\infty} \mathbb{E} F_{N,\beta} \leq \sqrt{2\alpha\ln 2}\beta - \ln 2 \tag{4.89}.$$

This proves the assertion for $\beta \geq \hat{\beta}$, and thus the theorem. □

5 Fluctuations of the Free Energy: Proof of Theorem 2.4

We first introduce some objects, which will be useful in the course of the proof. For $q \in \mathbb{N}$, let V^q be a linear space of dimension $\binom{N}{q}$ with an orthonormal basis $\{\psi_{\mathcal{I}}\}$ indexed by the subsets $\mathcal{I} \subset \{1, \dots, N\}$ of size q. Let $T^q = T^q_N : V^q \to V^q$ be a linear map, given by its matrix representation with respect to $\{\psi_{\mathcal{I}}\}$ by

$$T^q_{\mathcal{I},\mathcal{J}} = \sum_{\mu=1}^{M(N)} \xi^\mu_{\mathcal{I}} \xi^\mu_{\mathcal{J}}. \tag{5.1}$$

The proof of the variance estimate (2.31) is based on Burkholder's inequality for discrete martingales. Define a decreasing sequence of σ-algebras $\{\mathcal{F}_k\}_{k\in\mathbb{N}}$ by

$$\mathcal{F}_k = \sigma\left(\{\xi^\mu_i\}^{\mu\in\mathbb{N}}_{i\geq k}\right). \tag{5.2}$$

This allows to introduce a martingale difference sequence

$$\tilde{F}^k \equiv \mathbb{E}[F|\mathcal{F}_k] - \mathbb{E}[F|\mathcal{F}_{k+1}]. \tag{5.3}$$

Hence, $F - \mathbb{E}F = \sum_{k=1}^N \tilde{F}^k$ and, by the well-known Burkholder inequality

$$\mathbb{E}\left[(F-\mathbb{E}F)^2\right] = \mathbb{E}\left[\left(\sum_{k=1}^N \tilde{F}^k\right)^2\right] \leq \sum_{k=1}^N \mathbb{E}\left[\left(\tilde{F}^k\right)^2\right]. \tag{5.4}$$

It therefore remains to bound the individual terms in the above sum. A conventional strategy (see [PS], [B1]) is to introduce a family of Hamiltonians $\tilde{H}^k(\sigma,u)$, defined by

$$\tilde{H}^k(\sigma,u) = H(\sigma) + (1-u)\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\mu=1}^{M(N)} \sum_{\substack{\mathcal{I}\ni k \\ |\mathcal{I}|=p}} \xi^\mu_{\mathcal{I}} \sigma_{\mathcal{I}}. \tag{5.5}$$

This new Hamiltonian is equal to the original one for $u=1$, and independent of $\{\xi^\mu_k\}^{\mu=1,\dots,M}$ for $u=0$. Let

$$\tilde{Z}^k(u) = \mathbb{E}_\sigma[e^{-\beta\tilde{H}^k(\sigma,u)}], \tag{5.6}$$

and

$$G^k(u) = \frac{1}{N}\ln\tilde{Z}^k(u) - \frac{1}{N}\ln\tilde{Z}^k(0). \tag{5.7}$$

This latter quantity relates to $\tilde{F}^k$ via

$$\tilde{F}^k = \mathbb{E}[G^k(1)|\mathcal{F}_k] - \mathbb{E}[G^k(1)|\mathcal{F}_{k+1}] \tag{5.8}$$

Observe that $G^k(u)$ is convex in u, $G^k(0)=0$, and thus $|G^k(1)| \leq \max(|(G^k)'(1)|, |(G^k)'(0)|)$, where the prime denotes the derivative with respect to u. Moreover, since $\tilde{H}^k(\sigma, u=0)$ does not depend on σ_k, it follows that $(G^k)'(0)=0$, and hence we can use $|G^k(1)|\le|(G^k)'(1)|$. Explicitly, this is

$$|G^k(1)|\le|(G^k)'(1)|=\left|\frac{(p!)^{\frac{1}{2}}}{N^p}\mathcal{G}_{N,\beta}[\omega]\left(\sum_{\mu=1}^{M(N)}\sum_{\substack{\mathcal{I}\ni k\\|\mathcal{I}|=p}}\xi_{\mathcal{I}}^{\mu}\sigma_{\mathcal{I}}\right)\right|. \tag{5.9}$$

Let us now use this in (5.4). We observe that by (5.8), the properties of conditional expectations, Jensen's inequality (see also [B1] and [BGP2]), and (5.9) (in the last line),

$$\begin{aligned}\mathbb{E}[(\tilde{F}^k)^2]&=\mathbb{E}\left[(\mathbb{E}[G^k(1)|\mathcal{F}_k]-\mathbb{E}[G^k(1)|\mathcal{F}_{k+1}])^2\right]\\&=\mathbb{E}\left[(\mathbb{E}[G^k(1)-\mathbb{E}[G^k(1)|\mathcal{F}_{k+1}]|\mathcal{F}_k])^2\right]\\&\le\mathbb{E}\left[\mathbb{E}[(G^k(1)-\mathbb{E}[G^k(1)|\mathcal{F}_{k+1}])^2|\mathcal{F}_k]\right]\\&=\mathbb{E}\left[(G^k(1)-\mathbb{E}[G^k(1)|\mathcal{F}_{k+1}])^2\right]\\&=\mathbb{E}[(G^k(1))^2]-\left(\mathbb{E}[G^k(1)|\mathcal{F}_{k+1}]\right)^2\\&\le\mathbb{E}[(G^k(1))^2]\le\mathbb{E}[(G^k(1))'^2].\end{aligned} \tag{5.10}$$

We now use the definition of G^k, which yields

$$\mathbb{E}\left[(G^k(1))'^2\right]=\frac{\beta^2p!}{N^{2p}}\mathbb{E}\left[\mathcal{G}_{N,\beta}^{\otimes 2}\left[\left(\sum_{\mu=1}^{M(N)}\sum_{\substack{\mathcal{I}\ni k\\|\mathcal{I}|=p}}\sigma_{\mathcal{I}}^1\xi_{\mathcal{I}}^{\mu}\right)\left(\sum_{\nu=1}^{M(N)}\sum_{\substack{\mathcal{J}\ni k\\|\mathcal{J}|=p}}\xi_{\mathcal{J}}^{\nu}\sigma_{\mathcal{J}}^2\right)\right]\right]. \tag{5.11}$$

We separate the integrand in two terms,

$$\left(\sum_{\mu=1}^{M(N)}\sum_{\substack{\mathcal{I}\ni k\\|\mathcal{I}|=p}}\sigma_{\mathcal{I}}^1\xi_{\mathcal{I}}^{\mu}\right)\left(\sum_{\nu=1}^{M(N)}\sum_{\substack{\mathcal{J}\ni k\\|\mathcal{J}|=p}}\xi_{\mathcal{J}}^{\nu}\sigma_{\mathcal{J}}^2\right)=\sum_{\mathcal{I},\mathcal{J}\ni k}\sigma_{\mathcal{I}}^1\sum_{\mu=1}^{M(N)}\xi_{\mathcal{I}}^{\mu}\xi_{\mathcal{J}}^{\mu}\sigma_{\mathcal{J}}^2+\sum_{\substack{\mu,\nu=1\\\mu\neq\nu}}^{M(N)}\sum_{\mathcal{I},\mathcal{J}\ni k}\sigma_{\mathcal{I}}^1\sigma_{\mathcal{J}}^2\xi_{\mathcal{I}}^{\mu}\xi_{\mathcal{J}}^{\nu}. \tag{5.12}$$

The first term on the right is treated uniformly by a result from Part II. Indeed, consider $\sum_\mu\xi_{\mathcal{I}}^{\mu}\xi_{\mathcal{J}}^{\mu}$ as an element of the matrix $(T_{\mathcal{I},\mathcal{J}})_{\mathcal{I},\mathcal{J}}$ representing a linear map from an $\binom{N-1}{p-1}$ dimensional vector space onto itself. By Theorem 8.3, this yields ($\|\cdot\|_{\mathrm{op}}$ denoting the operator norm of a matrix)

$$\begin{aligned}N^{-2p}\sum_{\mathcal{I},\mathcal{J}\ni k}\sigma_{\mathcal{I}}^1\sum_{\mu=1}^{M(N)}\xi_{\mathcal{I}}^{\mu}\xi_{\mathcal{J}}^{\mu}\sigma_{\mathcal{J}}^2&\le N^{-2p}\left(\sum_{\mathcal{I}\ni k}(\sigma_{\mathcal{I}}^1)^2\right)\left\|\sum_{\mu=1}^{M(N)}\xi_{\mathcal{I}}^{\mu}\xi_{\mathcal{J}}^{\mu}\right\|_{\mathrm{op}}\\&\le N^{-2p}\binom{N}{p-1}\left(C_2N^{p-1}+C_3N^{2p-2}e^{-C_4N^{1/4}}\right)\\&\le C_4N^{-2}.\end{aligned} \tag{5.13}$$

To bound the second term, we use an approach which will be developed much further in the next chapter. Namely, we expand the Boltzmann factors appearing in the Gibbs measures. We start by writing

$$\mathbb{E}\left[\mathcal{G}^{\otimes 2}\left[\sum_{\substack{\mu,\nu=1\\\mu\neq\nu}}^{M(N)}\sum_{\mathcal{I},\mathcal{J}\ni k}\sigma_{\mathcal{I}}^1\sigma_{\mathcal{J}}^2\xi_{\mathcal{I}}^{\mu}\xi_{\mathcal{J}}^{\nu}\right]\right]=\sum_{\substack{\mu,\nu=1\\\mu\neq\nu}}^{M(N)}\sum_{\mathcal{I},\mathcal{J}\ni k}\mathbb{E}\left[\mathbb{E}_{\sigma^1,\sigma^2}\left[\frac{e^{-\beta H(\sigma^1)-\beta H(\sigma^2)}}{Z^2}\sigma_{\mathcal{I}}^1\sigma_{\mathcal{J}}^2\xi_{\mathcal{I}}^{\mu}\xi_{\mathcal{J}}^{\nu}\right]\right]. \tag{5.14}$$

We now treat the right-hand side term by term. The idea is to consider the part of H which is independent of both $\xi^\mu_\mathcal{I}$ and $\xi^\nu_\mathcal{J}$ as the main object, and the rest as a perturbation. To this aim, we introduce some notation. Let $\mathcal{P} = \{\mathcal{X} \subset \mathcal{N} : |\mathcal{X}| = p\}$, and define for any $\mathcal{K} \in \mathcal{P}$,

$$\mathcal{B}_\mathcal{K} = \{\mathcal{L} \subset \mathcal{P} : \mathcal{L} \cap \mathcal{K} \neq \emptyset\}, \tag{5.15}$$

and $\mathcal{B}^c_\mathcal{K} = \mathcal{P} \setminus \mathcal{B}_\mathcal{K}$. Consider now

$$\mathbb{E}_{\sigma^1,\sigma^2}\left[\frac{e^{-\beta H(\sigma^1)-\beta H(\sigma^2)}}{Z^2}\sigma^1_\mathcal{I}\sigma^2_\mathcal{J}\xi^\mu_\mathcal{I}\xi^\nu_\mathcal{J}\right] \tag{5.16}$$

as a function f of the variables $\{\xi^\mu_\mathcal{K}\}_{\mathcal{K}\in\mathcal{B}_\mathcal{I}}$ and $\{\xi^\nu_{\mathcal{K}'}\}_{\mathcal{K}'\in\mathcal{B}_\mathcal{J}}$ only. We now expand the this function about 0, up to third order, with a remainder of forth order. Taylor's theorem implies that

$$\begin{aligned}
f(\{\xi^\mu_\mathcal{K}\}_{\mathcal{K}\in\mathcal{B}_\mathcal{I}}, \{\xi^\nu_{\mathcal{K}'}\}_{\mathcal{K}'\in\mathcal{B}_\mathcal{J}}) \leq f(0) &+ \sum_{\mathcal{K}\in\mathcal{B}_\mathcal{I}} \xi^\mu_\mathcal{K}\frac{\partial}{\partial\xi^\mu_\mathcal{K}}f(0) + \sum_{\mathcal{K}\in\mathcal{B}_\mathcal{J}} \xi^\nu_\mathcal{K}\frac{\partial}{\partial\xi^\nu_\mathcal{K}}f(0) \\
&+ \frac{1}{2}\sum_{\mathcal{K},\mathcal{K}'\in\mathcal{B}_\mathcal{I}} \xi^\mu_\mathcal{K}\xi^\mu_{\mathcal{K}'}\frac{\partial^2}{\partial\xi^\mu_\mathcal{K}\partial\xi^\mu_{\mathcal{K}'}}f(0) + \frac{1}{2}\sum_{\mathcal{K},\mathcal{K}'\in\mathcal{B}_\mathcal{J}} \xi^\nu_\mathcal{K}\xi^\nu_{\mathcal{K}'}\frac{\partial^2}{\partial\xi^\nu_\mathcal{K}\partial\xi^\nu_{\mathcal{K}'}}f(0) \\
&+ \sum_{\mathcal{K}\in\mathcal{B}_\mathcal{I}}\sum_{\mathcal{K}'\in\mathcal{B}_\mathcal{J}} \xi^\mu_\mathcal{K}\xi^\nu_{\mathcal{K}'}\frac{\partial^2}{\partial\xi^\mu_\mathcal{K}\partial\xi^\nu_{\mathcal{K}'}}f(0) + \frac{1}{6}\sum_{\mathcal{K},\mathcal{K}',\mathcal{K}''\in\mathcal{B}_\mathcal{I}} \xi^\mu_\mathcal{K}\xi^\mu_{\mathcal{K}'}\xi^\mu_{\mathcal{K}''}\frac{\partial^3}{\partial\xi^\mu_\mathcal{K}\partial\xi^\mu_{\mathcal{K}'}\partial\xi^\mu_{\mathcal{K}''}}f(0) \\
&+ \frac{1}{6}\sum_{\mathcal{K},\mathcal{K}',\mathcal{K}''\in\mathcal{B}_\mathcal{I}} \xi^\mu_\mathcal{K}\xi^\mu_{\mathcal{K}'}\xi^\mu_{\mathcal{K}''}\frac{\partial^3}{\partial\xi^\mu_\mathcal{K}\partial\xi^\mu_{\mathcal{K}'}\partial\xi^\mu_{\mathcal{K}''}}f(0) \\
&+ \frac{1}{2}\sum_{\mathcal{K},\mathcal{K}'\in\mathcal{B}_\mathcal{I}}\sum_{\mathcal{K}''\in\mathcal{B}_\mathcal{J}} \xi^\mu_\mathcal{K}\xi^\mu_{\mathcal{K}'}\xi^\nu_{\mathcal{K}''}\frac{\partial^3}{\partial\xi^\mu_\mathcal{K}\partial\xi^\mu_{\mathcal{K}'}\partial\xi^\nu_{\mathcal{K}''}}f(0) \\
&+ \frac{1}{2}\sum_{\mathcal{K}\in\mathcal{B}_\mathcal{I}}\sum_{\mathcal{K}',\mathcal{K}''\in\mathcal{B}_\mathcal{J}} \xi^\mu_\mathcal{K}\xi^\nu_{\mathcal{K}'}\xi^\nu_{\mathcal{K}''}\frac{\partial^3}{\partial\xi^\mu_\mathcal{K}\partial\xi^\nu_{\mathcal{K}'}\partial\xi^\nu_{\mathcal{K}''}}f(0) \\
&+ \sup f^{(\text{iv})}\left(\sum_{\mathcal{K}\in\mathcal{B}_\mathcal{I}}\xi^\mu_\mathcal{K} + \sum_{\mathcal{K}'\in\mathcal{B}_\mathcal{J}}\xi^\nu_{\mathcal{K}'}\right),
\end{aligned} \tag{5.17}$$

where $\sup f^{(\text{iv})}$ stands for the supremum of the fourth derivatives of f. Straightforward calculations show that the only non-zero term are $\frac{\partial^2}{\partial\xi^\mu_\mathcal{I}\partial\xi^\nu_\mathcal{J}}f(0)$ and $\frac{\partial^3}{\partial\xi^\mu_\mathcal{I}\partial\xi^\nu_\mathcal{J}\partial\xi^\mu_\mathcal{K}}f(0)$ for any $\mathcal{K} \in \mathcal{B}_\mathcal{I}$, respectively $\frac{\partial^3}{\partial\xi^\mu_\mathcal{I}\partial\xi^\nu_\mathcal{J}\partial\xi^\nu_\mathcal{K}}f(0)$ for any $\mathcal{K} \in \mathcal{B}_\mathcal{J}$, as well as the error term.

Observe that the derivatives of f at 0 do not contain any of the variables ξ^μ_i, $i \in \mathcal{I}$, nor ξ^ν_j, $j \in \mathcal{J}$. Integrating (5.17) with respect to these variables thus only affects the monomials in $\xi^\mu_\mathcal{K}$ and $\xi^\nu_{\mathcal{K}'}$. Thus, by the Bernoulli nature of the variables ξ^μ_i,

$$\mathbb{E}f \leq \frac{1}{2}\mathbb{E}\,\frac{\partial^2}{\partial\xi^\mu_\mathcal{I}\partial\xi^\mu_\mathcal{I}}f(0) + \frac{1}{2}\mathbb{E}\,\frac{\partial^2}{\partial\xi^\nu_\mathcal{J}\partial\xi^\nu_\mathcal{J}}f(0) + \mathbb{E}\,\sup f^{(\text{iv})}\left(\sum_{\mathcal{K}\in\mathcal{B}_\mathcal{I}}\xi^\mu_\mathcal{K} + \sum_{\mathcal{K}'\in\mathcal{B}_\mathcal{J}}\xi^\nu_{\mathcal{K}'}\right) \tag{5.18}$$

Furthermore, observe that since f is a function on a compact set (a hypercube) any derivative of order q of f is bounded by a constant times $N^{q(1-p)}$. Hence,

$$\mathbb{E} f \leq C_6 N^{2-2p} + C_7 N^{4-4p} N^{2p-2} \leq C_8 N^{2-2p} \tag{5.19}$$

Using the above in (5.14), and summing over all allowed μ and ν, respectively $\mathcal{I}$ and $\mathcal{J}$, we obtain

$$N^{-2p} \mathbb{E} \left[\mathcal{G}^{\otimes 2} \left[\sum_{\substack{\mu,\nu=1 \\ \mu \neq \nu}}^{M(N)} \sum_{\mathcal{I},\mathcal{J} \ni k} \sigma^1_{\mathcal{I}} \sigma^2_{\mathcal{J}} \xi^{\mu}_{\mathcal{I}} \xi^{\nu}_{\mathcal{J}} \right] \right] \leq C_9 N^{-2} \tag{5.20}$$

Inserting (5.20) and (5.13) in (5.10), and using this in (5.4), we finally get

$$\mathbb{E} \left[(F - \mathbb{E} F)^2 \right] \leq C_1 \sum_{k=1}^{N} \mathbb{E} \left[\left(\tilde{F}^k \right)^2 \right] \leq C_{10} \sum_{k=1}^{N} N^{-2} = C_{10} N^{-1}. \tag{5.21}$$

This proves the first part of Theorem 2.4. Inequality (2.32) then follows by Chebyshev-Markov (with second moment). □

6 Condensation

6.1 Integration by Parts Formula: Proof of Theorem 2.5

Once again, a lengthy calculation will retrieve the Gaussian result, for which an integration by parts yields the result (2.37) almost immediately (see (2.30) in [T4]).[31] We start by evaluating the left-hand side of (2.37). By definition,

$$\beta \mathbb{E} \frac{\partial F_N}{\partial \beta} = -\beta \mathbb{E} \mathcal{G}_{N,\beta}[H] = \frac{\beta (p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\mu, \mathcal{I}} \mathbb{E}\mathbb{E}_\sigma \left[\frac{e^{-\beta H(\sigma)}}{Z_N} \xi_{\mathcal{I}}^\mu \sigma_{\mathcal{I}} \right]. \tag{6.1}$$

The idea is to isolate in the Hamiltonian the contribution from the term $\prod_{l \in \mathcal{I}} \xi_l^\mu \sigma_l$, expand the exponential of this quantity, and finally integrate with respect to these variables. Let $\mathcal{B} \subset \mathcal{P} = \{\mathcal{X} \subset \mathcal{N} : |\mathcal{X}| = p\}$ be defined by

$$\mathcal{B} = \mathcal{B}_{\mathcal{I}} = \{\mathcal{J} \in \mathcal{P} : \mathcal{J} \cap \mathcal{I} \neq \emptyset\}, \tag{6.2}$$

and let $\mathcal{B}^c = \mathcal{P} \setminus \mathcal{B}$. Define

$$H'[\omega](\sigma) = H'_{\mu,\mathcal{I}}[\omega](\sigma) = -\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu \neq \mu} \sum_{\mathcal{J} \in \mathcal{P}} \xi_{\mathcal{J}}^\nu \sigma_{\mathcal{J}} - \frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\mathcal{J} \in \mathcal{B}^c} \xi_{\mathcal{J}}^\mu \sigma_{\mathcal{J}} \tag{6.3}$$

Let $\mathcal{G}' = \mathcal{G}'_{\mu,\mathcal{I}}$ denote the Gibbs measure associated to the Hamiltonian H', and $Z'_{\mu,\mathcal{I}} = \mathbb{E}_\sigma e^{-\beta H'_{\mathcal{I}}}$ the corresponding partition function.

Proposition 6.1: *There exists a constant $C > 0$ such that*

$$\left| N^{1-p} \mathbb{E} \mathcal{G} \left[\xi_{\mathcal{I}}^\mu \sigma_{\mathcal{I}} \right] - \beta (p!)^{\frac{1}{2}} N^{2-2p} + \beta (p!)^{\frac{1}{2}} N^{2-2p} \mathbb{E} \mathcal{G}'^{\otimes 2}_{\mu,\mathcal{I}} \left[\sigma_{\mathcal{I}}^1 \sigma_{\mathcal{I}}^2 \right] \right| \leq C N^{1-2p}. \tag{6.4}$$

Proof: The key idea is to expand ad nauseam Boltzmann factors in the disorder variables (a first glimpse of which we already had in Chapter 5). Let $x_{\mathcal{J}} = N^{1-p} \xi_{\mathcal{I}}^\mu$, for $\mathcal{J} \in \mathcal{B}$. Define the function $f : \mathbb{R}^n \to \mathbb{R}$, with $n = \binom{N}{p} - \binom{N-p}{p}$ through

$$f((x_{\mathcal{J}})_{\mathcal{J} \in \mathcal{B}}) = \mathcal{G}\left[x_{\mathcal{I}} \sigma_{\mathcal{I}} \right]. \tag{6.5}$$

This function is C^∞ and therefore, by Taylor's Theorem, there exists a $C_1 > 0$ such that for any $x \in [-1, 1]^n$,

$$f(x) \leq f(0) + \sum_{k=1}^{5} \sum_{p:|p|=k} \frac{x^p D^p f(0)}{p!} + C_1 \left(\sum_{\mathcal{J} \in \mathcal{B}} x_{\mathcal{J}} \right)^6, \tag{6.6}$$

where we have used the usual multi-index notation, that is, p is a multi-index $(p_{\mathcal{J}})_{\mathcal{J} \in \mathcal{B}}$, $|p|$ is the sum of its components, $p! = \prod_{\mathcal{J}} p_{\mathcal{J}}$, and $x^p = \prod_{\mathcal{J}} x_{\mathcal{J}}^{p_{\mathcal{J}}}$, resp. $D^p = \prod_{\mathcal{J}} \frac{d^{p_{\mathcal{J}}}}{dx_{\mathcal{J}}^{p_{\mathcal{J}}}}$.

---

[31] This idea goes back to [ALR]

Let us start with the error term in (6.6). The $\xi_{\mathcal{I}}^{\mu}$ are centered Bernoulli random variables, and thus,

$$\mathbb{E}\left(\sum_{\mathcal{J}\in\mathcal{B}} x_{\mathcal{J}}\right)^6 \leq C_2 N^{6-6p}\mathbb{E}\left(\sum_{\mathcal{J}\in\mathcal{B}} \xi_{\mathcal{J}}^{\mu}\right)^6 \leq C_3 N^{6-6p} N^{3p-3} = C_3 N^{3-3p}. \tag{6.7}$$

We now turn to the lower order terms appearing in (6.6). The main part of the calculations will be to reduce the number of terms appearing. Observe first that

$$\mathbb{E}\left[x^p D^p f(0)\right] = \mathbb{E}\left[D^p f(0)\mathbb{E}_{\mu,\mathcal{I}}[x^p]\right] = 0, \tag{6.8}$$

if there exists an $i \in \mathcal{I}$ such that $|\{\mathcal{J} \in \mathcal{B} : \mathcal{J} \ni i \text{ and } p_{\mathcal{J}} > 0\}|$ is odd. Also,

$$D^p f(0) = 0 \tag{6.9}$$

whenever $p_{\mathcal{I}} = 0$ (since in this case, a factor $x_{\mathcal{I}}$ is left over after taking derivatives). In particular, $f(0) = 0$. This implies that the expectation of the left-hand side of (6.6) can be explicitly written as (two or more indices $\mathcal{J}_i$ in a sum are understood to be different)

$$\begin{aligned}
\mathbb{E} f(x) \leq\ & \mathbb{E}\left[x_{\mathcal{I}}^2 \frac{D_{\mathcal{I}}^2 f(0)}{2}\right] + \sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} f(0)\right] \\
&+ \mathbb{E}\left[x_{\mathcal{I}}^4 \frac{D_{\mathcal{I}}^4 f(0)}{4!}\right] + \sum_{\mathcal{J}\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}}^2 x_{\mathcal{J}}^2 \frac{D_{\mathcal{I}}^2 D_{\mathcal{J}}^2 f(0)}{4}\right] \\
&+ \sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0)\right] \\
&+ \sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3,\mathcal{J}_4\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} x_{\mathcal{J}_4} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} D_{\mathcal{J}_4} f(0)\right] \\
&+ \frac{1}{2} \sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}}^2 x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{I}}^2 D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0)\right] \\
&+ \frac{1}{2} \sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1}^2 x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{I}} D_{\mathcal{J}_1}^2 D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0)\right] \\
&+ \frac{1}{3!} \sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}}^3 x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{I}}^3 D_{\mathcal{J}_1} D_{\mathcal{J}_2} f(0)\right] + C_3 N^{3-3p}.
\end{aligned} \tag{6.10}$$

Let us treat term by term in the above expression. The first one is easy to calculate, and is

$$\mathbb{E}\left[x_{\mathcal{I}}^2 \frac{D_{\mathcal{I}}^2 f(0)}{2}\right] = \beta(p!)^{\frac{1}{2}} N^{2-2p} - \beta(p!)^{\frac{1}{2}} N^{2-2p}\mathbb{E}\mathcal{G}'^{\otimes 2}\left[\sigma_{\mathcal{I}}^1 \sigma_{\mathcal{I}}^2\right], \tag{6.11}$$

which is exactly the term on the left-hand side of (6.4). This means that we have to show that the remaining terms are small corrections (that is, we have mainly to worry about the powers of N). From now on, we will use extensively the fact that the derivatives of f are bounded uniformly by a constant.

Most of the remaining terms in (6.10) can be bounded uniformly. The third term in this expression satisfies

$$\left|\mathbb{E}\left[x_{\mathcal{I}}^4 \frac{D_{\mathcal{I}}^4 f(0)}{4!}\right]\right| \leq C_4 N^{4-4p}. \tag{6.12}$$

Similarly,

$$\left|\sum_{\mathcal{J}\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}}^2 x_{\mathcal{J}}^2 \frac{D_{\mathcal{I}}^2 D_{\mathcal{J}}^2 f(0)}{4}\right]\right| \leq C_5' N^{4-4p} \binom{N}{p-1} \leq C_5 N^{3-3p}. \tag{6.13}$$

Also,

$$\begin{aligned}\left|\sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}}^2 x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{I}}^2 D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0)\right]\right| &\leq C_6 N^{5-5p} N^{3p-4} \\ &= C_6 N^{1-2p},\end{aligned} \tag{6.14}$$

since $|(\mathcal{J}_1 \cup \mathcal{J}_2 \cup \mathcal{J}_3) \setminus \mathcal{I}| \leq 3p-4$, and

$$\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}}^3 x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{I}}^3 D_{\mathcal{J}_1} D_{\mathcal{J}_2} f(0)\right] \leq C_7' N^{5-5p} \binom{N}{p} \leq C_7 N^{5-4p}, \tag{6.15}$$

since $|\mathcal{J}_1 \cup \mathcal{J}_2 \setminus \mathcal{I}|$ can be at most equal to p. Finally,

$$\begin{aligned}\sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1}^2 x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{I}} D_{\mathcal{J}_1}^2 D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0)\right] &\leq C_8 N^{5-5p} N^{2p-1} \\ &= C_8 N^{4-3p}.\end{aligned} \tag{6.16}$$

The remaining terms have to be treated with more care. In fact, we will apply again a Taylor expansion.

Lemma 6.2: *There exists a number $C > 0$ (depending only on p and β) such that*

$$\left|\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} f(0)\right]\right| \leq C N^{2-2p} N^{-1}. \tag{6.17}$$

Proof: From (6.8), we know that to get a non-zero value in the expectation, the sets $\mathcal{J}_1$ and $\mathcal{J}_2$ have to fulfill the following condition: each $i \in \mathcal{I}$ has to be in exactly one of the sets $\mathcal{J}_i$. This leaves at most p elements in the set $\mathcal{I}' = (\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \mathcal{I}$. We distinguish two subsets of $\mathcal{I}'$. The set of $i \in \mathcal{I}'$ which are in exactly one of the sets $\mathcal{J}_i$, denoted by $\mathcal{J}_o$ and those that are in both of them, denoted by $\mathcal{J}_e$. Obviously, $|\mathcal{J}_o| + 2|\mathcal{J}_e| = p$, and both $|\mathcal{J}_o|$ and $|\mathcal{J}_e|$ are even. We now decompose the right-hand side of (6.17) as

$$\begin{aligned}&\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} f(0)\right] \\ &\qquad = \sum_{k=1}^{p} \sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I} \\ |\mathcal{K}|=k}} \sum_{\substack{\mathcal{L}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}) \\ |\mathcal{L}|=(p-k)/2}} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B} \\ \mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}} \mathbb{E}\left[N^{3-3p} \xi_{\mathcal{K}}^{\mu} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} f(0)\right] \\ &\qquad = \sum_{k=1}^{p} \sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I} \\ |\mathcal{K}|=k}} \sum_{\substack{\mathcal{L}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}) \\ |\mathcal{L}|=(p-k)/2}} \mathbb{E}\left[N^{3-3p} \xi_{\mathcal{K}}^{\mu} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B} \\ \mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} f(0)\right].\end{aligned} \tag{6.18}$$

If k is small enough, we can use a uniform bound. Indeed, let $k \leq p-4$. Then

$$\begin{aligned}\left|\sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I}\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{L}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\|\mathcal{L}|=(p-k)/2}}\mathbb{E}\left[N^{3-3p}\xi^{\mu}_{\mathcal{K}}\sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}\\\mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}}D_{\mathcal{I}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}f(0)\right]\right| &\leq C_1N^{3-3p}N^{k+\frac{p-k}{2}}\\ &= C_1N^{2-2p}N^{1-\frac{p-k}{2}}\\ &\leq C_1N^{-1}N^{2-2p}.\end{aligned} \tag{6.19}$$

The only remaining cases are $k=p$ and $k=p-2$. To bound them, let

$$\mathcal{B}'=\{\mathcal{X}\subset\mathcal{N}\setminus\mathcal{I} : |\mathcal{X}|=p \text{ and } \mathcal{X}\cup\mathcal{K}\neq\emptyset\}\cup\mathcal{K}, \tag{6.20}$$

and for all $\mathcal{J}\in\mathcal{B}'$, let $x_{\mathcal{J}}=N^{1-p}\xi^{\mu}_{\mathcal{J}}$. Furthermore, observe that

$$\sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}\\\mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}}D_{\mathcal{I}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}f(0) \tag{6.21}$$

is independent of $(x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}\cup\mathcal{I}}$. Define therefore

$$g_{\mathcal{I},\mathcal{K}}(\{x_{\mathcal{J}}\}_{\mathcal{J}\in\mathcal{B}'})\equiv x_{\mathcal{K}}\sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}\\\mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}}D_{\mathcal{I}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}f(0). \tag{6.22}$$

Let $k=p$. Then, by Taylor's theorem, the fact that we consider a bounded domain, and the same symmetry reasons as before,

$$\begin{aligned}\mathbb{E}\left[g_{\mathcal{I},\mathcal{K}}(\{x_{\mathcal{J}}\}_{\mathcal{J}\in\mathcal{B}'})\right]\leq\mathbb{E}\Big[&g_{\mathcal{I},\mathcal{K}}(0)+\sum_{\mathcal{J}\in\mathcal{B}'\setminus\mathcal{K}}x_{\mathcal{K}}x_{\mathcal{J}}D_{\mathcal{K}}D_{\mathcal{J}}g(0)+\frac{1}{2}x^2_{\mathcal{K}}D^2_{\mathcal{K}}g_{\mathcal{I},\mathcal{K}}(0)\\ &+\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'\setminus\mathcal{K}}x_{\mathcal{K}}x_{\mathcal{J}_1}x_{\mathcal{J}_2}D_{\mathcal{I}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}g_{\mathcal{I},\mathcal{K}}(0)+C_2(\sum_{\mathcal{J}\in\mathcal{B}}x_{\mathcal{J}})^4\Big].\end{aligned} \tag{6.23}$$

In the last expression, we bound every term uniformly. This yields,

$$\mathbb{E}\left[g_{\mathcal{I},\mathcal{K}}(\{x_{\mathcal{J}}\}_{\mathcal{J}\in\mathcal{B}'})\right]\leq 0+0+C_3N^{2-2p}+C_4N^{3-3p}N^p+C_5N^{4-4p}N^{2p-2}\leq C_6N^{3-2p}. \tag{6.24}$$

Summing over all allowed $\mathcal{K}$, we get for the last summand in (6.18)

$$\begin{aligned}\sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I}\\|\mathcal{K}|=p}}N^{2-2p}\mathbb{E}\left[x_{\mathcal{K}}\sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}\\\mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}}D_{\mathcal{I}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}f(0)\right] &\leq C_7N^{2-2p}N^pN^{3-2p}\\ &= C_7N^{2-2p}N^{3-p}\leq C_7N^{2-2p}N^{-1},\end{aligned} \tag{6.25}$$

since $p\geq 4$.

Let us turn to the case $k = p - 2$. By the same arguments as before, we get that (observe that the quadratic part now includes more terms),

$$\begin{aligned}\sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I}\\|\mathcal{K}|=p-2}}\ \sum_{\substack{\mathcal{L}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\|\mathcal{L}|=1}} & N^{2-2p}\mathbb{E}[g_{\mathcal{I},\mathcal{K}}(\{x_{\mathcal{J}}\}_{\mathcal{J}\in\mathcal{B}'})]\\ &\le \sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I}\\|\mathcal{K}|=p-2}}\ \sum_{\substack{\mathcal{L}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\|\mathcal{L}|=1}} N^{2-2p}\mathbb{E}\Big[g_{\mathcal{I},\mathcal{K}}(0)+\sum_{\substack{\mathcal{J}\in\mathcal{B}'\\\mathcal{J}\supset\mathcal{K}}} x_{\mathcal{K}}x_{\mathcal{J}}D_{\mathcal{K}}D_{\mathcal{J}}g(0)\\ &\quad+\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'} x_{\mathcal{K}}x_{\mathcal{J}_1}x_{\mathcal{J}_2}D_{\mathcal{K}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}g(0)+C_8(\sum_{\mathcal{J}\in\mathcal{B}}x_{\mathcal{J}})^4\Big].\end{aligned}\tag{6.26}$$

Suppose that $p \geq 6$ (the case $p = 4$ will be treated below). Using uniform bounds, we get

$$\begin{aligned}\sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I}\\|\mathcal{K}|=p-2}}\ \sum_{\substack{\mathcal{L}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\|\mathcal{L}|=1}} & N^{2-2p}\mathbb{E}[g_{\mathcal{I},\mathcal{K}}(\{x_{\mathcal{J}}\}_{\mathcal{J}\in\mathcal{B}'})]\\ &\le \sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I}\\|\mathcal{K}|=p-2}}\ \sum_{\substack{\mathcal{L}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\|\mathcal{L}|=1}} N^{2-2p}\big(0+C_9N^{2-2p}N^{p-k}\\ &\quad+C_{10}N^{3-3p}N^{2p-k}+C_{11}N^{4-4p}N^{2p-2}\big)\\ &\le C_{12}N^{2-2p}N^{p-1}(N^{4-2p}+N^{5-2p}+N^{2-2p})\\ &\le C_{13}N^{2-2p}N^{4-p}\le C_{13}N^{2-2p}N^{-2}.\end{aligned}\tag{6.27}$$

In the case $p = 4$, one sees from (6.27) that the only term which is not of the desired order in (6.26) is

$$\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'} x_{\mathcal{K}}x_{\mathcal{J}_1}x_{\mathcal{J}_2}D_{\mathcal{K}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}g(0)\tag{6.28}$$

To get the desired bound, we have to expand it again.

Lemma 6.3: *There exists a constant $C > 0$ such that*

$$\left|\sum_{\substack{\mathcal{K}\subset\mathcal{N}\setminus\mathcal{I}\\|\mathcal{K}|=p-2}}\ \sum_{\substack{\mathcal{L}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\|\mathcal{L}|=1}} \mathbb{E}\left[\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'} x_{\mathcal{K}}x_{\mathcal{J}_1}x_{\mathcal{J}_2}D_{\mathcal{K}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}g(0)\right]\right|\le CN^{-1}.\tag{6.29}$$

Proof: We proceed in exactly the same way. Note that in this case, $k = 2$ and $l = 1$. With the analogue definitions of $\mathcal{J}_e$ and $\mathcal{J}_o$, we get

$$\begin{aligned}
&\mathbb{E}\left[\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'} x_{\mathcal{K}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} g(0)\right] \\
&\quad = \sum_{k'=2}^{6} \sum_{l'=0}^{3-\frac{k'}{2}} \sum_{\substack{\mathcal{K}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\ |\mathcal{K}'|=k'}} \sum_{\substack{\mathcal{L}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}\cup\mathcal{K}')\\ |\mathcal{L}'|=l'}} \mathbb{E}\left[N^{3-3p}\xi^{\mu}_{\mathcal{K}'} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'\\ \mathcal{J}_o=\mathcal{K}',\mathcal{J}_e=\mathcal{L}'}} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} g(0)\right] \\
&\quad = \sum_{k'=2}^{6} \sum_{l'=0}^{3-\frac{k'}{2}} \sum_{\substack{\mathcal{K}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\ |\mathcal{K}'|=k'}} \sum_{\substack{\mathcal{L}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}\cup\mathcal{K}')\\ |\mathcal{L}'|=l'}} N^{2-2p}\mathbb{E}\left[x_{\mathcal{K}'} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'\\ \mathcal{J}_o=\mathcal{K}',\mathcal{J}_e=\mathcal{L}'}} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} g(0)\right] \\
&\quad = \sum_{k'=2}^{6} \sum_{l'=0}^{3-\frac{k'}{2}} \sum_{\substack{\mathcal{K}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\ |\mathcal{K}'|=k'}} \sum_{\substack{\mathcal{L}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}\cup\mathcal{K}')\\ |\mathcal{L}'|=l'}} N^{2-2p}\mathbb{E}\left[h_{\mathcal{K}',\mathcal{L}'}((x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}''})\right],
\end{aligned} \tag{6.30}$$

with the definitions

$$\mathcal{B}'' = \{\mathcal{X}\subset\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}) : |\mathcal{X}| = p \text{ and } \mathcal{X}\cap\mathcal{K}\neq\emptyset\}, \tag{6.31}$$

and

$$h_{\mathcal{K}',\mathcal{L}'}((x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}''}) = x_{\mathcal{K}'} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'\\ \mathcal{J}_o=\mathcal{K}',\mathcal{J}_e=\mathcal{L}'}} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} g(0). \tag{6.32}$$

Suppose that $k' = 2$ (the case $k' = 0$ does not appear). Then a uniform bound will suffice. Indeed,

$$\begin{aligned}
\sum_{\substack{\mathcal{K}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\ |\mathcal{K}'|=k'}} \sum_{\substack{\mathcal{L}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}\cup\mathcal{K}')\\ |\mathcal{L}'|=l'}} \mathbb{E}\left[N^{3-3p}\xi^{\mu}_{\mathcal{K}'} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'\\ \mathcal{J}_o=\mathcal{K}',\mathcal{J}_e=\mathcal{L}'}} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} g(0)\right]& \\
\le C_1 N^{k'+l'+3-3p}& \\
\le C_1 N^{3+\frac{k'}{2}+3-3p}&,
\end{aligned} \tag{6.33}$$

and thus

$$\begin{aligned}
\sum_{\substack{\mathcal{K}\in\mathcal{N}\setminus\mathcal{I}\\ |\mathcal{K}|=2}} \sum_{\substack{\mathcal{L}\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\ |\mathcal{L}|=1}} \sum_{\substack{\mathcal{K}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\ |\mathcal{K}'|=k'}} \sum_{\substack{\mathcal{L}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}\cup\mathcal{K}')\\ |\mathcal{L}'|=l'}} \mathbb{E}\left[N^{3-3p}\xi^{\mu}_{\mathcal{K}'} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'\\ \mathcal{J}_o=\mathcal{K}',\mathcal{J}_e=\mathcal{L}'}} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} g(0)\right]& \\
\le C_2 N^{6+\frac{k'}{2}+3-3p} \le C_2 N^{6-2p} \le N^{-2}&.
\end{aligned} \tag{6.34}$$

On the other hand, if $k' = 4, 6$, we expand the right-hand side of (6.30) up to third order. One obtains

$$\begin{aligned} &\mathbb{E}\left[N^{3-3p}\xi^{\mu}_{\mathcal{K}'} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}' \\ \mathcal{J}_o=\mathcal{K}',\mathcal{J}_e=\mathcal{L}'}} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} g(0) \right] \\ &\qquad \leq N^{2-2p}\mathbb{E}\Big[h(0) + \sum_{\substack{\mathcal{J}\in\mathcal{B}' \\ \mathcal{J}\supset\mathcal{K}'}} x_{\mathcal{K}'} x_{\mathcal{J}} D_{\mathcal{K}'} D_{\mathcal{J}} h(0) \\ &\qquad\qquad + \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}' \\ \mathcal{J}_1,\mathcal{J}_2\supset\mathcal{K}'}} x_{\mathcal{K}'} x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{K}'} D_{\mathcal{J}_1} D_{\mathcal{J}_2} h(0) \\ &\qquad\qquad + C_3 \big(\sum_{\mathcal{J}\in\mathcal{B}''} x_{\mathcal{J}}\big)^4 \Big] \\ &\qquad \leq C_4 N^{2-2p}\Big(0 + N^{2-2p+p-k'} + N^{3-3p+2p-k'} + N^{2-2p}\Big). \end{aligned} \tag{6.35}$$

Summing over all allowed $\mathcal{K}, \mathcal{L}, \mathcal{K}', \mathcal{L}'$ yields (since $k' + l' \leq 3 + \frac{k'}{2}$ and $k + l = 3$)

$$\begin{aligned} &\sum_{\substack{\mathcal{K}\in\mathcal{N}\setminus\mathcal{I} \\ |\mathcal{K}|=k}} \sum_{\substack{\mathcal{L}\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}) \\ |\mathcal{L}|=l}} \sum_{\substack{\mathcal{K}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}) \\ |\mathcal{K}'|=k'}} \sum_{\substack{\mathcal{L}'\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}\cup\mathcal{K}') \\ |\mathcal{L}'|=l'}} N^{2-2p}\mathbb{E}\left[h_{\mathcal{K}',\mathcal{L}'}((x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}''}) \right] \\ &\qquad \leq C_5 N^{3+k'+l'+2-2p}\Big(N^{3-p-k'} + N^{2-2p}\Big) \\ &\qquad \leq C_5\Big(N^{7-2p-\frac{k'}{2}} + N^{6-3p-\frac{k'}{2}}\Big) \leq C_5 N^{-1}. \end{aligned} \tag{6.36}$$

Finally, we sum over all allowed values for k' and l', using respectively the bounds (6.34) and (6.36). This yields the upper bound of (6.29). The lower bound follows by changing the sign of the error term in (6.35). □

Using (6.19), (6.25), and (6.27) or Lemma 6.3 in (6.18), we get

$$\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} f(0) \right] \leq C N^{2-2p} N^{-1}. \tag{6.37}$$

This is the upper bound in (6.17). To get the lower bound, we proceed as above, but change the sign of the quartic (error) term in (6.23), and (6.27) resp. This concludes the proof of Lemma 6.2. □

The fourth order term in (6.10) is taken care of by the following, analogous result.

Lemma 6.4: *There exists a constant $C > 0$ such that*

$$\left| \sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0) \right] \right| \leq C N^{-1} N^{2-2p}. \tag{6.38}$$

Proof: The proof is almost identical to the previous one. Define

$$\mathcal{J}_o = \mathcal{J}_o(\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3) = \left\{ i \in \bigcup_j \mathcal{J}_j \setminus \mathcal{I} : i \text{ is in an odd number of the } J_i\text{'s} \right\} \tag{6.39}$$

and

$$\mathcal{J}_o = \mathcal{J}_o(\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3) = \left\{ i \in \bigcup_j \mathcal{J}_j \setminus \mathcal{I} : i \text{ is in an even number of the } J_i\text{'s} \right\}. \tag{6.40}$$

Obviously, $|\mathcal{J}_o| + 2|\mathcal{J}_e| \leq 2p$, since otherwise at least one $i \in \mathcal{I}$ is not in any of the $\mathcal{J}_i$, and thus this term vanishes when integrated over $\xi^\mu_{\mathcal{I}}$. Also, $|\mathcal{J}_o|$ is even. Write the sum on the left-hand side of (6.38) as

$$\begin{aligned}
&\sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3 \in \mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0)\right] \\
&\qquad = \sum_{k=1}^{2p} \sum_{l=0}^{p-\frac{k}{2}} \sum_{\substack{\mathcal{K} \in \mathcal{N} \setminus \mathcal{I} \\ |\mathcal{K}| = k}} \sum_{\substack{\mathcal{L} \in \mathcal{N} \setminus (\mathcal{I} \cup \mathcal{K}) \\ |\mathcal{L}| = l}} \mathbb{E}\left[N^{4-4p} \xi^\mu_{\mathcal{K}} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3 \in \mathcal{B} \\ \mathcal{J}_o = \mathcal{K}, \mathcal{J}_e = \mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0) \right]
\end{aligned} \tag{6.41}$$

Suppose that $k = |\mathcal{K}| \leq 2p - 6$. Then, since the derivatives of f at 0 are bounded by some constant, and $l \leq p - \frac{k}{2}$,

$$\begin{aligned}
&\left| \sum_{\substack{\mathcal{K} \in \mathcal{N} \setminus \mathcal{I} \\ |\mathcal{K}| = k}} \sum_{\substack{\mathcal{L} \in \mathcal{N} \setminus (\mathcal{I} \cup \mathcal{K}) \\ |\mathcal{L}| = l}} \mathbb{E}\left[N^{4-4p} \xi^\mu_{\mathcal{K}} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3 \in \mathcal{B} \\ \mathcal{J}_o = \mathcal{K}, \mathcal{J}_e = \mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0) \right] \right| \\
&\qquad \leq C_1 \sum_{\substack{\mathcal{K} \in \mathcal{N} \setminus \mathcal{I} \\ |\mathcal{K}| = k}} \sum_{\substack{\mathcal{L} \in \mathcal{N} \setminus (\mathcal{I} \cup \mathcal{K}) \\ |\mathcal{L}| = l}} N^{4-4p} \leq C_2 N^{4-4p+k+l} \\
&\qquad \leq C_2 N^{4-3p+\frac{k}{2}} \leq C_2 N^{1-2p}.
\end{aligned} \tag{6.42}$$

This means that all these terms are already smaller than desired, and we are left with the cases $k = 2p$, $k = 2p-2$, and $k = 2p-4$ (since $\mathcal{K}$ has to be even). We treat them in the usual way. Let

$$\mathcal{B}' = \{\mathcal{X} \in \mathcal{N} \setminus \mathcal{I} : |\mathcal{X}| = p \text{ and } \mathcal{X} \cap \mathcal{K} \neq \emptyset\} \cup \mathcal{K}. \tag{6.43}$$

and $x_{\mathcal{J}} = N^{1-p} \xi^\mu_{\mathcal{J}}$, for $\mathcal{J} \in \mathcal{B}'$. Define furthermore

$$g((x_{\mathcal{J}})_{\mathcal{J} \in \mathcal{B}'}) = g_{\mathcal{I},\mathcal{K},\mathcal{L}}((x_{\mathcal{J}})_{\mathcal{J} \in \mathcal{B}'}) = x_{\mathcal{K}} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3 \in \mathcal{B} \\ \mathcal{J}_o = \mathcal{K}, \mathcal{J}_e = \mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} f(0). \tag{6.44}$$

We observe that unless $p = 4$ and $k = p - 4$, the term $x_{\mathcal{K}}$ does not appear in the derivatives of f at 0, and therefore $D^n_{\mathcal{K}} g(0) = 0$, for all $n \neq 1$. In this case (i.e., either $k \neq p - 4$ or $p \neq 4$), we have by Taylor's theorem, and the same arguments as before (that is, the conditions on the sets $\mathcal{J}_i$)

$$\begin{aligned}\mathbb{E}&\left[N^{4-4p}\xi^\mu_{\mathcal{K}}\sum_{\substack{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}\\ \mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}} D_{\mathcal{I}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}D_{\mathcal{J}_3}f(0)\right]\\ &\qquad= N^{3-3p}\mathbb{E}\left[g((x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}'})\right]\\ &\qquad\le N^{3-3p}\mathbb{E}\Big[g(0)+\sum_{\mathcal{J}\in\mathcal{B}'}x_{\mathcal{J}}D_{\mathcal{J}}g(0)\\ &\qquad\quad+\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'\setminus\{\mathcal{K}\}}x_{\mathcal{K}}x_{\mathcal{J}_1}x_{\mathcal{J}_2}D_{\mathcal{K}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}g(0)+C_3\Big(\sum_{\mathcal{J}\in\mathcal{B}'}x_{\mathcal{J}}\Big)^4\Big].\end{aligned}\tag{6.45}$$

Integrating with respect to $\xi^\mu_{\mathcal{K}}$, and uniform bounds on the remaining terms yields

$$\begin{aligned}N^{3-3p}\mathbb{E}\left[g((x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}'})\right]&\le C_4N^{3-3p}\Big(0+0+N^{3-3p+2p-k}+N^{2-2p}\Big)\\ &\le C_4N^{3-3p}\Big(N^{3-p-k}+N^{2-2p}\Big).\end{aligned}\tag{6.46}$$

Summing over all sets $\mathcal{K}$ of cardinality k, and all sets $\mathcal{L}$ of size l then gives (remember that $l\le p-\frac{k}{2}$)

$$\begin{aligned}\sum_{\substack{\mathcal{K}\in\mathcal{N}\setminus\mathcal{I}\\ |\mathcal{K}|=k}}\sum_{\substack{\mathcal{L}\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K})\\ |\mathcal{L}|=l}}\mathbb{E}&\left[N^{4-4p}\xi^\mu_{\mathcal{K}}\sum_{\substack{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}\\ \mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}} D_{\mathcal{I}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}D_{\mathcal{J}_3}f(0)\right]\\ &\le C_5N^{3-3p}N^l\Big(N^{3-p}+N^{2-2p+k}\Big)\\ &\le C_5N^{2-2p}\Big(N^{6-2p}+N^{3-p}\Big)\le C_5N^{-1}N^{2-2p},\end{aligned}\tag{6.47}$$

since $p\ge 4$.

Finally, if $p=4$ and $k=2p-4=4$, then there is an additional term in (6.45). Namely, we have to add $\frac{1}{2}x^2_{\mathcal{K}}D^2_{\mathcal{K}}g(0)$. This term, however, is also bounded uniformly by

$$\left|x^2_{\mathcal{K}}D^2_{\mathcal{K}}g(0)\right|\le C_6N^{2-2p}\tag{6.48}$$

and is therefore of the same order as the error term in (6.45). The above bound (6.47) is therefore also valid in this case.

Inserting the bounds (6.42) and (6.47) into the decomposition (6.41) proves the upper bound. The corresponding lower bound is obtained by changing the sign of the error term in (6.45). This concludes the proof of Lemma 6.4. □

Finally, the remaining term in (6.10) is treated in the same way by

Lemma 6.5: *There exists a constant $C>0$ such that*

$$\left|\sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3,\mathcal{J}_4\in\mathcal{B}}\mathbb{E}\left[x_{\mathcal{I}}x_{\mathcal{J}_1}x_{\mathcal{J}_2}x_{\mathcal{J}_3}x_{\mathcal{J}_4}D_{\mathcal{I}}D_{\mathcal{J}_1}D_{\mathcal{J}_2}D_{\mathcal{J}_3}D_{\mathcal{J}_4}f(0)\right]\right|\le CN^{-1}N^{2-2p}.\tag{6.49}$$

Proof: The proof is almost identical to the previous one. Define again J_o and J_e (compare (6.39) resp. (6.40)) by

$$\mathcal{J}_o = \mathcal{J}_o(\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4) = \left\{ i \in \bigcup_j \mathcal{J}_j \setminus \mathcal{I} : i \text{ is in an odd number of the } J_i\text{'s} \right\} \tag{6.50}$$

and

$$\mathcal{J}_o = \mathcal{J}_o(\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4) = \left\{ i \in \bigcup_j \mathcal{J}_j \setminus \mathcal{I} : i \text{ is in an even number of the } J_i\text{'s} \right\}. \tag{6.51}$$

Obviously, $|\mathcal{J}_o| + 2|\mathcal{J}_e| \leq 3p$, since otherwise at least one $i \in \mathcal{I}$ is not in any of the $\mathcal{J}_i$, and thus this term vanishes when integrated over $\xi^\mu_{\mathcal{I}}$. Also, $|\mathcal{J}_o|$ is even. Write the sum on the left-hand side of (6.49) as

$$\begin{aligned} &\sum_{\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3 \in \mathcal{B}} \mathbb{E}\left[x_{\mathcal{I}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} x_{\mathcal{J}_4} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} D_{\mathcal{J}_4} f(0)\right] \\ &\quad = \sum_{k=1}^{3p} \sum_{l=0}^{\frac{3p-k}{2}} \sum_{\substack{\mathcal{K} \in \mathcal{N} \setminus \mathcal{I} \\ |\mathcal{K}| = k}} \sum_{\substack{\mathcal{L} \in \mathcal{N} \setminus (\mathcal{I} \cup \mathcal{K}) \\ |\mathcal{L}| = l}} \mathbb{E}\left[N^{5-5p} \xi^\mu_{\mathcal{K}} \sum_{\substack{\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4 \in \mathcal{B} \\ \mathcal{J}_o = \mathcal{K}, \mathcal{J}_e = \mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} D_{\mathcal{J}_4} f(0) \right]. \end{aligned} \tag{6.52}$$

Suppose that $k \leq 3p - 8$. Then

$$\begin{aligned} &\left| \sum_{\substack{\mathcal{K} \in \mathcal{N} \setminus \mathcal{I} \\ |\mathcal{K}| = k}} \sum_{\substack{\mathcal{L} \in \mathcal{N} \setminus (\mathcal{I} \cup \mathcal{K}) \\ |\mathcal{L}| = l}} \mathbb{E}\left[N^{5-5p} \xi^\mu_{\mathcal{K}} \sum_{\substack{\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4 \in \mathcal{B} \\ \mathcal{J}_o = \mathcal{K}, \mathcal{J}_e = \mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} D_{\mathcal{J}_4} f(0) \right] \right| \\ &\qquad \leq C_1 N^{5-5p} N^{k+l} \leq C_1 N^{5-5p+\frac{3p+k}{2}} \leq C_1 N^{1-2p}. \end{aligned} \tag{6.53}$$

This means that all these terms are at most of the desired order. We are left with the cases $k = 3p, 3p-2, 3p-4, 3p-6$. Let again $\mathcal{B}'$ be the set of $\mathcal{J} \subset \mathcal{N} \setminus (\mathcal{I} \cup \mathcal{K})$ of size p plus the element $\mathcal{K}$, and also $x_{\mathcal{J}} = N^{1-p} \xi^\mu_{\mathcal{J}}$, for $\mathcal{J} \in \mathcal{B}'$. Then,

$$\begin{aligned} &\mathbb{E}\left[N^{5-5p} \xi^\mu_{\mathcal{K}} \sum_{\substack{\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4 \in \mathcal{B} \\ \mathcal{J}_o = \mathcal{K}, \mathcal{J}_e = \mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} D_{\mathcal{J}_4} f(0) \right] \\ &\qquad = N^{4-4p} \mathbb{E}\left[x_{\mathcal{K}} \sum_{\substack{\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3, \mathcal{J}_4 \in \mathcal{B} \\ \mathcal{J}_o = \mathcal{K}, \mathcal{J}_e = \mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} D_{\mathcal{J}_4} f(0) \right] \\ &\qquad = N^{4-4p} \mathbb{E}\left[g((x_{\mathcal{J}})_{\mathcal{J} \in \mathcal{B}'}) \right]. \end{aligned} \tag{6.54}$$

Again, we expand the last expression up to fifth order. Since $k \geq 3p-6$, we also have that $k > p$, and thus $D^n_{\mathcal{K}} g(0) = 0$ for all $n \neq 1$. Hence,

$$\begin{aligned} N^{4-4p}\mathbb{E}\left[g((x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}'})\right] \leq N^{4-4p}\mathbb{E}\Big[&\sum_{\mathcal{J}_1,\mathcal{J}_2\in\mathcal{B}'} x_{\mathcal{K}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} g(0) \\ &+ \sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}'} x_{\mathcal{K}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} g(0) \\ &+ \sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3,\mathcal{J}_4\in\mathcal{B}'} x_{\mathcal{K}} x_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} x_{\mathcal{J}_4} D_{\mathcal{K}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} D_{\mathcal{J}_4} g(0) \\ &+ \frac{1}{2} \sum_{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3\in\mathcal{B}'} x_{\mathcal{K}} x^2_{\mathcal{J}_1} x_{\mathcal{J}_2} x_{\mathcal{J}_3} D_{\mathcal{K}} D^2_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} g(0) \\ &+ C_2 \Big(\sum_{\mathcal{J}\in\mathcal{B}'} x_{\mathcal{J}}\Big)^6 \Big] \end{aligned} \tag{6.55}$$

Taking expectation with respect to $\xi^\mu_{\mathcal{K}}$ yields

$$\begin{aligned} N^{4-4p}\mathbb{E}\left[g((x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}'})\right] &\leq C_3 N^{4-4p}(N^{3-3p}N^{2p-k} + N^{4-4p+3p-k} + N^{5-5p+4p-k} \\ &\qquad + N^{5-5p+2p-k} + N^{3-3p}) \\ &\leq C_3 N^{2-2p}(N^{5-3p-k} + N^{6-3p-k} + N^{7-3p-k} + N^{5-5p}). \end{aligned} \tag{6.56}$$

Finally, we sum over all allowed sets $\mathcal{K}$ and $\mathcal{L}$. Since $l \leq 3$ and $k + l \leq 3p$, we get

$$\begin{aligned} \sum_{\substack{\mathcal{K}\in\mathcal{N}\setminus\mathcal{I} \\ |\mathcal{K}|=k}} \sum_{\substack{\mathcal{L}\in\mathcal{N}\setminus(\mathcal{I}\cup\mathcal{K}) \\ |\mathcal{L}|=l}} \mathbb{E}\left[N^{5-5p}\xi^\mu_{\mathcal{K}} \sum_{\substack{\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3,\mathcal{J}_4\in\mathcal{B} \\ \mathcal{J}_o=\mathcal{K},\mathcal{J}_e=\mathcal{L}}} D_{\mathcal{I}} D_{\mathcal{J}_1} D_{\mathcal{J}_2} D_{\mathcal{J}_3} D_{\mathcal{J}_4} f(0) \right] \\ \leq C_4 N^{2-2p}(N^{7-3p+l} + N^{5-5p+k+l}) \\ \leq C_4 N^{2-2p}(N^{10-3p} + N^{5-2p}) \leq C_4 N^{-2} N^{2-2p}. \end{aligned} \tag{6.57}$$

Inserting the bounds (6.53) and (6.57) into the decomposition (6.52) yields the upper bound. To get the lower bound, change the sign of the error term in (6.55). This proves Lemma 6.5. □

To finish the proof of Proposition 6.1, insert the bounds from Lemmas 6.2–6.5, together with the bounds (6.12)–(6.16) into (6.10). This gives the upper bound in (6.4).

The corresponding lower bound is obtained by changing the sign of the error term in (6.10). This concludes the proof of Proposition 6.1. □

Proposition 6.6: *There exists a constant $C > 0$ such that*

$$\left| \mathbb{E}\mathcal{G}^{\otimes 2}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}] - \mathbb{E}\mathcal{G}'^{\otimes 2}_{\mu,\mathcal{I}}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}] \right| \leq CN^{1-p}. \tag{6.58}$$

Proof: For the last time, we expand the Boltzmann factors in the variables $(x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}}$ (with $\mathcal{B}$ as in (6.2)). Let

$$g((x_{\mathcal{J}})_{\mathcal{J}\in\mathcal{B}}) = \mathcal{G}^{\otimes 2}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}]. \tag{6.59}$$

Then,

$$\begin{aligned}\mathbb{E}\mathcal{G}^{\otimes 2}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}] &\le \mathbb{E}\Big[g(0) + \sum_{\mathcal{J}\in\mathcal{B}} x_{\mathcal{J}} D_{\mathcal{J}} g(0) + C_1(\sum_{\mathcal{J}\in\mathcal{B}} x_{\mathcal{J}})^2\Big] \\ &= \mathbb{E}\mathcal{G}'^{\otimes 2}_{\mu,\mathcal{I}}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}] + 0 + C_2 N^{1-p}.\end{aligned} \tag{6.60}$$

On the other hand,

$$\begin{aligned}\mathbb{E}\mathcal{G}^{\otimes 2}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}] &\ge \mathbb{E}\Big[g(0) + \sum_{\mathcal{J}\in\mathcal{B}} x_{\mathcal{J}} D_{\mathcal{J}} g(0) - C_1(\sum_{\mathcal{J}\in\mathcal{B}} x_{\mathcal{J}})^2\Big] \\ &= \mathbb{E}\mathcal{G}'^{\otimes 2}_{\mu,\mathcal{I}}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}] + 0 - C_2 N^{1-p}.\end{aligned} \tag{6.61}$$

This proves the proposition. □

Proof of Theorem 2.5: By Proposition 6.1, resp. 6.6, there exist constants $C_1, C_2 > 0$ such that

$$\begin{aligned}&\left|\frac{1}{N^{p-1}}\mathbb{E}\mathcal{G}[\xi^\mu_{\mathcal{I}} s_{\mathcal{I}}] - \frac{\beta(p!)^{\frac12}}{N^{2p-2}} + \frac{\beta(p!)^{\frac12}}{N^{2p-2}}\mathbb{E}\mathcal{G}^{\otimes 2}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}]\right| \\ &\qquad \le \left|\frac{1}{N^{p-1}}\mathbb{E}\mathcal{G}[\xi^\mu_{\mathcal{I}} s_{\mathcal{I}}] - \frac{\beta(p!)^{\frac12}}{N^{2p-2}} + \frac{\beta(p!)^{\frac12}}{N^{2p-2}}\mathbb{E}\mathcal{G}'^{\otimes 2}_{\mu,\mathcal{I}}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}]\right| \\ &\qquad\quad + \frac{\beta(p!)^{\frac12}}{N^{2p-2}}\left|\mathbb{E}\mathcal{G}'^{\otimes 2}_{\mu,\mathcal{I}}[\sigma^1_{\mathcal{I}}s^2_{\mathcal{I}}] - \mathbb{E}\mathcal{G}^{\otimes 2}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}]\right| \\ &\qquad \le C_1 N^{1-2p} + C_2 N^{3-3p}.\end{aligned} \tag{6.62}$$

This implies that

$$N^{1-p}\left|\sum_{\mu,\mathcal{I}}\mathbb{E}\mathcal{G}[\xi^\mu_{\mathcal{I}}\sigma_{\mathcal{I}}] - \alpha\beta(p!)^{\frac12}\sum_{\mathcal{I}}\Big(1 - \mathbb{E}\mathcal{G}^{\otimes 2}[\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}]\Big)\right| \le C_3. \tag{6.63}$$

whence

$$\begin{aligned}-\beta\,\mathbb{E}\mathcal{G}[H] &= \frac{\beta(p!)^{\frac12}}{N^{p-1}}\mathbb{E}\mathcal{G}[\sum_{\mu,\mathcal{I}}\xi^\mu_{\mathcal{I}}\sigma_{\mathcal{I}}] \\ &\le \sum_{\substack{\mathcal{I}\subset\mathcal{N}\\|\mathcal{I}|=p}}\frac{\alpha\beta^2 p!}{N^{p-1}} - \frac{\alpha\beta^2 p}{N^{p-1}}\mathbb{E}\mathcal{G}^{\otimes 2}\Big[\sum_{\substack{\mathcal{I}\subset\mathcal{N}\\|\mathcal{I}|=p}}\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}\Big] + C_3 \\ &\le \alpha\beta^2 N - \alpha\beta^2 N\mathbb{E}\mathcal{G}^{\otimes 2}[R(\sigma^1,\sigma^2)^p] + C_4,\end{aligned} \tag{6.64}$$

respectively

$$\begin{aligned}-\beta\mathbb{E}\mathcal{G}[H] &= \frac{\beta(p!)^{\frac12}}{N^{p-1}}\mathbb{E}\mathcal{G}[\sum_{\mu,\mathcal{I}}\xi^\mu_{\mathcal{I}}\sigma_{\mathcal{I}}] \\ &\ge \sum_{\substack{\mathcal{I}\subset\mathcal{N}\\|\mathcal{I}|=p}}\frac{\alpha\beta^2 p!}{N^{p-1}} - \frac{\alpha\beta^2 p!}{N^{p-1}}\mathbb{E}\mathcal{G}^{\otimes 2}\Big[\sum_{\substack{\mathcal{I}\subset\mathcal{N}\\|\mathcal{I}|=p}}\sigma^1_{\mathcal{I}}\sigma^2_{\mathcal{I}}\Big] - C_3 \\ &\ge \alpha\beta^2 N - \alpha\beta^2 N\mathbb{E}\mathcal{G}^{\otimes 2}[R(\sigma^1,\sigma^2)^p] - C_5.\end{aligned} \tag{6.65}$$

Hence, since $\beta \mathbb{E} \frac{\partial F}{\partial \beta} = -\frac{\beta}{N} \mathbb{E} \mathcal{G}[H]$,

$$\left| \beta \mathbb{E} \frac{\partial F_N}{\partial \beta} - \alpha \beta^2 + \alpha \beta^2 \mathbb{E} \mathcal{G}^{\otimes 2} [R(\sigma^1, \sigma^2)^p] \right| \leq C_6 N^{-1}. \tag{6.66}$$

This proves Theorem 2.5. □

6.2 Condensation: Proof of Theorem 2.6

Theorem 2.6 is now a consequence of the convexity of the free energy. Suppose that $\beta < \beta_p$. Then

$$\limsup_{N \uparrow \infty} \mathbb{E} F_N = \frac{\alpha \beta^2}{2} \tag{6.67}$$

by the definition of β_p. As remarked after their definition in Chapter 2, $\mathbb{E} F_N$ is convex for all N. It then follows from a standard result in convex analysis ([Ro], Theorem 25.7) that

$$\limsup_{N \uparrow \infty} \mathbb{E} \frac{\partial F_N}{\partial \beta} = \frac{\partial}{\partial \beta} \limsup_{N \uparrow \infty} \mathbb{E} F_N = \alpha \beta. \tag{6.68}$$

Hence, from Theorem 2.5,

$$\mathbb{E} \mathcal{G}^{\otimes 2} [R^p] + \mathbb{E} \frac{\partial F_N}{\partial \beta} = \alpha \beta + \mathcal{O}(N^{-1}), \tag{6.69}$$

and thus, passing to the limit,

$$\limsup_{N \uparrow \infty} \mathbb{E} \mathcal{G}^{\otimes 2} [R^p] + \alpha \beta = \alpha \beta, \tag{6.70}$$

which in turn implies that

$$\mathbb{E} \mathcal{G}_N^{\otimes 2} [R^p] = 0. \tag{6.71}$$

Suppose now that

$$\limsup_{N \uparrow \infty} \mathbb{E} \frac{\partial F_N}{\partial \beta} < \alpha \beta. \tag{6.72}$$

Then it follows immediately from Theorem 2.5 that

$$\liminf_{N \uparrow \infty} \mathbb{E} \mathcal{G}^{\otimes 2} [R^p] = \alpha \beta - \limsup_{N \uparrow \infty} \mathbb{E} \frac{\partial F_N}{\partial \beta} > \alpha \beta - \alpha \beta = 0. \tag{6.73}$$

This proves (2.39). To see where the condition (6.72) actually holds, we observe first that by Lemma 4.10, it is satisfied for all

$$\beta > \hat{\beta}_p = \sqrt{\frac{2 \ln 2}{\alpha}}. \tag{6.74}$$

Furthermore, Theorem 5.5 in [Ro] implies that the function

$$f(\beta) = \limsup_{N \uparrow \infty} \mathbb{E} F_N \tag{6.75}$$

is a convex, bounded function on $\mathcal{U} = [0, \beta_p')$. By Theorem 25.3 in [Ro] it is thus differentiable on an open set $\mathcal{D} \subset \mathcal{U}$ which contains all but perhaps countably many points of $\mathcal{U}$, and its derivative f' is bounded on $\mathcal{D}$. Lebesgue's integrability criterion (see for instance [He], Theorem 199.3) then implies that

$$f(\beta) = f(\beta_p) + \int_{\beta_p}^{\beta} f'(u)du, \quad \forall \beta > \beta_p. \tag{6.76}$$

Now it is immediate that for all $\beta > \beta_p$ there must exist a set $I \subset (\beta_p, \beta)$ with strictly positive Lebesgue measure, on which f' is strictly less than $\alpha\beta$. Indeed, were this not the case, then $f \geq \frac{\alpha\beta^2}{2}$, which contradicts the definition of β_p.

Since β was arbitrary, the relevant condition (6.72) is satisfied on sets of positive Lebesgue measure arbitrarily close to β_p. □

6.3 Spin Glass Phase: Proof of Theorem 2.7

We first prove two auxiliary lemmas that estimate the value of the Hamiltonian in the vicinity of each pattern.

Lemma 6.7: *The Hamiltonian evaluated at the patterns satisfies*

$$\mathbb{P}\left[|H_N(\sigma = \xi^\mu)| \geq \frac{N}{(p!)^{\frac{1}{2}}} + zN\right] \leq C \begin{cases} e^{-\frac{z^2 N}{2\alpha}}, & \text{if } z \leq \beta_p', \\ e^{-\beta_p'(z - \frac{\alpha\beta_p'}{2})N}, & \text{otherwise.} \end{cases} \tag{6.77}$$

Proof: The Hamiltonian at the pattern ξ^μ is given by

$$\begin{aligned} H(\sigma = \xi^\mu) &= -\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\mathcal{I}} \xi^\mu_{\mathcal{I}} \xi^\mu_{\mathcal{I}} - \frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu\neq\mu} \sum_{\mathcal{I}} \xi^\nu_{\mathcal{I}} \xi^\mu_{\mathcal{I}} \\ &= -\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \binom{N}{p} - \frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu\neq\mu} \sum_{\mathcal{I}} \xi^\nu_{\mathcal{I}} \xi^\mu_{\mathcal{I}}, \end{aligned} \tag{6.78}$$

which implies that

$$-H_N(\xi^\mu) \leq \frac{N}{(p!)^{\frac{1}{2}}} + \frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu\neq\mu} \sum_{\mathcal{I}} \xi^\nu_{\mathcal{I}} \xi^\mu_{\mathcal{I}}. \tag{6.79}$$

We estimate the stochastic part in (6.79) with the same method used in the proof of Theorem 2.1. By Chebyshev's exponential inequality, independence of ξ^ν and ξ^μ (for $\nu \neq \mu$), and expansion of the exponential, we get for $z > 0$

$$\begin{aligned} \mathbb{P}\left[\left|\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu\neq\mu} \sum_{\mathcal{I}} \xi^\nu_{\mathcal{I}} \xi^\mu_{\mathcal{I}}\right| \geq zN\right] &\leq \inf_{t>0} e^{-tzN} \prod_{\nu\neq\mu} \mathbb{E}\left[\exp\left(t \frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\mathcal{I}} \xi^\nu_{\mathcal{I}} \xi^\mu_{\mathcal{I}}\right)\right] \\ &\leq \inf_{t>0} e^{-tzN} \prod_{\nu\neq\mu} \left(1 + \frac{t^2 p!}{2N^{2p-2}} \binom{N}{p}\right. \\ &\qquad \left. + \frac{t^3 (p!)^{\frac{3}{2}}}{3!\, N^{3p-3}} \mathbb{E}\left[\left|\sum_{\mathcal{I}} \xi^\nu_{\mathcal{I}}\right|^3 e^{t \frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \left|\sum_{\mathcal{I}} \xi^\nu_{\mathcal{I}}\right|}\right]\right). \end{aligned} \tag{6.80}$$

The error term can be written as

$$\frac{1}{N^{3p-3}}\mathbb{E}\left[\Big|\sum_{\mathcal{I}}\xi^\nu_{\mathcal{I}}\Big|^3 e^{t\frac{(p!)^{\frac12}}{N^{p-1}}|\sum_{\mathcal{I}}\xi^\nu_{\mathcal{I}}|}\right]=\frac{1}{N^{\frac{3p}{2}-3}}\mathbb{E}\left[\Big|N^{-\frac{p}{2}}\sum_{\mathcal{I}}\xi^\nu_{\mathcal{I}}\Big|^3 e^{t\frac{(p!)^{\frac12}}{N^{\frac{p}{2}-1}}|N^{-\frac{p}{2}}\sum_{\mathcal{I}}\xi^\nu_{\mathcal{I}}|}\right] \tag{6.81}$$

This latter term is exactly the same as in (3.2) (with β replaced by t). Hence, we get (compare (3.3))

$$\mathbb{P}\left[\frac{(p!)^{\frac12}}{N^{p-1}}\sum_{\nu\neq\mu}\sum_{\mathcal{I}}\xi^\nu_{\mathcal{I}}\xi^\mu_{\mathcal{I}}\geq zN\right]\leq\inf_{t\in(0,\beta'_p)}e^{-tzN+\frac{\alpha t^2N}{2}+C_1}. \tag{6.82}$$

Minimizing the exponent yields

$$\begin{aligned}\mathbb{P}\left[-H_N(\sigma=\xi^\mu)\geq\frac{N}{(p!)^{\frac12}}+zN\right]&\leq\mathbb{P}\left[\frac{(p!)^{\frac12}}{N^{p-1}}\sum_{\nu\neq\mu}\sum_{\mathcal{I}}\xi^\nu_{\mathcal{I}}\xi^\mu_{\mathcal{I}}>zN\right]\\&\leq C_2\begin{cases}e^{-\frac{z^2}{2\alpha}N}, & \text{if } 0<z\leq\alpha\beta'_p,\\ e^{-\beta'_p(z-\frac{\alpha\beta'_p}{2})N}, & \text{otherwise.}\end{cases}\end{aligned} \tag{6.83}$$

This proves the claim. □

The next result shows that the Hamiltonian does not fluctuate much around a pattern. This is in fact a result that was already proved by Newman [N1] for the Hamiltonian $\bar{H}$. Our case is even simpler. Define $B_\delta(\sigma)$ to be the $(N\delta)$-ball around the configuration σ in the Hamming distance. Then we have the following

Lemma 6.8: *If $\delta<\frac{1}{p}$, then there exists a constant $C>0$ such that*

$$\mathbb{P}\Big[\exists\sigma\in B_\delta(\xi^\mu):|H_N(\sigma)-H_N(\xi^\mu)|\geq(\frac{2^{p-1}}{(p!)^{\frac12}}\delta+z)N\Big]\leq Ce^{-N(f_\delta(z)+\delta\ln\delta+(1-\delta)\ln(1-\delta))}, \tag{6.84}$$

where

$$f_\delta(z)=\begin{cases}\frac{z^2}{2^p\alpha\delta}, & \text{if } z\leq 2^{p-1}\alpha\delta\beta'_p;\\ e^{-\beta'_pN(z-\frac{\alpha\beta'_p}{2(p-1)!})}, & \text{otherwise.}\end{cases} \tag{6.85}$$

Proof: By standard arguments (see also [N1], in particular inequality (2.3) and surrounding comments),

$$\begin{aligned}&\mathbb{P}\Big[\exists\sigma\in B_\delta(\xi^\mu):|H_N(\sigma)-H_N(\xi^\mu)|\geq(\delta+z)N\Big]\\&\qquad\leq\sum_{q=1}^{\lfloor\delta N\rfloor}\binom{N}{q}\mathbb{P}[|H_N(\zeta^q)-H_N(\xi^\mu)|\geq(\delta+z)N],\end{aligned} \tag{6.86}$$

where

$$\zeta^q_i=\begin{cases}-\xi^\mu_i, & \text{if } i\leq q;\\ \xi^\mu_i, & \text{if } i\geq q+1.\end{cases} \tag{6.87}$$

We start by calculating the difference $|H(\zeta^q) - H(\xi^\mu)|$. Let $\mathcal{J} = \mathcal{J}_q = \{1, \dots, q\}$. One obtains

$$\begin{aligned} H(\zeta^q) - H(\xi^\mu) &= -\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu=1}^{M(N)} \sum_{\mathcal{I}} (\zeta^q_{\mathcal{I}} \xi^\nu_{\mathcal{I}} - \xi^\mu_{\mathcal{I}} \xi^\nu_{\mathcal{I}}) \\ &= -\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu=1}^{M(N)} \sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\,\mathrm{odd}} (\zeta^q_{\mathcal{I}} \xi^\nu_{\mathcal{I}} - \xi^\mu_{\mathcal{I}} \xi^\nu_{\mathcal{I}}) \\ &= 2\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu=1}^{M(N)} \sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\,\mathrm{odd}} \xi^\mu_{\mathcal{I}} \xi^\nu_{\mathcal{I}} \\ &= 2\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\mathrm{odd}} 1 + 2\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu\neq\mu} \sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\,\mathrm{odd}} \xi^\mu_{\mathcal{I}} \xi^\nu_{\mathcal{I}} \end{aligned} \tag{6.88}$$

Explicitly, this is

$$H(\zeta^q) - H(\xi^\mu) = 2\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{r=1,\,\mathrm{odd}}^{p-1} \binom{N-q}{p-r}\binom{q}{r} + 2\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu\neq\mu} \sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\,\mathrm{odd}} \xi^\mu_{\mathcal{I}} \xi^\nu_{\mathcal{I}} \tag{6.89}$$

Let us treat the stochastic term in (6.89) first. By the usual procedure, we get

$$\begin{aligned} &\mathbb{P}\left[\left|\frac{(p!)^{\frac{1}{2}}}{N^{p-1}} \sum_{\nu\neq\mu} \sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\,\mathrm{odd}} \xi^\mu_{\mathcal{I}} \xi^\nu_{\mathcal{I}}\right| \geq zN\right] \\ &\quad\leq 2\inf_{t>0} e^{-tzN} \prod_{\nu\neq\mu} \left\{1 + \frac{t^2 p!}{2N^{2p-2}} \sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\,\mathrm{odd}} 1 \right. \\ &\qquad \left. + \frac{t^3 (p!)^{\frac{3}{2}}}{3!\, N^{3p-3}} \mathbb{E}\left[\left|\sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\,\mathrm{odd}} \xi^\mu_{\mathcal{I}} \xi^\nu_{\mathcal{I}}\right|^3 \exp\left(\frac{t(p!)^{\frac{1}{2}}}{N^{p-1}} \left|\sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\,\mathrm{odd}} \xi^\mu_{\mathcal{I}} \xi^\nu_{\mathcal{I}}\right|\right)\right]\right\} \\ &\quad\leq 2 \inf_{t\in(0,\beta'_p)} e^{-tzN} \prod_{\nu\neq\mu} \left\{1 + \frac{t^2 p!}{2N^{2p-2}} \sum_{r=1\,\mathrm{odd}}^{p-1} \binom{N-q}{p-r}\binom{q}{r} \right. \\ &\qquad \left. + C_1 N^{3-\frac{3p}{2}} \right\}. \end{aligned} \tag{6.90}$$

The last line follows from the usual bound on the error term (see the proof of Theorem 2.1 in Chapter 3; in fact, t can even be chosen somewhat larger than β'_p, since the sum over sets $\mathcal{I}$ contains fewer terms than we had there).

To treat products of binomial coefficients in last expression, observe that if $q \leq \lfloor \delta N \rfloor < \frac{N}{2}$, then the following inequality holds,

$$\begin{aligned} p! \sum_{r=1,\,\mathrm{odd}}^{p-1} \binom{N-q}{p-r}\binom{q}{r} &\leq \sum_{r=1,\,\mathrm{odd}}^{p-1} \binom{p}{r} (N-q)^{p-r} q^r \\ &\leq (N-q)^{p-1} q \sum_{r=1,\,\mathrm{odd}}^{p-1} \binom{p}{r} = 2^{p-1}(N-q)^{p-1} q. \end{aligned} \tag{6.91}$$

Using (6.91) in (6.90) yields

$$\begin{aligned} \mathbb{P}\Big[\Big|\frac{(p!)^{\frac{1}{2}}}{N^{p-1}}\sum_{\nu\neq\mu}\sum_{\mathcal{I}:|\mathcal{I}\cap\mathcal{J}|\text{ odd}}\xi^{\mu}_{\mathcal{I}}\xi^{\nu}_{\mathcal{I}}\Big|\geq zN\Big] \\ \leq 2\inf_{t\in(0,\beta'_p)}e^{-tzN}\exp\left(\frac{\alpha t^2}{2N^{p-1}}2^{p-1}(N-q)^{p-1}q+C_1\right). \end{aligned} \tag{6.92}$$

The deterministic term in (6.89) is given by (again using (6.91))

$$\begin{aligned} \frac{(p!)^{\frac{1}{2}}}{N^{p-1}}\sum_{r=1,\text{ odd}}^{p-1}\binom{N-q}{p-r}\binom{q}{r} &\leq \frac{1}{(p!)^{\frac{1}{2}}N^{p-1}}\sum_{r=1,\text{ odd}}^{p-1}\binom{p}{r}(N-q)^{p-r}q^r \\ &\leq \frac{2^{p-1}}{(p!)^{\frac{1}{2}}N^{p-1}}(N-q)^{p-1}q. \end{aligned} \tag{6.93}$$

If $\delta<\frac{1}{p}$, then the last line is bounded by the term for the maximum q. That is

$$\frac{(p!)^{\frac{1}{2}}}{N^{p-1}}\sum_{r=1,\text{ odd}}^{p-1}\binom{N-q}{p-r}\binom{q}{r}\leq\frac{2^{p-1}}{(p!)^{\frac{1}{2}}N^{p-1}}(N-\lfloor\delta N\rfloor)^{p-1}\lfloor\delta N\rfloor\leq\frac{2^{p-1}}{(p!)^{\frac{1}{2}}}N\delta. \tag{6.94}$$

Collecting (6.92) and (6.94), we get

$$\begin{aligned} \mathbb{P}\big[|H(\zeta^q)-H(\xi^\mu)|\geq\frac{2^{p-1}}{(p!)^{\frac{1}{2}}}\delta N+zN\big] \\ \leq 2\inf_{t\in(0,\beta'_p)}e^{-tzN}\exp\left(\frac{\alpha t^2}{2N^{p-1}}2^{p-1}(N-q)^{p-1}q^r+C_1\right). \end{aligned} \tag{6.95}$$

Plugging this into (6.86) gives

$$\begin{aligned} \mathbb{P}[\exists\sigma\in B_\delta(\xi^\mu):|H_N(\sigma)-H_N(\xi^\mu)|\geq(\frac{2^{p-1}}{(p!)^{\frac{1}{2}}}\delta+z)N] \\ \leq 2\sum_{q=1}^{\lfloor\delta N\rfloor}\binom{N}{q}\inf_{t\in(0,\beta'_p)}e^{-tzN}\exp\left(\frac{\alpha t^2}{2N^{p-1}}2^{p-1}(N-q)^{p-1}q+C_1\right). \end{aligned} \tag{6.96}$$

It is straightforward to check that under our assumptions on δ and for fixed t, the ratio between two consecutive terms in the above sum is larger than 2, and therefore the whole sum is at most twice the maximum term,

$$\begin{aligned} \mathbb{P}[\exists\sigma\in B_\delta(\xi^\mu):|H_N(\sigma)-H_N(\xi^\mu)|>(\frac{2^{p-1}}{(p!)^{\frac{1}{2}}}\delta+z)N] \\ \leq 4\binom{N}{\lfloor\delta N\rfloor}\inf_{t\in(0,\beta'_p)}e^{-tzN}\exp\left(\frac{2^{p-1}\alpha t^2}{2}N\delta+C_1\right). \end{aligned} \tag{6.97}$$

Minimizing with respect to t and using Stirling's formula for the binomial factor concludes the proof of Lemma 6.8. $\square$

Proof of Theorem 2.7: We observe the following elementary fact. By the definition of the free energy

$$F_N(\beta) \leq \frac{\beta}{N} \sup_\sigma |H_N(\sigma)|. \tag{6.98}$$

Hence, by Theorem 2.4, for any $\beta, z > 0$ there exists $C > 0$ such that

$$\mathbb{P}[\frac{1}{N} \sup_\sigma |H_N(\sigma)| < \frac{1}{\beta} \mathbb{E} F_N(\beta) - z] \leq \mathbb{P}[F_N(\beta) < \mathbb{E} F_N(\beta) - z] \leq CN^{-1} \tag{6.99}$$

Suppose that $\alpha\beta_p(\alpha) > \frac{1}{(p!)^{\frac{1}{2}}}$. Then there exists $\beta > 0$ such that

$$\alpha > \frac{1}{(p!)^{\frac{1}{2}}(\beta_p - \frac{\beta_p^2}{2\beta})}, \tag{6.100}$$

which is equivalent to

$$\frac{1}{(p!)^{\frac{1}{2}}} < \frac{1}{\beta}(\alpha\beta\beta_p - \frac{\alpha\beta_p^2}{2}) \leq \frac{1}{\beta} \mathbb{E} F_N(\beta) + C_1 N^{-1}. \tag{6.101}$$

The second inequality follows from the convexity of $F_N(\beta)$ (see (4.87)) and the definition of β_p. But then we can find $\delta \in (0, \frac{1}{p})$ and $z > 0$ such that (for all N sufficiently large)

$$\frac{2^{p-1}}{(p!)^{\frac{1}{2}}}\delta + 3z < \frac{1}{\beta} \mathbb{E} F_N(\beta) - \frac{1}{(p!)^{\frac{1}{2}}}, \tag{6.102}$$

and (with the definition of f_δ from Lemma 6.8)

$$f_\delta(z) + \delta \ln \delta + (1-\delta)\ln(1-\delta) > 0. \tag{6.103}$$

By Lemma 6.7, resp. 6.8, for any $m > 0$, we can find an $\bar{N} \in \mathbb{N}$ such that for all $N \geq \bar{N}$

$$\begin{aligned} \mathbb{P}[\exists \sigma \in \bigcup_{\mu=1}^{M(N)} B_\delta(\xi^\mu) : |H_N(\sigma)| \geq N(\frac{1}{(p!)^{\frac{1}{2}}} + \delta + 2z)] \\ \leq \mathbb{P}[\exists \sigma \in \bigcup_{\mu=1}^{M(N)} B_\delta(\xi^\mu) : |H_N(\sigma) - H_N(\xi^\mu)| \geq N(\delta + z)] \\ + \mathbb{P}[\sup_\mu |H_N(\xi^\mu)| \geq N(\frac{1}{(p!)^{\frac{1}{2}}} + z)] \\ \leq N^{-m}. \end{aligned} \tag{6.104}$$

On the other hand, the inequality (6.99) implies that

$$\mathbb{P}[\sup_\sigma |H_N(\sigma)| \leq N \frac{\mathbb{E} F_N(\beta)}{\beta} - zN] \leq CN^{-1}, \tag{6.105}$$

so that finally, by standard arguments,

$$\begin{aligned}\mathbb{P}[\arg\sup|H_N(\sigma)| \in \bigcup_{\mu=1}^{M(N)} B_\delta(\xi^\mu)] &\leq \mathbb{P}[\exists\sigma \in \bigcup_{\mu=1}^{M(N)} B_\delta(\xi^\mu) : |H_N(\sigma)| \geq N(\frac{1}{(p!)^{\frac{1}{2}}} + \delta + 2z)] \\ &\quad + \mathbb{P}[\sup_\sigma |H_N(\sigma)| \leq N\frac{\mathbb{E}\, F_N(\beta)}{\beta} - zN] \\ &\leq C'N^{-1},\end{aligned} \tag{6.106}$$

for all N large enough.

To show the existence of an α_{sp}, we observe that the bounds (2.25) and (2.26) on the critical β imply that the quantity $\alpha\beta_p(\alpha) \sim \sqrt{\alpha}$ and is thus eventually larger than any fixed number. This concludes the proof of Theorem 2.7. □

7 Proofs of the Results for the Second Interaction

7.1 Annealed Free Energy

The proof of Theorem 2.1' is almost identical to the calculation of the annealed free energy for the Hamiltonian H. In fact, since we deal with powers of sums of Bernoulli variables, we do not have to prove an equivalent to Lemma 3.1, but can use directly the standard bounds on empirical averages of Bernoulli variables.

Proof of Theorem 2.1': We first observe the following two facts. Since the ξ_i^μ are mutually independent, the expectation of the exponential decouples into the product of the expectations. Furthermore, the distribution of $Y_N^\mu = (\frac{1}{\sqrt{N}}\sum_i \xi_i^\mu \sigma_i)^p$ is equal to $X_N = Y_N^1 = (\frac{1}{\sqrt{N}}\sum_i \xi_i^1)^p$. We can thus write

$$\begin{aligned} \mathbb{E}[\exp(-\beta \bar{H}_N[\omega](\sigma))] &= \prod_{\mu=1}^{M(N)} \mathbb{E}\Big[\exp(\frac{\beta N^{1-\frac{p}{2}}}{s_p}\bar{Y}_N^\mu[\omega](\sigma)\Big] \\ &= \prod_{\mu=1}^{M(N)} \mathbb{E}\Big[\exp(\frac{\beta N^{1-\frac{p}{2}}}{s_p}\bar{X}_N)\Big], \end{aligned} \tag{7.1}$$

where $\bar{X}_N = X_N - \mathbb{E}X_N$.

In each factor of the left-hand side of (7.1) we apply the usual bound for the exponential function, and obtain

$$\begin{aligned} \Big|\mathbb{E}\Big[\exp\Big(\frac{\beta N^{1-\frac{p}{2}}}{s_p}\bar{X}_N\Big)\Big] - 1 - \frac{\beta N^{1-\frac{p}{2}}}{s_p}\mathbb{E}[\bar{X}_N] - \frac{(\beta N^{1-\frac{p}{2}})^2}{2s_p^2}\mathbb{E}[\bar{X}_N{}^2]\Big| \\ \leq \frac{(\beta N^{1-\frac{p}{2}})^3}{3!\, s_p^3}\mathbb{E}\big[|\bar{X}_N|^3\exp(\frac{\beta N^{1-\frac{p}{2}}}{s_p}|\bar{X}_N|)\big]. \end{aligned} \tag{7.2}$$

The linear term on the left-hand side is equal to zero, since $\bar{X}_N$ is centered. The quadratic term is equal to the variance of X_N,

$$\mathbb{E}[(X_N - \mathbb{E}X_N)^2] = (2p-1)!!(1+\mathcal{O}(N^{-1})) = s_p^2(1+\mathcal{O}(N^{-1})). \tag{7.3}$$

We bound the error term on the right-hand side of (7.2). Let us show that

$$\mathbb{E}[|\bar{X}_N|^3\exp(\frac{\beta N^{1-\frac{p}{2}}}{s_p}|\bar{X}_N|)] \leq C \tag{7.4}$$

uniformly in N. Since $|\bar{X}_N| \leq |X_N| \leq N^{p/2}$, we can write

$$|\bar{X}_N| = |\bar{X}_N|^{\frac{2}{p}}|\bar{X}_N|^{\frac{p-2}{p}} \leq |\bar{X}_N|^{\frac{2}{p}}N^{\frac{p-2}{2}}. \tag{7.5}$$

Then we see that the powers of N cancel in the exponent of (7.4). Moreover, $|\bar{X}_N|^{\frac{2}{p}}$ behaves "almost like" the square of a Gaussian variable. To make this precise, use (7.5) in the exponent of (7.4), which leads to

$$\mathbb{E}\Big[|\bar{X}_N|^3\exp\big(\frac{\beta N^{1-\frac{p}{2}}}{s_p}|\bar{X}_N|\big)\Big] \leq \mathbb{E}\Big[|\bar{X}_N|^3\exp\big(\frac{\beta}{s_p}|\bar{X}_N|^{\frac{2}{p}}\big)\Big]. \tag{7.6}$$

To bound the last expectation uniformly in N, we cut it into two pieces with the help of a cut-off γ, which satisfies

$$\gamma > (2p-1)!! = \limsup_N \mathbb{E}\, X_N > 0, \tag{7.7}$$

and is independent of N. Thus,

$$\begin{aligned}\mathbb{E}\Big[|\bar X_N|^3 \exp\Big(\frac{\beta}{s_p}|\bar X_N|^{\frac{2}{p}}\Big)\Big] &= \mathbb{E}\Big[|\bar X_N|^3 \exp\Big(\frac{\beta}{s_p}|\bar X_N|^{\frac{2}{p}}\Big)\mathbb{1}_{\{|\bar X_N|\le\gamma\}}\Big]\\ &\quad + \mathbb{E}\Big[|\bar X_N|^3 \exp\Big(\frac{\beta}{s_p}|\bar X_N|^{\frac{2}{p}}\Big)\mathbb{1}_{\{|\bar X_N|\ge\gamma\}}\Big].\end{aligned} \tag{7.8}$$

The first term is easily seen to be bounded by a constant uniformly in N. We therefore turn to the second. We partition $[\gamma,\infty)$ into pieces of length l. Clearly, since the function $|x|^3\exp(\beta|x|^{2/p})$ is increasing in $|x|$,

$$\begin{aligned}&\mathbb{E}\Big[|\bar X_N|^3 \exp\Big(\frac{\beta}{s_p}|\bar X_N|^{\frac{2}{p}}\Big)\mathbb{1}_{\{|\bar X_N|\ge\gamma\}}\Big]\\ &\qquad = \sum_{i\ge 0}\mathbb{E}\Big[|\bar X_N|^3 \exp\Big(\frac{\beta}{s_p}|\bar X_N|^{\frac{2}{p}}\Big)\mathbb{1}_{\{|\bar X_N|\in[\gamma+il,\gamma+il+l)\}}\Big]\\ &\qquad \le \sum_{i\ge 0}|\gamma+il+l|^3\exp\Big(\frac{\beta}{s_p}|\gamma+il+l|^{\frac{2}{p}}\Big)\mathbb{P}\Big[|\bar X_N|\in[\gamma+il,\gamma+il+l)\Big]\\ &\qquad \le \sum_{i\ge 0}|\gamma+il+l|^3\exp\Big(\frac{\beta}{s_p}|\gamma+il+l|^{\frac{2}{p}}\Big)\mathbb{P}\Big[|\bar X_N|\ge\gamma+il\Big],\end{aligned} \tag{7.9}$$

if the limit on the right exists. To show this, we bound the probability appearing in the last expression by

$$\begin{aligned}\mathbb{P}\Big[|\bar X_N|\ge\gamma+il\Big] &= \mathbb{P}\Big[|X_N-\mathbb{E}X_N|\ge\gamma+il\Big]\\ &= \mathbb{P}\Big[X_N\ge\gamma+il+\mathbb{E}X_N\Big]+\mathbb{P}\Big[X_N\le-\gamma-il+\mathbb{E}X_N\Big].\end{aligned} \tag{7.10}$$

Since $\gamma > \mathbb{E}\,X_N$ for all N, the second probability is zero for all $i\ge 0$ (remember that $X_N\ge 0$). Then, by the usual exponential bound for sums of i.i.d. Bernoulli variables,

$$\begin{aligned}\mathbb{P}\Big[X_N\ge\gamma+il+\mathbb{E}X_N\Big] &= \mathbb{P}\Big[\frac{1}{\sqrt N}\sum_{i=1}^N\xi_i^1 > (\gamma+il+\mathbb{E}X_N)^{\frac{1}{p}}\Big]\\ &\le \exp\Big(-\frac{1}{2}(\gamma+il+\mathbb{E}X_N)^{\frac{2}{p}}\Big).\end{aligned} \tag{7.11}$$

Using (7.11) in (7.9), we obtain

$$\begin{aligned}&\mathbb{E}\Big[|\bar X_N|^3 \exp\Big(\frac{\beta}{s_p}|\bar X_N|^{\frac{2}{p}}\Big)\mathbb{1}_{\{|\bar X_N|\ge\gamma\}}\Big]\\ &\quad\le \sum_{i\ge 0}|\gamma+il+l|^3\exp\Big(\frac{\beta}{s_p}|\gamma+il+l|^{\frac{2}{p}}-\frac{1}{2}(\gamma+il+\mathbb{E}\,X_N)^{\frac{2}{p}}\Big)\\ &\quad\le \sum_{i\ge 0}(\gamma+il+l)^3\exp\Big((\gamma+il)^{\frac{2}{p}}\Big(\frac{\beta}{s_p}(1+\frac{l}{\gamma+il})^{\frac{2}{p}}-\frac{1}{2}(1+\frac{\mathbb{E}X_N}{\gamma+il})^{\frac{2}{p}}\Big)\Big).\end{aligned} \tag{7.12}$$

Choose $l < \mathbb{E}X_N$. If $\beta < \frac{s_p}{2}$, then the second factor in the exponent is negative. The series is thus summable, and its value is some constant C depending on γ, l, but not on N. Using (7.12) in (7.8) shows that (7.4) holds indeed uniformly in N.

To finish the proof, we use (7.3) and (7.4) in (7.2) and obtain

$$\left|\mathbb{E}\left[\exp\left(\frac{\beta N^{1-\frac{p}{2}}}{s_p}\bar{X}_N\right)\right] - 1 - \frac{(\beta N^{1-\frac{p}{2}})^2}{2}\right| \leq \frac{\beta^3 N^{3-\frac{3p}{2}}}{3!}C. \tag{7.13}$$

We now proceed as in (3.3) and (3.4) which concludes the proof of Theorem 2.1'. □

We also have the analogue of Corollary 3.2.

Corollary 7.1: *Assume that $\beta < \bar{\beta}'_p = \frac{s_p}{2}$. Then following holds for all large enough N*

$$\mathbb{E}[\exp(\beta H_N[\omega](\sigma))] = e^{\alpha N \frac{\beta^2}{2}(1+\mathcal{O}(N^{-1}))}. \tag{7.14}$$

Proof: Completely analogous to the proof of Theorem 2.1'. □

7.2 Critical Temperature

Here again, the proof is almost identical to the Hamiltonian H. We start by stating and proving the analogues of the results in Chapter 4.1.

Lemma 7.2: *If $\beta < \frac{s_p}{4} = \frac{\bar{\beta}'_p}{2}$, then*

$$\mathbb{E}e^{-\beta\bar{H}_N(\sigma)-\beta\bar{H}_N(\sigma)} = e^{\alpha N\beta^2(1+g_p(R(\sigma,\sigma')))(1+\mathcal{O}(N^{-1}))}. \tag{7.15}$$

Observe that again the variance and the covariance of the Hamiltonian appear in the exponent.

Proof: We start in the same way as in the proof of Theorem 2.1'. The expectation of the exponential can be decomposed into a product over all patterns μ. We can therefore restrict our attention to one generic factor. Using the same bound for the exponential function as in equation (7.2) , we get

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\frac{\beta N^{1-\frac{p}{2}}}{s_p}(\bar{Y}^\mu_N(\sigma)+\bar{Y}^\mu_N(\sigma'))\right)\right] \\
&\quad\leq 1 + \frac{\beta N^{1-\frac{p}{2}}}{s_p}\mathbb{E}\left[(\bar{Y}^\mu_N(\sigma)+\bar{Y}^\mu_N(\sigma'))\right] + \frac{(\beta N^{1-\frac{p}{2}})^2}{2s_p^2}\mathbb{E}\left[(\bar{Y}^\mu_N(\sigma)+\bar{Y}^\mu_N(\sigma'))^2\right] \\
&\qquad + \frac{(\beta N^{1-p/2})^3}{3!\,s_p^3}\mathbb{E}\left[\left|\bar{Y}^\mu_N(\sigma)+\bar{Y}^\mu_N(\sigma')\right|^3\exp\left(\frac{\beta N^{1-\frac{p}{2}}}{s_p}|\bar{Y}^\mu_N(\sigma)+\bar{Y}^\mu_N(\sigma')|\right)\right].
\end{aligned} \tag{7.16}$$

The linear term in the above expression is zero. For the quadratic term we get

$$\frac{1}{s_p^2}\mathbb{E}\left[(\bar{Y}^\mu_N(\sigma)+\bar{Y}^\mu_N(\sigma'))^2\right] \leq \mathbb{E}\left[\frac{\bar{Y}^\mu_N(\sigma)^2}{s_p^2}\right] + \mathbb{E}\left[\frac{\bar{Y}^\mu_N(\sigma')^2}{s_p^2}\right] + 2\mathbb{E}\left[\frac{1}{s_p^2}\bar{Y}^\mu_N(\sigma)\bar{Y}^\mu_N(\sigma')\right]. \tag{7.17}$$

The first and second term on the right-hand side in the above inequality are given by the variance of $\bar{H}_N$, that is, they are each equal to one. The third term is equal to $2g_p(R(\sigma,\sigma'))$.

Finally, the error term is treated as in the proof of Lemma 4.1. Namely, using Cauchy-Schwarz to separate the terms containing respectively σ and σ', we get

$$\begin{aligned}
&\mathbb{E}\left[\left|\bar{Y}_N^\mu(\sigma)+\bar{Y}_N^\mu(\sigma')\right|^3 \exp\left(\frac{\beta N^{1-\frac{p}{2}}}{s_p}|\bar{Y}_N^\mu(\sigma)+\bar{Y}_N^\mu(\sigma')|\right)\right] \\
&\quad\le \sum_{j=0}^{3}\binom{3}{j}\mathbb{E}\left[\left|\bar{Y}_N^\mu(\sigma)\right|^j\left|\bar{Y}_N^\mu(\sigma')\right|^{3-j}\exp\left(\frac{\beta N^{1-\frac{p}{2}}}{s_p}\left(\left|\bar{Y}_N^\mu(\sigma)\right|+\left|\bar{Y}_N^\mu(\sigma')\right|\right)\right)\right] \\
&\quad\le \sum_{j=0}^{3}\binom{3}{j}\left(\mathbb{E}\left[\left|\bar{Y}_N^\mu(\sigma)\right|^{2j}\exp\left(\frac{2\beta N^{1-\frac{p}{2}}}{s_p}\left|\bar{Y}_N^\mu(\sigma)\right|\right)\right]\right)^{1/2} \\
&\qquad\times\left(\mathbb{E}\left[\left|\bar{Y}_N^\mu(\sigma')\right|^{6-2j}\exp\left(\frac{2\beta N^{1-\frac{p}{2}}}{s_p}\left|\bar{Y}_N^\mu(\sigma')\right|\right)\right]\right)^{1/2}.
\end{aligned} \tag{7.18}$$

The factors appearing on the right-hand side of (7.18) can now be bounded uniformly in σ and N by some constant C_2, if $2\beta < \frac{s_p}{2} = \bar{\beta}_p$ (by exactly the same argument as used in the proof of Theorem 2.1'). To conclude, we use the standard bounds for $\ln x$ as in the proof of Theorem 2.1 (see relation (4.5)). □

Define the truncated partition function $\tilde{Z}$ by (the bar over $\tilde{Z}$ will be omitted for typographic reasons)

$$\tilde{Z}_{N,\beta}(c) \equiv \mathbb{E}_\sigma\left[e^{-\beta\bar{H}_N[\omega](\sigma)}\mathbb{1}_{\{-\beta\bar{H}_N(\sigma)\le c\alpha\beta^2N\}}\right]. \tag{7.19}$$

Again, this truncation has no influence on the expected value by the following result whose proof is completely similar to the proof of Lemma 4.2 is therefore omitted.

Lemma 7.3: *For all $\beta>0$, $c>1$, such that $\beta c<\bar{\beta}'_p$, there exist constants $K,K'>0$ such that*

$$\begin{aligned}
\mathbb{E}\tilde{Z}_{N,\beta} = \mathbb{E}\mathbb{E}_\sigma\left[e^{-\beta\bar{H}_N(\sigma)}\mathbb{1}_{\{-\beta\bar{H}_N(\sigma)\le c\alpha\beta^2N\}}\right] &\ge (1-Ke^{-(c-1)^2K'N})\mathbb{E}\mathbb{E}_\sigma\left[e^{-\beta\bar{H}_N(\sigma)}\right] \\
&\ge (1-Ke^{-(c-1)^2K'N})\mathbb{E}\bar{Z}_N.
\end{aligned} \tag{7.20}$$

Proof: See the proof of Lemma 4.2. □

The square of the truncated partition function is decomposed as in (4.16). Let

$$S(b) \equiv \mathbb{E}_{\sigma,\sigma'}\left[e^{-\beta(\bar{H}_N(\sigma)+\bar{H}_N(\sigma'))}\mathbb{1}_{\{|R(\sigma,\sigma')|<b\}}\right], \tag{7.21}$$

and

$$T(c,b,b') \equiv \mathbb{E}_{\sigma,\sigma'}\left[e^{-\beta(\bar{H}_N(\sigma)+\bar{H}_N(\sigma'))}\mathbb{1}_{\{|R(\sigma,\sigma')|\in[b,b']\}}\mathbb{1}_{\{-\beta\bar{H}_N(\sigma)-\beta\bar{H}_N(\sigma')\le 2c\alpha\beta^2N\}}\right], \tag{7.22}$$

where explicit reference to the system size N is omitted. Then,

$$\tilde{Z}_{N,\beta}(c)^2 \leq S(b) + T(c,b,1), \tag{7.23}$$

for all $b \in (0,1)$. We now control each of these terms separately, starting with $S(b)$. It is here that we first encounter the difficulties brought about by the 2-spin interactions which are contained in $\bar{H}$. Namely, since the covariance $g_p(\mathbb{R})$ contains a quadratic term, we have to add an extra condition on β in the analogue of Lemma 4.3. The result is then as follows.

Lemma 7.4: *Suppose*

$$\beta < \min\left(\frac{\bar{\beta}_p'}{2}, s_p\sqrt{\frac{2}{\alpha p^2(p-1)^2(p-3)!!}}\right), \tag{7.24}$$

and b is such that

$$\gamma(b) = \alpha\beta^2\frac{g_p(b)}{b^2} < \frac{1}{2}. \tag{7.25}$$

Then for all $\varepsilon \in (0, \frac{1}{2}-\gamma)$ there exists $N_\varepsilon \in \mathbb{N}$ such that for all $N > N_\varepsilon$,

$$\mathbb{E}S(b) \leq \frac{1}{\sqrt{1-2(\gamma+\varepsilon)}}e^{\alpha\beta^2 N}. \tag{7.26}$$

Remark: The second condition on β in (7.24) assures that there exists a b verifying (7.25). In fact, since g_p contains a quadratic term, for (7.25) to hold the coefficient of this term in g_p has to be strictly less than $\frac{1}{2\alpha\beta^2}$. This is the second condition in (7.24). If this is the case, then one can always choose b small enough, such that the higher order terms are less than any positive number. We show the first steps of the proof, and will leave the rest, which is completely similar to the proof of Lemma 4.3, to the reader.

Proof: If β satisfies the above condition, we can apply Lemma 7.1 to the integrand of the right-hand side of (7.21). This gives

$$\mathbb{E}\left[e^{-\beta(\bar{H}_N(\sigma)+\bar{H}_N(\sigma'))}\mathbb{1}_{\{|R(\sigma,\sigma')|<b\}}\right] \leq \mathbb{1}_{\{|R(\sigma,\sigma')|<b\}}e^{\alpha\beta^2 N(1+g_p(R(\sigma,\sigma'))+CN^{-1})}. \tag{7.27}$$

Thus,

$$\begin{aligned}\mathbb{E}S(b) &\leq \mathbb{E}_{\sigma,\sigma'}\left[e^{\alpha\beta^2 N(1+g_p(R(\sigma,\sigma'))+CN^{-1})}\mathbb{1}_{\{|R(\sigma,\sigma')|<b\}}\right]\\ &\leq \mathbb{E}_{\sigma,\sigma'}\left[e^{\alpha\beta^2 N(1+R(\sigma,\sigma')^2 g_p(b)/b^2+CN^{-1})}\mathbb{1}_{\{|R(\sigma,\sigma')|<b\}}\right]\\ &= e^{\alpha\beta^2 N}\mathbb{E}_{\sigma,\sigma'}\left[e^{\alpha\beta^2 N(R(\sigma,\sigma')^2 g_p(b)/b^2+CN^{-1})}\right].\end{aligned} \tag{7.28}$$

If b is small enough such that the condition (7.25) is satisfied (and by the preceding remark, this exists always if β satisfies the second bound in (7.24)), then for all $\varepsilon > 0$, there exists $N_\varepsilon \in \mathbb{N}$ such that

$$\alpha\beta^2 N\big(R(\sigma,\sigma')^2\frac{g_p(b)}{b^2} + CN^{-1}\big) < (\gamma(b)+\varepsilon)NR(\sigma,\sigma')^2, \tag{7.29}$$

for all $N \geq N_\varepsilon$.

We are now in the setting of (4.21). Since the rest of the proof of Lemma 4.3 did not depend on H anymore (but only on the a priori measure $\mathbb{P}$), we can simply follow the remaining steps (4.22)–(4.24). This proves Lemma 7.4. $\Box$

We now turn to the term $T(c, b, 1)$ in (7.23).

Lemma 7.5: *Let $I(t)$ be the Cramèr Entropy as defined in (2.24). Suppose that there exist $c > 1$, $d > 0$, such that*

$$\forall t \in [b, b'], \quad 2\alpha\beta^2 c\Big(1 - \frac{c}{2(1+g_p(t))}\Big) \leq \alpha\beta^2 + I(t) - d. \tag{7.30}$$

Then, if

$$c < \min\big(\frac{\bar{\beta}'_p}{2\beta}, 1 + g_p(b)\big), \tag{7.31}$$

there exists $\bar{N} \in \mathbb{N}$ such that for all $N \geq \bar{N}$,

$$\mathbb{E}T(c, b, b') \leq e^{\alpha\beta^2 N} e^{-\frac{Nd}{2}}. \tag{7.32}$$

Proof: Completely analogous to the proof of Lemma 4.4 (just change the covariance terms, that is, replace every t^p by $g_p(t)$). $\Box$

The preceding lemmata imply the following result.

Proposition 7.6: *Suppose that $\beta < \check{\beta}_p$. Then there exist constants $C > 0$ and $c > 1$ such that*

$$\mathbb{E}[\tilde{Z}_{N,\beta}(c)^2] \leq C(\mathbb{E}\tilde{Z}_{N,\beta}(c))^2. \tag{7.33}$$

Furthermore,

$$\mathbb{P}[\tilde{Z}_{N,\beta}(c) > \frac{1}{2}\mathbb{E}\tilde{Z}_{N,\beta}(c)] \geq \frac{3}{4C}. \tag{7.34}$$

Proof: The proof is exactly the same as the proof of Proposition 4.5. All we have to check is that the additional condition on β in Lemma 7.4 is satisfied. Indeed, it is easily verified that

$$\check{\beta}_p < s_p\sqrt{\frac{2}{\alpha p^2(p-1)^2(p-3)!!}}, \tag{7.35}$$

for all values of $p \geq 4$. In particular, for $p \geq 6$, the right-hand side is even greater than the upper bound $\hat{\beta} = \sqrt{\frac{2\ln 2}{\alpha}}$. If $p = 4$, the right-hand side in (7.35) is equal to $\sqrt{\frac{4}{3}}$. However,

$$I(t)\frac{1+g_p(t)}{g_p(t)}\Big|_{t=\frac{1}{2}} < \frac{4}{3}, \tag{7.36}$$

which shows that the infimum certainly satisfies the condition (7.35). $\Box$

Finally, the proof of the lower bound on the critical β is easier than for the interaction H. This is due to the fact that $\bar{H}$ is a convex function of the disorder variables ξ, which allows us to use a strong concentration of measure result due to Talagrand instead of Corollary A.4.

Proof of Theorem 2.2', lower bound: By Proposition 7.6, if $\beta < \check{\beta}_p$, then,

$$\mathbb{P}\,[\tilde{Z}_{N,\beta}(c) > \frac{1}{2}\mathbb{E}\,\tilde{Z}_{N,\beta}(c)] \geq C_1. \tag{7.37}$$

On the other hand, by Theorem 2.1', if

$$\mathbb{E}\,\bar{F}_N < \alpha\beta^2/2 - N^{k-1} \tag{7.38}$$

infinitely often for some $k > \frac{1}{2}$, then (compare to (4.47))

$$\mathbb{P}\left[\tilde{Z}_N \geq \frac{1}{2}\mathbb{E}\,\tilde{Z}_N\right] \leq \mathbb{P}\left[\bar{F}_N - \mathbb{E}\,\bar{F}_N \geq N^{k-1} - C_2 N^{-1}\right]. \tag{7.39}$$

We now apply the following concentration of measure result, which is proved as Theorem 6.6 in [T1].

Theorem 7.7: (Talagrand) *Consider a real-valued function f defined on $[-1,1]^N$. We assume that, for each real number a, the set $\{f \leq a\}$ is convex. Consider a convex set $B \subset [-1,1]^N$, consider $\sigma > 0$ and assume that the restriction of f to B has a Lipschitz constant at most σ, that is*

$$\forall x, y \in B, |f(x) - f(y)| \leq \sigma\|x - y\|, \tag{7.40}$$

where $\|\cdot\|$ denotes the Euclidean norm. Consider independent random variables $(X_i)_{i\in\mathbb{N}}$ valued in $[-1,+1]$ and consider the random variable

$$h = f(X_1, \ldots, X_N). \tag{7.41}$$

Then, if M is a Median of h, we have, for all $t > 0$, that

$$\mathbb{P}[|h - M| > t] \leq 4b + \frac{4}{1-2b}\exp\left(-\frac{t^2}{16\sigma^2}\right), \tag{7.42}$$

where

$$b = \mathbb{P}[(X_1, \ldots, X_N) \in B^c] < \frac{1}{2}. \tag{7.43}$$

We show that $\bar{F}_N$ is indeed a convex function of the disorder variables. We start by proving that $-\bar{H}_N$ is convex. With abuse of notation, we identify ω with $\xi[\omega]$, and get for any fixed configuration σ

$$\begin{aligned} &-\left(\lambda\bar{H}_N[\omega](\sigma) + (1-\lambda)\bar{H}_N[\omega'](\sigma)\right) \\ &\quad = -N\sum_{\mu=1}^{M(N)} \mathbb{E}\,m^\mu(\sigma)^p + N\sum_{\mu=1}^{M(N)}\left(\lambda m^\mu[\omega](\sigma)^p + (1-\lambda)m^\mu[\omega'](\sigma)^p\right). \end{aligned} \tag{7.44}$$

Since p is even, the p^{th} power is convex function. Furthermore, $m^\mu[\omega](\sigma)$ is a linear function of ω. Hence,

$$\begin{aligned}
-&\big(\lambda \bar{H}_N[\omega](\sigma) + (1-\lambda)\bar{H}_N[\omega'](\sigma)\big) \\
&\leq -N \sum_{\mu=1}^{M(N)} \mathbb{E} m^\mu(\sigma)^p + N \sum_{\mu=1}^{M(N)} \big(\lambda m^\mu[\omega](\sigma) + (1-\lambda) m^\mu[\omega'](\sigma)\big)^p \\
&= -N \sum_{\mu=1}^{M(N)} \mathbb{E} m^\mu(\sigma)^p + N \sum_{\mu=1}^{M(N)} \big(m^\mu[\lambda\omega + (1-\lambda)\omega'](\sigma)\big)^p \\
&= -\bar{H}_N[\lambda\omega + (1-\lambda)\omega'](\sigma).
\end{aligned} \tag{7.45}$$

By its definition, $N\bar{F}_N[\omega] = \ln \mathbb{E}_\sigma \, e^{-\beta \bar{H}_N[\omega](\sigma)}$. Applying (7.45) and Hölder's inequality hence yields

$$\begin{aligned}
N\bar{F}_N[\lambda\omega + (1-\lambda)\omega'] &= \ln \mathbb{E}_\sigma \, e^{-\beta \bar{H}_N[\lambda\omega + (1-\lambda)\omega'](\sigma)} \\
&\leq \ln \mathbb{E}_\sigma \, e^{-\beta\lambda \bar{H}_N[\omega] - \beta(1-\lambda)\bar{H}_N[\omega']} \\
&\leq \ln \left\{ \left(\mathbb{E}_\sigma \, e^{-\beta \bar{H}_N[\omega]}\right)^\lambda \left(\mathbb{E}_\sigma \, e^{-\beta \bar{H}_N[\omega']}\right)^{1-\lambda} \right\} \\
&= \lambda N \bar{F}_N[\omega] + (1-\lambda) N \bar{F}_N[\omega'].
\end{aligned} \tag{7.46}$$

This shows that $\bar{F}_N$ is indeed convex in ω.

Moreover, the existence of a set B with the desired properties of the Lipschitz norm of $\bar{F}_N$ can be easily deduced from Lemma 4.8. Indeed, in the proof of that result, we actually wrote H_N as a sum of Hamiltonians $\bar{H}_N$ for different values of p and showed that for each of these contribution, the desired bound on the Lipschitz norm holds.

It is also well known [MS, Le] that in the case where such a concentration of measure result holds, median and expected value of a function are very close to each other (in our case, one shows easily that their difference is at most of the order $N^{-1/2}$).

Thus, Theorem 7.7 implies that that the right-hand side of (7.39) vanishes exponentially in N. This contradicts (7.37), and we therefore reject the assumption (7.38). This proves the lower bound (2.29). □

Proof of Theorem 2.2', upper bound: Inspection of the proof for the Hamiltonian H_N shows that only Theorem 2.1 and Corollary 3.2 enter. In the present case, we simply replace them by Theorem 2.1' respectively Corollary 7.1. Since each of them contains exactly the same conclusion as its counterpart for H_N, we conclude directly that the upper bound is the same as for the interaction H_N. □

7.3 Fluctuations of the Free Energy

Before we start with two preparatory lemmas, we define the following random matrices. For $j = 1, \ldots, p-1$, let

$$A_j \equiv A_j[\omega] \equiv \big(A_j^{\mu\nu}[\omega]\big)_{\mu,\nu=1,\ldots,M(N)} \quad , \tag{7.47}$$

with matrix elements

$$A_j^{\mu\nu}[\omega] \equiv M^{-1} \sum_{i_1,\dots,i_j=1}^{N} \xi_{i_1}^{\mu} \cdot \dots \cdot \xi_{i_j}^{\mu} \xi_{i_1}^{\nu} \cdot \dots \cdot \xi_{i_j}^{\nu} \tag{7.48}$$

Lemma 7.8: *For any $\omega, \omega' \in \Omega$,*

$$\bar{H}_N[\omega'](\sigma) - \bar{H}_N[\omega](\sigma) = \sum_{i_1,\dots,i_p=1}^{N} w_{i_1,\dots,i_p} \sigma_{i_1} \cdot \dots \cdot \sigma_{i_p}, \tag{7.49}$$

with

$$\sum_{i_1,\dots,i_p=1}^{N} w_{i_1,\dots,i_p}{}^2 \leq C \sum_{j=1}^{p-1} N^{1-p} \|A_j\|_{op} \|\xi[\omega] - \xi[\omega']\|_2^{p-j}, \tag{7.50}$$

for some $C > 0$.

Proof: Let $\bar{\xi}_i^{\mu} \equiv \xi_i^{\mu} - \xi'^{\mu}_i$. The decomposition (7.49) is nothing more than some elementary algebra. Indeed, the definition of H_N imply

$$\begin{aligned} \bar{H}_N[\omega'](\sigma) - \bar{H}_N[\omega](\sigma) &= N^{1-p} \sum_{\mu=1}^{M(N)} \left\{ \left(\sum_{i=1}^{N} \xi_i^{\mu} \sigma_i + \sum_{i=1}^{N} \bar{\xi}_i^{\mu} \sigma_i \right)^p - \left(\sum_{i=1}^{N} \xi_i^{\mu} \sigma_i \right)^p \right\} \\ &= N^{1-p} \sum_{\mu=1}^{M(N)} \sum_{j=0}^{p-1} \binom{p}{j} \left(\sum_i \xi_i^{\mu} \sigma_i \right)^j \left(\sum_i \bar{\xi}_i^{\mu} \sigma_i \right)^{p-j} \\ &= N^{1-p} \sum_{\mu=1}^{M(N)} \sum_{j=0}^{p-1} \binom{p}{j} \sum_{i_1,\dots,i_p=1}^{N} \xi_{i_1}^{\mu} \cdot \dots \cdot \xi_{i_j}^{\mu} \bar{\xi}_{i_{j+1}}^{\mu} \cdot \dots \cdot \bar{\xi}_{i_p}^{\mu} \sigma_{i_1} \cdot \dots \cdot \sigma_{i_p}. \end{aligned} \tag{7.51}$$

Define therefore the quantities $w_{i_1,\dots,i_p}$ by

$$w_{i_1,\dots,i_p} \equiv N^{1-p} \sum_{j=0}^{p-1} \binom{p}{j} \sum_{\mu=1}^{M(N)} \xi_{i_1}^{\mu} \cdot \dots \cdot \xi_{i_j}^{\mu} \bar{\xi}_{i_{j+1}}^{\mu} \cdot \dots \cdot \bar{\xi}_{i_p}^{\mu}, \tag{7.52}$$

which yields (7.49).

We now prove the bound (7.50). Applying Cauchy-Schwarz yields immediately ($C \leq \binom{p}{\frac{p}{2}}^2$)

$$\begin{aligned} \sum_{i_1,\dots,i_p} w_{i_1,\dots,i_p}{}^2 &= N^{2-2p} \sum_{i_1,\dots,i_p} \left(\sum_{j=0}^{p-1} \binom{p}{j} \sum_{\mu=1}^{M(N)} \xi_{i_1}^{\mu} \cdot \dots \cdot \xi_{i_j}^{\mu} \xi_{i_{j+1}}^{\mu} \cdot \dots \cdot \bar{\xi}_{i_p}^{\mu} \right)^2 \\ &\leq N^{2-2p} C \sum_{i_1,\dots,i_p} \sum_{j=0}^{p-1} \left(\sum_{\mu=1}^{M(N)} \xi_{i_1}^{\mu} \cdot \dots \cdot \xi_{i_j}^{\mu} \bar{\xi}_{i_{j+1}}^{\mu} \cdot \dots \cdot \bar{\xi}_{i_p}^{\mu} \right)^2. \end{aligned} \tag{7.53}$$

Expanding the square gives (using the definition (7.48) of A_j)

$$\begin{aligned}\sum_{i_1,\dots,i_p}&\Big(\sum_{\mu=1}^{M(N)}\xi^\mu_{i_1}\cdot\dots\cdot\xi^\mu_{i_j}\bar\xi^\mu_{i_{j+1}}\cdot\dots\cdot\bar\xi^\mu_{i_p}\Big)^2\\ &=\sum_{i_{j+1},\dots,i_p}\sum_{\mu,\nu}\bar\xi^\mu_{i_{j+1}}\cdot\dots\cdot\bar\xi^\mu_{i_p}\left(\sum_{i_1,\dots,i_j}\xi^\mu_{i_1}\cdot\dots\cdot\xi^\mu_{i_j}\xi^\nu_{i_1}\cdot\dots\cdot\xi^\nu_{i_j}\right)\bar\xi^\nu_{i_{j+1}}\cdot\dots\cdot\bar\xi^\nu_{i_p}\\ &=M\sum_{i_{j+1},\dots,i_p}\sum_{\mu,\nu}\bar\xi^\mu_{i_{j+1}}\cdot\dots\cdot\bar\xi^\mu_{i_p}A^{\mu\nu}_j[\omega]\bar\xi^\nu_{i_{j+1}}\cdot\dots\cdot\bar\xi^\nu_{i_p}.\end{aligned}\tag{7.54}$$

Consider now $\bar\xi^\mu_{i_{j+1}}\cdot\dots\cdot\bar\xi^\mu_{i_p}$ as the μ^{th} component of a vector in $\mathbb{R}^{M(N)}$. Then, by definition of the operator norm of a matrix, (7.53) and (7.54) imply

$$\sum_{i_1,\dots,i_p}{w_{i_1,\dots,i_p}}^2\le C\sum_{j=0}^{p-1}MN^{2-2p}\sum_{i_{j+1},\dots,i_p}\|A_j[\omega]\|_{op}\sum_{\mu=1}^{M(N)}\left(\bar\xi^\mu_{i_{j+1}}\cdot\dots\cdot\bar\xi^\mu_{i_p}\right)^2.\tag{7.55}$$

Finally,

$$\begin{aligned}\sum_{\mu=1}^{M(N)}\sum_{i_{j+1},\dots,i_p}\left(\bar\xi^\mu_{i_{j+1}}\cdot\dots\cdot\bar\xi^\mu_{i_p}\right)^2&=\sum_{\mu=1}^{M(N)}\sum_{i_{j+1},\dots,i_p}{\bar\xi^\mu_{i_{j+1}}}^2\cdot\dots\cdot\bar\xi^{\mu\,2}_{i_p}\\ &=\sum_{\mu=1}^{M(N)}\left(\sum_{i=1}^N\bar\xi^{\mu 2}_i\right)^{p-j}\le\left(\sum_{\mu=1}^{M(N)}\sum_{i=1}^N\bar\xi^{\mu 2}_i\right)^{p-j}\\ &=\|\bar\xi\|_2^{2(p-j)},\end{aligned}\tag{7.56}$$

since the summands $\sum_i\bar\xi^{\mu 2}_i$ are positive quantities. This proves the lemma. □

The second lemma will be used together with Lemma 7.8.

Lemma 7.9: *For all $\omega\in\Omega$, and $\{w_{i_1,\dots,i_p}\}_{i_1,\dots,i_p=1,\dots,N}$ defined as in Lemma 7.8, the following lower bound holds:*

$$\begin{aligned}\mu[\omega]\Big(\exp\big(\beta\sum_{i_1,\dots,i_p=1}^N\sigma_{i_1}\cdot\dots\cdot\sigma_{i_p}w_{i_1,\dots,i_p}\big)\Big)&\\ \ge\exp\Big(-\beta N^{\frac p4}\big(\sum_{i_1,\dots,i_p=1}^N w^2_{i_1,\dots,i_p}\big)^{\frac12}&\big(\mu[\omega]\otimes\mu[\omega](N^{\frac p2}R(\sigma,\sigma')^p)\big)^{\frac12}\Big).\end{aligned}\tag{7.57}$$

Proof: By Jensen's inequality,

$$\mu[\omega]\Big(\exp\big(\beta\sum_{i_1,\dots,i_p}\sigma_{i_1}\cdot\dots\cdot\sigma_{i_p}w_{i_1,\dots,i_p}\big)\Big)\ge\exp\Big(\beta\sum_{i_1,\dots,i_p}w_{i_1,\dots,i_p}\,\mu[\omega](\sigma_{i_1}\cdot\dots\cdot\sigma_{i_p})\Big).\tag{7.58}$$

Applying Cauchy-Schwarz then yields

$$\mu[\omega]\Big(\exp\big(\beta\sum_{i_1,\dots,i_p}\sigma_{i_1}\cdot\ldots\cdot\sigma_{i_p}w_{i_1,\dots,i_p}\big)\Big) \geq \exp\Big(-\beta\big(\sum_{i_1,\dots,i_p}w^2_{i_1,\dots,i_p}\big)^{\frac{1}{2}}\big(\sum_{i_1,\dots,i_p}\{\mu[\omega](\sigma_{i_1}\cdot\ldots\cdot\sigma_{i_p})\}^2\big)^{\frac{1}{2}}\Big). \tag{7.59}$$

We now apply what might be called the *replica trick*, namely, to write the square of a measure as the product measure over two copies of the system, that is,

$$\sum_{i_1,\dots,i_p}\{\mu[\omega](\sigma_{i_1}\cdot\ldots\cdot\sigma_{i_p})\}^2 = \sum_{i_1,\dots,i_p}\mu[\omega]\otimes\mu[\omega](\sigma_{i_1}\cdot\ldots\cdot\sigma_{i_p}\sigma'_{i_1}\cdot\ldots\cdot\sigma'_{i_p}). \tag{7.60}$$

To finish, we observe that

$$\begin{aligned}\sum_{i_1,\dots,i_p}\mu[\omega]\otimes\mu[\omega](\sigma_{i_1}\sigma'_{i_1}\cdot\ldots\cdot\sigma_{i_p}\sigma'_{i_p}) &= \mu[\omega]\otimes\mu[\omega]\Big(\big(\sum_{i=1}^{N}\sigma_i\sigma'_i\big)^p\Big)\\ &= \mu[\omega]\otimes\mu[\omega]\Big(N^pR(\sigma,\sigma')^p\Big),\end{aligned} \tag{7.61}$$

from which the statement of the lemma follows immediately. □

We are now ready to prove the "lower fluctuations" of Theorem 2.4'.

Proof of Theorem 2.4': The main idea of the proof is to find a set $B\subset\Omega$ on which F_N is near its annealed valued (and which has additional nice properties) such that it has probability greater than some constant (i.e. its probability is bounded from below uniformly in N). We then use general results due to Talagrand ([T1], Section 6) to show that[32] if $\bar{F}_N[\omega']$ is less than some small enough constant, then there exists $\omega\in B$ such that $\|\xi[\omega]-\xi[\omega']\|_2\leq(-\ln\mathbb{P}[B])^{1/2}$. This is used together with the preceding lemmas and a bound on the operator norm of the matrices A_j to extract the claimed statement.

Let

$$C(u)\equiv\{\omega\in\Omega:\bar{Z}_{N,\beta}[\omega]\leq e^{-u}\mathbb{E}\bar{Z}_{N,\beta}\} \tag{7.62}$$

be the set we are interested in. Define also the following auxiliary subsets of Ω_N (whose dependence on N is omitted).

$$\begin{aligned}B_1&\equiv\{\omega\in\Omega:\bar{Z}_{N,\beta}[\omega]\geq\frac{1}{4}\mathbb{E}\bar{Z}_{N,\beta}\}\\ B_2=B_2(K_1)&\equiv\{\omega\in\Omega:\mu[\omega]\otimes\mu[\omega](N^{p/2}R(\sigma,\sigma')^p)<K_1\}\\ B'_j=B'_j(K_2)&\equiv\begin{cases}\{\omega\in\Omega:\|A_j[\omega]\|_{op}\leq K_2N^{\frac{j-1}{2}}\}, & \text{if } j \text{ is odd}\\ \{\omega\in\Omega:\|A_j[\omega]\|_{op}\leq K_2N^{\frac{j}{2}}\}, & \text{if } j \text{ is even}\end{cases}\\ B=B(K_1,K_2)&\equiv B_1\cap B_2(K_2)\cap\left(\bigcap_{j=0}^{p-1}B'_j(K_2)\right).\end{aligned} \tag{7.63}$$

[32] This strategy was used by Talagrand to deduce the same result in the $p=2$ Hopfield model. See [T3], Section 2.

In a first step, we bound $\mathbb{P}[B]$ from below.

Proposition 7.10: *There exist constants $K_1^*, K_2^*, C^* > 0$, and $N^* \in \mathbb{N}$, such that*

$$\mathbb{P}[B(K_1, K_2)] > C^* > 0, \tag{7.64}$$

for all $K_1 > K_1^*$, $K_2 > K_2^*$ and $N > N^*$.

Proof: By the definitions of the sets,

$$\mathbb{P}[B] \geq \mathbb{P}[B_1] - \mathbb{P}[B_1 \cap B_2^c] - \sum_{j=0}^{p-1} \mathbb{P}[B_j'(K_2)] \tag{7.65}$$

The first term can be bounded by (using the definition (7.19) for $\tilde{Z}_N$, with an appropriate cut-off c),

$$\mathbb{P}[B_1] = \mathbb{P}[\bar{Z}_N[\omega] \geq \frac{1}{4}\mathbb{E}\,\bar{Z}_{N,\beta}] \geq \mathbb{P}[\tilde{Z}_{N,\beta}[\omega] \geq \frac{1}{4}\mathbb{E}\,\bar{Z}_{N,\beta}] \tag{7.66}$$

From Lemma 7.3, we know that there exists an $\bar{N} \in \mathbb{N}$, such that $\mathbb{E}\,\tilde{Z}_{N,\beta} \geq \frac{1}{2}\mathbb{E}\,\bar{Z}_{N,\beta}$ for all $N \geq \bar{N}$. This, and the result (7.37) imply that

$$\mathbb{P}[B_1] \geq \mathbb{P}[\tilde{Z}_{N,\beta}[\omega] \geq \frac{1}{4}\mathbb{E}\,\bar{Z}_{N,\beta}] \geq \mathbb{P}[\tilde{Z}_{N,\beta}[\omega] \geq \frac{1}{2}\mathbb{E}\,\tilde{Z}_{N,\beta}] \geq C^* > 0, \tag{7.67}$$

for all $N > \bar{N}$.

We now show that the second term ($\mathbb{P}[B_1 \cap B_2^c]$) is less than half of C^* if β is less than lower bound of the critical temperature (2.29).

Lemma 7.11: *Suppose $\beta < \check{\beta}_p$ (as in (2.29)). Then there exists a constant K_1^*, such that for all $K_1 > K_1^*$,*

$$\mathbb{P}[B_1 \cap B_2^c(K_1)] \leq \frac{C^*}{2} \leq \frac{1}{2}\mathbb{P}[B_1]. \tag{7.68}$$

Proof: The main idea is that we can control the Laplace transform of the quantity $NR(\sigma,\sigma')^2$ in a way similar to the proof of Lemma 4.3.

Using the definition, Chebyshev's inequality and the inequality $x^q \leq q!\,e^x$ for positive x,

$$\begin{aligned}
\mathbb{P}[B_1 \cap B_2^c] &= \mathbb{E}\,[\mathbb{1}_{B_1}\mathbb{1}_{\{\mu[\omega]\otimes\mu[\omega](N^{p/2}R(\sigma,\sigma')^p)\geq K_1\}}] \\
&\leq \mathbb{E}\,[\mathbb{1}_{B_1}K_1^{-1}\,\mu[\omega]\otimes\mu[\omega](N^{p/2}R(\sigma,\sigma')^p)] \\
&\leq K_1^{-1/2}\mathbb{E}\,[\mathbb{1}_{B_1}\,\mu[\omega]\otimes\mu[\omega]((K_1^{-1/p}NR(\sigma,\sigma')^2)^{p/2})] \\
&\leq K_1^{-1/2}\left(\frac{p}{2}\right)!\,\mathbb{E}\left[\mathbb{1}_{B_1}\mu[\omega]\otimes\mu[\omega](e^{K_1^{-1/p}NR(\sigma,\sigma')^2})\right].
\end{aligned} \tag{7.69}$$

Let $\delta \equiv K_1^{-1/p}$ for the rest of this proof. We rewrite the inner term (i.e. the integral with respect to the Gibbs measure) in the following way,

$$\begin{aligned}
\mu[\omega]\otimes\mu[\omega](e^{K_1^{-1/p}NR(\sigma,\sigma')^p}) &= \bar{Z}_{N,\beta}^{-2}\mathbb{E}_{\sigma,\sigma'}[e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+N\delta R(\sigma,\sigma')^2}] \\
&\leq \bar{Z}_{N,\beta}^{-2}\mathbb{E}_{\sigma,\sigma'}[\mathbb{1}_{\{-\beta\bar{H}(\sigma)\leq cq\}}\mathbb{1}_{\{-\beta\bar{H}(\sigma)\leq cq\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+N\delta R(\sigma,\sigma')^2}] \\
&\quad + \bar{Z}_{N,\beta}^{-2}\mathbb{E}_{\sigma,\sigma'}[(\mathbb{1}_{\{-\beta\bar{H}(\sigma)>cq\}} + \mathbb{1}_{\{-\beta\bar{H}(\sigma)>cq\}})e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+N\delta R(\sigma,\sigma')^2}] \\
&\equiv \bar{Z}_{N,\beta}^{-2}(U_1+U_2),
\end{aligned} \tag{7.70}$$

where $q = \alpha\beta^2 N$, and c is a cut-off parameter chosen as in Lemma 7.5. Thus,

$$\mathbb{E}[\mathbb{1}_{B_1}\mu[\omega]\otimes\mu[\omega](e^{\delta NR(\sigma,\sigma')^2})] \leq \mathbb{E}[\mathbb{1}_{B_1}\bar{Z}_{N,\beta}^{-2}U_1] + \mathbb{E}[\mathbb{1}_{B_1}\bar{Z}_{N,\beta}^{-2}U_2]. \tag{7.71}$$

Let us treat the second term first. Using the obvious bound $|R| \leq 1$, we get

$$U_2 \leq 2e^{\delta N}\bar{Z}_{N,\beta}\mathbb{E}_\sigma[\mathbb{1}_{\{-\beta\bar{H}(\sigma)>cq\}}e^{-\beta\bar{H}(\sigma)}]. \tag{7.72}$$

Hence, by definition of B_1 and the proof of Lemma 7.3 (that is, from the estimate on the analogue of the quantity (4.9)), there exist $C_1, C_2 > 0$ such that

$$\begin{aligned}\mathbb{E}[\mathbb{1}_{B_1}\bar{Z}_{N,\beta}^{-2}U_2] &\leq 2e^{\delta N}\mathbb{E}\mathbb{E}_\sigma[\mathbb{1}_{B_1}\mathbb{1}_{\{-\beta\bar{H}(\sigma)>cq\}}\bar{Z}_{N,\beta}^{-1}e^{-\beta\bar{H}(\sigma)}] \\ &\leq 8e^{\delta N}(\mathbb{E}\bar{Z}_{N,\beta})^{-1}\mathbb{E}\mathbb{E}_\sigma[\mathbb{1}_{\{-\beta\bar{H}(\sigma)>cq\}}e^{-\beta\bar{H}(\sigma)}] \\ &\leq C_1 e^{\delta N}(\mathbb{E}\bar{Z}_{N,\beta})^{-1}e^{-NC_2(c-1)^2}\mathbb{E}\bar{Z}_{N,\beta}.\end{aligned} \tag{7.73}$$

This implies that for all $\delta < C_2(c-1)^2$, there exists $C_3 > 0$ such that

$$\mathbb{E}[\mathbb{1}_{B_1}\bar{Z}_{N,\beta}^{-2}U_2] \leq C_1 e^{-NC_2(c-1)^2-\delta} \leq C_1 e^{-C_3 N}. \tag{7.74}$$

This means that this term will be completely irrelevant, if we choose δ sufficiently small, or equivalently, K_1 sufficiently large.

We now turn to the first term in (7.71). This is in fact by far the main part. Nevertheless, it will turn out that it is bounded by some constant. The presence of the indicator function $\mathbb{1}_{B_1}$ implies readily that

$$\begin{aligned}&\mathbb{E}[\mathbb{1}_{B_1}\bar{Z}_{N,\beta}^{-2}U_1] \\ &\quad\leq 16(\mathbb{E}\bar{Z}_{N,\beta})^{-2}\mathbb{E}\mathbb{E}_{\sigma,\sigma'}[1_{\{-\beta\bar{H}(\sigma)\leq cq\}}\mathbb{1}_{\{-\beta\bar{H}(\sigma')\leq cq\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+\delta NR(\sigma,\sigma')^2}].\end{aligned} \tag{7.75}$$

As in the proof of the lower bound on the critical β, we split the integrand into two parts, namely one where $|R|$ is small, respectively large. We write

$$\begin{aligned}&\mathbb{E}\mathbb{E}_{\sigma,\sigma'}[\mathbb{1}_{\{-\beta\bar{H}(\sigma)\leq cq\}}\mathbb{1}_{\{-\beta\bar{H}(\sigma')\leq cq\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+\delta NR(\sigma,\sigma')^2}] \\ &\quad\leq \mathbb{E}\mathbb{E}_{\sigma,\sigma'}[\mathbb{1}_{\{|R|<b\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+\delta NR(\sigma,\sigma')^2}] \\ &\qquad+\mathbb{E}\mathbb{E}_{\sigma,\sigma'}[\mathbb{1}_{\{|R|\geq b\}}\mathbb{1}_{\{-\beta\bar{H}(\sigma)\leq cq\}}\mathbb{1}_{\{-\beta\bar{H}(\sigma')\leq cq\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+\delta NR(\sigma,\sigma')^2}],\end{aligned} \tag{7.76}$$

for some b to be chosen later. Then, by Lemma 7.2, the first summand is bounded by (using also that there exists a $C_4 > 0$ such that $t^2 \leq C_4 g_p(t)$ for all $t \in [0,1]$)

$$\begin{aligned}&\mathbb{E}\mathbb{E}_{\sigma,\sigma'}[1_{\{|R|<b\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+\delta NR(\sigma,\sigma')^2}] \\ &\quad\leq \mathbb{E}_{\sigma,\sigma'}[1_{\{|R|<b\}}e^{\delta NR(\sigma,\sigma')^2+\alpha\beta^2N(1+g_p(R(\sigma,\sigma')))(1+C_3N^{-1})}] \\ &\quad\leq \mathbb{E}_{\sigma,\sigma'}[1_{\{|R|<b\}}e^{\delta NC_4 g_p(R(\sigma,\sigma'))+\alpha\beta^2N(1+g_p(R(\sigma,\sigma')))(1+C_3N^{-1})}] \\ &\quad= e^{\alpha\beta^2N+C_5}\mathbb{E}_{\sigma,\sigma'}[\mathbb{1}_{\{|R|<b\}}e^{\alpha g_p(R(\sigma,\sigma'))(\beta^2+\delta C_4/\alpha)}].\end{aligned} \tag{7.77}$$

We now want b such that it satisfies the hypothesis of Lemma 7.4 with $\beta'^2 = \beta^2 + \frac{\delta C_4}{\alpha}$, that is,

$$\beta^2 + \frac{\delta C_4}{\alpha} < \min\left(\frac{\bar{\beta}_p'^{\,2}}{4}, \frac{2s_p^2}{\alpha p^2(p-1)^2(p-3)!!}\right), \tag{7.78}$$

and

$$\gamma(b) = \alpha\left(\beta^2 + \frac{\delta C_4}{\alpha}\right)\frac{g_p(b)}{b^2} < \frac{1}{2}. \tag{7.79}$$

If $\beta < \bar{\beta}_p$, we can always find a δ^* such that the first inequality is satisfied for all $\delta < \delta^*$. Then we can choose b according to the second inequality. In a similar way to the proof of Lemma 7.4, it is then straightforward to see to that for ε small enough, there exists an $N_\varepsilon \in \mathbb{N}$ such that

$$\mathbb{E}\mathbb{E}_{\sigma,\sigma'}\left[1_{\{|R|<b\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+\delta NR(\sigma,\sigma')^2}\right] \leq e^{\alpha\beta^2 N}\frac{C_6}{\sqrt{1-2(\gamma+\varepsilon)}}, \quad \forall N > N_\varepsilon. \tag{7.80}$$

The second term in (7.76) is treated analogously to $\mathbb{E}T(c,b,b')$ in the proof of Lemma 7.5, respectively Lemma 4.4. Indeed, Fubini's theorem, the obvious inequality $|R| \leq 1$, and (7.32) yield (it is here that we actually use that we are in the low β region)

$$\begin{aligned} &\mathbb{E}\mathbb{E}_{\sigma,\sigma'}[\mathbb{1}_{\{|R|\in[b,1]\}}\mathbb{1}_{\{-\beta\bar{H}(\sigma)\leq cq\}}1_{\{-\beta\bar{H}(\sigma')\leq cq\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')+\delta NR(\sigma,\sigma')^2}] \\ &\qquad \leq \mathbb{E}_{\sigma,\sigma'}e^{\delta N}\mathbb{E}[\mathbb{1}_{\{|R|\in[b,1]\}}\mathbb{1}_{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')\leq 2cq\}}e^{-\beta\bar{H}(\sigma)-\beta\bar{H}(\sigma')}] \\ &\qquad \leq e^{N(\alpha\beta^2(1+C_5N^{-1})+\delta-d)} \leq e^{\alpha\beta^2N-dN/2}, \end{aligned} \tag{7.81}$$

whenever $\delta \leq \frac{d}{2}$ and N larger than some N_2.

Relations (7.80) and (7.81) then bound the right-hand side of (7.75) by (choose ε small enough)

$$\begin{aligned} \mathbb{E}[\mathbb{1}_{B_1}\bar{Z}_{N,\beta}^{-2}U_1] &\leq 16(\mathbb{E}\bar{Z}_{N,\beta})^{-2}(\frac{C_4}{\sqrt{1-2(\gamma+\varepsilon)}} + e^{-dN/2})e^{\alpha\beta^2 N} \\ &\leq 16e^{2C_6}(\frac{C_4}{\sqrt{1-2(\gamma+\varepsilon)}} + e^{-dN/2}) \\ &\leq C_7, \quad \forall N > N_\varepsilon. \end{aligned} \tag{7.82}$$

To finish the proof of Lemma 7.11, insert (7.82) and (7.74) in (7.71), and use the latter in (7.69). Thus,

$$\begin{aligned} \mathbb{P}[B_1 \cap B_2^c] &\leq K_1^{-1/2}(\frac{p}{2})!\,\mathbb{E}[\mathbb{1}_{B_1}\mu[\omega]\otimes\mu[\omega](e^{K_1^{-1/p}NR(\sigma,\sigma')^2})] \\ &\leq K_1^{-1/2}(\frac{p}{2})!\,(C_1e^{-C_3N} + C_7), \end{aligned} \tag{7.83}$$

whenever $K_1^{-1/p} < \min\{(c-1)^2\gamma, \frac{d}{2}, 2\alpha(\frac{1}{16}-\beta^2)\}$. Define therefore

$$K_1^* \equiv \left(\max\left\{(c-1)^{-2}\gamma^{-1}, \frac{2}{d}, \frac{16}{2\alpha(16\beta^2-1)}, \frac{9}{(C^*)^2}\left(\left(\frac{p}{2}\right)!\,(4e^{-C_2N}+C_7)\right)^2\right\}\right)^p. \tag{7.84}$$

Then for all $K_1 > K_1^*$ and all N large enough,

$$\mathbb{P}[B_1 \cap B_2^c] \leq \frac{C^*}{3}. \tag{7.85}$$

This proves Lemma 7.11. □

The terms for $j \geq 2$ in the remaining sum on the right-hand side of inequality (7.65) are bounded by the following result which is proved in Part II as Theorem 8.1.

Theorem 8.1: *Let $M = M(N) = \alpha N^{q'}$ with $\alpha > 0$ and $q' \geq q$. Then, for each $q \geq 2$, there exist constants $C_q, K > 0$, $l \in (0, \frac{1}{2})$, and $\tilde{N} \in \mathbb{N}$, and subsets $\Omega_N \subset \Omega$, such that for all $N \geq \tilde{N}$, the measure of Ω_N is at least $1 - e^{-KN^l}$, and*

$$\|A_q\|_q \leq \begin{cases} C_q N^{\frac{q-1}{2}}, & \text{if } q \text{ is odd;} \\ C_q N^{\frac{q}{2}}, & \text{if } q \text{ is even.} \end{cases} \tag{7.86}$$

for all $\omega \in \Omega_N$.

The case $j = 1$ corresponds to the matrix appearing in the $p = 2$ Hopfield model (see e.g. [BG1], Section 4). Although the size of the matrix is much larger (αN^{p-1} instead of αN), this is compensated by the factor M in the definition. The matrix can thus be bounded by the same methods as in [BG1]. This gives also an exponential bound on the probability of the set $B_1(K_2)$ (for K_2 large enough). The case $j = 0$ is trivial since this is just the identity matrix, whose norm is always equal to one.

Finally, the estimates (7.67), and (7.86) from the previous theorem, as well as the preceding lemma imply that

$$\mathbb{P}[B(K_1, K_2)] \geq C^* - \frac{C^*}{3} - (p-1)e^{-CN^l} \geq \frac{C^*}{2}, \tag{7.87}$$

if $K_1 > K_1^*$, $K_2 > K_2^*$ and N is large enough. This finishes the proof of Proposition 7.10. □

It follows from general results in [T1] that there exists $\omega \in B$ and $\omega' \in C(u)$ such that

$$\|\bar{\xi}\|_2 = \|\xi[\omega] - \|\xi[\omega']\|_2 \leq C_1(-\ln(\mathbb{P}[C(u)]))^{1/2}, \tag{7.88}$$

where C_1 is independent of N. Indeed, we know that $\bar{F}_N$ is a convex function (Section 7.2, (7.44)–(7.46)), which implies that the level set $C(u)$ is convex. Hence, for any $\omega \in B$,

$$d(C(u), \omega) \leq 2f(C(u), \omega), \tag{7.89}$$

where d denotes the Euclidean distance and f is the function defined in Chapter 6 of [T1]. Suppose that for all $\omega \in B$,

$$2t \leq d(C(u), \omega) \leq 2f(C(u), \omega). \tag{7.90}$$

Then $B \subset \{\omega' \in \Omega : f(C(u), \omega') > t\}$, and by Theorem 6.1 in [T1] and Proposition 7.10,

$$C_2 \leq \mathbb{P}[B] \leq \mathbb{P}[f(C(u), \cdot) > t] \leq \mathbb{P}[C(u)]^{-1} e^{-\frac{t^2}{4}}, \tag{7.91}$$

which yields immediately,

$$t \leq 2(-\ln \mathbb{P}[C(u)])^{\frac{1}{2}} + 2(-\ln C_2)^{\frac{1}{2}}. \tag{7.92}$$

However, $C(u) \subset B^c$ and hence $\mathbb{P}[C(u)] \leq 1 - C_2$. This implies

$$t \leq C_1(-\ln \mathbb{P}[C(u)])^{\frac{1}{2}}. \tag{7.93}$$

This shows the claim (7.88). Now

$$\begin{aligned} \bar{Z}_{N,\beta}[\omega'] &= \mathbb{E}_\sigma \left[e^{-\beta \bar{H}_N[\omega]} e^{-\beta \bar{H}_N[\omega'] + \beta \bar{H}_N[\omega]} \right] \\ &= \bar{Z}_{N,\beta}[\omega] \mu_{N,\beta}[\omega] \left[e^{-\beta \bar{H}_N[\omega'] + \beta \bar{H}_N[\omega]} \right]. \end{aligned} \tag{7.94}$$

Since $\omega \in B$,

$$\bar{Z}_{N,\beta}[\omega'] \geq \frac{\mathbb{E}\bar{Z}_{N,\beta}}{4} \exp\left(-\beta N^{\frac{p}{4}} \Big(\sum_{i_1,\dots,i_p} w^2_{i_1,\dots,i_p} \Big)^{\frac{1}{2}} \Big(\mu[\omega] \otimes \mu[\omega] (N^{p/2} R(\sigma,\sigma')^p) \Big)^{\frac{1}{2}} \right). \tag{7.95}$$

Apply now Lemma 7.8 and Lemma 7.9 to the integrand. This yields

$$\begin{aligned} \bar{Z}_{N,\beta}[\omega'] &\geq \frac{\mathbb{E}\bar{Z}_{N,\beta}}{4} \exp\left(-\beta C_2 N^{\frac{p}{4}} \Big(\sum_{j=1}^{p-1} N^{1-p} \|A_j\|_{op} \|\xi\|_2^{2(p-j)} \Big)^{\frac{1}{2}} \right) \\ &\geq \frac{\mathbb{E}\bar{Z}_{N,\beta}}{4} \exp\left(-\beta C_3 (\sum_{j \text{ odd}} N^{\frac{1+j-p}{2}} \|\bar{\xi}\|_2^{2(p-j)} + \sum_{j \text{ even}} N^{\frac{2+j-p}{2}} \|\bar{\xi}\|_2^{2(p-j)})^{\frac{1}{2}} \right). \end{aligned} \tag{7.96}$$

On the other hand, $\omega' \in C(u)$, and thus by (7.88)

$$\begin{aligned} \left(\frac{u - 2\ln 2}{\beta C_3} \right)^2 &\leq \sum_{j \text{ odd}} N^{\frac{j+1-p}{2}} \|\bar{\xi}\|_2^{2(p-j)} + \sum_{j \text{ even}} N^{\frac{j+2-p}{2}} \|\bar{\xi}\|_2^{2(p-j)} \\ &\leq \sum_{j \text{ odd}} N^{\frac{j+1-p}{2}} \Big(-\ln(\mathbb{P}[C(u)]) \Big)^{p-j} + \sum_{j \text{ even}} N^{\frac{j+2-p}{2}} \Big(-\ln(\mathbb{P}[C(u)]) \Big)^{p-j}. \end{aligned} \tag{7.97}$$

From this one concludes that

$$\mathbb{P}[C(u)] \leq \exp\left(-\max\left\{ \max_{j \text{ odd}} \left\{ \Big(\frac{u-2\ln 2}{\beta C_4} \Big)^{\frac{2}{p-j}} N^{\frac{p-1-j}{2}} \right\}, \max_{j \text{ even}} \left\{ \Big(\frac{u-2\ln 2}{\beta C_3} \Big)^{\frac{2}{p-j}} N^{\frac{p-2-j}{2}} \right\} \right\} \right). \tag{7.98}$$

The only terms that can achieve the maximum (for large N) are those whose exponents of N are equal to zero. Thus

$$\mathbb{P}[C(u)] \leq C_5 \exp\left(-\max\left\{ (\frac{u-2\ln 2}{\beta C_3})^2, \frac{u-2\ln 2}{\beta C_3} \right\} \right). \tag{7.99}$$

This shows (2.33).

The bound on the fluctuations above the annealed free energy follow immediately from Chebyshev's inequality with first mean. Indeed

$$\begin{aligned}\mathbb{P}\left[\bar{F}_N \geq \frac{1}{N}\ln \mathbb{E}\,\bar{Z}_N - \frac{u}{N}\right] &= \mathbb{P}\left[\bar{Z}_N \geq e^u \mathbb{E}\,\bar{Z}_N\right] \\ &\leq e^{-u}.\end{aligned} \tag{7.100}$$

This concludes the proof of Theorem 2.4'. □

7.4 Replica Overlap: Proof of Theorem 2.6'

Apply Jensen's inequality, and split the integrand in the same way as in the proof of Lemma 7.11 (inequality (7.70)). We obtain (let $\delta = 2\gamma$)

$$\begin{aligned}\mu\otimes\mu[e^{\delta N R(\sigma,\sigma')^2}] &\leq (\mu\otimes\mu[e^{\gamma N R(\sigma,\sigma')^2}])^{1/2} \\ &\leq (\bar{Z}_{N,\beta}^{-2} U_1 + \bar{Z}_{N,\beta}^{-2} U_2)^{1/2} \\ &\leq \bar{Z}_{N,\beta}^{-1} U_1^{1/2} + \bar{Z}_{N,\beta}^{-1} U_2^{1/2},\end{aligned} \tag{7.101}$$

where U_1, U_2 are as in (7.70). Hence, using Cauchy-Schwarz,

$$\begin{aligned}\mathbb{E}\,[\mu\otimes\mu[e^{\delta N R(\sigma,\sigma')^2}]] &\leq \mathbb{E}\,[\bar{Z}_{N,\beta}^{-1} U_1^{1/2}] + \mathbb{E}\,[\bar{Z}_{N,\beta}^{-1} U_2^{1/2}] \\ &\leq (\mathbb{E}\,[\bar{Z}_{N,\beta}^{-2}])^{1/2} (\mathbb{E}\,U_1)^{1/2} + (\mathbb{E}\,[\bar{Z}_{N,\beta}^{-1}])^{1/2} (\mathbb{E}\,[\bar{Z}_{N,\beta}^{-1} U_2])^{1/2}.\end{aligned} \tag{7.102}$$

The first summand is bounded by splitting U_1 into the two terms $S(b)$ and $T(c,b,1)$ from Section 7.2 and treating them exactly as in (7.75) and following, implying the bounds (7.80), resp. (7.81). For the second term, we use (7.72) and Lemma 7.3. This yields (for γ, resp. δ small enough)

$$\mathbb{E}\,[\mu\otimes\mu[e^{\delta N R(\sigma,\sigma')^2}]] \leq (\mathbb{E}\,[\bar{Z}_{N,\beta}^{-2}])^{1/2} C_1\,\mathbb{E}\,\bar{Z}_{N,\beta} + C_2 (\mathbb{E}\,[\bar{Z}_{N,\beta}^{-1}])^{1/2} e^{-C_3 N}\,\mathbb{E}\,\bar{Z}_{N,\beta}\,. \tag{7.103}$$

Theorem 2.4' implies that $\mathbb{E}\,[\bar{Z}_{N,\beta}^{-2}] \leq C_4 (\mathbb{E}\,\bar{Z}_{N,\beta})^{-2}$. Indeed,

$$\begin{aligned}\mathbb{E}\,[\bar{Z}_{N,\beta}^{-2}] &= \int_0^\infty \mathbb{P}[\bar{Z}_{N,\beta}^{-2} > x]\,dx \\ &\leq \int_0^{(\mathbb{E}\,\bar{Z}_{N,\beta})^{-2}} 1\,dx + \int_{(\mathbb{E}\,\bar{Z}_{N,\beta})^{-2}}^\infty \mathbb{P}[\bar{Z}_{N,\beta}^{-2} > x]\,dx \\ &\leq (\mathbb{E}\,\bar{Z}_{N,\beta})^{-2} + \int_{(\mathbb{E}\,\bar{Z}_{N,\beta})^{-2}}^\infty \mathbb{P}[F_N < -\frac{1}{2}\ln x]\,dx.\end{aligned} \tag{7.104}$$

Substituting $y = \ln \mathbb{E}\,\bar{Z}_{N,\beta} + \frac{1}{2}\ln x$ yields (with f from Theorem 2.4')

$$\begin{aligned} \mathbb{E}\,[\bar{Z}_{N,\beta}^{-2}] &\leq (\mathbb{E}\,\bar{Z}_{N,\beta})^{-2}\Big(1 + 2\int_0^\infty \mathbb{P}[F_N < \ln \mathbb{E}\,\bar{Z}_{N,\beta} - y]e^{2y}\,dy\Big) \\ &\leq (\mathbb{E}\,\bar{Z}_{N,\beta})^{-2}\Big(1 + 2\int_0^\infty e^{-f(y)}e^{2y}\,dy\Big). \end{aligned} \tag{7.105}$$

Since $f(y)$ grows as y^2 for large y, this shows that

$$\mathbb{E}\,[\bar{Z}_{N,\beta}^{-2}] \leq (\mathbb{E}\,\bar{Z}_{N,\beta})^{-2}C_4. \tag{7.106}$$

Also, by Jensen's inequality

$$\mathbb{E}\,[\bar{Z}_{N,\beta}^{-1}] \leq (\mathbb{E}\,[\bar{Z}_{N,\beta}^{-2}])^{1/2}. \tag{7.107}$$

Using (7.106) and (7.107) in (7.103) gives

$$\mathbb{E}[\mu \otimes \mu[e^{\delta N R(\sigma,\sigma')^2}]] \leq C_8 + C_9 e^{-C_3 N} \leq K. \tag{7.108}$$

This proves Theorem 2.6'. □

Part II

Random Matrices

8 Bounds on the Norm

In this second part we prove estimates on the norms of certain random matrices. These results were used in previous proofs on the fluctuations of the free energies in Chapter 5 and Section 7.3..

We define the sets of random matrices we will consider. Let $\{\xi_i^\mu\}_{i,\mu\in\mathbb{N}}$ be a family of i.i.d. Bernoulli random variables on some probability space $(\Omega,\mathcal{F},\mathbb{P})$, taking values $+1$ and -1 with equal probability. Construct the $M\times M$ random matrix $A_{N,q}=(A_{N,q}^{\mu\nu})_{\mu,\nu=1,\dots,M}$ according to (we omit the explicit reference to N and M for clarity of presentation)

$$A_q^{\mu\nu}\equiv M^{-1}\left(\sum_{i=1}^N \xi_i^\mu\xi_i^\nu\right)^q = M^{-1}\sum_{(i_l)_{l=1,\dots,q}}\prod_{l=1}^q \xi_{i_l}^\mu\xi_{i_l}^\nu. \tag{8.1}$$

The matrix A_q is symmetric by definition, and its diagonal elements are constant and equal to one. However, the off-diagonal elements are not independent. Let $\|\cdot\|$ denote the operator norm for linear maps from an M dimensional vector space into itself, that is, the maximum of the absolute values of its eigenvalues. Then we have the following estimate on $\|A_q\|$.

Theorem 8.1: *Let $M=M(N)=\alpha N^{q'}$ with $\alpha>0$ and $q'\geq q$. Then, for each $q\geq 2$, there exist constants $C_q,K>0$, $l\in(0,\frac{1}{2})$, and $\tilde{N}\in\mathbb{N}$, and subsets $\Omega_N\subset\Omega$, such that for all $N\geq\tilde{N}$, the measure of Ω_N is at least $1-e^{-KN^l}$, and*

$$\|A_q\|\leq C_q\begin{cases} N^{\frac{q-1}{2}}, & \text{if } q \text{ is odd;}\\ N^{\frac{q}{2}}, & \text{if } q \text{ is even.}\end{cases} \tag{8.2}$$

for all $\omega\in\Omega_N$.

Let us consider two variants of the above matrix. As before, M is the size of the matrix. Define $B_{N,q}=(B_{N,q}^{\mu\nu})_{\mu,\nu=1,\dots,M}$ by (we omit again the indication of N)

$$\begin{aligned} B_q^{\mu\nu} &\equiv M^{-1}\sum_{1\leq i_1<\ldots<i_q\leq N}\xi_{i_1}^\mu\cdot\ldots\cdot\xi_{i_q}^\mu\xi_{i_1}^\nu\cdot\ldots\cdot\xi_{i_q}^\nu \\ &= M^{-1}(q!)^{-1}\sum_{\substack{\{i_l\}_{l=1,\dots,q}\\ \text{all different}}}\prod_{l=1}^q\xi_{i_l}^\mu\xi_{i_l}^\nu, \end{aligned} \tag{8.3}$$

representing a linear map from an M dimensional vector space into itself.

Finally, we consider two closely related matrices, whose definitions require some more detail. Let V be a $\binom{N}{q}$ dimensional space with an orthonormal basis $\{\psi_\mathcal{I}\}$, indexed by the sets $\mathcal{I}\subset\{1,\dots,N\}$, $|\mathcal{I}|=q$. Define the map $B:V\to V$ by its matrix representation in the basis $\{\psi_\mathcal{I}\}$. That is, by the $\binom{N}{q}\times\binom{N}{q}$ matrix with elements

$$B'^{\mathcal{I},\mathcal{J}}={B'_{N,q}}^{\mathcal{I},\mathcal{J}}\equiv M^{-1}\sum_{\mu=1}^M\prod_{\substack{l\in\mathcal{I}\\ l'\in\mathcal{J}}}\xi_l^\mu\xi_{l'}^\mu. \tag{8.4}$$

Then we have the following estimates.

Theorem 8.2: *Suppose that $M(N) = \alpha N^{q'}$, and $q' > q \geq 2$. Then, for all $\varepsilon > 0$, and $C_1 > 1$, there exist $C_2 > 0$ and $\bar{N} \in \mathbb{N}$, such that for all $N \geq \bar{N}$,*

$$\mathbb{P}\left[\|B_q\| \geq C_1\right] \leq e^{-C_2 N^{\frac{1}{2}-\varepsilon}}, \tag{8.5}$$

and

$$\mathbb{P}\left[\|B'_q\| \geq C_1\right] \leq e^{-C_2 N^{\frac{1}{2}-\varepsilon}}, \tag{8.6}$$

where $\|\cdot\|$ is the operator norm.

Motivated by the application we have in mind (see Chapter 5), we also look at the following restriction of B'. For all $i \in \{1, \ldots, N\}$, let V_i be the linear space spanned by the basis elements $\{\psi_{\mathcal{I}}\}_{\mathcal{I} \ni i}$. Consider the map $T_i : V_i \to V_i$ given by the restriction of B' to V_i, that is

$$(T_i)_{\mathcal{I},\mathcal{J}} = \begin{cases} (B')_{\mathcal{I},\mathcal{J}}, & \text{if } \mathcal{I}, \mathcal{J} \ni i; \\ 0, & \text{otherwise.} \end{cases} \tag{8.7}$$

Then the following bound holds for $\|T_i\|$.

Theorem 8.3: *Suppose that $M(N) = \alpha N^{q'}$, and $q' \geq q \geq 2$. Then, for all $\varepsilon > 0$, and $C_1 > 1$, there exist $C_2 > 0$ and $\bar{N} \in \mathbb{N}$, such that for all $N \geq \bar{N}$,*

$$\mathbb{P}\left[\|T_i\| > C_1\right] \leq e^{-C_2 N^{\frac{1}{2}-\varepsilon}}. \tag{8.8}$$

The proofs of our results are based as usual on bounds on the expectation of traces of high powers of our matrices [FK] from which the estimates follow by Chebyshev's inequality. The hard work consists in the solution of the ensuing combinatorial problems which are rather different from those encountered in the classical cases.

The remainder of this part is organized as follows. In Section 9.1, we present a graphical representation of the expectation of powers of our matrices. In Section 9.2, we prove Theorem 8.1, while in Section 9.3 Theorem 8.2 respectively 8.3 are shown.

9 Proof of the Estimates

9.1 Graph Representation of the Trace

We start with a representation of the trace. As pointed out above, the main step of the proof is to bound the expectation of a high even power of A (q being a fixed parameter, we will in the sequel omit it where this is possible and no confusion arises). Let thus k be an even integer. Then, using the definition and elementary algebra to rearrange the sums, respectively the products, the trace of A^k can be written as (addition in the indices is understood to be modulo the largest allowed value)

$$\begin{aligned} \mathrm{tr} A^k = \sum_{\mu_1,\dots,\mu_k} \prod_{t=1}^{k} A^{\mu_t \mu_{t+1}} &= M^{-k} \sum_{\mu_1,\dots,\mu_k} \prod_{t=1}^{k} \Big(\sum_{(i_l)_{l=1,\dots,q}} \prod_{l=1}^{q} \xi_{i_l}^{\mu_t} \xi_{i_l}^{\mu_{t+1}} \Big) \\ &= M^{-k} \sum_{\mu_1,\dots,\mu_k} \sum_{(i_l^t)_{l=1,\dots,q}^{t=1,\dots k}} \prod_{l=1}^{q} \prod_{t=1}^{k} \xi_{i_l^t}^{\mu_t} \xi_{i_l^t}^{\mu_{t+1}} \end{aligned} \tag{9.1}$$

For the matrix B_q, the corresponding expression is

$$\mathrm{tr} B^k = M^{-k} (q!)^{-k} \sum_{\mu_1,\dots,\mu_k} \sum_{\substack{(i_l^t)_{l=1,\dots,q}^{t=1,\dots k} \\ \text{different}}} \prod_{l=1}^{q} \prod_{t=1}^{k} \xi_{i_l^t}^{\mu_t} \xi_{i_l^t}^{\mu_{t+1}}, \tag{9.2}$$

where *different* indicates that the sum runs only over those sets $\{i_l^t\}_{l=1,\dots,q}^{t=1,\dots k}$ such that for all t, the set $\{i_l^t\}_{l=1,\dots,q}$ has size q (that is, or each t, and all $l \neq l'$, $i_l^t \neq i_{l'}^t$).

We want to rearrange the sums on the right-hand side of (9.1), resp. (9.2) in a more transparent form. This is done in the following way. Sum first over all possible sequences $\mu_1, \dots, \mu_k$ which have a given range $\mathcal{R}$. Then sum over all $\mathcal{R} \subset \{1, \dots, M\}$ such that $|\mathcal{R}| = r$, and finally sum over all possible cardinalities for sets $\mathcal{R}$, i.e. $r = 1, \dots, k$. Split then the sum over all i_l^t in the corresponding way.

In the sequel, denote by $(\mu_1, \dots, \mu_k)$ the sequence of indices, and by $\{\mu_1, \dots, \mu_k\}$ the corresponding range. Similarly, (i_l^t) denotes the sequence of elements in $\mathcal{N}$, whereas $\{i_l^t\}$ is its range. Then the above splitting of the sums reads

$$\mathrm{tr} A^k = M^{-k} \sum_{r=1}^{k} \sum_{\substack{\mathcal{R} \subset \mathcal{M} \\ |\mathcal{R}|=r}} \sum_{\substack{(\mu_1,\dots,\mu_k): \\ \{\mu_1,\dots,\mu_k\}=\mathcal{R}}} \sum_{s=1}^{kq} \sum_{\substack{\mathcal{S} \subset \mathcal{N} \\ |\mathcal{S}|=s}} \sum_{(i_l^t):\{i_l^t\}=\mathcal{S}} \prod_{l=1}^{q} \prod_{t=1}^{k} \xi_{i_l^t}^{\mu_t} \xi_{i_l^t}^{\mu_{t+1}}. \tag{9.3}$$

The analogue decomposition for the matrix B is

$$\mathrm{tr} B^k = M^{-k} (q!)^{-k} \sum_{r=1}^{k} \sum_{\substack{\mathcal{R} \subset \mathcal{M} \\ |\mathcal{R}|=r}} \sum_{\substack{(\mu_1,\dots,\mu_k): \\ \{\mu_1,\dots,\mu_k\}=\mathcal{R}}} \sum_{s=1}^{kq} \sum_{\substack{\mathcal{S} \subset \mathcal{N} \\ |\mathcal{S}|=s}} \sum_{\substack{(i_l^t):\{i_l^t\}=\mathcal{S} \\ \text{different}}} \prod_{l=1}^{q} \prod_{t=1}^{k} \xi_{i_l^t}^{\mu_t} \xi_{i_l^t}^{\mu_{t+1}}, \tag{9.4}$$

different having the same meaning as before.

In the spirit of [FK], we construct a graph representation of the product $\prod_{l=1}^{q}\prod_{t=1}^{k}\xi_{i_l^t}^{\mu_t}\xi_{i_l^t}^{\mu_{t+1}}$ in (9.3), respectively (9.4). Let $\mathcal{K}$ be the complete bipartite graph with vertex sets $\mathcal{R}$ and $\mathcal{S}$. Then each ξ_i^μ corresponds to the edge (μ, i) of $\mathcal{K}$. Each product in (9.3), resp. (9.4) corresponds to a walk on $\mathcal{K}$ in the obvious way. Conversely, given $\mathcal{K}$ and the sequence $(\mu_1, \dots, \mu_k)$, any walk on $\mathcal{K}$ which visits the μ's in the correct order corresponds to possible product. It will from now on be implicitly understood that a *walk* on $\mathcal{K}$ visits the μ's in the given order.

However, we are interested in the expectation of the product. By the i.i.d. Bernoulli nature of the random variables, it is clear that the only way to get a non-zero contribution is for each variable ξ_i^μ to appear an even number of times, or, in the graph picture, the associated walk has to use every edge of $\mathcal{K}$ an even number of times (including zero). Given $\mathcal{K}$, we define therefore a walk w on it *admissible for* A, if it uses each edge of $\mathcal{K}$ an even number of times. Our problem of calculating the expectation of the trace is therefore transformed into the task of counting the admissible walks for given k and q. That is, taking expectation of (9.3),

$$\mathbb{E}\operatorname{tr} A^k = M^{-k}\sum_{r=1}^{k}\sum_{\substack{\mathcal{R}\subset\mathcal{M}\\|\mathcal{R}|=r}}\sum_{\substack{(\mu_1,\dots,\mu_k):\\\{\mu_1,\dots,\mu_k\}=\mathcal{R}}}\sum_{s=1}^{kq}\sum_{\substack{\mathcal{S}\subset\mathcal{N}\\|\mathcal{S}|=s}}\sum_{\substack{(i_l^s)\\\{i_l^s\}=\mathcal{S}}}\delta_{\mu_1,\dots,\mu_k,i_1^1,\dots,i_q^k}, \tag{9.5}$$

where

$$\delta_{\mu_1,\dots,\mu_k,i_1^1,\dots,i_q^k} = \begin{cases} +1, & \text{if there is an admissible walk;}\\ -1, & \text{if there is none.}\end{cases}$$

It will turn out that for a given sequence $(\mu_1, \dots, \mu_k)$, this quantity depends only on $\mathcal{S}$, and not on the exact order of its elements. We will therefore denote by

$$W_{\mu_1,\dots,\mu_k,i_1^1,\mathcal{S}} = \sum_{\substack{(i_l^s)\\\{i_l^s\}=\mathcal{S}}}\delta_{\mu_1,\dots,\mu_k,i_1^1,\dots,i_q^k}. \tag{9.6}$$

The main task will be to obtain a good bound on $W_{\mu_1,\dots,\mu_k,\mathcal{S}}$ and to show that the sum over s does in general not extend up to kq, but merely to some $s' = s'(\mu_1, \dots, \mu_k)$.

In the case of the matrix B, there is an additional restriction, namely the constraint that $i_l^t \neq i_{l'}^t$, for all t and all $l \neq l'$. For a given graph $\mathcal{K}$, we call a walk w *admissible for* B if it uses each edge of $\mathcal{K}$ an even number of times, and for all t and l, l', the points i_l^t and $i_{l'}^t$ are different.

We now define two additional graphs. Define first the following (undirected) graph $\mathcal{G}$. The vertex set $\mathcal{V}(\mathcal{G})$ is $\mathcal{R}$, and for each t, put an edge between μ_t and μ_{t+1} (μ_{k+1} being identified with μ_1) and label it by e_t. From this (multi)graph, we can reconstruct the sequence $(\mu_1, \dots, \mu_k)$ through the labels t. Construct a second graph $\mathcal{G}^q$ by replacing each edge e_t of $\mathcal{G}$ by q edges $\{e_{l,t}\}_{l=1,\dots,q}$. Now, the projection onto the graph $\mathcal{G}^q$ of any walk corresponding to a product obeys the lexicographic order of the labels of the edges.

We would like to have a characterization of admissible walks in terms of their projections onto $\mathcal{G}$, resp. $\mathcal{G}^q$. This is provided by the following results. We first define some useful notions. Define π_A to be the function which maps any walk (respecting the order in which the set $\mathcal{R}$ is visited) $w = (\mu_1, \dots, \mu_k, i_1^1, \dots, i_q^k)$ on $\mathcal{K}$ onto the set of partitions of $\mathcal{E}(\mathcal{G}^q)$ via

$$\pi_A : w \mapsto \mathcal{C}(w) = \{C_i(w)\}_{i\in\mathcal{S}}, \tag{9.7}$$

where $C_i = \{e_{l,t} \in \mathcal{E}(\mathcal{G}^q) : i_l^t = i\}$. Similarly, define π_B to be the function which maps any walk (respecting the order in which the set $\mathcal{R}$ is visited) w on $\mathcal{K}$ onto the family of subsets of $\mathcal{E}(\mathcal{G})$ via

$$\pi_B : w \mapsto \mathcal{C}(w) = \{C_i\}_{i\in\mathcal{S}}, \tag{9.8}$$

where $C_i = \{e_t \in \mathcal{E}(\mathcal{G}) : \exists l : i_l^t = i\}$.

We denote by ρ the mapping from the partitions of $\mathcal{E}(\mathcal{G}^q)$ onto the set $\mathcal{P}(\mathcal{G})$ of subgraphs of $\mathcal{G}$, induced by the projection $e_{l,t} \mapsto e_t$.

We also define the following concepts. A *circuit* is a graph in which every vertex has an even number of incident edges (observe that contrary to custom, we do not suppose that the graph is connected). A *cycle* is a connected graph whose vertices have all exactly two incident edges. A *circuit cover of a graph* $\mathcal{G}$ is a partition of its edge set into disjoint circuits. A *q-circuit cover of a graph* $\mathcal{G}$ is a collection of subgraphs of $\mathcal{G}$ such that each of these subgraphs is a circuit, and every edge of $\mathcal{G}$ is in exactly q of these subgraphs. *Cycle covers* and *q-cycle covers* are defined analogously. A *loop* is an edge whose endpoints coincide. The *size* of a circuit, resp. cycle C is the number of edges it contains. A cycle of size n is called an *n-cycle*.

We then have the following characterization of admissible walks.

Lemma 9.1: *A walk w on K is admissible for A if and only if $\pi_A(w)$ is a circuit cover of $\mathcal{G}^q$.*

Proof: It is clear that $\pi_A(w)$ is a disjoint cover of $\mathcal{G}^q$ by construction. Suppose now that for some walk, there exists an element C_i of the partition which is not a circuit. Then C_i contains at least one vertex v that has an odd number of incident edges belonging to C_i. But this means that the edge $(i,s) \in \mathcal{E}(\mathcal{K})$ is used an odd number of times, and thus the walk w is not admissible for A.

Conversely, we have to show that for any circuit decomposition $\mathcal{C}$ of $\mathcal{G}^q$, the walk $w(\mathcal{C}) = \pi_A^{-1}(\mathcal{C})$ is admissible for A. Consider an arbitrary $i \in \mathcal{S}$, its associated circuit $C_i \in \mathcal{C}$, and any vertex $v \in \mathcal{V}(C_i)$. The incident edges of C_i at v have other endpoints $\{v_1, \dots, v_h\}$. Since C_i is a circuit, h is even, regardless of whether C_i contains a loop based at v or not. This implies that the edge $(v,i) \in \mathcal{K}$ is indeed used an even number of times by the walk $w(\mathcal{C})$. Since i and v are arbitraty, this proves the lemma. □

Lemma 9.2: *A walk w on $\mathcal{K}$ is admissible for B, if and only if $\pi_B(w)$ is a q-circuit cover of $\mathcal{G}$.*

Proof: If w is admissible for B, it is also admissible for A, and thus $\pi_A(w)$ is a circuit cover of $\mathcal{G}^q$ by the same argument as in the above proof. Moreover, $\pi_B(w) = \rho \circ \pi_A(w)$, and since by construction, for all $i \in \mathcal{S}$, all edges $e_{l,t} \in C_i$ have different indices t, this implies that C_i is also a circuit of $\mathcal{G}$. Also, every edge has to be covered by exactly q circuits. This proves the *only if* part.

Conversely, every q-circuit cover of $\mathcal{G}$ has at least one preimage under ρ which is a circuit cover of $\mathcal{G}^q$ (in fact, there are at most $(q!)^k$). Their preimages under π_A are admissible walks for A by the second part of the proof of Lemma 9.1, and by construction, they also satisfy the condition *different*. Thus, they are admissible for B. This proves the lemma. □

9.2 Proof of Theorem 8.1

In this section, we calculate the number of admissible walks for A. Since the mapping π_A is bijective, we obtain readily the following corollary to Lemma 9.1.

Corollary 9.3: *The number of admissible walks for A, given sequence a $(\mu_1, \dots, \mu_k)$ and a set $\mathcal{S}$, is bounded by the sum over all circuit covers of $\mathcal{G}^q$ of the number of surjective maps from this circuit cover to $\mathcal{S}$.*

The key ingredient in the proofs is an optimal bound on the size of $\mathcal{S}$ for a given sequence $(\mu_1, \dots, \mu_k)$. This bound is expressed in terms of the following quantities. For any $\nu \in \mathcal{R}$, and any sequence $(\mu_1, \dots, \mu_k)$, let $n_\nu = \#\{t : \mu_t = \nu\}$ be the number of appearances of ν in the sequence. Similarly, $n'_\nu = \#\{t : \mu_t = \mu_{t+1} = \nu\}$.

Lemma 9.4: *If q is odd, then for any given sequence $(\mu_1, \dots, \mu_k)$, a necessary condition on $\mathcal{S}$ for the existence of an admissible walk is that*

$$|\mathcal{S}| \le s'(\mu_1, \dots, \mu_k) \equiv \sum_{\nu \in \mathcal{R}} (n_\nu - n'_\nu) \frac{q-1}{2} + \sum_{\nu \in \mathcal{R}} (n_\nu - n'_\nu - 1) + \sum_{\nu \in \mathcal{R}} q n'_\nu + 1. \qquad (9.9)$$

Remark: If $|\mathcal{R}| = k$, that is, if all μ_i are distinct, then the above condition is also sufficient and simplifies to

$$|\mathcal{S}| \le \frac{k(q-1)}{2} + 1 \equiv s_k. \qquad (9.10)$$

Observe also that for given $|\mathcal{R}| = r$, $s'(\mu_1, \dots, \mu_k)$ is maximum if and only if for all $\nu \in \mathcal{R}$, the identity $n_\nu - n'_\nu = 1$ holds. In this case, (9.9) simplifies similarly to (9.10). Namely,

$$\max_{|\mathcal{R}|=r} s'(\mu_1, \dots, \mu_k) = \frac{r(q-1)}{2} + (k-r)q + 1 \equiv s_r. \qquad (9.11)$$

Proof: Observe first the trivial fact that the maximum of $|\mathcal{S}|$ is achieved for cycle covers, each cycle being associated to a different site i in $\mathcal{S}$. It is also clear that each loop is a cycle, and thus contained in any cycle cover. They contribute $\sum_{\nu \in \mathcal{R}} q n'_\nu$ to the maximum number of cycles. We may therefore assume the graph $\mathcal{G}$ to be free of loops.

Claim: There exist walks for which the bound (9.9) is assumed.

Indeed, consider the associated graph $\mathcal{G}^q$ and cover it in the following way. For each $t = 1, \dots, k$, cover $q-1$ edges of $(e_{l,t})_l$ with 2-cycles. Without restricting the generality, we may assume that these are the edges $(e_{l,t})_{l=1,\dots,q-1}$. This yields $\sum_{\nu \in \mathcal{R}} (n_\nu - n'_\nu)(q-1)/2$ cycles (since loops have been removed). The remaining uncovered edges are isomorphic to the graph $\mathcal{G}$ (with all loops removed). Since a cycle is closed, $\mathcal{G}$ can be covered by at most $1 + \sum_{\nu \in \mathcal{R}} (n_\nu - n'_\nu - 1)$ cycles. This proves the claim.

We now prove that no cycle cover can have more cycles than the one constructed above. Indeed, suppose that $\mathcal{C} = \{C_1, \dots, C_n\}$ is not of the form described above. Then there exists a

$t \in \{1, \ldots, k\}$ and an odd number $p \geq 3$ such that there are p edges labeled (without restricting the generality) $\{e_{l,t}\}_{l=1,\ldots,p}$ with different endpoints (ν, ν') such that for each l, the cycle $\tilde{C}_l$ containing $e_{l,t}$ contains only edges with a different label $t' \neq t$.

Construct a new cover $\mathcal{C}'$ as follows. Leave all cycles $C \in \mathcal{C} \setminus (\bigcup_{l=2}^{p} \tilde{C}_l)$ unchanged. For each $j = 1, \ldots, \frac{p-1}{2}$, let

$$C'_{2j} = \{e_{t,2j}, e_{t,2j+1}\}, \tag{9.12}$$

and

$$C'_{2j+1} = \{e \in (\tilde{C}_{2j} \cup \tilde{C}_{2j+1}) \setminus C'_{2j}\}. \tag{9.13}$$

The new cover $\mathcal{C}'$ contains now all cycles in $\mathcal{C} \setminus (\bigcup_{l=2}^{p} \tilde{C}_l)$, the 2-cycles $\{C'_{2j}\}$, and all cycles obtained from decomposing the circuits $\{C'_{2j+1}\}$. This operation covers the edges $e_{1,t} \ldots, e_{q,t}$ with the maximum number of 2-cycles, namely $\frac{q-1}{2}$. Moreover, the number of cycles does not decrease.

Repeating this procedure for each $t = 1, \ldots, k$ yields a cycle cover of the type constructed in the first step of the proof, for which the bound (9.9) holds. This proves the lemma. □

Lemma 9.4': *If q is even, then for any given sequence $(\mu_1, \ldots, \mu_k)$, a necessary condition on $\mathcal{S}$ for the existence of an admissible walk is that*

$$|\mathcal{S}| \leq s'(\mu_1, \ldots, \mu_k) \equiv \sum_{\nu \in \mathcal{R}} (n_\nu + n'_\nu)\frac{q}{2} \tag{9.14}$$

Proof: The proof is immediate as soon as one recognizes that the graph $\mathcal{G}^q$ can be covered with 2-cycles only. □

For the rest of the calculations, we concentrate on the case of odd q. The following lemmas are just bounds on the different sums in (9.3). We now know how big $\mathcal{S}$ can be. That is, the sum over s in (9.5) does in general not extend to kq, but only to some smaller value s'. As shown in Lemma 9.4, respectively Lemma 9.4', s' is a function of the sequence $(\mu_1, \ldots, \mu_k)$, that is, $s' = s'((\mu_1, \ldots, \mu_k))$. Let us thus count the number of admissible walks for a given sequence $(\mu_1, \ldots, \mu_k)$ and a fixed $\mathcal{S}$, with $|\mathcal{S}| = s \leq s'$. From Lemma 9.1 we know that this means that we have to decompose the graph $\mathcal{G}$ into $s + j$ elementary cycles (with $j \geq 0$), and assign to each of them an element of $\mathcal{S}$. To estimate the number of circuit covers, we will use a uniform bound.

Lemma 9.5: *The number $W_{\mu_1,\ldots,\mu_k,\mathcal{S}}$ of admissible walks for a given sequence $(\mu_1, \ldots, \mu_k)$ and a given set $\mathcal{S}$, with $s = |\mathcal{S}| \leq s'$ is bounded by*

$$W_{\mu_1,\ldots,\mu_k,\mathcal{S}} \leq \Gamma(\mathcal{G}^q) \sum_{j=0}^{s'-s} s! \binom{s+j}{s} s^j \leq \Gamma(\mathcal{G}^q) \begin{cases} 2s! \binom{s'}{s} s^{s'-s}, & \text{if } s \geq 2; \\ \frac{s'(s'+1)}{2}, & \text{if } s = 1; \end{cases} \tag{9.15}$$

where

$$\Gamma(\mathcal{G}^q) = 2^{-q\sum_{\nu \in \mathcal{R}}(n_\nu - n'_\nu)} \prod_{\nu \in \mathcal{R}} \frac{(2q(n_\nu - n'_\nu))!}{(q(n_\nu - n'_\nu))!}. \tag{9.16}$$

Proof: It is clear from the above discussion that

$$W_{\mu_1,\dots,\mu_k,\mathcal{S}} = \sum_{j=0}^{s'-s} \#\{\text{arrangements of } s+j \text{ elements in } s \text{ cells, each being occupied}\} \\ \times \#\{\text{number of elementary partitions of } \mathcal{G} \text{ into } s+j \text{ cycles}\}. \tag{9.17}$$

The first number is bounded from above by $s!\binom{s+j}{s}s^j$. The first factor counts the arrangements of the cells. Then we choose s elements amongst $s+j$ and put each of them in a cell. Finally, the remaining j elements are distributed on the cells, but with no restriction.

We still have to estimate the number of elementary partitions. We do this uniformly, by using the following lemma. First, observe that since each loop is an elementary cycle, and thus appears in any elementary partition, we only have to consider the graph $\mathcal{G}'$, which is obtained from $\mathcal{G}^q$ by removing all loops.

Lemma 9.6: *The graph $\mathcal{G}'$ is an even Euler graph. The number $\Gamma(\mathcal{G}')$ of partitions into (not necessarily elementary cycles) is given by*

$$\Gamma(\mathcal{G}') = 2^{-q\sum_{\nu\in\mathcal{R}}(n_\nu - n'_\nu)} \prod_{\nu\in\mathcal{R}} \frac{(2q(n_\nu - n'_\nu))!}{(q(n_\nu - n'_\nu))!}. \tag{9.18}$$

Proof: The property of $\mathcal{G}'$ is obvious by construction. The formula (9.18) has been obtained by [K] and [Be] (see also [F]). In fact, it follows easily from the observation that each partition determines uniquely a pairing of the edges incident at each vertex. Since there are $2q(n_\nu - n_{\nu'})$ incident edges at each vertex, the number of pairings at each $\nu \in \mathcal{R} = \mathcal{V}(\mathcal{G}')$ is given by $(2q(n_\nu - n'_\nu) - 1)!! = 2^{-q}\frac{(2q(n_\nu - n'_\nu))!}{(p(n_\nu - n'_\nu))!}$. Equation (9.18) follows. □

This gives us the prefactor $\Gamma(\mathcal{G})$ in (9.15). We now sum over all allowed s. Suppose that $s \geq 2$. Then, the ratio $r(s,j)$ between two consecutive terms on the right-hand side of (9.15) is bounded by

$$r(s,j) = \frac{(s+j+1)!\,j!}{(s+j)!\,(j+1)!}s = \frac{s+j+1}{j+1}s \geq 2, \quad \forall s \geq 2. \tag{9.19}$$

Hence, the whole sum is less than twice the maximum term, which is the last one. If $s = 1$, then

$$\sum_{j=0}^{s'-s} \binom{s+j}{s} s^j = \sum_{j=0}^{s'-1} \binom{j+1}{1} = \sum_{j=0}^{s'-1} j = \frac{s'(s'+1)}{2}. \tag{9.20}$$

This proves Lemma 9.5. □

From the preceding lemma, we immediately get an estimate of the number of admissible walks for a given sequence $(\mu_1, \dots, \mu_k)$. Indeed, this number is bounded by

$$\sum_{s=1}^{s'} \sum_{\substack{\mathcal{S}\subset\mathcal{N}\\|\mathcal{S}|=s}} W_{\mu_1,\dots,\mu_k,\mathcal{S}} \leq \sum_{s=2}^{s'} 2s!\binom{N}{s}\Gamma(\mathcal{G}^q)\binom{s'}{s}s^{s'-s} \quad + N\Gamma(\mathcal{G}^q)\frac{s'(s'+1)}{2}. \tag{9.21}$$

Lemma 9.7: *If $k \leq N^{\frac{1}{2}-\varepsilon}$, for some $\varepsilon > 0$ and all N large enough, then for a fixed sequence $(\mu_1, \dots, \mu_k)$ and arbitrary $\mathcal{S}$, the number $\tilde{W}_{\mu_1,\dots,\mu_k}$ of admissible walks is bounded by*

$$\tilde{W}_{\mu_1,\dots,\mu_k} = \sum_{s=1}^{s'} \sum_{\substack{\mathcal{S}\subset\mathcal{N} \\ |\mathcal{S}|=s}} W_{\mu_1,\dots,\mu_k,\mathcal{S}} \leq 4\Gamma(\mathcal{G}^q)s'! \binom{N}{s'}. \tag{9.22}$$

Proof: Let us calculate the ration between two consecutive terms in the sum in (9.21). We get (for $2 \leq s \leq s'$),

$$\frac{\binom{N}{s+1}\binom{s'}{s+1}(s+1)!\,(s+1)^{s'-s-1}}{\binom{N}{s}\binom{s'}{s}s!\,s^{s'-s}} \geq \frac{(N-s)}{s(s+1)}. \tag{9.23}$$

If $s' \leq N^{\frac{1}{2}-\varepsilon}$, for some $\varepsilon > 0$ and all N large enough, then the right-hand side is greater than (say) 3 for all N large enough, uniformly in all allowed s. This means that the sum over all $s \geq 2$ is less than three halves of the largest summand, since this sum is dominated by

$$\binom{N}{s'}s'!\sum_{s=2}^{s'} 2^{s-s'} \leq \frac{3}{2}\binom{N}{s'}s'!. \tag{9.24}$$

Finally, the term with $s = 1$ is bounded by

$$\frac{s'(s'+1)}{2}s'!N \leq \frac{1}{2}\binom{N}{s'}s'!, \tag{9.25}$$

for all N large enough (and under the hypothesis on s'). Adding (9.24) and (9.25) thus gives

$$\sum_{s=1}^{s'} s!\binom{N}{s}\binom{s'}{s}s^{s'-s} \leq 2\binom{N}{s'}s'!, \tag{9.26}$$

that is, the whole sum is less than twice the maximum term.

We observe that since $s' \leq kq$, the above hypothesis on s' is always satisfied if $k \leq N^{\frac{1}{2}-\varepsilon'}$, for some $\varepsilon' > 0$ and N large enough. This proves the lemma. □

We now check that given a set $\mathcal{R}$ and multiplicities n_ν, the above estimate is maximum for those sequences for which $n_\nu - n'_\nu - 1 = 0$, for all $\nu \in \mathcal{R}$, that is, for sequences in which all multiple appearances are sub-sequential. This means that the additional number of partitions of $\mathcal{G}^q$ we get never wins against the loss of s'.

Lemma 9.8: *Suppose that $k \leq N^{\frac{1}{2}-\varepsilon}$, for some $\varepsilon > 0$. Then, given a sequence $(\mu_1, \dots, \mu_k)$ with $|\{\mu_1, \dots, \mu_k\}| = r$, there exists an $\bar{N} \in \mathbb{N}$ such that the number of admissible walks is always less than*

$$W'_{k,r} \equiv 2^{2-qr} \prod_{\nu\in\mathcal{R}} \frac{2q!}{q!}\binom{N}{\tilde{s}}\tilde{s}!, \tag{9.27}$$

for all $N \geq \bar{N}$.

Proof: What the lemma says is that for given set $\mathcal{R}$ and multiplicities $\{n_\nu\}_{\nu\in\mathcal{R}}$, the number of admissible walks is increasing in all n'_ν. Since they can be chosen independently once $\mathcal{R}$ and the multiplicities are fixed, this means that the number of admissible walks is maximized for maximum n'_ν. More precisely, if $(\mu_1,\dots,\mu_k)$ is a sequence such that $n_\nu - n'_\nu - 1 = 0$, for all $\nu\in\mathcal{R}$, then for any permutation of the indices $1,\dots,k$, the number of admissible walks cannot be greater, i.e.

$$\tilde{W}_{\mu_1,\dots,\mu_k} \geq \tilde{W}_{\mu_{\pi(1)},\dots,\mu_{\pi(k)}}, \tag{9.28}$$

where π is any element of the permutation group of k elements.

This is what one would expect, of course. Indeed, the factor $s'!\binom{N}{s'}$ is approximately (i.e. with correction terms of lower order in N) equal to $N^{s'}$. From (9.9) and the remark following that lemma, we see that s' is increasing in each n'_ν, that is , each time n'_ν is increased by one, we win some power of N. On the other hand, the combinatorial factor $\Gamma(\mathcal{G}^q)$ does not depend on N. Increasing n'_ν by one, we loose at most a power of k, which is not comparable to the gain if k is of order $o(N)$.

Let us make this arument rigorous. We compare the number of admissible walks $\tilde{W}_{\mu_1,\dots,\mu_k}$ and $\tilde{W}_{\tilde{\mu}_1,\dots,\tilde{\mu}_k}$ for two sequences $(\mu_1,\dots,\mu_k)$ and $(\tilde{\mu}_1,\dots,\tilde{\mu}_k)$ with same range $\mathcal{R}$ and same multiplicities $\{n_\nu\}_{\nu\in\mathcal{R}}$, and same numbers n'_ν except for one. That is, there exists $\rho\in\mathcal{R}$, such that $n'_\nu = \tilde{n}'_\nu$ for all $\nu\in\mathcal{R}\setminus\rho$, and $\tilde{n}'_\rho = n'_\rho + 1$.

Then the ratio of the associated combinatorial factors $\Gamma(\tilde{\mathcal{G}}q)$ and $\Gamma(\mathcal{G}^q)$ is given by

$$\begin{aligned}\frac{\Gamma(\tilde{\mathcal{G}}')}{\Gamma(\mathcal{G}^q)} &= 2^{-q\sum_{\nu\in\mathcal{R}}(\tilde{n}_\nu-\tilde{n}'_\nu)+q\sum_{\nu\in\mathcal{R}}(n_\nu-n'_\nu)}\prod_{\nu\in\mathcal{R}}\frac{(2q(\tilde{n}_\nu-\tilde{n}'_\nu))!}{(q(\tilde{n}_\nu-\tilde{n}'_\nu))!}\frac{(q(n_\nu-n'_\nu))!}{(2q(n_\nu-n'_\nu))!}\\ &= \prod_{j=0}^{q-1}(2q(n_\nu-n'_\nu)+2j+1)^{-1}.\end{aligned} \tag{9.29}$$

On the other hand, the difference of maximum sizes of the sets $\mathcal{S}$ are given by (9.9), and are equal to

$$\tilde{s}' - s' = (\tilde{n}'_\nu - n'_\nu)\frac{q-1}{2} = \frac{q-1}{2}, \tag{9.30}$$

so that

$$\frac{\tilde{s}'!\binom{N}{\tilde{s}'}}{s'!\binom{N}{s'}} = \frac{(N-s')!}{(N-\tilde{s}')!} \geq (N-\tilde{s}')^{\tilde{s}'-s'} \geq N^{\frac{q-1}{2}}(1-\mathcal{O}(\frac{kq^2}{N})). \tag{9.31}$$

If k is at most of order $o(N)$, then this shows that increasing each n'_ν to its maximum value $n_\nu - 1$ maximizes the number of admissible walks. Clearly, our hypothesis $k\leq N^{\frac{1}{3}-\varepsilon}$, for some $\varepsilon>0$ and N large enough, is more than enough for this to be true. □

Lemma 9.9: *Suppose that $k\leq N^{\frac{1}{2}-\varepsilon}$, for some $\varepsilon>0$ and all $N\geq\bar{N}$. Then the number of all admissible walks for given k and q is bounded by*

$$\begin{aligned}\mathcal{W}_k &\leq \sum_{r=1}^{k}\sum_{\substack{\mathcal{R}\subset\mathcal{M}\\|\mathcal{R}|=r}}\sum_{\substack{(\mu_1,\dots,\mu_k):\\\{\mu_1,\dots,\mu_k\}=\mathcal{R}}}\tilde{W}_{\mu_1,\dots,\mu_k}\\ &\leq 8((2q-1)!!)^k\binom{M}{k}k!\binom{N}{\frac{k(q-1)}{2}+1}(\frac{k(q-1)}{2}+1)!,\end{aligned} \tag{9.32}$$

for all N larger than $\bar{N}$.

Proof: For fixed $r \geq 2$, the number of sets $\mathcal{R}$ with cardinality r is clearly $\binom{M}{r}$. The number of possible sequences $(\mu_1, \dots, \mu_k)$ with given range $\mathcal{R}$ is less than $\binom{k}{r} r^{k-r} r!$, by the same argument we used already in the proof of Lemma 9.7. Thus, using Lemma 9.8,

$$\begin{aligned}\sum_{\substack{\mathcal{R}\subset\mathcal{M}\\|\mathcal{R}|=r}} \sum_{\substack{(\mu_1,\dots,\mu_k):\\ \{\mu_1,\dots,\mu_k\}=\mathcal{R}}} \tilde{W}_{\mu_1,\dots,\mu_k} &\leq \sum_{\substack{\mathcal{R}\subset\mathcal{M}\\|\mathcal{R}|=r}} \binom{k}{r} r^{k-r} r!\, W'_{k,r} = \binom{M}{r}\binom{k}{r} r^{k-r} r!\, W'_{k,r} \\ &\leq 4\binom{M}{r}\binom{k}{r} r^{k-r} r!\, 2^{-qr}\left(\frac{(2p)!}{p!}\right)^r \binom{N}{\tilde{s}_r},\end{aligned} \tag{9.33}$$

where s_r defined in (9.11) is the maximum value of s' for a given r.

If one shows that the ratio between two consecutive terms (as functions of r) is greater than e, then the sum over all terms is less than $\frac{e}{e-1}$ times the last term (i.e. the one for which $r = k$). For $2 \leq r \leq k-1$, we prove this by bounding the derivative of the logarithm from below by 1. Indeed, if

$$Q(r) \equiv \ln\left[\binom{M}{r}\binom{k}{r} r^{k-r} r!\, 2^{-qr}\left(\frac{(2p)!}{p!}\right)^r \binom{N}{s_r}\right], \tag{9.34}$$

then

$$Q'(r) = \ln(M-r) - 2\,\ln r + \ln(k-r) + \frac{k-r}{r} - \frac{q+1}{2}\ln(N-s_r) + c_1(q). \tag{9.35}$$

Using the hypothesis on k and the fact that $q' \geq q \geq 2$, this can be bounded from below by

$$\begin{aligned}Q'(r) &\geq \ln(M-k) - 2\,\ln k - \frac{q+1}{2}\ln N + c_1(q) \\ &\geq \ln M - 2\,\ln k - \frac{q+1}{2}\ln N + \ln(1-\frac{k}{M}) + c_1(q) \\ &\geq \varepsilon\ln N + \ln(1-\frac{k}{M}) + c_1(q) \\ &> c_2 > 1, \quad N \geq \bar{N}.\end{aligned} \tag{9.36}$$

The ratio between the terms with $r = k$ and $r = k-1$ is easily shown to be larger than e for all N large enough (this is due to the growth of M). This shows that when summing the terms in (9.33) over $2 \leq r \leq k$, each summand is at least e times as large as its predecessor. Thus,

$$\begin{aligned}\sum_{r=2}^{k}\sum_{\substack{\mathcal{R}\subset\mathcal{M}\\|\mathcal{R}|=r}} W'_{k,r} &\leq 4\,W'_{k,r}\sum_{r=2}^{k} e^{r-k} \leq 4\frac{e}{e-1}W'_{k,r} \\ &= 4\frac{e}{e-1}\binom{M}{k} k!\,((2q-1)!!)^k\binom{N}{s_k} s_k!.\end{aligned} \tag{9.37}$$

Finally, we consider the term for $r = 1$. In this case, $\mathcal{G}$ is composed of only one vertex and kq loops attached to it. Thus there are $N^{kq}M$ walks (one N for each loop times the number of sets $\mathcal{R}$, with $|\mathcal{R}| = 1$). A comparison with the right-hand side of (9.37) shows that for N large enough,

$$N^{qk}M \leq 4\frac{e-2}{e-1}\binom{M}{k} k!\,((2q-1)!!)^k\binom{N}{s_k} s_k!, \tag{9.38}$$

and hence,

$$\sum_{r=1}^{k}\sum_{\substack{\mathcal{R}\subset\mathcal{M}\\|\mathcal{R}|=r}} W'_{k,r} \le 8((2q-1)!!)^k \binom{M}{k} k! \binom{N}{s_k} s_k! \tag{9.39}$$

This proves the lemma. □

Lemma 9.9 expresses that fact that the overwhelming contribution comes from the terms with $r = k$. We collect the preceding lemmas in a proposition.

Proposition 9.10: *Suppose that q is odd and $k \le N^{\frac{1}{2}-\varepsilon}$ and even. Then*

$$\mathbb{E}\,\mathrm{tr}\,A^k \le 8((2q-1)!!)^k \binom{M}{k} k! \binom{N}{s_k} s_k! \quad . \tag{9.40}$$

Proof: The proof for odd q follows from (9.5), Lemma 9.1 and Lemma 9.9. For even q, one shows with exactly the same calculations that (9.40) still holds, with s_k now given by (9.14) in Lemma 9.4'. □

Proof of Theorem 8.1: As mentioned in the previous chapter, we use the fact that the trace of the kth power of A_q is an upper bound for the kth power of the norm, and Chebyshev's inequality, which yield

$$\begin{aligned} \mathbb{P}[\|A_q\| \ge cN^{(q-1)/2}] &\le \mathbb{P}[\mathrm{tr}{A_q}^k \ge c^k N^{k(q-1)/2}] \\ &\le \frac{1}{c^k N^{k(q-1)/2}} \mathbb{E}\,\mathrm{tr}\,A_q^k. \end{aligned} \tag{9.41}$$

Suppose that q is odd, and that c satisfies

$$\frac{c}{\prod_{j=0}^{(q-3)/2}(q-2j)} \equiv C > 1. \tag{9.42}$$

Then, using the Proposition 9.10, we bound the latter by

$$\begin{aligned} \frac{1}{c^k N^{\frac{k(q-1)}{2}}} \mathbb{E}\,\mathrm{tr}\,A_q^k &\le \frac{8}{c^k N^{\frac{k(q-1)}{2}} M^k} \frac{M!}{(M-k)!} \frac{N!}{(N - \frac{k(q-1)}{2} - 1)!} \prod_{j=0}^{\frac{q-3}{2}} (q-2j)^k \\ &\le \frac{8}{c^k N^{\frac{k(q-1)}{2}}} N^{\frac{k(q-1)}{2}+1} \prod_{j=0}^{\frac{q-3}{2}} (q-2j)^k \\ &= 8N \frac{\prod_{j=0}^{\frac{q-3}{2}} (q-2j)^k}{c^k} < 8C^{-k} N. \end{aligned} \tag{9.43}$$

Choose $k = N^{l'}$, with $l' < \frac{1}{2}$. Then,

$$\begin{aligned} \mathbb{P}[\|A_q\| \ge c_q N^{(q-1)/2}] &\le e^{-k \ln C + \ln N + \varepsilon} \\ & e^{-N^l \ln C}, \end{aligned} \tag{9.44}$$

for all $l < l'$ and N large enough. This proves the bound for odd q.

For even q, the result follows by exactly the same argument. The additional powers of N are needed to compensate for the greater maximum size s_k of the sets $\mathcal{S}$. □

9.3 Proofs of Theorem 8.2 and Theorem 8.3

The proof of Theorem 8.2 is similar, but simpler than the previous one. We first prove the statement for the matrix B_q. It is intuitively clear that the norm of the matrix B_q must be much smaller than the norm of A_q, since the main contribution to $\mathbb{E}\,\mathrm{tr}\,A^k$ came from walks which are obviously not allowed in the present case.

Lemma 9.11: *The maximum number of circuits forming a q-circuit cover of $\mathcal{G}$ is equal to q times the maximum number of cycles in a simple cycle cover of $\mathcal{G}$. This latter number is bounded by $1 + \sum_{\nu\in\mathcal{R}}(n_\nu - 1)$.*

Proof: It is obvious that the maximum is achieved for a cycle decomposition (otherwise, decompose all remaining circuits into cycles, thereby increasing the number).

Next, we show that each q-fold cycle cover $\mathcal{C}$ can be arranged into q simple covers. The claim is trivial for $q = 1$ and for the case that $\mathcal{G}$ is itself a cycle. Assume it to be true for $q - 1$. Choose an arbitrary edge $e_0 \in \mathcal{E}(\mathcal{G})$ and a cycle $C_0 \subset \mathcal{G}$ containing e_0. Since $C_0 \neq \mathcal{G}$, there exists an edge $e_1 \in \mathcal{E}(\mathcal{G}) \setminus \mathcal{E}(C_0)$. There is also a cycle C_1 containing e_1 such that $\mathcal{E}(C_0) \cap \mathcal{E}(C_1) = \emptyset$. Indeed, if this were not true, there would exist an edge $f \in \mathcal{E}(C_0)$ that is covered at least $q + 1$ times. Repeat this procedure until $\mathcal{G}$ is simply covered by the edge disjoint collection $\{C_1, \dots, C_n\}$.

The remaining cycles $\mathcal{C} \setminus \{C_0, \dots, C_n\}$ form a $q - 1$ cycle cover $\mathcal{G}$, for which the assertion is true by the induction hypothesis. This proves that the maximum number of circuits forming a q-fold cover of $\mathcal{G}$ is equal to q times the maximum number in a simple cycle cover. This number is now easily calculated. Each cycle is closed by definition. Therefore, to form a cycle, there must be a ν appearing twice in the sequence $(\mu_1, \dots, \mu_k)$. Thus, we get at most $\sum_\nu (n_\nu - 1)$ cycles, plus one which accounts for the fact that the walk returns from μ_k to μ_1. This proves the lemma. □

Corollary 9.12: *For given $r = |\mathcal{R}|$, the maximum number of circuits forming a q-fold cover of any associated $\mathcal{G}$ is given by $s' = s'(k, q) = q(k - r + 1)$.*

Proof: By the previous lemma, we want to maximize $\sum_\nu (n_\nu - 1)$ under the constraint $\sum_\nu n_\nu = k$. But $\sum_\nu (n_\nu - 1) = k - \sum_\nu 1 = k - r$ □

Lemma 9.13: *The number $W_{\mu_1,\dots,\mu_k,\mathcal{S}}$ of B-admissible walks, given a sequence $(\mu_1, \dots, \mu_k)$ and a set $\mathcal{S}$, with $s = |\mathcal{S}| \le q(k - r + 1)$, is bounded by*

$$
\begin{aligned}
W_{\mu_1,\dots,\mu_k,\mathcal{S}} &\le (\Gamma(\mathcal{G}))^q \sum_{j=0}^{s'-s} s! \binom{s+j}{s} s^j (q!)^k \\
&\le (q!)^k \Gamma(\mathcal{G})^q \begin{cases} 2s!\,\binom{s'}{s} s^{s'-s}, & \text{if } s \ge 2; \\ \frac{s'(s'+1)}{2}, & \text{if } s = 1; \end{cases}
\end{aligned}
\tag{9.45}
$$

where

$$\Gamma(\mathcal{G}) = 2^{-\sum_{\nu\in\mathcal{R}}(n_\nu - n'_\nu)} \prod_{\nu\in\mathcal{R}} \frac{(2(n_\nu - n'_\nu))!}{(n_\nu - n'_\nu)!} = \prod_{\nu\in\mathcal{R}} ((2(n_\nu - n'_\nu) - 1)!!). \tag{9.46}$$

Proof: The proof is almost identical to the proof of Lemma 9.5. To obtain the number of admissible walks, we count the number of q-cycle covers with $s+j \leq s'$ cycles, and assign to each an element of $\mathcal{S}$. As remarked in the proof of Lemma 9.2, there are at most $(q!)^k$ preimages of this circuit cover under the map ρ^{-1}. We thus get the second part on the right-hand side of (9.45). The number of simple circuit covers of $\mathcal{G}$ is again bounded as in Lemma 9.6 (inequality (9.18)), applied this time, however, to the graph $\mathcal{G}$ (which means removing the factor q everywhere). The final bound is again the same calculation as in the proof of Lemma 9.5. □

Corollary 9.14: *If $k \leq N^{\frac{1}{2}-\varepsilon}$, for some $\varepsilon > 0$ and all N large enough, then for a fixed sequence $(\mu_1, \dots, \mu_k)$ the number $\tilde{W}_{\mu_1,\dots,\mu_k}$ of admissible walks is bounded by*

$$\tilde{W}_{\mu_1,\dots,\mu_k} = \sum_{s=1}^{s'} \sum_{\substack{\mathcal{S}\subset\mathcal{N} \\ |\mathcal{S}|=s}} W_{\mu_1,\dots,\mu_k,\mathcal{S}} \leq 4\,(\Gamma(\mathcal{G}))^q\,(q!)^k s'! \binom{N}{s'}, \tag{9.47}$$

where $s' = q(k-r+1)$. Moreover, for all sequences $(\mu_1, \dots, \mu_k)$ with the same range $\mathcal{R}$, the number of admissible walks $W_{k,r}$ is bounded by

$$W_{k,r} \leq 4((2(k-r)-1)!!)^q (q!)^k N^{s'(k,r)}. \tag{9.48}$$

Proof: The proof of (9.47) is the same as the proof of Lemma 9.7. The second bound is implied by the fact that $\Gamma(\mathcal{G})$ is bounded by $(2(k-r)-1)!!$ □

The proof of Theorem 8.2 is similar to the proof of Theorem 8.1.

Proof of Theorem 8.2: As in the case of the matrix A (compare (9.33)), we have

$$\begin{aligned} \mathbb{E}\,\mathrm{tr} B^k &= M^{-k}(q!)^{-k} \sum_{r=1}^{k} \sum_{\substack{\mathcal{R}\subset\mathcal{M} \\ |\mathcal{R}|=r}} \sum_{\substack{(\mu_1,\dots,\mu_k): \\ \{\mu_1,\dots,\mu_k\}=\mathcal{R}}} \tilde{W}_{\mu_1,\dots,\mu_k} \\ &\leq M^{-k}(q!)^{-k} \sum_{r=1}^{k} \sum_{\substack{\mathcal{R}\subset\mathcal{M} \\ |\mathcal{R}|=r}} \binom{k}{r} r^{k-r} r! W_{k,r}. \end{aligned} \tag{9.49}$$

Inserting the bound (9.48) yields,

$$\begin{aligned} \mathbb{E}\,\mathrm{tr} B^k &\leq M^{-k}(q!)^{-k} \sum_{r=1}^{k} \binom{M}{r} N^{s'(k,r)} \binom{k}{r} r! r^{k-r} (q!)^k ((2(k-r)-1)!!)^q \\ &\leq \sum_{r=1}^{k} M^{r-k} N^{q(k-r)+q} \binom{k}{r} r^{k-r} ((2(k-r)-1)!!)^q. \end{aligned} \tag{9.50}$$

Since $q' > q$, the ratio between two consecutive terms is always greater than one for N large enough, if $k \le N^{\frac{1}{2}-\varepsilon}$. Thus, the whole sum is less than a constant times the largest term,

$$\mathbb{E}\,\mathrm{tr}\,B^k \le C_1 N^{s'(k,k)} = C_1 N^q. \tag{9.51}$$

As in the proof of Theorem 8.1, we conclude with the Chebyshev's inequality, that is, for all $c > 1$,

$$\mathbb{P}\Big[\|B\| \ge c\Big] \le \mathbb{P}\Big[\mathrm{tr}B^k \ge c\Big] \le c^{-k}\mathbb{E}\,\mathrm{tr}\,B^k. \tag{9.52}$$

Use (9.51) and choose $k = N^{\frac{1}{2}-\varepsilon}$, for some $\varepsilon > 0$. Hence, for all $C_3 > 0$, there exists an $\bar{N} \in \mathbb{N}$ such that for all $N \ge \bar{N}$,

$$\begin{aligned}\mathbb{P}[\|B\| \ge c] \le \frac{C_1 N^q}{c^k} &= e^{-k\ln c + q\ln N + \ln C_1}\\ &\le e^{-C_2 N^{\frac{1}{2}-\varepsilon}}.\end{aligned} \tag{9.53}$$

This proves (8.5). To establish the result for the matrix B'_q, observe that

$$\mathrm{tr}{B_q}^k = \mathrm{tr}{B'_q}^k, \tag{9.54}$$

which follows at once from the explicit formula for the two expressions. The bound on the largest eigenvalue of B'_q then follows immediately. □

Proof of Theorem 8.3: The proof of this result follows easily from the above bound. As before, the key to the result is an upper bound for the size of $\mathcal{S}$. Since each edge of $\mathcal{G}$ has to be covered by C_i (this is the circuit associated to the site i), C_i itself is a 1-cover of $\mathcal{G}$. This implies that $|\mathcal{S} \setminus \{i\}|$ is bounded by $(q-1)(k-r+1)$.

The combinatorial calculations that follow are exactly the same as in the previous proof. We therefore obtain the following analogue to (9.50),

$$\mathbb{E}\,\mathrm{tr}\,B^k \le \sum_{r=1}^{k} M^{r-k} N^{(q-1)(k-r)+q}\binom{k}{r} r^{k-r}((2(k-r)-1)!!)^q. \tag{9.55}$$

Since $q' \ge q$, the ratio between two consecutive terms is always greater than one if N is large enough and $k \le N^{\frac{1}{2}-\varepsilon}$. Thus,

$$\mathbb{E}\,\mathrm{tr}\,B^k \le C_1 N^{q-1}. \tag{9.56}$$

The assertion of the theorem now follows by Chebyshev's inequality and the identity (9.52). □

Part III

The Hopfield Model with Gaussian Patterns

10 Main Results

In the last part of this thesis, we turn to the Gaussian variant of the Hopfield model, introduced in section 1.4. Let us recall its definition.

The configuration space is $\Sigma^N = \{-1,+1\}^N$ (resp. $\Sigma^\infty = \{-1,+1\}^\infty$ for the infinite system), equipped with the product topology. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space and let $\xi_i^\mu[\omega]$, $i \in \mathbb{N}$, $\mu = 1,2$, denote a family of i.i.d. standard Gaussian variables. We will write $\xi^\mu[\omega]$ for the N-dimensional vector whose ith component is given by $\xi_i^\mu[\omega]$; such a vector is called a *pattern*. On the other hand, we will write $\xi_i[\omega]$ for the two dimensional vector with components (ξ_i^1, ξ_i^2). When we write $\xi[\omega]$ without indices, we consider it as a $2 \times N$ matrix, whose transpose will be denoted by ξ^T.

Throughout the remaining chapters, $(\cdot,\cdot)$ denotes the scalar product, without indication of the space where its arguments lie. We define random maps $m_N^\mu[\omega](\sigma) : \Sigma_N \to [-1,+1]$ (conventionally called *overlap parameters*) through

$$m_N^\mu[\omega](\sigma) \equiv \frac{1}{N}\sum_{i=1}^N \xi_i^\mu[\omega]\sigma_i. \tag{10.1}$$

The Hamiltonian is now defined as

$$\begin{aligned} H_N[\omega](\sigma) &\equiv -\frac{N}{2}\sum_{\mu=1,2}\left(m_N^\mu[\omega](\sigma)\right)^2 \\ &= -\frac{N}{2}\|m_N[\omega](\sigma)\|_2^2, \end{aligned} \tag{10.2}$$

where $\|\cdot\|_2$ denotes the l_2-norm in $\mathbb{R}^2$.

As already remarked in the introduction (Section 1.5), the distribution of the disorder variables is invariant under the map ${\xi'}_i^1 = {\xi^\alpha}_i^1 = \xi_i^1\cos(\alpha) + \xi_i^2\sin(\alpha)$ and ${\xi'}_i^2 = {\xi'}_i^\alpha = \xi_i^1\sin(\alpha) - \xi_i^2\cos(\alpha)$. Moreover, the Hamiltonian has the same form in the original and the primed variables. However, this transformation is a *statistical* symmetry, as opposed to the spin-flip symmetry which is an exact symmetry for any given realization of the disorder.

Through this Hamiltonian, the finite volume Gibbs measures on Σ_N are defined by

$$\mathcal{G}_{N,\beta}[\omega](\sigma) \equiv 2^{-N}\frac{e^{-\beta H_N[\omega](\sigma)}}{Z_{N,\beta}[\omega]}, \tag{10.3}$$

and the induced distribution of the overlap parameters by

$$\mathcal{Q}_{N,\beta}[\omega] \equiv \mathcal{G}_{N,\beta}[\omega] \circ m_N[\omega]^{-1}. \tag{10.4}$$

The normalizing factor in (10.3), called the *partition function*, is explicitly given by

$$Z_{N,\beta}[\omega] \equiv 2^{-N}\sum_{\sigma\in\Sigma_N} e^{-\beta H_N[\omega](\sigma)} \equiv \mathbb{E}_\sigma\, e^{-\beta H_N[\omega](\sigma)}. \tag{10.5}$$

We are mainly interested in the concentration behavior of $\mathcal{Q}_{N,\beta}$ as $N \to \infty$. It will be convenient to do this by considering the auxiliary measure $\widetilde{\mathcal{Q}}_{N,\beta} \equiv \mathcal{Q}_{N,\beta} \star \mathcal{N}_2(0, \frac{1}{\beta N}\mathbb{I})$ obtained by a convolution with a Gaussian measure, its so-called *Hubbard-Stratonovich transform*. Since, for N large, $\mathcal{N}_2(0, \frac{1}{\beta N}\mathbb{I})$ converges in the weak sense rapidly to the Dirac measure at zero, the two measures have asymptotically the same properties. For details see e.g. [BGP1]. $\widetilde{\mathcal{Q}}_{N,\beta}$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^2$ and has the density

$$\frac{e^{-\beta N \Phi_{N,\beta}[\omega](z)}}{Z_{N,\beta}[\omega]}, \tag{10.6}$$

where $\Phi_{N,\beta}$ is given by

$$\Phi_{N,\beta}[\omega](z) = \frac{1}{2}\|z\|_2^2 - \frac{1}{\beta N}\sum_{i=1}^{N} \ln\cosh\beta(\xi_i[\omega], z). \tag{10.7}$$

As usual in mean-field models, we construct the extremal Gibbs measures by tilting the Hamiltonian (10.2) with an *external magnetic field*.[33] That is, we define a more general Hamiltonian

$$H_N^h[\omega](\sigma) \equiv -\frac{N}{2}\|m_N[\omega](\sigma)\|_2^2 - N(h, m_N[\omega](\sigma)), \tag{10.8}$$

where $h = (b\cos(\vartheta), b\sin(\vartheta)) \in \mathbb{R}^2$. The corresponding measures on the spins and on $\mathbb{R}^2$ are denoted by $\mathcal{G}_{N,\beta}^h[\omega]$ and $\mathcal{Q}_{N,\beta}^h[\omega]$, respectively. We then take the limits $\lim_{b\to 0}\lim_{N\to\infty}$, for all values of $\vartheta \in [0, 2\pi)$.

We are now able to give a precise formulation of our main results.

Theorem 10.1: *Let $h = (b\cos\vartheta, b\sin\vartheta)$. Then*

$$\lim_{b\to 0}\lim_{N\to\infty} \mathcal{Q}_{N,\beta}^h = \delta_{(r^*\cos\vartheta, r^*\sin\vartheta)}, \tag{10.9}$$

where r^* is the largest solution of the equation

$$r^* = \frac{1}{\sqrt{2\pi}}\int dx\, e^{-\frac{x^2}{2}} x\tanh(\beta x r^*). \tag{10.10}$$

From the form of (10.10), it is easy to see that $r^* = 0$ is always a solution. It is also straightforward to check that there exists a β^*, $0 < \beta^* < \infty$, such that the largest solution r^* is non-zero whenever $\beta > \beta^*$.

Theorem 10.1 shows that there is an uncountable number of extremal limiting induced measures, indexed by the circle. The following Corollary shows that to each of them corresponds a distinct limiting Gibbs measure on the spins.

Corollary 10.2: *For any finite set $I \subset \mathbb{N}$, and $\mathbb{P}$-almost all ω,*

$$\mathcal{G}_{\infty,\beta}^h[\omega]\Big(\{\sigma_I = s_I\}\Big) \equiv \lim_{b\to 0}\lim_{N\to\infty} \mathcal{G}_{N,\beta}^h[\omega]\Big(\{\sigma_I = s_I\}\Big) = \prod_{i\in I}\frac{e^{\beta s_i(\xi_i[\omega], m)}}{2\cosh(\beta(\xi_i[\omega], m))}, \tag{10.11}$$

[33] For a general discussion on the issue of limiting Gibbs states in mean-field models, see [BG1], Sect. 2.4 or [BG3], Sect. 2.

where $m = (r^* \cos(\vartheta), r^* \sin(\vartheta))$*, and* r^* *is as in (10.10).*

In Theorem 10.1 and Corollary 10.2 convergence is almost sure due to the presence of the tilting field. The situation changes if we set $b = 0$ first and take the infinite volume limit later. The result for the induced measures is given in

Theorem 10.3: *Let* $\mathcal{Q}_{N,\beta}$ *be as in (10.4) and* $m = m(\vartheta) = (r^* \cos\vartheta, r^* \sin\vartheta)$*, where* $\vartheta \in [0, \pi)$ *is a uniformly distributed random variable. Then*

$$\mathcal{Q}_{N,\beta} \xrightarrow{\mathcal{D}} \frac{1}{2}\delta_{m(\vartheta)} + \frac{1}{2}\delta_{-m(\vartheta)} \equiv \mathcal{Q}_{\infty,\beta}[m]. \tag{10.12}$$

Furthermore, the (induced) AW-metastate is the image of the uniform distribution of ϑ *on* $[0, 2\pi)$ *under the measure-valued map* $\vartheta \mapsto \mathcal{Q}_{\infty,\beta}[m(\vartheta)]$.

From this, we get the description on the level of the Gibbs measures:

Corollary 10.4: *Let* $I \subset \mathbb{N}$ *be finite. Then the following holds:*
(i) Let $\{g_i\}_{i\in I}$ *be a family of i.i.d. random variables, distributed as* $\mathcal{N}(0, r^*)$*. Then*

$$\lim_{N\uparrow\infty} \mathcal{G}_{N,\beta}(\{\sigma_I = s_I\}) \xrightarrow{\mathcal{D}} \frac{1}{2}\prod_{i\in I} \frac{e^{\beta s_i g_i}}{2\cosh \beta g_i} + \frac{1}{2}\prod_{i\in I} \frac{e^{-\beta s_i g_i}}{2\cosh \beta g_i}. \tag{10.13}$$

(ii) The AW-metastate is the image of the uniform distribution on ϑ *under the measure-valued map* $\vartheta \mapsto \mathcal{G}_{\infty,\beta,m(\vartheta)}[\omega]$ *where*

$$\mathcal{G}_{\infty,\beta,m}[\omega](\{\sigma_I = s_I\}) = \frac{1}{2}\prod_{i\in I} \frac{e^{\beta s_i(\xi_i[\omega],m)}}{2\cosh \beta(\xi_i[\omega], m)} + \frac{1}{2}\prod_{i\in I} \frac{e^{-\beta s_i(\xi_i[\omega],m)}}{2\cosh \beta(\xi_i[\omega], m)}. \tag{10.14}$$

Statement (ii) of Corollary 10.4 motivates the notion of metastates. Whereas on the level of the induced measures $\mathcal{Q}_{N,\beta}$ one cannot see any influence by the conditioning, this is clearly the case on the level of the Gibbs measures on the spins.

The remaining chapters are mainly devoted to the proofs of the two theorems (the corollaries are standard consequences (see e.g. [BGP1] or [BG3] for proofs of analogous statements in more complicated situation). They are organized as follows. In Chapter 11 we prove the necessary concentration estimates on the measures $\mathcal{Q}_{N,\beta}$, respectively $\mathcal{Q}^h_{N,\beta}$. This will yield immediately Theorem 10.1. In the case $h = 0$ we will show that the measure concentrates near the absolute minima of some random process, and in Section 12 we will analyze the properties of these minima. In particular we will prove that these converge in distribution to one-point sets. This will allow us to prove Theorem 10.3. In Section 13 we discuss some further consequences on the chaotic volume dependence, the empirical metastate and the superstate.

11 Concentration of the Induced Measures

In this chapter we show the concentration properties of the measures $\widetilde{\mathcal{Q}}_N$ for large β. These imply the same concentration results for the measures $\mathcal{Q}_N$ by standard arguments that have been developed in much more complicated situations, see e.g. [BG2]. The estimates presented here are mostly similar, and often much simpler, to those that can be found e.g. in [BG2], but we decided to present some parts in detail where some care is required.

We start with the more delicate case $h = 0$ that will be relevant for the proof of Theorem 10.3 (which will be given at the end of chapter 13). We are interested in the concentration behavior of the measures $\widetilde{\mathcal{Q}}_{N,\beta}$. The following two lemmata each give a partial answer. The first one asserts that $\widetilde{\mathcal{Q}}_{N,\beta}$ is concentrated exponentially about a circle around the origin, whereas the second one tells us that even on this circle, only a small part really contributes to the total mass.

Lemma 11.1: *Let $\{\xi_i^\mu\}_{i\in\mathbb{N},\mu=1,2}$ be i.i.d. standard Gaussian variables, and define $\Phi_{N,\beta}(z)$ as*

$$\Phi_N(z) \equiv \frac{1}{2}\|z\|_2^2 - \frac{1}{\beta N}\sum_{i=1}^{N} \ln\, \cosh \beta(\xi_i, z). \tag{11.1}$$

Let furthermore $\delta_N = N^{-\frac{1}{10}}$. Then there exist strictly positive constants K, K', l, l' such that (r^* is the largest solution in (10.10))

$$\frac{\int\limits_{|\,\|z\|-r^*|\geq\delta_N} e^{-\beta N\Phi_N(z)}\,dz}{\int\limits_{|\,\|z\|-r^*|<\delta_N} e^{-\beta N\Phi_N(z)}\,dz} \leq Ke^{-KN^l}, \tag{11.2}$$

on a set of $\mathbb{P}$-measure at least $1 - K'e^{-K'N^{l'}}$.

The second result needs an additional definition. Let

$$g_N(\vartheta) \equiv \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \ln\, \cosh(\beta r^*\zeta_i \cos(\vartheta - \varphi_i)), \tag{11.3}$$

where (ζ_i, φ_i) are the polar coordinates of the two dimensional vector ξ_i.

Lemma 11.2: *Assume the hypotheses of Lemma 11.1. Let $a_N = N^{-\frac{1}{25}}$. Then there exist strictly positive constants K_1, K_2, C_1, C_2 such that on a set of $\mathbb{P}$-measure at least $1 - K_1e^{-N^{\frac{1}{25}}}$, the following bound holds,*

$$\frac{\int\limits_{\mathcal{A}'_N} e^{-\beta N\Phi_N(z)}\,dz}{\int\limits_{\mathcal{A}_N} e^{-\beta N\Phi_N(z)}\,dz} \leq C_1e^{-N^{\frac{2}{5}}}, \tag{11.4}$$

where

$$\begin{aligned}\mathcal{A}_N &= \Big\{(r,\vartheta)\in\mathbb{R}_0^+\times[0,2\pi)\Big||r-r^*|<\delta_N, g_N(\vartheta)-\min_{\vartheta} g_N(\vartheta) < a_N\Big\},\\ \mathcal{A}'_N &= \Big\{(r,\vartheta)\in\mathbb{R}_0^+\times[0,2\pi)\Big||r-r^*|<\delta_N, g_N(\vartheta)-\min_{\vartheta} g_N(\vartheta) \geq a_N\Big\}.\end{aligned} \tag{11.5}$$

Combining these two lemmata and using the Borel-Cantelli lemma, we get immediately the following result.

Proposition 11.3: *Assume the hypotheses of Lemma 11.1. Then there exist strictly positive constants K, K', l, such that*

$$\mathbb{P}\left[\frac{\int_{\mathcal{A}_N^c} e^{-\beta N\Phi_N(z)}\,dz}{\int_{\mathcal{A}_N} e^{-\beta N\Phi_N(z)}\,dz} > Ke^{-K'N^l},\ i.o.\ in\ N\right] = 0, \tag{11.6}$$

where $\mathcal{A}_N$ is as in Lemma 11.2.

To see why the preceding results should be expected, we must consider the function $\Phi_{N,\beta}$. Note that the expectation of this function,

$$\mathbb{E}\Phi_N(z) = \frac{1}{2}\|z\|_2^2 - \frac{1}{\beta}\mathbb{E}\ \ln\cosh\beta(\xi_1, z). \tag{11.7}$$

depends only on the modulus of its argument. It is useful to observe that if $z = (r\cos\theta, r\sin\theta)$, we can represent $\mathbb{E}\Phi_N(z)$ as

$$\mathbb{E}\Phi_N(z) = \frac{1}{2}r^2 - \mathbb{E}_\varphi\mathbb{E}_\zeta\ \ln\cosh(\beta r\zeta\cos(\varphi)) \tag{11.8}$$

where ζ, ϕ are the representation of the polar decomposition of a two dimensional normal vector, i.e. ζ is distributed with density $xe^{-x^2/2}$ on $\mathbb{R}^+$, and φ uniformly on $[0, 2\pi)$.

From (11.8), choosing $\theta = 0$, it follows that $\mathbb{E}\Phi_N(z)$ takes its minimum on the circle with radius $r^*(\beta)$, where r^* is defined in Theorem 10.1. As remarked after the statement of Theorem 10.1, there exists $0 < \beta^* < \infty$, such that $r^*(\beta) > 0$ if and only if $\beta > \beta^*$.

It is also straightforward to check that $\mathbb{E}\Phi$ is sufficiently smooth to guarantee that it is bounded from above by a quadratic function (of $\|z\|$) in some neighborhood containing r^*.

Proof of Lemma 11.1: We start with the numerator. We decompose the domain of integration into an "inner" part $\mathcal{I}$, and an "outer" part $\mathcal{O}$:

$$\begin{aligned}\left\{z \in \mathbb{R}^2 : |\|z\| - r^*| \geq \delta\right\} &= \left\{z \in \mathbb{R}^2 : \|z\| - r^* \geq \delta\right\} \cup \left\{z \in \mathbb{R}^2 : \|z\| - r^* \leq -\delta\right\} \\ &\equiv \mathcal{O} \cup \mathcal{I}.\end{aligned} \tag{11.9}$$

Consider the integral on $\mathcal{O}$. We write it as

$$\int_{\mathcal{O}} e^{-N\Phi_N(z)}\,dz = \int_{\mathcal{O}} e^{-\beta N\mathbb{E}\,\Phi_N(z)}e^{-\beta N(\Phi_N(z) - \mathbb{E}\,\Phi_N(z))}\,dz, \tag{11.10}$$

and observe that $\mathbb{E}\Phi_N$ can also be bounded below by a quadratic function $C(\|z\| - r^*)^2$. We are left with the task of estimating the term $\Phi_N(z) - \mathbb{E}\Phi_N(z)$. This is accomplished by the following lemma.

Lemma 11.4: *Let* $f_N(z) = \frac{1}{\beta N}\sum_{i=1}^N \ln\cosh\beta(\xi_i, z)$ *and*

$$\mathcal{O} = \left\{ z \in \mathbb{R}^2 : \|z\| > r^* + \delta \right\}. \tag{11.11}$$

Then, for δ *small enough, such that* $\delta^2/16 \le \delta/2\sqrt{2}$*, there exist strictly positive constants* C_1*,* C_2*,* K_1*,* K_2 *such that*

$$\mathbb{P}\left[\sup_{z\in\mathcal{O}} |f_N(z) - \mathbb{E}f_N(z)| \ge \frac{C}{2}(\|z\| - r^*)^2\right] \le K_1 e^{-K_2 N} + C_1\delta^{-2} e^{-C_2\delta^4 N} N^{-\frac{1}{2}}. \tag{11.12}$$

Proof: Define $\bar{f}_N(z) = f_N(z) - \mathbb{E} f_N(z)$. The left-hand side of (11.12) is bounded from above by

$$\begin{aligned} &\mathbb{P}\left[\sup_{z\in\mathcal{O}} |f_N(z) - \mathbb{E}f_N(z)| \ge \frac{C}{2}(\|z\| - r^*)^2\right] \\ &\qquad \le \mathbb{P}\left[\sup_{z'\in\mathcal{W}_r\cap\mathcal{O}} \left|\bar{f}_N(z')\right| \ge \frac{C}{4}\left(\|z'\| - r^*\right)^2\right] \\ &\qquad + \mathbb{P}\left[\sup_{z'\in\mathcal{W}_r\cap A}\ \sup_{z\in B_r(z')} \left|\bar{f}_N(z) - \bar{f}_N(z')\right| \ge \frac{C}{4}\left(\|z'\| - r^*\right)^2\right], \end{aligned} \tag{11.13}$$

where $\mathcal{W}_r$ is the grid with spacing r in $\mathbb{R}^2$, and $z' \in \mathcal{W}_r$ is chosen such that $0 \le \|z\| - \|z'\| < \sqrt{2}\, r$.

The argument of the second term can be uniformly bounded. Using e.g. Lemma 6.10 of [BG1], we get that

$$|f_N(z) - f_N(z')| \le \|z - z'\|_2 \|A\|^{\frac{1}{2}}, \tag{11.14}$$

where A is the matrix $\frac{1}{N}\xi^T\xi$. Similarly,

$$|\mathbb{E} f_N(z) - \mathbb{E} f_N(z')| \le \|z - z'\|_2 (\mathbb{E}\|A\|)^{\frac{1}{2}}. \tag{11.15}$$

A trivial computation shows that

$$\mathbb{E}\|A\| \le 1 + C/\sqrt{N} \tag{11.16}$$

and using (for instance) the same argument as in Section 4 of [BG1], but replacing Talagrand's concentration estimate for bounded r.v.'s by the standard Gaussian concentration inequality (see e.g. [LT], Chapter 1), one shows easily that

$$\mathbb{P}\left[|\|A\| - 1| \ge x\right] \le C e^{-Nx^2/C}. \tag{11.17}$$

Therefore,

$$\begin{aligned} &\mathbb{P}\left[\sup_{z'\in\mathcal{W}_r\cap A}\ \sup_{z\in B_r(z')} |\bar{f}_N(z) - \bar{f}_N(z')| \ge \frac{C}{4}(\|z'\| - r^*)^2]\right] \\ &\qquad \le \mathbb{P}\left[r(\|A\|^{\frac{1}{2}} + (\mathbb{E}\|A\|)^{\frac{1}{2}}) \ge \frac{C}{4}(\|z'\| - r^*)^2\right] \\ &\qquad \le \mathbb{P}\left[(\|A\|^{\frac{1}{2}} + 2) \ge \frac{C\delta^2}{4r}\right], \end{aligned} \tag{11.18}$$

Choosing the grid parameter r such that $r \leq C\delta^2/16$, the right-hand side of (11.18) is bounded by $\mathbb{P}[\|A\| > 4] \leq Ce^{-9N/C}$ This takes care of the second term in (11.13). Let us now treat the first term. The probability that the supremum over all lattice points of some function exceeds some given value is transformed into a summable series of probabilities that at each lattice point the function is greater than this value. More precisely, we have

$$\begin{aligned}\mathbb{P}\left[\sup_{z'\in\mathcal{W}_r\cap\mathcal{O}}\left|\bar{f}_N(z')\right| \geq \frac{C}{4}\left(\|z'\|-r^*\right)^2\right] &\leq \sum_{z'\in\mathcal{W}_r\cap\mathcal{O}}\mathbb{P}\left[\left|\bar{f}_N(z')\right| \geq \frac{C}{4}\left(\|z'\|-r^*\right)^2\right] \\ &\leq \sum_{z'\in\mathcal{W}_r\cap\mathcal{O}} e^{-KC^2\left(\|z'\|-r^*\right)^4 N},\end{aligned} \tag{11.19}$$

by Chebyshev's inequality. Then

$$\begin{aligned}\sum_{z'\in\mathcal{W}_r\cap\mathcal{O}} e^{-KC^2\left(\|z'\|-r^*\right)^4 N} &= r^{-2}\sum_{z'\in\mathcal{W}_r\cap\mathcal{O}} r^2 e^{-KC^2\left(\|z'\|-r^*\right)^4 N} \\ &\leq r^{-2}\int_{\mathbb{R}^2\setminus B_0(r^*+\delta-\sqrt{2}\,r)} e^{-KC^2\left(\|z'\|-r^*\right)^4 N}\,dz \\ &\leq r^{-2}e^{-K\frac{C^2}{16}\delta^4 N}\int_{\mathbb{R}^2\setminus B_0(r^*+\delta/2)} e^{-K\frac{C^2}{16}\left(\|z'\|-r^*\right)^4 N}\,dz \\ &\leq r^{-2}2\pi e^{-K\frac{C^2}{16}\delta^4 N}N^{-\frac{1}{2}}\int_{\delta/2}^{\infty} ze^{-\widetilde{K}z^4}\,dz \\ &\leq K'r^{-2}e^{-K\frac{C^2}{2}\delta^4 N}N^{-\frac{1}{2}},\end{aligned} \tag{11.20}$$

where K' stands for an upper bound for the integral, which is independent of N (assuming $\delta > 2\sqrt{2}r$). Combining this and (11.18), and choosing δ small enough such that $C\delta^2/16 \leq \delta/(2\sqrt{2})$ concludes the proof of Lemma 11.4. $\square$

Therefore, on a set of measure at least $1 - C_1 e^{-C_2 N\delta^4}$, the integral (11.10) can be bounded by

$$\begin{aligned}\int_{\mathcal{O}} e^{-\beta N\mathbb{E}\,\Phi_N(z)}e^{-\beta N(\Phi_N(z)-\mathbb{E}\,\Phi_N(z))}\,dz &\leq \int_{\mathcal{O}} e^{-\beta N\frac{C}{2}(\|z\|-r^*)^2}\,dz \\ &\leq 2\pi\int_{r^*+\delta}^{\infty} re^{-\beta NC(r-r^*)^2}\,dz \\ &\leq 2\pi e^{-N\frac{C}{4}\delta^2}\int_0^{\infty} re^{-\beta N\frac{C}{4}r^2}\,dr \\ &= 2\pi\frac{2}{\beta NC}e^{-\beta N\frac{C}{4}\delta^2}.\end{aligned} \tag{11.21}$$

We now turn to the integral on the "inner" part $\mathcal{I}$. Again, we have to control the term

$$\Phi_N(z) - \mathbb{E}\,\Phi_N(z). \tag{11.22}$$

Since $\mathcal{I}$ is compact, we can do this uniformly by using the following lemma.

Lemma 11.5: *Let $f_N(z) = 1/(\beta N)\sum_{i=1}^N \ln\cosh\beta(\xi_i, z)$ and $\mathcal{D} \subset \mathbb{R}^2$ a bounded set. Then there exist strictly positive constants K_1, K_2, C_1, C_2 such that*

$$\mathbb{P}\left[\sup_{z\in\mathcal{D}} |f_N(z) - \mathbb{E} f_N(z)| > \varepsilon\right] \le K_1 e^{-K_2 N} + C_1\varepsilon^{-2} e^{-C_2\varepsilon^2 N}. \tag{11.23}$$

Proof: The proof is similar (if not simpler) to the proof of Lemma 11.4. Define again $\bar f_N(z) \equiv f_N(z) - \mathbb{E} f_N(z)$. Let $\mathcal{W}_r$ be collection of grid points with spacing r. Then

$$\begin{aligned}\mathbb{P}\left[\sup_{z\in\mathcal{D}} |\bar f_N(z)| > \varepsilon\right] &\le \mathbb{P}\left[\sup_{z'\in\mathcal{W}_r\cap\mathcal{D}}\ \sup_{z\in B_r(z')} |\bar f_N(z) - \bar f_N(z')| > \frac{\varepsilon}{2}\right]\\ &\quad + \mathbb{P}\left[\sup_{z'\in\mathcal{W}_r\cap\mathcal{D}} |\bar f_N(z')| > \frac{\varepsilon}{2}\right].\end{aligned} \tag{11.24}$$

The first term is treated similarly to the second summand in inequality (11.13) in the proof of Lemma 11.2. Thus, with the same definition of the matrix A as in the proof of Lemma 11.4,

$$\begin{aligned}\mathbb{P}\left[\sup_{z'\in\mathcal{W}_r\cap\mathcal{D}}\ \sup_{z\in B_r(z')} |\dots|\right] &\le \mathbb{P}\left[\beta r(\|A\|^{\frac12} + (\mathbb{E}\|A\|)^{\frac12}) > \frac{\varepsilon}{2}\right]\\ &\le \mathbb{P}\left[\beta r(\|A\|^{\frac12} + 10) > \frac{\varepsilon}{2}\right].\end{aligned} \tag{11.25}$$

Choosing the grid parameter $r = \frac{\varepsilon}{40\beta}$, this gives

$$\begin{aligned}\mathbb{P}\left[\beta r(\|A\|^{\frac12} + 10) > \frac{\varepsilon}{2}\right] &\le \mathbb{P}\left[\|A\|^{\frac12} > 10\right]\\ &\le K_1 e^{-K_2 N},\end{aligned} \tag{11.26}$$

again by the standard result concerning random matrices (see [Ge]).

We now turn to the second term in (11.24). Using the "trick" to transform the supremum into a sum, as well as Chebyshev's exponential inequality, we get

$$\begin{aligned}\mathbb{P}\left[\sup_{z'\in\mathcal{W}_r\cap\mathcal{D}} |f_N(z') - \mathbb{E} f_N(z')| > \frac{\varepsilon}{2}\right] &\le \sum_{z'\in\mathcal{W}_r\cap\mathcal{D}} \mathbb{P}\left[|f_N(z') - \mathbb{E} f_N(z')| > \frac{\varepsilon}{2}\right]\\ &\le \sum_{z'\in\mathcal{W}_r\cap\mathcal{D}} C' e^{-CN\frac{\varepsilon^2}{4}}\\ &= |\mathcal{W}_r\cap\mathcal{D}| C' e^{-CN\frac{\varepsilon^2}{4}}.\end{aligned} \tag{11.27}$$

Since $\mathcal{D}$ is bounded, there exists $R > 0$ such that $\mathcal{D}$ is contained in the ball $B_R(0)$. Thus (see [BG2]),

$$|\mathcal{W}_r\cap\mathcal{D}| \le |\mathcal{W}_r\cap B_R(0)| \le \left(\frac{R}{r}\right)^2 \left(\frac{9\pi e}{2}\right)^{\frac12}. \tag{11.28}$$

We finally get from (11.26), (11.27) and (11.28)

$$\mathbb{P}\left[\sup_{z\in\mathcal{D}}|f_N(z)-\mathbb{E}\,f_N(z)|>\varepsilon\right]\leq K_1e^{-K_2N}+C_1\varepsilon^2e^{-C_2N\varepsilon^2}, \tag{11.29}$$

which is the statement of Lemma 11.5. □

Lemma 11.5 implies that

$$\begin{aligned}\int_{\mathcal{I}} e^{-\beta N\Phi_N(z)}\,dz &\leq e^{\varepsilon N}e^{-\beta N\mathbb{E}\,\Phi(r^*)}\int_{\mathcal{I}} e^{-\beta N\mathbb{E}\,\Phi_N(z)}\,dz\\ &\leq e^{\varepsilon\ bN}e^{-\delta^2C\beta N}\pi r^{*2},\end{aligned} \tag{11.30}$$

using the fact that $\mathbb{E}\,\Phi_N(\|z\|)-\mathbb{E}\,\Phi(r^*)$ can be bounded uniformly on $\mathcal{I}$ by its value for $\|z\|=r^*-\delta$.

Finally, the denominator in (11.2) can be bounded from below, using the second order Taylor expansion of $\mathbb{E}\,\Phi_N(\|z\|)$

$$\begin{aligned}\int_{|\,\|z\|-r^*|<\delta} e^{-\beta N\Phi_N(z)}\,dz &\geq e^{-\beta N\mathbb{E}\,\Phi(r^*)}\int_{|\,\|z\|-r^*|<\delta} e^{-NC(\|z\|-r^*)^2-NC'(\|\tilde{z}\|-r^*)^3-N\varepsilon}\,dz\\ &\geq 2\pi\frac{1}{\beta NC}e^{-\varepsilon\beta N}e^{-\beta NC'\delta^3}e^{-\beta N\mathbb{E}\,\Phi(r^*)}\left(1-\delta e^{-\beta NC\delta^2}\right),\end{aligned} \tag{11.31}$$

on a set of measure at least $1-Ke^{-KN}-C\varepsilon^{-2}e^{-CN\varepsilon^2}$ (this error term can be estimated by Lemma 11.5). Collecting (11.21), (11.30) and (11.31), we get that on a set of measure exponentially close to one,

$$\begin{aligned}\frac{\int_{|\|z\|-r^*|\geq\delta} e^{-\beta N\Phi_N(z)}\,dz}{\int_{|\|z\|-r^*|<\delta} e^{-\beta N\Phi_N(z)}\,dz} &\leq e^{\varepsilon\beta N}e^{\beta NC'\delta^3}(2\pi)^{-1}\beta NC\left(1-\delta e^{-\beta NC\delta^2}\right)^{-1}\\ &\quad\times\left\{e^{\varepsilon\beta N}e^{-\beta NC\delta^2}\pi r^{*2}+2\pi e^{-\beta N\frac{C}{4}\delta^2}\frac{2}{\beta NC}\right\}\\ &= Ke^{-\beta N(C\delta^2-2\varepsilon-C'\delta^3)}\beta N\left(1-\delta e^{-NC\delta^2}\right)^{-1}\\ &\quad+Ke^{-\beta N(\frac{C}{4}\delta^2-\varepsilon-C'\delta^3)}N\left(1-\delta e^{-\beta NC\delta^2}\right)^{-1}.\end{aligned} \tag{11.32}$$

Now let us choose $\delta_N=N^{-\frac{1}{10}}$, $\varepsilon_N=N^{-\frac{1}{4}}$. Then (11.32) implies that

$$\begin{aligned}\frac{\int_{|\,\|z\|-r^*|\geq\delta_N} e^{-\beta N\Phi_N(z)}\,dz}{\int_{|\,\|z\|-r^*|<\delta_N} e^{-\beta N\Phi_N(z)}\,dz} &\leq KNe^{-\beta N^{\frac{4}{5}}(C-2N^{-\frac{1}{20}}-C'N^{-\frac{1}{10}})}\\ &\quad+KNe^{-N^{\frac{4}{5}}(\frac{C}{4}-N^{-\frac{1}{20}}-C'N^{-\frac{1}{10}})},\end{aligned} \tag{11.33}$$

on a set which is exponentially close (in N) to 1. This concludes the proof of Lemma 11.1. $\square$

We now turn to the proof of Lemma 11.2 which is a little more delicate than the previous one.

Proof of Lemma 11.2: Let us write $I(B)$ for the integral $\int_B e^{-\beta N \Phi_N(z)}\,dz$. We will prove the concentration behavior by a strategy similar to the one used in Lemma 11.1. Namely we replace the function Φ_N by its expectation $\mathbb{E}\,\Phi_N$ and control the error.

Write the fluctuation term $\Phi_N - \mathbb{E}\,\Phi_N$ as

$$\begin{aligned}\Phi_N(z) - \mathbb{E}\,\Phi_N(z) &= \frac{1}{\beta N}\sum_{i=1}^N \{\ln\cosh\beta(\xi_i, z) - \mathbb{E}\ln\cosh\beta(\xi_i, z)\} \\ &= \frac{1}{\beta N}\sum_{i=1}^N \{\ln\cosh\beta(\xi_i, z) - \ln\cosh\beta(\xi_i, z') \\ &\qquad - \mathbb{E}\ln\cosh\beta(\xi_i, z) + \mathbb{E}\ln\cosh\beta(\xi_i, z')\} \\ &\quad + \frac{1}{\beta N}\sum_{i=1}^N \{\ln\cosh\beta(\xi_i, z') - \mathbb{E}\ln\cosh\beta(\xi_i, z')\}.\end{aligned} \tag{11.34}$$

Let $z' = z'(z) = r^* \frac{z}{\|z\|}$, that is, z' is the radial projection of z onto $S^1(r^*)$. Define the two functions

$$\begin{aligned}h_N(z) \equiv \frac{1}{\sqrt{N}}\sum_{i=1}^N &\{\ln\cosh\beta(\xi_i, z) - \ln\cosh\beta(\xi_i, z') \\ &- \mathbb{E}\ln\cosh\beta(\xi_i, z) + \mathbb{E}\ln\cosh\beta(\xi_i, z')\},\end{aligned} \tag{11.35}$$

with z' defined as above, and

$$g_N(z) \equiv \frac{1}{\sqrt{N}}\sum_{i=1}^N \{\ln\cosh\beta(\xi_i, z) - \mathbb{E}\ln\cosh\beta(\xi_i, z)\}. \tag{11.36}$$

Then the fluctuation term takes the form

$$N(\Phi_N(z) - \mathbb{E}\,\Phi_N(z)) = \frac{\sqrt{N}}{\beta}(h_N(z) - g_N(z')). \tag{11.37}$$

It is the term g_N that determines the concentration behavior of the measure. To see this we first bound the term h_N uniformly on the "annulus of concentration" $\mathcal{A}_N \cup \mathcal{A}'_N$. We have the following result.

Lemma 11.6: *Let $\{\xi_i\}_{i\in\mathbb{N}}$ be i.i.d. Gaussian variables with mean zero and variance one. Let h_N be as in (11.35), and $\mathcal{A}_N, \mathcal{A}'_N$ as in (11.5). Then for any $\varepsilon > 0$,*

$$\mathbb{P}\left[\sup_{z\in\mathcal{A}_N\cup\mathcal{A}'_N} |h_N(z)| \geq \varepsilon\right] \leq K N^2 e^{-N^{\frac{1}{10}}(\varepsilon - K N^{-1/10})}. \tag{11.38}$$

Proof: Let us write

$$f_i(z) \equiv \ln\cosh\beta(\xi_i, z), \tag{11.39}$$

and

$$\bar{f}_i \equiv \ln\cosh\beta(\xi_i, z) - \mathbb{E}\ln\cosh\beta(\xi_i, z). \tag{11.40}$$

We also keep the notation $z' = z'(z)$ defined above. Introduce a polar grid $\mathcal{W}_N$ in $\mathbb{R}^2$, i.e. a discrete set of points $x_{i,j}$ whose polar coordinates are given by $(\rho_i, \alpha_j) \in \mathbb{R}^+ \times [0, 2\pi)$, such that $\Delta_N \alpha \equiv |\alpha_i - \alpha_j| = CN^{-\frac{1}{2}}$ and $\Delta_N \rho \equiv |\rho_i - \rho_j| = CN^{-\frac{1}{2}}$, for some appropriate constant C. Note that for any point z in a bounded domain $\mathcal{U} \subset \mathbb{R}^2$, the distance to the closest grid point is less than $KN^{-\frac{1}{2}}$.

For any $z \in \mathbb{R}^2$, define $x = x(z) \in \mathcal{W}_N$ to be the grid point closest to z, and $y = y(z) \in \mathcal{W}_N$ the grid point closest to $z' = z'(z)$. One can easily convince oneself, that $x' = y'$, i.e. the two points x and y lie on the same ray starting at the origin. Then we can decompose the function $h_N(z)$ as

$$\begin{aligned} h_N(z) &= \frac{1}{\sqrt{N}}\sum_{i=1}^N \{\bar{f}_i(z) - \bar{f}_i(z')\} \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^N \{\bar{f}_i(z) - \bar{f}_i(x)\} + \frac{1}{\sqrt{N}}\sum_{i=1}^N \{\bar{f}_i(x) - \bar{f}_i(y)\} \\ &\quad + \frac{1}{\sqrt{N}}\sum_{i=1}^N \{\bar{f}_i(y) - \bar{f}_i(z')\}. \end{aligned} \tag{11.41}$$

Denote by $I_1(z,x)$, $I_2(x,y)$, $I_3(y,z')$ respectively the first, second and third sum on the right-hand side of (11.41). We can then write (let $\mathcal{A}_N = \mathcal{A}_N \cup \mathcal{A}'_N$, the “annulus of concentration”)

$$\begin{aligned} \mathbb{P}\left[\sup_{z\in\mathcal{A}_N} |h_N(z)| \geq \varepsilon\right] &= \mathbb{P}\left[\sup_{z\in\mathcal{A}_N} |I_1(z,x) + I_2(x,y) + I_3(y,z')| \geq \varepsilon\right] \\ &\leq \mathbb{P}\left[\sup_{x\in\mathcal{W}_N\cap\mathcal{A}_N}\ \sup_{z\in B_{KN^{-\frac{1}{2}}}(x)} |I_1(z,x)| \geq \frac{\varepsilon}{3}\right] \\ &\quad + \mathbb{P}\left[\sup_{x\in\mathcal{W}_N\cap\mathcal{A}_N}\ \sup_{\substack{y\in\mathcal{W}_N\cap\mathcal{A}_N \\ y'=x'}} |I_2(x,y)| \geq \frac{\varepsilon}{3}\right] \\ &\quad + \mathbb{P}\left[\sup_{y\in\mathcal{W}_N\cap\mathcal{A}_N}\ \sup_{z'\in B_{KN^{-\frac{1}{2}}}(y)} |I_3(y,z')| \geq \frac{\varepsilon}{3}\right]. \end{aligned} \tag{11.42}$$

The first and the third term (they are equal) can be uniformly bounded by an estimate analogous to the proof of Lemma 11.5. In fact, for any u, v, we have

$$\left|\frac{1}{\sqrt{N}}\sum_{i=1}^N \{\bar{f}_i(u) - \bar{f}_i(v)\}\right| \leq \sqrt{N}\beta(\|A\|^{\frac{1}{2}} + (\mathbb{E}\|A\|)^{\frac{1}{2}})\|u - v\|_2. \tag{11.43}$$

If $\|u - v\|_2 \leq 4\varepsilon' N^{-\frac{1}{2}}/\beta$, we have the following exponential bound.

$$\begin{aligned} \mathbb{P}\left[|\frac{1}{\sqrt{N}}\sum_{i=1}^N \{\bar{f}_i(u) - \bar{f}_i(v)\}| \geq \varepsilon'\right] &\leq \mathbb{P}\left[\|A\|^{\frac{1}{2}} + (\mathbb{E}\|A\|)^{\frac{1}{2}} \geq \frac{\varepsilon' N^{-1/2}}{\beta\|u - v\|_2}\right] \\ &\leq \mathbb{P}\left[\|A\| \geq 4\right] \leq Ce^{-CN}. \end{aligned} \tag{11.44}$$

Thus we get for the first term in (11.42),

$$\mathbb{P}\left[\sup_{x\in\mathcal{W}_N\cap\mathcal{A}_N}\ \sup_{z\in B_{KN^{-\frac{1}{2}}}(x)}|I_1(z,x)|\geq\frac{\varepsilon}{3}\right]\leq\sum_{x\in\mathcal{W}_N\cap\mathcal{A}_N}\mathbb{P}\left[\sup_{z\in B_{KN^{-\frac{1}{2}}}(x)}|I_1(z,x)|\geq\frac{\varepsilon}{3}\right]$$
$$\leq\sum_{x\in\mathcal{W}_N\cap\mathcal{A}_N}\mathbb{P}\left[\|A\|\geq 4\right]\leq CN^{\frac{1}{10}}N^{-1}e^{-CN}, \tag{11.45}$$

since we know that $\|x-z\|=KN^{-\frac{1}{2}}$ by the remark preceding (11.41), and the number of grid points in $\mathcal{A}_N$ is bounded by $N\delta_N^{-1}$ times some constant. The same estimate is valid for the term containing I_3 (since they are equal).

Let us now consider the term containing I_2. We know that $\|x-y\|\leq 2\delta_N$, since those two points are supposed to lie on the same "ray". Again, we can turn the supremum into a sum,

$$\mathbb{P}\left[\sup_{x\in\mathcal{W}_N\cap\mathcal{A}_N}\ \sup_{\substack{y\in\mathcal{W}_N\cap\mathcal{A}_N\\ y'=x'}}|I_3(x,y)|\geq\frac{\varepsilon}{3}\right]\leq\sum_{x,y}\mathbb{P}\left[|I_3(y,z')|\geq\frac{\varepsilon}{3}\right], \tag{11.46}$$

where x,y on the right-hand side satisfy the same conditions as on the left-hand side. By Chebyshev's inequality, we get that for any u, v

$$\mathbb{P}\left[\sum_{i=1}^{N}\{\bar f_i(u)-\bar f_i(v)\}\geq\sqrt{N}\varepsilon'\right]\leq\inf_{s>0}e^{-s\varepsilon'\sqrt{N}}\mathbb{E}\left[e^{s\sum_{i=1}^{N}\{\bar f_i(u)-\bar f_i(v)\}}\right]$$
$$=\inf_{s>0}e^{-s\varepsilon'\sqrt{N}}\prod_{i=1}^{N}\mathbb{E}e^{s\{\bar f_i(u)-\bar f_i(v)\}}. \tag{11.47}$$

Now we use the series expansion of the exponential function, the fact that the exponent in the right-hand side of (11.47) is a centered random variable, and some obvious inequalities for each term of the expansion, to get

$$\mathbb{E}e^{s\{\bar f_i(u)-\bar f_i(v)\}}\leq\left\{1+\frac{s^2}{2}\mathbb{E}\left[(\bar f_i(u)-\bar f_i(v))^2e^{s|\bar f_i(u)-\bar f_i(v)|}\right]\right\}. \tag{11.48}$$

To evaluate the expectation term, we use the inequality

$$|f_i(u)-f_i(v)|\leq\beta|(\xi_i,u-v)|. \tag{11.49}$$

Then the expectation term in (11.48) is bounded by

$$\mathbb{E}\left[(\bar f_i(u)-\bar f_i(v))^2e^{s|\bar f_i(u)-\bar f_i(v)|}\right]\leq\left(\mathbb{E}\left[(\bar f_i(u)-\bar f_i(v))^4\right]\right)^{\frac{1}{2}}\left(\mathbb{E}e^{2s|\bar f_i(u)-\bar f_i(v)|}\right)^{\frac{1}{2}}$$
$$\leq 4\left(\mathbb{E}\left[(f_i(u)-f_i(v))^4\right]\right)^{\frac{1}{2}}\left(\mathbb{E}e^{2s|f_i(u)-f_i(v)|}\right)^{\frac{1}{2}}$$
$$\times e^{s\mathbb{E}|f_i(u)-f_i(v)|}, \tag{11.50}$$

where the first inequality follows by Cauchy-Schwarz, and the second one is a consequence of the inequality $(a+b)^2 \leq 2(a^2+b^2)$ (applied twice to the first factor), respectively the trivial fact that $|a-b| \leq |a|+|b|$. All quantities in (11.50) can be bounded easily using (11.49). One gets (by calculating explicit Gaussian integrals)

$$\mathbb{E}\left[(f_i(u)-f_i(v))^4\right] = 3\|u-v\|_2^4, \tag{11.51}$$

$$\mathbb{E}\, e^{2s|f_i(u)-f_i(v)|} \leq 2e^{2s^2\|u-v\|_2^2}, \tag{11.52}$$

$$e^{s\mathbb{E}\,|f_i(u)-f_i(v)|} \leq e^{s\sqrt{2/\pi}\|u-v\|_2}. \tag{11.53}$$

Inserting (11.51)–(11.53) into (11.50), gives

$$\frac{s^2}{2}\mathbb{E}\left[(\bar{f}_i(u)-\bar{f}_i(v))^2 e^{s|\bar{f}_i(u)-\bar{f}_i(v)|}\right] \leq 2\sqrt{6}s^2\|u-v\|_2^2 e^{2s^2\|u-v\|_2^2+s\sqrt{2/\pi}\|u-v\|_2}. \tag{11.54}$$

We use the above bound (11.54) in (11.48). Together with the inequality $1+x \leq e^x$, and the fact that $\|x-y\|_2 \leq \delta_N = KN^{-1/10}$, we thus get the following estimate for the right-hand side of (11.47)

$$\mathbb{P}\left[\sum_{i=1}^N \{\bar{f}_i(u)-\bar{f}_i(v)\} \geq \sqrt{N}\varepsilon'\right] \leq \inf_{s>0} e^{-s\varepsilon'\sqrt{N}+Ks^2N^{4/5}e^{2s^2N^{-1/5}+\sqrt{2/\pi}N^{-1/10}}}. \tag{11.55}$$

Choosing $s = N^{-2/5}$, this gives

$$\mathbb{P}\left[\sum_{i=1}^N \{\bar{f}_i(u)-\bar{f}_i(v)\} \geq \sqrt{N}\varepsilon'\right] \leq \widetilde{K}e^{-N^{\frac{1}{10}}(\varepsilon'-KN^{-1/10})}. \tag{11.56}$$

The same bound applies to

$$\mathbb{P}\left[\sum_{i=1}^N \{\bar{f}_i(u)-\bar{f}_i(v)\} \leq -\sqrt{N}\varepsilon'\right]. \tag{11.57}$$

Inserting (11.56) and (11.57) into the left-hand side of (11.46) gives

$$\mathbb{P}\left[\sup_{x\in\mathcal{W}_N\cap\mathcal{A}_N}\ \sup_{\substack{y\in\mathcal{W}_N\cap\mathcal{A}_N\\ y'=x'}} |I_2(x,y)| \geq \varepsilon'\right] \leq KN^{\frac{1}{2}}N^{\frac{1}{10}}e^{-N^{\frac{1}{10}}(\varepsilon'-K'N^{-1/10})}, \tag{11.58}$$

since the number of terms in the sum does not exceed a constant times $N^{\frac{1}{2}}$ (the number of allowed x) times $N^{\frac{1}{10}}$ (the number of allowed y). Using (11.45) and (11.58), (11.42) gives

$$\mathbb{P}\left[\sup_{z\in\mathcal{A}_N} |h_N(z)| \geq \varepsilon\right] \leq KN^2e^{-K'N^{\frac{1}{10}}\varepsilon}. \tag{11.59}$$

This concludes the proof of Lemma 11.6. □

Note that we can choose ε as a function of N, and still get an exponential bound. For example, choose $\varepsilon = \varepsilon_N \equiv (\ln N)^2 N^{-\frac{1}{20}}$. Lemma 11.6 then reads

Lemma 11.7: *Let $\{\xi_i\}_{i\in\mathbb{N}}$ be i.i.d. Gaussian variables with mean zero and variance one. Let h_N be as in (11.35), and $\mathcal{A}_N, \mathcal{A}'_N$ as in (11.5). Then,*

$$\mathbb{P}\left[\sup_{z\in\mathcal{A}_N\cup\mathcal{A}'_N} |h_N(z)| \geq N^{-\frac{1}{20}}(\ln N)^2\right] \leq KN^2 e^{-N^{\frac{1}{20}}((\ln N)^2 - K'N^{-\frac{1}{20}})}. \tag{11.60}$$

Furthermore,

$$\mathbb{P}\left[\sup_{z\in\mathcal{A}_N\cup\mathcal{A}'_N} |h_N(z)| \geq N^{-\frac{1}{20}}(\ln N)^2, \text{i.o. in } N\right] = 0. \tag{11.61}$$

Proof: The first statement is a straightforward consequence of Lemma 11.6. Equation (11.61) then follows by the first Borel-Cantelli Lemma. □

Let us now estimate the integral $I(\mathcal{A}'_N)$. Using the bound for h_N from Lemma 11.6, we get explicitly

$$\begin{aligned}
\int_{\mathcal{A}'_N} e^{-\beta N\Phi_N(z)}\,dz &= \int_{\mathcal{A}'_N} e^{-\beta N\mathbb{E}\,\Phi_N(z)} e^{-\sqrt{N}h_N(z)} e^{-\sqrt{N}g_N(z'(z))}\,dz \\
&\leq \int_{|r-r^*|<\delta_N} re^{-\beta N\mathbb{E}\,\Phi_N(r^*)} e^{\sqrt{N}\varepsilon}\,dr \\
&\quad\times \int_{g_N(\vartheta)-\min g_N > a_N} e^{-\sqrt{N}g_N(\vartheta)}\,d\vartheta \\
&= 2e^{-\beta N\mathbb{E}\,\Phi_N(r^*)} e^{\sqrt{N}\varepsilon} \int_{|r-r^*|<\delta_N} r\,dr \\
&\quad\times \int_{g_N(\vartheta)-\min g_N > a_N} e^{-\sqrt{N}g_N(\vartheta)}\,d\vartheta \\
&\leq 4e^{-\beta N\mathbb{E}\,\Phi_N(r^*)} e^{\sqrt{N}\varepsilon} r^*\delta_N \\
&\quad\times 2\pi e^{-\sqrt{N}a_N} e^{-\sqrt{N}\min g_N}.
\end{aligned} \tag{11.62}$$

The last line follows from the crude estimate

$$\int_{r^*-\delta_N}^{r^*+\delta_N} r\,dr = 2\delta_N r^*. \tag{11.63}$$

Thus,

$$\int_{\mathcal{A}'_N} e^{-\beta N\Phi_N(z)}\,dz \leq Ke^{-\beta N\mathbb{E}\,\Phi_N(r^*)} e^{\sqrt{N}\varepsilon}\delta_N r^* e^{-\sqrt{N}a_N}. \tag{11.64}$$

We now turn to the integral $I(\mathcal{A}_N)$. Using standard estimates for Gaussian integrals, a quadratic upper bound of g_N about its minima, and the fact that $\mathbb{E}\,\Phi(\|z\|)$ can be bounded from above by a quadratic function in some neighborhood containing r^*, we get

$$\begin{aligned}\int_{\mathcal{A}_N} e^{-\beta N \Phi_N(z)}\,dz \geq\; & e^{-\beta N \mathbb{E}\,\Phi_N(r^*)} e^{-\sqrt{N}\varepsilon} \int_{|r-r^*|<\delta_N} r e^{-\beta N C'(r-r^*)^2}\,dr \\ & \times \int_{g_N(\vartheta)-\min g_N \leq a_N} e^{-\sqrt{N} g_N(\vartheta)}\,d\vartheta \\ \geq\; & K e^{-\beta N \mathbb{E}\,\Phi_N(r^*)} e^{-\sqrt{N}\varepsilon} (r^*-\delta_N) \left(\frac{\pi}{NC'}\right)^{\frac{1}{2}} \\ & (1-e^{-NC'\delta_N}) \left(\frac{\pi}{K\sqrt{N}}\right)^{\frac{1}{2}} (1-e^{-\sqrt{N}K' a_N}).\end{aligned} \tag{11.65}$$

We get finally for the ratio $I(\mathcal{A}'_N)/I(\mathcal{A}_N)$

$$\frac{I(\mathcal{A}'_N)}{I(\mathcal{A}_N)} \leq K \frac{r^*}{r^*-\delta_N} N^{3/4} e^{-\sqrt{N}(a_N-2\varepsilon)}. \tag{11.66}$$

Lemma 11.7 allows us to choose $\varepsilon = \varepsilon(N) = N^{-\frac{1}{20}}(\ln N)^2$. Inserting this choice, together with $a_N = N^{-1/25}$, into (11.66), gives

$$\frac{I(\mathcal{A}'_N)}{I(\mathcal{A}_N)} \leq K N^{3/4} e^{-N^{23/50}(1-K'(\ln N)^2 N^{-1/100})}. \tag{11.67}$$

This statement is true for all $\omega \in \Omega$, for which Lemma 11.6 respectively 11.7 holds, that is on a set of $\mathbb{P}$-measure at least $KN^2 e^{-N^{\frac{1}{20}}((\ln N)^2 - K' N^{-\frac{1}{20}})}$. This proves Lemma 11.2. $\square$

Let us now turn to the proof of Theorem 10.1. We again state first a result about the concentration of the induced measure $\widetilde{\mathcal{Q}}^h_{N,\beta}$.

Proposition 11.8: *Let $\{\xi_i^\mu\}_{i\in\mathbb{N},\mu=1,2}$ be i.i.d. standard Gaussian variables, and define*

$$\Phi^h_{N,\beta}(z) \equiv \frac{1}{2}\|z\|_2^2 - \frac{1}{\beta N}\sum_{i=1}^N \ln\cosh\beta(\xi_i, z+h). \tag{11.68}$$

Let furthermore $\delta_N = N^{-1/5}$. Then there exist strictly positive constants K, K', l such that

$$\mathbb{P}\left\{ \frac{\int_{\|z-\tilde{r}^h\|\geq\delta_N} e^{-\beta N \Phi^h_{N,\beta}(z)}\,dz}{\int_{\|z-\tilde{r}^h\|<\delta_N} e^{-\beta N \Phi^h_{N,\beta}(z)}\,dz} \geq K e^{-K' N^l},\ \text{i.o. in } N \right\} = 0, \tag{11.69}$$

where $\tilde{r}^h$ is the unique minimum of the function

$$\mathbb{E}\,\Phi^h_{N,\beta}(z) = \frac{1}{2}\|z\|_2^2 - \frac{1}{\beta}\mathbb{E}\ln\cosh\beta(\xi_1, z+h). \tag{11.70}$$

Proof: Let us decompose $\Phi^h_{N,\beta}$ in the usual way

$$\Phi^h_{N,\beta}(z) = \mathbb{E}\,\Phi^h_{N,\beta}(z) + \Phi^h_{N,\beta}(z) - \mathbb{E}\,\Phi^h_{N,\beta}(z). \tag{11.71}$$

We first treat the denominator appearing in (11.69). $\mathbb{E}\,\Phi^h_{N,\beta}$ can be bounded from below by some quadratic function $C\|z-\tilde{r}^h\|_2^2$ on the set $\|z-\tilde{r}^h\| \geq \delta_N > 0$. The fluctuation term can be controlled by the following analogue of Lemma 11.4.

Lemma 11.9: *Let $f_N = \frac{1}{\beta N}\sum_{i=1}^N \ln\cosh\beta(\xi_i, z+h)$. Then for δ small enough, such that $C\delta^2/80 < \delta/2$, there exist strictly positive constants C_1, C_2, K_1, K_2 such that*

$$\begin{aligned} p_N \equiv \mathbb{P}\Big[\sup_{z:\|z-\tilde{r}^h\|_2\geq\delta} |f_N(z) - \mathbb{E}\,f_N(z)| \geq \frac{C}{2}\|z-\tilde{r}^h\|_2^2\Big] \\ \leq K_1 e^{-K_2 N} + C_1 N^{\frac{1}{2}}\delta^{-2}e^{-C_2 N}. \end{aligned} \tag{11.72}$$

Proof: The proof is completely analogous to the proof of Lemma 11.4, and is left to the reader. □

Therefore, with probability greater than $1-p_N$, the quantity $\sup(\Phi^h_{N,\beta} - \mathbb{E}\Phi^h_{N,\beta}(z)$ does not exceed one half of the lower bound of the deterministic part, which implies that

$$\begin{aligned} \int\limits_{\|z-\tilde{r}^h\|\geq\delta_N} e^{-\beta N\Phi^h_{N,\beta}(z)}\,dz \leq e^{-\beta N\mathbb{E}\,\Phi^h_{N,\beta}(\tilde{r}^h)} \int\limits_{\|z-\tilde{r}^h\|\geq\delta_N} e^{-\beta N\frac{C}{2}\|z-\tilde{r}^h\|_2^2}\,dz \\ \leq e^{-\beta N\mathbb{E}\,\Phi^h_{N,\beta}(\tilde{r}^h)} e^{-\beta N\frac{C}{4}\delta_N^2 K}. \end{aligned} \tag{11.73}$$

We now turn to the denominator in (11.69). The probability that the fluctuation term exceeds an $\varepsilon > 0$ is bounded by Lemma 11.5:

$$q_N \equiv \mathbb{P}\Big[\sup_{\|z-\tilde{r}^h\|<\delta_N} |f_N(z) - \mathbb{E}\,f_N(z)| \geq \varepsilon\Big] \leq K_1 e^{-K_2 N} + C_1\varepsilon^{-2}e^{-C_2\varepsilon^2 N}. \tag{11.74}$$

Using the Taylor expansion of $\mathbb{E}\,\Phi^h_{N,\beta}(z)$ about $\tilde{r}^h$ up to order 2, with an error term of order 3, we get that with probability higher than $1-q_N$,

$$\begin{aligned} \int\limits_{\|z-\tilde{r}^h\|<\delta_N} e^{-\beta N\Phi^h_{N,\beta}(z)}\,dz \geq e^{-\beta N(\Phi^h_{N,\beta}(\tilde{r}^h + C''\delta_N^3+\varepsilon))} \int\limits_{\|z-\tilde{r}^h\|<\delta_N} e^{-\beta NC'\|z-\tilde{r}^h\|_2^2}\,dz \\ \geq e^{-\beta N(\Phi^h_{N,\beta}(\tilde{r}^h + C''\delta_N^3+\varepsilon))} KN^{-\frac{1}{2}}(1 - e^{-\beta N\frac{C'}{2}\delta_N^2}). \end{aligned} \tag{11.75}$$

Combining (11.73) and (11.75) gives

$$\frac{\int\limits_{\|z-\tilde{r}^h\|\geq\delta_N} e^{-\beta N\Phi^h_{N,\beta}(z)}\,dz}{\int\limits_{\|z-\tilde{r}^h\|<\delta_N} e^{-\beta N\Phi^h_{N,\beta}(z)}\,dz} \leq Ke^{-\beta N(\frac{C}{2}\delta_N^2 - \varepsilon - C''\delta_N^3)} \tag{11.76}$$

with probability greater than $1-(q_N+p_N)$. Choosing $\delta_N = N^{-1/5}$, $\varepsilon = N^{-1/5}$, implies that $\sum_N(p_N+q_N) < \infty$. The statement of Proposition 11.8 then follows by Borel-Cantelli. □

Theorem 10.1 is now obvious:

Proof of Theorem 10.1: Let f be a bounded continuous function. Then

$$\begin{aligned} \mathcal{Q}^h_{N,\beta}(f) = f(\tilde{r}^h \mathcal{Q}^h_{N,\beta}(\mathbb{1}_{\{\|z-\tilde{r}^h\|\leq\delta_N\}}) + \mathcal{Q}^h_{N,\beta}((f(\tilde{r}^h - f)\mathbb{1}_{\{\|z-\tilde{r}^h\|\leq\delta_N\}}) \\ + \mathcal{Q}^h_{N,\beta}(f\mathbb{1}_{\{\|z-\tilde{r}^h\|>\delta_N\}}). \end{aligned} \tag{11.77}$$

Taking the limit $N\uparrow\infty$, we can replace $\mathcal{Q}^h_{N,\beta}$ by $\widetilde{\mathcal{Q}}^h_{N,\beta}$ and use Proposition 11.8. Since f is bounded, the third term on the right-hand side of (11.77) converges to zero, and since it is continuous, the second term also vanishes too. These statements are true $\mathbb{P}$-a.s. Finally we let $b=\|h\|_2\to 0$. Again by continuity of f, $f(\tilde{r}^h)\to f(r^*(\cos\vartheta,\sin\vartheta))$. This proves the Theorem. □

12 Uniqueness of Extrema of Certain Gaussian Processes

In the previous chapter we have seen that the measures $\widetilde{\mathcal{Q}}_{N,\beta}$ concentrate on a circle of radius r^* at the places where the random function $g_N(\vartheta)$ takes its minimum. Here we will show that these sets degenerate to a single point, a.s. in the limit $N \uparrow \infty$. To do so we first prove a uniqueness theorem for the absolute minimum of a certain class of strongly correlated Gaussian processes. Then we show convergence in distribution of $g_N(\vartheta)$ to such a process and finally we show that this implies also the desired convergence in distribution of our measures. We begin with the following general result.

Proposition 12.1: *Suppose $\chi(t)$ is a real stationary Gaussian process which is periodic with period T. Suppose furthermore that its covariance function $r(s,t) = r(s-t)$ is even, $\in C^\infty[0,T]$, and $r(\tau)$ is less than $r(0)$ for all $\tau \in (0,T)$. Then there exists an equivalent process $\eta(t)$ having almost surely infinitely differentiable sample paths. Moreover, the probability that there exist two or more maxima with equal height in $[0,T)$ is zero.*

Proof: Without restricting the generality, we can assume that $\mathbb{E}[\chi(t)] = 0$ and $\kappa = \mathbb{E}[\chi(t)^2] = 1$. By its continuity properties, $r(\tau)$ can be expanded about the origin as

$$r(\tau) = 1 - \frac{\lambda_2}{2!}\tau^2 + O(\tau^4). \tag{12.1}$$

The first assertion then follows from the following result due to Cramér and Leadbetter (see [CL], Chapter 9.2).

Lemma 12.2: *Suppose that for some $a > 3$,*

$$r(\tau) = 1 - \frac{\lambda_2}{2}\tau^2 + O\left(\frac{\tau^2}{|\ln|\tau||^a}\right), \tag{12.2}$$

where λ_2 is a constant. Then there exists a process $\eta(t)$ equivalent to $\chi(t)$ and possessing, with probability one, a continuous derivative $\eta'(t)$.

Proof: See Cramér/Leadbetter [CL]. □

It is easily checked that by (12.1), $r(\tau)$ satisfies the condition (12.2) in Lemma 12.2, which proves the statements about continuity and existence of a continuous derivative.

Consider now the process $\chi'(t)$. Its covariance function $\tilde{r}(\tau)$ is given by $\tilde{r}(\tau) = -r''(\tau)$ (see for example Leadbetter et al. [LLR], p. 161, Chapter 7.6). Then it can be expanded about the origin as

$$\tilde{r}(\tau) = \lambda_2 - \frac{\lambda_4}{2}\tau^2 + O(\tau^4). \tag{12.3}$$

Then $\tilde{r}(\tau)$ also verifies condition (12.2) in Lemma 12.2. Repeating this argument implies, together with the Borel-Cantelli Lemma, that there exists an equivalent process $\eta(t)$ having, with probability one, infinitely differentiable sample paths.

From now on, we assume that $\chi(t)$ itself has the above continuity properties. We want to find the probability that there are not two maxima with equal height in $[0,T)$, i.e.

$$\mathbb{P}\left[\exists s,t \in T \times T : |s-t| \neq kT, |\chi(t) - \chi(s)| = 0, |\chi'(t)| = |\chi'(s)| = 0\right] = 0. \tag{12.4}$$

We first show that for any $\vartheta > 0$,

$$\mathbb{P}\left[\exists s,t \in T\times T: \left|kT - |s-t|\right| \geq \vartheta, |\chi(t)-\chi(s)| = 0, |\chi'(t)| = |\chi'(s)| = 0\right] = 0 \tag{12.5}$$

Let us choose a collection of grid points $t_i \in T$, separated by some distance $\varepsilon > 0$. By the continuity properties, χ and χ' are Lipschitz-continuous with a.s.-finite constants C_0, C_1. Consider the set $\tilde{\Omega}_C \subset \Omega$ such that C_0 and C_1 are bounded by some number $C > 0$. Then, by Lipschitz-continuity, $\chi'(t) = 0$, $t \in [t_i, t_{i+1})$ implies that (for some $x \in [t_i, t]$)

$$|\chi'(t_i)| \leq C\varepsilon. \tag{12.6}$$

Similarly, $|\chi(t) - \chi(s)| = 0$ implies

$$|\chi(t_i) - \chi(t_j)| \leq 2C\varepsilon \tag{12.7}$$

where $t - t_i < \varepsilon$, $s - t_j < \varepsilon$. Then we can estimate the probability of the event in (12.5) (on $\tilde{\Omega}$) by

$$\begin{aligned}\mathbb{P}&\left[\exists s,t \in T\times T: \left|kT - |s-t|\right| \geq \vartheta, |\chi(t)-\chi(s)| = 0, |\chi'(t)| = |\chi'(s)| = 0\right] \\ &\leq \mathbb{P}\Big[\exists t_i, t_j : \left|kT - |s-t|\right| \geq \vartheta, |\chi(t_i)-\chi(t_j)| \leq 2C\varepsilon, |\chi'(t_i)| \leq C\varepsilon, \\ &\qquad |\chi'(t_j)| \leq C\varepsilon\Big].\end{aligned} \tag{12.8}$$

Let us denote the event appearing on the left-hand side of (12.8) by $\mathcal{A}_\vartheta$, and the event appearing on the right-hand side by $\mathcal{B}_{\vartheta,\varepsilon}$. The probability $\mathbb{P}[\mathcal{B}_{\vartheta,\varepsilon}]$ can be estimated by the standard bound

$$\mathbb{P}\left[\mathcal{B}_{\vartheta,\varepsilon}\right] \leq \sum_{|kT - |t_i - t_j|| \geq \vartheta} \mathbb{P}\left[|\chi(t_i)-\chi(t_j)| \leq 2C\varepsilon, |\chi'(t_i)| \leq C\varepsilon, |\chi'(t_j)| \leq C\varepsilon\right]. \tag{12.9}$$

Now, for any fixed i, j,

$$(\chi(t_i) - \chi(t_j), \chi'(t_i), \chi'(t_j)) \tag{12.10}$$

is a Gaussian vector, and due to the condition on $|t_i - t_j|$ and the assumption concerning $r(\tau)$, its distribution is non-degenerate. Therefore, each term in the sum on the right-hand side of (12.9) can be bounded by

$$\mathbb{P}\left[|\chi(t_i)-\chi(t_j)| \leq 2C\varepsilon, |\chi'(t_i)| \leq C\varepsilon, |\chi'(t_j)| \leq C\varepsilon\right] \leq K\varepsilon^3 C^3 (2\pi\sigma_{i,j})^{-1}, \tag{12.11}$$

where $\sigma_{i,j}$ is the determinant of the non-degenerate covariance matrix of the random vector (12.10). Since the t_i, t_j are chosen in a compact set, this quantity can be bounded uniformly in i, j. We thus get

$$\mathbb{P}\left[|\chi(t_i)-\chi(t_j)| \leq 2C\varepsilon, |\chi'(t_i)| \leq C\varepsilon, |\chi'(t_j)| \leq C\varepsilon\right] \leq K(\vartheta)\varepsilon^3 C^3. \tag{12.12}$$

Finally, the number of allowed pairs (i, j) in the sum in equation (12.9) does not exceed $T^2\varepsilon^{-2}$, which implies that

$$\begin{aligned}\mathbb{P}[\mathcal{A}_\vartheta] \leq & \mathbb{P}[\mathcal{B}_{\vartheta,\varepsilon}] + \mathbb{P}\left[\tilde{\Omega}_C^c\right] \\ \leq & K(\vartheta)T^2\varepsilon^{-2}\varepsilon^3 + \mathbb{P}\left[\tilde{\Omega}_C^c\right],\end{aligned} \tag{12.13}$$

keeping track of the set $\tilde{\Omega}_C^c$ on which the above estimates are not valid. Now choose $C = C(\varepsilon) = o(\varepsilon^{-1/3})$, and observe that due to the continuity properties

$$\begin{aligned}\lim_{\varepsilon\to 0} \mathbb{P}\left[\tilde{\Omega}_{C(\varepsilon)}^c\right] = & \mathbb{P}\left[\bigcap_{n\in\mathbb{N}} \{C \geq n\}\right] \\ = & 0.\end{aligned} \tag{12.14}$$

Finally, letting ε tend to zero in (12.13) gives that the probability (12.5) is zero. This holds for any $\vartheta > 0$ and thus shows that local maxima are separated with probability one. In particular, there are no constant pieces and no accumulation points of maxima. This concludes its proof. □

Corollary 12.3: *Suppose $\chi(t)$ satisfies the conditions in Proposition 12.1. Then $\chi(t)$ has a.s. only one global maximum in any interval $[s, s+t]$, $t < T$.*

To see that Proposition 12.1 is relevant for our problem, we will next show that the process $g_N(\vartheta)$ converges to a process of the type covered by this proposition. In fact we have

Proposition 12.4: *Let $g : \mathbb{R} \to \mathbb{R}^+$, $g \in C^\infty$ be an aperiodic even function. Suppose also that $\chi_i(\vartheta)$, $\vartheta \in [0, 2\pi]$ is the stochastic process given by*

$$\chi_i(\vartheta) = g\left(r\zeta_i \cos(\vartheta - \phi_i)\right), \tag{12.15}$$

where r is a positive constant, $\{\zeta_i\}_{i\in\mathbb{N}}$, $\{\phi_i\}_{i\in\mathbb{N}}$ are two mutually independent families of i.i.d. random variables, distributed as cxe^{-x^2} (ζ_i), and uniformly (ϕ_i). Then the process η_N given by

$$\eta_N(\vartheta) \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \{\chi_i(\vartheta) - \mathbb{E}\chi_i(\vartheta)\} \tag{12.16}$$

converges in distribution to a strictly stationary Gaussian process $\eta(\vartheta)$ having a.s. continuously differentiable sample paths. Furthermore, $\eta(\vartheta)$ has a.s. only one global maximum on any interval $[s, s+t]$, $t < \pi$.

Remark: We will use this proposition of course with $g(\cdot) = \ln\cosh(\beta\cdot)$. Then the proposition implies that the process $g_N(\vartheta) - \mathbb{E}g_N(\vartheta)$ converges to a Gaussian process with the above properties.

Proof: As $\xi_i(\vartheta)$ are i.i.d. stationary processes on the circle which are infinitely differentiable, the convergence of the process to a stationary Gaussian process on the circle is a simple application of the central limit theorem in Banach spaces (see e.g. [LT]). A computation shows that the covariance of the limiting process is given by

$$\begin{aligned} f(s,t) &= \mathbb{E}[(\chi_1(s) - \mathbb{E}\chi_1(s))(\chi_1(t) - \mathbb{E}\chi_1(t))] \\ &= \mathbb{E}[g(r\zeta_1\cos(\varphi_1))\, g(r\zeta_1\cos(t-s-\varphi_1))] - (\mathbb{E}[g(r\zeta_1\cos(\varphi_1))])^2 \end{aligned} \tag{12.17}$$

We see that this function is even, and is in C^∞ as a function of $\tau = t - s$. Moreover, it is easily checked that the covariance function $f(\tau)$ is strictly smaller than $f(0)$, whenever $\tau \neq k\pi$. Proposition 12.1 and Corollary 12.3 then imply the assertions about continuity and non-existence of more than one global maximum. This concludes the proof of Proposition 12.4. □

We now check some intuitive properties of the position of the minimum of the Gaussian process from Proposition 12.1 (for those ω such that the minimum exists and is unique).

Proposition 12.5: *Suppose that the conditions of Proposition 12.1 are satisfied. Define the space $(\Omega', \mathcal{F}', \mathbb{P}')$ to be the restriction of $(\Omega, \mathcal{F}, \mathbb{P})$ to all ω such that the conclusions of Proposition 12.1 are true. Then the position of the minimum*

$$\vartheta^*[\omega] \equiv \arg\min_{\vartheta \in [0,\pi)} \chi[\omega](\vartheta) \tag{12.18}$$

of the sample path $\chi[\omega]$ is a random variable with uniform distribution on $[0,\pi)$.

Proof: To prove that $\vartheta^*[\omega]$ is a random variable, it is enough to show that for all intervals $\mathcal{U} = (a,b) \subseteq [0,\pi)$, the set $\vartheta^{*-1}(\mathcal{U})$ is in $\mathcal{F}'$. We note that by the continuity of χ on $[0,\pi)$ for all $\omega \in \Omega'$,

$$\begin{aligned} \vartheta^{*-1}(\mathcal{U}) &\equiv \{\omega \in \Omega : \chi[\omega](\cdot) \quad \text{assumes its minimum in} \quad \mathcal{U}\} \\ &= \{\omega \in \Omega' : \exists t \in \mathcal{U} \cap \mathbb{Q} \quad \text{such that} \quad \forall s \in \mathcal{U}^c \cap \mathbb{Q}, \chi(t) < \chi(s)\}. \end{aligned} \tag{12.19}$$

The second line can be written as

$$\bigcup_{t \in \mathcal{U}\cap\mathbb{Q}} \bigcap_{s \in \mathcal{U}^c \cap \mathbb{Q}} \{\omega \in \Omega' : \chi(t) < \chi(s)\}, \tag{12.20}$$

which clearly is in $\mathcal{F}'$.

Equation (12.20), together with the strict stationarity (since it is a real stationary process) of the process χ, implies the uniformity of the distribution. This proves Proposition 12.5. □

Finally, to get some information about the convergence of functions of the position of the minimum, we use the following two results.

Lemma 12.6: *Let $\mathcal{P}([0,\pi))$ be the space of π-periodic, continuous functions, having only one minimum, together with the supremum norm. Suppose we have a sequence of π-periodic, continuous functions (f_n) converging uniformly to $f \in \mathcal{P}([0,\pi))$. Then the positions of the global minima of f_N converge to the position of the global minimum of f.*

Proof: Suppose that there exists a sequence (f_n) of periodic, continuous functions, converging uniformly to $f \in \mathcal{P}([0,\pi))$, but such that infinitely many of the f_n have global minima whose positions do not converge to the position of the unique global minimum ϑ^* of f. Then we can choose a subsequence (f_{n_k}) with global minima $\vartheta^*_{n_k}$ such that $|\vartheta^*_{n_k} - \vartheta^*| > \delta > 0$, $\forall k$.

Since by assumption, $\vartheta^*_{n_k}$ is a global minimum of f_{n_k}, we have that

$$f_{n_k}(\vartheta^*_{n_k}) - f_{n_k}(\vartheta^*) \leq 0, \tag{12.21}$$

On the other hand, for any $\varepsilon > 0$,

$$\begin{aligned} f_{n_k}(\vartheta^*_{n_k}) - f_{n_k}(\vartheta^*) &= f_{n_k}(\vartheta^*_{n_k}) - f(\vartheta^*_{n_k}) + f(\vartheta^*_{n_k}) - f(\vartheta^*) + f(\vartheta^*) - f_{n_k}(\vartheta^*) \\ &\geq -\varepsilon + f(\vartheta^*_{n_k}) - f(\vartheta^*) - \varepsilon, \end{aligned} \tag{12.22}$$

for all k large enough, since f_{n_k} is assumed to converge uniformly to f. Choosing ε small enough, the right hand side of (12.22) can be made positive if indeed $|\theta^* - \theta^*_{n_k}| > \delta > 0$, contradicting (12.21). This implies the lemma. □

The following result is crucial to link the weak convergence of the process $g_N(\vartheta)$ to the weak convergence of the measures $\mathcal{Q}_{N,\beta}$.

Proposition 12.7: *Define the random sets*

$$L_N[\omega] = \left\{\vartheta \in [0,\pi) : \eta_N[\omega](\vartheta) - \min_{\vartheta'} \eta_N[\omega](\vartheta') \leq \varepsilon_N\right\} \tag{12.23}$$

with ε_N some sequence converging to zero. Then

$$L_N \xrightarrow{\mathcal{D}} \vartheta^* \tag{12.24}$$

Proof: The random processes η_n, η lie a.s. in the space of π-periodic C^∞ functions. This space, together with the sup-norm topology, is separable due to Weierstrass' approximation theorem. In this situation the *method of a single probability space* (see [Shi], Chapter 3, Section 8, Theorem 1) ensures the existence of a probability space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ and random processes η^*_N, η^*, such that

$$\eta^*_N \to \eta^*, \quad \mathbb{P}^* - a.s., \tag{12.25}$$

and

$$\eta^* \stackrel{\mathcal{D}}{=} \eta, \qquad \eta^*_N \stackrel{\mathcal{D}}{=} \eta_N. \tag{12.26}$$

Introduce the random level sets

$$L^*_N[\omega^*] = \left\{\vartheta \in [0,\pi) : \eta^*_N[\omega^*](\vartheta) - \min_{\vartheta'} \eta^*_N[\omega^*](\vartheta') \leq \varepsilon_N\right\},$$

Then L_N and L^*_N have the same distribution. But since $\eta^*_N[\omega]$ converges almost surely to $\eta^*[\omega] \in \mathcal{P}([0,\pi))$, one sees that due to Lemma 12.6, $L^*_N[\omega]$ converges $\mathbb{P}^*$-a.s. to the position of the unique absolute minimum of $\eta^*[\omega^*]$. This minimum has the same distribution as that of η, which is the uniform distribution by Proposition 12.5. Therefore, L_N converges in distribution to a uniformly distributed point on $[0,\pi)$. $\square$

We have finally all tools available to prove Theorem 10.3.

Proof of Theorem 10.3: We have to check convergence of measures ν on the following type of functions $F : \mathcal{M}(\mathbb{R}^2) \to \mathbb{R}$

$$F(\nu) = \widetilde{F}(\nu(f_1), \dots, \nu(f_k)), \tag{12.27}$$

where $\widetilde{F}$ is a polynomial function, and $f_1, \dots, f_k$ are bounded continuous functions from $\mathbb{R}^2 \to \mathbb{R}$. Convergence in law then means that

$$\lim_{N\uparrow\infty} \mathbb{E}\Big[F(\mathcal{Q}_{N,\beta}[\omega])\Big] = \frac{1}{\pi}\int_0^\pi F(\frac{1}{2}\delta_{(m^*\cos\vartheta, m^*\sin\vartheta)} + \frac{1}{2}\delta_{(m^*\cos\vartheta+\pi, m^*\sin\vartheta+\pi)})\, d\vartheta. \tag{12.28}$$

The left-hand side of (12.28) is explicitly written as

$$\lim_{N\uparrow\infty} \mathbb{E}\Big[\widetilde{F}(\mathcal{Q}_{N,\beta}[\omega](f_1), \dots, \mathcal{Q}_{N,\beta}[\omega](f_k))\Big]. \tag{12.29}$$

We now treat the individual arguments of $\widetilde{F}$ in (12.29). Let $A_N[\omega]$ (the level sets in the previous lemmata) be decomposed into its $2l'$ connected components $A_{N,j}[\omega]$. As a consequence of Lemma 12.7, there exists $N[\omega]$ which is finite a.s. such that for all $N \geq N(\omega)$, $l = 1$, and the two corresponding connected components are symmetric with respect to the origin. Now choose arbitrary points $x_{N,j}[\omega] \in A_{N,j}[\omega]$. Then we can decompose

$$\begin{aligned}\widetilde{\mathcal{Q}}_{N,\beta}[\omega](f_i) = &\sum_j f_i(x_{N,j})\widetilde{\mathcal{Q}}_{N,\beta}[\omega](\mathbb{1}_{A_{N,j}}) + \sum_j \widetilde{\mathcal{Q}}_{N,\beta}(\mathbb{1}_{A_{N,j}}(f_i(x_{N,j}) - f_i)) \\ &+ \widetilde{\mathcal{Q}}_{N,\beta}(\mathbb{1}_{A_N^c} f_i).\end{aligned} \tag{12.30}$$

Expanding $\widetilde{F}$ using the decomposition (12.30), we get a sum consisting of two different types of terms: (i), summands that are products of the first sum on the right-hand side of (12.30) only, and (ii), summands where at least one of the second and third term from the right-hand side of (12.30) enter. Proposition 11.3 and Proposition 12.7, and the continuity and boundedness of the f_i's imply that the terms of type (ii) vanish $\mathbb{P}$-a.s., as $N \uparrow \infty$. In the limit, the only terms left are of type (i), which together sum up to

$$\widetilde{F}\left(\sum_j f_1(x_{N,j})\widetilde{\mathcal{Q}}_{N,\beta}[\omega](\mathbb{1}_{A_{N,j}}), \dots, \sum_j f_k(x_{N,j})\widetilde{\mathcal{Q}}_{N,\beta}[\omega](\mathbb{1}_{A_{N,j}})\right) \tag{12.31}$$

All arguments of $\widetilde{F}$ in (12.31) converge in distribution to

$$\frac{1}{2} f_i((m^*\cos\vartheta, m^*\sin\vartheta)) + \frac{1}{2} f_i((m^*\cos\vartheta + \pi, m^*\sin\vartheta + \pi)), \quad \forall i = 1, \dots, k \tag{12.32}$$

where ϑ is a uniformly distributed r.v. on $[0,\pi)$, by Proposition 12.7. But convergence in distribution means by definition that

$$\begin{aligned}\lim_{N\uparrow\infty}\mathbb{E}\Big[\widetilde{F}\big(\sum_{j_N} f_1(x_{N,j_N})\widetilde{\mathcal{Q}}_{N,\beta}[\omega](A_{N,j_N}),\dots,\sum_{j_N} f_2(x_{N,j_N})\widetilde{\mathcal{Q}}_{N,\beta}[\omega](A_{N,j_N})\big)\Big]\\ =\frac{1}{\pi}\int_0^\pi \widetilde{F}\big(\frac{1}{2}f_i((m^*\cos\vartheta, m^*\sin\vartheta))+\frac{1}{2}f_i((m^*\cos\vartheta+\pi, m^*\sin\vartheta+\pi))\,d\vartheta,\end{aligned} \tag{12.33}$$

which in turn is by definition equal to

$$\frac{1}{\pi}\int_0^\pi F\big(\frac{1}{2}\delta_{(m^*\cos\vartheta, m^*\sin\vartheta)}+\frac{1}{2}\delta_{(m^*\cos\vartheta+\pi, m^*\sin\vartheta+\pi)}\big)\,d\vartheta. \tag{12.34}$$

This proves the convergence in law (10.12) in Theorem 10.3. To obtain the identification of the metastate, just note that the process $\eta_N(\vartheta)[\omega]$ actually converges to the same Gaussian process under any of the conditional laws $\mathbb{P}[\cdot|\mathcal{F}_n]$, where $\mathcal{F}_n$ is the sigma-algebra generated by the random variables $\xi_i, i\leq n$. $\square$

13 Volume Dependence, Empirical Metastates, Superstates

We conclude this paper with the discussion of some more sophisticated concepts that have been proposed by Newman and Stein [NS2] and Bovier and Gayrard [BG3] and that should capture in more detail the actual asymptotic volume dependence of the Gibbs measures. In fact, the first question one may ask is whether for a fixed realization as the volume grows the finite volume Gibbs states really explore all the possibilities in the support of the metastate. One way of stating that this is the case is the following

Theorem 13.1: *There exist (deterministic) sequences $N_k \uparrow \infty$ such that the empirical metastate*

$$\frac{1}{k}\sum_{\ell=1}^{k} \delta_{\mathcal{Q}_{N_k,\beta}}, \tag{13.1}$$

converges almost surely to the law of $\mathcal{Q}_{\infty,\beta}$.

Proof: We have seen that the measure $\mathcal{Q}_{N_k,\beta}$ is sharply concentrated on the circle of radius r^* and at the angle where the process $g_{N_k}(\vartheta)$ (defined in (11.3) takes its absolute minimum. The idea is to choose N_k in such a way that these angles will be virtually independent for different k. Now note that we can write

$$g_{N_k}(\vartheta) = \widetilde{g}_k(\vartheta) + R_k(\vartheta), \tag{13.2}$$

where

$$\widetilde{g}_k(\vartheta) = \frac{1}{N_k} \sum_{i=N_{k-1}+1}^{N_k} \ln\cosh(\beta(r^*\zeta_i \cos(\vartheta - \varphi_i))), \tag{13.3}$$

are independent for different k by construction and

$$R_k(\vartheta) = \frac{1}{N_k} \sum_{i=1}^{N_{k-1}} \ln\cosh(\beta(r^*\zeta_i \cos(\vartheta - \varphi_i))). \tag{13.4}$$

Now by standard estimates identical to those presented in Chapter 12, one shows easily that there is a constant $C < \infty$ such that

$$\mathbb{P}\left[\sup_{\vartheta\in[0,\pi)} |R_k(\vartheta) - \mathbb{E}R_k(\vartheta)| \geq x\frac{N_{k-1}}{N_k}\right] \leq C\exp\left(-x^2/C\right). \tag{13.5}$$

Thus we can always choose N_k growing sufficiently rapidly (e.g. $N_k = k!$) such that R_k is totally negligible compared to $\widetilde{g}_k$ for large k, and the position of the absolute minimum of $g_{N_k}(\vartheta)$ is asymptotically equal to that of $\widetilde{g}_k(\vartheta)$. This allows us to approximate for large k the random measures $\delta_{\mathcal{Q}_{N_k,\beta}}$ by independent measures and from this the asserted result follows from the law of large numbers. □

Remark: Theorem 13.1 says that that the empirical metastate constructed with sparse subsequences converges to the Aizenman-Wehr metastate, a.s.. This is a special example of a general theorem due to Newman and Stein [NS2] (where however they require possibly subsequences ℓ_i in the definition (13.1)).

Rather than considering the empirical metastate with sparse subsequences one may be interested in the volume dependence as the volume grows at its natural pace. To capture this, the idea put forward in [BG3] is to construct a measure valued stochastic process

$$\mu_\beta^t \equiv \lim_{N\uparrow\infty} \mu_{\beta,[tN]}, \tag{13.6}$$

with $t \in (0,1]$ and to consider either the (conditional) probability distribution of this process (the "superstate" [BG3]) or the (conditional) empirical distribution of the process (the "empirical metastate" [NS2]). Let us see what this entails in our context. The reader who has been following the exposition of the last two chapters will easily be convinced that this problem amounts to study the quantity

$$\vartheta_t^* \equiv \arg\min_{\vartheta\in[0,\pi)} \left(\chi_t(\vartheta)\right), \tag{13.7}$$

where $\chi_t(\vartheta)$ is the distributional limit of the process

$$\chi_N^t(\vartheta) \equiv g_{[tN]}(\vartheta) - \mathbb{E}g_{[tN]}(\vartheta). \tag{13.8}$$

where $g_N(\vartheta)$ is defined in (11.3). By completely standard arguments one shows that the following invariance principle holds:

Lemma 13.2: *The process $\chi_N^t(\vartheta)$ converges in distribution, as $N\uparrow\infty$ to the Gaussian process $\chi_t(\vartheta)$, $t\in(0,1], \vartheta\in[0,\pi)$ with mean zero and covariance*

$$C(\vartheta,\vartheta',t,t') \equiv \frac{t\wedge t'}{\sqrt{tt'}} f(\vartheta,\vartheta'), \tag{13.9}$$

where

$$f(\vartheta,\vartheta') = \mathbb{E}\left[\ln\cosh\left(\beta r\zeta_1\cos(\varphi)\right)\ln\cosh\left(\beta r\zeta_1\cos\left(\varphi-(\vartheta-\vartheta')\right)\right)\right]. \tag{13.10}$$

$\chi_t(\vartheta)$ is a rather curious Gaussian process: as a function of t, for fixed ϑ it is (normalized) Brownian motion, while for fixed t as a function of ϑ it is the C^∞ process discussed in the previous section. The question is then: what can be said about the process ϑ_t^*, defined by (13.7)?

Some facts follow easily. For instance, the process is almost surely single valued for all $t\in(0,1]$ except possibly on some Cantor set of zero Lebesgue measure. On the other hand, it seems natural that such an exceptional set will exist and that a typical realization will have continuous pieces and "jumps". Also, for t going to zero, the process "circles" around rapidly since χ_t and χ_s become uncorrelated as $s\downarrow 0$. But otherwise we do not see any immediate more specific characterization of the process or its path-properties.

Appendix: A Deviation Inequality

The aim of this appendix is to prove a deviation inequality due to Ledoux [Le] under slightly weaker conditions, and which is used in the course of the proof of Theorem 2.2. Ledoux starts by proving the following Log-Sobolev inequality (his Theorem 1.2, resp. inequality (1.6) with the optimal constant).[34]

Theorem A.1: *Let g be smooth function on $\mathbb{R}^n$ such that $\ln g^2$ is separately convex ($g^2 > 0$). Then, for any product probability $\mathbb{P}$ on $[-1,1]^n$,*

$$\mathbb{E}\left[g^2 \ln g^2\right] - \mathbb{E}\left[g^2\right] \ln \mathbb{E}\left[g^2\right] \leq 8\,\mathbb{E}\left[|\nabla g|^2\right]. \qquad (A.1)$$

From this the deviation inequality follows as a corollary (compare [Le] and references therein).

Theorem A.2: *Let f be a separately convex Lipschitz function on $\mathbb{R}^n$ with Lipschitz constant $\|f\|_{Lip} \leq 1$. Then, for every $t \geq 0$,*

$$\mathbb{P}\left[f \geq \mathbb{E} f + t\right] \leq e^{-\frac{t^2}{2}}. \qquad (A.2)$$

Unfortunately, in the application we have in mind, the uniform bound on the Lipschitz constant is not uniformly satisfied. However, it is violated only on a set of exponentially small probability. In this situation, we would like to have a tool similar to Theorem 6.6 in [T1], which handles this inconvenience in the case of convex functions. It is clear that one needs some additional integrability conditions on f, to make up for its weaker Lipschitz properties. The conditions we present are adapted to what we can prove about the free energy of the p-spin Hopfield model in chapter 4. The proof is not very original and follows essentially the lines of Ledoux [Le].

Theorem A.3: *Suppose that $G_N : [-1,1]^{L(N)} \to \mathbb{R}$ are smooth positive functions, separately convex, and satisfy the following conditions: there exist constants $c^*, t^*, \kappa, \alpha > 0$, and $\bar{N} \in \mathbb{N}$ such that for all $N \geq \bar{N}$,*

- *(i) for all $c > c^*$: $\mathbb{P}[\|G_N\|_{Lip} > cN^{-\frac{1}{2}}] \leq e^{-\delta(c_1)N}$, for some increasing function δ;*
- *(ii) for all $t > t^*$: $\mathbb{P}[|G_N| > tN] \leq e^{-\kappa N(t-\alpha)}$;*
- *(iii) the Lipschitz constant as a function of N is uniformly bounded by some polynomial function p of N.*

Then there exist constants $K_1, \ldots, K_5 > 0$ such that

$$\mathbb{P}[NG_N \geq N \int G_N dP + tN] \leq \begin{cases} e^{-\frac{Nt^2}{K_1}}, & \text{if } t < K_2, \\ K_3 e^{-K_4 N(t-K_5)} & \text{otherwise.} \end{cases} \qquad (A.3)$$

Thus, fluctuations above the mean are of the order of the square root of N. This is stated more elegantly in the following immediate corollary, which is used in Chapter 4 as Theorem 4.6.

---

[34] Note that we adapt the results to functions defined on $[-1,1]$, respectively $[-1,1]^n$.

Corollary A.4: *Under the conditions of Theorem A.3, there exists for each $t > 0$ and $k \in [\frac{1}{2}, 1)$ constants $C > 0$ and $\bar{N} \in \mathbb{N}$ such that*

$$\mathbb{P}[NG_N \geq N \int G_N dP + tN^k] \leq e^{-\frac{N^{2k-1}t^2}{C}}, \tag{A.4}$$

for all $N \geq \bar{N}$.

Proof of Theorem A.3: Suppose first that $\mathbb{P}$ is absolutely continuous with respect to Lebesgue measure on $\Omega_N = [-1,1]^N$. We observe that under our conditions, the Log-Sobolev inequality (A.1) holds for $g^2 = e^{\lambda N G_N}$,

$$\int g^2 \log g^2 d\mathbb{P} - \int g^2 d\mathbb{P} \log \int g^2 d\mathbb{P} \leq 8 \int |\nabla g^2| d\mathbb{P}. \tag{A.5}$$

This is equivalent to the following differential inequality for the function $\tilde{G}(\lambda) = \int e^{\lambda N G_N} d\mathbb{P}$,

$$\lambda \frac{d}{d\lambda} \tilde{G}(\lambda) - \tilde{G}(\lambda) \ln \tilde{G}(\lambda) \leq 2\lambda^2 N^2 \int |\nabla G_N|^2 e^{\lambda N G_N}] d\mathbb{P}. \tag{A.6}$$

To integrate this inequality, we seek a good upper bound for its right-hand side. We first observe that since $\mathbb{P}$ is absolutely continuous with respect to Lebesgue measure, the set where $|\nabla G_N| > c$ has the same measure as the set in hypothesis (i). We now decompose the integral as

$$\begin{aligned} \int |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P} = & \int_{|\nabla G_N| \leq c_1 N^{-\frac{1}{2}}} |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P} + \int_{\substack{|\nabla G_N| > c_1 N^{-\frac{1}{2}} \\ |G_N| \leq c_2}} |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P} \\ & + \int_{\substack{|\nabla G_N| > c_1 N^{-\frac{1}{2}} \\ |G_N| > c_2}} |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P}, \end{aligned} \tag{A.7}$$

where $c_1 > c^*$ and $c_2 > t^*$ (and will be determined later). The first term is bounded by

$$\int_{|\nabla G_N| \leq c_1 N^{-\frac{1}{2}}} |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P} \leq c_1^2 N^{-1} \int e^{\lambda N G_N} d\mathbb{P} = c_1^2 N^{-1} \tilde{G}(\lambda). \tag{A.8}$$

In the second term, we use a uniform bound on $|\nabla G_N|$, resp. G_N, and observe that since G_N is positive, $\tilde{G}(\lambda)$ is greater than one. Hence,

$$\begin{aligned} \int_{\substack{|\nabla G_N| > c_1 N^{-\frac{1}{2}} \\ |G_N| \leq c_2}} |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P} & \leq p(N)^2 e^{\lambda c_2 N} \mathbb{P}[|\nabla G_N| > c_1 N^{-\frac{1}{2}}] \\ & \leq p(N)^2 e^{\lambda c_2 N} e^{-\delta(c_1) N} \leq p(N)^2 e^{\lambda c_2 N} e^{-\delta(c_1) N} \tilde{G}(\lambda) \end{aligned} \tag{A.9}$$

by hypothesis (i), resp. (iii). This means that if $\lambda c_2 \leq \delta(c_1)$, this term vanishes exponentially. To treat the remaining term in the decomposition (A.7), we write it as follows,

$$\begin{aligned}
\int\limits_{\substack{|\nabla G_N|>c_1 N^{-\frac{1}{2}} \\ |G_N|>c_2}} |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P} &\leq p(N)^2 \int\limits_{\substack{|\nabla G_N|>c_1 N^{-\frac{1}{2}} \\ |G_N|>c_2}} e^{\lambda N G_N} d\mathbb{P} \\
&= p(N)^2 \sum_{t_i} \int\limits_{\substack{|\nabla G_N|>c_1 N^{-\frac{1}{2}} \\ |G_N|\in(t_i,t_{i+1}]}} e^{\lambda N G_N} d\mathbb{P} \\
&\leq p(N)^2 \sum_{t_i} e^{\lambda N t_{i+1}} \min\{\mathbb{P}[|\nabla G_N| > c_1 N^{-\frac{1}{2}}], \mathbb{P}[G_N > t_i]\},
\end{aligned} \tag{A.10}$$

where $\{t_i\}_{i\in\mathbb{N}}$ is the decomposition of (c_2,∞) into unit intervals. By condition (ii),

$$\mathbb{P}[G_N > t_i] \leq K e^{-N\kappa(t_i-\alpha)}, \tag{A.11}$$

which is less than $\mathbb{P}[|\nabla G_N| > c_1 N^{-\frac{1}{2}}]$ whenever $t_i > \alpha + c_2/\kappa$. We therefore choose $c_3 > \alpha + c_2/\kappa$. Thus, (again using the fact that $\tilde{G} > 1$)

$$\begin{aligned}
\int\limits_{\substack{|\nabla G_N|>c_1 N^{-\frac{1}{2}} \\ |G_N|>c_2}} |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P} &\leq K p(N)^2 \sum_{t_i} e^{\lambda t_{i+1} N - N\kappa(t_i-\alpha)} \\
&\leq K p(N)^2 e^{N\kappa\alpha - N\lambda} \sum_{t_i} e^{-N t_i(\kappa-\lambda)} \\
&\leq K p(N)^2 e^{N\kappa\alpha - N\lambda + N c_2(\kappa-\lambda)} \left(1 - e^{-N(\kappa-\lambda)}\right)^{-1} \tilde{G}(\lambda).
\end{aligned} \tag{A.12}$$

Therefore, if

$$c_2 > \frac{\kappa\alpha + \lambda}{\kappa - \lambda}, \tag{A.13}$$

the coefficient of $\tilde{G}(\lambda)$ on the right-hand side of (A.12) will tend to zero exponentially fast.

Collecting the bounds for the three parts in the decomposition (A.7), we get that there exist constants $c_1^*, c_3, c_4, \lambda^* > 0$ such that

$$\int |\nabla G_N|^2 e^{\lambda N G_N} d\mathbb{P} \leq \left(c_1^2 N^{-1} + c_3 e^{-c_4 N}\right) \tilde{G}(\lambda) \tag{A.14}$$

for all $c_1 > c_1^*$, and $\lambda \in (0, \lambda^*)$.

We are now ready to integrate the differential inequality (A.6) for $\tilde{G}(\lambda)$. Inserting (A.14) yields

$$\lambda \frac{d}{d\lambda} \tilde{G}(\lambda) - \tilde{G}(\lambda) \ln \tilde{G}(\lambda) \leq \left(2\lambda^2 N^1 c_1^2 + 2 c_3 \lambda^2 N^2 e^{-c_4 N}\right) \tilde{G}(\lambda). \tag{A.15}$$

Let $H(\lambda) \equiv \frac{1}{\lambda} \log \tilde{G}(\lambda)$. Then (A.15) reduces to

$$H'(\lambda) \leq 2 c_1^2 N + 2 c_1^2 N^2 e^{-c_4 N}. \tag{A.16}$$

Moreover,

$$H(0) = \lim_{\lambda \downarrow 0} \frac{1}{\lambda} \tilde{G}(\lambda) = \frac{\tilde{G}'(0)}{\tilde{G}(0)} = N \int G_N d\mathbb{P}. \tag{A.17}$$

Thus, for all $\lambda \in (0, \lambda^*)$,

$$H(\lambda) \leq N \int G_N d\mathbb{P} + 2c_1^2 N\lambda + 2c_1^2 N^2 e^{-c_4 N} \lambda, \tag{A.18}$$

which in turn implies that

$$\tilde{G}(\lambda) \leq \exp\left(\lambda N \int G_N d\mathbb{P} + \lambda^2 (2Nc_1^2 + 2c_3 N^2 e^{-c_4 N})\right). \tag{A.19}$$

Let now $\mathbb{P}$ be arbitrary. Then any smooth convolution of $\mathbb{P}$ will satisfy the above inequality. Since G_N is supposed to be continuous, the same is true for $\mathbb{P}$ itself.

To finish the proof of the Theorem, we use the exponential Chebyshev inequality, i.e.

$$\begin{aligned} \mathbb{P}[NG_N > N \int G_N d\mathbb{P} + Nt] &\leq e^{-\lambda t N - \lambda N \int G_N d\mathbb{P}} \int e^{N\lambda G_N} d\mathbb{P} \\ &= e^{-\lambda t N - \lambda N \int G_N d\mathbb{P}} \tilde{G}(\lambda), \end{aligned} \tag{A.20}$$

for all $\lambda > 0$. Using (A.19), we get

$$\mathbb{P}[NG_N > N \int G_N d\mathbb{P} + Nt] \leq \exp\left(-\lambda t N + \lambda^2 (2Nc_1^2 + 2c_3 N^2 e^{-c_4 N})\right). \tag{A.21}$$

Optimizing with respect to λ then yields (uniformly over all $N \geq \bar{N}$)

$$\mathbb{P}[NG_N > N \int G_N d\mathbb{P} + Nt] \leq C \begin{cases} e^{-\frac{Nt^2}{8c_1^2}}, & \text{if } t \leq 4c_1^2 \lambda^*, \\ e^{-\lambda^* t N - 2\lambda^{*2} N c_1^2}, & \text{otherwise,} \end{cases} \tag{A.22}$$

which is the statement of Theorem A.3. □

References

[ALR] M. Aizenman, J. L. Lebowitz, D. Ruelle, *Some rigorous results on the Sherrington-Kirkpatrick model*, Commun. Math. Phys. **112**, 3–20 (1987)

[AW] M. Aizenman, J. Wehr, *Rounding effects of quenched randomness of first-order phase transitions*, Comm. Math. Phys. **130**, 489–528 (1990).

[Ba] Z. D. Bay, *Methodologies in spectral analysis of large dimensional random matrices, a review*, Stat. Sin. **9**, 611–677 (1999).

[BaY] Z. D. Bay, Y. Q. Yin, *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix*, Ann. Prob. **16**, 1729–1741 (1988).

[Be] K. A. Berman, *On the number of balanced orientations of even graphs*, in: Proc. of the Tenth Southeastern Conf. on Comb., Graph Theory, and Comput., (Boca Raton, 1979), Vol. I, Congr. Numer. XXIII, 193–198 (1979).

[BCKM] J.-P. Bouchaud, L. F. Cugliandolo, J. Kurchan, M. Mézard, *Out of equilibrium dynamics in spin-glasses and other glassy systems*, in *Recent progress in random magnets*, A. P. Young (ed.), World Scientific, Singapore (1997).

[B1] A. Bovier, *Self-averaging in a class of generalized Hopfield Models*, J. of Physics **A 27**, 7069–7077 (1994).

[B2] A. Bovier, *Some remarks on the REM and the p-spin SK-model*, unpublished (1999).

[B3] A. Bovier, *Statistical mechanics of disordered systems*, Lecture Notes, TU Berlin (2000).

[BvEN] A. Bovier, A. C. D. van Enter, B. Niederhauser, *Stochastic symmetry breaking in a Gaussian Hopfield model*, J. Stat. Phys. **95**, 181–213 (1999).

[BG1] A. Bovier, V. Gayrard, *Hopfield models as generalized random mean field models*, in *Mathematical aspects of spin glasses and neural networks*, A. Bovier and P. Picco (eds.), Progress in Probability, Birkhäuser, Boston-Basel-Berlin (1998).

[BG2] A. Bovier, V. Gayrard, *The retrieval phase of the Hopfield model: a rigorous analysis of the overlap distribution*, Prob. Theory Rel. Fields **107**, 61–98 (1997).

[BG3] A. Bovier, V. Gayrard, *Metastates in the Hopfield model in the replica symmetric regime*, Math. Phys. Anal. Geom. **1**, 107-144 (1998).

[BG4] A. Bovier, V. Gayrard, *An almost sure central limit theorem for the Hopfield model*, Markov Proc. Related Fields **3**, 151–173 (1997).

[BGP1] A. Bovier, V. Gayrard, P. Picco, *Gibbs states of the Hopfield model in the regime of perfect memory*, Prob. Theory Rel. Fields **100**, 329–363 (1994).

[BGP2] A. Bovier, V. Gayrard, P. Picco, *Gibbs states of the Hopfield model with extensively many patterns*, J. Stat. Phys. **79**, 395–414 (1995).

[BGP3] A. Bovier, V. Gayrard, P. Picco, *Distribution of overlap profiles in the one-dimensional Kac-Hopfield model*, Commun. Math. Phys. **186**, 232–379 (1997)

[BF] A. Bovier, J. Fröhlich, *A heuristic theory of the spin-glass phase*, J. Stat. Phys. **44**, 347–391 (1986).

[BKL] A. Bovier, M. Löwe, I. Kurkova, *Fluctuations of the free energy in the REM and the p-spin SK models*, `preprint.`

[BM] A. Bovier, D. Mason, *Extreme value behaviour in the Hopfield model*, to appear in Ann. of

Appl. Prob. (2000).

[BdMS] A. Boutet de Monvel, M. V. Shcherbina, *On the norm of random matrices*, Math. Notes **57**, 475–484 (1995).

[COP] M. Cassandro, E. Orlandi, P. Picco, *Typical configurations for one dimensional random field Kac model*, Ann. Prob. **27**, 1414–1467 (1999).

[Cho] D. Chowdhury, *Spin glasses and other frustrated systems*, World Scientific, Singapore (1986).

[CH] R. Courant, D. Hilbert, *Methoden der mathematischen Physik I*, 3rd ed., Springer, Berlin-Heidelberg-New York (1968).

[CL] H. Cramér, M. R. Leadbetter, *Stationary and related stochastic processes*, Wiley (1967).

[D1] B. Derrida, *Random energy model: limit of a family of disordered models*, Phys. Rev. Letts. **45**, 79–82 (1981).

[D2] B. Derrida, *Random energy model: An exactly solvable model of disordered systems*, Phys. Rev. B **24**, 2613–2626 (1984).

[Do] R. L. Dobrushin, *Description of a random field by means of conditional probabilities and conditions for its regularity*, Theor. Prob. Appl. **15**, 197–224 (1968).

[vE] A. C. D. van Enter, *Stiffness exponent, number of pure states, and Almeida-Thouless line in spin-glasses* J. Stat. Phys. **60**, 275-279 (1990).

[vEFS] A. C. D. van Enter, R. Fernandez, A. D. Sokal, *Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory*, J. Stat. Phys. **72**, 879–1167 (1993).

[FH1] D. S. Fisher, D. A. Huse, *Ordered phase of short-range spinglasses*, Phys. Rev. Lett. **56**, 1601–1604 (1986).

[FH2] D. S. Fisher, D. A. Huse, *Equilibrium behavior of the spin-glass ordered phase*, Phys. Rev. B **38**, 386–411 (1988).

[FH3] D. S. Fisher, D. A. Huse, *Pure states in spin glasses*, J. Phys. A **20**, L997-1003 (1987).

[FH4] D. S. Fisher, D. A. Huse, *Absence of many states in realistic spin glasses*, J. Phys. A **20**, L1005–1010 (1987).

[FP1] L. A. Pastur, A. L. Figotin, *Exactly soluble model of a spin glass*, Sov. J. Low Temperature Physics **3**(**6**), 378–383 (1977).

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

[F] H. Fleischner, *Eulerian graphs and related topics, Part 1, vol. 2*, Annals of discrete mathematics **50**, North-Holland, Amsterdam (1991).

[FK] Z. Füredi, J. Komlos, *The eigenvalue of random symmetric matrices*, Combinatorica **1**, 233–241 (1981).

[Ge] S. Geman, *A limit theorem for the norms of random matrices*, Ann. Prob. **8**, 252–261 (1980).

[GL] B. Gentz, M. Löwe, *Fluctuations in the Hopfield model at the critical temperature*, Markov Proc. Related Fields **5**, 423–449 (1999).

[Geo] H.-O. Georgii, *Gibbs measures and phase transitions*, de Gruyter Studies in Mathematics **9**, Walter de Gruyter, Berlin-New York (1988).

[GG] S. Govindarajan, R. A. Goldstein, *Evolution of model proteins on a foldability landscape*, Proteins **29**, 461–466 (1997).

[H] D. O. Hebb, *The organization of behavior*, Wiley, New York (1949).

[He] H. Heuser, *Lehrbuch der Analysis, Teil 2*, 5th ed., Teubner, Stuttgart (1990).

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

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

[K] A. Kotzig, *On finite graphs oriented in equilibrium*, Časopis Pěst Mat. **84**, 31–45 (1959).

[Ku1] C. Külske, *Metastates in disordered mean field models: random field and Hopfield models*, J. Stat. Phys. **88**, 1257–1293 (1997).

[Ku2] C. Külske, *Limiting behavior of random Gibbs measures: metastates in some disordered mean field models*, in *Mathematical aspects of spin glasses and neural networks*, A. Bovier, P. Picco (eds.), Progress in Probability 41, Birkhäuser, Boston-Basel-Berlin (1998).

[Ku3] C. Külske, *Metastates in disordered mean-field models II: the superstates*, J. Stat. Phys. **91**, 155–176 (1998).

[LR] O. E. Lanford, D. Ruelle, *Observables at infinity and states with short range correlations in statistical mechanics*, Comm. Math. Phys. **13**, 194–215 (1969).

[LLR] M. R. Leadbetter, G. Lindgren, H. Rootzén, *Extremes and related properties of random sequences and processes*, Springer, Berlin-Heidelberg-New York (1983).

[Le] M. Ledoux, *On Talagrand's deviation inequalities for product measures*, ESAIM: Probability **1**, 63–87 (1996).

[LT] M. Ledoux and M. Talagrand, "Probability in Banach spaces", Springer, Berlin-Heidelberg-New York (1991).

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

[MPR] E. Marinari, G. Parisi, J. J. Ruiz-Lorenzo, *Numerical simulations of spin glass systems*, in *Spin glasses and random fields*, A. P. Young (ed.), World Scientific, Singapore (1998).

[MCP] W. S. McCullough, W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, Bull. Math. Biophys. **5**, 115–133 (1943).

[MPV] M. Mézard, G. Parisi, M. A. Virasoro, *Spin glass theory and beyond*, World Scientific, Singapore (1987).

[MS] V. D. Milman, G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture notes in Mathematics **1200**, Springer, Berlin-Heidelberg-New York (1986).

[N1] C. M. Newman, *Memory capacity in neural network models: rigorous lower bounds*, Neural Networks **1**, 223–238 (1988).

[N2] C. M. Newman, *Topics in disordered systems*, Birkhäuser, Boston-Basel-Berlin (1997).

[NS1] C. M. Newman, D. L. Stein, *Non-mean-field behavior of realistic spin glasses* Phys. Rev. Lett. **76**, 515-518 (1996).

[NS2] C. M. Newman, D. L. Stein, *Thermodynamic chaos and the structure of short-range spin glasses*, in *Mathematical aspects of spin glasses and neural networks*, A. Bovier and P. Picco (eds.), Progress in Probability 41, Birkhäuser, Boston-Basel-Berlin (1998).

[NS3] C. M. Newman, D. L. Stein, *Spatial inhomogeneity and thermodynamic chaos*, Phys. Rev. Lett. **76**, 4821–4824 (1996).

[NS4] C. M. Newman, D. L. Stein, *Metastate approach to thermodynamic chaos*, Phys. Rev. E **55**, 5194–5211 (1997).

[NS5] C. M. Newman, D. L. Stein, *Multiple states and thermodynamic limits in short ranged Ising spin glass models*, Phys. Rev. B **46**, 973–982 (1992).

[NS6] C. M. Newman, D. L. Stein, *Simplicity of state and overlap structure in finite-volume realistic spin glasses*, Phys. Rev. E **57**, 1356–1369 (1998).

[PPRT] A. Pagnani, G. Parisi, F. Ricci-Tersenghi, *Glassy transition in a disordered model for the RNA secondary structure*, Phys. Rev. Lett. **84**, 2026–2029 (2000).

[PS] L. Pastur, M. Shcherbina, *Absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model*, J. Stat. Phys. **74**, 1161–1183 (1994).

[PN] P. Peretto, J. J. Niez, *Long term memory storage capacity of multiconnected neural networks*, Biolog. Cybernetics **39**, 53–63 (1986).

[Ro] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton (1972).

[ST] M. Shcherbina, B. Tirozzi, *The free energy for a class of Hopfield models*, J. Stat. Phys. **72**, 113–125 (1992).

[S] D. Sherrington, *Spin glasses*, in *Physics of novel materials*, M. P. Das (ed.), World Scientific, Singapore (1999).

[SK] D. Sherrington, S. Kirkpatrick, *Solvable model of a spin-glass*, Phys. Rev. Lett. **35**, 1792–1796 (1975).

[Shi] A. N. Shiryaev, *Probability*, 2^{nd} ed., Graduate Texts in Mathematics **95**, Springer, Berlin-Heidelberg-New York (1996).

[Si] J. W. Silverstein, *Eigenvalues and eigenvectors of large sample covariance matrices*, Contemporary Math. **50**, 153–159 (1986).

[SiSo] Ya. G. Sinai, A. B. Soshnikov, *A refinement of Wigner's semicircle law in a neighbourhood of the spectrum edge for random symmetric matrices*, Funct. Anal. Appl. **32**, 114–131 (1998).

[So] A. B. Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys. **207**, 697-733 (1999).

[Sty] L. Stryer, *Biochemistry*, 4^{th} ed., Freeman and Company, New York (1995).

[T1] M. Talagrand, *A new look at independence*, Ann. Prob. **24**, 1–34 (1996).

[T2] M. Talagrand, *The Sherrington-Kirkpatrick model: a challenge for mathematicians*, Prob. Theory Rel. Fields **110**, 109–176 (1998).

[T3] M. Talagrand, *Rigourous results for the Hopfield model with many patterns*, Prob. Theory Rel. Fields **110**, 177–276 (1998).

[T4] M. Talagrand, *Rigorous low temperature results for the mean field p-spin interaction model*, Prob. Theory Rel. Fields **117**, 303–360 (2000).

[Wi1] E. Wigner, *Characteristic vectors of bordered matrices with infinite dimension*, Ann. Math. **62**, 548–564 (1955).

[Wi2] E. Wigner, *On the distribution of the roots of certain symmetric matrices*, Ann. Math. **67**, 325–328 (1958).

[Wi3] E. Wigner, *Random matrices in physics*, SIAM Review **9**, 1–23 (1967).

[YBaK] Y. Q. Yin, Z. D. Bai, P. R. Krishnaiah, *On the limit of the largest eigenvalue of the large dimensional sample covariance matrix*, Prob. Theory Rel. Fields **78**, 509–521 (1988).

[Yu] V. V. Yurinskii, *Exponential inequalities for sums of random vectors*, J. Multivariate Analysis **6**, 473–499 (1976).

[Z] B. Zegarlinski, *Random spin systems with long-range interactions*, in *Mathematical aspects of spin glasses and neural networks*, A. Bovier and P. Picco (eds.), Progress in Probability, Birkhäuser, Boston-Basel-Berlin (1998).